Thèse
Nouveau régime
Pour obtenir le grade de

Docteur en Sciences du Langage
Discipline : Linguistique Théorique, Descriptive et Automatique

Présentée et soutenue publiquement par

Eduardo Correa Soares

le 15 décembre de l'année MMXVII

Anaphors in Discourse

Anaphoric Subjects in Brazilian Portuguese

Directeur de thèse :

Philip Harold Miller

Composition du jury :

Anne Abeillé	CNRS, Université Paris Diderot-Paris 7 (présidente)
Barbara Hemforth	CNRS, Université Paris Diderot-Paris 7 (co-directrice)
Jeffrey Runner	University of Rochester (rapporteur)
Patricia Cabredo-Hofher	CNRS, Université Paris 8-Vincennes-Saint-Denis (membre)
Philip Harold Miller	Université Paris Diderot-Paris 7 (directeur)
Scott Schwenter	Ohio State University (rapporteur)
Sergio de Moura Menuzzi	CNPq, Universidade Federal do Rio Grande do Sul (co-directeur)

Résumé

La présente thèse porte sur l'utilisation et l'interprétation des sujets nuls et pronominaux en portugais brésilien. Son objectif est de comprendre les facteurs sémantiques et discursifs qui peuvent être pertinents pour le choix entre ces expressions anaphoriques en brésilien et la façon dont ce choix s'articule avec la théorie générale de la résolution de l'anaphore. Le point de départ de cette thèse a été la recherche sur les sujets nuls et réalisés sous la perspective de la grammaire générative, en particulier la théorie paramétrique. Cette thèse démontre que l'analyse proposée dans cette perspective ne peut pas rendre compte des données observées. Par exemple, la généralisation sur la « pauvreté » de la morphologie verbale directement liée à l'absence, ou à la fréquence réduite, de sujets nuls est contestée avec les données expérimentales ainsi qu'avec la distribution de la fréquence relative des sujets nuls au sein des personnes discursives dans le corpus. Une explication alternative présentée dans la littérature, à savoir l'importance des caractéristiques sémantiques des antécédents – l'Animacité et le Specificité –, semble mieux expliquer la distribution constatée. Cette explication n'est cependant pas suffisante pour comprendre le choix des sujets anaphoriques en brésilien, puisque le nombre relatif de sujets nuls animés et spécifiques est relativement plus élevé que dans les langues à expression obligatoire des sujets. Par conséquent, cette thèse soutient que les facteurs discursifs semblent jouer un rôle crucial dans l'utilisation des sujets nuls et réalisés en brésilien. Les principaux facteurs identifiés ici sont le statut évident de l'antécédent et le caractère contrastif de l'information d'arrière-plan et l'information nouvelle. Le premier est un facteur standard dans la littérature sur la résolution de l'anaphore (exprimé par différents termes comme l'accessibilité, la familiarité, la thématicité, etc.), qui permet l'hypothèse d'une relation inverse entre le degré de saillance de l'antécédent et degré explicitation nécessaire dans l'expression anaphorique : plus l'antécédent est saillant, moins l'anaphore doit être explicite, le pronom nul constituant le degré le plus faible d'explicitation. Le second facteur, le contraste, constitue la

principale contribution nouvelle de cette thèse : Comme pour d'autres niveaux d'analyse linguistique et d'autres phénomènes dans le langage, le choix de l'expression anaphorique en portugais brésilien semble être orienté vers l'efficacité. Plus précisément, lorsque l'information d'arrière-plan (« background ») et l'information assertée (focalisée) dans un énoncé contrastent, il est plus probable qu'un sujet nul soit utilisé. Les caractéristiques d'une grammaire permettant de traiter ces diverses caractéristiques est esquissée : on propose une grammaire probabiliste scalaire à plusieurs niveaux dont les contraintes sémantiques et discursives agissent en parallèle à travers un principe de correspondance probabiliste. Il est ainsi démontré que les sujets nuls sont probables dans certains contextes de co-référence discursive, puisque dans ces contextes, leurs antécédents sont plus évidents et contrastent plus avec l'information d'arrière-plan. Une contre-preuve apparente à la proposition esquissée ici est analysée : l'interprétation générique des sujets nuls. Cependant, on montre que les mêmes contraintes sémantiques appliquées à d'autres constructions génériques dans plusieurs langues peuvent produire des sujets nuls génériques en portugais brésilien, étant donné l'échec de la mise en arrière-plan prédite par l'approche proposée ici. Enfin, les résultats de trois expériences de mouvements oculaires en lecture, qui étudient l'utilisation et l'interprétation des sujets nuls et pronominaux, sont présentés. Ces résultats corroborent de façon convaincante l'hypothèse selon laquelle les sujets nuls et réalisés ainsi que leur interprétation peuvent être expliqués par la théorie proposée ici, qui les traite en termes de contraintes d'interprétation plutôt qu'en termes de légitimation syntaxique.

Mots-clés : paramètre pro-drop, résolution de l'anaphore, caractéristiques sémantiques, structure du discours, analyse de corpus, expériences psycholinguistiques.

Abstract

The present dissertation is concerned with the use and interpretation of null and pronominal subjects in Brazilian Portuguese. This investigation examines these phenomena in an attempt to disentangle the semantic and discursive factors that can be relevant for the choice between these anaphoric expressions in Brazilian Portuguese and the way in which this choice is related to the general theory of anaphora resolution. The starting point of this dissertation was the research looking into null and overt subjects from the perspective of Generative Grammar, especially the Parametric Theory. Throughout the present dissertation, however, the analyses proposed in this perspective were shown not to account for the data at stake. The generalization that poor verbal morphology is directly related to the absence or reduced frequency of null subjects, for example, is challenged through experimental data and an investigation of the relative frequency of null subjects across discourse persons in corpora. An alternative explanation presented in the previous literature, namely the importance of the antecedents' features of Animacy and Specificity, seems to better account for the attested distribution. However, this explanation is not sufficient for understanding the choice between null and overt subjects in Brazilian Portuguese, since the number of animate and specific null subjects is still relatively higher than in languages with obligatory expression of subjects. Therefore, it is argued that discourse factors seem to play a crucial role in the use of null and overt subjects in Brazilian Portuguese. The main factors identified here are Obviousness and Contrast. The first is a standard feature in the literature about anaphora resolution (expressed by a variety of terms, such as Salience, Familiarity, Accessibility, etc.), which is part of the reverse mapping hypothesis according to which the more accessible the subject is, the less explicit the co-referential form is allowed to be. The second factor, Contrast, is the main finding of the present dissertation: as is the case for other levels of linguistic analyses and other phenomena in language, the choice of anaphoric expression in Brazilian Portuguese seems to be driven by ef-

ficiency. In the present case, this means that, when the backgrounded information and the asserted (focused) information in an utterance contrast the most, it is more likely that a null subject will be used. The design of a grammar that deals with these multiple features is sketched, specifically, a multi-layered scalar probabilistic grammar is proposed, whose semantic and discourse constraints act in parallel through a probabilistic mapping. It is, thus, shown that null subjects are likely in discursive co-reference, since in these contexts their antecedents are more obvious and the focused information contrasts the most with the background. An apparent counter-example to the proposal sketched here is analyzed: the generic interpretation of null subjects. However, it is shown that the same semantic constraints cross-linguistically applied to other generic constructions can produce generic null subjects in Brazilian Portuguese, given the failure to be grounded predicted by the approach proposed here. Finally, on-line evidence for the analysis of the use and interpretation of null and pronominal subjects is provided. The results found in three eye-tracking while reading experiments provide striking evidence in favor of the proposal put forward here, according to which null and overt subjects and their interpretation can be accounted for in terms of constraints on interpretation rather than licensing.

Keywords: pro-drop parameter, anaphora resolution, semantic features, discourse structure, corpus analysis, psycholinguistic experiments.

Resumo

A presente dissertação diz respeito ao uso e à interpretação de sujeitos nulos e pronominais em português brasileiro. Esta investigação analisa esses fenômenos com o objetivo de examinar fatores semânticos e discursivos que podem ser relevantes para a escolha entre essas expressões anafóricas no português brasileiro e a maneira como essa escolha pode ser articulada com a teoria geral da resolução anafórica. O ponto de partida desta dissertação foi a literatura que analisou sujeitos nulos e plenos na perspectiva da Gramática Gerativa, especialmente a Teoria Paramétrica. Ao longo do presente trabalho, no entanto, mostra-se que a análise proposta nessa perspectiva não dá conta dos dados em questão. A generalização segundo a qual uma morfologia verbal pobre está diretamente relacionada à ausência, ou à frequência reduzida, de sujeitos nulos, por exemplo, é desafiada por dados experimentais e pela distribuição entre pessoas do discurso da freqüência relativa de sujeitos nulos em corpus. Uma proposta alternativa apresentada na literatura precedente, a saber, a importância das características dos antecedentes – Animação e Especificidade – parece explicar melhor essa distribuição. No entanto, essa explicação não é suficiente para entender a escolha de sujeitos anafóricos em português brasileiro, uma vez que o número relativo de sujeitos nulos animados e específicos ainda é relativamente maior do que em línguas com realização obrigatória do sujeito. Portanto, argumenta-se que fatores discursivos parecem desempenhar um papel crucial no uso de sujeitos nulos e plenos em português brasileiro. Os principais fatores identificados aqui são Obviedade e Contraste. A primeira é uma característica apontada em muitos trabalhos sobre a resolução anafórica (de fato, em termos diferentes, tais como Saliência, Familiaridade, etc.), que faz parte da hipótese de mapeamento reverso, segundo a qual quanto mais óbvio o sujeito é, menos explícita a forma co-referencial pode ser. A última é a principal descoberta da presente dissertação: como em outros níveis de análise linguística e outros fenômenos da linguagem, a escolha da expressão anafórica no português brasileiro parece ser voltada para eficiência e, portanto, quanto

mais as informações no "background" e as informações "afirmadas" (focalizadas) contrastam, mais provável é a utilização de um sujeito nulo. O design de uma gramática que trata dessas restrições múltiplas é esboçado: propõe-se uma gramática probabilística escalar em múltiplas camadas, cujas restrições semânticas e discursivas atuam em paralelo através de um mapeamento probabilístico. É, portanto, demonstrado que os sujeitos nulos são prováveis em coreferências discursivas, pois, nesses contextos, seus antecedentes são mais óbvios e a informação focalizada contrasta mais com o "background". Uma aparente contra-evidência à proposta esboçada aqui é analisada: a interpretação genérica de sujeitos nulos. No entanto, é demonstrado que as mesmas restrições semânticas aplicadas em outras línguas a outras construções genéricas podem produzir sujeitos nulos genéricos no português brasileiro, dada a falha em ligar a referência do sujeito ao "background", conforme a abordagem proposta aqui prediz. Finalmente, são fornecidas evidências "on-line" para a análise do uso e da interpretação de sujeitos nulos e pronominais. Os resultados encontrados em três experimentos de rastreamentos oculares durante a leitura são evidência definitiva de que sujeitos nulos e plenos e sua interpretação podem ser explicados na teoria proposta aqui, que os trata em termos de restrições de interpretação em vez de licenciamento.

Palavras-chave: parâmetro pro-drop, resolução anafórica, traços semânticos, estrutura do discurso, análise de corpus, experimentos psicolinguísticos.

Dedication

Às minhas avós Maria e Erena, à minha esposa Maria e o nosso bebê, para sempre...

Acknowledgements

Contents

Breakdown of Experimental Evidence

Experiment 1

Main Target: Verbal Marking

Methodology: Acceptability Judgment Task

Main Factors: Subject (Null or Overt); Verb (Ambiguous or Exclusive)

Experiment 2

Main Target: Animacy of the Antecedent

Methodology: Acceptability Judgment Task

Main Factors: Subject (Null or Overt); Animacy (Animate or Inanimate)

Experiment 3a

Main Target(s): Specificity of the Antecedent and Underlying Syntatic Structure

Methodology: Acceptability Judgment Task; Closed Question Interpretation Task

Main Factors: Main_Clause_Subject (Specific or Non-specific); Secondary_Predication (with_SP or without_SP)

Experiment 3b

Main Target(s): Specificity of the Antecedent and Underlying Syntatic Structure

Methodology: Acceptability Judgment Task; Closed Question Interpretation Task

Main Factors: Embedded_Subject (Null or Overt); Secondary_Predication (with_SP or without_-SP)

Experiment 4

Main Target(s): Contrast and Contexts of Use

Methodology: Acceptability Judgment Task

Main Factors: Subject (Null or Overt); Non-at-Issue Adjunct (with_NaIEl or without_NaIEl); Question_Under_Discussion (Yes-No_Answer or Wh-Unary_Answer)

Experiment 5

Main Target(s): Generics and Co-referentiality (Position of Adjuncts), Last Resort Strategy

Methodology: Acceptability Judgment Task; Closed Question Interpretation Task

Main Factors: Adjunct_Position (*In_Situ* or Fronted); Secondary_Predication_Gender (Masculine or Feminine)

Experiment 6

Main Target(s): Generics, Adjunct_Position and Contrast (in non-open polarity contexts)

Methodology: Eye-tracking Reading Task

Main Factors: Adjunct_Position (*In_Situ* or Fronted); Verb (Modal or Episodic); Subject (Null or Overt)

Experiment 7

Main Target(s): Generics, Adjunct_Position and Contrast (in object relative clauses)

Methodology: Eye-tracking Reading Task

Main Factors: Verb (Modal or Episodic); Subject (Null or Overt)

Experiment 8

Main Target(s): Generics, Antecedent Prominence and Relative_Pronoun_Interpretation (in subject free relative clauses)

Methodology: Eye-tracking Reading Task

Main Factors: Verb (Modal or Episodic); Antecedent (Subject or Non-Subject); Relative_Pronoun_Interpretation (What$_{object}$ or Those_Who$_{subject}$)

Abbreviations

AFMH Advantage of First-Mention Hypothesis

BP Brazilian Portuguese

DGB Dialogue Gameboard

EP European Portuguese

EPP Extended Projection Principle

ILH Information Load Hypothesis

MaxCoNAH Maximal Contrast Null Arguments Hypothesis

NaI Non-at-Issue

NonMaxCoNAIH Non-Maximally Contrastive Null Arguments Interpretation Hypothesis

NURC-RJ Corpus da Norma Urbana Culta da cidade do Rio de Janeiro "Urban Formal Norm Corpus from Rio de Janeiro"

ORC Object Relative Clause

PAH Position of the Antecedent Hypothesis

QUD Question Under Discussion

TT Tense Type

vYNA verbal Yes-No answer

E esse silêncio tem sido a fonte de minhas palavras. E do silêncio tem vindo o que é mais

precioso que tudo: o próprio silêncio.

"And this silence has been the source of my words. And from the silence what is more precious

than everything has been coming: the own silence."

Clarice Lispector, in Crónicas no "Jornal do Brasil"

Отношение слова к мысли и образование новых понятий есть такой сложный,

таинственный и нежный процесс души.

"The relation of the word to thought and the formation of new concepts is such a complex,

mysterious and tender process of the soul."

Лев Толстой, in Педагогические сочинения

Introduction

The present dissertation is concerned with the use and interpretation of null and pronominal overt subjects in Brazilian Portuguese (henceforth, BP). One of the main questions for linguistic theory for at least the past 40 years has been how to interpret non-pronounced elements, also called "null" constituents, and in which contexts they are allowed and/or preferred. How can one know that in a sentence such as *João₁ disse que _₁ vai sair de noite* "João₁ said that he₁'s going out at night", João₁ is the person who will go out and not someone else who might have been referred to? Throughout this dissertation, a theory of how the mechanisms of the grammar are used to produce and interpret null and overt pronominal subjects is proposed, focusing on data from BP. This dissertation brings together two different theoretical traditions that look at the same set of phenomena from different points of view: first, the theories and analyses proposed in the Generative Grammar perspective will be analyzed, since the literature about the use and interpretation of null and overt subjects in BP is mostly couched in that perspective; second, theories and analytical devices aiming to deal with cross-linguistic anaphora resolution and production will be presented and examined. Overall, the present dissertation aims to contribute methodologically, empirically and theoretically to both these traditions and to account for their findings in a comprehensive way, while dealing with null and overt subjects in a broader discourse-oriented perspective.

In the 1980s, the use of null and pronominal subjects was given especial prominence within Generative Theory with the study of so-called "empty categories" (see, for example, Chomsky 1981, chapter 4): certain languages are claimed to allow non-pronounced abstract elements to express parts of the meaning of given utterances. Cross-linguistically, it was observed that languages vary between the obligation and the optionality of the phonetic realization of certain contents that may be recovered from the context, such as in (1) and (2) in English and Spanish respectively:

(1)   John₁ said that he₁/*_₁ loves Lisa.

(2)   Eduardo₁ dijo que él₁/_₁ ama María.

       "Eduardo₁ said that [he₁] loves María."

In English, as in French and German, the use of the pronoun *he* for the co-reference with John₁ is obligatory in such a context. The absence of the pronoun makes sentence (1) strongly unacceptable. In Spanish, as in European Portuguese (henceforth, EP) and Italian, the equivalent subject pronoun – *él* "he" – can be omitted without decreasing the acceptability or the comprehensibility of

the parallel sentence given in (2).[1] In the Generative literature, this difference among languages is known as the "Pro-drop" Parameter (*i. e.*, "pronoun drop"). English is said to be a "non-pro-drop" language, while Spanish is said to be a "pro-drop" language, given the difference in the requirements of the phonological expression of subjects. Despite Chomsky (1981)'s footnote, according to which languages as Japanese can drop constituents much more freely,[2] in the first formulation of the pro-drop parameter, the main observation was that languages with rich agreement morphology are more likely to drop anaphoric and agreement-inferable subjects (see Taraldsen 1980, Rizzi 1980, Chomsky 1981, Jaeggli 1986, Jaeggli & Safir 1989, among many others).

Following Chomsky 1981's observation, Huang (1982, 1984) shows that Chinese, despite not having a rich subject-verb agreement morphology, also allows null subjects. Notice, for instance, the contrast in verbal agreement morphology in Chinese and Italian respectively. In examples (3) and (4), null subjects are used to answer questions about 3rd persons singular and plural. No variation in verbal morphology is observed in answer (3b): the bare verb *láile* "come-PRF" is an appropriate answer for a question about Zhangsan$_1$ or they$_2$. In Italian, adjustments in verbal morphology are required in the corresponding context: when the question is about two referents *Maria e Paolo* "Maria and Paolo", the verb in the answer must be in 3rd person plural – *sono arrivati* "be.PRS.3PL arrived.PL"; when the question is about one single referent Paolo$_1$, the verb in the answer must be in 3rd person singular – *è arrivato* "be.PRS.3SG arrived.SG".

(3) a. A – Zhāngsān$_1$/tāmen$_2$ lái-le   ma?
         Zhangsan/they   come-PERF Q
      "Has Zhangsan$_1$/Have they$_2$ come?"

   b. B – _$_1$/$_2$ lái-le.
              come-PRF
      "Zhangsan$_1$ has/they$_2$ have come."

(adapted from Paul 2016, 5, ex. 6)

---

[1]   Spanish is discussed in Chapter 2, following Filiaci (2010), Filiaci et al. (2013), de la Fuente & Hemforth (2013), de la Fuente et al. (2016), de la Fuente (2016), who show that there are differences between Romance standard pro-drop languages.
[2]   Namely, "The principle suggested is fairly general, but does not apply to such languages as Japanese in which pronouns can be missing much more freely" (Chomsky 1981, 284, fn 47).

(4) a. i. A – Sono   arrivati   [Maria e   Paolo]₁?

          be.PRS.3PL arrived.PL Maria and Paolo

          "Have [Maria and Paolo]₁ arrived?"

      ii. B – Sì, _₁ sono   arrivati.   / * _₁ è   arrivato.

          Yes,   be.PRS.3PL arrived.PL   be.PRS.3SG arrived.SG

          "Yes, [they]₁ have arrived / [he] has arrived."

   b. i. A – È   arrivato   Paolo₁?

          be.PRS.3SG arrived.SG Paolo

          "Has Paolo₁ arrived?"

      ii. B – Sì, _₁ è   arrivato.   / * _₁ sono   arrivati.

          yes   be.PRS.3SG arrived.SG   be.PRS.3PL arrived.PL

          "Yes, [he] has arrived."

(adapted from Paul 2016, 5, ex. 7)

Putting aside the case of languages like Chinese for now, much of the literature has been devoted to Romance and Germanic languages. In fact, the generalization about the relation between verbal agreement morphology and the use of null and overt subjects seemed robust and intuitively plausible. Based on diachronic data, previous literature have argued that Old English and Medieval French had null subjects and that as the verbal morphology gradually impoverished, the possibility of using null subjects in these languages simultaneously disappeared (see Roberts 1993, 2014, Vance 1997, Adams 1987*a,b*, *inter alia* on diachronic data from French, and Visser 1963, Traugott 1992, Rusten 2010, van Gelderen 2013, among others, on diachronic data from English).

BP has been brought into play as regards null subjects, since, according to Lira (1982) and Tarallo (1983), it is in the process of becoming "non-pro-drop" (*i.e.*, a language that does not allow null subjects). A sample of this tendency is observed in example (5), in which the attested example (5a), from the NURC-RJ corpus[3], is more natural than the corresponding example with a null subject (5b).

---

[3]   NURC-RJ is a spoken corpus, composed of different sociolinguistic interviews, college lessons and dialogues recorded in the 70s and the 90s. More details about the corpus are given in section 1.6 in Chapter 1 and in section 2.7 in Chapter 2.

(5) a. então a   gente lê   pra ele$_1$ sentado   ali... ele$_1$ gosta. . .
         So   the people read.PRS.3SG for him seat.PTCP there he   like.PRS.3SG

         "So there we read for him$_1$ when seated down and he$_1$ likes that."

                                                          (NURC-RJ, "Inquiry_011")

     b. ??então a   gente lê   pra ele$_1$ sentado   ali... _$_1$ gosta. . .
         So   the people read.PRS.3SG for him seat.PTCP there he like.PRS.3SG

         "So there we read for him$_1$ when seated down and [he$_1$] likes that."

However, BP data are not so clear with respect to this distribution, as shown throughout the present dissertation. In fact, in BP, there seem to be a number of intervening factors, such as discourse structure and semantic features, which restrict and govern null subjects, the use of pronouns and the interpretation of these elements (which will be called "null and pronominal anaphora in discourse"). Many of these factors have been reported in the literature about anaphora resolution and have been extensively studied in a variety of languages. Studies dealing with these factors with respect to BP data are however quite scarce.[4]

The general aim of this dissertation is to study these factors and to find out whether the hypothesis concerning (the change in) the structure of the grammar of BP can account for the synchronic data in both the production and interpretation of null and pronominal subjects. It is however worth noticing that this claim is not unchallenged, since the current most favored hypothesis is that BP is a "partial null subject language", whose syntactic and morphological features constrain the use of null subjects. After refuting the claim that the only reason for the use of null and pronominal subjects is the impoverishment of the verbal morphology in Chapter 1, using data from previous literature, data from a corpus studied in this dissertation, and experimental data, the main proposal is fully spelled out in Chapter 2, where the design of a grammar that can account for the data is sketched. It is claimed that a multi-layered probabilistic scalar grammar is necessary to model the synchronic knowledge of BP speakers about the null and pronominal subjects that alternate in

---

[4] Among the different definitions of the term, anaphora is defined by Huang (2000, 1) as "a relation between two linguistic elements, wherein the interpretation of one (called an anaphor) is in some way determined by the interpretation of the other (called an antecedent)" (see also Lust 1986, Wasow 1986, Huang 1994). Notice that this definition excludes one of the cases under examination in the present dissertation, namely the "generic null subject". This case is studied as an extension of the main proposal and the main topic of the present dissertation (the anaphoric subject), but it will not be referred to as "anaphoric" (nor "anaphora/anaphor"). These terms are reserved for co-referential and deictic subjects.

many similar contexts. This grammar is needed in order to model different factors playing different roles in different contextual circumstances and how they interact or not. All in all, the main conclusion is that BP discourse-anaphora grammar is designed towards efficiency and optimality of interpretation, being more economical when referring to discourse referents and more explicit to refer to less salient and less obvious referents (as in the approaches proposed by Hawkins 1994, 2004, 2014, Gibson 1998, 2010, *inter alia*). However, this claim must be correctly and carefully understood, since throughout the present dissertation apparently contradictory facts are presented. Especial attention will be given to the apparently contradictory case of generic null subjects in Chapter 3, in which the main proposal is simply that the generic interpretation emerges given appropriate semantic conditions (which are the same as in any other generic sentences) in the context of non-optimal discourse conditions for the co-referential interpretation. These claims are further supported by the results of eye-tracking while reading experiments in Chapter 4.

While proposing an exhaustive description of the use and interpretation of null and overt subjects in BP is beyond the purposes of the present dissertation, a key proposal about the design of a grammar that accounts for null and over subject in BP is put forward. Moreover, methodological and theoretical suggestions for future research are formulated here. The scope of the analysis in this dissertation is the present day system of BP, but extensions to other languages are possible. In this endeavor, it is not expected that exactly the same constraints that apply to BP data will account for data from other languages. Nonetheless, the style of the analysis, the design of the grammar and the methodology used in the present dissertation may be replicated in order to further understand linguistic phenomena and human language in general.

This study is based on corpus investigations and experimental psycholinguistic procedures. Both these methodologies are taken to be complementary in the present dissertation, because the corpus is assumed to provide access to spontaneous data in natural environments, while the experiments are focused on controlled samples, which facilitates isolating factors for testing their relevance. Many protocols have been developed while this dissertation was being produced. Before each section in which empirical data are presented, methodological procedures are described. Some introductory clarification about the methodological procedures are given below, as well as the way the term "acceptability" is understood in the present dissertation. Before these clarifications, in what follows, the present dissertation will be briefly outlined and the relevance of each

chapter with respect to the main proposal will be sketched.

Outline of the dissertation

In Chapter 1, the departure point of the present dissertation is presented and analyzed. As previously mentioned, the main claim about the use and interpretation of null and overt subjects in BP and elsewhere is that a rich agreement verbal morphology "*licenses*" the use of null subjects, while pronominal subjects are claimed not to need such a licensing mechanism, because they are independently allowed (possibly for other syntactic positions in which no relation with the verb is needed).[5] The specific claims about BP null subjects are scrutinized and a puzzle concerning the verbal paradigm of present day BP and the distribution of null and overt subjects across discourse persons is presented. Namely, as noticed in the literature about the pro-drop parameter in BP, the distribution of null and overt pronouns is not constant across discourse persons and surprisingly the lowest number of overt subjects is found in 3rd persons. Approaches that expect the structure of the verbal morphology to have an influence on the choice between overt and null subjects fail to explain this imbalance, while approaches based on the functional assumption that more ambiguous verbal forms tend to be more explicit regarding pronominal choices are also incompatible with data on the multiply ambiguous 3rd person marked verbal agreement. Based on an under-explored proposal about the direction of the change in BP, an alternative analysis is examined: it is proposed that null subjects are preferred when retrieving antecedents which are at lower levels on a natural, cognitive scale of *Referentiality*. The explanation given by this hypothesis accounts for the data in both the corpus research and the experimental evidence.

Chapter 1 ends by concluding that the reduced number of null subjects is not directly related to a problem of *licensing*, but rather of *identification* (following Rizzi 1986's null constituent theory and terminology). However, the semantics of the antecedents does not seem to be the only factor playing a role in the use and interpretation of null and overt subjects in BP. Rather, statistical analysis shows that their semantic features are indeed significant factors, but this is only a piece of

---

[5] "License" here is taken to be a very specific term, which refers to formal mechanisms that allow a sentence not to be strongly unacceptable at first sight (see Chomsky 1981, 1986, Rizzi 1986, among others). Comparing, for instance, the null subject sentences in (1) and (2), Spanish is taken to license null subjects, while English is not. More details about acceptability are given in in what follows.

the puzzle and the overall picture to be drawn is more complex. For instance, the relative number of 3rd person null subjects that are at the highest level in the Referential Hierarchy, as well as the 1st and 2nd persons (all of which are animate and specific), is still lower than in languages with obligatory overt subjects. One way to further understand this distribution is to take the mechanisms of identification of co-referential forms (anaphora resolution) and to study the contextual relation between possible antecedents and anaphoric forms. Taking contexts in which subjects are co-referent (even if deictic at the same time), the syntactic and discourse-structure factors that can drive the use of null and overt subjects are studied in Chapter 2.

A number of syntactic contexts that favor the use of null subjects have been reported in the literature about the pro-drop parameter in BP. In Chapter 2, these contexts are analyzed and a theoretical analysis to account for them is proposed. Being an issue directly related to identification of the referent of null or overt subjects, the theory and analysis of anaphoric subjects in BP can be directly studied from the point of view of anaphora resolution and production. In Chapter 2 and in the subsequent chapters, the integration of the generative approach and an anaphora-resolution approach is pursued. The starting point of this investigation is the use of null subjects in verbal answers to direct Yes-No questions in BP. This case has been reported to be a context of obligatory null subjects in BP, but it is not well-studied, although it is highly relevant and can shed light on the general distribution of null subjects. These verbal answers are taken here to be the simplest discourse and syntactic contexts in which null and pronominal subjects indeed alternate. Chapter 2 presents a review of the literature on these cases, and an alternative analysis based on discourse constraints on elliptical elements is proposed. Moreover, unary verbal answers to Wh-questions are integrated into the investigation as an immediately more complex context for null and overt subjects in BP. Basically, in both cases studied, null subjects are likely in Maximal Contrastive contexts (Maximal Contrastive Null Argument Hypothesis [MaxCoNAH]). The experimental data reported in Chapter 2 suggest that the discourse-based proposal accounts for the data in BP straightforwardly, while a syntactic-structural account would have to further stipulate constraints and additional machinery to predict the data. After having analyzed these two simple contexts, extensions to other contexts are carried out. For this extension, however, a theory of the prominence of the antecedent and how it is related to the MaxCoNAH is indeed needed. Some theories of the choice of referential expressions according to their antecedent's Accessibility are

thus presented and some of their main observations are incorporated into the hypothesis presented here. Additional statistical regressions on the corpus data taking into consideration multiple factors are reported. They suggest that a multiply factorial analysis is necessary to account for the distribution of overt and null subjects across several contexts in BP. Most of the previous research on anaphora resolution argues that the use of a given form of co-reference and its interpretation are driven by the presence of the referent in the preceding context in a certain prominent status. However, only a few articles take into consideration the syntactic and discourse status of the sentence in which the anaphoric form appears. This factor seems to be significant in both corpus and experimental data. In the present dissertation, both the discourse status of the antecedent, as an Obvious candidate, and the status of the anaphoric element, as Maximally Contrastive, are brought together to compose the optimal context for discursive co-reference. All in all, thus, it is proposed here that the main principle underlying the use and interpretation of null subjects is that they are likely to co-refer within the discourse. This principle manifests as, for example, the four graded factors identified as relevant in this dissertation: the top-most entities in each of the four scales are discourse referable and, thus, more likely to be retrieved by null subjects; while the bottom-most entities in the scales are picked up from the non-linguistic context and more likely to be referred to by overt subjects.

After pointing out some factors that may contribute to the choice between overt and null subjects in BP, the main proposal is put forth in Chapter 2. The design of a grammar that can account for multiple factors and predict how they interact is sketched and the predictions and generalizations are explained. There are three outstanding features of the proposal here that diverge from previous proposals: (i) the grammar is multi-layered in nature; (ii) the constraints are probabilistic and (iii) the mapping to/from semantics and discourse is graded (in the sense that semantic and pragmatic features cannot be analyzed in clearcut binary categories, and the boundary between one category and another is fuzzy). The main idea is that this design predicts tendencies and generalizes over constraints thought to be categorical, which in fact turn out to be violable. In order to account for some these contexts in which the constraints are violable, mainly cases where certain interpretations of null subjects are achieved in non-Maximally Contrastive utterances, a second hypothesis is put forth (Non-Maximally Contrastive Null Arguments Interpretation Hypothesis [NonMaxCoNAIH]), according to which, when the null subject is not in a Maximally Contrastive

context, it is likely to be interpreted intra-discursively.

Chapter 3 presents an extension of the proposal put forth in Chapter 2 to tackle a specific possible interpretation of null subjects in BP, namely the generic interpretation. Generic null subjects are taken to be a main defining feature of partial pro-drop languages, since either non-pro-drop or standard pro-drop resort to an overt strategy to produce this interpretation. Moreover, since the referent is not established via discourse co-reference, they could be argued to pose a *prima facie* challenge to the hypothesis pursued in the present dissertation. Previous literature on generic null subjects and its implications to the theory of co-referential null subjects are thus briefly presented and discussed in the context of the present dissertation. In general, most of the possible analyses face difficulties in dealing with the corpus data presented in the present dissertation. Some of them could be possibly applied to the data gathered from the corpus, especially those proposals which attempt to explain the generic interpretation as a last resort to save the sentence or propose *ad hoc* syntactic structures to analyze this interpretation. However, none of the approaches seem to cover the full story. The experimental data presented in section 3.7.2 suggests that the generic interpretation is indeed not preferential in the context of optimal use of co-referent null subjects, but it emerges elsewhere and surely not as a last resort: where the generic interpretation is preferential, the sentence is not significantly less acceptable than an optimal context for a co-referential null subject. In contexts in which speakers really must use a last resort to solve anaphoric co-reference, the effects on the acceptability of the sentence are much stronger than in the case of the generic null subject. Some reservations are necessary, however, since this interpretation is subject to variation among speakers. However, the experimental data show that, when speakers do not achieve this interpretation and force the null subject to be co-referential in a given context, a significant decline in acceptability is observed. This decline is seen here as a failure to ground the interpretation of the null subject in the current discourse (shared common ground). In this sense, generic null subjects are not different from other generic sentences reported in the literature. The generic interpretation of null subjects thus emerges from a failure to interpret the null subject as co-referential, which provides additional evidence for the hypothesis that co-referential null subjects are tightly related to the optimal context described by a probabilistic multi-layered grammar, as proposed in Chapter 2.

All the data from the experiments in Chapters 1, 2 and 3 constitute final judgments and inter-

pretations obtained in an off-line fashion. Also the corpus data is a post-production analysis of real time processing, interpretation and production of null and overt subjects. Chapter 4 investigates whether the factors that have been shown to affect pronominal and null subject interpretation and production in previous experiments exert their effects during on-line sentence processing. In this vein, three eye-tracking experiments were carried out and are reported in Chapter 4. In all these experiments, co-referential null subjects alternate with generic null subjects. In Experiments 6 and 7, the baseline cases were sentences with overt subjects. Experiment 6 is intended to clarify the judgments on a context whose analysis is left under-explored in Chapter 2, namely, sentences where polarity is not open. The other two experiments were designed to investigate null subjects in the contexts of the so-called cross-over effects. Experiment 7 aims to study object relative clauses in which co-referential and generic null subjects alternate with overt co-referential subjects. Experiment 8 investigates subject free-relatives. In each sentence the subject could be either generic or co-referential with a topic subject or a less prominent possible antecedent in the preceding sentence. The interpretation of the relative pronoun could be either animate (in which case it is the subject of the verb in the free-relative) or inanimate (where it is the object of the verb in the free-relative), creating a garden-path effect. The results suggest that there is no strong difference in complexity of processing, apart from the case in which the co-referential null subject is not optimal, that is, where it is not in a Maximally Contrastive context. Therefore, the hypothesis put forward in the present dissertation generalizes over all cases of increased complexity in the processing these sentences. Some additional findings in these experiments were the following. (i) Neither Topicality of the antecedent nor syntactic prominence were relevant in the processing of co-referential null subjects out of the the focus-background relation of their utterances: if the clause in which the null subject appears is not backgrounded, the co-referential interpretation is acceptable; while in backgrounded clauses, either the generic interpretation emerges or the co-referential interpretation is less easy to process, regardless of the Salience of the antecedent. (ii) If syntactic processing and licensing constraints are taken to exert effects in the early processing of sentences, it is unlikely that null subjects are licensed or subject to any kind of syntactic property in the contexts tested. Finally, (iii) garden-path effects have a longer duration in non-optimal contexts for the interpretation of co-referential null subjects.

Putting together the results of all the experiments and corpus research presented, this disser-

tation suggests that constraints thought to be structural in nature in the previous literature can be derived by semantics and discourse structure. The starting point of many analyses which attempted to account for null and overt anaphoric subjects was the impoverishment of verbal morphology. In the present dissertation, the features Animacy and Specificity take over the purported role of the impoverishment of verbal morphology in BP and explain straightforwardly the imbalanced number of null subjects across discourse persons. However, as shown throughout this dissertation, these two features are not sufficient on their own. Discourse constraints, especially related to the Max-CoNAH, are also relevant factors in the choice between null and overt anaphoric subjects. Furthermore, this same constraint seems also to play a role in the interpretation of null subjects, favoring the generic interpretation in non-Maximally Contrastive contexts, in which the co-referential interpretation fails to be grounded. When the co-referential interpretation is made implausible by the discourse context and the generic one is not available because of the semantic interpretation of the sentence, a decline in acceptability or an increase in processing complexity is consistently observed. On the whole, this dissertation suggests that a fundamental principle governs the choice between overt and null subjects in BP: null subjects are internal discourse "referers", while overt pronominal subjects are more likely to refer to elsewhere entities either deictically or in long distance co-reference. This choice is, however, graded and many natural scales of "discursiveness" are relevant. In the present dissertation four scales/factors were identified and tested significant in several, methodologically different empirical tests. It is possible that more factors can be found to be relevant, but this does not invalidate the empirical findings of the present study. The present dissertation concludes by drawing general theoretical conclusions and by stating future lines of research both methodologically and theoretically.

Methodological Considerations

The present dissertation is based on empirical observations. The main sources of data are corpus research and psycholinguistic experiments. Occasionally, corpus examples were manipulated in order to produce relevant contrasts in interpretation and acceptability. However, as a methodological principle, I have avoided creating contexts that are not observed in corpora or experimentally tested.

Examples, interpretations and acceptability judgments reported in the previous literature are

reproduced here, mostly when they are close to undisputed. Dubious cases are tested through corpus investigation and in experiments. Overall, the data provided here aims to be empirically based and original. The sentences tested in Experiments 1-8, for instance, have not been tested in the literature about BP before, and are rarely tested in other languages.[6] Unfortunately, there was not enough time to test every relevant context. However, this dissertation opens a path for further investigation beyond what is reported in the last page of Chapter 4. Additional experiments and data checking will be required to broaden the scope of the results provided here and to test the present theory and analysis against a larger empirical domain.

Acceptability and Interpretation

Acceptability (often referred to as "Grammaticality") is sometimes considered a binary judgment on a stimulus (see the reviews Myers n.d., Schütze & Sprouse 2013, Schütze 2016, and the abundant literature cited there). However, it is clear that in many cases acceptability rates vary in degree from the acceptable to the unacceptable cases. Although in many theories, these degrees are claimed to be the consequences of performance or processing effects, in the present dissertation they are taken to reflect the in-depth grammatical knowledge of speakers, as explicitly proposed in Chapter 2. Starred examples throughout this dissertation are not thought to represent absolutely unacceptable sentences, but less acceptable sentences in contrast with a more acceptable option. Starred interpretations are not intended to represent an impossible interpretation either, but rather a bias toward another possible interpretation, which is possibly more natural for most speakers (where "most" is understood as a relevantly larger number and, when statistically tested, a significantly larger number, rather than 50% plus one).

A decision was made to present and analyze mostly the raw data, rather than z-scores or other conventional rescaled data.[7] This decision was couched in both theoretical and methodological

---

[6]   All experiments in this thesis were realized as Latin Square designs. For each experiment, lists were created corresponding to the number of conditions in each experiment, such that each participant saw the same number of items in each condition but never the same item in more than one condition. The number of items was adjusted to each experiment with at least three items per condition. The number of participants varies considerably from one experiment to the other due to participant accessibility at the time of the experiment. The experiments are not presented in the chronological order according to which they were run.

[7]   The suggestion of normalizing the data using z-scores was stressed out to me by Scott Schwenter in his *pré-raport*, since it is possible that not all participants in the experiments use the whole scale of acceptability judg-

reasons: methodologically, it has been pointed out that (i) z-scores, for example, can create or inflate certain effects for a given speaker who does not really notice that there is a difference across conditions and (ii) difference among speakers are accounted for in the mixed-effects models, which include Items and Participants as random Factors (see in this regard Myers n.d. and literature cited there); theoretically, as proposed throughout the present dissertation, acceptability varies in degree as multiple constraints are violated: since, in most experiments, only one constraint violation is tested (keeping the other factors at the top of the scales of acceptability as much as possible), not strongly declined rates due to unacceptability are expected; rather the raw data presented corroborate the hypotheses put forth here according to which each factor contributes individually to the degree of acceptability of a given sentence; the fact that they are significant (or not) shows that they exert an effect (or possibly not), but they do not turn the sentences completely unacceptable unless multiple constraints are violated at the same time (see also Chapter 2, especially section 2.8).

Two disclaimers about the phenomenon at stake in the present dissertation are required. First, the use of null subjects is taught at school and can be tightly related to the level of education of BP speakers, since contact with older stages of BP and with EP, in which the relative frequency of null subjects is higher, is established throughout years of formal education (see Magalhães 2000, Laperuta 2004). Second, in recent research proposals, it has been observed that there is regional variation in the acceptability and in the interpretation of null subjects in BP.[8] The systematic study of this sociolinguistic variation remains to be done. In the present dissertation, the group of speakers who took part in Experiments 1 to 8, as well as the speakers recorded and analyzed in the corpus research, are taken to be more or less homogeneous. Beyond this, most of the corpus results re-

---

ments when performing the task, and thus (ii) they may actually not be using the same scale when responding the questionnaires. I agree that this can be the case, but such a divergence in performing this task is accounted for in the statistical analysis, by taking "Participants" as a random Factor. Jeffrey Runner also pointed out that, in the analysis of Experiment 1, the model might have been under-powered and a z-score transformation might make the effects clearer. However, since the result of Experiment 1 goes in the same direction suggested in the analysis of the corpus data, I kept the analysis of the raw data as evidence of a weak effect (see section 1.7.2).

[8] I have recently become aware of the existence of a research project, called *Condições de interface em dependência sintáticas* "Interface conditions in syntactic dependencies" developed in the LAPAL (*Laboratório de Psicolinguística e Aquisição da Linguagem* "Laboratory of Psycholinguistics and Language Acquisition" PUC-RJ – Pontifical Catholic University of Rio de Janeiro), which aims to identify dialectal variation in the acceptability of null subjects across regions in Brazil. As the description of the project suggests, in some regions null subjects seem to behave as syntactic anaphors, while in others they pattern as overt pronouns. Psycholinguistic experiments will be carried out to identify such differences.

ported here are similar to those found in other regions of Brazil (see Negrão 1990, for São Paulo, Laperuta 2002, for Londrina, among others), if not in absolute numbers, at least in the systematic variation across factors.

Chapter 1

The Impoverishment of Verbal Inflection, the Semantics of the Referents and the Morphology-Semantics Interface

1.1 Introduction

This chapter aims to study the interaction of overt morphological verbal markings and semantic properties of referents in BP as regards their effects on the realization of overt and null subjects. It is assumed here, as a starting point for the discussion, that there are two general approaches to pro-drop in BP, one diachronic, the other synchronic, which can be taken as complementary to each other: (i) there is an ongoing diachronic change in this parameter towards the obligatory phonological realization of subject pronouns (Lira 1982, Tarallo 1983, among others); and (ii) at this synchronic stage, BP is a partial pro-drop language (Holmberg et al. 2009, Biberauer et al. 2010, *inter alia*). When BP is compared to other standard pro-drop Romance Languages (for instance, EP, Spanish and Italian), the linguistic contexts in which null subjects are allowed are indeed scarcer (Duarte 1995, Kato 1999, Barbosa, Duarte & Kato 2005, *inter alia*). In (6a) below, for example, the overt realization of the pronominal subject seems obligatory. These restrictions on the use of null subjects are taken to be a by-product of the impoverishment of morphological markings in BP, illustrated by the verbs in the relative clauses in (6b) and (6c) below.[1]

(6) a. O   que o Rafa$_1$ ia   fazer com aquele livro... se ??(ele$_1$) estava
       what   the Rafa go.IMP.3SG do.INF with that   book   if (he)   be.IMP.3SG
       morto... se ??(ele$_1$) ESTÁ   morto?
       dead   if (he)   be.PRS.3SG dead
       "What was Rafa$_1$ going to do with that book... if he$_1$ was dead... if he$_1$ IS dead?"

   b. _$_2$ não eram   coisas soltas que ??(você$_3$) escreveu.
         not be.IMP.3PL things loose that (you)   write.PST.3SG
       "Those$_2$ weren't aleatory things that (you$_3$) wrote."

   c. _$_2$ eram   coisas que ??(a gente$_4$) viveu   junto,   coisas que ??(a
         be.IMP.3PL things that (the people) live.PST.3SG together things that (the
       gente$_3$) escreveu   junto.
       people) write.PST.3SG together.
       "These$_2$ were things that (we$_4$) lived together, things that (we$_4$) wrote together."

                                                        (OPUScorpora, "Entre Nós")

---

[1]   "OPUSCorpora" is a multilingual parallel corpus, constituted of movie subtitles (see Lison & Tiedemann 2016 for further information). This corpus is not extensively studied in the present dissertation, but occasional examples are used here to illustrate some descriptive points. This corpus was studied in a previous paper (Correa Soares 2016), where most of the examples used here were gathered.

Although the interpretation of the subjects in the relative clauses in (6b) and in (6c) are "you" and "we" respectively, the verbal inflection is the same for both the examples and is parallel to (6a), *i. e.* 3rd person singular. In all these examples, the omission of these subject pronouns would produce less acceptable sentences. However, in (6b) and (6c), the sentences start with perfectly acceptable null pronouns, which, in the context from which these data were extracted, refer to *anotações e memórias* "notes and memories". This sort of examples can be used to justify the *ongoing* change and the *partial* pro-drop status of BP. And this is where the problem arises. The cases above show both null and pronominal subjects. Though in present day spoken BP some configurations do not allow null subjects, in other constructions and linguistic contexts null subjects are permitted or even preferred. The above cases in the beginning of (6b) and (6c) illustrate a context in which the null subjects are preferred but not obligatory, although, at least in the dialect of the southern region of Brazil, the use of the corresponding pronoun *elas* "they.FEM" is much less acceptable. The main aim of this chapter is to start setting up the relevant properties which can favor the use of null pronouns in given contexts. The preference differences illustrated in (6) will be discussed throughout the present dissertation.

In this chapter, the main proposal is to compare two apparently contradictory hypotheses put forth in the recent literature about syntactic change and partial pro-drop in BP: (i) BP is on the way to becoming a non-pro-drop language because of the impoverishment of verbal overt morphology, the so-called "Taraldsen's generalization" (Taraldsen 1980, Rizzi 1986, Roberts 2014, Simonenko et al. 2015, 2017) and (ii) the contexts where null pronouns continue to be preferred in BP are those where their referents are lower in a natural, cognitive scale of salience, called "Referential Hierarchy" (Cyrino et al. 2000, Kato et al. 2006, Kato & Duarte 2014).[2] The apparent contradiction lies in the fact that these hypotheses predict different distributions: (i) the morphologically unmarked verb form 3rd person singular is syncretic with the 2nd person singular and 1st person plural with *a gente* (literally, "the people") and in some tenses with the 1st person singular too, and the 3rd person plural is also syncretic with the 2nd person plural, being thus predicted to be preceded by more overt subjects; and (ii) the 3rd persons are the *loci* for less referential subjects,

---

[2]   Throughout this dissertation, the proposal based on the Referential Hierarchy is further discussed in the light of the theory of the Accessibility (Ariel 1990, 1994, 2001).

such as indefinite, non-specific and inanimate referents, *i. e.* they are expected to be more likely to be dropped according to the Referential Hierarchy. Indeed, as shown throughout this chapter, 3rd discourse person null subjects are much more frequent than in 1st and 2nd discourse persons, even though the 3rd person marking covers a broader range of discourse persons. In this chapter, these two hypotheses will be compared, and the relative weight of each one in the current system of BP will be worked out.

Outline of Chapter 1

This chapter is organized as follows. In section 1.2, the verbal inflection and subject pronoun systems in present day BP are presented. In section 1.3, some proposals that attempt to address the use of null subjects either cross-linguistically or specifically in BP are briefly summarized. In section 1.4, the facts about BP are analyzed and used to test both the typology and the theory of the pro-drop parameter, as well as the general theory of anaphora resolution, as regards the inherent semantic features of the antecedents. In section 1.5, a summary of the literature is presented. In section 1.6, corpus data addressing the main question of the chapter (the discourse persons and the semantic features of the antecedents) are presented. In section 1.7, the results of 4 experiments designed to test whether the verbal inflection and the semantic features of the antecedents are relevant in the use and interpretation of null subjects in present day BP are reported. Finally, section 1.8, discusses the impact of the data presented in this chapter on the proposals of the previous literature and paves the way for the further investigations that will be carried out in the following chapters.

1.2 The System in Present Day Brazilian Portuguese

BP has substantially modified its inflectional and pronoun systems when compared to previous stages of this language, to other varieties of Portuguese or to other Romance Languages (Italian and Spanish, for instance) (see Duarte 1993, 1995, Kato 1999, Galves 1987, 1993, 1997, 1998, Kato & Negrão 2000, *inter alia*). In this section, the organization of the BP verbal inflectional system and the pronoun paradigm of subject position in present day BP are presented. In subsection 1.2.1, it is shown how the verbal paradigm is organized in present day colloquial BP. In subsection 1.2.2,

the new organization of the subject pronominal system in BP is introduced. Finally, in subsection 1.2.3, the interaction of the use of null subjects with regard to both new paradigms is presented.

1.2.1 A New Verbal Paradigm in Brazilian Portuguese

Table 1.1 below summarizes the paradigm of verbal inflectional markings in present day BP. In Table 1.2, the verbal paradigm of older stages of BP is presented for the sake of comparison (these forms were used in the beginning of the XIXth century).[3]

Table 1.1 – The Verbal Paradigm in Present Day BP

Person and Number	TT1 Verb (present)	TT2 Verb (imperfect past)
1st singular	*falo*	*falava*
2nd singular	*fala*	*falava*
3rd singular	*fala*	*falava*
1st plural (a gente)	*fala*	*falava*
1st plural (nós)	*falamos*	*falávamos*
2nd plural	*falam*	*falavam*
3rd plural	*falam*	*falavam*

Table 1.2 – The Verbal Paradigm in Older Stages of BP

Person and Number	TT1 Verb (present)	TT2 Verb (imperfect past)
1st singular	*falo*	*falava*
2nd singular	*falas*	*falavas*
3rd singular	*fala*	*falava*
1st plural	*falamos*	*falávamos*
2nd plural	*falais*	*falavais*
3rd plural	*falam*	*falavam*

As shown in Table 1.1, the present day colloquial BP presents three forms for present tense and two for imperfect past tense if, for the 1st person plural, only the form *a gente fala/va* (literally "the people speak/spoke"), which is typical of colloquial speech and triggers 3rd person singular

---

3   See the project Tycho Brahe (available at http://www.tycho.iel.unicamp.br/ tycho/prfpml/fase2/index.html) for an overview of the multiple changes in BP during the past two centuries.

agreement, is taken into consideration. The main difference between this paradigm and older stages of BP, written prescriptive BP and EP is that 2nd person singular and plural markings (-*s* for singular and -*is* for plural), as well as 1st person singular (-*mos*), are standard in those varieties.[4] The impoverishment in verbal morphology of present day BP is well reported in the literature (see Duarte 1993, 1995, Kato 1999, Galves 1987, 1993, 1997, 1998, Kato & Negrão 2000, among others). However, it is rarely pointed out that there is variation in the 1st person singular according to the verbal tense. There is a unique, distinctive verbal marking of 1st person singular in the present (*falo*) and past perfect (*falei*) tenses and "compound" tenses which involve the present and past perfect tenses of the auxiliary verbs "ir" and "ter"(*vou falar* and *tenho falado*, for example); all other tenses are not exclusively marked for 1st person singular, as in, for example, imperfect past (*falava*), conditional (*falaria*), periphrastic conditional (*ia falar*) in indicative mode, present (*fale*), future (*falar*) and imperfect (*falasse*) in subjunctive mood and all the remaining compound forms with the auxiliary *ter* "have" (*tinha falado*, *teria falado*, etc). The first group of inflected verbs will be called here TT1 [tense type 1] or "exclusive" and the latter, TT2 [tense type 2] or "ambiguous"/"non-exclusive". In the next subsection, the possible agreement combinations between subject pronouns and inflected verbs are presented.

1.2.2 A New Pronominal System in Brazilian Portuguese

Simultaneously to the change in the verbal paradigm, BP also changed with respect to the pronominal system, as shown in Table 1.3 below. This system is more complex than the verbal paradigm, since the interaction between morphosyntactic characteristics and semantic features of the referents of the pronouns presents many mismatches, shown in the second and third columns.[5]

---

[4]   The forms in older stages of BP shown in Table 1.2 mostly overlap with EP and prescriptive written BP. It is sufficient to know that, despite a few exceptions (e. g., syncretic forms), there exists one exclusive marking for each person and number.

[5]   Most of the verbal forms in present day BP are syncretic, that is, as shown in Table 1.3, they can combine with more than one discourse person. For historical reasons, I will refer to the syncretic verbal forms as 3rd person singular, even if it can be used with, for instance, 1st person plural (*a gente*) and 2nd person singular (*você*), and as 3rd person plural, even when referring to 2nd person plural (*vocês*).

Table 1.3 – The Subject Pronominal System in Present Day BP

Pronoun	Discourse Person and Number	Person and Number Agreement	Gender Agreement
eu	1st singular	1st singular	Masc/Fem
você	2nd singular	3rd singular	Masc/Fem
(*tu*)	2nd singular	3rd/2nd singular	Masc/Fem
ele	3rd singular	3rd singular	Masc
ela	3rd singular	3rd singular	Fem
a gente	1st plural	3rd singular	Masc/Fem
(*nós*)	1st plural	1st plural	Masc/Fem
vocês	2nd plural	3rd plural	Masc/Fem
eles	3rd plural	3rd plural	Masc
elas	3rd plural	3rd plural	Fem

As observed in Table 1.3, the discourse and grammatical person and number do not match for the pronouns *Você* (and *Tu*) "You.SING", *A gente* "the people/we" and *Vocês* "You.PL". Besides, the 1st and 2nd person singular and plural can be combined with both grammatical genders – masculine or feminine.

Table 1.3 is based on properties of agreement "close" to the subject. For example, as regards the pronoun *a gente* "we", only the verbal agreement marker and the reflexive SE were taken into consideration. When using the form *a gente* (literally "the people") to refer to discourse 1st person plural, the verb inflection and the reflexive SE marker are the same as in the 3rd person singular, as in *A gente₁ se₁ viu.* "The people SE see.PST.3SG" (meaning "*We saw ourselves/each other.*"). The use of a 1st person plural verb or a reflexive makes the sentence less acceptable, as in *A gente₁ *nos₁ *vimos.* "The people ourselves see.PST.1PL". However, when using a long-distance co-referent form, the anaphoric pronoun can be *nós* "us" – the 1st person plural pronoun – as in *A gente₁ viu uma cobra atrás de nós₁.* "We saw a snake behind us" (see the discussion about these properties in Lopes 1999, Menuzzi 2000, Taylor 2009).

It is worth observing in Table 1.3 that, in the dialects that preserve the pronoun *tu* "you" for 2nd person singular, in the colloquial register, verbal inflection is 3rd person singular, but the anaphoric form of the reflexive marker is *te* "yourself" (2nd person singular reflexive).[6] For this reason, it can

---

6   According to Sergio Menuzzi, younger generations are starting to accept *se* as a reflexive for the subject *tu* –

be considered a mixed form, that can be accounted for in a grammar that has a multiple types of agreement (see, for instance, Dowty & Jacobson 1989, Pollard & Sag 1994, Kathol 1999, Chung 2000, Wechsler & Zlatič 2001, *inter alia*). However, in the present dissertation, the form *tu* is not taken into consideration due to its low frequency of occurrence in the corpus research reported here and to the fact that none of experiments had sentences using it.

It should also be noted that this pronominal system could be enriched with further types of pronouns that can be used instead of these standard forms. For instance, *o(s) senhor(es)/a(s) senhora(s)* "the lord(s)/the lady/ies" can be used instead of *você(s)* "you.SG/PL" in formal contexts. Discursively and grammatically these forms exhibit exactly the same patterns as *você(s)*. Another case that could be incorporated into Table 1.3 are the demonstratives *isto/isso* "this" and *aquilo* "that", which are the preferred forms to co-refer to an event or other higher order entity (such as a proposition). These pronouns are not akin to any other form in the paradigm, since they exhibit a neuter semantic feature, although they are grammatically compatible with masculine grammatical gender. However, they also have adjectival gender and number marked forms *este(a)(s)/esse(a)(s)* "this/these" and *aquele(a)(s)* "these/those", which are more restricted in their occurrences in corpora, as mentioned in section 1.6.

1.2.3   The Null Subject in Present Day Brazilian Portuguese

The null subject in current spoken BP can replace any of the pronominal forms described in section 1.2.2 above, given appropriate conditions. Kato (1999), for instance, has identified four contexts in which null subjects are quite frequent in present day spoken BP, three of which are relevant to the present investigation. According to her, BP presents null expletives (which are beyond the scope of the present dissertation, as they do not replace any pronominal forms), has null subjects with an arbitrary generic interpretation (a case discussed in Chapter 3), null bound subjects (example 7a), and "anaphoric" embedded null subjects (examples 7b and 7d).[7] However, according to her,

---

unlike *se* bound by *eu*.

[7]   As Jeffrey Runner pointed out, the fact that BP allows null expletives might be a topic to be analyzed in the present dissertation. Although the use of overt expletives is a matter of controversy in previous literature, non-argumental null subjects in meteorological predicates and other standard contexts are acceptable and frequent in corpus. The first point to be stressed out is that primary interest here are the referential subjects (deictic and co-referential). Second, this fact by itself makes the claim that BP is not a non-null subject language (which does

the 1st and 2nd persons (examples 7b and 7d respectively) and the referential 3rd are almost never null. Some other researchers have reported a number of uses in the contexts of verbal answers to yes-no questions, such as in (7c-ii) (see Kato & Tarallo 1993, Urbano et al. 1993, Magalhães & Santos 2006, Holmberg 2016, for instance). In these well-known cases, the null element can indeed substitute any person and number, as (7) shows.

(7)   a. todo mundo$_1$ que entra   [no   curso] sabe   o   que   _$_1$
          all   world   that enter.PRS.3SG in.the course know.PRS.3SG that which
          quer   fazer.
          want.PRS.3SG do.INF

          "Everybody$_1$ who enrolls for [the course] knows what they$_1$ want to do."

      b. Eu$_1$ não sei   se _$_1$ tô   fugindo um pouco do   assunto,
          I   not know.PRS.1SG if   be.PRS.1SG escaping a   little from.the topic

          "I$_1$ don't know if I$_1$'m going off topic a little ..."

      c. i. A – Bem, e   depois, pra ir   pra faculdade, você$_1$ fez   algum
                  Well and after   to go.INF to faculty   you   DO.PRS.3SG any
              curso depois que _$_1$ concluiu   o   técnico?
              course after   that   finish.PST.3SG the technical

              A – "Well, and after, to go to the college, did you$_1$ attend any course after [you$_1$]

              finished the technical one?"

         ii. B – _$_1$ Fiz.
                       do.PST.1SG

              B – "Yes. (=I attended.)"

      d. quer dizer, tudo bem você$_1$ aprender como, como _$_1$ calcula   [uma laje]
          want say,   all   well you   learn.INF how   how _$_2$ calculate.PRS.3SG a   slab
          "that is, it's ok for you$_1$ to learn how _$_1$ to calculate [a rooftop]."

(NURC-RJ, "Inquiry_ac_1")

In (7), the null subject replaces different antecedents regardless of discourse person. Given proper contexts, such as the embedded co-referent (as in 7a, 7b and 7d) or a positive verbal answer (as in 7c-ii), null subjects are thus perfectly acceptable in present day BP.[8]

---

not license null subjects altogether) absolute senseless. Finally, if null subjects are to be accounted for in the present dissertation, the Referential Hierarchy would predict this cases neatly, since expletives are in the lowest level of this hierarchy, as explained in what follows (see especially section 1.3.5 below).

[8]   In much previous literature, it has been claimed that cases such as (7a) are an especial kind of construction

Many researchers have reported, however, that the change in BP has not equally affected each discourse person. Negrão (1990), based on a corpus of oral production in a public school in São Paulo, first pointed this out. Duarte (1993, 1995, 2000), based on the study of popular written plays, shows an asymmetry across discourse persons over the period in which BP has become different from EP. Duarte's hypothesis is that the impoverishment of the inflectional paradigm, shown in Table 1.1, intrinsically related to the new pronoun paradigm and its agreement properties, shown in Table 1.3, along with the deactivation of the "Avoid Pronoun Principle", has caused the decline in the number of null subjects from 80% in the second quarter of the XIXth century to 26% in the 90s (see Duarte 1993, *inter alia*);[9] however, she points out that this decrease is far more drastic in the 1st and 2nd person than in the 3rd person, as summarized in Figure 1.1 below.

Figure 1.1 – Percentage of Null Subjects according to Discourse Person
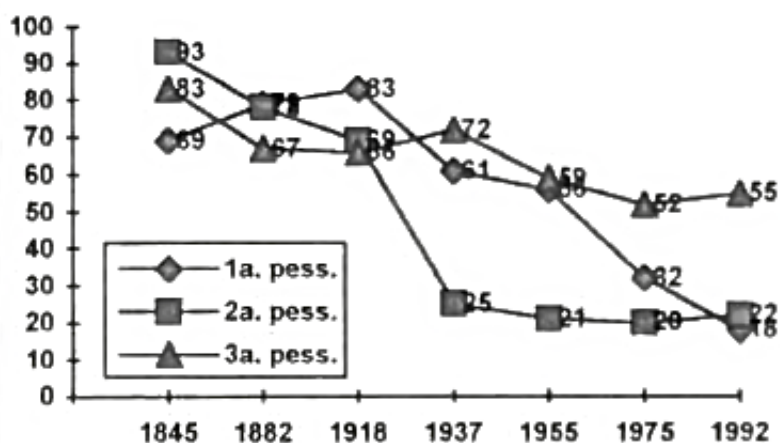
(Duarte 1995, p. 20)

This distribution poses a problem for some theories of change in the "Pro-drop Parameter" (see Biberauer et al. 2010, among others), since the 3rd person is the most ambiguously marked verb form, which has historically covered the 2nd and, in some cases, the 1st persons.[10] Moreover, it is

---

(bound variable) in which null subjects are obligatory in some languages (Spanish, for instance). In the present dissertation, these constructions are not taken to have an especial status (or to be derived by an especial syntactic device), since it could be accounted for by the general principles proposed in the present dissertation and, in some of the examples presented here, no c-command relation (which is claimed to be a condition for bound variables) is established.

[9] The Avoid Pronoun Principle is a principle which states that pronouns should not be used whenever they are not required (Chomsky 1981).

[10] As Jeffrey Runner has pointed out, clearly the use of null subjects with 2nd person has dropped precipitously

unlikely that there is a higher proportion of 3rd persons subjects in the syntactic contexts described by Kato (1999) as favoring null subjects, or in verbal answers, which also favor null subjects. Yet this idea has been claimed to account for the higher number of null subjects in 3rd person. This hypothesis will be considered later in this chapter.

According to Duarte (1995, p. 48), based on a study on the same corpus used in the present dissertation, the impoverishment in the inflectional/pronominal paradigm does not produce gradual and uniform effects across each discourse person. Rather, the process seems to be much faster in the 1st and 2nd persons than in the 3rd person. However, her claim that this is a by-product of the on-going change and that this imbalance is a trace of the previous status of BP as a pro-drop language does not explain why such an asymmetry appears, so that an additional assumption would be required. Notice, moreover, that the additional hypothesis put forth by Duarte (1995), according to which the "Avoid Pronoun Principle" is deactivated in BP also does not explain the relative distribution of null subjects across discourse persons, and rather suggests that there should be some kind of symmetry. If this principle is somehow formulated in terms of redundancy, it can be tested in both corpus and experiments with the exclusive marking in 1st person singular, as shown in what follows (see Fernandez-Soriano 1989, in this regard).

1.3   The Null Subject and the Discourse Person Puzzle

The pro-drop parameter, Taraldsen's generalization and the use of null subjects have drawn much attention in the linguistic literature. In this section, some previous accounts that are directly related to the discourse person puzzle studied in this chapter are summarized. After summing up the historical importance of this parameter and the role of BP in this discussion in subsection 1.3.1, the relation between the impoverishment of verbal inflectional markings and the null subject is presented in subsection 1.3.2. In subsections 1.3.3 and 1.3.4, some approaches aiming to deal with

---

since 1918. This coincides with the reduction (and virtual disappearance) of inflectional morphological marking in verbs in 2nd persons. However the drop in the relative number of 1st person, which clearly happened from 1955 on, may not be associated with any impoverishment, since at least the 1st person singular has maintained the same inflection patterns. What would be expected by an impoverishment-based approach to this change is a fall in the relative number of null subjects in the 3rd person, which has not taken place, as shown in what follows.

the imbalance between the number of null subjects across discourse persons are summed up. In section 1.3.5, an alternative proposal is presented, aiming to explain the imbalance in the relative percentage of null and overt subjects in terms of semantic features of the antecedents.

1.3.1   The Pro-Drop Parameter Historically

The discussion about how to account for null subjects and the (im)possibility of using them in a given language goes back to the early 80s (Taraldsen 1980, Rizzi 1980, Chomsky 1981, Rizzi 1982*a*). By comparing English, as a non-null-subject language, and Italian, as a standard null-subject language, many researchers proposed that a rich agreement morphology allows speakers of languages such as Italian to omit phonologically overt subject pronoun, since such languages have a distinctive form for almost each person-number combination in every tense. However, also in the 80s, this claim was weakened by researchers working on Chinese and similar languages (such as Huang 1984, 1989, *inter alia*), which lack distinctive inflectional markings but allow null subjects. Despite such cases, the generalization about the impoverishment of the agreement paradigm and the absence of null subjects appears robust across many languages (see Holmberg 2005, Roberts & Holmberg 2010, Roberts 2016, for overviews). Not surprisingly, diachronic research has also given evidence across periods of time in favor of this correlation, especially as regards the change from Middle (pro-drop) to Modern (non-pro-drop) French (see the discussion posited by Adams 1987*a*,*b*, among others).

BP has played a major role in this discussion since the 80s, when Lira (1982) and Tarallo (1983) claimed that it was undergoing a change in the pro-drop parameter. They found only around 20% of null subjects in their corpus researches. During the 80s and 90s, much research further contributed to this claim, suggesting that BP is a "live sample" of a parametric change (Duarte 1993, 1995, Galves 1987, 1993, 1997, 1998, Kato & Negrão 2000, among many others). In the 2000s, another claim about BP was added to the general discussion about the pro-drop parameter: BP along with some others languages (for instance Finnish, Marathi and Russian, according to Holmberg et al. 2009, Biberauer et al. 2010) were claimed to be examples of "partial pro-drop languages" (see Rodrigues 2004, Holmberg 2005, Holmberg et al. 2009, Biberauer et al. 2010). As already mentioned, these two ideas converge into two fundamental and complementary claims about the

pro-drop parameter in BP: (i) diachronically BP is on the way to becoming a non-pro-drop language, supposedly like English or French (Duarte 1993, 1995, 2000, Galves 1987, 1993, 1997);[11] and (ii) synchronically BP is a partial pro-drop language (Rodrigues 2004, Figueiredo-Silva 2000, Holmberg 2005, Holmberg et al. 2009, Roberts & Holmberg 2010, Biberauer et al. 2010).[12] Along with each one of these claims, different theories and analyses have been proposed, approaching the facts and the data with two different strategies: (i) diachronically researchers have been trying to quantify and qualify the contexts in which null subjects are still present in BP as a trace of previous stages of the language; and (ii) given the partial pro-drop character of BP, researchers have been trying to identify and isolate the morphosyntactic contexts in which null subjects are still possible or obligatory as exceptions that can be listed or explained by other principles. However, a study that combines both strategies, going from the systematic study of specific constructions to a quantification of their occurrence or vice-versa, is necessary. In the present study, the use and interpretation of null subjects is approached using both strategies, although no diachronic analysis is put forward here.

1.3.2   The Impoverishment of Verbal Inflection

In general, the reason that has been claimed to be crucial in the change of BP toward non-pro-drop status is the impoverishment of verbal inflectional morphology, as described in Section 1.2. The recent literature proposes two possible explanations as to how this impoverishment has affected the possibility of having a null subject in BP: (i) because it is no longer distinctive and strong, the inflectional marking does not have the proper features to license the null pronoun in the subject position (Duarte 1995, Holmberg 2005, Holmberg et al. 2009); (ii) differently from standard pro-drop languages, in which the inflectional marking is the subject and satisfies the syntactic requirements of the sentence, the inflectional marking in BP is no longer sufficient to check the features of sub-

---

[11]   The status of French pronouns is not undisputed, see for example Legendre et al. (2010), who see them as verbal affixes, which makes French a pro-drop language.

[12]   As Sergio Menuzzi has pointed out to me, though these two positions do not exclude each other, they were not put forward as compatible in the literature and a precise way in which they might be a single combined diachronic and synchronic hypothesis has not been proposed. Specifically, researchers defending the "partial pro-drop" analysis have mainly argued for this position against the idea that BP was becoming a non-null subject language of the English or French type.

ject position, such as the EPP feature, and an overt pronoun is obligatory to fill the position and check the relevant features (Kato 1999, Barbosa et al. 2005).[13] Assuming either of these positions as such supposes that the requirements apply across the board. That is, it is not expected that any person is more likely to be expressed by overt pronouns than any other.

1.3.3 Not all null subjects are born equal

A few researchers propose an account for the observation that the number of null subjects does not decrease uniformly in many languages. Holmberg (2005), for example, proposes that null anaphoric pronouns – that is, the co-referential 3rd person – in partial null-subject languages, including Finnish (and possibly BP), are D[eterminer]-less ϕPs, as in consistent pro-drop languages; whereas the 1st and 2nd persons are full DPs which are deleted.[14]

However, Finnish is different from BP as regards the fact that in Finnish the 1st and 2nd persons are optionally null in all environments while the 3rd person can only be null when bound by an antecedent in the next clause up (Holmberg 2005, Vainikka & Levy 1999). In BP, on the other hand, the 1st and 2nd person overt subjects seem to be mostly required (or at least more frequent than 3rd person subjects). As in BP, though, 3rd person agreement is less marked than 1st and 2nd person agreement in some tenses and moods (see Holmberg et al. 1993, Holmberg & Nikanne 2002): (i) 3rd person is null in the past tense and in the conditional mood; and (ii) in many varieties of colloquial Finnish there is no distinction between 3rd person singular and plural. So, exactly as in BP, in Finnish there is variation in the number of distinctive markings according to the tense and mood of the verb and according to the level of formality; and the 3rd person is the least distinctive form. However, as shown in the previous section, 1st and 2nd person null subjects in BP are quantitatively much less frequent than 3rd person null subjects, while in Finnish these persons are easier to omit (unfortunately Holmberg 2005 does not provide quantitative information and I am not aware of any research which does). It is possible to suppose, though, that the 1st and 2nd person

---

[13]   The EPP feature is the Minimalist Program version of Chomsky (1981, 1986)'s Extended Projection Principle, according to which every sentence must have a subject.

[14]   Holmberg (2005)'s main evidence comes from Finnish expletives. Expletives are, however, beyond the scope of the present dissertation. On the other hand, his discussion of the importance of morphological verb marking is summarized in this section. Further details of his proposal, which can be found in the source, are not relevant for the purposes of this section.

null subjects in BP are like the 3rd person subjects in Finnish, and 3rd person subjects in BP are like the 1st and 2nd person null subjects in Finnish. This hypothesis predicts the following: (i) since the 1st and 2nd null subjects would have to value the feature "Agr", they would require inherently valued phi-features ("interpretable" features in the Minimalist Theory); however, they would lack the substructure required for a definite, referential category; as a consequence, null subjects would not be able to refer to an individual or group, neither independently/deictically nor under co-reference with an independently referring DP and, as a last resort, they could be generic. And (ii) the 3rd person null subject in BP would be a DP which could be deleted given certain requirements (as proposed by Barbosa 2011). Assumption (i) is inherently implausible, since the BP 1st and 2nd null subjects can evidently refer to individuals. Assumption (ii) needs more careful consideration.

1.3.4 Null subjects are null NPs

Barbosa (2011) puts forth a hypothesis based on assumption (ii) and on the fact that BP passes all of Li & Thompson (1976)'s diagnostics for being considered a topic prominent language. She points out that Modesto (2008*a*,*b*), in a comparative study of BP and Finnish, argues that the definite/anaphoric null subject in Finnish and BP is in topic position — *i. e.*, it is a null topic in the spirit of Huang (1984). Under her "null NP anaphora" approach, 3rd person null subjects are bare NPs and, when they raise to a topic position, the individual (definite/anaphoric) reading becomes available (see Kuno 1973, for arguments that topichood signals definiteness in Japanese), and then the definite/anaphoric interpretation of the bare NP requires the application of the operation of Type Shifting to an individual. She suggests that this operation is triggered under Agree by T bearing a D-feature (as in Holmberg 2005). This hypothesis will be analyzed throughout this dissertation, and evidence that shows that this cannot be an explanation for BP data will be provided.

1.3.5 Null Subjects and the Features of Antecedents

The literature reported so far tends to approach the null subject in BP from the point of view of isolating the constructions in which it is still present as "exceptions" somehow allowed by the synchronic grammar. A different perspective is taken by Cyrino et al. (2000), Kato et al. (2006), Duarte, Mourão & Santos (2012), Kato & Duarte (2014), Duarte (2015), who try to explain the

asymmetry in the use of null subjects in BP across different persons and numbers in diachronic terms. The authors propose that the change that is taking place in BP is governed by a "Referential Hierarchy", as shown in (8). According to this hierarchy, languages tend to use overt pronouns for picking up more referential entities, that is, those that are higher in the hierarchy.[15]

(8) Referential Hierarchy

non-argument	proposition	[-human]	[+human]

3rd 2nd 1st person

-specif   +specif

[-ref] <————————————————————————————> [+ref]
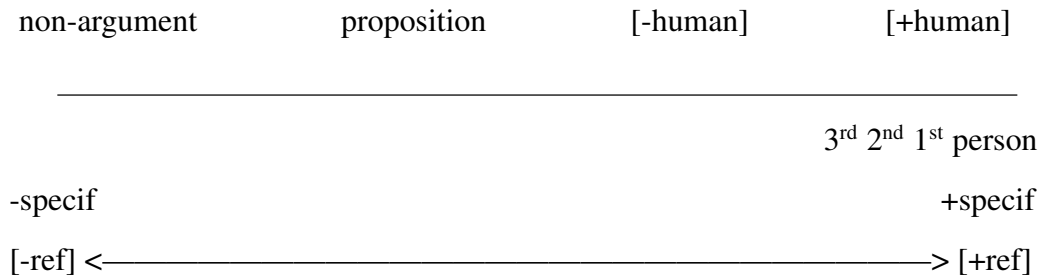
(9) The Implicational Mapping Hypothesis:

a. The more referential the subject is, the greater the possibility of it being expressed by a non-null pronoun is.

b. A null variant at a specific point on the scale implies null variants to its left in the Referential Hierarchy.

(Cyrino et al. 2000, p.39)

Both parts of the implicational mapping hypothesis in (9) shed some light on the distribution of null subjects in BP. In fact, since BP allows null subjects for any person and number, as pointed out in section 1.2, assumption (9b) is straightforwardly true about BP.[16] Assumption (9a) is, though, more counterintuitive despite being also more predictive regarding the data studied in the present dissertation. It predicts that more referential subjects are likely to be expressed by overt pronouns. In the literature on anaphora resolution, for example, the correlation is taken to be that the more accessible/salient the antecedent, the less explicit the anaphoric element needs to be, with null items

---

[15] According to (Cyrino et al. 2000, 59), "[+N +human] arguments are the highest in the Referential Hierarchy, while non-arguments [(expletives)] are the lowest. For pronouns, since the speaker (*eu* 'I') and the addressee (*você* 'you') are inherently human, first and second person pronouns are the highest in the hierarchy, while the third person pronoun which refers to a proposition is the lowest, with the [-animate] entity in between. The feature [± specific] interacts with all the other features."

[16] Cyrino et al. (2000) analyze other languages in which assumption (9b) makes stronger predictions. Since these languages are not at stake here, the reader is advised to consult the original paper for further information.

being the least explicit (see section 1.4.2 below). This leads to predictions diverging from those made by the Referential Hierarchy. Specifically, much literature assumes that less referential antecedents (e.g. higher order entities, less specific entities, etc.) are inherently less accessible/salient than first order entities, including humans, and specific entities, so that you would expect more null subjects for the latter than for the former (Gundel 1988, Gundel et al. 1993, Ariel 1990, 1994, *inter alia*).[17] See section 1.4.2 below for further discussion. Assumption (9a), however, converges with the idea advanced by Negrão & Müller (1996), Menuzzi (2002), Müller (2003), according to which BP has an especial function for null pronouns. Following these authors, the null pronoun behaves much like a reflexive (a syntactic anaphor), being identified when bound by operators as a bound variable.[18] One of the pieces of evidence for this claim are cases in which a null subject designates a non-specific antecedent, that is, a co-reference with a less referential entity, as predicted by hypothesis (9a).

1.4   Null Subjects Elsewhere

This section aims to present the facts about pro-drop in BP and to compare them to what is reported in previous literature about other languages. Especial attention is given to unexpected factors that seem to be at stake in BP, but are apparently contradictory with what is observed in other languages. In this vein, null subjects are presented from a parametric and typological point of view. The possibility of using null subjects in a given language and how null subjects are interpreted has been the focus of much research. In fact this topic is one of the most widely studied parameters of cross-linguistic variation. However, the very status of this parameter remains highly controversial (as pointed out by Biberauer et al. 2010). The discussion here is however limited to the cases that are either unexpected or contradictory when compared to BP. For the sake of exposition, this

---

[17]   Notice that the Accessibility/Salience based hypotheses make no predictions for impersonal subjects, since they do not have an antecedent (so that it makes no sense to talk about their antecedent's Accessibility); on the other hand they allow for more fine grained distinctions between antecedents located at the same position in the Referential Hierarchy: for instance, one of two human referents can be more accessible than the other (if it was more recently mentioned, for example).

[18]   Negrão & Müller (1996), Menuzzi (2002), Müller (2003) diverge in the details, which can be decisive in terms of predictions. The concept of bound variable and how it is implemented in the grammar can be different in these papers. The readers are referred to the original papers for details.

section is split into two parts: in subsection (1.4.1), the discussion covers the possibility of using null subjects in certain languages; in subsection (1.4.2), given this possibility, the study of the different interpretation preferences of anaphoric subjects is discussed.

1.4.1 The Parametric Theory and Typological Distribution

The first typological descriptions of the use and interpretation of null and overt subjects come from traditional grammars. For instance, in the beginning of the XXth century, before the appearance of modern linguistics, Bennett (1895), Greenough & Allen (1888/1903), Gildersleeve & Lodge (1905) observe that the subject in Latin can be omitted when the verb makes explicit the referent of the subject. In a similar vein, linguists like Jespersen (1922) also claimed that Latin, Italian, Hebrew, Finnish, among others, may not express the pronominal subject, unless it is necessary to stress the reference of the subject. In fact, these observations come from the ancient Greek-Roman grammar tradition, where occasional references to this idea are found: Apollonius Dyscolus, in the IInd century AD, pointed out that, in Greek, the "nominative" can be understood through the verb, and that it is definite in the 1st and 2nd and indefinite in the 3rd, because of the unrestricted number of possible referents (Discole 1997). In the context of this initial description of the facts, BP thus appears to be an unexpected case, since, when the subject can be understood through the verb (in, for instance, the case of exclusively marked 1st person singular), the number of null subjects is lower, and, despite having unlimited possible referents in the 3rd persons, 3rd person null subjects are far more frequent than 1st and 2nd person null subjects.

The study of the pro-drop parameter was incorporated into the Generative tradition seminally by Perlmutter (1971), based on the null subject languages Spanish, Italian, Serbo-Croatian, Arabic, Hebrew, Hausa, Warlpiri and Basque. The distinction between pro-drop and non-pro-drop languages was, for at least some languages and some approaches, claimed to be a difference as to how this position is filled: a pronoun in subject position which is either deleted or unpronounced in null-subject languages or obligatorily pronounced in non-pro-drop languages (see, e.g., Chomsky 1981). The main assumption is that, based on English facts, every sentence should have a subject. However, as observed by Dryer (2013) on a study about 711 languages, languages in which a pronoun in subject position is obligatory are quite infrequent. In other words, the analyses of other

languages looks underlyingly like English, despite the evident superficial differences. Van Valin & LaPolla (1997, 331) propose that, in languages where the verbs themselves can convey the expression of pronominal subjects, the affixes are the real subjects of the clause, even in clauses in which there is a separate nominal subject. As mentioned in section 1.3.2, Kato (1999), Barbosa et al. (2005), Barbosa (2009) propose a similar account for the European Portuguese facts, although Holmberg (2005), Roberts (2010) explicitly ruled this possibility out. Note that regardless of the theoretical problems involved in this line of analysis in one or another framework, it does not explain the imbalance in the distribution of null and overt subjects in BP. Moreover, the very existence of a language such as BP challenges the work of the typologist, since the classification of such a language is not clear at all (in Gilligan 1987 for instance, BP is classified as a language that allows co-referential null subjects across the board).

Even historically, BP has been a challenging case. While in some articles similarities between BP and Medieval French have been observed (Duarte 1993, 1995, Kato 1999, Barbosa et al. 2005), explicit claims against this comparison have recently emerged (Kaiser 2006, 2009, Roberts 2014, Correa Soares 2016). The link between the impoverishment of verbal morphology and the disappearance of null subjects in French has also long been studied in much previous research (e. g. Ewert 1943, Vennemann 1993, Rizzi 1986, Adams 1987*a*,*b*, Alexiadou & Anagnostopoulou 1998, Roberts 2010, Sheehan to appear), but recently it has also been questioned by Roberts (2014), Simonenko et al. (2015, 2017), who show that quantitatively this generalization does not hold straightforwardly from written data collected over many centuries of linguistic change.

In recent years, the study of different types of pro-drop has been advanced. As early as the 80s, the correlation between rich agreement morphology and the parameter pro-drop was called into question (see Huang 1982, 1984 and Gilligan 1987, for a overview of this discussion in the 80s). Basically, the existence of languages in which co-referential null subjects are permitted with verbs whose agreement morphology is completely absent, such as Alyawarra, Angami, Burmese, Caviteno, Garo, Indonesian (Betawi), Japanese, Malagasy, Malayalam, Mandarin, Murut, Rao, Sre, Thai, Vietnamese and Yessan-Mayo, challenges the parametric coocurrence of rich verbal morphology and null co-referential subjects. Huang (1982, 1984) proposed that not only languages with rich verbal agreement morphology allow null subjects, but also languages that lack verbal morphology altogether. This idea has been reformulated a number of times in terms of rich-

ness of agreement morphology (Rizzi 1982*a*), morphological uniformity (Jaeggli & Safir 1989), undistinctive morphology (Rohrbacher 1994) and syncretism (Rizzi 2002). However, mixed systems, such as that of BP, always create a problem for such approaches, and certainly counting the number of syncretic forms, as proposed in some of the literature, does not provide a reasonable account for the observed data. Other languages also show some imbalances across tenses in terms of agreement. Russian has different forms for the three persons and two numbers in present tense, but only gender and number in past tense (see Muller 2005, among others). Hebrew has the full paradigm in future and past, but not in present tense (see Berman 1980). Finnish also does not seem to be uniform concerning the verbal paradigm (Rodrigues 2004, Holmberg 2005, *inter alia*). As mentioned before, languages with such an imbalance are taken to be "partial" pro-drop languages, since they seem to allow null co-referential subjects in some contexts but not in others. These cases are apparently counterexamples for the theory of parameters, which is claimed to be binary in most of its implementations.

On the basis of the issues arising from the large number of different languages involving subtly different realizations of the pro-drop parameter, Newmeyer (2005, to appear) calls into question the explanatory power of the parametric theory. Based on Gilligan (1987) and the evident typological discrepancies across languages, Newmeyer (2004, 2005) claims that the parametric correlation between richness of agreement morphology and the null realization of subjects (among other parametric correlations) cannot be sustained. Roberts & Holmberg (2010) agree with some of the observations made by Newmeyer (2004, 2005), but simply suggest that the data is rather inconclusive. In an attempt to integrate these various divergent patterns into the framework of the parametric theory, Biberauer et al. (2010), Roberts & Holmberg (2010), Holmberg et al. (2009), Holmberg & Phimsawat (2015) posit a more complex notion of pro-drop, according to which languages are be distributed into several (sub-)groups. The notion of "partial" pro-drop emerges from this proposal: while Chinese and Japanese are claimed to be "radical" pro-drop languages and Italian, Spanish and EP are said to be "standard" pro-drop languages, Finnish, Russian, Marathi and BP are hypothesized to be "partial" pro-drop languages, whose main defining features are (i) pronominal subjects can be null under specific conditions determined by both morphological and syntactic contexts and (ii) the generic interpretation of subjects is achieved by a covert strategy. As discussed throughout the present dissertation, however, if these are the defining features

of a "partial" pro-drop language, BP may not be required to fall under this classification, since the morphological and syntactic constraints proposed in the literature can be explained in terms of semantic and discourse constraints on the interpretation of anaphoric subjects, and the generic interpretation is taken to be a by-product of this theory. These claims fit in well with what is proposed by Newmeyer (2004, 2005, 2006, to appear), according to whom the parsing-based theory of Hawkins (1994, 2004, 2014) accounts neatly for many of the generalizations that, in the past, had been attributed to differences in parameter settings.

It is not the main aim of this dissertation to discuss the plausibility and the coverage of the theory of parameters. However, it is clear that BP provides a fundamental piece of evidence to test this theory. Throughout this dissertation, the position taken as starting point is that the grammar is concerned with efficiency. In this sense, the approach pursued here is close to Hawkins' and Newmeyer's reasoning: "Many patterns of interaction, even in grammars, can be explained when viewed from an efficiency and ease of processing perspective" (Hawkins 2014, 202).

The starting point of the investigation here are the theories that take languages to be what languages are intended to be: meaningful devices for efficient communication. In this regard, BP might provide counter-evidence to this reasoning, given the facts presented so far regarding the distribution of null and overt subjects across discourse persons. The fact that 3rd person and, if the hypothesis of the Referential Hierarchy presented in the previous section holds, non-specific and inanimate antecedents are the most likely to be null could be interpreted as a potential problem for a number of proposals based on the claim that obvious and salient referents are retrieved by less specific forms. Some of this research is summarized an discussed in section 1.4.2 below.

1.4.2 Different Subjects, Different Anaphora Resolution

Previous research has found that less referential antecedents are relatively more frequently retrieved by null subjects than by overt pronouns in BP (Cyrino et al. 2000, Berlinck et al. 2009, Duarte, Mourão & Santos 2012, Kato & Duarte 2014, *inter alia*). This fact could be viewed as counter-evidence for a standard assumption in the literature about pronoun resolution (Kuno 1972, 1978, Halliday 1967, Halliday & Hasan 1976, Chafe 1976, Prince 1981, Gundel 1988, Ariel 1990, 1994, 2001, Gundel et al. 1993, Grosz & Snider 1986, Grosz et al. 1995, Walker & Prince 1996, Al-

mor 1996, 1999, Gordon et al. 1993, Carminati 2002, among many others): it is generally assumed that anaphora resolution is guided by a reversed mapping principle between antecedent Salience and anaphor Explicitness – more salient antecedents are retrieved by less complex and less informative anaphoric forms. However, it turns out that BP is only an apparent counter-argument to this principle. To see this, it is necessary to clarify the notion of antecedent Salience. In previous research, different definitions and applications of the idea have been proposed, which ultimately lead to contradictory predictions. In what follows, a brief summary of the idea of antecedent Salience is given. The notion of Salience is split into two different subtypes: discourse Salience and inherent semantic Salience. In this section, the focus is on the latter. Lately, in Chapter 2, these two subtypes will be compared and incorporated into a coherent theory in order to understand how they can be seen to describe a single principle.

To make a long story short, for more than 60 years anaphora resolution has been described in terms of a preference for using less specified forms to refer back to the most salient, familiar or active antecedents in linguistic theory, philosophy, computer science and psychology (Lewis 1970, 1979, Lyons 1977, Mitkov 1999, 2002, Garnham 2001, Büring 2005, *inter alia*). Evidence in favor of this hypothesis has come from very different sources, from corpus research to experimental studies (see McEnery 2000, Garnham 2001, among many others, for overviews). Earlier functional approaches observed that in English, for example, unstressed personal pronouns are more likely to refer to the subject of the preceding sentence than full NPs, as shown in (10).

(10)   Adele went to the university library.

      a. She urgently needed a book for her exams.

      b. This institution was one of the oldest in the country.

(Falk 2014, ex. 2)

In her particularly influential book of 1990, Ariel proposes the following detailed relationship between the Complexity/Explicitness of the referential form and the Accessibility of the antecedent, which she calls the Accessibility Marking Scale (Ariel 1990, 73).

Figure 1.2 – Ariel (1990)'s Scale of Accessibility from Arnold (1998)

	Marking Scale	Examples
Low accessibility	Full name + modifier	Joan Smith, the president
	Full ('namy') name	Joan Smith
	Long definite description	The tall and authoritative president
	Short definite description	The president
	Last name	Smith
	First name	Joan
	Distal demonstrative + modifier	that hat we bought last year
	Proximal demonstrative + modifier	this hat we bought last year
	Distal demonstrative + NP	that hat
	Proximate demonstrative +NP	this hat
High accessibility	Distal demonstrative	that
	Proximate demonstrative	this
	Stressed pronoun + gesture	SHE (plus gesture)
	Stressed pronoun	SHE
	Unstressed pronoun	she
	Cliticized pronoun	(no examples in English)
	Extremely High Accessibility Markers	gaps, including pro, PRO and *wh*-traces, reflexives, and Agreement

As shown in Figure 1.2, Ariel (1990)'s Scale of Accessibility fairly accommodates example (10): the intuitively most salient referent *Adele* in sentence (10) is retrieved by an unstressed, lower ranked pronoun in continuation (10a); in the second continuation in (10b), a less prominent, distant and less salient antecedent is recovered by a proximate demonstrative+NP, which is higher in the hierarchy for being more Explicit and Complex.

As for the BP facts, the four lowest levels are at stake. Basically, Ariel (1990) states that null pronouns are less Complex and Explicit than overt pronouns. So far so good: this claim seems absolutely uncontroversial and finds resonance regardless of the approach (see, for instance, Cardinaletti & Starke 1994 in Minimalist Theory and Burzio 1991, 1996, 1998 in Optimality Theory,

for similar ideas).

As for the semantic-discourse definition of Salience, the literature is highly controversial. In a nutshell: the different notions of Salience of the antecedent make different, sometimes, contradictory predictions. In what follows, different notions of Salience are briefly presented. It is shown that, if Salience is taken to be a scale of cognitively innate referentiality and BP null subjects are biased toward lower levels of referentiality, the data would challenge the predictions borne out by the Accessibility Theory; but, if Salience is understood as within the discourse, as proposed by Ariel (1990), the preferences established for BP null subjects do not contradict the inverse relation hypothesis. Ultimately, BP is a key piece of evidence for understanding the linguistic notion of Salience, which might well turn out to be different from the general cognitive notion of Salience.

Previous research has studied semantic features of the antecedent as relevant for anaphora resolution. Most of them propose a hierarchy such as (11) below, taken from Silverstein (1976):

(11) HUMAN > ANIMATE > INANIMATE

(Silverstein 1976)

Dahl & Fraurud (1996) show that this hierarchy is a strong predictor for the choices between pronouns and NPs in a Swedish written corpus study: humans are more likely to be pronominalized, while non-humans are more likely to be retrieved by NPs or demonstratives. Bittner (2007) and Gagarina (2007) crossed Animacy with the Syntactic Function of the antecedent in several interpretation experiments with children in German and Russian respectively. Both studies tested null subjects (ungrammatical in German), personal pronouns and demonstratives. Some of the main findings are that (i) crossing the semantic feature of animacy with the syntactic feature of grammatical function, there is no unified notion of prominence that can be proposed, (ii) younger children seems to be more sensitive to animacy, while older children seem to rely more on syntactic salience; and (iii) there is a primacy of syntactic salience over animacy salience in these languages in the contexts that were tested, since the less complex forms were significantly biased toward the subject, although animacy also exerted effects on many conditions and age groups. The results found for German are difficult to compare to the BP facts, since German does not allow null subjects, but the results for Russian are especially interesting and will be addressed in the General Discussion of this Chapter. However, it is clear that younger children tend to use animacy as the

basis for Salience, while older children bias their interpretation according to more "linguistic" features (e. g. subjecthood).

Based on Montalbetti (1984), some psycholinguistic research regarding the specificity of the antecedent has also been carried out. Carminati (2002) tested the bound variable behavior of null subjects which co-refer with quantified antecedents.[19] She found a correlation between the use of null subjects in Italian and their interpretation as bound by quantified subject antecedents (as proposed by Montalbetti 1984). In a more theoretical non-experimental perspective, Menuzzi & Lobo (2016) make similar considerations about Spanish and Portuguese. According to them, in Romance languages null subjects tend to co-refer with a non-specific antecedent in c-commanding configuration, especially when the antecedent is in subject position. In section 1.6, this observation is extended to more syntactic configurations using corpus research. In section 1.7, these predictions are tested in BP. Before advancing on this topic, some considerations about the notion of Salience are due.

Many researches have pointed out that the null forms seem to be biased toward discourse internal antecedents, while overt forms are freer to refer either deictically or to discursively less salient antecedents (see Mayol 2010, *inter alia*). In this sense, the notion of Salience proposed by Ariel (1990) accounts for speaker's linguistic knowledge about anaphora resolution. According to her, the Salience of the antecedent is defined according to its degree of discursiveness, cf. Figure 1.3 below:

Figure 1.3 – Ariel (1990)'s Salience of the Antecedent

Most accessible   Linguistic context
        ↓   Physical context
Least accessible   Encyclopaedic context
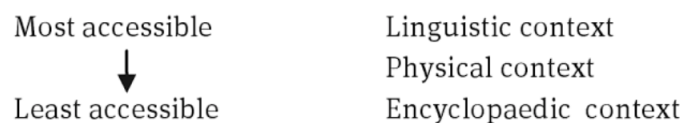
In section 1.8 of this chapter, it will be argued that the relevant notion of Salience is discursively defined. In Chapter 2, the scales of Animacy and Specificity based on Ariel's proposal will be presented as regards the design of a grammar that accounts for the data presented here. So far, two

---

[19]   The behaviour of null subjects when refering back (or bound) by quantified antecedents is a matter of much discussion in the literature (see, among others, Grodzinsky & Reinhart 1993, Heim 1998, Reuland 2001).

ideas must be borne in mind: (i) there are multiple levels of Salience that a given antecedent can be subject to: some are inherent to the semantics of the referent (Animacy, Specifity, for instance), which will be discussed throughout this chapter, and some are relative to the sentence in which the referent occurs (Subjecthood, Topicality, Centrality, among others), which will be addressed in Chapter 2; and (ii) different notions of Salience lead to different and perhaps contradictory predictions concerning the inverse correspondence principle of anaphora resolution.

In the next two sections, new empirical research will be presented, which aims to investigate the relation between inflectional verbal morphology, the semantic features of the referential subjects and the use of null and overt subjects in BP. In section 1.6, the results of a corpus investigation are reported. These results suggested ideas for a set of experiments, the results of which are reported in section 1.7.

1.5 Summary of the Literature

In Table 1.4 below, a summary of the relevant literature briefly reviewed so far is provided, along with their predictions concerning the data which will be analyzed. This first section of literature review aims to quickly set out the problems. More detailed comments about some of the approaches are made in the discussion of the results in each chapter and in specific sections when the discussion is relevant.

Table 1.4 – Summary of the Literature

Foundations	Main Representatives	Proposed Devices	Languages at stake	Predictions about BP
Government and Binding / Parametric Variationist Theory / Diachronic	Duarte (1993, 1995, 2000)	Avoid Pronoun Principle / Taraldsen Generalization	BP (compared to older stages and to other Romance languages)	BP tends to avoid null subjects altogether.
Minimalist Program / Parametric Theory	Kato (1999), Barbosa et al. (2005)	feature checking [EPP]/ deficient AgrP	BP and European Portuguese (compared to French)	3rd person null subjects are due to *ad hoc* logophoric binding.
Minimalist Program / Parametric Theory / Synchronic Partial Pro-drop Approach	Holmberg (2005), Holmberg et al. (2009)	ϕP deficient constituent / DP deletion	Finnish (and BP and Marathi)	The imbalance across persons is produced by different syntactic mechanisms.
Minimalist Program / Parametric Theory / Synchronic Partial Pro-drop Approach	Barbosa (2009, 2011)	Null NP-Anaphora / Covert Topic Movement	BP and European Portuguese (compared to Chinese)	Topical subjects are null NPs.
Parametric Variacionist Theory / Diachronic	Cyrino et al. (2000)	Referential Hierarchy and Mapping Hypothesis	BP (compared to Finnish)	Lower degrees of Referentiality are slower regarding the change in BP.
Parametric Theory (Partial Pro-drop Approach)	Biberauer et al. (2010), Holmberg et al. (2009)	partial pro-drop / isolating morphosyntactic contexts for null subjects	cross-linguistic	Isolated morphosyntactic contexts allow null subjects.
Typological Theory	Gilligan (1987), Dryer (2013)	–	cross-linguistic	BP is pro-drop.
Anaphor Resolution / Cognitive Approach	Ariel (1990), Silverstein (1976), Dahl & Fraurud (1996), Carminati (2002), Gagarina (2007)	Scales of Complexity of Referential Forms and Scales of Salience	English, Danish, Italian, Russian, and others	Higher Number of Null Subjects with Animates; Specificity coincides with other parameters of Salience

1.6   Reassessing Corpus Data

In this section, the results of reassessing a corpus of oral interviews, called NURC-RJ, are reported. This corpus was previously analyzed by Duarte (1995). A new analysis was carried out for two main reasons: (i) the criteria used to exclude some data in previous research seemed too restrictive, as for instance no "contrastive" subjects were taken into consideration, and ended up excluding cases that for the purposes of the present dissertation are crucial to understanding the nature of the null subject in BP; and (ii) with new theories and analytical toolkits, such as new statistical packages and automatic annotators, more relevant factors and correlations might be discovered. Nine interviews carried out in the 70s and nine interviews from the 90s (of which twelve were with the same person during the two relevant periods, that is, six people participate twice), were analyzed. Overall 8032 inflected clauses in which the subject was either co-referential or generic (the same criteria used in Duarte 1995, cf. p. 36) were gathered. The null expletive, such as in existential or presentative clauses, was then excluded. However, what Duarte (1995, p. 37) calls "*ser* 'to-be' existential constructions" were taken into account here and were carefully analysed in what follows.

1.6.1   Methodological Considerations

The corpus research carried out here obeyed several coarse-grained, independent steps in order to guarantee unbiased results. Firstly, eighteen texts were selected from the NURC-RJ corpus. The criterion used to selected these texts was the availability of the audio file, so that in dubious cases the primary source of the data would be available to check (about 0.5% of the data was double-checked in such a fashion). The second step was to automatically annotate the transcriptions (also available on the NURC-RJ website). The automatic parser VISL (Visual Interactive Syntax Learning) was used to fully annotate the eighteen texts.[20] In these texts, the sentences with null subjects

---

[20]   More precise information about the annotation tool is available on the website of the project: http://visl.sdu.dk/visl/about/. This tool is part of the project *Floresta Sintá(c)tica* ("Syntactic Treebank"), a collaboration between Linguateca and the VISL project. It contains texts in Portuguese (from Brazil and Portugal) annotated (analyzed) automatically by the parser PALAVRAS (Bick 2000) and reviewed by linguists.

were manually gathered.[21] Afterwards, the sentences with full pronominal forms were collected through searches on each form (*eu*, *tu/você*, etc). The decision was then made to begin the investigation using the criteria proposed as relevant in the literature which can be unambiguously evaluated, including "verbal inflectional form", "antecedent distance". By focusing on these criteria, this study avoids commitment to any specific theory and theory dependent criteria such as "c-command", "unaccusativity", etc. These data were descriptively analyzed in qualitative and quantitative terms. Finally, a inferential analysis was carried out using logistic regressions with the glmer function of the LanguageR package in R applying the logit linking function.[22]

1.6.2 Analysis

Based on the claim that the impoverishment of the inflectional paradigm is the trigger for the decrease in the number of null subjects, the analysis started by splitting the cases according to the standard classification by discourse person of the subject, following Table 1.1. This was followed by a classification of the verbal inflection according to its "T[ense]T[ype]" and Grammatical Person and Number, taking into consideration only the current explicit agreement markers (cf. Table 1.1). Finally, the semantic features of the antecedents were analyzed by sorting them into specific vs. non-specific, sentential vs. non-sentential and animate vs. inanimate, based on the Referential Hierarchy in (8).

For practical reasons, when dealing with whether the verb has an exclusive marking for 1st singular, incidental lexically ambiguous past verb forms (such as *soube* "know.PST.1SG/3SG", *disse* "say.PST.1SG/3SG", *quis* "want.PST.1SG/3SG" and *trouxe* "take.PST.1SG/3SG") were not considered to be part of the ambiguous verb tense group, because they only affect a very small number of cases (< .1% of the data).

Secondly, animacy is clear in BP. Animals are taken to be animate in BP, since they can be combined with almost any predicate typical of animate humans, such as intentional predicates, as *morder* "to bite", or sentience predicates, such as *sentir* "to feel". As regards the distinction

---

[21] I collected the clauses with null subjects on my own. Each clause was further sorted according to twelve independent criteria, as explained in what follows. That means that each case was re-analyzed twelve times (guaranteeing quality control check).

[22] See Bates & Maechler (2009), Bates et al. (2011, 2015) for details.

between sentential vs. non-sentential, most of the predicates attributed to a co-referring sentential subject are those which can only be applied to a constituent that corresponds to a (finite or not) clause in the previous context, such as *o que não pode é [funcionário público formar sindicato]₁. [Isso]₁ é proibido.* "*what is not allowed is that civil servants create a union.₁ That₁ is forbidden.*".[23]

The definition of Specificity raises problems. As shown in the literature on the topic (Enç 1991, Abbott 1995, von Heusinger 2002, Kagan 2006, von Heusinger 2011, Falco 2002), despite the notion's intuitive simplicity, it is difficult to come to a consensus on a formal definition of specificity, and it is certainly beyond the scope of this dissertation to attempt to do so. For the present purposes, the following operational criterion will be used.

(12)   Operational criterion for *Specificity*:

Given a NP denotation $\parallel \alpha \parallel$ in a predicate $\parallel \beta \parallel$, the denotation of $\parallel \alpha \parallel$ is *specific* iff:

$$\forall x \in \parallel \alpha \parallel [\parallel \beta \parallel (x)] \quad \rightarrow \quad \neg \diamond \exists y \in \parallel \alpha \parallel \neg [x \otimes y] \wedge [\parallel \beta \parallel (y)]$$

If for any individual *x* that belongs to the denotation $\parallel \alpha \parallel$ such that the proposition $\parallel \beta \parallel$ applies to *x*, it is not possible that there is at least one individual *y*, which does not overlap *x*, and

---

[23]   As pointed out by Philip Miller, the criterion used here to distinguish sentential and non-sentential antecedents is not strictly semantic. Indeed, the fact that a higher-order entity is encoded into a clause makes a strong difference in the way this entity is recovered in by an anaphoric subject in BP. In this respect, BP seems to have a distinction as that in English anaphoric subjects below (examples and judgments are Philip's).

(1)   Null pronouns have been hypothesized to be disappearing in BP. This/??It has turned out to be false.

(2)   The hypothesis that null pronouns are disappearing in BP was put forward by X. This/It has turned out to be false.

(3)   Null pronouns have been hypothesized to be disappearing in BP. Though this hypothesis has been often repeated in the literature, it/??this has turned out to be false.

What happens in English is that it is not simply the semantic type of the antecedent that comes into play but also the form with which it is expressed. In all of these cases, the relevant referent is the same abstract entity, namely a hypothesis, so the semantic properties of the antecedent cannot differentiate between the three cases. When the antecedent is expressed by a clause, as in (1), it cannot be referred to using "it", a demonstrative is required. When introduced by a complex NP as in (2), both the demonstrative and "it" are possible. When it has been referred to by a simple NP, as in (3), "it" is definitely preferred. Something similar is happening in BP. It is worth highlighting, however, that it does not contradicts the primary decision taken here, according to which the classification should not be theory-internal or theory-dependent.

the proposition $\| \beta \|$ also applies to *y*. In (12), a mereological definition of *individual* is assumed, and also of *completely overlap* (see Link 1983, Krifka 1998*b*, *inter alia*). This definition is a simplification of what was proposed by Kagan (2006). It classifies as non-specific some quantifiers and quantified NPs (such as *[tudo] foi penhorado.* "*Everything* was pawned"), some indefinite NPs (such as *então, [uma pessoa que tá querendo fazer eletrônica] vai ter cálculo.* "Then, *a person who wants to do Eletronics* will have Math."), some mass nouns (such as *[barulho como existe no Rio de Janeiro] eu acho que dificilmente se encontrará em outra cidade* "Noise as it exists in Rio de Janeiro I think that one will hardly find in any other city.") and some plural NPs (such as *há [professores] dando doze horas diárias* "There are *teachers* giving twelve daily hours [of classes]). It also includes as non-specific certain predicates that trigger a non-specific reading for NPs due to tense or modality, as in, for example, *[O indivíduo] tinha até uma escala profissional.* "The individual had even a professional scale". It also classifies negative NPs, such as *[ninguém] quer pensar nisso* "nobody wanna think about this", as non-specific, because they falsify both the sides of the conditional definition. But it does allow for some negative quantified NPs, such as *eu não peguei [nenhuma dessas professoras]* "I haven't had *any of these teachers*," to be specific. It also excludes "generic uses" of 1st and 2nd person, such as in *Se [tu] vai casar, pede um empréstimo.* "If *you*'re gonna marry, *you* ask for a loan" from the set of specific referents. On the other hand, it correctly includes proper names, some definite descriptions (*Conheço mais ou menos [o sindicato dos professores]* "I know more or less *the teacher's union*"), indefinite NPs, such as *tenho parente inclusive em essa situação, que é [um indivíduo que trabalhava com mecânica de automóveis]* "I have a relative in this situation, who is *an individual that worked on auto mechanics*", and quantified elements, such as in *[todos os cursos que anunciavam no Diário de Notícias] receberam a comunicação de que tinham que comparecer lá.* "*all the courses that advertise in the Diário de Notícias* received the notification that [they] had to go/attend there.".[24]

---

[24] Certainly to take a binary operational criteria is not the optimal way to understand the effect of Specificity on the choice of overt and null subjects. As explained in what follows, many issues are related to the controversial notion of Specificity (which includes the fact that bound variables can be understood as non-specific, but have an especial status; the degrees of specificity; whether the interpretation of a quantified NP can be taken to be specific or not; e. g.) can be risen. However, different degrees of Specificity are left for further research. Here, the aim is to study and establish whether non-specific and specific in a first approximation can exert any influence in the use and interpretation of overt and null subjects in BP.

1.6.3 Results

The results in the present study are generally compatible with those found in previous literature (Negrão 1990, Duarte 1995, e.g.). Overall, a slight preference for overt over null subjects in both the periods analyzed was found, cf. Figure 1.4. The higher quantity of null subjects found in the present study compared to those found by Duarte (1995) is likely due to the data that were excluded in the previous study. In the present corpus investigation, some contexts different from those considered by Duarte (1995) for example are taken into consideration, as (i) non-initial coordinated sentences, since they can optionally have an overt subject, which are thus a motivated choice (see also Rodrigues 2004, 77-79, and section 2.7 about this matter);[25] (ii) affirmative answers, which are not categorically null subject clauses (contra Duarte 1995, Kato & Tarallo 1993, *inter alia*); and (iii) main clauses with epistemic verbs, excluded in Duarte (1995) for being too frequent. "Frozen expressions" and "discursive markers" were also taken into consideration, although they were teased apart from the set of data analyzed in the present dissertation.[26]

---

[25] Although the cases of non-initial coordinated sentences might be considered ambiguous between sentence and VP coordination (which is actually possible even in non-Null Subject Languages), the final result of the analysis would be the same, except by a displacement of where the speaker must make a choice of the form of expression (either in the (non-)expression of the subject or in a VP vs. sentence coordination). In fact, there are cases in the corpus at stake in which a co-referential overt subject is found in such structures. If null subjects are much more frequent in such structures, this is a matter for analysis, not for exclusion. Moreover, being a corpus based on oral interviews, the criteria used for punctuation in the transcriptions is completely arbitrary. Deciding about when a coordinate sentence is initial or non-initial is a choice of the person who made the transcript, rather than a scientific decision. Excluding these cases would be an unjustified choice. Finally, languages described as non-pro-drop allow null subjects in non-initial coordinate sentences. However, it is well know that other subject "drops" are allowed in, for instance, colloquial English and German. If all null subjects that are allowed in one or another non-pro-drop language were excluded, probably only a few cases would be analyzed, not constituting a reliable sample of the language.

[26] For the sake of comparison, Barbosa et al. (2005) found very similar figures in their corpus study (a written corpus consisting of newspaper interviews) for BP (44% of null subjects), and a much higher relative frequency of null subjects in EP (78% in 162 cases analyzes). In Spanish, Otheguy et al. (2007) also found a lower relative frequency of pronominal subjects in different dialects of Latin American and Caribean Spanish, ranging from 19% in Mexico to 41% in Dominican Republic.

Figure 1.4 – Percentage of Null and Overt Subjects in NURC-RJ corpus
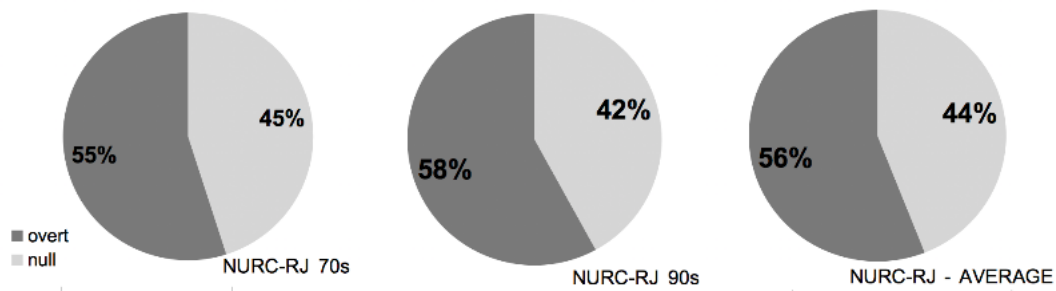
Discourse Persons and Phi-Features Verbal Markings

As Figure 1.5 shows, the percentages of null and overt subjects are not equally distributed across discourse persons. In this, the data presented here absolutely align with those presented in previous research (Negrão 1990, Duarte 1993, 1995, 2000, Cyrino et al. 2000, Barbosa et al. 2005, Duarte 2012, Kato & Duarte 2014, Duarte 2015, *inter alia*), despite the fact that here the data is also sorted into singular and plural: 3rd person subjects (from the total, 3740 finite clauses have 3rd person plural and singular subjects – 46,5% of total number of clauses) are much more frequently null than the others. About a half of the 3rd person plural subjects and even fewer of the 3rd person singular subjects are overt. The other discourse persons range between 1st person plural, the lowest, 2nd person plural, the highest, in percentage of overt subjects (with the 1st person singular and the 2nd person singular in between them, with quite similar distribution). To compare preferences for overt and null subject pronouns, the logistic regression model showed the following effects: the six conditions depicted in Figure 1.5 were compared, taking the condition with the highest number of overt subject pronouns (2nd person plural) as the baseline. Statistically, overt subjects pronouns occur significantly less frequently for the 3rd discourse person (for plural, z-value: -7.10/p-value:1.25e- 12 and for singular, z-value: -11.818/p-value: < 2e-16) than the baseline. Overt subjects were also less frequent for 1st person plural (z-value: 7.841/p-value: 4.48e-15), while for the 1st and 2nd person singular discourse persons, the decrease in numbers of overt subject pronouns compared to 2nd plural approached significance (z-value: 1.869/p-value: 0.0617 and z-value: 1.882/p-value:0.0598).[27]

---

[27] The 2nd person plural was taken as the intercept in this model.

48

Figure 1.5 – Percentage Overt Subjects in NURC-RJ according to Discourse Persons
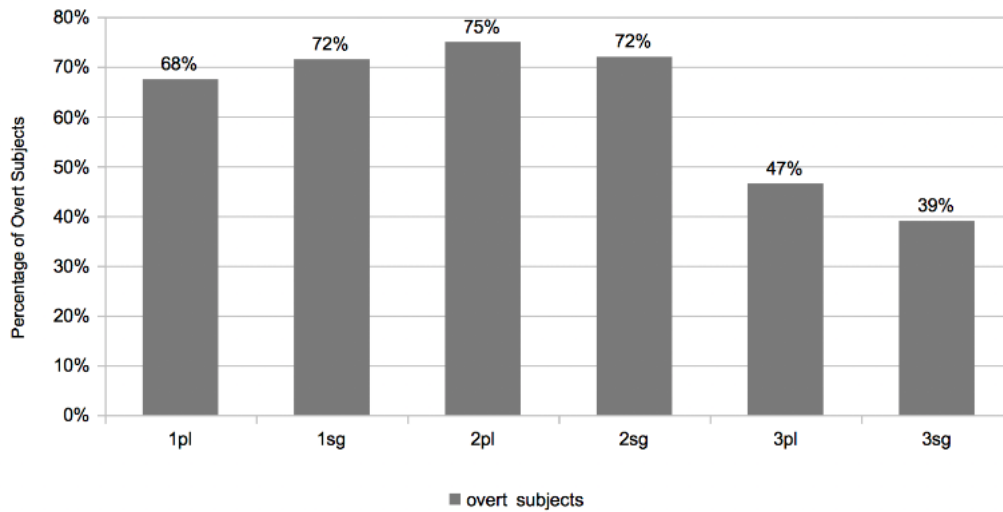
However, as discussed in section 1.2.3, these data do not shed much light on what is taking place in BP. Rather, the distribution of the relative percentages of overt subjects across discourse person challenges the relation between the impoverishment of verbal inflection and the general tendency of BP to express the subject of inflected verbs overtly. The low relative number of overt subjects in the 3rd discourse person is especially challenging to explain by means of an impoverished inflection-based approach, since this person is multiply ambiguous. On the other hand, the high number of 2nd persons, which is historically the first inflectional marking to virtually disappear in BP, might be taken to favor this approach (cf. Duarte 1993, 1995, 2000, *inter alia*). The crucial case for evaluating the effects of the impoverishment of verbal morphology on the choice between overt and null subjects in the present day BP is the 1st person singular inflectional marking system: as shown in Table 1.1, some tenses have an exclusive marking for 1st person singular, and some of them do not. If the verbal inflectional marking is a significant factor for the choice between an overt or a null subject, the verb tense group should be significant for the 1st person singular, but should not interact with the 3rd person singular and plural, since they are always semantically ambiguous either between 2nd person plural and 3rd person plural when 3rd person plural marked or between ambiguously marked 1st person singular, 1st person plural (*a gente* "the people"), 2nd person singular and 3rd person singular. The results obtained in this study partially support this prediction, as shown in Figure 1.6 below.

Figure 1.6 – Percentage Overt Subjects according to the Overt Inflectional Marking
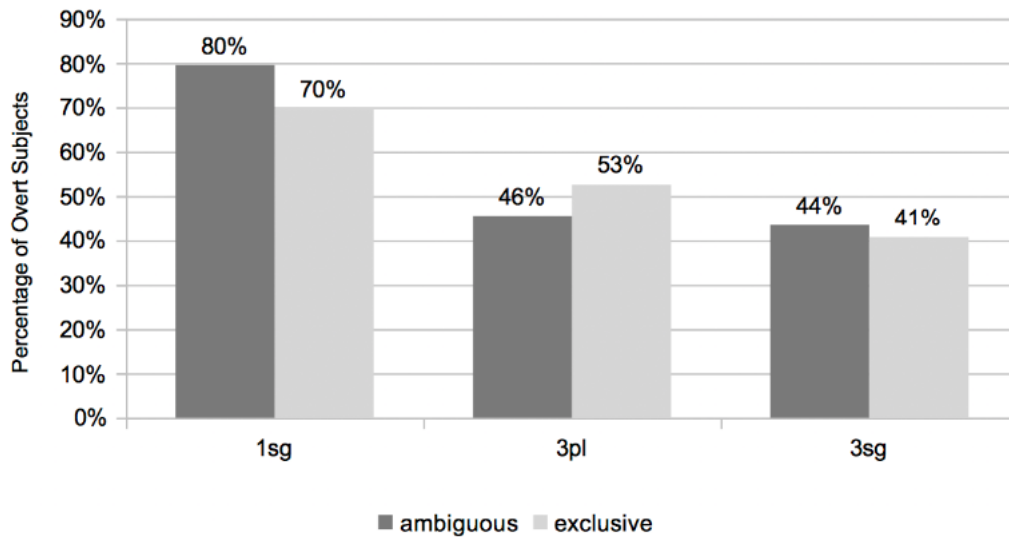
In Figure 1.6, only the current inflectional system of BP is presented; that is, only the overt markings of verbs (cf. Table 1.1) are taken into consideration. The number of 1st person singular overt subjects are overall higher than the 3rd person singular and plural. Ambiguous marking also seems to slightly influence the overall number of overt pronouns: about 5% more subjects were overt in ambiguous verbal tense than in exclusive verbal tense. Noticeably, the effect of the ambiguous tense marking is not uniform across inflectional markings: it is inverted in 3rd person plural, weak in the 3rd person singular and really relevant only for the 1st person singular: 1st person singular subjects were more frequently overt when followed by a verb in an ambiguously marked tense than when followed by a verb in a non-ambiguously marked tense. This 10% difference shows that the ambiguously marked tense group favors the use of overt subjects in the 1st person singular, although it does not transparently affect the 3rd persons in the same way. A logistic regression model with Discourse Person and Verbal Tense Type as fixed Factors indicates that the person marking contributes in a significant way to the choice between overt and null subjects (1st person singular vs. 3rd person plural: z-value: -6.850 and p-value: 7.38e-12; 1st person singular vs. 3rd person singular: z-value: -12.180 p-value: < 2e-16). Factor Verb Tense Group is also significant (z-value: -4.349 p-value:1.37e-05) with fewer overt pronouns for exclusively marked tenses. The interaction between the verb tense group and the inflectional marking has an effect on the use of overt and null subjects, presumably because verb tense group showed an inverse

pattern for 3rd person plural as regards the tense group (more overt subjects in exclusively marked verbal tenses). Finally, the intercept, which was the 1st person singular in ambiguously marked tenses, was significant (z-value: 12.327 p-value: < 2e-16), showing a preference for the use of overt subjects.

Examples (13) below nicely illustrate the quantitative data presented so far: when the verb is ambiguously marked (in the imperfective in 13a or the conditional in 13b) the 1st person singular pronoun appears; while when the verb has an exclusive unambiguous marking (the past perfect in both 13a and 13b), the overt subject is less frequent.[28]

(13) a. como ex-funcionário do   Tribunal de Contas,   na   época inclusive que $\_1$
          as   ex-employee of.the court   of accounts, in.the time   inclusive that
          dei   a   primeira entrevista, eu$_1$ estava   no   Tribunal de Contas
          give.PST.1SG the first   interview   I   be.IMP.1SG in.the court   of accounts
          do   Estado do   Rio de Janeiro...
          of.the state   of.the Rio de Janeiro
          "As a former employee of the Court of Auditors, at the time at which (I$_1$) gave the first
          interview, I$_1$ was at the Court of Auditors of the State of Rio de Janeiro"

       b. mas não $\_1$ tive   oportunidade e   depois   também a   idade
          but not   have.PST.1SG opportunity   and afterwards also   the age
          vai   chegando, preguiça, aquela coisa. Eu até   poderia   ter,
          go.PRS.3SG arriving   laziness   that   thing I even can.COND.1SG have.INF,
          mas $\_2$ perde   o   elã...
          but   lose.PST.3SG the enthusiasm
          "But I$_1$ haven't had the opportunity, and then also the age is coming, laziness, that
          thing. I$_1$ could have it, but one$_2$ loses the enthusiasm..."

(NURC-RJ, "Inquiry_r_52")

---

[28]   It is important to point out that the overt subject is still probably more frequent in similar contexts than, for instance, in EP. However, a similar investigation would have to be carried out in order to provide quantitative data from EP.

Semantic Features of Antecedents

The result reported in the previous subsection supports the impoverished-inflection-based approach.[29] However, the fact about the high number of null subjects in the 3rd person singular and plural compared to other discourse persons remains unexplained. The data were reanalyzed and a further statistical regression was run to provide a possible explanation for the high number of null subjects. In this analysis, criteria based on semantic properties of the antecedent were taken into consideration, specifically (i) the Referential Hierarchy in (8), proposed by Cyrino et al. (2000), Kato et al. (2006), Duarte, Mourão & Santos (2012), Duarte, Mourão & Guimarães (2012), Kato & Duarte (2014), Duarte (2015) and (ii) specificity proposed by Negrão & Müller (1996), Menuzzi (2002), Müller (2003), Menuzzi & Lobo (2016), who claim that antecedents of null subjects tend to exhibit lower specificity in BP. In Figure 1.7 below, the percentages of 3rd person overt subjects divided according to their animate or inanimate and specific or non-specific antecedents were plotted. In this analysis, only co-referential (non-arbitrary) and entity-referring (non-sentential) null and overt subjects were taken into consideration (total: 2882 clauses).

---

[29] As pointed out to me by Sergio Menuzzi, this confirms what would be expected at first-glance in a functionalist approach: the more potentially ambiguous the verb form, the higher the frequency of overt subjects. However, the numbers attest to a difference of 10% vs. 7% vs. 3% – that is, a variable effect on a small proportion of the overall set of occurrences. So, it seems that what is picked up by these number is rather a functional factor operating on the overall use of the available forms, and not a strong structural constraint. This observation goes along with what is proposed in the next sections.

Figure 1.7 – Percentage 3rd Person Overt Subjects according Animacy and Specificity
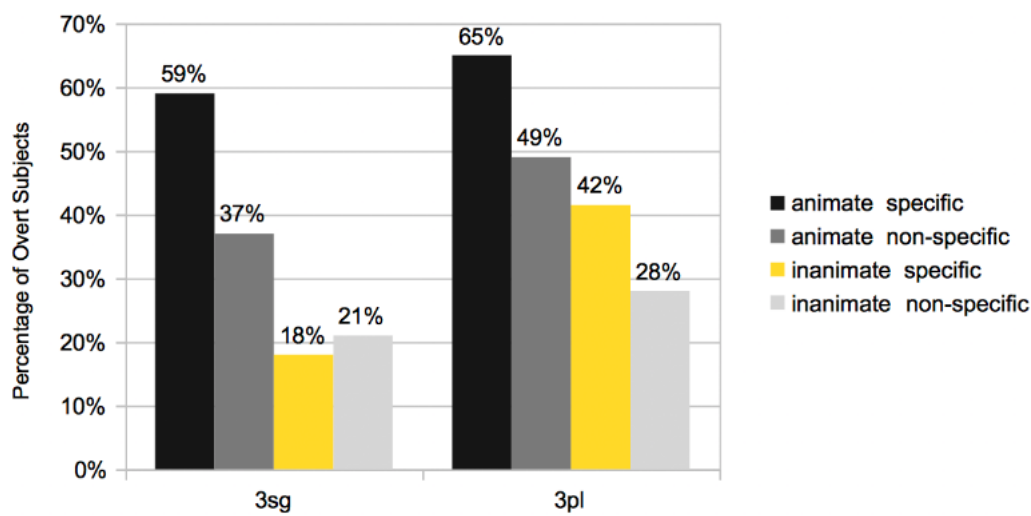
In Figure 1.7, there is a clear tendency: animate and specific antecedents are mostly retrieved by overt subjects while inanimate and non-specific antecedents are preferably recovered by null subjects.[30] The 3rd person singular subject is more frequently overt when its antecedent is animate and specific, followed by the animate non-specific antecedent, and more or less at the same level the inanimate specific and non-specific antecedents; in the 3rd person plural, animate specific overt subjects are close in frequency to other discourse persons, followed by animate non-specific subjects, by inanimate specific subjects at a close frequency level and, at the bottom of the scale, inanimate non-specific subjects. In the three way interaction model, taking into consideration the features inflection (plural vs. singular), animacy (animate vs. inanimate) and specificity (specific vs. non-specific), each one of them was individually significant (inflection z-value: -2.907 and p-value: 0.00365; animacy z-value: -3.266 and p-value: 0.00109; specificity z-value: 2.958 and p-value: 0.00310). No interaction was even slightly significant but the three way interaction was (z-value: -2.111 and p-value: 0.03476), which must be due to the number of specific inanimate overt subjects being lower than non-specific inanimate overt subjects in 3rd person singular.

In the examples below, the quantitative data reported in this section are illustrated by some

---

[30]   The same conditioning is found with null objects, as pointed out in much previous research (see Cyrino et al. 2000, Creus & Menuzzi 2004, *inter alia*).

instances of the use of null subjects in relation to the semantic features of the antecedents. In (14a), the null subject retrieves a specific inanimate antecedent *o suporte* "the support". In (14b), the null subject refers back to the non-specific animate antecedent *as pessoas* "the people". Finally, in (14c), the antecedent of the null subject is an inanimate non-specific entity *sopa* "the soup".

(14) a. A casa estava   velha. Esse suporte$_1$ caía   toda hora... uma vez _$_1$
          the house be.IMP.3SG old   this support fall.IMP.3SG all time one time
          caiu   na   orelha da   empregada, _$_1$ quase tira   a   orelha
          fall.PST.3SG in.the ear   of.the maid,   almost take.PST.3SG the ear
          fora...
          out
          "The house was old. This support$_1$ fell all the time... once it$_1$ fell on the ear of the maid, it$_1$ almost took the ear off..."

      b. as pessoas$_1$ comem   tanto... _$_1$ comem   milho... paçoca... pamonha...
          the people   eat.PST.3PL so.much _$_1$ eat.PST.3PL corn   paçoca   pamonha
          "People$_1$ eat so much... they$_1$ eat corn... paçoca... pamonha..."

      c. Sopa$_1$ saía   quente da   cozinha e   quando _$_1$ chegava   lá
          soup leave.IMP.3SG hot   from.the kitchen and when   arrive.IMP.3SG there
          na   sala   _$_1$ já   chegava   fria.
          in.the living.room   already arrive.IMP.3SG cold
          "The soup$_1$ left the kitchen warm and when it$_1$ arrived there in the living room it$_1$ already arrived cold."

                                        (NURC-RJ, "Inquiry_011")

As exemplified by (15) below, specific animate antecedents, such as *a cozinheira* "the cook", tend to be retrieved by overt subjects.

(15) eu$_1$ cabia   direitinho debaixo da   pedra e   _$_1$ ficava   conversando
        I   fit.IMP.3PL right   below   of.the stone and _$_1$ stay.IMP.1SG talking
        com a   cozinheira$_2$... ela$_2$ ficava   cheia porque _$_2$ diz   que eu$_1$ não
        with the cook...   she stay.IMP.3SG full   because _$_2$ say.PST.3SG that I   not
        parava   de falar...
        stop.IMP.1SG of talk.INF
        "I$_1$ fit right under the rock and I$_1$ kept talking to the cook$_2$... she$_2$ was fed up because she$_2$ says that I$_1$ didn't stop talking..."

                                        (NURC-RJ, "Inquiry_011")

As for the quantitative analysis, notice that not only personal pronouns have been taken into consideration, but also demonstrative pronouns, which may have increased the number of overt subjects in the sample. However, the tendency to use null subjects to retrieve inanimates seems so strong that the demonstratives cannot have deeply impacted the final results (as appears clearly when comparing the results presented in this dissertation – with demonstratives – to those found by Duarte (1995) without them). On the other hand, including demonstratives led to a new observation: although demonstrative pronouns, such as *esse/a(s)* and *isso* (and the almost non occurring *este/a(s)*) and *isto*) "this(these)" as well as *aquele/a(s)* and *aquilo* "that(those)" are used to refer to propositions and other higher order entities (especially when they are sentential) and, less frequently, to first order entities: the standard nominative pronouns (*ele/a(s)* "s/he/it(they)") typically refer to first order entities, rather than propositions and other higher order entities, and cannot refer to sentential antecedents (in the corpus research reported here, no single co-reference between a sentential antecedent and a standard gender-marked nominative pronoun was found). Such a specialization of standard gender-marked nominative pronouns is not found in the distribution of null subjects: they can refer to both sentential and non-sentential entities. The gender marked demonstratives, though, seem to be used preferentially in contexts of exclusiveness, as in (16) below.

(16)   Aqui no Rio, por exemplo, nós temos uns quatro ou cinco$_1$ que não... não passam daquilo nunca, compreende? Ao passo que os outros$_2$, por serem donos de uma grande torcida como é o caso do Flamengo, do Vasco, do Fluminense, do Botafogo, do América... Então, *esses*$_2$ conseguem estar sempre na crista. Os outros$_1$, infelizmente, muito embora $\_1$ tenham um patrimônio e tudo, eles$_1$ não conseguem armar nada.

"Here in Rio, for example, we have four or five$_1$ that do not ... never go beyond that, do you understand? Whereas the others$_2$, because they$_2$ own a great crowd as Flamengo, Vasco, Fluminense, Botafogo, América ... So, these$_2$ can always be on the top. The others$_1$, unfortunately, although they$_1$ have a heritage and everything, they$_1$ cannot put together anything."

(NURC-RJ, "Inquiry_052")

As observed by Duarte (1995), Duarte, Mourão & Guimarães (2012), Duarte (2015), the demonstratives seem to be used quite often to retrieve a sentential content. The quantitative data

points in this direction: almost a half of the inanimate subjects referring to sentential antecedents were overt, while the relative percentage of overt subjects referring back to non-sentential inanimate entities is much lower, as shown in Figure 1.8 below. The quantity of overt subjects (demonstrative pronouns) referring to this sort of antecedent is in fact significantly greater than for other sorts of inanimate subjects (z-value: 7.266, p-value: 3.7e-13). This result is absolutely unexpected under the Referential Hierarchy hypothesis in (8), since "propositional" is the lowest degree in this hierarchy and thus the type predicted to be most likely to be recovered by null subjects. Previous research has also already observed similar figures (see, for instance, Kato & Duarte 2014, 8), without explaining them.

Figure 1.8 – Percentage Non-Sentential and Sentential 3rd Person Inanimate Overt Subjects
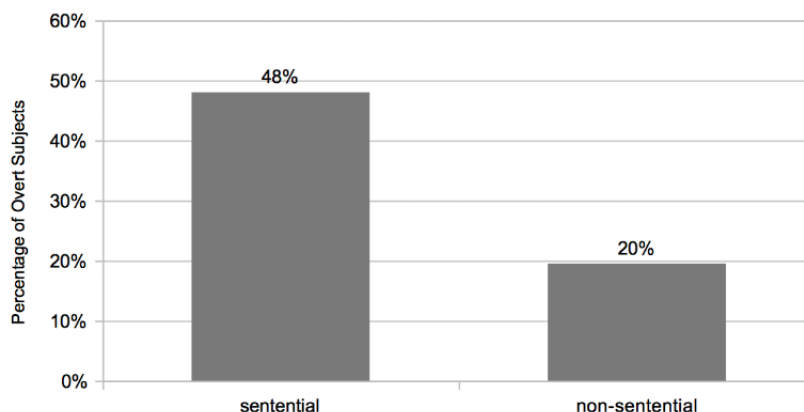
Just as sentential antecedent is lowest level on the scale of Animacy, indefinite genderless pronouns are the lowest level on the scale of Specificity. Though a few cases of overt subjects corefering with *alguém* "somebody" were found, no overt subject was co-referent with *ninguém* "nodody", although null subjects commonly co-refer with this sort of antecedent, as illustrated by (17) below.[31]

---

[31]   The term "corefer" is used loosely here when talking about quantificational NPs and bound variables (see Montalbetti 1984, Menuzzi & Lobo 2016, for a different view on the relation between null subjects and these antecedents).

(17)   Ninguém₁ revela realmente tudo que ₁ sente.

"Nobody₁ really reveals everything they₁ feel."

<div align="right">(NURC-RJ, "Inquiry_071")</div>

1.6.4   Discussion

The results of the corpus research reported here converge in many aspects with those found in previous research. Although the frequency of null subjects found in this study is higher than in those cited above (Lira 1982, Tarallo 1983, Duarte 1993, 1995, 2000, 2012, *inter alia*), the distribution across discourse persons is quite similar to that found previously. This is a crucial point for understanding the nature of the use of null and overt subjects in BP, since much literature has been devoted to the claim that the impoverishment of verbal inflectional markings has triggered the preference for (co-)referential overt pronominal subjects.

As observed in the literature review and in the corpus research, this claim does not explain the higher relative number of 3rd person null subjects in BP, since the verbal marking for this person is the least informative. However, in this corpus investigation, an effect of inflectional marking was found: when the 1st person singular is used in a TT that does not exclusively and explicitly reveal the discourse person of the subject, the overt form is preferred to the null form. In this, the corpus findings suggest that impoverishment has led to a bias in favor of maximum informativeness in present day BP, taking into account the paradigmatic knowledge of the verbal system marking along with the syntagmatic context in which the subject is used. Statistically, the combination of 1st person singular together with an ambiguous TT was significant, but so were the multiply ambiguous 3rd person markings. Overall, the importance of discourse person is unclear: the high relative number of null subjects in 3rd persons is not explained by the impoverished verbal marking system. Any hypotheses which rely on some type of relation between verbal marking and the licensing of null subjects require further assumptions for dealing with such a distribution. However, as suggested by the proposals based on the Referential Hierarchy, the semantic features of the antecedent can explain the high relative frequency of 3rd person null subjects.

Two other features have been shown to be relevant among those proposed in the previous literature, as the lower relative number of overt subjects also reaches statistical significance, namely

inanimate and non-specific antecedents favor null subjects. Because such antecedents are strongly linked to the 3rd person, these features might have triggered the figures observed in this and in previous corpus studies (except by Barbosa et al. 2005, who found 43% of inanimate overt subjects in their written newspaper interview corpus). As suggested by the quantitative and qualitative analyses carried out here, the semantic features [-animate] and [-specific] seem to be the loci of null subjects in the current grammar of BP, by being preferentially and sometimes obligatorily realized by an anaphoric null subject.

In the next section, experimental evidence will be provided in order to verify the robustness of the primary observations reported in this section, leading to a deeper theoretical understanding of the phenomena. Each of the predictors that reached significance in the corpus research will be individually tested. With this procedure, it is attempted to (i) ensure that, in a controlled linguistic environment, these individual predictors can also be significant; (ii) eliminate possible confounding factors that might have influenced the choice between overt and null subjects in the interviews (for instance, specific syntactic or discourse contexts); and (iii) provide evidence from comprehension, rather than from production, that these features play a role in the grammar of the anaphoric subject in the current system of BP.[32]

1.7   Experimental Evidence

Given the results obtained in the corpus research described in the previous section, three experiments were carried out in order to check whether the factors established in the corpus research can be considered individually relevant. In the first experiment, exclusive vs. non-exclusive ("ambiguous") verbal marking was tested by varying the tense of a subordinate clause verb whose subject was always 1st person singular. In the second, the effect of Animacy on the null and overt subject acceptability was tested by a comparison between animate vs. inanimate antecedents. Finally, in the third experiment, specific vs. non-specific null subject preference was at stake. The results

---
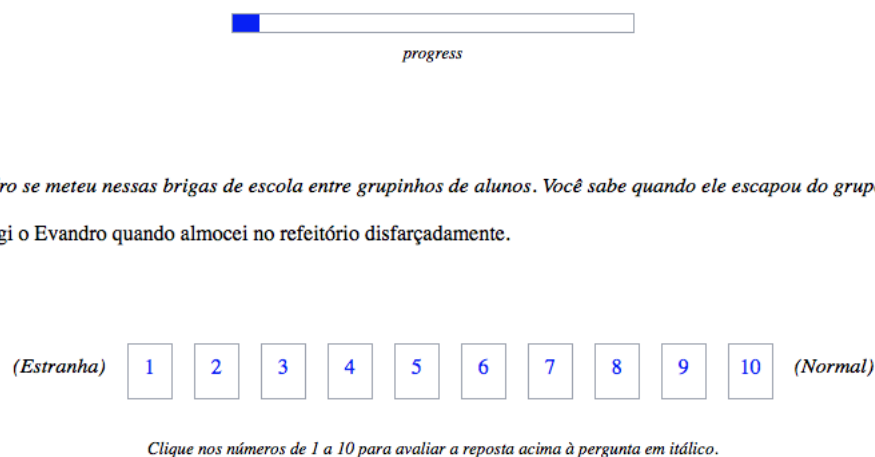
[32]   Interactions between the predictors found in this chapter and other factors are initially taken into consideration in Chapter 2. It is surely necessary to further study these interactions, providing a mathematical model that can precisely predict the use and interpretation of null subjects in BP. Here the primary aim is to establish which principles play a role in the system.

of the third experiment indicated further developments of the importance of Specificity, so it was rerun with slightly different materials in order to better understand the structure and the interpretation of non-specific null and overt subjects. The results of the experiments confirm the predictions made from the corpus analysis and in some of the preceding literature.

1.7.1 Methodological Considerations

In all the experiments reported in this section, the methodology was the same: participants read a sequence of two turns, that is, a short dialogue. They were asked about the acceptability of the answer in the relevant context on a scale from 1 to 10 in a judgment task, cf. Figure 1.9. They were told to use the full scale according to how natural "Normal" or strange "Estranha" the answer seemed in the context of the question. After judging them, the participants were asked about the interpretation of the relevant subject – null or overt – in a closed yes-no question task, cf. Figure 1.10.

Figure 1.9 – Screen sample – Judgment Task

progress

O Evandro se meteu nessas brigas de escola entre grupinhos de alunos. Você sabe quando ele escapou do grupo rival?

Eu protegi o Evandro quando almocei no refeitório disfarçadamente.

(Estranha) 1 2 3 4 5 6 7 8 9 10 *(Normal)*

Clique nos números de 1 a 10 para avaliar a reposta acima à pergunta em itálico.

Figure 1.10 – Screen Sample – Closed Question Task

progress

Era o Evandro quem tinha almoçado no refeitório disfarçadamente?

1. Não.

2. Sim.
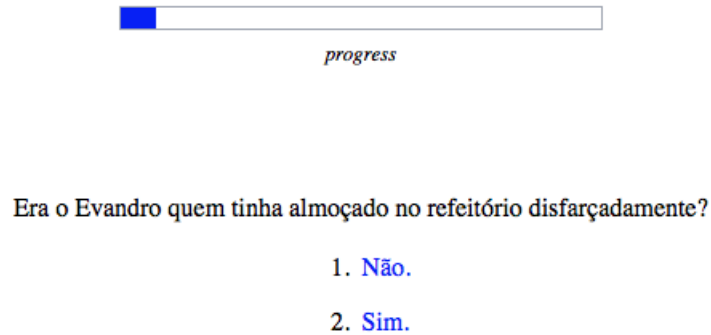
All participants voluntarily participated in the experiments on the IbexFarm platform (http://spellout.net/ibexfarm). They filled in a basic information form that included a declaration of written consent and had 4 sentences to practice before starting the experiments, which took them around 30 minutes to complete.

Some of the experiments were run together, so that an item of one experiment could be a distractor for the items of another experiment.[33] Among the items, four perfectly acceptable control sentences were inserted. Four control sentences that violate strong grammatical or pragmatic constraints were inserted at the end of the experiment, in order to (i) compare to the relevant items of the experiment, and (ii) insure that participants were attentive until the end of the experiment.

1.7.2 Experiment 1 – Exclusive vs Non-exclusive Marking

This experiment was designed to check whether BP speakers rely on verbal marking to identify the subject of the verb. In all the sentences, the subject was informationally and structurally salient both in the context and in the preceding main clause.[34] Two experimental Factors were tested: null

---

[33] Due to the availability of participants at the moment in which the experiments were run, some of them were run together. I acknowledge that this is not the optimal set-up for experiments such as those related to null and overt subjects, since there is mounting evidence on the effects of frequency. However, because of multiple practical restrictions, this solution was the most suitable way of getting enough participations in many experiments.

[34] Informatively and structurally salient is taken to be the topic and subject of the current segment of the discourse (in this case, the clause immediately preceding the clause at stake). More details about Salience of the antecedent are given in Chapter 2, where this notion is discussed and motivation for the notions assumed here are given.

vs. overt subject (Factor Subject) and exclusive vs. ambiguous marking (Factor T[ense]T[ype]). The hypotheses were the following: (i) if the participants rely on verbal marking, the ambiguous TT along with a null subject would be judged worse than the other three conditions; (ii) if the participants take overt subjects to be redundant with verbal exclusive marking, either there would be no difference regarding Factor Subject with the ambiguous marking or the overt subject would be judged worse for being redundant (over-informative) with the exclusive marking, as predicted by the "Avoid Pronoun Principle" (Chomsky 1981, Duarte 1995); that is, differently from (i), the exclusive marking should make the Overt Condition less acceptable and the ambiguous marking should not interact with the null or overt subject; and (iii) if the participants take the system as a whole to be impoverished and thus structurally and informatively deficient, the main effect of Factor Subject would be significant (overt subjects must be preferred), but no interaction with Factor TT would be expected.

Material Design

Twenty-four items were created for this experiment, based on a previous experiment carried out for different purposes (Fernandes et al. 2016, to appear). The context provided for each item is considered heavily biased toward the subject of the main clause (it averaged 80% of antecedent choices co-referring to the subject of the matrix clause, see Experiment 1 in Fernandes et al. 2016, to appear for further evidence). Differently from the previous experiment, however, the subject was always 1st person singular. Also, a context question was provided, cf. (18a), in order to make the sequence sound as natural as possible. This context question was composed of two sentences: an introductory context and an indirect temporal question introduced by (*quando*) asking when an event took place. The answer to the question displayed the four conditions in a temporal subordinate clause, also introduced by *quando*, directly addressing the relevant question, cf. (18b) and (18c) below.

The four conditions were the following: in items such as (18b), the verb in the temporal subordinate clause had an exclusive marking and it was either preceded by the overt pronoun *eu* "I" or not (that is, a null 1st person singular subject should be recovered from the context or from the verbal marking); in items such as (18c), the relevant verb was ambiguously marked and was also preceded by the overt pronoun *eu* "I" or not, in which case the null 1st person singular subject

61

should be interpreted from the (syntactic) context.

Afterwards, the participants had to answer an interpretation question indirectly questioning whether the relevant subject was either *eu* "I" (in which case the expected answer was *Sim.* "Yes") or the direct object of the main clause (in which case the expected answer was *Não.* "No"), cf. (19).

(18)   a. A – A Maria estava muito nervosa. Você sabe quando ela ficou mais calma?

"Mary was very nervous. Do you know when she's got calmer?"

b. B – Eu_1 tranquilizei   a   Maria quando $\text{eu}_1/\_1$ divulguei   os
        I   calm.down.PST.1SG the Maria when I   publish.PST.1SG the
resultados do   exame.
results   of.the examen

"I calmed Mary down when I published the results of the exam."

c. B – Eu tranquilizei   a   Maria quando $\text{eu}_1/\_1$ ia   divulgar   os
        I   calm.down.PST.1SG the Maria when I   would publish.INF the
resultados do   exame.
results   of.the examen

"I calmed Mary down when I'd publish the results of the exam."

(19)   Era a Maria/eu quem tinha divulgado/ia divulgar os resultados do exame?

"Was it Mary/me who had published/would publish the results of the examen?"

a. Sim.

"Yes."

b. Não.

"No."

Given the results of the corpus study, the empirical predictions are the following: items such as (18b) will be better rated without the overt subject, since the subject is redundant with the exclusive inflectional marking if the "Avoid Pronoun Principle" still has some effect in the realization of anaphoric subjects in BP; items such as (18c) will be considered better when preceded by an overt subject than when the subject is null, if participants rely on verbal explicit marking. Interpretations should not diverge across conditions.

Participants

Twenty-seven participants took part in this experiment, all of them highly educated (minimally under-graduation students) living in the southern region of Brazil (Rio Grande do Sul and Santa Catarina). They were invited to participate via email and Facebook, and gave their written consent to use the data. Their age averaged 38,5 years (ranging from 32 to 45 years). They work or study in different fields (all the 4 big areas of knowledge described by the Brazilian government were included).[35] In the analysis, the results of participants who either scored below 80% in the interpretation task (which had correct and incorrect answers) or scored ungrammatical control sentences above eight were discarded (three participants). Only right answers were taken into consideration (only 4 trials were excluded from the group of participants analyzed).

Results

As shown in Figure 1.11 below, the experimental results correlate with those found in the corpus research reported in Section 1.6.3: exclusive marking is not the main factor modulating the acceptability of null or overt pronouns. Rather the combination of an ambiguous TT with a null Subject reduces the acceptability of the sentence (averaging 6.7). When the ambiguous TT verb is preceded by an overt subject, acceptability is much higher (mean: 7.5) and close to the the acceptability the participants attributed to the exclusively marked verb preceded by an overt subject (mean: 7.6). The most acceptable sentences were those whose verb shows an exclusive TT with a null subject (averaging 7.7).[36]

---

[35] The Brazilian Government splits higher education into four big areas: (i) Math and Engineering, (ii) Humanities, (iii) Linguistics and Arts, and (iv) Earth and Life Sciences. For the sake of analysis, these areas were taken as a reference.

[36] In Figure 1.11 and in all the following graphs, error bars represent standard deviations.

Figure 1.11 – Mean Acceptability Judgments according to Tense Type and Subject Factors
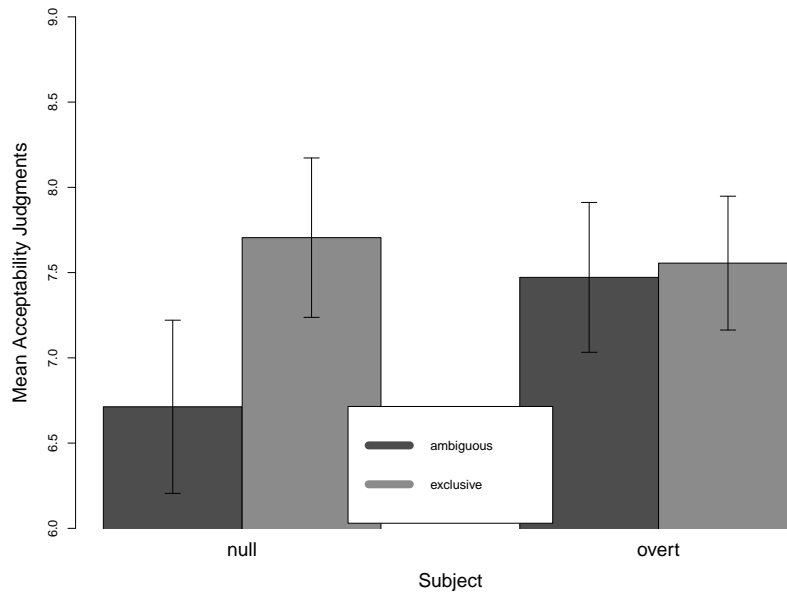
For the inferential statistical analysis, mean acceptability judgments were entered into a log-linear mixed-effects model analysis containing two Factors (TT and Subject) with two levels and random effects (Participants and Items) along with random slopes (maximal model) (Barr et al. 2013). The outcome of the model is summarized in Table 1.5 below. Main Factor TT was individually significant, but this effect may partly be due to the interaction effect. Main Factor Subject and the interaction showed marginal effects.

Condition	Estimate	Std. Error	T-value	Pr(>|t|)
(Intercept)	6.7708	0.3855	17.564	1.78e-15 ***
Subject	0.6767	0.3608	1.875	0.07642 .
TT	0.8901	0.3023	2.945	0.00854 **
Subject:TT	-0.7163	0.4042	-1.772	0.08460 .

Table 1.5 – Log-linear mixed-effects model for Experiment 1

Discussion

The result of Experiment 1 goes in the same direction as the result from the corpus analysis: BP speakers tend to disprefer the ambiguity generated by the lack of an exclusive verbal marking preceded by a null subject in the 1st person singular. This claim is supported by the higher number

of overt subjects found in the corpus in the 1st person singular in ambiguously marked tenses and by the drop in acceptability in Experiment 1 when the verb was not exclusively marked and not preceded by an overt subject. Notice, however, that both the effects are marginally significant: (i) in the corpus study, the difference between ambiguous and exclusive tense groups was 10%, while the difference across person marking was 20% between 1st person singular and 3rd person plural and 30% between 1st person singular and 3rd person singular; and (ii) the range of different acceptabilities was not very broad (the lowest mean was 6.7 and the highest 7.7, while the ungrammatical control sentences averaged 2.8 in the same acceptability task). It thus appears that BP speakers do not disprefer redundant marking. Ambiguity, on the other hand, is slightly dispreferred: non-exclusive verb forms with a null subject are judged at 6.8, but it is still very far from the lower part of the scale and close to the highest average judgment of 7.7 for exclusive verb forms with a null subject. For this reason, the interaction of factors was only marginally significant in Experiment 1. These results thus suggest that BP speakers have paradigmatic knowledge that the ambiguously marked construction with a null subject is not optimal for comprehension, and that the most informative and efficient form in this case is the overt subject. The preference for overt subjects is less strong when the 1st person singular verb shows an exclusive marking and the choice between an overt or a null subject seems to be motivated by other factors, as it is when a multiple ambiguous inflection marking is used (3rd person plural or 3rd person singular, for instance). In the next section, a second experiment that aims to work out some of these additional factors is presented.

1.7.3 Experiment 2 – Inanimate and Animate Antecedents

This second experiment was designed to verify whether 3rd person singular null and overt subjects show any preference for particular semantic types of antecedents. In all the sentences the subject was informationally and structurally salient both in the context and in the question under discussion. Two binary Factors were tested: overt vs. null subject (Factor Subject) and inanimate vs. animate antecedent (Factor Animacy). The hypotheses were the following: (i) if Animacy plays a role in the use of null subjects in BP (cf. the Referential Hierarchy in (8), for example), a significant interaction between the conditions is expected; the null subject should be better rated when referring to an inanimate antecedent and less acceptable in the case of animate antecedents; (ii) if

the Animacy of the antecedent has a relevant effect on the choice of overt or null subjects, but it is not as predicted by the Referential Hierarchy in (8), either no significant interaction is expected, but the main effect of Animacy would be significant, or the null subject can be better rated when retrieving animate antecedents and the overt when co-referring to inanimate ones, and a significant effect of the interaction would come up; and (iii) if Animacy plays no significant role in the use of null and overt subjects, the overt subject must be preferred regardless of the semantic type of antecedent, since BP is generally taken to favor overt subjects over null subjects in the current stage of the language.[37]

Material Design

Twenty-four items were created, based on verbs that were found with a null or overt inanimate subject in a corpus search. A Google search confirmed that they were equally used with animate subjects. As in the previous reported experiment, a context sentence was provided, such as (20) below. Following this sentence, an indirect question asking what happened either to an animate referent or to an inanimate one was provided, as in Table 1.6. The answer could have either a null subject or an overt gender-marked subject pronoun, cf. Table 1.6 (masculine and feminine genders were counterbalanced across items). Afterwards, participants were indirectly asked if the subject of the relevant verb was either *uma pessoa* "a person" or *um objeto* "an object", cf. (21) below.

(20)   A Maria estava muito irritada depois da reforma no apartamento.

"Maria was very stressed out after the flat refurbishment."

---

[37]   As pointed out by Scott Schwenter, the frequency of combination of a given verb with a specific kind of referent might have influenced the results: supposing that a verb, such as *cair* "to fall", is by far more frequent with animate subjects than with inanimate ones, given the general frequency of overt subjects in BP, a collocation overt subject + *cair* could be at stake in the results found in this experiment. I agree that the role of frequency may be further studied as regards the realization of overt and null subjects with some verbs, but this possible intervening effect is accounted for in the present analysis as a random Factor ("Item") in the mixed-effects model below. In future studies, once the frequency of given combinations is established, "frequency" can be run as main (intervening) Factor in the model.

Animacy	Subject	Question	Answer
(A) animate	null	A –Você sabe o que aconteceu com a colega de quarto dela lá? A –Do you know what happened to her roommate there?	B – Caiu da bancada. B –She fell from the stand.
(B) inanimate	null	A –Você sabe o que aconteceu com a televisão dela lá? A –Do you know what happened to her television there?	B – Caiu da bancada. B –It fell from the stand.
(C) animate	overt	A –Você sabe o que aconteceu com a colega de quarto dela lá? A –Do you know what happened to her roommate there?	B – Ela caiu da bancada. B –She fell from the stand.
(D) inanimate	overt	A –Você sabe o que aconteceu com a televisão dela lá? A –Do you know what happened to her television there?	B – Ela caiu da bancada. B –It fell from the stand.

Table 1.6 – Materials – Experiment 2

(21)   Então, foi uma pessoa/um objeto que caiu?

"Was it a person/an object that fell?"

a. Sim.

"Yes."

b. Não.

"No."

Based on the results from the corpus study, the empirical predictions were the following: null subjects are preferentially used to retrieve inanimate antecedents (Condition B) over animate antecedents (Condition A); on the other hand, overt subjects are more acceptable when they pick up an animate antecedent (Condition C) than when they refer back to an inanimate one (Condition D). A significant interaction between Factors (Subject and Animacy) is thus expected.

Participants

Twenty-nine participants took part in this experiment. They were on average 37.1 years old (more or less 7 years). In the result analysis, five participants were discarded because they either scored

below 80% in the interpretation task (which had correct and incorrect answers) or scored ungrammatical control sentences above eight. Otherwise, the procedure was the same as in the previous experiment.

Results

The results of this experiment provide clear corroboration for the findings of the corpus study. As shown in Figure 1.12 below, the null subject is preferred when referring to an inanimate antecedent (averaging 8.5/10), while the overt subject is less acceptable (mean: 7.45/10). On the other hand, the overt subject is better rated when the antecedent is animate (mean: 8/10) than when the antecedent is inanimate (7.5/10).

Figure 1.12 – Mean Acceptability Judgments according to Animacy and Subject Factors
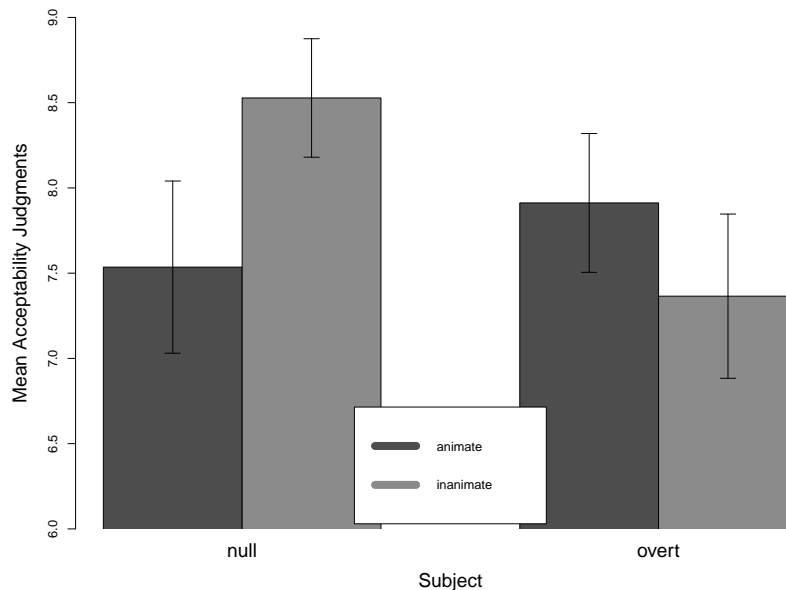
As for the inferential statistical analysis, as with Experiment 1, the data were analyzed using a log-linear mixed-effects model analysis containing two Factors (Animacy and Subject) with two levels each and random effects (Participants and Items) (Barr et al. 2013). The full model is summarized in Table 1.7 below. The intercept was considered the null subject referring back to the animate antecedent. The interaction between both Factors was significant. Factor Animacy is marginally significant and Factor Subject was not significant at all. The positive Estimate shows

that inanimate antecedents were generally judged slightly better than animate antecedents.

Condition	Estimate	Std. Error	T-value	Pr(>|t|)
(Intercept)	7.6099	0.3580	21.256	< 2e-16***
Subject	0.4313	0.2974	1.451	0.16165
Animacy	0.7787	0.4316	1.804	0.08397.
Subject:Animacy	-1.2624	0.4149	-3.043	0.00517**

Table 1.7 – Log-linear mixed-effects model for Experiment 2

Discussion

As in the corpus study reported in section 1.6.3, in Experiment 2, Animacy of the antecedent seems to modulate the acceptability of overt or null subjects. The statistical analysis of Animacy turned out to be significant in two different approaches and methodologies. These results shed light on the relative frequency of null subjects across discourse persons: this unclear relation between the high number of null subjects in the 3rd person singular and the multiply ambiguous verbal marking in this discourse person is explained by a factor that is independent from inflection, namely the inherent semantic features of the antecedent. Further experiments will be necessary to test in more detail whether this difference is related to grammatical gender features or purely semantic conceptual features.[38] The next section will focus on another semantic feature that can also facilitate the use of null subjects in BP, which has been also shown to be relevant in the corpus study.

1.7.4   Experiment 3a – Non-specific *vs.* Specific Antecedent

This third experiment was designed to test whether the semantic feature Specificity plays a significant role in the choice between null or overt 3rd person singular subjects. For this experiment, however, a slightly different methodology was used: the interpretation question which follows the judgment task is crucial (rather than a question for checking attention/accuracy from participants), since both the interpretations for the overt or null subject would theoretically be possible. Con-

---

[38]   Creus & Menuzzi (2004) advance an hypothesis in this sense, claiming that the lack of semantic gender is the explanation for the use of null objects in BP. In the final section of their paper, it is suggested that the same analysis could be extended to the use of null subjects in BP. It is not clear so far that this analysis can account for the data found in this dissertation. Further research is necessary to ground this extension on empirical data.

sidering this design, the kind of evidence that this experiment would eventually bring about is related to anaphora resolution and the bias toward specific or non-specific antecedents. In particular, this experiment brings together two important aspects that could be possibly related. First, the potential bias of null subjects toward non-specific antecedents. Second, the observation that the null subject can refer to the discourse topic, as first pointed out (to the best of my knowledge) by Moreira da Silva (1984), and reported by Modesto (2000*b*, 51-52) and Ferreira (2004, 19). According to the last author, null subjects are somehow linked to a null element in the topic syntactic position (TopP), in a similar way to Wh-movement or constituent topicalization (extraction). Following his reasoning, null subjects are subject to the same constraints that apply to Wh or NP extraction, namely island constraints (see also Sheehan 2006, for a good synthesis, and Chapter 4 in the present dissertation, where island constraints are discussed in more details). In this case, the referent of the null pronoun can come from the discourse context, as in (22) below, mainly when it is the topic of a question. However, Modesto (2000*b*, 51-52) observes it is not always possible for the referent of a null subject to be retrieved from the discourse context, as in (23) below.

(22) A– E o Paulo$_2$?

"What's up with Paulo$_2$?"

a. B– O Pedro$_1$ disse que $\_\_2$ ganhou na loto.

"Pedro$_1$ said that he$_2$ won the lottery"

b. B– $\_\_2$ trabalha na universidade.

"He$_2$ works at the university"

(23) Ninguém$_1$ disse que $\_\_1$/*ele$_1$ ganhou na loto.

"Nobody$_1$ said that he$_1$ won the lottery"

(Modesto 2000*b*, 51-52)

Modesto (2000*b*, 51-52) points out that the main clause subject *Ninguém* "Nobody" is preferably interpreted as the antecedent of the null embedded subject in (23), as opposed to (22a), in which the main clause subject does not interfere in the co-reference with the discourse salient antecedent. What Modesto (2000*b*, 51-52) does not mention, however, is that in the context of (22)

the sentence in (23) is perfectly acceptable when an overt pronoun referring to the discourse salient antecedent is used. The point is thus that null and overt subjects seem to be in some sort of complementary distribution in such cases. Moreover, a further observation from the corpus data is that overt and null subjects diverge in their interpretation bias when the subject of the matrix clause is either a specific or a non-specific referent, cf. (24).

(24)   não , o único$_2$ que$_2$ eu vi e que$_2$ achei um projeto maravilhoso ... e ele$_2$ teve uma mixuricagem de votos foi o Sérgio Bernardes$_2$. (...) todo mundo$_1$ diz que _$_1$/*_$_2$/*ele$_1$/ele$_2$ é sonhador.

"No, the only one$_2$ that$_2$ I saw and whose$_2$ project I found wonderful and he$_2$ had a bit of votes was Sérgio Bernardes$_2$. (...) everyone$_1$ says they$_1$/he$_2$'s a dreamer. "

(NURC-RJ, Inquiry_r_347)

In the original example in (24), the overt pronoun unambiguously refers back to the contextual salient referent *Sérgio Bernardes*$_2$ and probably no null pronoun could refer to this antecedent in this context. With a null pronoun in the embedded sentence, the subject would clearly refer to the main clause subject *Todo mundo*$_1$ "Everybody$_1$" and possibly no overt pronoun could be used to refer to that antecedent in such a context. Modesto (2000*b*, 14) also mentions a similar example, when arguing against the loss of the "Avoid Pronoun Principle" in BP. Notice, thus, that the null and the overt subjects tend to be in complementary distribution in such contexts, rather than the overt pronoun replacing the null subject.

One approach to explaining these biases in the distribution of null or overt subjects is put forth by Modesto (2000*b*, 51-52) and Ferreira (2004, 19), who explain the behavior of null subjects by their relation with the topic position. According to them, quantified antecedents, such as *Ninguém* "Nobody" in (23) in *Todo mundo* "Everybody" in (24), are in topic position, either in the syntactic or semantic level because of so-called "quantifier raising" (May 1977, 1985, Huang 1982, 1984, *inter alia*). Interestingly, it has been proposed in the literature that other kinds of non-specific referents could be analyzed in a similar fashion (see Diesing 1992, von Fintel 1994). Taking these proposal into consideration, the present experiment also investigates the possible interaction with the underlying syntactic structure of such constructions by adding a secondary predication to the materials, as detailed below.

71

Given the assumptions and facts presented so far, the present experiment was designed in order to check whether (i) the different interpretations modulated by the specificity of the main clause subject are shared by naive BP speakers, (ii) a gender marked adjective can bias the interpretation toward one of the possible antecedents (also excluding the possibility of analyzing this structure as a by-product of ellipsis of a constituent bigger than VP, see Holmberg 2007, 2016, and Chapter 2 of the present dissertation) (iii) there is some difference in acceptability when the subject comes from context or from the same sentence, and (iv) the acceptability of the discourse-bound subject is different from the syntactic bound subject when combined with a secondary predication. If (iv) is relevant, it may indicate that the null pronouns in one or another interpretation are indeed syntactically different.[39]

Two binary Factors were tested: specific or non-specific (Factor "M[ain] C[lause] S[ubject]") and with or without secondary predication (Factor "Secondary_Predication"). The hypotheses were the following: (i) if Specificity plays a role in the interpretation of null subjects in BP (cf. what was observed in the corpus study), Factor Main Clause Subject will be significant: the null subject should be biased toward the discourse antecedent when the main clause subject is specific while the non-specific main clause subject should interfere in the co-reference with the discourse accessible antecedent; (ii) if the interpretation is helped by the agreement features, when the embedded verb is followed by a secondary predication whose gender is different from the gender of the main clause subject, the antecedent would be easily recovered; and (iii) if the syntactic structure plays a role, a significant interaction between the conditions in the acceptability judgment task would emerge, favoring the use of a secondary predication when the subject is co-referential with the main clause subject.

Materials

Twenty items were created based on the observations reported in the previous section. In order to have variation of possible referents and gender across items, *Todo X* "Every X" (X = a profession)

---

[39]   As mentioned in section 1.6, the binary criteria used in the corpus study is a first approximation. In this and in the next experiment, a cleaner case involving Specificity is tested. It is left for future work other kinds/degrees of Specificity.

instead of *ninguém* "nobody" or *todo mundo* "everybody", as in (23) and in (24), was used for the non-specific antecedent. As in the previous reported experiments, a context sentence was provided, such as (25) below. In this context sentence, an indirect question about the topic of the discourse makes a specific referent available for co-reference. This indirect question might have one of the four sentences in Table 1.8 as an answer. The relevant answer to the question is a null subject embedded clause whose main clause subject was either an specific or a non-specific referent (in A-B and C-D respectively). Following the verb that addresses the question, a gender-marked adjective was provided in (B) and (D), while not in (A) and (C). After judging this short dialogue composed of these two turns, participants were asked if the subject of the main clause assures/says/confirms that the discourse antecedent performs the action described by the question, cf. (26).

(25)   A – O João$_1$ é um dos principais advogados da firma de contabilidade. Você sabe se o João$_1$ trabalha na empresa até as 20h?

"A – João$_1$ is one of the leading lawyers of the accounting firm. Do you know if João$_1$ works at the company until 20h?"

	MCS	Adjective	Answer
(A)	specific	without	B – A Maria$_2$ diz que _$_1$/$_2$ trabalha. B – Maria$_2$ tells that (s)he$_1$/$_2$ does.
(B)	specific	with	B – A Maria$_2$ diz que _$_1$ trabalha lá sempre trancado. B – Maria$_2$ tells that he$_1$ works there always locked in.
(C)	non-specific	without	B – Todo advogado$_3$ diz que _$_1$/$_3$ trabalha. B – Every lawyer$_3$ tells that he$_1$/they$_3$ do(es).
(D)	non-specific	with	B – Todo advogado$_3$ diz que _$_1$/$_3$ trabalha lá sempre trancado. B – Every lawyer$_3$ tells that he$_1$/they$_3$ work(s) there always locked in.

Table 1.8 – Materials – Experiment 2

(26) A Maria$_2$/Todo advogado$_3$ afirma que o João$_1$ trabalha na empresa até as 20h?

"Does Maria$_2$/every lawyer$_3$ assures that João$_1$ works in the in the company until 8pm?"

   a. Sim.

     "Yes."

   b. Não.

     "No."

Based on the results from the corpus study and the previous reported literature, the empirical predictions were the following: when the main clause subject is specific, the null subject is preferentially used to retrieve the discourse salient antecedent (Conditions A and B) over the main clause antecedent, *i. e.* participants would answer more often "yes" in the interpretation task; on the other hand, when the main clause subject is non-specific, null subjects are more likely to pick up the main clause non-specific antecedent (Conditions C and D) than the discourse salient antecedent, *i. e.* participants would mostly answer the question negatively.[40] The secondary predication should have a significant effect when a specific antecedent is in the main clause, biasing the co-reference toward the discourse antecedent because of gender marking. As for the acceptability judgment task, significant interaction between Factors (MCS and Adjective) is thus expected if the two null subjects are syntactically different: when the discourse subject is picked up, the secondary predication should make the sentence less acceptable if this is a syntactic topicalization, because of both the linear and the structural length of the dependency, as shown in the structurally analogous examples in (27a) and (27b).

(27)   a. [Que advogado$_2$/O João$_2$ [a Maria$_1$ diz [que _$_2$ trabalha lá sempre [ec$_2$ trancado]]]]?

     "Which lawyer/João does Maria say works there always locked in?"

  b. [Que advogado$_2$/Todo advogado$_2$ [_$_2$ diz [que ec$_2$ trabalha lá sempre [ec$_2$ trancado]]]]?

     "Which lawyer/Does every lawyer say(is) that they/he work(s) there always locked in?"

---

[40]   The plausibility of the non-specific interpretation could be called into question: why should a speaker answer about lawyers in general in the context of a question about *João* in (materialsexp3judg)? However, this sort of answer is not infrequent when the speaker tries to skip directly addressing the question.

Although the structure in (27a) and (27b) will be slightly different according to which theory and analyses are assumed, the empty category to which the adjective is attributed is much further away and its antecedent should be harder to retrieve when the antecedent comes from the discourse if syntactic topicalization is assumed as the analysis for this co-reference (see Hornstein 1999, 2001, 2002, Marušič et al. 2003, among others). Notice that the interpretation task and the judgment task are expected to give contradictory results according to this hypothesis: the adjective should make the interpretation easier when the subject comes from the discourse (because the matrix clause subject has a different gender), but the sentence should be rated less acceptable.

Overall, null subjects should also be better rated when they are co-referent with the matrix clause subject, so Factor MCS should also be significant. Factor Secondary_Predication may also be significant, if participants take the sentences to be less acceptable because they provide additional information (namely that contained in the secondary predication) beyond the relevant answer to the polar question which is the question under discussion in (25).

Participants

One-hundred-fifty-two participants took part in this experiment. They averaged 30.6 years old (ranging from 24 to 37). In the data analysis, eight participants were discarded because they rated ungrammatical control sentences above eight. Otherwise, the procedure was as in the previous experiments reported here.

Results

The results partially confirm the hypotheses. As regards the interpretation of the null embedded subject, the participants interpreted it more often as co-referring with the main clause subject when it was non-specific than when it was specific (coded 1: main_clause_subject and 0: discourse_subject), cf. Figure 1.13 below. The main clause subject was picked up 30% of the times when it was specific, and 75% when it was a non-specific referent. The secondary predication, however, caused an unexpected effect: the unambiguous context, when the gender of the adjective unambiguously refer to the discourse subject, disturbed the co-reference with the discourse salient antecedent (increasing by 10% the negative answers to the interpretation question in the specific

MCS Condition).

Figure 1.13 – Percentage of Choices of MCS as the Antecedent of the Null Subject according to Main Clause Subject and Secondary Predication Factors
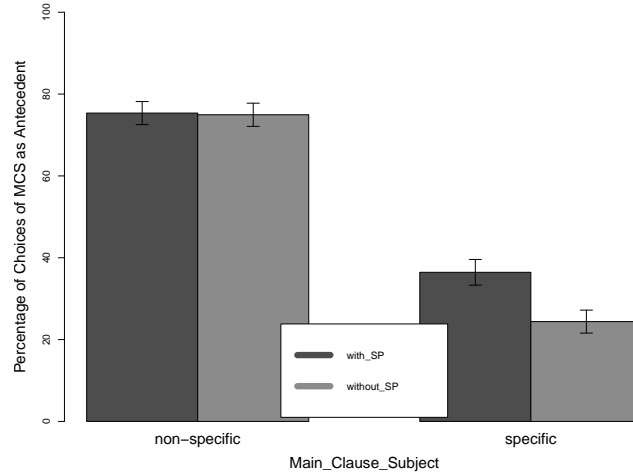
The results were entered in a mixed logit regression (Jaeger 2008). The binomial outcomes (coded 1: main_clause_subject and 0: discourse_subject) were regressed to analyze the choices of discourse topic vs. main clause subject as the null subject antecedent using *glmer* model with the optional 'logit' link function. The model included all main effects and interactions with a maximal-random structure (Barr et al. 2013). The full model reported here includes Subject and Secondary_Predication as fixed effects and the random effects (Participant and Item).

As shown in Table 1.9, Factor Subject and the interaction between both Factors were significant in the choices of the antecedent for the null embedded subject.

Condition	Estimate	Std. Error	Z-value	Pr(>|z|)
(Intercept)	1.7802	0.2122	8.390	< 2e-16 ***
Subject	-2.6716	0.2211	-12.085	< 2e-16 ***
Adjective	0.3778	0.2498	1.513	0.13
Subject:Adjective	-1.2583	0.2551	-4.933	8.09e-07 ***

Table 1.9 – Log-linear mixed-effects model for Answers in Experiment 3

Given the results obtained in the interpretation task, it would be worth considering the Interpretation as an additional Factor when analyzing the judgment task. This decision was taken after a first preliminary result analysis, since it is noticeable that an additional parameter could further

clarify the relation between the interpretation of the null subject and the mean acceptability of the sentences. This analysis is thus summarized in the Figures 1.14 and 1.15 below.

Figure 1.14 – Mean Acceptability Judgments according to Main Clause Subject and Secondary Predication Factors (Expected Answers)
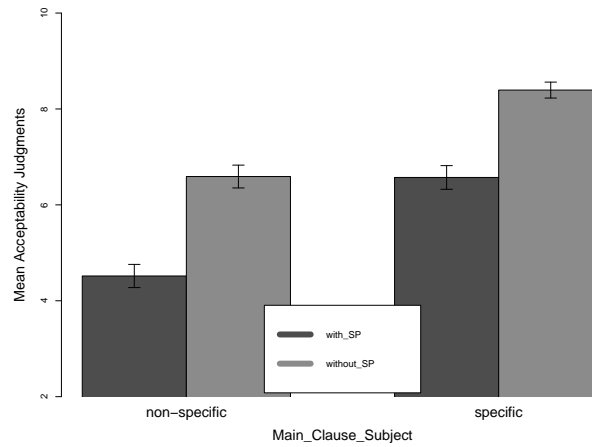
Figure 1.15 – Mean Acceptability Judgments according to Main Clause Subject and Secondary Predication Factors (Unexpected Answers)
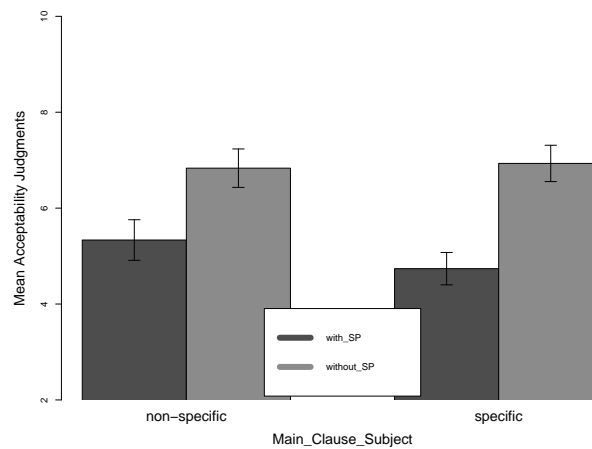
In Table 1.14, the expected answers are shown (72,5% of the answers), that is, when the null subject was considered the main clause subject in the non-Specific MCS Conditions and when the null subject was interpreted as coming from discourse in the Specific MCS Conditions. By comparing these two pairs of mean acceptability judgments, the sentences were judged more acceptable when (i) they directly addressed the question, that is, when the subject interpretation came from

discourse, and (ii) when they do not go any further than directly answering the question (without a secondary predication). However, almost the same difference of the mean acceptability judgments is observed in the two pairs of conditions: in the first case (when the non-specific MCS was interpreted as co-referential with the null subject), the sentences averaged 4.73/10 with a secondary predication and 6.6/10 without (difference: 1.87); in the second case (when the main clause subject was a proper name and the interpretation of the null subject came from the discourse), they averaged 6.57/10 with a secondary predication and 8.4/10 without it (difference: 1.83). As shown in Figure 1.15, this difference is indeed almost constant even in the non-expected answers, across the two types of MCS and Interpretations. In the case that the MCS is specific and the interpretation is co-referent with the main clause subject, the sentences were judged less acceptable with a secondary predication (4.5/10) and without it more acceptable than with the non-specific interpretation (6.93/10) (difference: 2.43). The case in which the null subject was interpreted as coming from discourse and the MCS was non-specific is not relevant here as it involves only a small number of cases in the interpretation task (less than 25%) and because it maintains the same difference as other conditions (difference: 1.5).

Both previously established Factors plus Factor Interpretation were thus entered as fixed effects in a linear-mixed model, taking the intercept and the individual fixed effects to interact with the random Factors (Participant and Item). The full model thus took into consideration 3 fixed Factors with two levels (MCS, Secondary_Predication and Interpretation) and two random Factors (Participant and Item). This model is reported in Table 1.10 below.

Condition	Estimate	Std. Error	T-value	Pr(>|t|)
(Intercept)	6.8928	0.2295	30.028	< 2e-16 ***
MCS	-0.2446	0.2340	-1.045	0.2973
Secondary_Predication	-1.9023	0.2644	-7.194	1.68e-11 ***
Interpretation	1.4489	0.2334	6.207	1.14e-09 ***
MCS:Secondary_Predication	-0.2834	0.2387	-1.187	0.2352
MCS:Interpretation	-1.3630	0.2755	-4.947	8.17e-07 ***
Secondary_Predication:Interpretation	0.1801	0.2538	0.710	0.4779
MCS:Secondary_Predication:Interpretation	0.8724	0.3482	2.506	0.0123 *

Table 1.10 – Log-linear mixed-effects model for Acceptability Judgments in Experiment 3a

As shown in Table 1.10, Factor MCS is not significant individually. Main Factors Interpretation and Secondary_Predication were significant. No significant interaction was found between MCS

and Secondary_Predication and Secondary_Predication and Interpretation, but a significant effect in the interaction MCS and Interpretation came up. Last, a significant effect was found in the three-way interaction among the main Factors.

Discussion

This experiment partially corroborates the findings in the corpus study suggesting that the feature Specificity plays a role in the interpretation of null subjects.[41] It is clear in the results of the interpretation task that non-specific antecedents somehow intervene in even highly salient, topical specific antecedent co-reference with null embedded subjects, as previously observed by Modesto (2000*a*,*b*) and Ferreira (2004). The secondary predication played an unexpected role in the interpretation question, though, decreasing the number of co-references to the discourse salient antecedent when only this antecedent was possible due to gender agreement. This unexpected result will be further addressed in the discussion of the next experiment.

The structural differences predicted by the topicalization-based approach turned out not to be relevant: rather the addition of a secondary predication has a constant effect across conditions and interpretations, except when the null subject was interpreted as the main clause specific subject. In that case, the secondary predication strongly decreased the acceptability of the sentences, possibly due to the above-mentioned gender mismatch. This gender mismatch may be the main reason for which this case is judged to be the least acceptable and for the three-way interaction found in the inferential statistical analysis. In general, the main significant factors were the semantic and pragmatic effects, that is, when the sentences directly addressed the question and did not provide information going beyond it, they were significantly more acceptable.

1.7.5 Experiment 3b – Re-running Specificity of the Antecedent

Taking into consideration the results obtained in Experiment 3a and the observations from the corpus, such as example (24), Experiment 3a was re-run with slightly different materials. Instead of

---

[41] An alternative interpretation of these results is to take the bias to be related to the quantified antecedent (see in this regard the eye-tracking study by Cunnings et al. 2014). This analysis will be discussed in more detail in section 1.8 below.

having only null embedded subjects and varying whether the main clause subject is specific or not, only non-specific main clause subjects were provided along with null and overt subjects in the embedded clauses. With this set-up, it is expected to observe different preferential interpretations for the overt pronoun, which is supposed to prefer specific antecedents, given the corpus observations, such as in example (24). The interpretation of the null subject must still preferentially be the non-specific main clause subject. The secondary predication was kept as in the previous version of this experiment, in order to compare sentences with an overt pronoun and without it as regards their underlying syntactic structure.

Materials

Twenty items were created based on those reported in the previous section. As in the previous experiment, a context sentence was provided, such as (28) below. In this context sentence, an indirect question about the topic of the discourse makes a specific referent available for co-reference, as before. This indirect question might have one of the four sentences in Table 1.11 as an answer. The relevant answer to the question was an embedded clause either preceded by a null subject or by an overt pronoun (in A-B and C-D respectively). Following the verb that addresses the question, a gender-marked adjective was provided in (B) and (D), while not (A) and (C). After judging the short dialogue composed of these two turns, participants were asked if the subject of the main clause assures/says/confirms whether either the discourse antecedent performs the action described by the question or the non-specific antecedent does, cf. (29).

(28)   A – O João$_1$ é um dos principais advogados da firma de contabilidade. Você sabe se o João$_1$ trabalha na empresa até as 20h?

"A – João$_1$ is one of the leading lawyers of the accounting firm. Do you know if João$_1$ works at the company until 20h?"

	Subject	Adjective	Answer
(A)	null	without	B –Todo advogado$_2$ diz que $\_{1/2}$ trabalha. B –Every lawyer$_2$ says that he$_1$/they$_2$ do(es).
(B)	null	with	B –Todo advogado$_2$ diz que $\_{1/2}$ trabalha lá sempre trancado. B –Every lawyer$_2$ says that he$_1$/they$_2$ work(s) there always locked in.
(C)	overt	without	B –Todo advogado$_2$ diz que ele$_{1/2}$ trabalha. B –Every lawyer$_2$ says that he$_1$/they$_2$ do(es).
(D)	overt	with	B –Todo advogado$_2$ diz que ele$_{1/2}$ trabalha lá sempre trancado. B –Every lawyer$_2$ says that he$_1$/they$_2$ work(s) there always locked in.

Table 1.11 – Materials – Experiment 3b

(29) Todo advogado$_2$ afirma que o João$_1$/$\_2$ trabalha na empresa até as 20h?

"Does every lawyer$_2$ assert that João$_1$/$\_2$ works in the in the company until 8pm?"

a. Sim.

"Yes."

b. Não.

"No."

Based on the results from the corpus study and from Experiment 3a, the empirical predictions were the following: given a non-specific main clause subject, a null subject is preferentially used to retrieve such a non-specific antecedent (Conditions A and B) over a discourse salient antecedent, *i. e.* participants are expected to provide more "no" answers in the interpretation task when it has a proper name and more "yes" answers when it does not; on the other hand, when the embedded subject is overt, it is more likely to pick up the discourse salient antecedent than to refer back to the main clause non-specific antecedent (Conditions C and D), *i. e.* participants would mostly answer "yes" to the question with a proper name and "no" to the null subject version. The secondary predication should not have a significant effect in the interpretation task, since the gender does not solve the ambiguity. As for the judgment task, no significant interaction between Factors (Subject and Secondary_Predication) is expected if the null and the overt subjects are not syntactically different; now, regardless of interpretation, both the linear and the structural length of the co-reference is constant if the empty category is like a pronoun, as shown in the structural analogous examples in (30a) and (30b). If there is a structural difference, it is likely that the empty category is not like a pronoun and it could be generated by "movement out of case domains" (as in Ferreira

2004, Rodrigues 2004, built on Hornstein 1999, 2001, 2002). In this case, a significant difference in acceptability is expected, since linear and structural length would be different, favoring the acceptability of the overt subject over the null subject.

(30)   a.   [O João$_1$/Todo advogado$_2$ [$\__2$ diz [que ele$_1$/$_2$ trabalha lá sempre [ec_1/$_2$ trancado]]]].

       b.   [O João$_1$/Todo advogado$_2$ [$\__2$ diz [que ec_1/$_2$ trabalha lá sempre [ec_1/$_2$ trancado]]]].

            "João$_1$/every lawyer$_2$ says that he$_1$/they$_2$ work(s) there always locked in."

Notice that the interpretation task and the judgment task are expected to have convergent results in this experiment: the adjective should not interfere in the interpretation of the embedded subject (because the matrix clause subject and the discourse antecedent have the same gender) and the sentences should be rated less acceptable with it, as in the previous experiment. Overall, overt subjects should be better rated since they are supposed to address the discourse-salient referent of the question. However, Factor Secondary_Predication is expected to show a main effect, since participants take the sentences to be less acceptable for going beyond the minimal information to answer the question in (28).

Participants

As in Experiment 1.

Results

The results summarized in the Figure 1.16 below partially contradicted some of the hypotheses. Although a clear inversion of preference of the overt subject was found when the verbs were not followed by a secondary predication, this element disturbed the biases. When the embedded subject was an overt pronoun, it was taken to be co-referent with the discourse subject in 80% of the total answers; but this bias decreased to 65% when the verb was followed by a Secondary Predication. The null pronoun was considered co-referent with the discourse salient antecedent in 35% of the total of answers, but the figure rose to the chance level (50%) when it was followed by a secondary predication. This pattern is different from the previously reported experiment, where the gender removed the ambiguity in the discourse-antecedent-biased condition. Notice also that

the Secondary Predication had the strongest effect in the condition in which absolutely no effect was found in the previous experiment.

Figure 1.16 – Percentage of Choices of Discourse Salient Referent as Antecedent according to Embedded_Subject and Secondary_Predication Factors
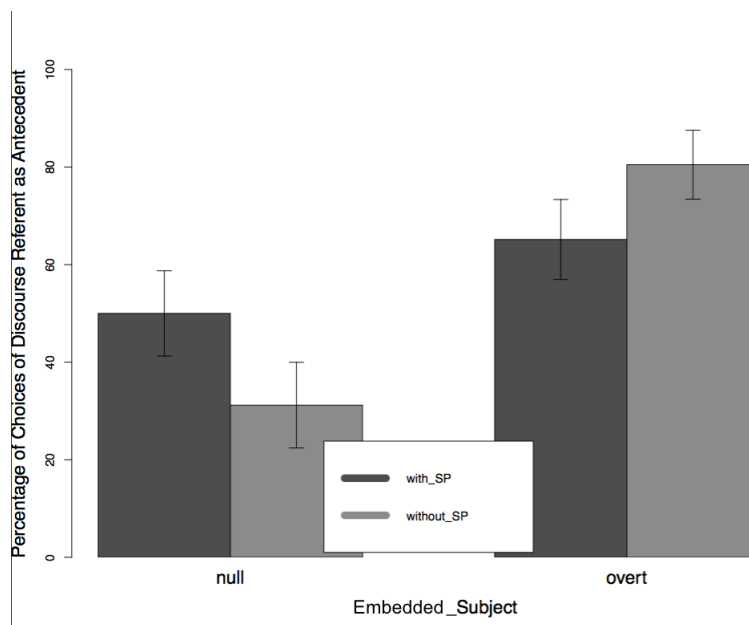
This result needs to be carefully disentangled. In fact, since the interpretation was checked indirectly by the questions in (29), the type of question proposed could have had an effect. And indeed it does: the interpretation was disturbed in the null pronoun plus secondary predication only when the question had the format *Todo advogado$_2$ afirma que $\_2$ trabalha na empresa até as 20h?* "Does every lawyer$_2$ assert that $\_2$ works in the company until 8pm?" (63% of unexpected negative answers vs. 84% of expected answers in the other type of question); the interpretation had also a less strong bias in Condition Overt Subject plus Secondary Predication only when the proper name was in the question, such as in *Todo advogado$_2$ afirma que o João$_1$ trabalha na empresa até as 20h?* "Does every lawyer$_2$ assert that João$_1$ works in the company until 8pm?" (50% of unexpected negative answers vs. 70% of expected answers in the other type of question). Looking back to the question, the problem is that the Secondary Predication seems to cancel a VP-ellipsis interpretation that is possible when the adjectival phrase is not present, such as in (32a) vs. (32b) for the Null Subject Condition. Since the question presupposed that the whole VP content would be recovered, the participants with more strict pragmatic interpretation answered negatively

to the question when followed by a Secondary Predication, as in (32b), because of the lack of the temporal adjunct content.

(31) A –O João é um dos principais advogados da firma de contabilidade. Você sabe se o João$_1$ trabalha na empresa até as 20h?

"A –João is one of the leading lawyers of the accounting firm. Do you know if João$_1$ works at the company until 20h?"

(32) a. Todo advogado$_2$ diz que $\_2$ trabalha [na empresa até as 20h].

"Every lawyer$_2$ says that they$_2$ do [= work at the company until 8pm]."

b. Todo advogado$_2$ diz que $\_2$ trabalha lá sempre trancado.

"Every lawyer$_2$ says that they$_2$ work there always locked in."

Similar assumptions can be made about the Overt Subject Condition. Given this unexpected orthogonal effect, the type of question was also entered as a predictor in the inferential statistical analysis. So, as in the previous experiment, the results were entered in a mixed logit regression (Jaeger 2008). The binomial outcomes (coded 0: main_clause_subject and 1: discourse_subject) were regressed to analyze the choices of discourse topic vs. main clause subject as the embedded subject antecedent using *glmer* model with the optional 'logit' link function. The model included the main effects and interactions with a maximal-random structure (Barr et al. 2013). The full model reported includes Subject, Secondary_Predication and Q[uestion]_T[ype] as fixed effects. The random effects were Participants and Items.

Condition	Estimate	Std. Error	Z-value	Pr(>|z|)
(Intercept)	-0.32522	0.27037	-1.203	0.22903
Subject	2.03956	0.37807	5.395	6.86e-08 ***
Secondary_Predication	-1.38922	0.56849	-2.444	0.01454 *
Q_T	-0.03855	0.44055	-0.087	0.93028
Subject:Secondary_Predication	-0.34248	0.68617	-0.499	0.61770
Subject:Q_T	-0.50357	0.63087	-0.798	0.42474
Secondary_Predication:Q_T	2.32769	0.78313	2.972	0.00296 **
Subject:Secondary_Predication:Q_T	-0.79327	0.91371	-0.868	0.38529

Table 1.12 – Log-linear mixed-effects model for Answers in Experiment 3b

As shown in Table 1.12, the main effect of the type of subject is a significant predictor for the

interpretation of the embedded subject. Another significant effect was the Secondary_Predication; however, since the interaction between Secondary_Predication and Q[uestion]_T[ype] was also significant, it is likely that this effect comes from the non-recoverability of the VP-content when the Secondary Predication was present and the question presupposed this content to be true. No other significant effect was found.

As for the judgment task, taking into account the results of the interpretation task, only expected answers were taken into account in the analysis (70.2% of the total amount of data), since it would be hard to incorporate the VP recoverability as a Factor. So, in Figure 1.17 below only the judgments about the embedded null subjects interpreted as the main clause non-specific antecedent and the overt subjects taken to be co-referent with the discourse salient antecedent are shown.

Figure 1.17 – Mean Judgments according to Embedded_Subject and Secondary_Predicate Factors
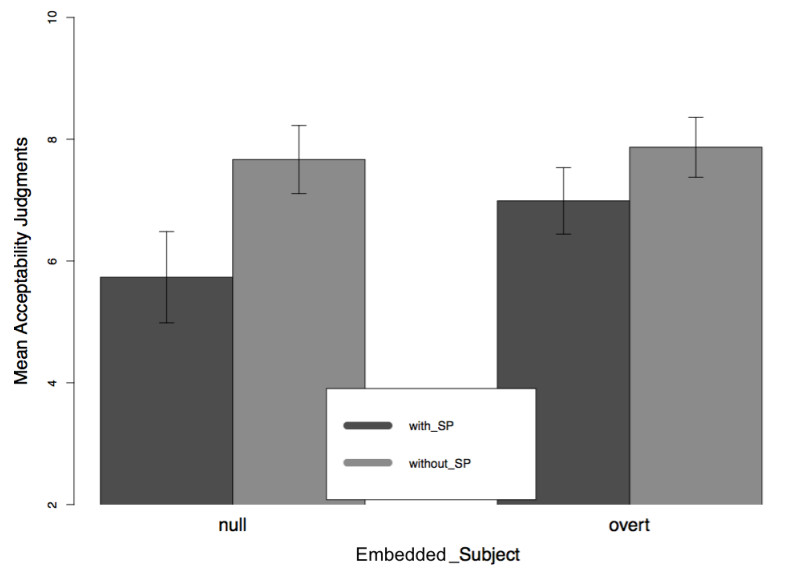
As shown in Figure 1.17, the mean acceptability judgments do not diverge too much from those found in Experiment 3a. The Overt Subject Condition was judged more acceptable without the Secondary Predication (averaging 7.86/10), while slightly less acceptable with a Secondary Predication (mean: 6.78/10). The Null Subject Condition was slightly less acceptable, but the difference between with Secondary Predication (mean: 5.73/10) and without it (7.6/10) is similar.

As for the inferential statistical analysis, as in the previously reported experiments, the out-

come judgments were entered in a linear mixed model. The main Factors were Subject and Sec-ondary_Predication, and the random Factors were Participants and Items. The maximal model is summarized in Table 1.13 below.

Condition	Estimate	Std. Error	T-value	Pr(>|t|)
(Intercept)	5.7365	0.5123	11.197	1.07e-11 ***
Subject	1.3609	0.4332	3.141	0.00545 **
Secondary_Predicate	1.7383	0.6399	2.717	0.01182 *
Subject:Secondary_Predicate	-0.9320	0.6225	-1.497	0.14728

Table 1.13 – Log-linear mixed-effects model for Judgments in Experiment 3b

As observed in Table 1.13, main Factors Subject and Secondary_Predication are significant. Crucially, the interaction between main Factors did not reach significance.

Discussion

Apart from the unexpected VP-recoverability effect, the results provide strong confirmation for the previously established predictions. Embedded overt subjects, in this context, favored the specific antecedent ("jumping" over the main clause non-specific subject), while embedded null subjects maintained their bias toward the main clause non-specific subject. This effect was significant across the board, while the effect of the Secondary_Predication was modulated by the Question_-Type, which can be explained by the VP-recoverability effect. In fact, the same problem can explain the 10% difference in Condition "MCS Specific" with a Secondary_Predication in Exper-iment 3a. The secondary predicate does not seem to have an effect on the interpretation of the embedded subject, beyond the fact that it disturbs expected positive answers to the questions in the experiment; however, further investigation needs to be carried out in order to check whether the effect of Factor Subject does not disappear in a non-VP-ellipsis context.

Overall, Factor Secondary_Predication had a significant effect in the judgment task, as pre-dicted, which can be attributed to the fact that these sentences went beyond the information that was required to answer the question. Overt subjects were judged better than null subjects, possi-bly because they addressed the question more directly, but this effect seems to be independent of the syntactic structure of the sentence. It seems that the acceptability judgments were more af-fected by pragmatic problems in the sentences with respect to how they address the questions than

by underlying syntactic differences across sentences. Crucially, regardless of interpretation and possible length dependence variation, the acceptability of the sentence remained constant across conditions and no significant interaction between the kind of subject and the use of a secondary predication was found. This finding strongly favors an hypothesis according to which null and overt subjects have similar syntactic structures, and are equally able to bind secondary predications. Further investigation about the nature of syntactic restrictions on the use of null subjects are reported in Chapters 2 and 4. In this experiment, the crucial finding is the variable bias in interpretation depending on the different available antecedents and the use of different anaphoric options, as previously observed in the literature.

1.8 General Discussion

Impoverishment. Since the 80s, the relation between the impoverishment of verbal morphology and the availability of referential null subjects in a given language has been observed and assumed to be quite direct (Taraldsen 1980, Chomsky 1981, Rizzi 1980, 1982*a*, 1986, Huang 1984, Picallo 1984, Jaeggli & Safir 1989, Contreras 1991, Rohrbacher 1994, Speas 1994, Cardinaletti 1997, *inter alia*). Some diachronic data, as well as sociolinguistic data has been taken to support this correlation (Roberts 1993, Roberts & Kato 1993, Vance 1997, among others). BP has been considered a key piece of evidence to support such claims (Lira 1982, Tarallo 1983, Galves 1987, 1993, 1997, Duarte 1993, 1995, 2000, 2012, among many others). However, no single analysis has succesfully accounted for all observations that emerge from the empirical data.

In this chapter, further corpus and experimental evidence that challenges the correlation between the impoverishment of verbal morphology and the drop in the number of null subjects in BP has been presented: in particular, 3rd person null subjects, whose verbal morphology agreement is widespread and multiply ambiguous, are more frequent than, for instance, 1st person singular null subjects, even when the 1st person singular verb is marked with an exclusive inflection, as shown in Figure 1.6. However, a small effect of verbal inflection is still observed: when the 1st person singular is not marked by an exclusive agreement inflection, the decrease in the number of null subjects in corpora and in the acceptability of the sentences without overt subjects in Experiment 1 seems to be relevant.

In the frameworks assumed by the previous literature, these results could be approached through two different accounts: (i) the 1st person singular exclusive agreement marker is a trace of the previous organization of BP system and, for this reason, the high number and the higher acceptability of null subjects with these verbal markings are also a trace of the previous system (Lira 1982, Duarte 1993, 1995, 2000, Cyrino et al. 2000, Kato & Duarte 2014); or (ii) in the current system of BP, not all the null subjects are the same (Figueiredo-Silva 1994, Negrão & Müller 1996, Modesto 2000*b*, Ferreira 2004, Rodrigues 2004, Holmberg 2005, Barbosa et al. 2005, Barbosa 2009), and specific verbal inflections can license the null subject (Barbosa 2011, based on Ritter l995, Shlonsky 2009). Note, moreover, that the two hypotheses suggested in this paragraph do not exclude each other, although they have been taken as mutually exclusive by most of the literature about BP.

The research based on assumption (i) faces a substantial difficulty in explaining the data presented here. Kroch (1989) shows evidence from various linguistic change phenomena that the time course of syntactic change is constant across a set of linguistic contexts and based on the grammar of the changing language.[42] He shows that the context might have generally differed from one to another at each period of time in the degree of preference for one form over another, but they do not diverge in the rate of the change.[43] Evidence presented in the present dissertation are from three periods of time, the 70s and 90s, in the corpus research, and the 2010s in the experimental trials. In all of them the effect remained constant.

If the effect of the 1st person singular exclusively marked on the relative number of null subjects was a trace from the previous system of BP, it should have been weakened or have disappeared at least in Experiment 1 (also due to the fact that a reading and judgment task is much less sensitive to subtle manipulation than a production task). However, clearly the effect is present and constant throughout the past fifty years.[44] Also, Kroch (1989) shows that the constant rate of the change is not observable on surface properties, such as the form of a word or morpheme, but it is rather

---

[42] See Correa Soares (2016) for further discussion of diachronic data and for a hindcast statistical analysis of the change in BP and French. The focus of the present dissertation is rather the synchronic BP system.

[43] Kroch (1989) analyzes evidence over different periods: when analyzing the replacement of *have* by *have got* in British English, he refers to examples about 50 years distant in time. In order to further understand the diachronic path of null subjects in BP, it is worth looking at data taken from a larger time interval, such as most of the data provided by Kroch (1989).

[44] Similar figures are found in other diachronic research, such as Cyrino et al. (2000), Berlinck et al. (2009), Correa Soares (2016), *inter alia*.

based on a shared in-depth abstract grammatical analysis on the part of the speakers.[45]

So, the apparently superficial effect of the morphologically overt exclusive inflectional marking and the relative low frequency of 3rd person null subjects should be expected to have a deeper foundation in the grammar of current BP. In the present dissertation, this is accounted for through the different levels of Complexity in Ariel (1990)'s scale of Complexity/Explicitiness. In her scale, null anaphors and inflectional markings are not at the same level of Complexity: null subjects that are related to agreement markers, such as those of standard pro-drop Romance languages, are higher (more Complex and Explicit) than actual "zero expressions", such as those of pro-drop languages that lack agreement markers altogether (Chinese Mandarin, for instance) (Ariel 1994, 2001, Arnold 1998, de la Fuente et al. 2016, *inter alia*). It is assumed here that BP has a mixed system, which explicitly and exclusively marks some 1st persons (these are, thus, higher in Ariel 1990's scale), while the other null subjects are, as in Chinese Mandarin, real "zero expressions". This explains the weak effect of exclusive marking found in the present dissertation in both experimental and corpus data. As for the 3rd persons, the unequal change observed across persons and the relevant semantic types (as proposed by Cyrino et al. 2000, Kato & Duarte 2014) must also be considered not only a direction of change and/or a focus of resistance of the previous grammar, but also as a consequence of deep grammatical knowledge of BP speakers, who appear to tend to a synchronic specialization of function ("division of labor") for each form in competition, as in the research based on (ii).

The authors who assume (ii) usually take the syntactic contexts to favor the use of 3rd person null subjects, considering them pronominal anaphors (Figueiredo-Silva 1994, 2000, Barbosa et al. 2005), bound variables (Negrão & Müller 1996, Müller 2003, Muller 2005), silent variables (Modesto 2000*b*,*a*, 2008*a*,*b*) or even traces of movement (Ferreira 2004, Rodrigues 2004), but none of them has considered the effect of the 1st subject exclusive vs. ambiguous agreement in the choice between overt and null subjects. Although this effect has no drastic consequences against these analyses, most of them do not account for this effect altogether.

The only article that assumes (ii) in which a possible explanation for this effect is provided

---

[45] More details about the data analyzed by Kroch (1989) and specific claims about grammatical diachronic change can be checked in the original paper. For the purposes of the present dissertation, the consequences of his theory for the synchronic system are enough.

is Barbosa (2011), building on (Modesto 2008*a*,*b*). As mentioned in section 1.3.4, her account is based on the fact that null subjects are minimally specified NPs (much like a bare noun), whose definite reading must be triggered compositionally by checking a D-feature (as in Holmberg 2005). Mentioning data from Hebrew and Russian (languages that also have exclusive inflectional markings for some tenses but not for others), she suggests that a type shift operation takes place when the D-feature is checked. In a brief outline of future work, the author cites Ritter (l995), who proposed that verbal agreement in Hebrew has a D feature only in Past and Futures tenses, and Shlonsky (2009), who claims that 1st and 2nd person agreement morphemes in Hebrew are incorporated subject clitics and that 3rd person agreement enters an underspecified person slot. These proposals seem *mutatis mutandi* to provide a possible account for the data observed in this chapter.[46] However, following Barbosa's reasoning, the type-shifting in BP is triggered by movement to a topic position (as in Modesto 2008*a*,*b*, *inter alia*), rather than by a specific tense-agreement combination. The data provided throughout this dissertation suggest that the covert-topic-based approach does account for the use and interpretation of null subjects in BP. The question of topicalization will be detailed in what follows, but before the 3rd person null subjects must be addressed.

Regardless of the specific analysis proposed, most of the approaches that assume (ii) attempt to explain the higher number of 3rd person null subjects by the possible constructions in which they can appear. Based on Kato (1999) and Holmberg (2005), Barbosa et al. (2005), for example, argue that when they are logophorically related to another referent in a previous linguistic context, 3rd person null subjects appear more frequently than other persons.[47] Rodrigues (2004) discards the logophoric interpretation and claims that the 3rd person null subjects are syntactically anaphoric and are exceptions in the non-pro-drop grammar of BP, generated by "movement out of case-domains", such as obligatory control structures (see also Ferreira 2004, *inter alia*). Modesto (2000*b*, 2008*a*,*b*) also argues for an analysis based on topicalization, much in the spirit of Huang (1982, 1984, 1989), who proposes a variable in the topic position co-indexed with the empty subject position. Notice that, independently of the analysis assumed, in these proposals the number of

---

[46]   The reader is referred to Barbosa (2011) for further details.
[47]   According to Holmberg (2005), co-referential null subjects are distinct lexical items, similar to logophoric pronouns (Sells 1987). They are not deleted, but rather they block the deletion of their non-logophoric counterpart. Barbosa et al. (2005), thus, proposes that, in BP, null subjects prefer a c-commanding antecedent in syntax; when such an antecedent is not found, co-reference can be logophoric (Landau 2000).

3rd person null subjects are a by-product of the constructions they appear in, rather than an inherent property of the 3rd person. However, statistically, it is unlikely that 3rd person null subjects are more frequently in, for instance, embedded clauses (or any other kind of c-command syntactic environment), than 1st and 2nd person null subjects. All these researchers are, thus, forced to assume that 1st and 2nd person null subjects are syntactically different from 3rd person null subjects, which, besides being a stipulation, only displaces the discussion about the statistical difference of their relative frequency to another level of the grammar, but does not shed much light on the relation between discourse person and the use of null subjects. Rather, these approaches assume that this difference is a statistical coincidence.

Semantic Features. The statistical significance of the number of inanimate null subjects in BP and the significant difference in acceptability of null vs. overt subjects referring back to inanimate or animate antecedents is absolutely not addressed by any of the previously mentioned approaches (with the exception of those which assume the Referential Hierarchy). Rather, some of them would predict the contradictory outcome (see Kato 1999, Barbosa et al. 2005, Modesto 2008*a,b*, Barbosa 2011, *inter alia*): since inanimate antecedents tend not to be topical, and according to them null subjects do tend to be topical, the statistical correlation should be the inverse (more overt subjects to refer to non-topical inanimate antecedents). However, if only 3rd person animate subjects are considered, the number of overt subjects is parallel to the number of overt subjects in 1st and 2nd persons (as previously observed by Creus & Menuzzi 2004, Berlinck et al. 2009), when taking into consideration only specific antecedents. This suggests that there is no way to derive this effect from, for instance, the discourse-pragmatic features of person, by arguing that the addressee or the speaker have an especial syntactic status in the left periphery, or in the CP-domain Sigurdsson as recently suggested by 2004, 2012, Sigurdsson & Holmberg as recently suggested by 2008.[48]

---

[48] As suggested by Anne Abeillé, it is possible to propose that 1st and 2nd person pronouns started as discourse topics with a null subject, and then became reanalyzed as subjects. So there would be two constructions: (i) one with a dislocated topic (and a null subject) maybe reserved for some matrix clauses (with no other topic) and 1st and 2nd person subjects and (ii) another with a true pronominal subject (and maybe another discourse topic). Although that is a possible analysis and it is not unheard of in the literature (see, for example, Kato 1999), being a topic does not account for many facts about null and overt subjects in BP, as detailed in the sections about verbal Yes-No answers in BP in Chapter 2 (see especially sections 2.4.1 and 2.5). To rule out this specific analysis based on dislocated topics goes beyond the purposes of the present dissertation and requires further research about dislocated (and duplicated) subjects in BP in both corpus and experimental data.

The initial observation about the relevance of the "degree of referentiality" in the relative frequency of null and overt subjects is taken to be a predictive directional theory of language change (Cyrino et al. 2000, Kato & Duarte 2014), that is, the path by which the change is affecting the whole system. Once again, however, it is worth remembering the data and claims advanced by Kroch (1989): the change occurs at a regular rate across linguistic contexts according to the speakers' in depth grammatical knowledge. Considering the corpus data from two periods and the significant results from Experiment 2 (whose participants speak present day southern BP), the hypothesis suggested in this chapter is that the "animacy" feature has a synchronic effect on the choice between overt and null subjects in BP, besides its possible diachronic effect, favoring null subjects when their antecedents are inanimate.

Beyond this semantic feature, Cyrino et al. (2000), Kato & Duarte (2014) also claim that diachronically Specificity is a reason for the higher number of null subjects in 3rd persons. This has already been observed by other researchers who claim that null subjects exhibit a bound variable behavior (see Negrão & Müller 1996, Müller 2003, Muller 2005, Modesto 2000*b*, among others) or a syntactically anaphoric pattern (Menuzzi 2002, Branco 2007, Menuzzi & Lobo 2016). The higher frequency of non-specific null subjects gathered in the corpus study was indeed significant. The comparison between the preferential interpretations in Experiments 3a and 3b is also striking evidence for the synchronic relevance of Specificity in the null and overt subject use and interpretation. However, some considerations are due in order to correctly interpret the importance of the findings reported here as regards the approaches proposed in the previous literature. These considerations are discussed in what follows.

As pointed out before, the findings of the present dissertation are surprising for some of the theories of anaphora resolution based on the reversing mapping hypothesis. Depending on the concept of "Salience" of the antecedent taken into account and its empirical coverage, non-specific and inanimate antecedents should be taken to be at lower levels of Salience scales (see Silverstein 1976, Dahl & Fraurud 1996, among others). The fact that they are retrieved by less complex forms provides counter-evidence against an anaphora-resolution-based approach for null subjects in BP and against the universality of the notions of salience/prominence generally accepted in the relevant literature. However, as predicted by Ariel (1990)'s notion of Accessibility, the linguistically relevant notion of Salience must be established within the discourse. Accepting this notion

of Salience, antecedents of these semantic kinds are expected to be at the top of their respective hierarchies (of Animacy and Specificity). In Chapter 2, this theory is incorporated into the general approach proposed in the present dissertation.

Deriving Semantic Features from Syntax.   First of all, consider the hypothesis that 3rd person null subjects are generated by movement either by (a) obligatory control movement out of case domains (Ferreira 2004, Rodrigues 2004, *inter alia*) or by (b) a variable in the left periphery (either TopP or [spec, AgrP]) coindexed with a null argument (Modesto 2000*b*, 2008*a,b*, Barbosa 2011). Let's start with (a). One of the main arguments for this analysis are locality effects. Ferreira (2004) shows that a closer antecedent is preferred by the null subject, while an overt one can co-refer with a more distant antecedent, even if the latter is non-specific, cf. (33) below.

(33)   a. Nenhum menino$_1$ disse que a Maria$_2$ acha que *$_{-1}$/$_{-2}$ é inteligente.

      b. Nenhum menino$_1$ disse que a Maria$_2$ acha que ele$_1$ é inteligente.

         "No boy$_1$ said that Maria thinks [he$_1$] is intelligent."

                                            (Ferreira 2004, 42)

There is no doubt that there is a strong preference for the closest antecedent in (33a) and co-reference with the non-specific higher NP is marginal. It is also clear that the overt pronoun in this case can co-refer with the non-specific subject, mainly because of gender agreement. But if the genders match, the preference for the closest referent is also observed when the overt pronoun is used, as in sentence (34a). Moreover, if both antecedents have the same gender and both are specific, what is observed is a complementarity of reference, with the null subject biased toward the closest and the overt pronoun toward the more distant antecedent, as in (34b).

(34)   a. Nenhuma menina$_1$ disse que a Maria$_2$ acha que ?ela$_1$/ela$_2$ é inteligente.

         "No girl said that Maria thinks she$_{?1}$/$_2$ is intelligent."

      b. A Joana$_1$ disse que a Maria$_2$ acha que ela$_1$/?ela$_2$/?$_{-1}$/$_{-2}$ é inteligente.

         "Joana$_1$ said that Maria$_2$ thinks (she)$_1$/$_2$ is intelligent."

                                      (adapted from Ferreira 2004, 42)

Also, as the results of Experiment 3a have shown, no effect of locality can be found when the antecedent of the null subject is the most salient in the discourse context and both the antecedents are specific, such as in Conditions A and B in Table 1.8. In fact, when dealing with these contexts, the analysis based on (a) assumes some sort of structure like (b) (see Ferreira 2004, 9, for instance). Furthermore, besides multiplying the number of necessary mechanisms to generate null subjects, this proposal only displaces the problem to another level of the analysis, but it does not cast any light on the biases of interpretation of null and overt subjects. In Experiments 3a and 3b, for example, according to these accounts, interpreters need to decide which syntactic structure is the underlying representation of each sentence. However, the relation between Specificity, locality and the syntactic underlying representation is not straightforward. To conclude about locality: it is clear that it is relevant for the interpretation of the anaphoric subject, however it seems that it is not the only factor that plays a role; rather, multiple factors that generate some sort of imbalance or contrast between possible antecedents constrain the possible interpretations of null and overt subjects in these contexts.

Modesto (2000*b*) proposes an account that could possibly explain the interpretation preferences in Experiment 3a and 3b. He proposes that the null subject in embedded clauses is a genuine *pro* (in the sense of Chomsky 1981, 1986), whose restrictions of interpretation are tighter in BP than in other Romance languages. For the BP *pro* to acquire an interpretation, it must be co-indexed with a nominal element in an A-bar position (Modesto 2000*b,a*). Assuming that the non-specific subjects in Experiment 3a and 3b are moved somehow, co-reference with the discourse subject is disrupted. The discourse interpretation is indeed achieved by a null Topic in the left periphery. Also, by assuming that any non-specific referent has to move to some position in the left periphery of the sentence, either in syntax or in Logical Form, Modesto's theory could also explain the high number of non-specific null subjects found in corpora.[49] At first sight, the only objection that could be made based on the data provided in the present dissertation is that Modesto's proposal does not account for the effects of animacy. On the other hand, observing island effects on null objects in BP, Bianchi & Figueiredo Silva (1994) propose an account according to which animate

---

[49]   Presumably additional assumptions will be required to get the right scope for non-specific entities in the semantic representation. This hypothesis will not be explored here, but it is not unheard of in the literature.

null objects are bound variables (as in Ferreira 2004's analysis of null subjects) while inanimate ones are genuine pronominal forms. This analysis does not seem very convincing because, once again, it displaces the problem rather than accounting for it. Indeed there is nothing inherent to such syntactic devices that predicts the possible antecedents. Barbosa (2011) could account for the animacy effect through her null NP anaphora hypothesis, but the requirement of movement to a topic position does not fit the results presented in the present dissertation, as explained below.

One of the arguments against any analysis based on either (a) or (b) is the possible effect of the length of the dependency in cases with Secondary Predication. It is well known in the literature about psycholinguistics that, for instance, the length of wh-gap-dependencies shows significant effects on sentence comprehension (Crain & Fodor 1985, Frazier 1987, Frazier & Flores d'Arcais 1989, Stowe 1986), on sentence production (Crain & Fodor 1993), on ERPs (Felser et al. 2003, Phillips et al. 2005), on judgments of perceived complexity (Gibson 1998) and acceptability judgments (Phillips et al. 2005). If null subjects in BP are somehow generated by a similar mechanism (that is to say, either movement in the case of (a) or co-indexing the A-bar element to a gap when assuming (b)), an effect on acceptability would be expected when the obligatorily controlled secondary predication is used to refer to either the discourse topic subject or the main clause subject in Experiment 3a, as well as some effect either on the pronominal option or the null subject in Experiment 3b. Such an effect was not found, however. This suggests that the null subject in subordinate clauses is more similar to a pronoun than to another kind of empty category, such as a trace or a variable.

A Proposal and Further Investigation.   Having said that no previous account seem to fully account for the data presented in this chapter, a way to further investigate the null subject in BP will be proposed. What seems to govern the distribution and interpretation of null and overt subjects is a "local specialization": speakers seem to try to optimize contrast taking into account not only *in presentia* elements but also the paradigmatic possible structures that are available in the language but not used in a given context (following Baumann et al. 2014, de la Fuente et al. 2016, among others).[50] In this chapter, three sorts of contrast were identified: paradigmatic verbal marking,

---

[50]   Similar ideas have also been proposed in previous literature (see Levinson 1987, 1991, Safir 2004).

animacy and specificity. The present survey does not intend to be exhaustive, though. Probably there will be more sources of contrast among possible antecedents and between these two anaphoric options, which can possible include distance effects, syntactic function, discourse factors, among others. This position supports the theory advanced by Müller (2003), Negrão & Müller (1996) that null subjects in BP have developed some sort of specialization rather than disappearance. However, an important remark on any of these proposals is that whatever constraints are proposed on the use of null subjects, these constraints cannot be formulated as hard binary properties: although the investigation conducted here has taken into account these three factors as binary, it seems that they are in fact scalar. This scalar status is reflected in the non-categorical numbers in corpora and in, for example, the divergent judgments reported by Müller (2003) and Menuzzi (2002) regarding non-specific antecedents, as in (35) below.

(35) a. Um milionário$_1$ sempre acha que __$_1$/?ele$_1$ é muito poderoso.

"A millionaire$_1$ always thinks he$_1$ is very powerful."

b. Um milionário$_1$ sempre precisa acreditar que os outros$_2$ acham que ele$_1$ é muito poderoso.

"A millionaire$_1$ always needs to believe that the others$_2$ think he$_1$ is very powerful."

c. Todo milionário$_1$ acha que __$_1$/??ele$_1$ é muito poderoso.

"All millionaires$_1$ think they$_1$ are very powerful."

d. Nenhum político admite que __$_1$/??ele$_1$ recebe dinheiro por fora.

"No politician admits that they receive money on the side [=illegally]."

e. Ninguém admite que __$_1$/*ele$_1$ recebe dinheiro por fora (para aprovar...)

"Nobody admits that they receives money on the side [=illegally]."

(adapted from Menuzzi 2002)

For Müller (2003), (35a) is acceptable with both the null and the (stressed) overt pronoun; while Menuzzi (2002) indicates that overt variant may be a little marked in contrast with the null subject, but that it certainly does not need to be focused to be acceptable – as Müller (2003) claims. However, Menuzzi (2002) straightforwardly accepts (35b), even without particular intonation. (35c) and (35d) are generally considered to be less acceptable with the overt embedded subject, and finally (35e) is generally considered odd. Also, in the corpus study reported here, no

overt pronoun co-referring to "nobody" was found, though a few cases were found when referring to other non-specific antecedents.[51] Similar assumptions can be made about Animacy.

The main aim of this chapter was to understand the problematic relation between the impoverishment of the verbal inflectional paradigm and the lower number of null subjects in BP than in other Romance languages. In particular, the main goal was to disentangle the relation between discourse persons and the distribution of null subjects across discourse persons, with especial regard toward the number of null subjects in the more ambiguous forms. In this endeavor, the findings presented in this study align with three claims in the previous literature: (i) BP does not show the effects of a direct relation between the impoverishment of verbal morphology and the reduced frequency of null subjects (Taraldsen's generalization) in a straightforward way, but rather seems to be changing on a multiply triggered path (Negrão 1990, 1997, Modesto 2000*b*, 2008*a*,*b*); (ii) the lower number of null subjects in BP does not seem to be a by-product of the lack of licensing properties, but rather a specialization in the way the null subject is identified (Rizzi 1986, Negrão 1990, 1997, Modesto 2000*b*,*a*, 2008*a*,*b*, Müller 2003, Branco 2007, Menuzzi & Lobo 2016); and (iii), as a corollary of (ii), the study of null subjects in BP must bring together two apparently divergent areas, that have been recently reconciled by a number of researchers, as in Corrêa (1998), Costa et al. (1998), Carminati (2002), Luegi (2012), Filiaci et al. (2013), Luegi et al. (2014), Almor et al. (2017), namely anaphora resolution and the parametric theory. If the change in BP is a matter of identification of the null subject by any possible mechanism, the study of this identification has to be integrated into the anaphora resolution literature. Inevitably the anaphora resolution research has also had to deal with the possibly divergent anaphoric options cross-linguistically, and so the study of a language such as BP, whose null anaphoric option is undergoing a change, is crucial to understanding the nature of the anaphoric relations. With this study, BP could possibly be integrated into a typologically more precise category.

---

[51]   In Experiments 3a and 3b, an intermediary degree of specificity was tested: quantified non-specific antecedents, such as (35c). As Patricia Cabredo-Hofherr suggested, the results found in these experiments might be interpreted as a result of the so-called bound variable behavior of null subjects (see, among others, Montalbetti 1984). Here, quantified non-specific antecedents are taken to be one degree of the scale of Specificity, which favors co-reference with null subjects. However, in order to insure that this is not an especial pattern of quantified phrases, it is worth testing non-quantified non-specific antecedents, such as (35a), in future research.

1.9 Interim Summary

In this chapter, it was argued that approaches based on the assumption that impoverished verbal morphology triggers a ban on null subjects (Taraldsen's Generalization) cannot account for the relative frequency distribution of null and overt subjects across discourse persons in BP. Some attempts in the literature to deal with this distribution were summarized, but most of them cannot explain the puzzling quantitative differences in the BP system, found in both corpus and experimental research. It was claimed that the very existence of a language such as BP is intriguing from a typological, parametric and cognitive point of view. Based on previous literature, especially on empirical corpus research and on the Referential Hierarchy presented in section 1.3.5, three features were studied: Verbal Marking, Animacy and Specificity. It was observed that BP speakers tend to avoid ambiguity in the expression of the subject, but they do not avoid redundancy as much. This result was found both in the corpus and in the experimental research. The decisive factors that govern the distribution of null and overt subjects appear to be the inherent semantic features of the antecedents, namely Animacy and Specificity. These two factors were also significant in both corpus and experimental research, though different methodologies were used to check their relevance. Although some of the empirical research procedures were carried out in a binary fashion, these two factors are thought to be scalar in nature, as in the Referential Hierarchy. The full proposal about how to incorporate these features into a coherent grammar of anaphoric subjects will be spelled out in Chapter 2. It will be argued, as in the present chapter, that rather than a licensing mechanism, the BP subject anaphoric system relies on general properties of identification, which are tightly related to anaphora resolution.

Though the relevance of certain factors has been demonstrated, some problematic issues remain unanswered: are there other factors that might additionally favor the use of null or overt anaphoric subjects in BP? Syntactic contexts seem to be a relevant factor according to the previous literature both on BP and on other languages. Discourse factors, such as Relative Salience, initially discussed in section 1.4.2, are also reported to influence anaphoric choices cross-linguistically. These issues will be addressed in what follows. In Chapter 2 below, the full model for accounting for anaphoric null and overt subjects will be presented, along with its comparison with other cognitive models, especially with those which account for anaphora resolution in other languages.

Chapter 2

Out of Focus – Beyond the Topicality of Null and Overt Subjects in Brazilian Portuguese in the Syntax-Discourse Interface

2.1 Introduction

This chapter addresses the use of null and overt subjects in BP according to the discourse and syntactic context in which the sentence occurs. It starts from the investigation of the simplest and most controllable dialogue pair in BP with respect to information structure, namely, a Yes-No question followed by what will be called a verbal Yes-No answer [vYNA]:

(36)   a.   A – você$_1$ leu [o livro]$_2$?

"Have you read the book?"

   b.   B – _$_1$ Li   _$_2$.
        _ read.PST.1SG _
        "Yeah[, I have]."

(OPUScorpora, "Entre Nós")

In (36a), Speaker A asks a Yes-No question, namely whether Speaker B has read the book, Speaker B can answer affirmatively by uttering only an appropriately inflected finite verb, in this case the 1st person singular, as in (36b).[1] This is the simplest discourse context in which the alternation between overt and null subjects can be found in BP.[2] This chapter aims to understand the constraints associated with this context, to extend them to an immediately more complex context, namely, unary wh-questions, and to analyze further extensions based on data from corpus. An account based mainly on the underlying organization of the discourse (information structure) is put forward, under the view that the discourse is structured to answer underlying questions in a coherent way (Question under Discussion approach) (Roberts 1996, 2012, Ginzburg 1996, 2012, Büring 2003, 2005, *inter alia*). The main claim of this chapter is that a principle of Maximal Contrast, along with a mapping constraint, constrains null subjects in present day BP to contexts in which the referent

---

[1]   Although vYNAs might appear to be similar to VP-ellipsis in English, there are some differences both in the contexts they can be used (which will be explored throughout the present chapter) and in the form they can take (vYNAs can be constituted of an auxiliary and a main verb, for example).

[2]   For some authors, verbal answers to Yes-No questions in BP are considered a context of categorical null subjects (see Simões 1997, Magalhães 2000, among others). It could also be argued that, when a subject is used, the verbal answer is no longer a strict answer to a yes-no question, but rather goes further than only addressing the polarity of the question (by clarifying the topic, for example). In the present dissertation, however, the empirical notion of vYNA is taken to have a broader coverage, as any sentential answer used to address a yes-no question.

of the subject is obvious, non-contrastive and backgrounded, and the focused material contrasts the most with it, that is, the whole overt linguistic material is focused.

After proposing an account in terms of the structure of the discourse for the use of null and pronominal subjects in vYNAs in BP, extensions of this proposal are analyzed in the corpus. The results suggest that the discourse based proposal can generalize over other accounts and bring together many observations made in the literature about the pro-drop parameter and anaphora resolution. Finally, the design of a grammar that accounts for the use and interpretation of null and overt subjects in BP is drawn: a multiply constrained probabilistic scalar grammar is proposed and discussed as regards the constraints observed in Chapters 1 and 2.

Outline of Chapter 2

This chapter is organized as follows. In section 2.2, some reasons to take vYNAs into account as evidence to discuss the pro-drop parameter in BP are presented. In section 2.3, in the first three subsections, previous analyses proposed to deal with the use of null and overt subjects in vYNAs are outlined and, in the fourth subsection, the main proposal of the present dissertation is presented. In the last subsection, the discussion is summarized. In section 2.4, the mechanisms proposed in the literature are applied to a set of data from NURC-RJ and the previous theories are contrasted with the proposal put forward here. In the second subsection, the discussion is summarized and the design for an experiment about vYNAs and their contribution to the study of pro-drop is discussed. In section 2.5, an experiment that tests different discursive contexts and syntactic structures is reported. In section 2.6, the previous proposals concerning anaphora resolution cross-linguistically are compared to the proposal presented here. The significance of the effects predicted by different theories of anaphora resolution are tested in a corpus statistical survey in section 2.7. Finally, in section 2.8, a summary of the relevant factors found here is given and the design of a grammar that can deal with these multiple factors is proposed.

2.2 The Role of Verbal Answers for a Grammar of Null Subjects

In this section, it will be argued that vYNAs must be well understood in order to clarify the use of null subjects in BP. Henceforth the discussion presented here largely follows that of Magalhães & Santos (2006), who propose that vYNAs are a relevant set of data both for a general theory of pro-drop and for the acquisition of subjects in BP. This approach will be discussed in the light of a broader context of research when necessary.

2.2.1 The General Importance of Verbal Answers for the Theory of Pro-Drop Parameter in Brazilian Portuguese

It is *a priori* unexpected that verbs not preceded by an overt subject pronoun are the standard way to answer Yes-No questions in BP, since BP is usually taken to be a language in which null subjects are becoming scarcer, as mentioned in the previous chapter and in much of the previous literature (Tarallo 1983, Galves 1987, 1993, 1997, Duarte 1993, 1995, 2000, Kato 1999, 2000, Modesto 2000*b*, Ferreira 2004, Holmberg et al. 2009, *inter alia*). If only the contrast between non-pro-drop and standard pro-drop languages is taken into account, it is possible to observe a contrast between BP and English or French (whose subjects are required in similar contexts) and no difference between BP and European Portuguese (see Martins 2004). In the literature about the pro-drop parameter, however, vYNA have been neglected (see, for instance, Duarte 1995, Kato 1999).

According to Magalhães & Santos (2006), previous research about null subjects in BP generally excludes verbal answers from their database. This is an unjustified analytical decision, because vYNAs are not only examples of the use of null subjects in (almost all) pro-drop languages (see Martins 2004), but they are also constructions in which null subjects are maintained in "partial" pro-drop languages, such as BP and Finnish (see Holmberg 2001, 2013). Magalhães & Santos (2006) observe that verbal answers are excluded in, for instance, Simões (1997), Lopes (2003), Magalhães (2000, 2003), because they would categorically be preceded by null subjects. Surprisingly, the research referred to by Magalhães & Santos (2006) also counts the lack of literature on

vYNAs as an argument to set them apart (see also Duarte 1995). Therefore, most of this literature claim that including this case could lead to a misunderstanding of both the relative proportion of null subjects and the analysis of the pro-drop phenomenon in BP.

Magalhães & Santos (2006) propose that this context must be taken into consideration and that it confirms their hypothesis that null subjects in BP are not disappearing, but rather that they are restricted to more specific contexts, as proposed in Chapter 1 and elsewhere in the literature (see Negrão 1990, Negrão & Müller 1996, Müller 2003, Branco 2007, Menuzzi & Lobo 2016, *inter alia*). They provide three arguments to support this idea: (i) vYNAs in BP are instances of null subjects in adults' grammar (this argument will be summarized in section 2.2.2 below). (ii) The exclusion of this set of data would only be justified if vYNAs were utterances without an internal structure. However, the literature about this case suggests that these constructions either are created by or, at least, contain instances of VP ellipsis (cf., for instance Matos 1992, Martins 1994*a*, see also discussion in section 2.3 below). If this analysis holds, vYNAs in BP would be full sentences in which the verb is pronounced and the rest of the sentence is covert (this proposal will be analyzed throughout this chapter).[3] Finally, (iii) when comparing pro-drop and non-pro-drop languages, in a non-pro-drop language (for instance, English), vYNAs have obligatorily overt subjects, while pro-drop languages typically allow null subjects.

Lastly, Magalhães & Santos (2006) makes two claims that will be further pursued in this chapter. First, verbal answers are not a context in which the omission of the subject is categorical, as opposed to what is suggested by Simões (1997), Lopes (2003), Magalhães (2000, 2003). And second, overt subjects in vYNAs raise discourse effects in languages in which a null subject is allowed: pronominal subjects, for instance, favor a contrastive, non-exhaustive reading (see sections 2.6, 2.7 and 2.8).

2.2.2 Brief Discussion of the Theory of Acquisition of Answers

This subsection is a very brief discussion about the acquisition of vYNAs and the pro-drop parameter in BP and EP following the work of Magalhães & Santos (2006). The aim of this subsection

---

[3] The fact that vYNAs have some internal structure also rules out the possibility of analyzing them as bare verbs (possibility mentioned by Patricia Cabredo-Hofherr).

is to support the idea that, even if the use of null subjects in this context can be taken as very construction specific, it is part of the inner competence of native speakers, regardless of their level of education or influence from contact with other varieties of Portuguese (for instance, the European Portuguese literature taught at school).[4] There is abundant literature about the acquisition of null subjects in BP (see the research cited herein and Magalhães 2007, for an overview). The discussion will not be further developed in this chapter. Some consequences of the theory and analysis proposed in this chapter for the acquisition of null subjects in BP is left for future work.[5]

According to Magalhães & Santos (2006), since the earliest stages of the process of acquisition, there is evidence to claim that verbal answers conform to the target adults' grammar. They claim that even verbal answers uttered during first stages of acquisition are not mere repetitions.[6] They base this claim upon the three following arguments.

(i) vYNAs occur independently from the position in which the verb is in the question — a rightmost overt element vs. other positions. (37b) could be analyzed as a simple repetition of the question in (37a). However, it is observed that children choose the right target of the question in their answers, as in (38), in this case the rightmost verb of the parent's question. If children only target the rightmost verb, as a preferential antecedent for repetition, the answer recovering the main verb (which is leftmost) in (39) would be unexpected. Following Magalhães & Santos (2006)'s reasoning, uses like this show children's attempts to recognize the syntactic structure of the question or identify the VP as the domain to be recovered in the answer.

(37) a. MÃE   - você₁ espera   eu aqui?
          MOTHER - you   wait.for.PRS.3SG I   here
          "Do you₁ wait for me here?"

---

[4]   Some research has claimed that null subjects are mainly taught at school to BP speakers, directly or indirectly by contact with EP literature or formal prescriptive written language (see, for instance, Magalhães 2003, Simões 1997). In this section, however, evidence that preschool-aged children productively use null subjects in vYNAs is given.

[5]   Due to the restricted availability of BP native speaking children in France, experiments and other sorts of data analysis about acquisition were not possible. This area of linguistic research will be explored in further research.

[6]   There is a vast literature about repetitions in discourse and conversational analysis (see, among many others, Wong 2000, Sidnell 2010), as well as in language acquisition (e. g. Masur 1995, Gathercole 2006). The answers analyzed here could be a product of the process of acquisition or some sort of linguistic strategy of repair. However, it is claimed throughout this chapter that this is not the case for BP vYNAs.

b. RAQ: _$_1$ espero.
   CHILD: wait.for.PRS.1SG

   "Yes(, I$_1$ do)."

(adapted from Magalhães 2007, 105, ex. 7b)

(38) a. TEL   – você$_1$ não quer... não vai   dar   comida para ela?
       MOTHER - you   not want   not go.PRS.3SG give.INF food   to   her

       "you$_1$ don't want... you$_1$ aren't going to feed her?"

   b. ANA   - _$_1$ vou.
       CHILD -   ir.PRS.1SG

       "Yes(, I$_1$ am)."

(adapted from Magalhães & Santos 2006, 184, ex. 7)

(39) a. MÃE   - _$_1$ Vamos   ver   se a gente$_1$ acha   a cabeça?
       MOTHER -   go.PRS.1PL see.INF if the people find.PRS.3SG the head

       "Are we$_1$ going to see whether we$_1$ find the head?"

   b. RAQ   - $_1$ Vão.
       CHILD - go.PRS.3PL

       "Yes(, we$_1$ are.)"

(adapted from de Oliveira 1996, 5, ex. 13)

(ii) Answers with a focal element copied from the question are rare (contra Kato & Tarallo 1993). According to Kato & Tarallo (1993), children could build their answers over any focused element present in an interrogative utterance, such as the demonstrative pronoun in (40).

(40) a. TEL: você que(r)   esse mesmo?
       MOTHER -   you   want.PRS.3SG this same

       "do you really want this?"

   b. ANA   - Esse.
       CHILD - this

       "Yes(, I$_1$ do)."

(adapted from Magalhães & Santos 2006, 186, ex. 15)

However, Magalhães & Santos (2006) claim that children actually produce mostly vYNAs, whose foci are not a specific overt constituent but the assertion as a whole, such as those in (37)-(39). Constituent focused answers are exceptions and quantitatively rare, as shown in Table 2.1.

Table 2.1 – verbal answers vs. answers with another focused constituent

Child	Verbal answers (the total out of the five first sessions)	Constituent-focused answers (the total out of the five first sessions)
Tomás	86	1
Inês	55	4
(*Inês M.*)	94	3
Raquel	19	0
Ana	41	3

(adapted from Magalhães & Santos 2006, 186, Table 4)

(iii) Comparing a BP-speaking child and an EP-speaking child in Tables 2.2 and 2.3 below, to take into account vYNA or to exclude vYNA data highlights the difference between a classical pro-drop language (EP) and BP.

Table 2.2 – Referential Null Subjects by Raquel (BP)

Sections	Null Subjects(including vYNAs)	Null Subjects(excluding vYNAs)
01	82.3%	73.8%
02	42.7%	36.2%
03	60.1%	51.4%
04	55.7%	45.8%
05	55.2%	48.1%
06	49.3%	33.9%
average	57.5%	48.2%

(adapted from Magalhães & Santos 2006, 189, Table 6)

Table 2.3 – Referential Null Subjects by João (EP)

Sections	Null Subjects(including vYNAs)	Null Subjects(excluding vYNAs)
01	83.3%	78.5%
02	46.4%	43.1%
03	77.8%	77.7%
04	67.7%	64.3%
05	70.8%	68.4%
06	77.8%	75.4%
average	70.6%	67.9%

(adapted from Magalhães & Santos 2006, 186, Table 4)

The results presented on Table 2.2 show that the BP-speaking child uttered an average of 57.5% of sentences with null subjects when verbal answers are taken into account. If they are excluded, the average number of null subjects is reduced by 9.3%. When the verbal answers are excluded from the set of data produced by the EP-speaking child, the decrease in the percentage of null subjects is almost irrelevant, since it undergoes a much smaller fall of 2.7%.

What is observed in the discussion summarized above is that (i) children's vYNAs are produced by complex linguistic analysis, rather than mere repetitions, and conform to the target adult grammar; (ii) their productions of null subjects (when the analysis includes this context) are neither equivalent to those of a speaker of a non-pro-drop language, nor to those of a child acquiring a standard pro-drop language;[7] and (iii) vYNAs with null subjects are part of the internal grammar of native speakers of BP since early stages of the acquisition, and not a result of contact with other variants or with Portuguese taught at school. So, vYNAs are a robust argument against the idea that null subjects are disappearing in BP, as previously proposed in Chapter 1; they point to the conclusion that, at most, null subjects are becoming more constrained – but the relevant constraints do not seem to affect the possibility of having (*i. e.*, the licensing of) null subjects in vYNAs.

---

[7]   Magalhães & Santos (2006) take this fact for granted, although they do not systematically study productions of any non-pro-drop-language speaker (see Hyams et al. 2015, for an overview). From their argument, it is possible to understand that the claim is that children acquiring BP do not produce vYNAs similarly to adult speakers of non-pro-drop languages (e. g. English).

2.3   Accounting for Verbal Answers in BP

In this section, earlier approaches on the use of null and pronominal subjects in vYNAs in BP are introduced. In fact, there is no consensus on relevant issues, which raises a debate on the topic. In subsection 2.3.1, earlier analyses of vYNAs are summarized. From there, two positions are developed: a pragmatic-based account and a syntactic-based account. In subsection 2.3.2, the first position is presented: it takes pro-drop in vYNAs to be pragmatic in terms of expectancy of the answer or topicality of the referent of the subject. In subsection 2.3.3, the second position, whose central claim is that the null and overt subject alternation in vYNAs is a result of structural constraints, namely some sort of deletion in syntax, is briefly summarized and discussed. Finally, in section 2.3.4, an alternative account is developed: by formalizing the structure of the discourse and the relation between turns in a dialogue, a hypothesis for accounting for the use of null and pronominal subjects in question-answer pairs is proposed.

2.3.1   Early Proposals: Structural vs. Pragmatic Accounts

Analyzing dialogues in English, Stubbs (1983) proposes a radical position according to which there is no syntactic relation between turns. According to him, links between turns are established at the propositional level or purely by means of illocutionary force. However, for the study pursued in the present dissertation, this position cannot be maintained. Corpus and experimental data inconsistent with this claim will be presented throughout this chapter. First, verbal answers are likely to be one of the clearest contexts of the pragmatics-syntax interface. If absolutely no syntactic relation exists between turns, verbal answers should not obey any kind of syntactic restrictions imposed by the previous contexts. This is however claimed to be false here. For example, BP does not easily allow a verb in the past tense whose subject is not overt, and whose antecedent is animate and specific, as in (41a) below (as discussed in Chapter 1). But in the context of the answer in (41b), the bare verb is perfectly acceptable and highly frequent in corpora (see also Table 2.4).

(41)   a. A – E o vendedor$_1$/*$_1$ soube $_2$ depois?

       "and did the seller know [that the lottery ticket won the prize] after?"

   b. B – $_1$ soube $_2$.
           _ know.PST.3SG _
       "Yeah[, he did]."

                                           (adapted from NURC-RJ, Inquiry_r_52)

Moreover, as discussed in section 2.2.1, vYNAs are studied as regards general principles of the grammar and are claimed to obey underlying rules that also apply to full sentences. This is observed in the fact that, while pro-drop languages allow null subjects in these answers, in non-pro-drop languages an overt subject is absolutely obligatory (as in English VP-ellipsis). A third reason presented so far to be skeptical of Stubb's position is children's data: the deep linguistic analysis that is required to target the right verb in, for example, the contexts of main-embedded questions, such as in (39), shows that there is some sort of syntactic knowledge, which is present from the first stages of acquisition, driving the kind of answers that are possible given a certain question.

Analyzing BP data from a different position, Kato & Tarallo (1993) deal with yes-no answers that are direct speech acts (excluding answers related to questions by implicature). These direct speech acts can be formed by an assertive particle (*sim* "yes" or *não* "no") and/or by a sentential form – that is, a vYNA. According to them, in fact particles have no functional or argument structure, which avoids the possibility of attributing reference to possible arguments related to the arguments in the question (see Holmberg 2016, Ginzburg 2012, for different positions). The meaning of the answer is indeed claimed to be related to the question only at a semantic/propositional level. However, Kato & Tarallo (1993) claim that, when using a vYNA, the syntactic relation is clear and direct and, against Stubbs (1983), this relation must be established at a syntactic/structural level.

Kato & Tarallo (1993) observe some evidence for postulating that vYNAs have an internal syntactic structure. The first observation is that the essential element – *i. e.*, the verb – is an element that carries the inflection.[8] The second is that vYNAs have a specific internal structure,

---

[8]   Kato & Tarallo (1993) also empirically observe that this element must be a phonological word. According

which in BP seems to lack of a subject and an object. Concerning this internal structure, they carried out a corpus survey, in which only the usage of overt pronominal subjects were counted. This corpus analysis is summarized in Table 2.4 below.

As shown in Table 2.4, the particle *sim* "yes" followed by a verb was only found in null subject constructions; without an affirmative particle, the verb was preceded by an overt subject in only 3.4% of the cases. Crosslinguistically, Kato & Tarallo (1993) mention that Italian, for example, has categorical null subjects in contexts where the antecedent is such obvious, but BP vYNAs are variable (despite the percentage around 96% of null subjects in vYNAs). They conclude that, even if BP is not a classical pro-drop language, it is much closer to a pro-drop language when only vYNAs are taken into consideration. Unfortunately, they do not provide any further explanation for this proximity.

Table 2.4 – The internal structure of vYNAs

index	SE–arg	SE+arg	P+SE–arg	P+SE+arg
number of tokens	169	6	6	0
relative percentage	96,6%	3,4%	100%	0%

SE = sentential form, *i. e.* vYNAs
P+SE = particle *sim* "yes"+ sentential form , *i. e.* vYNAs
+/– arg = overt or null subject

(adapted from Kato & Tarallo 1993, 273, Table 3)

---

to them, in languages with clitics, it is not possible to reduce the VP to V, because the arguments are mor-phophonologically attached to the verb as clitics. In the corpus analysis carried out in the present dissertation, however, examples such as (1) below, which contradict this hypothesis, were found.

(1) a. A   - Agora uma coisa: o   senhor$_1$ se lembra   do,   do   nascimento do   seu
        DOC - Now one thing the lord   SE remember.PRS.3SG of.the of.the birth   of.the your
        pri... primeiro filho$_2$?
        []   first   child
        "Now, one thing: do you[formal]$_1$ remember your first child's birth$_2$?"
     b. B – $_1$ lembro $_2$, perfeitamente.
              remember.PRS.1SG ,   perfectly.
        "Yeah[, I$_1$ do], perfectly"

(NURC-RJ, Inquiry_071)

Urbano et al. (1993) take questions and answers to be the minimal dialogue unit. Following Stubbs (1983)[9], they claim that the two-way relation between these pairs is undeniable. Questions anticipate and semantically constrain answers and seem to depend on them. Answers, in turn, depend on questions even more. According to them, questions impose illocutionary restrictions on answers, making explicit whether a possible answer is adequate or not; appropriate answers satisfy these conditions.

As regards Yes-No questions specifically, referring to Stubbs (1983), Urbano et al. (1993) claim they are expected to receive a positive answer *sim* "yes" or a negative *não* "no". They observe that other elements whose meaning is similar can also be used, as in the following dialogue:

(42) a. A – Inf.   : (...) Hair (...) você$_1$ não assistiu?   Você$_1$ assistiu,
         Interviewed :   Hair   you   not watch.PST.3SG you   watch.PST.3SG
       né?
       [DM]
       "Hair, didn't you$_1$ [informal] watch [something]? You$_1$ [informal] watched [something], didn't you?"

    b. B1 – Doc1.   : uhn   uhn.
         Interviewer1 : [DM]
       "Yeah[, I have]."

    c. B2 – Doc2.   : _$_1$ Assisti.
         Interviewer2 :   watch.PST.1SG
       "Yes(, I$_1$ did)."

(adapted from Urbano et al. 1993, 77, ex. 16)

In dialogue (42) above, both interviewers answer the question positively, but with different expressions: Doc1 utters an agreement/assent marker in (42b); Doc2 uses a verb in (42c). Urbano et al. (1993) claim that there is a preference for elliptical answers in BP. Moreover, because elliptical answers (and discourse particles like *sim* "yes" and *não* "no") are used as answers to questions, and questions can only be expressed linguistically, they cannot be used as the initial sentence in a discourse segment. For this reason, they are rarely found outside of dialogical contexts (spoken or

---

[9]   In fact, Urbano et al. (1993) reference a Spanish version, published in 1987.

written), except for rhetorical questions (cf. Stubbs 1983, 116).[10] Based upon Moeschler (2003), Urbano et al. (1993) claim that the illocutionary sequential restrictions ("satisfying conditions") are neither necessary nor sufficient, but rather impose a scale of adequacy on possible answers. Urbano et al. (1993) propose that the conditions are the following:

(i) topic maintenance: an answer needs to be thematically related to the question;

(ii) propositional content adequacy: an answer must semantically refer to the question, by means of paraphrase, implication or opposition;

(iii) illocutionary force compatibility: the illocutionary type of an answer must be compatible with the illocutionary type of the question;

(iv) argumentative orientation similarity: the argumentative orientation of an answer needs to be the same as the argumentative orientation of the questions.

Urbano et al. (1993) apply these conditions to the following dialogues:

(43) a. A – Você$_1$ pode   sair   hoje à   noite?
              You   can.PRS.3SG leave.INF today to.the night
          "Can you$_1$[informal] hang out tonight?"

     b. B1 – Sim.

          "Yes."

     c. B2 – Não, _$_1$ não posso.
              No   not can.PRS.1SG
          "No, I$_1$ can't."

     d. B3 – Você$_1$ quer   me levar pra sair?
              You   want.PRS.3SG me take to   leave.INF
          "Do you$_1$[informal] want take me out?"

(adapted from Urbano et al. 1993, 83, ex. 28)

In dialogue (43), B1's answer in (43b) satisfies all the conditions. As mentioned above, another possible, similar answer to this question would be _$_1$ *Posso.* "_ can.PRS.1SG". B2's answer in (43c) satisfies the same conditions, except (iv). B3's answer in (43d) (which in fact is a question)

---

[10]   Urbano et al. (1993) also note the occurrence of null subject verbal answers with post-verbal negation. About this matter, see, e. g., Schwenter (2005).

only satisfies (i) and (ii). Only B3's answer has an overt subject, so that, apparently, "illocutionary force compatibility" plays a role in the use of overt subjects. This is unpredicted for an approach based upon only structural constraints. This point will be further discussed in the following parts of this chapter.

2.3.2   Purely Pragmatic Accounts

Martins (1994*a,b*, 2004) provides a more elaborated account for the analysis of vYNAs in Portuguese. In Martins (2004), vYNAs in EP are analyzed as consisting of a VP-ellipsis structure in which the subject is inside the verbal projection (VP) and the verb itself moves from T (the tense projection) to Σ (the polarity projection).[11] For her, after the verb moves, the remnant VP containing the subject is elided. For the purposes of the present dissertation, this is a possible analysis for EP, since it does not have to check the EPP-feature in TP, and the subject can stay inside VP; however, this analysis cannot be applied to BP, in which subjects have to move to Spec of TP, and would not be affected by VP-ellipsis. The best derivation that this proposal would predict for the data at stake here is a post verbal subject in Spec of TP, but it would not account for null subjects without additional assumptions.[12]

In Martins (1994*a*, 182), EP pre-verbal subjects occupy [Spec, ΣP] and are interpreted as unmarked topics. In fact, in Martins (2012), she assumes that "bare verb answers have a certain degree of structural autonomy with respect to the antecedent-question". Eventually she ends up claiming that, in verbal answers, the post-verbal (VP-internal) constituents are structurally recovered under (syntactic) identity with the antecedent in the question (VP-ellipsis), but preverbal (VP-external) constituents may be pragmatically recovered. According to her, vYNAs do not have to interpretatively depend on the whole antecedent and, actually, besides recovering the subject, they may be unable to pragmatically recover certain preverbal adverbs.

---

[11]   According to Martins (2004), the spell out of the bare verb is triggered by the LCA and morphological reanalysis of the verb in Σ. According to Kayne (1994), the LCA (Linear Correspondence Axiom) is a principle ordering terminal nodes, which establishes that, if a non-terminal category A c-commands a non-terminal category B, all terminal nodes dominated by A must precede all terminal nodes dominated by B.

[12]   It is generally accepted in the Minimalism-based proposals that in BP the EPP feature must be checked in a spec position higher than TP, since the Verb-Subject order when the verb is intransitive or transitive is quite restricted (see, for instance, Kato 2000, Barbosa et al. 2005).

Since the present dissertation is primarily interested in the use of preverbal overt subjects and null subjects in vYNAs, the set of Martins' papers is taken to be a proposal to deal with null subjects at pragmatic level. Martins (2004) is thus considered an exception and, as pointed out, the analysis proposed in her paper does not account for BP, because of the standard assumptions about EPP in the theory she assumes. To sum up, Martins (1994*b*, 2012)'s proposal is viewed as an attempt to deal with null subjects as unmarked topics (cf. Martins 1994*b*), with certain degree of pragmatic recoverability (following Martins 2012).

Armstrong (2008) mentions that structurally vYNAs are well studied in BP (by Kato & Tarallo 1993, Martins 1994*a*), but a more developed pragmatic analysis is necessary. According to her, Portuguese has an "echo system" (following Sadock & Zwicky 1985's typology), in which "no especial answer words are used".[13] The standard way to answer to polar questions is thus simply the use of a verb.

Following Sorjonen (2001), who analyzes the Finnish echo system, Armstrong (2008) claims that speakers choose different responses based on pragmatic restrictions as they convey several meanings, such as a basic confirmation, an offer of new information or a confirmation of an assumption. In the same spirit as Sorjonen (2001), she explains how the use of these responses is triggered by specific pragmatic conditions which creates a "multiplicity in meaning", as regards both information structure and cognitive aspects.

According to Armstrong (2008), overt subjects are used in contexts in which some "unexpected" contribution is made by the answer, cf. (44):

(44) a. A – Você$_1$ realmente acha   que o Lula vai   vencer?
              You   really   think.PRS.3SG that the Lula go.PRS.3SG win.INF
          "Do you$_1$ really think that Lula will win?"

      b. B – Eu acho.
              I   think.PRS.1SG
          "Yes, I$_1$ do."

(adapted from Armstrong 2008, 292, ex. 5')

---

[13]   Indeed *sim* "yes" and *não* "no" are also possible answers to direct questions, which can be considered a little more formal. Therefore, BP seems to have a mixed system of yes-no answers.

According to Armstrong (2008), unexpectedness is visible in this example by the use of the word *realmente* "really". Following her reasoning, Speaker B's evaluation of A's belief allows the overt subject vYNA as a marking of the "unexpectedness" of B's response. B then evaluated A's belief about the proposition and also A's expectation about his/her answer. The use of an overt subject marks the unexpectedness of the answer with respect to this evaluation. When there is no evidence about "unexpectedness", such as in a sentence without the adverb *realmente* "really", it would be an inference from the "shape" of the sentence (e. g. prosody).[14]

Thus, according to Armstrong (2008, 292), "some unexpected contribution must be made" in cases of overt-subject vYNAs. This is the second degree of her scale, which is preceded by the "neutral" context in terms of expectation and beliefs (vYNAs with null subjects) and followed by the "refutation" degree (verb plus post-verbal affirmative particle *sim* "yes"). According to her, "[w]hile the answer used to convey unexpectedness, the S[ubject]P[ersonal]P[ronoun]+S[imple]VE[rbal] response (*i. e.*, the overt-subject vYNAs) requires evidence and belief, the PVS [the post-verbal *sim* 'yes'] response requires strong evidence and belief, as it is used to refute a proposition".[15] To evaluate expectedness and beliefs, a speaker uses evidence to form his/her own beliefs, or to infer the beliefs of others. Such evidence may come from the discourse or exist independently of the discourse. However, in fact, *acho* "think.PRS.1SG" would not be inappropriate, as also noticed by Armstrong (2008). What is also interesting is the fact that the neutral "default" context triggers

---

[14]   Armstrong (2008) also points out some prosodic saliency in several examples. For some reason, no particular intonation is needed to pronounce the overt-subject vYNA in (44).

[15]   This scale itself, proposed by Armstrong (2008), could be disputed. Martins (2012) presents subjunctive questions to which null subject vYNAs are possible answers as are also null-subject vYNAs followed by an affirmative particle, but never vYNAs with an overt subject, as in (1) below. This would be impossible according to Armstrong (2008)'s scalar proposal, since there is no proposition being refuted in (1) below.

(1)   a. A – O João$_1$ talvez saia   do   hospital hoje?
               the João   maybe leave.SUBJ.3SG from.the hospital today
         "Will João$_1$ perhaps leave the hospital today?"

    b. B – _/*Ele sai   (sim).
             He   leave.PRS.3SG yes
         "Yes, he will leave the hospital today."(Possible continuation: "They've just called from the hospital.")

                          (adapted from Martins 2012, 8, ex. 26)

null-subject vYNAs. Why, if in other contexts the overt subject is supposedly a neutral option, would they be pragmatically neutral only in vYNAs? This question will be analyzed in section 2.3.4.

Finally, Martins (2012) presents a series of "disagreement" examples, that is, "refuting" vYNAs. These examples show that it does not matter whether the expectations regarding the answer are in the same rhetorical direction or not: even if the question and the answer are divergent, null-subject vYNAs are possible, such as in (45) below.

(45) a. i. A – Ele$_1$ hoje não vai   sair?
                  he   today not go.PRS.3SG leave.INF
           "He$_1$ is not going out today?"

          ii. B – _$_1$ Vai.
                    go.PRS.3SG
           "Yes, he$_1$ is."

  b. i. A – Ele$_1$ hoje não vai   sair,   pois   não?
                  he   today not go.PRS.3SG leave.INF, POIS-confirmative no
           "He is not going out today, is he?"

          ii. B – _$_1$ Vai.
                    go.PRS.3SG
           "Yes, he$_1$ is."

(adapted from Martins 2012, 4, ex. 12 and 14)

In both negative question (45a-i) and negative tag-question (45b-i), clearly A expects a confirmatory negative answer. B's disagreement answers in both (45a-ii) and (45b-ii) to these questions are vYNAs without an overt subject. These examples shows that Armstrong's proposal are not sufficient to explain the use of null and overt subjects with vYNAs. Besides the empirical problems, Armstrong (2008)'s proposal is really difficult to implement theoretically in a formal and decidable way: "(un)expectedness" is a difficult concept, which can hardly be encoded into a clear formalization or clear-cut notion. In subsection 2.3.4, a way formalize and account for her intuitions about "expectedness" is proposed.

2.3.3   Structural Analyses with Semantic-Pragmatic Consequences

Holmberg (2001) has a different account of the use of subjects in vYNAs. His proposal primarily addresses Finnish data. However, Finnish is similar to BP: the standard affirmative answers to Yes/No questions are repetitions of the bare finite verb of the question; common negative answers are negations followed by tensed verbs.

Moreover, in Holmberg (2007, 2013), he extends his proposal to deal with BP data. Holmberg (2001, 2007, 2013, 2014, 2016, to appear) suggests that null subject vYNAs are derived from full sentences by an especial mechanism: the movement of at least the verb or auxiliary (plus negation) and optionally other sentence constituents out of IP, which eventually undergoes ellipsis (*i. e.*, deletion under identity). He proposes then that IP-ellipsis occurs only in vYNAs; in this case, the antecedent of IP-ellipsis must be a polar question. According to him, the crucial property of these questions is the fact that they do not have a fixed polarity.

Holmberg (2001) claims that there are two conditions for IP-ellipsis to arise: (i) the sentence which undergoes the IP-ellipsis must have "polarity focus", Σ, and (ii) the antecedent must have an open polarity.[16] The elided IP would then be a presupposition and, being covert in a given sentence, it has to be copied from the previous clause antecedent.[17] Such an antecedent must be open regarding polarity, otherwise the assertion does not find a variable to apply to, and the sentence ends up violating "Full Interpretation".[18] According to Holmberg (2001), the only kinds of sentences which conform to this requirement are polar questions and polarity focus constructions.[19]

---

[16]   According to Holmberg (2001, 2007, 2013, 2014, 2016, to appear), "polarity focus" is an operator Σ which "takes two arguments: a clause with variable polarity, the presupposition, and a clause which picks out one of the two possible values that the polarity variable can have, the assertion". Actually, this definition of polarity focus makes the second condition completely dispensable. The second condition is maintained in the text for the sake of fidelity to Holmberg (2001)'s proposal and of easiness of the exposition.

[17]   Holmberg (2001) uses a concept of presupposition different from that adopted in the present dissertation. For him, "presupposition" is a fully structured sentence, which serves as an "antecedent" for another sentence. In the section 2.3.4, the notion of presupposition assumed in the present dissertation is explained, following Stalnaker (1978), among others. See also similar discussion in Schwenter (2005).

[18]   Holmberg (2001, 2007, 2013, 2014, 2016, to appear) follows Chomsky (1986, 98ff), who proposes that the principle of full interpretation requires that every element of PF and LF must receive an appropriate interpretation. None can be simply disregarded. In the case at issue, the variable would not receive an appropriate interpretation.

[19]   A polarity focus construction is a sentence whose focus is over the truth of the statement being made. Cf. Holmberg (2001), polarity focus constructions may be the antecedent of IP-ellipsis, including those which exhibit themselves IP-ellipsis.

As a consequence, the antecedent of IP-ellipsis must be one of such sentence types.

A further point in his proposal, which has been pointed out concerning Martins (2004)' VP-ellipsis proposal, is how the EPP is checked in null-subject vYNAs in Finnish and (by extension) in BP. The same question could be raised about Holmberg (2001, 2007, 2013, 2014, 2016, to appear)'s IP-ellipsis proposal. He proposes that a verb projection containing the subject and the verb and its complement(s), which he calls vP, moves to [spec, TopP], satisfying the EPP in vYNA.[20]

According to Holmberg (2001), the strongest evidence for his IP-ellipsis analysis is the fact that overt co-referential 3rd person subjects are never required in vYNAs, although they must be generally overt in other contexts. As mentioned in Chapter 1, in Finnish, 1st and 2nd person subjects are generally null; however, co-referential 3rd person subjects are generally overt in, for instance, independent clauses. This difference is neutralized in the context of vYNAs. In BP, the 3rd person singular has been resisting the general tendency to increase the use of overt subjects, as pointed out in Chapter 1. Consequently, there is also an imbalance in the relative frequency of null subjects across discourse persons in BP. And, indeed, as in Finnish, the difference among 1st and 2nd person, and 3rd person subjects also seems to be neutralized in the context of vYNAs in BP.

Holmberg (2001)'s second argument for his proposal is that in Finnish some adverbs cannot cooccur with the bare verb if the subject is dropped. He specifically points out that adverbial constituents that are internal to IP cannot appear in null subject vYNAs due to the fact that the whole IP is deleted. According to him, the use of adverbial expressions without the subject in vYNAs are represented by a structure in which only parts of constituents would have been elided. Sentences with an overt subject and adverbial expressions would then be generated by standard VP-ellipsis.

However, as pointed out by Martins (1994*a*,*b*, 2004, 2012) and mentioned in section 2.3.2, preverbal elements (either subjects or adverbial constituents) have a certain degree of structural autonomy in Portuguese and seem mainly pragmatically recovered, cf. (46) below.

---

[20]   According to Holmberg (2007, 2013, 2014, 2016, to appear), EPP is checked in TopP in Finnish. According to him, a vP can function as a topic, since as long as it includes the subject, vP is a referential expression, denoting an event or a state, and is therefore a possible topic. It becomes a predicate, a property-denoting expression, only as a result of movement of an argument (usually the subject) out of vP. A predicate is a lexical projection which contains an empty, A-bound argument position; this is an idea which goes back to Holmberg (2003)

(46) a. i. A – Ele$_1$ ontem   usou   gravata?
                  he   yesterday wear.PST.3SG tie
           "Did he$_1$ wear a tie yesterday?"

      ii. B – _$_1$ Usou.
                   wear.PST.3SG
           "Yes.(=Yes, he wore a tie yesterday)"

      b. i. A – Ele$_1$ mal   tocou   na   comida?
                  he   hardly touch.PST.3SG in.the food
           "Did he$_1$ hardly eat?"

      ii. B – _$_1$ *Tocou.
                   touch.PST.3SG
           "Yes." (=Yes, he hardly ate.)

<div align="right">(adapted from Martins 2012, 10, ex. 33 and 35)</div>

In (46a), a temporal adverb (*ontem* "yesterday") is recovered by the null subject vYNA. On the other hand, the null subject vYNA is not able to recover the adverb in (46b). Both the constituents would be considered VP-internal adverbs according to the theory proposed by Holmberg (2001). It could be argued that in this case the adverb *mal* "badly" does not mean "manner" and so it is not in its standard structural position; in (46b) it is indeed interpreted as an intensity adverbial element, such as *muito* "much" or *frequentemente* "frequently". However, these adverbs are also internal to VP, so that *mal* "badly" in the intensity interpretation should pattern exactly like other VP internal adverbs, such as *ontem* "yesterday".

Finally, Holmberg (2001) argues that IP-ellipsis has a restricted distribution. According to him, while VP-ellipsis is found in a variety of discourses, IP-ellipsis is limited to vYNAs and replies to polarity focus constructions. Among the contexts discussed by Holmberg (2001), the opposition between the polarity focus construction in (47) below and the "standard VP-ellipsis context" in (48) below is noteworthy.

(47) a. Ei Matti$_1$ Pariisissa ole   käynyt.
          not Matti   to.Paris   have.PRS.3SG be.PART
          "Matti$_1$ has NOT been to Paris."

      b. B1 – _$_1$ On   käynyt.
                     have.PRS.3SG be.PART
          "Yes, he$_1$ has."

c. B2 – _1 Ei ole.
        not have.PRS.3SG

"No, he1 hasn't."

d. B3 – _1 On.
        Ø   have.PRS.3SG

"Yes."

(adapted from Holmberg 2001, 160, ex. 44)

(48) a. A – Matti1 ei ole käynyt Pariisissa.

"Matti1 hasn't been to Paris."

b. B – On   *_1/ se1.
       have.PRS.3SG   he

"Yes, he1 has."

(adapted from Holmberg 2001, 157, ex. 36)

According to Holmberg (2001), since the focus in (47) is the negative polarity (marked by an especial word order in Finnish), this leads to an antecedent presupposition with an open variable. This presupposition is available to be copied by the elided IP in B's possible answers in (47b)-(47d). On the other hand, in (48b), the null subject vYNA is not acceptable, but the same sentence with an overt subject is. According to him, because the focus of (48a) is not the polarity itself, it does not provide an appropriate antecedent for IP-ellipsis, given condition (ii) above. Holmberg (2001) argues that the vYNA with an overt subject is generated by standard VP-ellipsis. The antecedent in this case need not have an open variable. A's sentence in (48a) then provides a good antecedent for VP-ellipsis.

In a cross-linguistic study, Holmberg (2007) provides another context in which VP-ellipsis and IP-ellipsis leads to opposite judgments of acceptability. In (49), null subject post-comma sentence is claimed to be produced by IP-ellipsis. The same construction with an overt subject is acceptable since it can result from VP-ellipsis, according to his proposal.

(49) a. Dizem   que o João₁ não fala   francês, mas *_₁/ele₁ fala.
         Say-p-3p that the João not speak-p-3s French, but Ø   he   speak-p-3s
         "They[indet] say that John₁ doesn't speak French, but he₁ does."

(adapted from Holmberg 2007, 215, ex. 4d)

A remaining question is raised by Holmberg (2016). Why, in a language that has pro-drop, for instance (Br)Portuguese, is IP-ellipsis necessary to produce null subjects in vYNAs, as in (45)? Answering this question, Holmberg (2016) proposes a test to verify whether a (partial) pro-drop language has IP-ellipsis. In his book, Holmberg (2016) compares his own proposal, which he calls the IP-ellipsis analysis, to VP-ellipsis plus null subject analyses (Martins 1994*a*, e.g.). According to him, his test is a straightforward way to show that, in a subset of languages, verb answers have the structure based on IP-ellipsis, but in other languages they conform to a pro-drop analysis. According to him, as mentioned before, 3rd person null subjects are highly restricted in standard sentences in Finnish, as in (50b), despite the fact that they could be produced by VP-ellipsis plus null subject. His argument against the pro-drop analysis, at least for Finnish, is that vYNAs are much more free to drop the subject than is the case in other contexts (Kato 2013 points out the same for BP). These patterns are very similar to BP.

(50) a. i. A – Tulee-   ko   Liisa₁ pian?
               Come.PRS.3SG [question.marker] Liisa   soon
            "Is Liisa₁ coming soon?"

       ii. B – _₁ Tulee.
               come.PRS.3SG
           "Yes."

    b. A – Liisa₁ ei   ole kotona. *_₁/Hän₁ tulee   pian
           Liisa   not is home   _₁/she₁   come.PRS.3SG soon
       "Liisa₁ isn't at home. She will get here soon."

(adapted from Holmberg 2013, 35-36, ex. 8-9)

Given this contrast in Finnish, Holmberg (2016)'s assumption is the following: as this test presents a pattern in Finnish and in Welsh (which he argues should be analyzed as IP-ellipsis languages), this pattern should be a diagnosis for IP-ellipsis in other languages whose nature is doubtful. Naturally, the analysis of vYNA in Finnish and Welsh could be *sub judice* as well.

However, let us assume that they are cases in which IP-ellipsis takes place for the sake of argument. His proposal is to distinguish between the IP-ellipsis and VP-ellipsis-plus-pro-drop derivations by using an existential indefinite subject in the antecedent question, because in Finnish indefinite referents can be dropped in vYNAs. According to him, if an answer to a question can be a bare verb or Yes+Verb, the derivation cannot involve pro-drop, since (a) the subject in the answer is identical to the subject in the question and (b) an existential indefinite pronoun cannot undergo pro-drop, but must rather be spelled out (at least in independent clauses in standard pro-drop languages).[21] In fact, it holds for Finnish, Welsh and (Br)Portuguese, as shown in (51) (see also Martins 1994a, for further arguments against the pro-drop analysis of a similar case in EP).

(51)   a. A – Alguém$_1$ / Ninguém$_1$ trouxe   açucar$_2$?
             Someone anyone   bring.PST.3SG sugar

        "Did someone$_1$/anyone$_1$ bring sugar$_2$?"

    b. B – $_1$ Trouxe $_2$. ($_2$ Está   aí.)
             bring.PST.3SG   (be.PST.3SG there)

        "Yes. It$_2$'s over there."

<div align="right">(adapted from Holmberg 2013, 48, ex. 41)</div>

Lastly, Holmberg (2013, to appear) uses his proposal to account for answers to negative polar questions in various languages. In these papers, he provides a way to deal with "expectations" formally. That is, Holmberg (2013, to appear) attempt to explain the phenomena that Armstrong (2008) takes to be pragmatic by means of a structural mechanism. According to him, as opposed to polar questions without negation, negative polar questions convey a presupposition concerning the

---

[21]   Holmberg (2016) argues that indefinites in pro-drop languages cannot be expressed by null subjects, as in example (1) below.

(1)   _ Può   controllare questo macchinario con una mano sola.
          can.PRS.3SG control.INF this   machine   with one hand only
        "He/she/*someone can control this machine with one hand."

<div align="right">(adapted from Holmberg 2016, 32, ex. 50)</div>

    Holmberg (2016) argues that "someone" is not a possible interpretation for the null subject in (1). It is doubtful for me, however, that no indefinite antecedent could be retrieved by the null subject in (1), given Carminati (2002)'s results about the "bound" variable interpretation in Italian. Further research on independent clauses in Italian should clarify this question.

truth of the proposition (*viz.*, that it is false), and answers to them depend on the expected answer derived from this presupposition. He thus proposes that the two possible answers in (52) below are derived by different mechanisms. The bare affirmative particle is due to IP-ellipsis; on the other hand, the standard "tag" answer is to be generated by VP-ellipsis.

(52) a. A – Isn't Mary$_1$ coming?

    b. i. B1 – ?Yes.

        ii. B2 – Yes, she$_1$ is.

<div align="right">(adapted from Holmberg 2013, 48, ex. 41)</div>

In response (52b-i), the antecedent copied by the covert IP is the presupposition of the negative question, which conveys A's expectation. It is infelicitous because the antecedent presupposition provides a negative polar feature to the elided IP, which clashes with the positive feature of the affirmative particle *yes*. The "tag" answer in (52b-ii), however, is felicitous because it is not produced by IP-ellipsis, but rather VP-ellipsis. According to Holmberg (2013, to appear), as the elided VP does not copy the polarity from the question, since only the VP constituent is copied, no feature clash takes place.

Since Holmberg (2001, 2007, 2013) bases his analyses for the use of subjects in vYNAs upon the consequences of VP- vs. IP-ellipsis, the proposal developed in Holmberg (2014, to appear) predicts that null subjects in positive vYNAs in BP should not be felicitous answers to negative polar questions, but that overt-subject vYNAs should be. That is, null subjects in vYNAs should be used to convey the expected answer to negative polar questions (confirmatory negative) by IP-ellipis (since polarity is included in the ellipsis), and vYNA with overt subjects should be used to express an unexpected answer (disconfirmatory positive) by VP-ellipsis (since polarity is not included in the elliptical material). However, this prediction is not correct for BP. In (45) in section 2.3.2 for example, repeated below as (53), Martins (2012)'s examples show that negative polar question, as well as negative "tag" questions, can be answered positively using vYNAs with null subjects.

(53) a. i. A – Ele$_1$ hoje não vai   sair?
              he   today not go.PRS.3SG go.out.INF
          "He$_1$ is not going out today?"

ii. B – _1 Vai.
        go.PRS.3SG

"Yes, he$_1$ is."

b. i. A – Ele$_1$ hoje não vai   sair,   pois   não?
      he   today not go.PRS.3SG go.out.inf, POIS.confirmative no

"He is not going out today, is he?"

ii. B – _1 Vai.
        go.PRS.3SG

"Yes, he$_1$ is."

(adapted from Martins 2012, 4, ex. 12 and 14)

Kato (2013) adapts ideas from Holmberg (2001, 2007, 2013, 2014, to appear), proposing a structural syntactic account with pragmatic consequences (in a "Cartographic Syntactic Structure" *à la* Rizzi 1997) to account for BP data. In her approach to vYNAs, the inflected verb in Pol (that is, "Polarity") moves to F ("Focus") in the left periphery, and the Remnant IP moves to GroundP, where it is interpreted as the sentence presupposition, and then deleted at PF. For her, this analysis explains why BP, a variety that is in the process of losing null subjects, can exhibit an apparent context where the null variant seems to be categorical.

Kato's 2013 proposal cannot, though, explain data pointed out by Armstrong (2008), as for example (44), repeated here as (54) below, since despite being an unexpected answer the overt subject is part of the background provided by the question.

(54) a. A – Você$_1$ realmente acha   que o Lula vai   vencer?
      You   really   think.PRS.3SG that the Lula go.PRS.3SG win.INF

"Do you$_1$ really think that Lula will win?"

b. B – Eu acho.
      I   think.PRS.1SG

"Yes, I$_1$ do."

(adapted from Armstrong 2008, 292, ex. 5')

2.3.4 The Structure of the Discourse

The phenomenon of the use of subjects in verbal answers is one case that shows the importance of the discourse-grammar interface. From the point of view developed in the present dissertation, it

is impossible to fully account for the constraints on null and overt subjects in this context without understanding the relation between the syntactic structure of these constructions and the pragmatic restrictions imposed on them. This context does not fit in the general rules of the use of null subjects as previously discussed, preferring the null form to the pronominal one. However, they are sentential answers, and consequently have internal grammatical structure (agreement marking, verbal tense semantics, for instance). Assuming by hypothesis that the relation between turns is somehow governed by pragmatic dialogue rules, there must clearly also be a pragmatic element governing the choices between overt and null subjects in the context of dialogue pairs, specifically in verbal answers. The main questions to be answered are thus the following: is this pragmatic element fully formally describable? How does is it interact with the syntactic level? Can the interactive structure of the discourse account for the use of null and overt subjects in BP (supposing that it is an obligatory level of the linguistic analysis, since languages are intended to be meaningful and interactive)?[22]

Information and Discourse Structure Roots

In the past five decades, the developments in formal pragmatics have brought about a theory that connects the organization of the discourse to the formal representation of the sentence at the phono-logical, morphological and syntactic levels (see Vallduví 1990, Lambrecht 1994, Vallduví & Engdahl 1996, among others, for overviews). According to this theory, the information import of sentences constrains the way that they are formally represented. The form of the utterances is used to guide the speakers in the contextual updates that they must carry out in order to build a coherent discourse (see Chafe 1976, Clark & Haviland 1977, Prince 1986, Lambrecht 1994). Following this theory, utterances that were previously taken to be semantically and pragmatically quite similar, despite formal differences, are not variations that speakers can use in the same discourse contexts. Rather they are different ways of expressing the same main meaning (the propositional content)

---

[22] Surely, the interactive structure of the discourse cannot account for the use of null and overt subjects in BP in and of itself. Otherwise, supposing that this structure is basically the same for all languages (see Grice 1975, Carlson 1983, Roberts 1996, Ginzburg 1996, Roberts 2012, Ginzburg 2012, among others), all of them would have the same options as BP. So, some sort of specific interface constraints must be proposed, as put forth in what follows.

in different *delivery offers* (from this idea, the expression "information package" is derived; see Chafe 1976 for the first formulation of this metaphor). The discourse context indicates which delivery option is the most or least appropriate for each part of the conversation. Taking the discourse requirements into consideration, speakers follow such communicative instructions when "packaging" the information they want to share. By violating these indications, a speaker produces less appropriate or infelicitous utterances in a given context or triggers implicatures.

Many different categories are assumed in the literature about information structure. Vallduví & Engdahl (1996) identify two informational articulations among the several proposals put forth in the literature: (i) those who divide the sentence into "ground" and "focus" (Halliday 1967, Jackendoff 1972, Halliday & Hasan 1976, Clark & Haviland 1977, Wilson & Sperber 1979, Prince 1986, Lambrecht 1994, Rochemont 1986, Rooth 1985, *inter alia*) and (ii) those which split the sentence into Topic and Comment (Givón 1983, Reinhart l981, Gundel 1988, Gundel et al. 1993, Gundel 2010, among others). There are actually those who also assume both the articulations (or a tripartite articulation), such as Vallduví (1990), Vallduví & Engdahl (1996), Sæbø (1997), Büring (2003), e. g. However, even when the assumed articulation is the same across proposals, the terminology, the concepts, and the empirical coverage of the relevant notions are absolutely not uncontroversial. Although the concepts used in the Information Structure literature do not have a theoretical status in the theory developed here, some of them will be used in the sketch of certain definitions and must consequently be clarified. Such concepts will be defined in next subsection in line with the approach assumed in the present dissertation.

The Game Theory of Discourse Structure – Question under Discussion and Dialogue Gameboard Approaches

In this dissertation, a slightly different approach is assumed. By the middle of the 1990s, two approaches to information structure were independently developed, although they share many points: Roberts (1996), republished as Roberts (2012), proposes the Question under Discussion [QUD] approach, while Ginzburg (1996) proposes a dynamic semantics for dialogue (Dialogue Gameboard [DGB] approach). Roberts (1996), based on Kadmon & Roberts (1986), points out that most of the information structure literature does not look beyond the sentence level to examine in any detail the contexts in which a certain utterance is indeed felicitous. As a consequence, Roberts (1996) devel-

ops a theory of information structure in which information structures the discourse and vice-versa ("on the inquiry pursued in discourse and the information which that inquiry yields"). She builds the notion of the information structure of a discourse by extending Stalnaker (1978)'s possible worlds idea of context. In her approach, questions and assertions are both taken into account and they impose a pragmatically-motivated structure on well-formed dialogues. Ginzburg (1996) also takes Stalnaker (1978)'s and Lewis (1979)'s idea that building a context is updating commonly accepted information to a shared set of possible worlds. His main point is to stress the structured view of context, according to which there is a set of questions currently under discussion (apart from other contextual information). The goal of the present dissertation is not to present and compare both approaches extensively, but rather to use them as ways to explain the structure of the discourse and account for the use of null and overt subjects straightforwardly without adding unexplained devices or *ad hoc* diacritics to the theory of language.

In both Roberts (1996) and Ginzburg (1996), the linguistic exchanges between speakers are viewed as a game, in which there are goals (efficient communication with less effort), moves (utterances), rules (the shared constraints to achieve good results) and strategies (evaluations of the best ways to achieve the goals). Following Stalnaker (1978), the common ground is taken to be the set of propositions which the interlocutors in a certain discourse share and believe to be true. A proposition in this theory is technically a set of possible worlds. The "Related Context Set", that is, the intersection of the common ground, is the set of worlds where all the propositions in the common ground are true. One of the goals of efficient communication is to reduce the context set to a singleton set, the actual world. The language rules are, according to Roberts (1996), linguistic conventions (syntactic, morphological, etc.) and conversational rules (such as Gricean Maxims). Conversational rules are general rational properties as regards the goal (The Cooperative Principle is a result of the common wish to communicate, the Maxim of Quality comes from the fact that the truth is the final goal, Quantity 1 from the attempt to maximize the payoff of a move bearing in mind the goal, cf. Grice 1975, *inter alia*, and from human cognitive limitations; see Sperber & Wilson 1986, Hawkins 2014, Gibson 1998, etc). Following Carlson (1983), Roberts (1996) proposes two types of moves: setup moves (questions) and payoff moves (assertions). Questions (different from the "imperative" speech acts proposed by Lewis 1970) are viewed as a set of alternative propositions, as proposed originally by Hamblin (1973) (similar to Alternative Semantics,

see Rooth 1985, 1992, Kratzer 1991, von Stechow 1991, Groenendijk & Stokhof 1984). Accepting a given question makes it the current, or immediate, question under discussion. Following Stalnaker (1978), Assertions provide the descriptive content that is added to the common shared ground assuming that the other participants do not object to it (see also Ginzburg 1996, 3). Roberts (1996, 6:6) highlights that, for Stalnaker (1978), assertions are choices among alternatives and, when accepted, they shrink the context set. For her, relevant answers in coherent question-answer pairs are choices among the alternatives proffered by the QUD. According to Roberts (1996)'s proposal, there is one Big Question – "What is the way things are?" – and all possible questions are subquestions, somehow entailed by this question.

Ginzburg (1996) goes a step further in the theory of dialogue, showing that the accepting-rejecting dichotomy cannot be maintained when dealing with actual dialogues. In Ginzburg (2012, to appear), the "possible moves" in his theory are fully developed. For present purposes, as a general overview of the DGB theory, it is enough to bear in mind that it is possible for a speaker not to directly address a given current QUD: the speaker can deliberately partially accept or discuss such a QUD. The QUD approach was also further developed by, *e. g.*, Büring (2003), Krifka (2007, 2008), Simons et al. (2010). The possibility of not addressing the current QUD is also incorporated in the QUD approach by the analysis of Not-at-Issue content [NaI] (see Potts 2015, Simons et al. 2010, Destrual et al. 2013, among others). The definition of at-Issue in the present dissertation is given later in this chapter.

Some important concepts must be defined in the context of these theories. As in Construction Grammar (Lambrecht 1994), Alternative Semantics (Rooth 1985, 1992, Kratzer 1991, von Stechow 1991), QUD approaches (Roberts 1996, 2012, Simons et al. 2010) and DGB approaches (Ginzburg 1996, 2012, to appear), "Focus" here is taken to be a semantic-pragmatic notion, which corresponds to the open part of a unary question and the asserted content in an Assertion. In the present theory, formally it corresponds to an open variable in the question, and to saturated variable in the answer. All the content that is in a current QUD and is not part of the focus is considered "Backgrounded" (see Stalnaker 1978, Lambrecht 1994, Rooth 1985, 1992, Kratzer 1991, von Stechow 1991, Roberts 1996, 2012, among others for similar ideas). Topic will be occasionally referred to as the entity that is part of the background and that the sentence is about (see Reinhart 1983, *inter alia*), but it has no formal correspondent in the present theory. The notion of Topic

and Topicality will be further discussed in section 2.6. Apart from language specific properties, such as morphological topic marking or syntactic topic position, "Topic" is usually stressed as an important concept in multiply opened questions, which are expressed by double stressed sentences (also know as Contrastive Topics; see Büring 1999, 2003). This case will also be addressed later in this chapter. For present purposes, the aim is to study the background-focus articulation and how it might possibly be a factor that is relevant for the use of null and overt subjects in BP.

Maximal Contrast - Maximal Contribution with Less Effort

Vallduví & Engdahl (1996), based on Vallduví (1990), make an observation about the interface between information structure and syntax. According to them, a sentence like *He LOVES it.* can be analyzed as an All-Focus sentence lacking a ground. Following their reasoning, the information structure of (55a-ii) and (55b-ii) is quite similar: both examples are answers to unary Wh-questions, such as (55a-i) and (55b-i).[23]

(55) a. i. A – How does John$_1$ feel about beer$_2$?

ii. B – [$_F$ He$_1$ LOVES it$_2$.]

b. i. A – What drink does John love?

ii. B – [$_F$ BEER.]

(Vallduví & Engdahl 1996, 15, ex. 31)

According to Vallduví & Engdahl (1996, 14), citing Vallduví (1990), both answers in (55) are minimal, since they are as elliptical as possible in the context of the respective questions. The presence of "weak proforms" in (55a-ii) is due to independent requirements of English grammar, in which null arguments are not allowed. They report that in Catalan, for instance, the canonical object slot would be empty.[24] According to them, the proforms in English are thus only grammatical place-holders.

---

[23] Following the traditional notation used in the literature on Information Structure, fully-capitalized words indicate the sentential focal prosodic stress (see Selkirk 1984, Rochemont 1986, Truckenbrodt 1995, Lambrecht 1994, Vallduví & Engdahl 1996, among others).

[24] Unfortunately, Vallduví & Engdahl (1996) do not provide an example, so I am not sure the empty slot is not linked to a clitic attached to the verb (as in Spanish).

It is clear in the above examples that the two sentences have similar information structures, given that both examples are proper answers to the unary wh-questions in (55a-i) and (55b-i), provided by Valldu ví & Engdahl (1996). Also, in BP, where both null subjects and objects are possible under certain contextual circumstances, a bare verb *Ama.* "love.PRS.3SG" is a possible answer in context (55a-i). However, a caveat is needed in this analysis: while example (55b-ii) is unequivocally all-focus, the all-focus status of (55a-ii) is disputable. Compare, for instance, the context in (55) above with that in (56) below.

(56)   [Context: Adam has murdered a woman (referred to by "She$_2$" below), who is *not* his wife.]

   a.   A – Look, I love my wife$_1$. I love my wife$_1$ very much. I ain't out for sexual toys.

   b.   B – What happened next, Adam?

   c.   A – She$_2$ bled to death. (...)

(COCA corpus, "SPOK: CBS 48 HOURS 10:00 PM EST")

Differently from the context in (55), in (56) the QUD is *And what happened next?*; thus answer (56c) is an actual all-focus sentence. Observe that, neither in question (56b) nor in the preceding turn in (56a), "She$_2$" is the topic. So, the proform in (56c) is really part of the asserted content and not in the background.[25] In this case, "She$_2$" is not only a grammatical placeholder. In the light of this, what is the information structure of the answers in (55) and how can they be analyzed? In the present dissertation, the answers in (55) are considered Maximally Contrastive, rather than all-focused sentences.[26]

---

[25]   Even if the antecedent of "She$_2$" was mentioned earlier in the conversation, it does not make its referent part of the background in the sense proposed in the present dissertation. While the opposition between "new" and "old" information would predict so, the opposition between "asserted" and "backgrounded" content (in the current QUD) does not imply that any proform is part of the current "background" as defined in 57 below.

[26]   Just as sentences with multiple focus create a problem for some of the (back)ground-focus-based theories, example (55a-ii) poses a problem for topic-comment-based approaches. If this sentence is not an all-focus sentence, as shown by the contrast with the example (56c), "it" has either no information status or the definition of topic ("entity which the predicate is about", following Reinhart 1983) must be changed, since both pronouns would be considered the topic of the sentence. Multiple solutions to this problem have been proposed in the literature. In the present dissertation, they are not explored, since the (back)ground-focus articulation is the primary interest.

Maximally Contrastive answers are sentences in which the backgrounded and the focused information contrast the most. Contrasting the most here is defined in the following terms: there is a backgrounded set of propositions, in which each single alternative in the set of propositions proffered by the *current* QUD contains the same backgrounded content, and the focused information directly addresses the current QUD. Assuming that every discourse is anchored in a background that includes at least the Big Question, at first sight, any all-focus sentence would be Maximally Contrastive; however, as the contrast between (55) and (56) shows, this claim does not hold: the set of possible propositions included in Big Questions, such as (56b), is formed by sentences which do not share the same background information, since either they do not have any backgrounded information or the backgrounded information is not part of the *current* QUD. In (55a-i), all possible propositions include the backgrounded entities *John* and *beer* and possible relations between these entities are in the set of alternative propositions : "John loves beer.", "John hates beer.", "John likes beer.", etc. The alternative propositions in the set proffered by the big question in (56b) (the *current* QUD) are completely unrelated to each other, and these propositions possibly do not share any common background: "The woman bled to death.", "Nothing.", "My dog killed a bird.", etc. Notice that the presence of a pronoun in English does insure that it is part of the background of the current QUD. The reference of "she$_2$" is somehow topical in the conversation in (56) and it may even be considered implied in question (56b), but it is not part of the *current* QUD at the point at which sentence (56c) is uttered.

In formal terms, Maximal Contrast is defined as follows (following the formal definition of "Focus" given in Stalnaker 1978, Rooth 1985, 1992, Roberts 1996; "QUD" in Roberts 1996, Ginzburg 1996, 2012; "Shared Context" in Stalnaker 1978, Roberts 1996, Ginzburg 1996, 2012; "Proposition" in Lyons 1977, Stalnaker 1978, Roberts 1996; and "Utterance" in Lyons 1977, Ginzburg 1996, 2012):

(57)   Maximal Contrast

Given a proposition α in a context Γ, such that Γ contains a set of shared accepted propositions *P* at a given discursive time-point *t*, the utterance of α is Maximally Contrastive iff:

a. α is asserted with a focused semantic part *x*, so that α_x is its discursive-semantic representation;

b. *y* is open in the *current* QUD q_y, such that it produces a set of alternative assertive propositions Q_y [$\alpha_x \subset Q_y \wedge Q_y \not\subset P$ at *t*];

c. $\forall \alpha_x \forall \alpha_z \subset Q_y \exists q_y[\text{y} = \text{x} \vee \text{z}]$;

d. $\forall \alpha_x \forall \beta_z \forall \beta_x \subset Q_y \, [\alpha_x = \beta_x \neq \beta_z]$.

In (57), Maximal Contrast is defined in discursive-semantic terms as a relation between focus of an assertion, the context and the QUD to which this assertion can correspond. Following Stalnaker (1978), Roberts (1996) and Ginzburg (1996, 2012), the definition starts by setting the context in which a structured set of accepted propositions (the common ground) occurs. A given propositional content α is a Maximally Contrastive utterance in a given discourse if and only if it fits in the four definitions below. (a) By (57a), there is a semantic focused part *x* in the assertion of α, such that the assertion is represented by α_x (that is, proposition α becomes an assertion α_x because of the existence of a focused element *x* within the propositional content).[27] At this point, it is not necessary to enter into details about how foci are obtained compositionally (see Rooth 1985, 1992, Kratzer 1991, von Stechow 1991, among other in this regard). The assertion symbols α_x stand for atomic entities, whose focus diacritic is defined independently (by standard compositional semantics). (b) By (57b), the current QUD q_y has a non-resolved semantic part that produces a set of alternative assertions Q_y, such that the assertion α_x is contained by it, but the set of alternative assertions Q_y is not contained in the current set of shared propositions (propositions are assumed to be entailed by assertions by definition, since all possible focus over a propositional content should be part of such a content).[28] (c) By (57c), focus is formally defined and restricted: for any assertion α_x and α_z contained in the set of alternative assertive propositions Q_y, there exists a question q_y such that the open value of *y* is either *x* or *z* or both (following the logical value of the disjunction). Finally, (d) by (57d), for any alternative assertive propositions α_x, β_x and β_z contained in the set Q_y, they differ only in the focused content.

Maximal Contrast in (57) captures many general principles observed cross-linguistically. First,

---

[27] Focus here is taken to be an exclusively discursive-semantic notion. The actual realization of focus is a matter of much debate in the literature and varies cross-linguistically (see Büring 2009, for an overview).

[28] The assumption that propositions contain all possible assertions is not completely true when it comes to, for example, focus sensitive particles and quantifiers. This discussion is far beyond the purposes of the present dissertation.

it is an attempt to formalize the Gricean Maxims of Quantity, by defining an optimal contribution to the context. In this vein, it is worth mentioning that Maximal Contrast stands as a definition of a level of informativeness (possibly the highest), but does not explain the whole scale. Lower levels in the scale are discussed further in this chapter. Second, it brings to the discursive-semantic level an idea that is fundamental in other levels of linguistic analysis: languages are organized in terms of contrast. In phonology for example, contrast is a central notion for explaining many phenomena and cross-linguistic variation. Last, Maximal Contrast formalizes an intuition about the nature of the human mind and communication: they are based on *contrast*. Several studies have shown that, in any language, one of the main sources productivity is contrast (Saussure 1983, Givón 1985, 1991, among many others). Exploring the possible relations between contrast and the theory of language goes further than the purpose of this section. Finally, it is worth mentioning that other researchers, including especially Ginzburg (1996, 2012, to appear), have also identified a similar notion, which he calls Fact-A_{Sit_0} in earlier papers, and ultimately MaxQUD. Maximal Contrast also shares some points with the idea of "Complete Answer" in Roberts (1996), but does not completely overlap (see also Destrual et al. 2013 for a similar idea in the QUD framework). Different formalizations are proposed in previous research (in Ginzburg 1996, 2012, to appear in Situation Semantics combined with a HPSG matrix representation and in Roberts 1996 in type-logics combined with traditional compositional semantic-syntatic rules in montagovian fashion). However, the notational variants are not important here, since the main idea is captured by the formal representation in (57). In what follows, Maximal Contrast is used to explain the use of null and overt subjects in verbal answers in BP.

Maximally Contrastive Verbal Utterances in BP

Verbs in (European or Brazilian) Portuguese bear at least three blocks of semantic information: (i) the event, that is, the description of what happened to the entities involved in the situation (event here is used in Davidsonian sense, that is, either a state or a dynamic event, which is probably bound by an existential quantifier)[29], (ii) the tense (including clues about aspectual information), how the event is located in time with respect to the time of the utterance, and (iii) the assertion

---

[29]   See Davidson (2001) for a compilation of his ideas.

of the (relative) truth of the proposition. For the sake of analogy, a verb in Portuguese conflates semantic information of the main verb and an auxiliary in English (in Portuguese, there are also some auxiliaries, but they are more scarce than in English; when an auxiliary is used, the tasks are split).

Following Vallduví & Engdahl (1996)'s idea, in a context in which both the subject and the direct object are backgrounded, a sentence like "He LOVES it." in English could be a bare verb in a language in which these arguments are not required for grammatical reasons. In the previous section, this sentence was analyzed as an example of Maximal Contrastive Utterance (or MaxQUD in Ginzburg's terms). Ginzburg (1996, 2012, to appear) claims that both answers to yes-no questions and unary wh-questions are MaxQUD answers. BP is the perfect sample to test both these claims, since BP allows both null subjects and objects in some contexts and BP verbs can be used to answer Yes-No questions. In this section and sections 2.3.5 and 2.4, only vYNAs are investigated. In the experimental section 2.5, both contexts are studied.

To use the definition of Maximal Contrast to predict the bare verbs in vYNAs and in unary Wh-questions, as mentioned before, some interface constraint must be formulated. There are many ways of proposing such a mapping constraint. Here it is formulated in the most conservative way, taking null arguments to be simply "allowed" and overt to be "required" in their proper contexts. Using other terms in the formulation of the hypothesis below could indicate affiliation to one or another syntactic (or interface) theory ("elide" from Merchant 2001's proposal, "deletion" from Barbosa 2011's approach, etc). Although the hypothesis below has some similarities with other approaches to missing arguments, it does not completely overlap with any of them, so that general terms are preferable. The mapping constraint that relates null arguments to Maximal Constrast is as follows:

(58)   Maximal Contrast Null Arguments Hypothesis [MaxCoNAH] (informal first version)
       Null arguments are allowed when they are backgrounded in an utterance that is Maximally
       Contrastive. Otherwise, overt arguments are required.

To properly understand how the MaxCoNAH and Maximal Contrast account for verbal answers in BP, an example is worked out below. In (59), a bare verb is used in (59b) to direct address the current QUD (59a). The following sentence in answer (59b) goes on addressing the same QUD,

but it fails to be Maximally Contrastive, since an assertion with the adjunct "at the same time" is not part of the set of assertive propositions produced by question (59a).

(59) a. A – mas professor$_1$ é   funcionário público?
        but   teacher   be.PRS.3SG servant   public
"But are teachers$_1$ civil servants?"

b. B – _$_1$ é   ... _$_1$ é   ... mas ele$_1$ ao   mesmo tempo ... ele$_1$ não
        be.PRS.3SG   be.PRS.3SG   but he$_1$ at.the same   time   he$_1$ not
pode   ser   funcionário se ele$_1$ se   sindicalizar   ...
can.PRS.3SG be.INF servant   if   he$_1$ SE unionize.SUBJ.FUT.3SG
"Yes... they$_1$ are... but they$_1$ at the same time they$_1$ cannot be a civil servant if they$_1$ unionize..."

(NURC-RJ, "Inquiry_164")

To make things more explicit, the discourse structure is decomposed in (60) below in the "cleaned" version of (59) (discourse markers, such as pauses and repetitions, are omitted):

(60) a. A – [$_q$ mas professor$_1$ [$_y$ é] funcionário público] [$y = 1 \lor 0$]

"But are teachers$_1$ civil servants?"

b. B – [$_\alpha$ _$_1$ [$_x$ é]] [$x = 1$]

"Yes, they$_1$ are"

c. B' – [$_\alpha$ _$_1$ [$_x$ não é]] [$x = 0$]

"No, they$_1$ aren't"

d. B – mas [$_\alpha$ ao mesmo tempo ele$_1$ [$_x$ não pode ser funcionário se ele$_1$ se sindicalizar]] [x = cannot be civil servant if...]

but at the same time they$_1$ cannot be a civil servant if they$_1$ unionizes..."

In (60), one example was added to the original from to corpus to make explicit both the possible Maximal Contrastive answers. In a nutshell: question (60a) produces a set of alternatives Q_y that contains the proposition "teachers$_1$ are civil servants" and "teachers$_1$ aren't civil servants" (the

focal possible values are 1 – true or 0 – false);[30] both alternatives are Maximally Contrastive, since they differ only in their focal value *x*. Given semantic-grammatical properties of verbs in Portuguese, they can encode the meaning of the focused parts of the assertive propositions in set Q_y. By the MaxCoNAH, the subject argument *professor*$_1$ "teachers" can be null, given that it is in Maximally Contrastive utterances in (60b) and (60c). (60d), however, does not fit in the definition of a Maximal Contrastive Utterance, because it does not differ from the other alternative answer only in its focal value. According to the MaxCoNAH, an overt pronoun is thus required.

Based on the MaxCoNAH as a possible account for the use of null and overt subjects in BP, the study of verbal utterances is carried out in the next two chapters. In section 2.4, context-type examples from NURC-RJ corpus are analyzed and the proposals to account for them are discussed. This first qualitative analysis provides evidence for the approach based on Maximal Contrast and sets up the scenario for Experiment 4. However, other approaches could possibly make similar predictions regarding the examples from the corpus. Based thus on the corpus study, in section 2.5, striking evidence in favor of the Maximal Contrast approach is provided: Experiment 4 is designed to test the Maximal Contrast approach and other approaches that make similar predictions regarding the corpus examples (especially the IP-ellipsis approach). Before going over the empirical study, a summary of the accounts for vYNAs is presented in section 2.3.5 below.

2.3.5 Summary

In this section, the debate about vYNAs in BP and the previous proposals to deal with the use and the omission of subjects in these constructions were introduced. Although the early proposals from 1983 to 1993 do not directly aim to explain null subjects in vYNAs, they observe that null subject vYNAs are the standard way to respond to polar questions in BP (cf. Kato & Tarallo 1993 and Urbano et al. 1993). Moreover, they initiate the dispute over the level of analysis that should be considered when dealing with the relation between turns and recovered referents. Against Stubbs (1983), Kato & Tarallo (1993) proposes that the analysis must be done at a syntactic level when treating vYNAs. Following Stubbs (1983), Urbano et al. (1993) argues for a pragmatic treatment.

---

[30] This is a simplification, since evidently the truth of the statements needs to be modalized. That is, in a full account, it should be relative to possible worlds.

The literature specifically discussing the use of null and overt subjects in vYNAs is very limited. So far, the initial debate leads to two possible sets of analyses of the phenomenon. The first set presented in this dissertation assumes that the relation between sentences in a dialogue is mainly pragmatic and that the conditions on the null and overt subjects are as well. As shown above, these analyses are not completely explicit and assume some flexibility concerning their criteria. They can be summarized as two main ideas: (i) according to Martins (1994*a*, 2012), null subjects must be unmarked topics in a pragmatic sense; and (ii) Armstrong (2008) points out that the use of overt subjects contributes to make a vYNA in some degree "unexpected".

The second set of analyses proposes that the relation between polar questions and vYNAs is structural and syntactically represented and that pragmatic effects are by-products of their relation. As pointed out by Holmberg (2016), two analyses to account for null subjects in vYNAs are possible. (i) Holmberg (2001, 2007, 2013, to appear, 2016) proposes an especial mechanism to deal with such constructions, namely IP-ellipsis. Holmberg (2016) also mentions a second possibility which is (ii) to derive null subjects in vYNAs by VP-ellipsis plus standard pro-drop (Martins 1994*a*,*b*, 2004, 2012). As far as this second analysis is concerned with the null and overt subjects in vYNAs, Martins (1994*a*,*b*, 2004, 2012) assumes a standard pro-drop analysis. Since the syntactic proposal in Martins (2004) is focused on EP and cannot be extended to BP data (as pointed out in section 2.3.2), Martins (1994*a*, 1994*b*, 2012)'s pro-drop analysis is taken here to be pragmatic. To sum up, Holmberg (2013) compares the IP-ellipsis analysis to the pro-drop analysis. The latter is a pragmatic account at least as far as it concerns vYNAs. Finally, Kato (2013) attempts to adapt Holmberg's ideas to a "cartographic" syntactic analysis. As shown above, this attempt, roughly speaking, translates some pragmatic notions into syntactic nodes and can be discussed in terms of pragmatic consequences. However, differently from what is proposed in the present dissertation, the translation into a "cartographic" structure does not take into account the set of alternative propositions and other paradigmatic comparisons that are crucial in the Maximal-Contrast-based proposal.

Finally, going back to pragmatics and taking syntax to interact with this level, an alternative proposal based on the fact that the discourse has its own structure was proposed. Maximal Contrast was formally defined based exclusively on standard assumptions in semantics and discourse structure. Following other authors' observations about the interface between discourse and syntax

(especially as concerns missing parts of the sentence), a hypothesis (the MaxCoNAH) for accounting for missing arguments in verbal utterances was informally drawn. According to this hypothesis, missing arguments are allowed for informational reasons, when they are in Maximally Contrastive contexts.

In the next section, these different approaches are tested qualitatively against corpus data from the NURC-RJ project. The sentences presented there are supposed to be context-types, that is, representative examples of recurrent dialogue pairs in BP. In this respect, exceptions do not constitute the focus of the present dissertation.

2.4   Corpus Assessment

In this section, the approaches dealing with null and pronominal subjects in vYNAs presented in the literature review above are discussed and some example-types extracted from NURC-RJ BP Corpus are provided in order to illustrate the discussion.[31] Moreover, a comparison with the proposal put forth in the subsection 2.3.4 and the superiority of the MaxCoNAH over other analyses is defended. At the end of this section, some points to be investigated in Experiment 4 are suggested. Mainly, the design of an experiment for testing the structural hypothesis (IP-ellipsis Hypothesis) and the pragmatic-discursive hypothesis (MaxCoNAH) is proposed. In most of the examples in this section, the subject refers to (specific) animate entities, in order to control as much as possible the effects of the inherent features of the antecedents.

2.4.1   Context-Types from a Brazilian Portuguese Corpus

According to the review presented in the previous section, it is generally accepted that standard (unmarked) vYNAs in BP have null subjects. This observation is further confirmed in the qualitative corpus study reported henceforth. In (61), the polar question is positively answered by a null-subject vYNA. Also, when answering negatively, the negation particle (*não* "No") is commonly followed by another negation particle and a bare verb, as in (62). Example (63) – a negative

---

[31]   In this section, no quantitative data are provided. In order to carry on a quantitative analysis of vYNAs, the corpus would have to be enlarged.

question followed by a positive disconfirmatory answer – is problematic for accounts such as those proposed by Armstrong (2008), based on the expectation of the answer or discourse rhetorical orientation. A negative question followed by a negative answer in (64) is predicted by some approaches, but Armstrong (2008) fails to deal with this case, as explained below.

(61)   a. A – Mas $\_1$ passou   na   primeira?
                But you_1 pass.PST.3SG in.the first?

             "A – But did you_1 pass in the first [attempt]?"

      b. B – $\_1$ Passei.
                   pass.PST.1SG

             "B – Yes, I did."

                                              (NURC-RJ – "Inquiry_ac_01")

(62)   a. A – $você_1$ manja   disso?
                you   understand.PRS.3SG of.this

             "A – Do you understand it? [=Do you have the knack of this?]"

      b. B – Não, $\_1$ não manjo.
                   no   no understand.PRS.1SG

             "B – No, I_1 don't."

                                              (NURC-RJ – "Inquiry_96")

(63)   a. A – Assaltado mesmo, $você_1$ nunca foi?
                Robbed   same   you   never be.PST.3SG?

             "A – but have you_1 never been really robbed?"

      b. B – Não, já, $\_1$ já   fui,   mas de carro.
                no   already   already be.PST.1SG but of car

             "B – No, I_1 have already [been really robbed], but while I was in a car."

                                            (NURC-RJ – "Inquiry_ac_02")

(64)   a. B – _ não sei   se   vocês sabem   ... $funcionário publico_1$ não
                   not know.PRS.1SG whether you   know.PRS.3PL   servant   public   not
                pode   [ter   sindicato]$_2$ ... $\_2$ é   proibido por lei ...
                can.PRS.3SG have.INF union   be.PRS.3SG forbidden by law

             "B – I don't know whether you know it... Civil servants cannot have a union... this is forbidden by the law..."

b. A – funcionário público$_1$ [não pode]?
    servant   public   not can.PRS.3SG

"A – Can't civil servants$_1$?"

c. B – é, $_1$ não pode   ...
    be.PRS.3SG   not can.PRS.3SG

"B – Yeah, (=they$_1$ cannot.)"

(NURC-RJ – "Inquiry_164")

All the theories discussed above explain the null subject in the verbal confirmatory answers in (61); however, the negative confirmatory answer in (64) is not accounted for by Armstrong (2008), since interlocutor A expresses surprise regarding the denial of the proposition, which is reasserted by interlocutor B in the following turn. In terms of beliefs and expectancy, it is clear that interlocutor A believed that the "civil servants could have a union", contrary to what was said by interlocutor B. In the DGB theory (Ginzburg 1996, 2012, to appear), the truth of the proposition remains under discussion until the last turn. That is, the asserted information in the last turn is still an open variable in the previous turn, because it was not accepted by interlocutor A as being part of the common ground. This makes the vYNA in (64) Maximally Contrastive. In (63), interlocutor A expects a negative utterance, since the question is negative. The correction is thus unexpected. The fact that interlocutor B has already been robbed disconfirms interlocutor A's previous beliefs. However, the QUD is open until interlocutor B's final answer, denying interlocutor A's previous expectations about the common ground. For Armstrong (2008), it should then be a sentence with an overt subject. Holmberg (2001, 2007, 2013, 2014, 2016, to appear) would analyze it as a case of VP-ellipsis plus pro-drop rather than a case of IP-ellipsis (supposing that he accepts that BP has both the mechanisms for producing null subjects). Since the negative adverb *Nunca* "Never" would share its negative feature, such a non-open polarity would lead to an incorrect interpretation of the correction if IP-ellipsis took place.[32] Therefore, Holmberg (2001, 2007, 2013, 2014, 2016, to appear) would have to suppose that BP has both mechanisms to generate null subjects in vYNAs. Martins (1994*a,b*, 2012)'s pragmatic pro-drop analysis can deal with (61)-(64), since the subjects

---

[32]   In fact, IP-ellipsis apparently produces a feature clash, since *já* "already" in BP seems to have a positive feature. The negative version of *já* is either *ainda não* "not yet" or *nunca* "never". This question does not affect the fact that Holmberg (2001, 2007, 2013, 2014, 2016, to appear)'s analysis for this case could only be based on VP-ellipsis plus pro-drop.

are all unmarked topics. Finally, the MaxCoNAH correctly predicts all the patterns above to be subjectless, since all of them are Maximally Contrastive. That is, all QUDs produce alternative assertions that diverge only in their focused content, which is, in these cases, the polarity of the answer. Being expected to be positive or negative does not affect the choice between a null or overt subject according to the MaxCoNAH.

Cases (61)-(64) are all answers that directly address the polarity of the question and do not provide any further information. However, if additional information is provided by the speaker or is used to correct some presupposition (possibly) implied by the question, even when the speaker gives a confirming answer to the question, an overt-subject answer is used, as in (65)-(69).

(65) a. A – Ah   é,   você gostou   mais de lá   do   que daqui?
           [DM] be.PRS.3SG you like.PST.3SG more of there of.the that of.here
       "Ah yeah, did you$_1$ like that place more than here?"

   b. B – Eu$_1$ gostei   de lá,   até porque, aqui né,   tem,   já
          I   like.PST.1SG of there even because here [DM] have.PRS.3SG already
      existia,   não com tanta   força, mas esse medo de assalto,
      exist.IMP.3SG not with so.much force but this fear of robbery
      "B – I liked that place, even because here there is, there already was, although not so

      strongly, this fear of robberies."

                                                             (NURC-RJ – "Inquiry_ac_2")

(66) a. A – ... Você$_1$ fez   todo o   curso lá?
              you   do.PST.3SG all   the course there
       "Did you do the whole course there?"

   b. B – Bom, eu fiz   o   primário, o   jardim de infância e   o   primário
          Well I do.PST.1SG the first   the garden of childhood and the first
      lá.
      there, [DM]
      "B – Well, I did primary school, kindergarten and primary school there."

                                                             (NURC-RJ – "Inquiry_096")

(67) a. A – você$_1$ se lembra   do... que havia   dentro da
              you   SE remember.PRS.3SG of.the that there.be.PST.IMP.3SG inside of.the
      casa da   sua avó?   (...)
      house of.the your grandma
      "And do you$_1$ remember what was inside your grandma's house..."

b. B – ah... olha,   eu$_1$ me lembro   uma coisa horrível...
     [DM] see.IMP.3SG I$_1$ SE remember.PRS.1SG a   thing horrible
    "B – Well, I$_1$ remember a horrible thing..."

                                                     (NURC-RJ – "Inquiry_011")

(68) a. A – se você$_1$ fosse   estar   nessa fase do   ano no   Rio Grande do
         if you   be.SUBJ.PST.3SG be.INF in.this phase of.the year in.the Rio Grande do
      Sul, você usaria   um pulôver   desse tipo? (...)
      Sul you$_1$ use.COND.3SG a   pull-over of.this type
      "A – If you were in Rio Grande do Sul in this period of the year, would you$_1$ use a

      pull-over like this?"

  b. B – Não, eu$_1$ usaria   um pulôver   mais, mais completo,
           no   I   use.COND.1SG a   pull-over more more complete
      "B – No, I$_1$ would use a more complete pull-over."

                                              (NURC-RJ – "Inquiry_096")

(69) a. A – o   técnico$_1$ só   funciona   durante a   partida?
         the coach   only work.PRS.3SG during   the match
      "Does the coach$_1$ only work during the match?"

  b. B – não... ele$_1$ funciona   durante... a   semana toda...
           no... he   work.PRS.3SG during   the week   whole
      "B – No... he$_1$ works during the whole week"

                                              (NURC-RJ – "Inquiry 52")

Examples (65)-(67) have positive answers while (68)-(69) have negative answers, followed in all cases by the addition of some information. In all of them the subject is overt. Martins' (1994a,b, 2012) analysis fails to deal with these cases, since the subjects are unmarked topics.[33] Kato (2013) also cannot account for them, since the subjects are part of the background provided by the questions. Armstrong (2008) might possibly predict the use of overt subjects in some of these sentences, that is, the corrections of expectations in (68)-(69), but it is unclear whether her approach predicts (65)-(67), since they exhibit positive answers that go beyond the information

---

[33]   Martins (1994*a*,*b*, 2012) is primarily concerned with EP; however she proposes to extend her analysis to BP, claiming that in vYNAs EP and BP do not diverge much.

required by the question. Her explanation for the answers that tighten the truth conditions of the questions, as in (66)-(67), is especially unclear, since the answers go in the same rhetorical direction, and are simply more specific. There are two possible accounts for the group of examples (65)-(69). Holmberg (2001, 2007, 2013, 2014, 2016, to appear)'s approach would predict that these cases cannot be IP-elliptical because the additional information is within the VP domain, that is, it is internal to IP, which consequently cannot be deleted under identity while leaving it behind. The MaxCoNAH also predicts that null subjects would not be allowed in (65)-(69): the additional information goes beyond the current QUD and thus does not fit in the definition of Maximal Contrast. That is, these utterances do not fit the criteria for being Maximally Contrastive since they are not part of the set produced by the current QUD. In (65), for example, the question is a comparison between "there" and "here". However, the answer is not a comparison and, by implicature, denies the assertion of the comparison. The answers in (66), (67) and (68) tighten the truth conditions given by the respective questions. So, they are not part of the current QUD. Finally, (69) is a correction of what is supposed in the question.

Also, when some element is fronted, overt subjects tend to be used preceding vYNAs even if the dislocated element is mentioned in the question, as in (70)-(73).

(70)   a.   A – A   senhora$_1$ conhece   a   Amazônia?
              The lady   know.PRS.3SG the Amazon
           "Do you$_1$ [formal] know the Amazon [rainforest]?"

       b.   B – Não, a   Amazônia eu$_1$ não conheço.
              No, the Amazon   I   not know.PRS.1SG
           "B – No, the Amazon [rainforest] I$_1$ don't know."

(NURC-RJ – "Inquiry_347r")

(71) a. A – agora... me   diz   uma coisa... você$_1$ só   joga?   você$_1$ não
               now   to.me say.PRS.3SG one thing   you   only play.PRS.3SG you   not
          vai   assistir?
          go.PRS.3SG watch.INF
          "Now, tell me, you only play? Don't you go to watch?"

      b. B – não... assistir   eu$_1$ vou...
               no... watch.INF I   go.PRS.1SG
          "B – No, I$_1$ go to watch."

(72) a. A – (...) agora _$_1$ não temos   mais [lacerdinhas] né?
               now   not have.PRS.1PL more [thrips]   [DM]?
          "A – Now we$_1$ don't have [thrips] anymore, do we?"

      b. B – Agora, nós$_1$ tivemos   aí   [lacerdinhas]
               Now   we   have.PST.1PL there [thrips]
          "B – Now, we$_1$ had [them] there."

(73) a. A – (...) _$_1$ não tem   medo de, de roubo não?
                   not have.PRS.3SG fear   of of robery not
          "A – Aren't you$_1$ afraid of robberies?"

      b. B – olha,   roubo   eu$_1$ tenho
               see.PRS.3SG, robbery I   have.PRS.1SG
          "B – Well, robbery, I$_1$ am [afraid of]."

In (70)-(73), overt-subject vYNAs are used to disagree with the respective questions (giving either a positive answer to negative questions or negative answers to non-negative questions). Also, as mentioned before, the fronted elements, as well as the subjects, are part of the question. Therefore, neither Kato (2013) nor Martins (1994*b*,*a*, 2012) account for these examples. The Max-CoNAH, however, accounts for these cases: for a subject to be null in a vYNA, it is not enough that its referent is backgrounded and/or topic, it must also be Maximally Contrastive, as predicted by the MaxCoNAH. Another possibility is to account for these examples in structural terms: as they are extracted from within the VP domain, IP-ellipsis cannot be applied to these cases, so the null

subject is not an option Holmberg (2001, 2007, 2013, 2014, 2016, to appear). Holmberg (2001, 2007, 2013, 2014, 2016, to appear) would thus appeal to standard assumptions about pro-drop in his framework to predict the data in (70)-(73).[34]

When the subject is somehow different in the questions and in the answers, the subject of the vYNA tends to be overt, as in (74)-(76) below. These cases are considered to be prototypical contrastive topics (see Büring 1999, 2003, *inter alia*). Most of the approaches that deal with null and overt subjects make some mention of these cases. They are of course not Maximally Contrastive, since, as proposed by Büring (1999, 2003), the related set of alternative is not monotonic. Rather, in the present approach, the overt subject picks up one of the possible backgrounds. In this respect, the definition of Maximal Contrast does not fit these cases, and thus the MaxCoNAH predicts an overt subject.

(74)   a. A – Normalmente as pessoas$_1$ têm   alergias, né?
            normally   the people   have.PRS.3PL allergies [DM]
        "Usually, people$_1$ have allergies, don't they?"

     b. B – Eu$_2$ não tenho...
            I   not have.PRS.1SG
        "B – I$_2$ don't."

                                              (NURC-RJ – Inquiry_096r)

(75)   a. A – Todo mundo$_1$ pratica   alguma coisa?
            every world   practice.PRS.3SG some   thing
        "Does everybody$_1$ practice something?"

     b. B – É,   embora a   minha idade, eu$_2$ ainda gosto   de praticar o
            be.PRS.3SG... despite the my   age   I   still   like.PRS.1SG of practice the
            esporte.
            sport
        "B – Yeah. Despite my age I$_2$ still like to practice sports."

                                              (NURC-RJ – Inquiry_052r)

---

[34]   It is unclear to me which explanation for the unacceptability of null subjects in (70)-(73) Holmberg (2001, 2007, 2013, 2014, 2016, to appear) would appeal to. It is assumed here that, not being generated by IP-ellipsis, he would claim that other syntactic constraints block the use of other syntactic types of null subjects that are proposed in his theory (see, in this regard, Chapter 1).

(76) a. A – (...) A   senhora$_1$ tem   caderneta de poupança?
                The lady   have.PRS.3SG account   of savings?

      "A – Do you$_1$[formal] have a savings account?"

      b. B – Ah, claro, caderneta de poupança a   gente$_2$ tem,   né?
                Ah, clear, account   of savings   the people have.PRS.3SG [DM]

      "B – Ah, of course, savings accouns we$_2$ have, right?"

(NURC-RJ – Inquiry_002)

To give further evidence for the MaxCoNAH, it is worth comparing vYNAs to a similar context, in which a set of alternatives is explicit, namely alternative "Or-questions", as in (77a) below. In this kind of question, the MaxCoNAH predicts a null subject if one of the alternatives is asserted. However, if the speaker gives an answer that is not in the set of alternatives, as in (77b) in below, the MaxCoNAH predicts an overt subject. And indeed this prediction holds in the corpus data. Approaches based only on the topicality of null subjects do not have much to say about these cases.

(77) a. A – A   moça$_1$ é   mais parecida com o   senhor$_2$ ou com a   sua
                the young.lady be.PRS.3SG more similar   with the lord   or with the your
            senhora$_3$?
            lady

      "A – Does the young lady$_1$ look more like you[formal]$_2$ or more like your wife$_3$?"

      b. B – A   moça$_1$ é   mais parecida com a   minha mãe$_4$.
                the lady   be.PRS.3SG more similar   with the my   mother

      "B – The young lady$_1$ looks more like my MOTHER$_4$."

(NURCRJ – Inquiry_148)

Finally, it is worth pointing out that even though yes-no questions are generally about truth relative to a world and the answers should be relative to this same world, it is possible to make modal corrections with null-subject vYNAs. In (78) and (79), the answers to the question are relative to possible worlds, which are not the current one: in (78), the world of interlocutor B's wishes, and in (79), a possible world where ties are created between employer and employee (*isso*, "this", in this context refers to relations between employers and self-employed workers who render some service to their employers). Notice that deletion under identity of the whole IP is not

146

possible, because the antecedent and the vYNA do not share the same structure (unless additional assumptions about the verbal tense and mood nodes are proposed).

(78) a. A – o   senhor₁ toca   algum instrumento?
          the lord   play.PRS.3SG any   instrument?
      "A – Do you[formal]₁ play an instrument?"

   b. B – Não, ₋₁ gostaria   de tocar.
          no   like.COND.1SG of play.INF
      "B – No, I₁ would like to."

(NURCRJ – Inquiry_r_052)

(79) a. A – isso₁ já   não cria   um vínculo não? (...)
          this   already not create.PRS.3SG a   bound   no
      "A – Doesn't this₁ create a bound already?"

   b. B – é   ... é   ... ₋₁ pode   criar   ... ₋₁ pode
          be.PRS.3SG   be.PRS.3SG   can.PRS.3SG create.INF   can.PRS.3SG
      criar   ...
      create
      "B – yeah... yeah... it₁ can... it₁ can..."

(NURCRJ – Inquiry_164)

2.4.2   Discussion and Directions

Previous Approaches vs. Maximal Contrast

It was shown that previous literature has proposed two strategies for analyzing the relation between Yes-No questions and vYNAs, a pragmatic account, proposed by Martins and Armstrong, for example, and a syntactic strategy, proposed by Holmberg, for instance. More specifically, Martins (1994*a*,*b*, 2012) proposes that the use of null subjects is due to the fact that they are unmarked topics; Armstrong (2008) claims that null subject in vYNAs are the unmarked option and the use of overt subjects is a strategy to highlight that the answer is "unexpected". Both approaches fail to deal with some of the example-types collected from the corpus. On the other hand, Holmberg (2001, 2007, 2013, 2014, 2016) proposes a structural account to deal with the absence of the subject, and suggests an especial mechanism (IP-ellipsis) for producing null subjects in vYNAs,

but assumes that VP-ellipsis plus standard pro-drop assumptions can also take place. His proposal can potentially deal with several example-types, but it raises many theoretical questions, which will be studied next. Finally, Kato (2013) argues for a structural account with pragmatic underlying import. Nonetheless, her proposal cannot account for many examples in which the subject is overt despite being in the background.

Holmberg (2001, 2007, 2013, 2014, 2016) has a powerful device for explaining the presence and absence of subjects in vYNAs. However, his analysis needs a specific mechanism to deal with subjectless vYNAs – IP-ellipsis. Since BP is taken to be a language in which overt subjects are getting scarcer, the availability of such a specific mechanism would not be surprising. However, as mentioned in section 2.2.2, from the earlier stages of acquisition, children's grammar is similar, in this respect, to their target (adults') grammar. The question thus arises as to how children learn the IP-ellipsis mechanism to generate null subjects in vYNAs so quickly, given the absence of positive evidence for it and the fact that vYNAs are not mere repetitions. As pointed out by Simões (1997), children learning BP behave in a more similar way to children learning non-pro-drop languages than standard pro-drop languages. Where does the use of null subjects in vYNAs come from then? Also, Holmberg (2001, 2007, 2013, 2014, 2016) argues that pro-drop cannot have as antecedent an indefinite subject.

However, vYNAs are not the only construction that can have a null subject retrieving an indefinite, but they are rather a context in which the null subject is highly likely by the convergence of multiple factors, as shown in section 2.8 below. As pointed out in Chapter 1, null subjects are the preferential option for retrieving this kind of referent. Moreover, as shown in example (80) below, null subjects in standard pro-drop contexts do retrieve indeterminate referents. Furthermore, given some conditions, null subjects can be the only option, since overt pronouns are less likely with "Nobody antecedents". Finally, the use of null subjects can produce a very specific reading that are hardly derived by ellipsis. In (80) for instance, the subsequent mentions of the antecedent "Nobody" actually mean "everybody", an interesting switch that cannot be predicted by theories based on (strucural) ellipsis of missing arguments without further assumptions.

(80) a. Agora, nas   férias,   ninguém$_1$ quer   pensar   nisso. _/??ele$_1$
          Now   in.the holidays, nobody   want.PRS.3SG think.INF in.this.
          Quer   esquecer até   que _/??ele$_1$ é   professor. _/??ele$_1$
          want.PRS.3SG forget   even that he   be.PRS.3SG teacher
          Quer   descansar.
          want.PRS.3SG rest.INF

          "Now, on holidays, nobody$_1$ wants to think about this. They$_1$ even want to forget that

          they$_1$ are teachers. They$_1$ want to rest."

Occam's razor would lead one to abandon Holmberg (2001, 2007, 2013, 2014, 2016)'s proposal unless it is correct to propose that null subjects in vYNAs are especial. The especial character of this construction, however, is a puzzle for studies of the acquisition of verbal answers and of null subjects in general. Another possible strategy is to assume that pragmatic concepts (on an informational structure level), such as "topic", "background", "expectedness" and "focus", are relevant and to attempt to refine this set of criteria in order to explain which of them are really decisive. These concepts can be encoded in a purely pragmatic way (as in Armstrong 2008) or in a structural way (as in Kato 2013). Also, some flexibility due to pragmatic factors or performance can be taken into account (as in Martins 1994*a,b*, 2012). As in Roberts (1996), this account is rejected here because of the fact that there is no terminological or conceptual agreement on which notions should be taken into account. Rather, it is proposed in this dissertation to understand the actual context in which these utterances are appropriate (following Roberts 1996, Ginzburg 1996).

Finally, an analysis that is truly focused on determining the discourse contexts in which a verbal utterance and the use of null and overt subjects are felicitous was proposed in 2.3.4, by defining the structure of the relevant discourse context (Maximally Contrastive) and sketching a mapping constraint (MaxCoNAH). This proposal deals with most of the context-types gathered from the corpus and generalizes over other previous analysis. However, in many cases, the IP-ellipsis approach and the MaxCoNAH converge to the same predictions. In order to disentangle the factors that are really playing a role in the use of overt and null subjects in verbal answers and to better understand these constructions, one experiment was carried out. The results are reported in the next section. Before moving on to the report on the experiment, however, a theoretical discussion is required.

Introducing (Non-)At-Issueness

One of the main factors claimed here to be relevant for the analysis of vYNAs is the syntactic and discourse structure of the verbal clause in which the null or overt subject occurs. In this vein, in BP some of the elements that can be easily manipulated to test the internal structure of these answers are adjuncts. Thus, a discussion of the syntactic type, way of attachment and position of adjuncts and their discursive-semantic status is required. Two types of adjuncts must be differentiated for the purposes of the present dissertation. (i) On the one hand, adjuncts generated internally to the VP, whether they are fronted or not, which are "scenery" adjuncts, that is, those which are taken to clearly scope over internal elements of the sentence (for instance the event), such as those of mode, time, and location. These adjuncts generally answer Wh-QUDs starting with *How*, *When* and *Where* and they are part of the propositional content (see Wechsler 2005 for an overview). (ii) On the other hand, (Left-most) external to IP-generated adjuncts or evaluative adjuncts: those are taken to scope over the whole sentence, by adding either the speaker's judgment about the propositional content or some kind of evaluation of the likelihood of the truth of the proposition (some examples are "unfortunately", "fortunately", "hopefully", "of course", "no doubt", "really", etc);[35] these adjuncts are generally taken to be Non-at-Issue content (see Simons et al. 2010, Destrual et al. 2013, Potts 2015). Many definitions of "At-Issue" and "Non-at-Issue" have been proposed. In the same vein as for the definition of Maximal Contrast, At-Issueness here is defined as follows.

(81)   Definition of Complete At-Issueness

Given a proposition α in a context Γ, such that Γ contains a set of shared accepted propositions P at a given discursive time-point t, the whole content of α is At-Issue iff:

a. α is asserted with a focused semantic part x, so that α_x is its discursive-semantic representation;

b. y is open in the *current* QUD q_y, such that it produces a set of alternative assertive propositions Q_y [$\alpha_x \subset Q_y \wedge Q_y \not\subset P$ at t];

c. $\forall \alpha_x \forall \alpha_z \subset Q_y \exists q_y [y = x \vee z]$;

---

[35]   Amaral & Schwenter (2003) observe that similar adverbs (which they claim to be speaker oriented) can facilitate or prevent the use of overt subjects is Portuguese and Spanish in other constructions beyond those tested in Experiment 4, given proper semantic and pragmatic conditions.

The definition of At-Issueness in (81) is a less restricted version of Maximal Contrast (Maximal Contrast, defined in 57, differs from 81 only in that the latter includes the additional clause 57d). As expected, it predicts that any Maximal Contrastive utterance is a Completely At-Issue utterance, while not every At-Issue Utterance is necessarily Maximal Contrastive. As before, by (81a), an assertion is a proposition of which some semantic part is focused. By (81b), the assertion is contained in the set of alternatives proffered by the hypothetical current QUD. And finally, by (81c), whatever value is attributed to the open semantic part of the question, it is a focused part of one or more of the assertions in the set of propositions generated by the question. To understand how it describes the discourse structure of a given utterance, the definition of Complete At-Issueness is applied to the example (82) below. By definition (81), there would be two possible ways of having an adverb, like *realmente* "really" in (82), as an At-Issue element: (i) either it is part of the set of alternatives generated by the question as a backgrounded element, and will consequently be present in the current QUD (by 81b) or (ii) it is focused, and is consequently one of the possible values for the open part of the current QUD by (81c). In example (82) below, the only element of the answer in (82b) that is not present in question (82a) is the adverb *realmente* "really". The set of asserted alternatives defined in (82a) clearly does not include *realmente*. However, it is deliberately included in answer (82b) by the speaker, who makes an additional contribution that goes beyond the current QUD(s).

(82)   a. A – O senhor$_1$ tem   um filho diplomata. O   senhor$_1$ viaja
              the mr.   have.PRS.3SG a   son diplomat The mr.   travel.PRS.3SG
              muito? Ou $_1$ não viajou?   Ou$_1$ gosta   de viajar? $_1$
              much or   not travel.PST.3SG or   like.PRS.3SG of travel.INF
              Tem   alguns planos... nesse sentido?
              have.PRS.3SG some   plans   in.this sense
              "A – You[formal]$_1$ have a son who is diplomat. Do you$_1$ travel often? Or didn't you$_1$ travel? Or do you$_1$ like to travel? Do you$_1$ have any plans in this respect?'

        b. B – Eu$_1$ gosto,   realmente, de viajar.
              I   like.PRS.1SG really   of travel.INF
              "B – I$_1$ really like traveling."

<div align="right">(NURC-RJ, "Inquiry_024")</div>

It is worth comparing example (82) with a different case, in which the adjunct is included in

<div align="center">151</div>

the question. In (83), the adverb *normalmente* "normally" is part of the current QUD, and is thus at-issue at the moment at which answer (83b) is uttered.

(83)   a. A – E   normalmente o, o   que   faziam   as moças$_1$ na, ne... nesse
                And normally   the what do.PST.IMP.3PL the ladies   in   in.that
          tempo?
          time
          "A – And normally what did the ladies do at that time?"

       b. B – _$_1$ Casavam-se.
                   marry.PST.IMP.3PL-SE
          "They$_1$ married."

(NURC-RJ, "Inquiry_076")

Simons et al. (2010, 323) define proposition at-issueness, by arguing that a proposition is at-issue if the speaker intends to address the QUD via such a proposition. This intention is felicitous only if the proposition is relevant to the QUD and the speaker can reasonably expect the addressee to recognize this intention. This definition is too strict for the purposes of the present dissertation, since it targets only propositions, rather than parts of a given assertion. The definition in (81) translates the definition made by Simons et al. (2010) into a property of parts of propositions (or assertions). The notion of relevance is translated into the clauses (b) and (c) (in set theoretical notation) and the acknowledgement of the speaker's intention is not relevant in the present dissertation. The latter property is a topic left for future research, though. For the purposes of the present chapter, the definition of Complete At-Issueness is enough, and will be used to explain the design and results of Experiment 4 presented in section 2.5 below.

2.5   Experimental Evidence
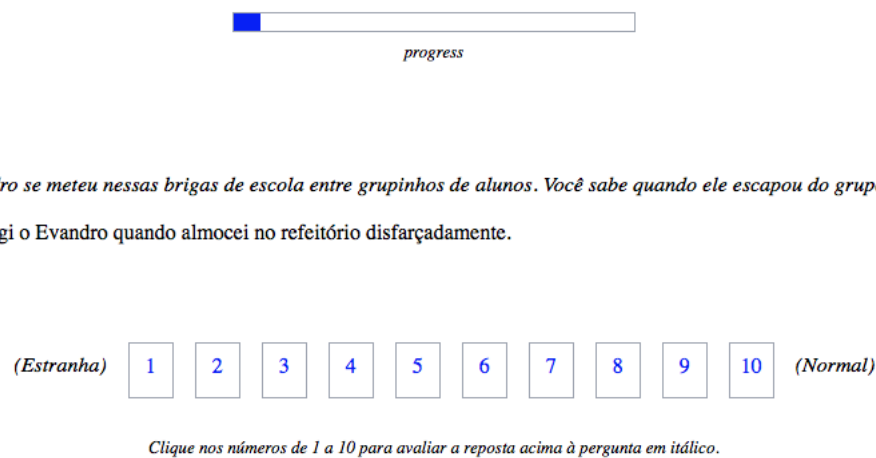
Given the observations from the corpus research described in the previous section and the theories that could possibly account for the use of null and overt subjects in vYNAs, an experiment was designed and carried out in order to attempt to evaluate the possible accounts for the use of overt and null subjects in vYNAs. In this experiment, two Maximally Contrastive contexts (Yes-No Answers and unary Wh-verbal answers) were tested; furthermore, adding a non-at-issue adjunct

before the verb allows one to check whether an element external to VP can influence the choice between overt and null subjects in the context of an obvious subject referent, by making the utterance non-Maximally Contrastive. The main finding is a significant interaction of factors between the overt or null subject (Factor Subject) and the use of a non-at-issue adjunct (Factor NaIAdj), but no significant interaction with the type of question, either a unary wh-question or a yes-no question (Factor QUD).

2.5.1 Methodological Considerations

In this experiment, as in the experiments reported in Chapter 1, participants read a sequence of two turns, that is, a short dialogue. They were asked to judge the acceptability of the answer in the relevant context on a scale from 1 to 10, cf. Figure 2.1. They were told to use the full scale according how natural "Normal" or strange "Estranha" the answer sounded in the context of the question. After judging the answer, participants were asked about the interpretation of the relevant subject – null or overt – in a closed yes-no question task, cf. Figure 2.2.

Figure 2.1 – Screen sample – Judgment Task

progress

O Evandro se meteu nessas brigas de escola entre grupinhos de alunos. Você sabe quando ele escapou do grupo rival?

Eu protegi o Evandro quando almocei no refeitório disfarçadamente.

(Estranha) 1 2 3 4 5 6 7 8 9 10 *(Normal)*

Clique nos números de 1 a 10 para avaliar a reposta acima à pergunta em itálico.

Figure 2.2 – Screen Sample – Closed Question Task
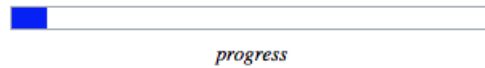
progress

Era o Evandro quem tinha almoçado no refeitório disfarçadamente?

1. Não.

2. Sim.

All participants voluntarily participated in the experiment, which was set up on the IbexFarm platform (http:// spellout.net/ibexfarm). They filled in a basic information form, which included a consent statement, and had 5 sentences to practice before starting the experiment, which took them around thirty minutes to complete. As in the previous experiments reported in section 1.7, highly acceptable control sentences were randomly presented among the experimental sentences, and unacceptable sentences were presented at the very end of the experiment.

2.5.2   Experiment 4 – Maximal *vs*. Non-Maximal Contrast

This experiment was designed to check whether, in contexts controlled for information structure, the use of NaI adjuncts has a significant effect on the acceptability of null and overt subjects in BP. In all sentences, the subject was informatively and structurally salient both in the context and in the question. Three two-level Factors were tested: null vs. overt subject (Factor "Subject") and presence vs. absence of a non-at-issue adjunct (Factor "N[on]-a[t]-I[ssue]Adj[unct]") and Yes-No question vs. unary Wh-question (Factor "Q[uestion]U[nder] D[iscussion]"). The hypotheses were the following: (i) if the approach based on IP ellipsis accounts for the use of null subjects in vYNAs, adding a non-at-issue external to VP adjunct should not interact significantly with the use of null subjects in this context; (ii) if IP or VP ellipses in these contexts require some sort of verbal structural identity, a significant interaction of Factor QUD with Factor Subject is expected, since unary Wh-questions do not provide the right antecedent for any of these ellipses; and (iii) if Maximal Contrast and the MaxCoNAH account for the use of null subjects in these contexts,

Factor QUD should not interact significantly with Factor Subject while the interaction between Factor NaIAdj and Factor Subject should be significant: verbal answers preceded by a NaI adjunct would be significantly more acceptable when the subject is overt than when the subject is null, since null subjects should be likely to appear mostly in Maximally Contrastive contexts.

Material Design

Twenty-four items were created for this experiment, based on the corpus research and on the previously described approaches. Each item started with a context sentence, such as (84a), followed either by a yes-no question, as in (84b), or by a unary wh-question, as in (84c). In the possible contexts and questions, the referent of the relevant subject is unambiguous, is activated and is completely obvious. For this reason, these contexts are taken to be controlled for Activation, Familiarity, Identifiability and all the other properties related to antecedent contextual Salience. In both the questions, the referent of the subject is also the only possible topic. The only manipulation in the context is due to the focus of the question: either the "truth" of the proposition in the yes-no question in (84b) or the "event" that took place in the unary wh-question in (84c). This manipulation in the context question is encoded as Factor QUD (Yes-No question vs. Wh question). Differently from the previous experiments, thus, in Experiment 4 the context question was one of the manipulations. Also differently, instead of only controlling for information structure, the question under discussion is a crucial factor in this experiment. Beyond testing the two different discourse contexts, possible underlying syntactic structures were manipulated, since it is only in the Yes-No context that the verb in the question is the same as in the answer; in the Wh-question context, the verb is the hyperonym *fazer* "to do", which can refer to basically any agent-patient interaction.

(84)   a. A – Só o Felipe$_1$ fez os exercícios com material esportivo hoje de manhã.

"Only Felipe$_1$ worked out with the sport equipment this morning."

b. O Felipe$_1$ guardou as bolas de basquete?

"Has Felipe stored the basketballs?"

c. O que o Felipe$_1$ fez com as bolas de basquete?

"What has Felipe done to the basketballs?"

155

The answers to the questions encoded other two Factors in a two by two design. They were the following: either the bare verb provided a satisfactory answer to the questions, as in (85a) below, or it could be preceded by a pronominal subject, as in (85b). Both these answers could also be preceded by a non-at-issue evaluative adjunct, as in (85c) and (85d). To sum up, the four possible answers were the following: (a) null subject plus verb, as in (85b); (b) pronominal subject plus verb, as in (85b); (c) evaluative adjunct followed by null subject plus verb, as in (85c); and (d) evaluative adjunct followed by pronominal subject plus verb, as in (85d). These four possible utterances were presented as answers to one of the two questions in (84b) and (84c), producing a total of eight conditions for the judgment task.[36]

(85) a. B – $\_1$ Guardou.
             store.PST.3SG
     "Yes, (he$_1$ has)." "He$_1$'s stored them."

    b. B – Ele$_1$ guardou.
           He   store.PST.3SG
     "Yes, (he$_1$ has)." "He$_1$'s stored them."

    c. B – Sem   dúvida $\_1$ guardou.
           Without doubt   store.PST.3SG
     "Yes, (he$_1$ has) no doubt (done so)." "He$_1$'s no doubt stored them."

    d. B – Sem   dúvida ele$_1$ guardou.
           Without doubt he   store.PST.3SG
     "Yes, (he$_1$ has) no doubt (done so)." "He$_1$'s no doubt stored them."

Afterwards, the participants had to answer an interpretation question indirectly questioning whether the action described by the verb in the answer was performed either by someone else (in which case the expected answer was *Não.* "No") or nobody else (in which case the expected answer was *Sim.* "Yes"), cf. (86).

(86) Alguém mais além do Felipe guardou as bolas de basquete?

    "Did someone else besides Felipe store the basketballs?"

---

[36]   In all the cases, the object is omitted. It is possible that in the Yes-No question Conditions, the omission is due to some kind of VP-ellipsis, however this cannot be the case of the Wh question Conditions.

a. Sim.

  "Yes."

b. Não.

  "No."

Following the observations drawn from the corpus study and the hypotheses in the present dissertation, the empirical predictions are the following: if the MaxCoNAH plays a role in the choice between overt and null subjects, null subjects should be more acceptable only with the bare verb than when preceded by an evaluative adverb; whereas overt subjects should be better rated than null subjects when preceded by an evaluative adjunct. This result should be reproduced in both possible contexts: when the context is a direct yes-no question and when it is a unary wh-question, since both of them are Maximally Contrastive. In other words, it is expected that answers (85a) and (85d) are more acceptable than (85c) and (85b), regardless of whether the question is (84b) or (84c). If, on the other hand, the hypothesis based on the syntactic structure of BP is correct and null subjects in such contexts are generated by IP-ellipsis, the null subject should be better rated when answering the yes-no question (84b) than when answering the unary wh-question (84c), given that only the first provides the appropriate syntactic antecedent for IP-ellipsis; no influence of the presence of the evaluative adjunct is expected under this hypothesis, since these elements are base-generated constituents external to VP. Answers to the interpretation task are not expected to diverge across conditions.

Participants

One-hundred-fifty-two participants took part in this experiment, all of them higher educated (minimally under-graduate students) living in the south region of Brazil (Rio Grande do Sul and Santa Catarina). They were invited to participate via email and Facebook, and gave their consent to use the data. Their age averaged 32,5 years old ($\pm$ 6,2 years). They work or study in several different fields (as in previous experiments, all the 4 big areas of knowledge described by the Brazilian government were included). In the analysis of the results, only cases with correct answers to the interpretation task were taken into consideration. This affected less than 10% of the raw data collected. Eight participants were excluded for having rated control sentences over 8 or answered

more than 40% of questions wrongly.

Results

As shown in Figures 2.3 and 2.4 below, the experimental results confirm the hypothesis that the MaxCoNAH, rather than structural IP-ellipsis, impacts on the acceptability of the null and overt subjects in BP: the acceptability of the null subject decreases in a non-maximally contrastive contexts, that is, when a NaIAdj is added to the sentence. As shown by the comparison between the two charts, the pattern of the mean acceptability is roughly the same in Yes-No questions and Wh questions. So, regardless of Factor QUD, when a verbal answer is preceded by a NaIAdj, the overt subject is better rated than the null subject (averaging -0.5 difference between null and overt subjects). When no adjunct was present, the null subject is better rated (averaging 9.17 in answers to Yes-No questions and 9.04 in answers to Wh-questions).

Figure 2.3 – Mean Acceptability Judgments with Yes-No questions depending on the Subject and NaIAdj Factors
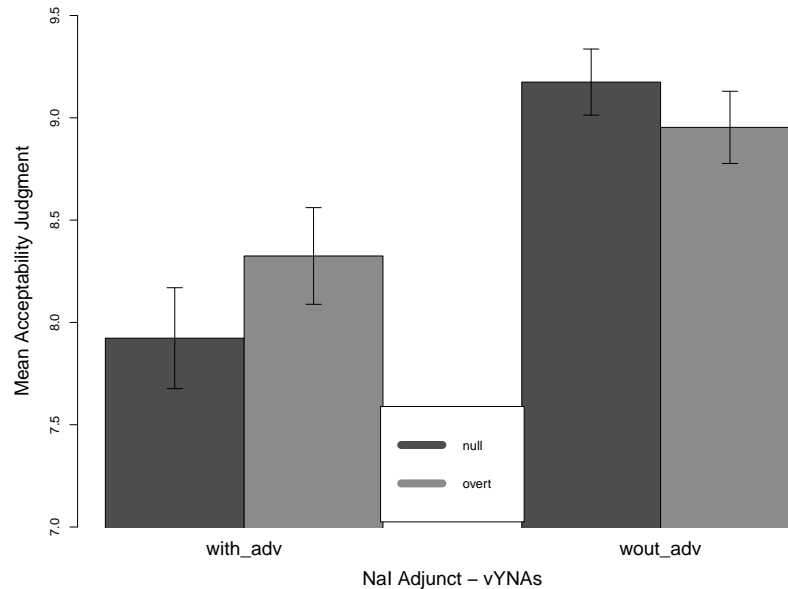
For the inferential statistical analysis, mean acceptability judgments were entered into a log-linear mixed-effect model analysis containing three Factors with two levels (Subject, NaIAdj and QUD) and random effects (Participants and Items, including random slopes) (Barr et al. 2013). Subject, NaIAdj and QUD, are individually significant. According to the graph, including adjuncts

Figure 2.4 – Mean Acceptability Judgments with WH questions depending on the Subject and NaIAdj Factors
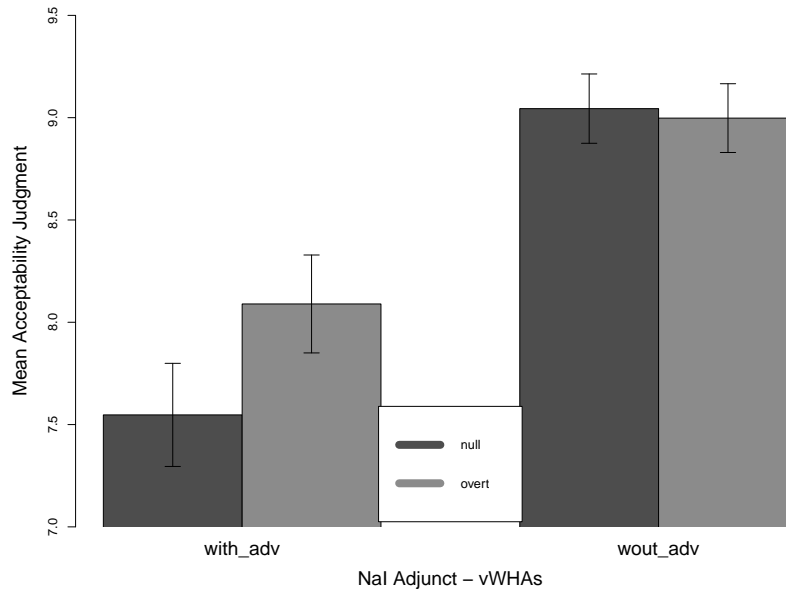
seems to decrease acceptability in general. This makes sense since these are non-optimal answers. A significant interaction NaIAdj and Subject was found. As for Factor QUD, no significant interaction with other two Factors and their interaction (Subject and NaIAdj, Subject:NaIAdj) came up.[37]

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	7.56445	0.18856	40.116	<2e-16 ***
Subject	0.46740	0.11822	3.954	9.96e-05 ***
NaIAdj	1.40293	0.16546	8.479	2.09e-13 ***
QUD	0.42338	0.14969	2.828	0.00644**
Subject:NaIAdj	-0.39300	0.15696	-2.504	0.01234*
Subject:QUD	-0.23974	0.15926	-1.505	0.13234
NaIAdj:QUD	-0.15551	0.15348	-1.013	0.31104
Subject:NaIAdj:QUD	-0.06988	0.21888	-0.319	0.74954

Table 2.5 – Log-linear mixed-effects model for Answers in Experiment 4

---

[37] The effect of the QUD alone seems to be due to the fact that, in the wh question Condition, all the judgments are slightly pushed down, possibly because in the Yes-No question Conditions the verb is the same in the question and in the answer while they are different in the Wh question Conditions (repetitions are generally taken to be more acceptable).

Discussion

The results of Experiment 4 suggest that the discourse structure of the sentence in which the null subject occurs is a relevant factor in the null subject's acceptability: BP speakers tend to rate null subject sentences better when they are Maximally Contrastive. This claim is supported by the examples collected from the corpus and by the evidence given by the results of the present Experiment. It was shown that fronting almost any structure before the answer to a direct question seems to favor the use of overt subjects; furthermore not only VP internal elements or moved constituents are able to favor the use of pronominal subjects. As shown in Experiment 4, base generated evaluative adjuncts external to VP can also favor the use of pronominal subjects rather than null subjects. Notice that the effect observed in sections (2.4) and (2.5.2) in both corpus and experimental data is completely orthogonal to the fact that the subject is a salient topic of the discourse. Therefore, when extending the generalizations made by Huang (1982, 1984, 1989) about the topic position and the bound variable behavior of null subjects in BP, previous research were missing one important constraint on the use of null subjects, as regards the information structure of the clause in which the null subject occurs (Ferreira 2000, Modesto 2000*a,b*, Rodrigues 2004, Modesto 2008*a,b*). Experiment 4 is striking in that it shows the effects of the information structure on the acceptability of null and overt pronouns regardless of structural conditions. It is impossible to predict the pattern of acceptability of null and overt subjects in the contexts used in Experiment 4 by any currently proposed structural account without further stipulation (see Amaral & Schwenter 2003, for similar observations about Portuguese and Spanish as regards the pragmatic import of left-most elements preceding null and overt subjects).[38] These results thus suggest that BP speakers seem to make use of the principle of Maximize Contrast, which can be understood to be a consequence of the Gricean Maxims of Relevance and Quantity. Speakers' paradigmatic knowledge seems to favor the construction that is optimal for comprehension, and as economic as possible (a maximal payoff in Roberts 1996's terms). Because it is highly salient in the preceding discourse and Maximally Contrastive as the only constituent out of the informational focus, the subject is dropped according

---

[38]   The observations made by Amaral & Schwenter (2003) did not clearly show that Spanish is sentive to Maximal Contrast, but their data suggest that the contrast between at-Issue and Non-at-Issue can play a role in the use of null and overt subjects in Spanish.

to the MaxCoNAH – a possible implementation of the Maxims of Quantity and Relevance (see Grice 1975). The preference for overt subjects increases as the sentence is less contrastive and the referent of the subject is less obvious. Sections 2.6 and 2.7 below will further investigate the interaction between prominence factors, such as the syntactic and discourse salience of the antecedent, and the proposal about the discourse structure of the clause in which the subject occurs.

2.6 Contrast and/vs. Accessibility - the Properties of the Antecedent and of the Anaphoric Subject

In Chapter 1, the hypothesis of a reverse relation between the Complexity and Explicitness of the anaphoric form and the Salience of the antecedent was briefly presented. In section 1.4.2, it was mentioned that the notion of Salience of the antecedent is highly controversial. Especially it was claimed that possibly contradictory predictions are borne out from different notions of Salience. In that section, the inherent Salience, given by the semantic features of the antecedent, was stressed as a factor that is relevant for the use and interpretation of the anaphoric subject. In the present section, the Salience given by the syntactic and discourse relation of the antecedent in the context in which it occurs is presented and studied. In subsection 2.6.1, different approaches to contextual Salience and Accessibility of the antecedent are summarized. In section 2.6.2, the notion of contextual Salience assumed in the present dissertation is incorporated into the MaxCoNAH.

2.6.1 The Discourse and Syntactic Salience of the Antecedent

As mentioned in Chapter 1, in the literature about anaphora resolution, it has been accepted that less Complex and Explicit anaphoric forms are biased toward more salient, prominent and accessible antecedents. Many different concepts of Complexity, Explicitness, Salience and Accessibility have been proposed and confronted. With respect to the more specific question at the center of the present dissertation, however, there has been general agreement that overt pronouns are more Complex and more Explicit than null pronouns. The concepts and applications that remain unclear about this principle are the relevant notion and empirical coverage of salience, prominence and Accessibility that are linguistically relevant.

The discussion about Accessibility of referents goes back to the philosophical tradition (for example, Frege 1989, Russell 1905). Uniqueness, Familiarity, Inclusiveness and Identifiability have been claimed to influence the degree of Accessibility of a given referent. Heim (1982) and Kamp (1981) independently developed theories to organize these multiple concepts into a more comprehensive theory of reference. Lambrecht (1994) remarkably claims that the cognitive concept of (non-)identifiability does not directly correspond to any linguistically relevant notion of Salience or Accessibility. His main claim is that the relevant notion must be intra-discursive: in his framework, it must be a salient "frame" (Fillmore 1982, 111). In Heim (1982)'s and Kamp (1981)'s theories, either a "file" or a "discourse referent" can be salient or not. Hawkins (1978) argues for a similar proposal in terms of speech act theory: by saying an NP, the speaker performs three speech acts, which are (i) introducing a referent to the addressee, (ii) instructing the addressee to locate the referent in a set of shared objects/entities and (iii) referring to the maximal set that this referring expression delimits. In many other theories, a "shared set" is proposed (Hawkins 1991 "P-set", Sperber & Wilson 1986 "mutual cognitive environment", among others). However, being part of the "shared knowledge" is not enough for being considered high salient. Lyons (1999) proposes to incorporate a notion of "Familiarity", according to which the addressee also knows which referent is being picked up. Lambrecht (1994, 77-78) conveys this idea in terms of "shared representation in the speaker's and the audience's mind at the time of an utterance", which he calls "identifiability". All these theories are relevant to the choice of the anaphoric expression, but they are still too rough to influence the distribution of null and overt pronouns. For instance, identifiable, familiar, shared discourse referents can be referred to by definite NPs, NPs with demonstratives, overt personal pronouns and null pronouns in virtually any language in which these constructions are allowed.

Ariel (1990) comes up with the idea that, beyond being identifiable, familiar and shared, a given referent can be more or less accessible. Based on much literature about anaphora resolution (Li & Thompson 1976, Givón 1976, Grosz & Snider 1986, Sandord & Garrod 1981, Gundel 1988, *inter alia*), she proposes that many factors can influence the Accessibility of an antecedent. In what follows, the syntactic and discourse variables taken to influence antecedent Accessibility are summarized. This is not intended to be exhaustive, but rather to give an overview of some factors that can be taken into consideration when computing the Accessibility of an antecedent. To the extent that they make similar predictions, various different proposals are collapsed here and minor

differences among them are omitted for the sake of readability and conciseness. For the purposes of this dissertation, I abstract away from various theoretical standpoints (whether anaphora resolution is analyzed at the cognitive level or in terms of computational treatment, for instance), and only the empirical linguistic predictions are taken into consideration.

Assumed Familiarity and Givenness

As briefly summarized above, a crucial factor in the salience of a referent is how familiar, known or part of the shared ground it is (see, for instance, Lambrecht 1994, for an overview). Prince (1981), for example, takes Familiarity to influence the salience of the antecedent. In her account, "Assumed Familiarity" is a by-product of the notion of "Givenness", *i. e.*, the relations between the parts of an utterance that are "new" and the parts that are "given/old" (see Halliday 1967, Halliday & Hasan 1976, Clark & Haviland 1977, Kuno 1972, 1978, among others). In brief, a given referent can be (non-)anchored "brand-new" (when it has not been previously introduced into the discourse), "inferable" (when the audience can build a possible mental representation of it) and "evoked" (when it is in the discourse or within the audience's perceptual environment). Gundel et al. (1993) further develop this idea by proposing a direct mapping between a "Givenness Hierarchy" and (co-)referential expressions in English. This hierarchy is summarized in Figure 2.5 below. According to this hierarchy, cf. Table 2.6 below, a given form is only appropriate when its antecedent fits the cognitive status described by the discourse-semantic level.

Figure 2.5 – Gundel et al. (1993)'s Givenness Hierarchy
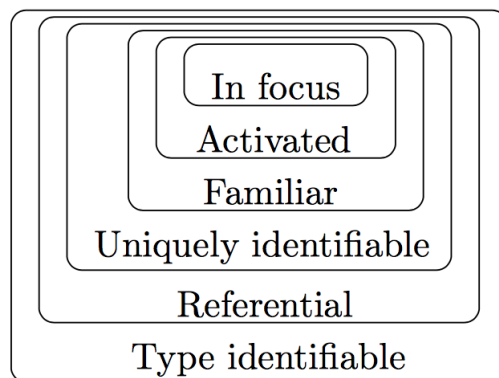
(adapted from Williams & Scheutz 2017)

163

Form	Description	Level
It	associate representation in focus of attention	(in focus)
this/that/this N	associate representation in working memory	(activated)
that N	associate representation in memory	(familiar)
the N	associate unique representation with NP	(uniq. identifiable)
Indefinite this N	associate unique representation	(referential)
a N	associate type representation	(type identifiable)

Table 2.6 – Correspondence Form-Level in the Givenness Hierarchy

Prince (1981)'s and Gundel et al. (1993)'s proposal are very similar, except for the fact that the Givenness Hierarchy is an inheritance hierarchy, in the sense that each higher level inherits all the properties of the lower level.

Under a Romance language based implementation of the scale of Complexity and Explicitness, for independent reasons, Luján (1985, 1986) proposes that null subjects in Italian and Spanish equal non-stressed pronouns in English (see also Cardinaletti & Starke 1994, Carminati 2002, for similar ideas). In the above hierarchy, null subjects in pro-drop languages would be at the very top of the spectrum, in the position of English *it*. However, as shown in the contrast between the English example given by Gundel (2010) and the structurally and discursively similar BP example, neither the "Givenness Hierarchy" nor "Assumed Familiarity" are fine-grained enough to explain the choice between BP null or overt subjects. Since both approaches are designed to deal with English (which does not allow null subjects), there is no place for a distinction between pronominal and null anaphors, at the top of the scale of these hierarchies.

(87) A restudy of pareiasaurs$_1$ reveals that they$_1$ are the nearest relatives of turtles.

(Gundel 2010, ex. 2)

(88) A mãe$_2$ do João$_1$ acha   [que *$_1$/ele$_1$ é bonito].
      The mother of.the João think.PRS.3SG that   he be.PRS.3SG handsome
      "João$_1$'s mother$_2$ thinks he$_1$ is handsome."

(Ferreira 2000, 35, ex. 43b)

164

Topicality, Distance and Competition

Givón (1976, 1983) proposes that the choice among different (co-)referential expressions is linked to the degree of "Topicality" of the referent of this expression in a given discourse. He mentions three criteria that can affect the Topicality of a referent: (i) Referential Distance or Recency, that is, the linear distance between the (co-)referential expression and the previous mention of the same referent; (ii) Persistence, or the maintenance of a given entity in subsequent parts of the discourse; and (iii) Potential Interference, *i. e.*, the number of potential referents that can be taken as antecedents for a (co-)referential expression. In his theory, these factors affect the degree of Topicality of a given referent on a scale that is statistically correlated to the options available in the language to (co-)refer to this entity. Speakers use this information to decide the role that they want to give to this referent in the upcoming discourse. Givón (1983)'s theory makes interesting predictions about the distribution of null and overt subjects in BP, as shown below.

First, Referential Distance seems to be a factor that influences the use of null and overt subjects in BP. In previous corpus studies, Barbosa et al. (2005) show that linear distance is more relevant in BP than in EP as regards the use of null and overt subjects, as Table 2.7 shows (although the authors do not stress this point):

Antecedent	BP	EP
Subject of the matrix clause	5/23 (22%)	1/40 (3%)
Subject of the previous adjacent sentence	20/48 (42%)	6/55 (11%)
Non-Subject in the previous adjacent sentence	13/23 (57%)	8/24 (33%)
Subject of a previous, non-adjacent sentence.	21/28 (75%)	8/28 (29%)

Table 2.7 – Overt subjects according to structural context

While in EP there seems to be an effect of the antecedent's syntactic function in the last two patterns, in BP the main effect in the distribution is clearly "Referential Distance": in the immediate preceding context (matrix clause), only a quarter of the subjects were overt; in the intermediate distal contexts, around a half of the subjects were overt (with a slight preference for null subjects when the antecedent is the subject, rather than when it is in another syntactic function); and in the most distant context, three quarters of the subjects were overt.

I am not aware of any research that has extensively studied persistence in BP or in any other other null-subject language. This criterion is partially analyzed in the corpus research presented in

section 2.7. Occasional mentions of the notions of "topic maintenance" as opposed to "topic shift" are made by Luján (1985, 1986). However, in these notions not only "persistence" is at stake, but also some kind of "potential interference", as shown in sequences (89) and (90) below.

(89) a. A – (...) vocês$_{1+2}$ têm   planos pro   futuro?
            you.PL   have.PRS.3PL plans   for.the future
      "Do$_{1+2}$ you have plans for the future?"

b. B – não. A gente$_{1+2}$ tem   e   _$_{1+2}$ não tem,   né,   quer dizer,
         No   the people   have.PRS.3SG and   not have.PRS.3SG (DM) (DM) (DM)
      a   gente$_{1+2}$ tinha,   aí de   uns tempos pra cá   (...) Ele$_2$ antes
      the people   have.PSR.IMP.3SG, then from a   time   to here   he   before
      queria   casar,   eu$_1$ não queria.   Agora eu$_1$
      want.PST.IMP.3SG marry.INF I   not want.PST.IMP.1SG now   I
      quero,   ele$_2$ não quer,   _ entendeu   .
      want.PRS.1SG, he   not want.PRS.3SG,   understand.PST.3SG

      "No, we$_{1+2}$ do and we$_{1+2}$ don't, that is, we$_{1+2}$ did, but then, from some time now (...)

      Before, he$_2$ wanted to marry, I$_1$ didn't. Now, I$_1$ do, he$_2$ doesn't, did you understand?"

                                                        (NURC-RJ, "Inquiry_ac_03")

(90) a. A - E   ... bem, você   falou   que o, você   tem   um irmão
            And   well you.SG say.PST.3SG that the you.SG have.PRS.3SG a   brother
      casado$_1$. Ele$_1$ tem   filhos$_2$?
      married he   have.PRS.3SG children
      "And, well, you said you have a married brother$_1$. Does he$_1$ have children$_2$?"

b. B – Não. Ele$_1$ tá   casado há   dois anos, _$_1$ é   mais novo
         No   he   be.PRS.3SG married there.be two years   be.PRS.3SG plus young
      do   que eu$_3$ até, _$_1$ tem   vinte ... vinte e   cinco, é,   _$_1$
      of.the that I   even   have.PRS.3SG twenty   twenty and five   (DM)
      tem   vinte e   cinco. Mas a   garota$_4$ é   novinha, ela$_4$ ainda
      have.PRS.3SG twenty and five   But the girl   be.PRS.3SG young   she still
      faz   faculdade, ela$_4$ tem   vinte e   dois eu acho, então, por
      do.PRS.3SG college   she be.PRS.3SG twenty and two I   think thus   for
      enquanto eles$_{1+4}$ não têm   _$_2$. _$_{1+4}$ Não tão   nem pensando em
      while   they   not have.PRS.3PL   not be.PRS.3PL nor thinking on
      filho.
      children

      "No. He$_1$ has been married for two years, he$_1$ is even younger than me$_3$, he$_1$ is twenty...

      twenty-five, yes, he$_1$ is twenty-five. But the girl$_4$ is young, she$_4$ is still an undergrad-

uate, she$_4$ is twenty-two, I think, thus, so far, they$_{1+4}$ don't have [children$_2$]. They$_{1+4}$ are not even thinking about children$_2$."

The two sequences of questions and answers in (89) and (90) make clear that "persistence" and "potential interference" (as well as "topic shift" and "topic maintenance") are two intimately related concepts. In answer (89b), the contrast among the three possible subjects drives the use of null and overt subjects: "we" starts the sequence, but it is opposed (topic-shifted) to "he" and "I", which are clearly opposed to each other in a contrastive alternation. As for "persistence" or "topic maintenance", answer (90b) shows that this can be a factor in BP, but it is also modulated by contrast: the answer starts with a sequence of four clauses about the married brother$_1$, it follows next with a sequence of three clauses about his wife (the girl$_4$), pronominalized twice, and it ends with two clauses about both of them (the married brother and the girl$_{1+4}$), once pronominalized (topic shift) and once null (topic maintenance), even though a plural noun (children$_2$) is a possible competitor for the reference in the last clause.

Givón (1983)'s approach establishes a scale according to which the more confusing, surprising, discontinuous or difficult it is to identify the topic, the more important the quantity of encoded information that the (co-)referential form provides will be. It partially predicts the data in (89) and (90), but possibly further refinements are necessary. These predictions are studied in the corpus research reported in section 2.7 below. It is worth noticing that Ariel (1990, 1994, 2001) incorporates these very same factors, along with others, as predictors in her Accessibility Scale ("Distance" and "Competition"). Almor (1996, 1999, 2000) also incorporates these ideas (in a more complex and refined fashion), when proposing the Information Load Hypothesis [ILH]. The ILH broadly establishes that more processing cost serves additional discourse functions. For the ILH, the cost is a product of the conceptual representation (based on the "semantic" distance between the anaphor and the representation of the antecedent) and of the anaphoric form. The discourse function is related to both identifying the referent and adding new information. The mapping process is given by a direct relation between cost and discourse function: the less specific the representation of the anaphor with respect to the representation of the antecedent is, the less costly the anaphor is to process. The acceptability of a given anaphor in a certain context is not a consequence of its

formal class but rather of the relation between its cost and the discourse function. The implications of the ILH will be further addressed in the present dissertation. At this point, it is important to set up the criteria that could be important for deciding in favor of each analysis. In this respect, the ILH seems to be very similar to Givón (1983)'s and Ariel (1990)'s proposals.

As Competition is very difficult to establish (because the number of sentences before or even after a competitor can be found would be a stipulation), in the corpus research only the Distance and Topic Maintenance properties are quantified and analyzed as predictors for null and overt subjects in BP. In the next subsection, two more predictors are discussed: syntax and position.

Order of Mention, Parallelism and Subjecthood

Gernsbacher & Hargreaves (1988), based on Gernsbacher (1989), propose that first mentioned entities have priority over other posteriorly introduced referents as regards Accessibility. According to their Advantage of First-Mention Hypothesis [AFMH], a coherent discourse is built upon laying a foundation, to which subsequent information is mapped. The first mentioned entity would so be preferred as a laying foundation to the next mentioned entities. In their papers (Corbett & Chang 2005, Von Eckardt & Potter 1985, Gernsbacher & Hargreaves 1988), however, first mention frequently coincided with the subject of the sentence, the topic of discourse or an agentive semantic-role, which are claimed to be independent factors of Salience. For this reason, in a series of experiments, Luegi (2012), Luegi et al. (2014) tested for ordering of mention in contrast to syntactic function of the antecedent in EP and BP. Later in this section, the results found by Luegi (2012), Luegi et al. (2014) are reported, but before this the hypotheses that address the syntactic function of the antecedent are outlined.

In many analyses, couched in different theoretical perspectives that discuss the saliency of antecedents, its syntactic properties are considered a fundamental factor as regards anaphora resolution. One of the strategies for explaining the bias of a certain anaphoric element for antecedents in a given syntactic function is the Parallel Function Hypothesis (Sheldon 1974, Caramazza et al. 1977, Cowan 1980, Smyth 1974). According to this hypothesis, roughly speaking, the antecedent of a certain anaphoric form is generally taken to be the referent whose syntactic function is the same as that of the anaphor. In the case at stake in the present dissertation, anaphoric subjects would prefer to retrieve subject antecedents. For this reason, this hypothesis will be taken to make

similar predictions to those which argue for the primacy of the subject, since they deal with the data presented here in a similar fashion.

Based on different theories and assumptions, much previous research proposes that less informative anaphoric elements are biased toward subject antecedents, since these are more prominent and salient. The idea that there exists a saliency hierarchy of syntactic functions is not uncommon in the linguistic literature and is widely used for different purposes (Keenan & Comrie 1977 for relatives, Jackendoff 1990, for argument realization mapping, Pollard & Sag 1994, for syntactic anaphors and reflexives, *inter alia*). In this section, two hypotheses will be presented: Centering Theory and the Position of the Antecedent Hypothesis.

According to the Centering Theory, certain entities, or "centers" (or discourse referents for Kamp 1981, Kamp & Reyle 1993 or "files" for Heim 1982), in a given utterance are more central than others, and thus they impose constraints on the possible set of next-reference options from which the speaker can choose one (Grosz et al. 1995, Walker & Prince 1996). These entities are ranked according to a syntactic scale of relative salience (SUBJ > DOBJ > IOBJ > OTHER). Centers can be forward-looking (Cf) (those which are available for future (co-)reference) or backward-looking (Cb) (those which refer back to an entity already mentioned) in a given utterance Ui in a discourse D. More salient Cfs are more likely to be retrieved by Cbs and to keep their salient status throughout the discourse. That is, the Cb, which represents the entity that the utterance is about (the topic of Ui) retrieves the highest-ranked element of Cf(Ui-1,D) realized in Ui. The preferred center (Cp), which is the highest-ranked Cf(Ui,D), is likely to be the Cb of the up-coming utterance. Importantly, the main factor taken to influence Salience is the fact that a certain Cf is higher in the syntactic prominence scale.

Building on the Salience Hierarchy proposed by Centering Theory, Carminati (2002) proposes the Position of the Antecedent Hypothesis [PAH], according to which the most salient antecedent is the DP [Determiner Phrase] in the top [Spec, IP] position. According to this hypothesis, null subjects in pro-drop languages (e.g., Italian) will consequently have a strong bias toward the highest c-commanding DP in [spec,IP], while overt subjects will prefer other structural positions. The prominence of the antecedent is thus structurally defined. The idea that the subject is a preferential antecedent for less complex forms is widespread in the anaphora resolution literature. The fact that a less complex form is preferably interpreted as co-referential with a subject antecedent in the

169

preceding sentence or clause was called the "Subject Bias Strategy" (e. g. Crawley et al. 1990, Grober et al. 1990), which Chafe (1976), for example, explains by the fact that the subject position often coincides with the agent semantic role and the topic of the sentence. What Carminati (2002) proposes, however, is to bring together both the literature about anaphora resolution and the classical Generative structural analysis for pro-drop languages by showing that the [spec,IP] position has a prominent status regardless of semantic features (agentiveness, e. g.). In her Experiment 4, Carminati (2002) tested for the pre-verbal dative subject of psych-verbs in Italian, which is taken to be a dative subject (Belletti & Rizzi 1988), in sentences like (91). In this experiment, only one Factor with two levels (null in (91a) vs. overt in (91b)) was taken into consideration. According to her, the sentences were pragmatically disambiguated, that is, the co-referential subject should be interpreted as Giovanni$_1$ in (91) because of the meaning of the second clause.

(91) a. Poiche   a Giovanni$_1$ non piace affatto Enzo$_2$, $\_1$ cerca   di
          Because to Giovanni no   please indeed Enzo   look.for.PRS.3SG of
          evitar-lo$_2$.
          avoid.INF-ACC.MASC.SG

       b. Poiche   a Giovanni$_1$ non piace affatto Enzo$_2$, lui$_1$ cerca   di
          Because to Giovanni no   please indeed Enzo   he look.for.PRS.3SG of
          evitar-lo$_2$.
          avoid.INF-ACC.MASC.SG
          "Because Enzo$_2$ does not please Giovanni$_1$ at all, he$_1$ tries to avoid him$_2$."

(adapted from Carminati 2002, 104)

According to Carminati (2002), participants judged (91a) significantly better than (91b) (1.61 difference; range of 1-5): this result supports the PAH in Italian. Two considerations are indeed relevant here: (i) the PAH is intended to deal only with the intra-sentential level, according to Carminati (2002, 308); and (ii) Carminati (2002) does not disentangle the complex relation between the subject position (spec of IP in her terminology) and the topic-focus relation (positions higher than [spec IP] in her words). Carminati (2002, 308) mentions in her conclusions that focused subjects probably are not retrieved preferentially by a null subject, because null subjects are usually used to recover shared information. These two factors are crucial in the present dissertation, since they seem to strongly influence the choice of the co-referential expression in BP and appear to be tightly related to the discourse structure of the clauses involved in the co-reference (as

shown by Arnold 1999, Kaiser 2011, Miltsakaki 2002, Colonna et al. 2014, Baumann et al. 2014, de la Fuente & Hemforth 2013, de la Fuente et al. 2016, de la Fuente 2016 in cross-linguistic data, including Romance languages).

The PAH, and especially the "division of labor" it implies, was further tested in other Romance languages, such as Spanish (Alonso-Ovalle et al. 2002, Filiaci 2010, Filiaci et al. 2013, de la Fuente & Hemforth 2013, Runner & Ibarra 2016, among others). As in Italian, null subjects in Spanish have a strong bias toward the pre-verbal subject position, but overt pronouns are less biased toward other structural positions. Alonso-Ovalle et al. (2002) showed that, in inter-sentential contexts, overt subjects are used to refer to the preceding subject at about chance levels (~50% of trials). De la Fuente & Hemforth (2013) tested Spanish materials with a sentence structure identical to Carminati (2002)'s Italian materials. In Figure 2.6 below, the interpretation biases of overt and null co-referential subjects in intra-sentencial contexts are plotted (data from Carminati 2002 and de la Fuente & Hemforth 2013).

Figure 2.6 – Italian vs. Spanish antecedent preferences for Co-Referential Null and Overt Subjects
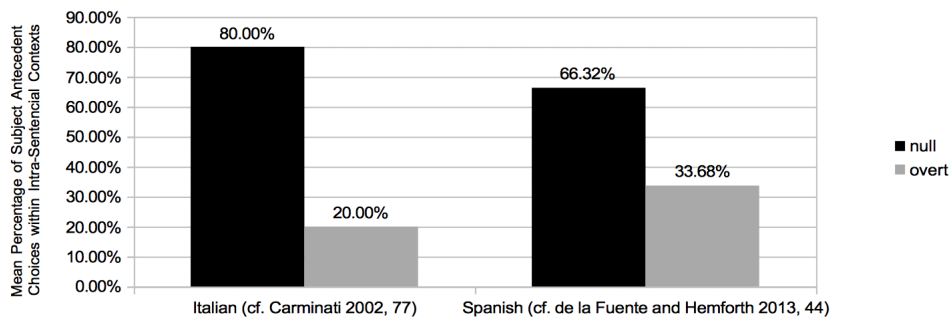
Filiaci (2010), Filiaci et al. (2013) also found a weaker bias for Spanish overt subjects in a series of self-pace reading experiments. While in Italian for overt subject pronouns their experiments replicate Carminati (2002)'s results (the forced co-reference between an overt subject and a subject antecedent yielded a significant processing penalty), in Spanish the same construction did not trigger any significant extra effort in terms of reading times. Filiaci (2010), Filiaci et al. (2013) thus state that, although the PAH is valid for Spanish, the "division of labor" implied by the proposal is not, and the null and overt subjects may each have different (non-complementary) biases.

Another Romance language that has been exhaustively studied as regards the structural prop-

erties of antecedent Salience is Portuguese (both European and Brazilian). Corrêa (1998) tested for the preferential interpretation of null and overt subjects in BP. In her experiment, she had three Factors (one with three levels and two with two levels): (i) syntactic link between clauses (independent in 92a vs. coordinated in 92b vs. subordinate in 92c); (ii) degree of Activation (+ or - activated) and (iii) subject (null vs. overt). After listening to the materials, the participants were asked about the interpretation of the relevant null or pronominal subject (Who questions, whose answers were recorded and, after, quantified). According to her, following the Parallel Function Hypothesis, the interpretation of null subjects should be biased toward the subject antecedent.

(92) a. Pedro$_1$ costuma ler jornal todo o dia. Ele$_1$ e Leonardo$_2$ conversam muito sobre política.

   "Peter$_1$ habitually reads the newspaper everyday. He$_1$ and Leonardo$_2$ talk about politics alot."

   i. Pedro$_{[+activated]1}$ avistou Leonardo$_2$ no escritório. _/Ele contou as últimas novidades e saiu.

   "Peter$_1$ saw Leonardo$_2$ at the office. _/He gave the latest news and _ went out."

   ii. Leonardo$_{[-activated]2}$ avistou Pedro$_1$ no escritório. _/Ele contou as últimas novidades e saiu.

   "Peter$_1$ saw Leonardo$_2$ at the office. _/He gave the latest news and _ went out."

   b. Emília$_1$ gosta de fazer exercícios. Ela$_1$ e Cristina$_2$ caminham na Lagoa.

   "Emília$_1$ likes working out. She and Cristina$_2$ walk along the Lake."

   i. Emília$_{[+activated]1}$ chamou Cristina$_2$ e _/ela atravessou o sinal vermelho.

   "Emília$_1$ called Cristina$_2$ and _/she crossed the red light."

   ii. Cristina$_{[-activated]2}$ chamou Emília$_1$ e _/ela atravessou o sinal vermelho.

   "Cristina$_2$ called Emília$_1$ and _/she crossed the red light."

   c. Alexandre$_1$ sempre estudou com interesse. Ele$_1$ e Rodrigo$_2$ vão todos os dias à faculdade.

   "Alexandre$_1$ always studied with interest. He and Rodrigo go everyday to the college."

   i. Alexandre$_{[+activated]1}$ viu Rodrigo$_2$ nos pilotis quando _/ele mostrou o novo calendário escolar.

172

"Alexandre$_1$ saw Rodrigo at the stilts when _/he showed the new academic schedule."

ii. Rodrigo$_{\text{[-activated]}2}$ viu Alexandre$_1$ nos pilotis quando _/ele mostrou o novo calendário escolar.

"Alexandre$_1$ saw Rodrigo$_2$ at the stilts when _/he showed the new academic schedule."

(Corrêa 1998)

As summed up in Table 2.8 below, Corrêa (1998) shows that null subjects have a strong bias toward the subject of the preceding sentence, specially when the antecedent is activated somewhere else in the discourse (as in Givón 1983, Carminati 2002, *inter alia*). Her results also suggest that the inter- and intra-sentential context is relevant for the choice of the antecedent (Miltsakaki 2002, Alonso-Ovalle et al. 2002, de la Fuente 2016).

	[+ activated]			[− activated]		
	Independent	Coordinated	Temporal	Independent	Coordinated	Temporal
null	98%	92%	96%	81%	83%	78%
overt	92%	40%	40%	67%	33%	38%

Table 2.8 – Mean Percentage of Subject Retrieval Corrêa (1998)

Other subsequent research has also demonstrated the preference of null subjects for subject antecedents in the preceding clause (Costa et al. 1998, Costa 2003, Costa et al. 2004, Melo & Maia 2005, Madeira et al. 2010, Luegi 2012), but in some cases dis-confirms the preference of the overt subject for other non-subject antecedents in Portuguese (in general, EP shows a stronger bias than BP; in particular, for the overt pronouns the difference between varieties is stronger). It thus seems that subjecthood does play a role in anaphora resolution cross-linguistically. This criterion is further analyzed in the corpus research presented in section 2.7 and, given the results obtained by Corrêa (1998), will be contrasted with other factors, such as Distance and Continuity (or "Activation"). As pointed out before, the Parallel Function Hypothesis makes similar predictions about the data presented here, since the scope of the present chapter is the co-referential subject. Lastly, "order of mention" could also be analyzed. However, this factor was studied by Luegi (2012) who suggests that it is not relevant in BP.

As a starting point, Luegi (2012) proposes that EP and BP have similar anaphoric options (null and overt subjects), which either show a "division of labor" as proposed by Carminati (2002), Alonso-Ovalle et al. (2002), *inter alia*, or, in BP, are distinct, given the strict syntactic requirements of null subjects (c-commanding antecedent, as proposed by Guesser 2007, 2008, among others) and the deactivation of the "Avoid Pronoun Principle". Luegi (2012) investigates some factors that influences the Salience of the antecedent by proposing three hypotheses: (i) the syntactic position salience hypothesis, according to which null subjects are biased toward the subject antecedent, while overt subjects are biased toward other syntactic functions, following Centering Theory and the PAH (Grosz & Snider 1986, Grosz et al. 1995, Carminati 2002, *inter alia*); (ii) the structural position hypothesis, according to which null subjects preferentially retrieve the highest entity in the syntactic structure, regardless of syntactic function, and the overt subjects prefer the others, following the AFMH (Gernsbacher & Hargreaves 1988, Gernsbacher 1989, *inter alia*); and (iii) the multi-factorial hypothesis, which is thought to be the interaction between the two previous hypotheses (Kaiser 2011, and colleagues).

To test these hypotheses, Luegi (2012) used sentences like those in (93) below. In these sentences, the interpretation of the anaphoric subject in a *when*-subordinate clause was manipulated by the gender of the participle at the end of the clause. Factors were the ordering of the sentence (SVO in 93a and OVS 93b), null or overt subjects (variation in 93i vs. 93ii) and forced co-reference (by varying the participle gender).

(93) a. i. O   João conversou   com a   Cláudia quando _ foi
            The João talk.PST.3SG with the Cláudia when   be.PST.3SG
            internado/a.
            admit.PTCP.MASC/FEM

       ii. O   João conversou   com a   Cláudia quando ele/ela foi
            The João talk.PST.3SG with the Cláudia when   he/she be.PST.3SG
            internado/a.
            admit.PTCP.MASC/FEM

   b. i. Com a   Cláudia conversou   o   João quando _ foi
            With the Cláudia talk.PST.3SG the João when   be.PST.3SG
            internado/a.
            admit.PTCP.MASC/FEM

       ii. Com a   Cláudia conversou   o   João quando ele/ela foi
            With the Cláudia talk.PST.3SG the João when   he/she be.PST.3SG

internado/a.
admit.PTCP.MASC/FEM

"João talked to Cláudia when he/she was admitted [in the hospital]."

(Luegi 2012, 117)

In a self-paced reading experiment (Experiment 1 in Luegi 2012), no significant difference across condition was found. In an eye-tracking while reading experiment with similar materials, again no significant difference was found. However, in an experiment with an off-line questionnaire and visual world paradigm (her Experiment 2), differences in the interpretation of the anaphoric subject came up with slightly different materials. In these experiments, Luegi (2012) tested for transitive verbs in the subordinate clause, which do not make explicit the reference of the subject, such as *quando _/ele recebeu a medalha de condecoração* "when he received the honor medal". She found a clear preference for subject antecedents when a null subject is used and an object bias for the overt subject, but these preferences were modulated by the manipulation of word order, favoring the hypothesis of an interaction of factors in EP. Experiment 3 is particularly relevant here, because European and Brazilian Portuguese were contrasted. With ambiguous materials as in her Experiment 2, Luegi (2012) used an off-line questionnaire with twenty-four EP speakers and twenty-four BP speakers. The participants read either an SVO or an OVS main clause whose subordinate clause had either a null or an overt subject. After reading the sentences, they had to answer an interpretation question about the referent of the subordinate clause co-referential subject, which could be either the subject or the oblique of the main clause. The contrasted results are presented in the Table 2.9 below.

	EP	BP
SVO_null	75%	68%
SVO_overt	29%	41%
OVS_null	65%	62%
OVS_overt	46%	49%

Table 2.9 – Mean Percentage of Subject Antecendent Answers from Luegi (2012, 166)

As shown in Table 2.9, in SVO order, EP speakers show a clear tendency to behavior like Italian speakers, splitting the preferences of interpretation. However, these preferences are modulated in the inverted order OVS. BP speakers show a weaker bias across conditions, with crucially no bias

175

at all in the OVS_overt Condition. Luegi (2012) mentions that no significant statistical difference between both conditions with null subject and OVS_overt was found in BP, and this last condition is not significantly different from the SVO order with overt subject. In the discussion of this experiment, Luegi (2012) claims that the only relevant factor in BP seems to be the preference of null subjects for a c-commanding antecedent, while the overt form does not seem to be in complementary distribution with the null form. Given the results and the discussion in Luegi (2012), Order of Mention is set aside for future research, since it does not seem to be a relevant factor in BP anaphoric system.

Teasing Apart Some (Semantic) Criteria

Although some semantic factors such as thematic roles are not studied in detail in the present dissertation, it is worth mentioning that some authors claim that agentivity plays a major role in antecedent prominence and that the subjecthood preference is probably related to a preference for antecedents that are both syntactically and semantically salient (Ferreira 1994, Kaiser 2011, Morgado 2012). Morgado (2012) tested for the effect of semantic role in intra-sentential contexts on subject anaphora resolution in EP by varying the prominence of the patient by a passive structure. In the passive (patient subject) antecedent, the overt subjects showed a strong bias toward the subject. In the Null Condition, both the agent subject and the patient subject were chosen as the preferential antecedent. Although testing these factors goes beyond the goals of the present dissertation, Morgado (2012)'s conclusion may be accommodated in the proposal presented here: null pronouns tend to be sensitive to discourse factors (such as subjecthood), while overt pronouns tend to be affected by other factors, in particular, extra-discursive factors (such as agent-patient relative cognitive prominence). In the present dissertation, semantic roles are not analyzed because in corpus studies it is very difficult to decide which role is attributed to a certain NP given the number of different proposals concerning the existence, number and the types of semantic roles (see Levin & Rappaport-Hovav 2005, and references cited there).

Implicity causality is also thought to be relevant for anaphora resolution cross-linguistically Costa (2003), Costa et al. (2004), de la Fuente (2016). Testing for this factor goes beyond the scope of this dissertation and will be left for future research, because the focus of the present chapter and of this dissertation in general is the morphology-semantic and syntactic-discourse interfaces. The

effects of specific semantic factors on discourse organization would take us a step further in the theory of interfaces pursued here.

Finally, temporal relations are also mentioned in de la Fuente (2016) as a possible factor to be taken into account in anaphora resolution. Again, this factor is set aside for future research.

2.6.2 Maximal Contrast – The Discourse and Syntactic Properties of the Antecedent of the Anaphoric Subjects in Brazilian Portuguese

As reviewed in section 2.6.1 above, many different syntactic and discourse criteria have been proposed in the literature on anaphora resolution. What is generally accepted, however, is that null subjects in Romance languages are likely to retrieve the highest antecedent in the Salience Hierarchy, defined in structural or discourse terms. The factors that play a role in the choice of null and pronominal subjects will be studied and scrutinized in section 2.7 below. Before going over this analysis, the well accepted idea that Salience plays a role in the use of null subjects will be integrated with the Maximal Contrast proposal and fully spelled out.

The discourse articulation between given information and new information is reasonably accepted in the literature on information structure (Halliday 1967, Jackendoff 1972, Halliday & Hasan 1976, Clark & Haviland 1977, Wilson & Sperber 1979, Prince 1986, Lambrecht 1994, Rochemont 1986, Rooth 1985, *inter alia*). This idea is built on the fact that a coherent discourse must be both anchored in what is shared between the speakers and their addressees and what is relevant to the context, that is, it must add new information to what is shared (Stalnaker 1978, Carlson 1983, Rochemont 1986, Roberts 1996, Ginzburg 1996, Stalnaker 2002, Ginzburg 2012, among others). In the present dissertation, this articulation is thought of in terms of Question under Discussion: at a given point of the discourse, an assertion is made in order to (i) be directly related to the context, (ii) take (a portion of) the shared information and (iii) update the context in a relevant way (as proposed in section 2.3.4). Underlying these three principles, it is assumed here that, at each point of a certain discourse, there is an underlying question to be answered by the assertion made by the speaker. This articulation is optimal if it fits the Gricean Maxims of Quantity and Relevance at the highest level. In this sense, an optimal contribution to the context is always Maximally Contrastive: it is directly related to the preceding context and, at the same

177

time, it exhausts the current QUD. As shown in previous sections, (verbal) answers to Yes-No questions and to unary wh-questions are Maximally Contrastive and are thus optimal contributions to the context, because everything else beyond the asserted content is already backgrounded and Obvious. Salience of a given antecedent, in this approach, is thus a property of the backgrounded entities, which will be spelled out in this section.[39] For the sake of exposition and clarity, the term "topic" will be used to refer to backgrounded entities that may anchor an assertion to the previous discourse, although it has no theoretical status in the present dissertation.

The MaxCoNAH (Maximally Contrastive Null Argument Hypothesis) predicts that the higher the discourse contrast between backgrounded and focused/asserted information is, the higher the formal contrast between the expression of what is backgrounded and what is asserted must be. Being a relational definition, the focused, or asserted, content can only be considered highly contrastive if the backgrounded information is present and Obvious. In informational structure terms, an entity which is highly Obvious is thought to have a high degree of "Topicality" and "Activation". In this sense, the Maximal Contrast Hypothesis incorporates many factors ("Familiarity", "Givenness", "Accessibility", "Activation" and "Topicality") in one single definition. Other important factors mentioned in the previous section have clear effects on how Obvious a given element is: (i) distance (the more distant, the less Obvious an antecedent is); (ii) competitors (the more possible antecedents, the less Obvious the backgrounded referent to anchor the new information is) and (iii) being a first-mentioned left-most syntactic subject, a referent tends to be topic, and so more likely to be backgrounded and obvious continuation of a discourse. Surely, "topic change" continuations are not Maximally Contrastive, since they signal to the addressee a switch from the current "file" to another new "file", which can be old or not. In this sense, the underlying QUD of an overt subject clause in BP may be as proposed by Büring (1999, 2003): a set of multiple proposition with different possible topics; that is, different backgrounds.

To explain the predictions and how Salience of the antecedent is incorporated in the present approach, one example is detailed in what follows. Example (94) below is a limited sample, as the

---

[39]   Jeffrey Runner has pointed out "Salience" means that something is very "apparent", "visible", while "backgrounded" means it is not as important. He thus wondered whether there might be a distinction within backgrounded material regarding the degree of salience. Throughout this dissertation, no evidence in favor of this hypothesis, apart from the fact that null subjects prefer highly obvious antecedents, especially when no competition among possible antecedents takes place.

annotation of QUD in corpus is still being developed. Hopefully, in the near future, corpus studies will benefit from the development of QUD-theory and clear criteria will be proposed and possibly quantified.

(94)   a.   A – Você estava falando do problema$_1$ dos professores$_2$...

"A – You were speaking about the problem$_1$ of teachers$_2$..."

b.   B – E ... esse problema$_1$ o sindicato$_3$ não pode resolver ... __$_3$ não pode ...

"B – Yes... this problem$_1$ the union$_3$ cannot solve... it$_3$ can't..."

c.   A – Quem$_4$ é que entra nisso$_1$ então ?

"A – Who$_4$ fits this$_1$ so?"

d.   B – Ninguém$_4$ ...

"B – Nobody$_4$..."

e.   A – E o professor$_5$?

"A – And what about the teachers$_5$?"

f.   B – __$_5$ Foi despedido ... __$_5$ Recebe a indenização a que __$_5$ tiver direito e pronto ... e o problema social$_6$ é criado ...

"B – They$_5$ are fired... They$_5$ receive the indemnization which they$_5$ have the right to ... and the social problem$_6$ is created..."

g.   A – Mas se ele$_5$ quiser criar um problema pro colégio$_7$... acionar o colégio$_7$ ... algum jeito de ganhar (...)

"A – But if they$_5$ want to create a problem for the school$_7$... to sue the school$_7$ ... some way to win (...)"

h.   B – Não há condições ... o colégio$_7$ diz... olha ... o problema é o seguinte ... eu não tenho condições mais de pagar esse professor$_5$ ...

"B – No, there are no conditions... the school$_7$ says... well... the problem is the following... I don't have conditions for paying this teacher$_5$ anymore... "

(NURC-RJ, "Inquiry_164")

This is a dialogue between interviewer A and interviewed B, but it develops more like a natural dialogue than a sociolinguitic interview (possibly because the interviewer and the interviewed are acquainted with each other, as noticeable from the content). In the first turn in (94a), A retrieves the topic of the conversation (already backgrounded), taking the topic to a salient position. The following turn starts with the acceptance of the topic and an assertion by the interviewed, who retrieves the topic with a "this NP" topicalized phrase *esse problema* "this problem", followed by a (new) subject referent, and a second assertion. At this point the assertion has not been backgrounded yet, and so the underlying QUD remains at issue: "as for this problem, can't the union solve it?" And so, a Maximally Contrastive sentence is used: $\_3$ *não pode* "it can't". Noticeably, the antecedent of the null subject is the closest syntactic subject antecedent, and not the topic of the previous sentence. However, in this specific case, the repetition of the modal verb might play a role, favoring the parallel function interpretation. In (94c), the interviewer asks for the referent to which the predicate is applied. But the interviewed asserts that this is an empty set in (94d), by attributing the predicate to the quantified *Ninguém* "Nobody". So, the interviewer introduces a new salient topic in (94e). The interviewed follows the new topic introduction by two Maximally Contrastive assertions in (94f), since the backgrounded entity, which the assertion is about, is not doubtful and is highly obvious, and all the information except this entity is new information in the discourse. At this point, the underlying QUD would be the following: "what happens to the teachers$_5$?" However, at the end of this segment, a new referent is introduced *o problema social$_6$* "the social problem". Being the closest salient antecedent (and possibly because null subjects are likely to retrieve inanimate antecedents), an overt subject is used to co-refer to the previously mentioned antecedent. In other words, there are two possible salient topics. The use of an overt subject is triggered by the fact that it is not highly contrastive, since the referent of the subject is not obvious. Notice that semantically the referent of the subject is clear, since "teachers$_5$" is the only possible referent in the context who can "want to create a problem for the school". So the use of an overt subject in this context cannot be explained by the assumption that the speaker tries to avoid ambiguity. That is rather a complex calculus that involves computing Relevance, Quantity, possible alternative formulations and Economy. Notice also that many factors probably play a role at the same time, since in this case Animacy and Specificity ("teachers$_5$" in this case is non-specific) are also computed. The dialogue follows in (94h) with other forms of (co-)reference.

To sum up, here the MaxCoNAH incorporates Salience in terms of "Obviousness", since it assumes a double articulation between background, which is computed as regards how obvious it is, and asserted content, which is the real "contrastive" part. Finally, the idea of discourse salience must be recapitulated. Ariel (1990) proposes a Salience Hierarchy in which discourse vs. world prominence is ranked, cf. Figure (1.3), repeated here as Figure (2.7).

Figure 2.7 – Ariel (1990)'s Salience of the Antecedent
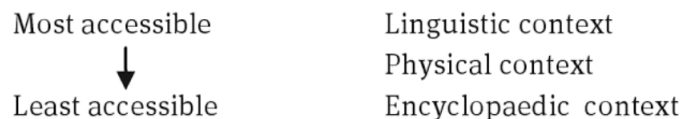
In Ariel (1990)'s proposal, salient antecedents in the linguist context are more accessible than those which are salient in the physical or encyclopaedic context. Here, it is assumed that Obviousness is also defined discursively, rather than in a supposedly cognitive world-related general scale of prominence. What is computed in Obviousness will be analyzed in what follows, according to the discussion of the literature presented in the previous section.

In section 2.7 below, multiple factors that can interfere with Obviousness and Contrast are assessed in a corpus study (once again the NURC-RJ, as in Chapter 1). By using inferential statistics, it is possible to obtain a first idea of the significant factors and how they interact. Most of the criteria used in this analysis have no theoretical status. As in the corpus analysis previously reported in the present dissertation, the decision was made to begin the investigation using the criteria proposed as relevant in previous literature, which can be unambiguously evaluated, including "clause type", "antecedent syntactic position". By focusing on these criteria, the corpus research carried out here avoids commitment to any specific theory and theory dependent criteria, such as "CP-position", "locality", etc. The combination of these criteria with inferential statistic significance can further favor one or another hypothesis or analysis within specific theoretical toolkits.

2.7   Corpus Reassessment

So far in this chapter, vYNAs were taken as a first case in which the information structure could be easily described, manipulated and thus studied as regards the use of null and overt subjects.

However, no language could possibly be restricted to only this context. The use of null and overt subjects must be understood in other contexts as well. For this purpose, in this section, other theory-independent criteria are checked in corpora and quantified. In corpora, however, only indirect evidence in favor of one or another hypothesis can be found. The main reason for the difficulties in finding clear evidence for one or another hypothesis is that clauses taken from corpora are subject to multiple factors acting on the relevant phenomena at the same time, including dialogue properties (repetitions, repairs, e. g.). For this reason, an inferential statistical analysis was carried out. In such a kind of analysis, many factors can be taken into consideration and the weight of each of them can be studied by deducing properties of underlying probability distributions. The aim in reassessing corpus data in such a way is to find characteristics that, being statistically significant, can indirectly favor one or another analysis.

Once more, in this section, the results of reassessing the corpus of oral interviews (NURC-RJ) are reported. As much as possible, in the discussion, these results will be compared to those found by Duarte (1995), Negrão (1990), Barbosa et al. (2005), who also report corpus results though on smaller scales. It was decided to carry out a second analysis on the same set of data (nine interviews carried out in the 70s and nine interviews in the 90s, whose audios are available, six of them given by the same person at two different moments / 8032 inflected clauses in which the subject was either co-referential or arbitrary), for the following reasons: (i) to check whether the principle of Maximal Contrast, proposed in the previous sections, can explain some of the previous findings in the literature and generalize over cases taken to be quantitatively relevant; (ii) to examine the interaction of this constraint with other factors proposed in the literature about anaphora resolution (such as those reported in subsection 2.6.1); and (iii) to open a possible new way of investigating interactions of multiple factors and their influence in grammatical phenomena taken to be categorical.

2.7.1 Methodological Considerations

As in the first corpus study reported in section 1.6, the corpus research carried out here obeyed several coarse-grained, independent steps in order to guarantee the isonomy of the results. The same initial steps concerning selection of the texts, automatic annotation and collection of data

were followed. Finally, the criteria from the previous literature taken into consideration are those which are directly visually accessible, such as "antecedent distance", "fronted element" etc. Once again, these criteria are intended to avoid affiliation to a specific theory and a subjective analysis, although the results provided are more indirect as a consequence. These data were descriptively analyzed in qualitative and quantitative terms. Finally, as in the previous corpus study, the inferential analysis was carried out through a binary statistical regression using the software R, by using a generalized linear model with the optional link "logit".

2.7.2 Analysis

In order to avoid any possible interaction effect with discourse person, only the results for 3rd persons (46,5% of total number of clauses) were analyzed in this section. Occasional mentions of other discourse persons are made in this chapter, especially when they are unexpectedly overt or null (for example, null 1st person singular in vYNAs, which in Chapter 1 is shown to be mostly overt). However, the focus of this part of the corpus research is to check whether there are syntactic-structural and discourse factors that can facilitate the use of null subjects and to understand how such factors can influence their relative probabilistic distribution.

Based on the claim that the Salience of the antecedent is one of the most important factors, four criteria were investigated: Subjecthood, Distance between the antecedent and the anaphoric element, being part of the Main_Clause or not and being Overt or not. All these criteria were taken to be binary, for the sake of analysis (even if Distance is a graded notion). Some practical decisions were thus necessary: (i) an antecedent was considered to be in subject position only when it was to the left side of the verb with which it agrees (independently of the number and type of intervening elements between the subject and the verb); (ii) distance was converted into the feature [+/– previous] clause (this means that when the antecedent was in a main clause followed by an embedded clause with a co-referential subject, it was considered [+ previous]); (iii) an antecedent was taken to be in a Main_Clause if either it is followed by a subordinate clause or by a another Independent or Coordinate Clause, and finally (iv) an antecedent was considered overt when it had explicit overt material.

Based on previous corpus research and on the hypothesis that the discourse structure of the

sentence in which the anaphoric subject appears plays a role in the choice of the anaphoric form, two other criteria were investigated. The position of the anaphoric subject and the type of clause were analyzed: (i) position of the anaphoric subject was analyzed in terms of a binary feature (whether it was the first element in the clause or not), regardless of which kind of material could precede the subject; in the cases in which only a conjunction (subordinating or coordinating) or a(n extra-clausal) discursive marker preceded the anaphoric subject, it was considered the first element of the sentence; (ii) the types of clauses were analyzed using seven levels: Asyndetic (independent clauses without conjunction), Syndetic (coordinate clauses with a conjunction, such as *e* "and", *ou* "or" and *então* "so"), Embedded (starting with a subordinate conjunction *que* "that" or *se* "whether"), Relative (started by relative pronouns *que* "that/which", *o/a qual(is)* "which" – free relatives and standard relatives were collapsed in this corpus investigation, although it was observed that they behave differently regarding null and overt subjects), Comparative (starting with *como* "as" and superiority, equality and inferiority adjectives followed by the comparitive *que* "than" and *qual* "as", as *maior... que* "bigger... than", *tal qual* "such as", *menos... que* "less than", etc), Adverbial (starting with all other subordinate conjunctions (*se* "if", *mesmo se* "even if", etc) and Main_Clause (the clause in a subordinative-subordinate relation with an embedded, a comparative or an adverbial clause).

In the results presented in section 2.7.3 below, generic null subjects and sentential subjects were excluded (2563 clauses, around one third of the corpus was analyzed): the first for potentially having no antecedent and the latter for being difficult to analyze in terms of the antecedent properties. In section 2.7.3 below, rather than stressing the absolute numbers, the focus is on the interaction of factors and thus on the regressions.

2.7.3   Results

As for the statistical analysis, pronominal subjects were coded 1 and null subjects 0, so that positive numbers mean higher probability of an overt subject. By running the outcomes into a generalized linear mixed-effect model with the optional function "logit", the features of the antecedent and the syntactic-discursive structure of the clause in which the anaphoric subject appears were analyzed in terms of relative weight for the use of null and overt subjects. The results are summarized in the

following subsections.

Antecedent

In the first regression, the features of the antecedent (Distance, Subject Antecedent, Main Clause and Overt Antecedent) were included in the models. The first model was run with all four factors. The outcome of the model is summarized in Table 2.10 below.

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	0.91629	0.48305	1.897	0.05784 .
Subj_Ant	1.48160	1.15036	1.288	0.19776
Distance	-2.12026	0.81650	-2.597	0.00941 **
Main_Clause	-1.02165	0.66667	-1.532	0.12540
Overt_Ant	0.24686	0.70415	0.351	0.72591
Subj_Ant:Distance	-1.35377	1.35870	-0.996	0.31907
Subj_Ant:Main_Clause	-0.52895	1.30122	-0.406	0.68438
Distance:Main_Clause	0.65701	1.05804	0.621	0.53462
Subj_Ant:Overt_Ant	-0.50469	1.36573	-0.370	0.71173
Distance:Overt_Ant	0.98937	0.98053	1.009	0.31296
Main_Clause:Overt_Ant	1.07490	0.88767	1.211	0.22593
Subj_Ant:Distance:Main_Clause	0.01569	1.57803	0.373	0.70924
Subj_Ant:Distance:Overt_Ant	0.58349	1.56479	-0.505	0.13234
Subj_Ant:Main_Clause:Overt_Ant	-0.30792	1.54598	-0.199	0.84212
Distance:Main_Clause:Overt_Ant	-1.26826	1.22768	-1.033	0.30158
Subj_Ant:Distance:Main_Clause:Overt_Ant	0.38516	1.80782	0.213	0.83129

Table 2.10 – Generalized Log-linear mixed-effects model for Null and Overt Subjects in the NURC-RJ

As shown in Table 2.10 above, the only main effect that is significant in the model is Distance.

Null-Overt Subject Clause

Two criteria were taken into consideration as regards the syntax and discourse structure of the clause in which the anaphoric subject appears. First, the kind of clause was included in the model. The outcome of the model is summarized in Table 2.11.

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	0.0354	0.1331	0.266	0.79019
Asyndetic	-0.6979	0.1485	-4.699	2.62e-06 ***
Comparative	-1.8812	0.6353	-2.961	0.00307 **
Embedded	0.3231	0.2200	-1.468	0.14198
Main_Clause	0.2669	0.3464	-0.770	0.44107
Relative	1.1686	0.2681	4.359	1.31e-05 ***
Syndetic	-0.5255	0.1774	-2.961	0.00306 **

Table 2.11 – Generalized Log-linear mixed-effects model for Null and Overt Subjects in the NURC-RJ

As shown in Table 2.11, four kinds of clauses are significant in the choice between overt and null subjects (the two coordinate – Asyndetic and Syndetic – Comparative and Relative). Notice, however, that the effect is not uniform: while Comparative, Asyndetic and Syndetic have a negative effect (leading to more null subjects), Relative had a significant positive effect, which means that relative clauses are more likely to have overt subjects. In Chapter 4, syndetic and (object and free) relative clauses are studied in more detail. Comparative clauses, however, are left for future research, as mentioned in the conclusion of the present dissertation, because they seem to be a further step in the theory of Maximal Contrast. Adverbial clauses are left for future research because of the complexity of their syntactic and dscourse structures (see, in this regard, de la Fuente et al. 2016, who showns that some sorts of adverbial clauses are in a c-command relation with their main clauses, while others do not, and that different sorts of semantic and discourse relations between adverbial and main clauses affect anaphora resolution).

In order to run a more comprehensive model, which incorporates some additional effects of the discourse structure, Factor First_Pos[ition] was incorporated to Factor Types of Clause. The results of this second model are summed up in Table 2.12 below.

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	0.904456	0.287460	3.146	0.001653 **
First_Pos	-1.157292	0.327063	-3.538	0.000402 ***
Asyndetic	-0.238460	0.331156	-0.720	0.471474
Comparative	-0.904456	1.040497	-0.869	0.384708
Embedded	-0.211309	0.544903	-0.388	0.698170
Main_Clause	-0.150684	0.516194	-0.292	0.770352
Relative	-0.561881	0.388599	1.446	0.148201
Syndetic	-0.632523	0.371070	-1.705	0.088270 .
First_Pos:Asyndetic_Ant	-0.460346	0.373900	-1.231	0.218247
First_Pos:Comparative_Ant	-1.675922	1.471649	-1.139	0.254784
First_Pos:Embedded_Ant	-0.008762	0.599170	-0.015	0.988333
First_Pos:Main_Clause	-0.001945	0.754037	-0.003	0.97942
First_Pos:Relative	-1.407658	0.917613	-1.534	0.125019
First_Pos:Syndetic	0.133942	0.426207	0.314	0.753320

Table 2.12 – Generalized Log-linear mixed-effects model for Null and Overt Subjects in the NURC-RJ

In Table 2.12, the only significant effect is the First_Pos of the anaphoric subject, which facilitates the use of null subjects. Also, the type of clause Syndetic approached significance.

Interactions

Finally, in this section, some interactions between the syntactic and discourse structure of the antecedent and of the anaphoric subject are investigated. The first model controlled Factor First Position (only anaphoric subjects in the first position were taken into consideration) and two Factors were run: type of sentence and Distance of the antecedent. In this model, again Distance came out significant and Condition Embedded Clause also was considered significant. The interaction between Distance and the type of Embedded clause was also significant. Finally the intercept term (Adverbial clause : antecedent in a non-preceding sentence) was also significant. The outcome of the model is summarized in the Table 2.13 below.

Ultimately, the Factors that came out significant across the board were entered into a model, in order to check whether they modulate each other. Main Factors Distance and First_Pos were individually significant, but the interaction did not approach significance. The intercept term (First position subject whose antecedent is not in the previous clause) was significant, encouraging the use of an overt subject. The outcome of this model is summarized in Table 2.14 below.

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	1.94591	0.47809	4.070	4.70e-05 ***
Preceding	-2.22732	0.50050	-4.450	8.58e-06 ***
Asyndetic	-0.84730	0.51748	-1.637	0.10156
Comparative	-1.94591	1.49284	-1.303	0.19241
Embedded	-1.57819	0.56797	-2.779	0.00546 **
Main_Clause	-0.84730	0.94617	-0.896	0.37052
Relative	0.57982	0.87609	-0.662	0.50809
Syndetic	-0.55962	0.60618	-0.923	0.35591
Preceding:Asyndetic_Ant	-0.17101	0.54346	0.315	0.75301
Preceding:Comparative_Ant	0.03010	1.67510	0.018	0.98566
Preceding:Embedded_Ant	1.23314	0.62771	1.965	0.04947 *
Preceding:Main_Clause	1.25387	1.02111	1.228	0.21946
Preceding:Relative	0.61789	0.92364	0.669	0.50351
Preceding:Syndetic	0.03147	0.63815	0.049	0.96067

Table 2.13 – Generalized Log-linear mixed-effects model for Null and Overt Subjects in the NURC-RJ

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	2.1001	0.3058	6.867	6.57e-12 ***
Distance	-1.6242	0.3259	-4.984	6.23e-07 ***
First_Pos	-1.2872	0.3435	-3.747	0.000179 ***
Distance:First_Pos	-0.2828	0.3675	-0.770	0.441467

Table 2.14 – Generalized Log-linear mixed-effects model for Null and Overt Subjects in the NURC-RJ

2.7.4 Discussion

The results suggest that two Factors are the main predictors of the use of null and overt subjects in BP in the sample analyzed here: Distance and First Position. It might be argued that these are very superficial factors not related to the factors central to this chapter, such as Obviousness, Contrast, and At-Issueness. However, it will be argued here that these results favor the hypothesis that Obviousness and Contrast are the decisive factors in the choice of the anaphoric form in BP. In the approach proposed in the present dissertation these two factors are captured by Maximal Contrast and At-Issueness: not being in the first position of the clause, it is likely that the co-referential subject is not in a Maximally Contrastive context; and not being in the preceding clause, it is probable that the antecedent is not obvious or perhaps even not At-Issue. In highly contrastive and obvious contexts, null subjects are more likely to be used. Surprisingly, no effect of Subjecthood

was found. Moreover, the effect of fronting (turning the subject into a non-first element of the clause) generalizes over the type of clause (taken to be relevant elsewhere in the literature about null subjects in BP, as, for instance, in Duarte 1995). The only relevant significant interaction between main Factors and Type of Clause was the interaction between Embedded Clause and Distance, which clearly has to do with the fact that, in the immediate preceding clause, the subject of the main clause is very likely to be retrieved by a null subject in the embedded clause (notice that this effect disappears once the model incorporates Factor First Position).

In section 2.6.1, a series of factors considered to be cross-linguistically relevant for anaphora resolution were summed up. Some of them probably do not shed much light on the distribution of null and overt subjects in BP, given the results presented in the previous section. For example, subjecthood does not seem to play a relevant role when dealing with large amounts of data (contra previous claims in the literature about anaphora resolution, Corrêa 1998, Carminati 2002, Alonso-Ovalle et al. 2002, among many others). Certainly, there is a general tendency to go on with the same subject across many segments of a certain discourse. This tendency can overlap with the general distribution of null subjects in a given language. However, by itself, it does not seem to be a factor that facilitates the use of null over pronominal subjects. Order of mention also does not seem to play an important role in the use of null and overt subjects in BP. Other criteria, such as Givenness, Familiarity, Topicality and Competition, seem to be more important in BP. These ideas are indirectly favored by the importance of the main predictor Distance. The present approach incorporates this factor in one single notion: Obviousness. In terms of logical representation, Obviousness is a feature of the background on which the discourse structure is built. Recapitulating the fourth clause of the definition of Maximal Contrast, an obvious antecedent is one present in all the possible assertive alternatives generated by a given current QUD, cf. (95d).

(95) Maximal Contrast

Given a proposition α in a context Γ, such that Γ contains a set of shared accepted propositions *P* at a given discursive time-point *t*, the utterance of α is Maximally Contrastive iff:

a. α is asserted with a focused semantic part *x*, so that α_x is its discursive-semantic representation;

b. *y* is open in the *current* QUD q_y, such that it produces a set of alternative assertive

189

propositions Q_y $[\alpha_x \subset Q_y \wedge Q_y \not\subset P$ at $t]$;

c. $\forall \, \alpha_x \, \forall \, \alpha_z \subset Q_y \, \exists q_y[\text{y} = \text{x} \vee \text{z}]$;

d. $\forall \, \alpha_x \, \forall \, \beta_z \, \forall \, \beta_x \subset Q_y \, [\alpha_x = \beta_x \neq \beta_z]$.

When more than one antecedent or a possible intervening NP is in between the antecedent and the anaphor, different possible backgrounds are available. This finding can be further formalized as a general cognitive constraint similar to the ILH, proposed by Almor (1996, 1999, 2000), Almor et al. (2017). In the general perspective assumed in this dissertation, this idea can be formally represented as D-trees in Büring (1999, 2003)'s theory of discourse structure. The main point to be stressed here is that clause (95d) is violated in such a context, thus triggering the use of overt subjects rather than null subjects.

Moreover, there are many ways in which a fronting can violate one of the clauses of the definition of Maximal Contrast in (95). The evaluative adjuncts in Experiment 4, for example, violate it in being non-at-issue. However, simple fronting of at-issue elements can produce different backgrounded contents from which the fronted element picks up the relevant one (again as formally represented in Büring 1999, 2003's theory of D-trees). In this respect, a fronted at-issue adjunct is quite similar to an overt contrastive pronoun which solves the ambiguity between possible antecedents. Among the possible backgrounds that are generated by the current QUD, which in such cases are not monotonic (violating 95d), one is picked up as relevant for focus to be applied to. Again, the prediction is that this triggers the use of overt subjects, and, according to the corpus research carried out here, the prediction holds.

2.8 The Design of the Subject Anaphoric Grammar

In this section, the design of a grammar to account for the use and interpretation of anaphoric null and pronominal subjects in BP is presented. In subsection 2.8.1, the interaction between the semantic features of the antecedent, presented in Chapter 1, and the discourse-based proposal, put forward in the present chapter, is discussed. In subsection 2.8.2, the design of a grammar is outlined. Along with much previous literature, it is proposed that this grammar must be multi-layered, probabilistic and scalar. Finally, in section 2.8.3, it is explained how this grammar is able

to make predictions about the data studied in the present dissertation.

2.8.1   Multiple Constraints

As shown in the previous sections, the use and interpretation of null and pronominal anaphoric subjects is possibly multiply constrained. Consider, for instance, the dialogue below.

(96)   a. A – E   hoje, eh, se te... se tem,   se dá   um nome e   se
                And today   SE   SE have.PRS.3SG SE give.PRS.3SG a   name and SE
            oferecem   atividades diferentes pra crianças de três, quatro anos?
            offer.PRS.3PL activities   different   for children of three four   years
            "A – And nowadays one has, gives [them] a name and offers different activities to three

            or four year old children?"

    b. B – Bom, teoricamente, _ oferecem,   né?
                Well   theoretically   offer.PRS.3PL [DM]
            "B – Well, theoretically they do [= offer different activities to three or four year old

            children]."

(NURC-RJ, "Inquiry_076")

Example (96) above is the prototypical context of a vYNA preceded by a NaIAdj as in Experiment 4. To recap, in that experiment when the verb was preceded by a NaIAdj, and so the answer was not Maximally Contrastive, a null subject was dis-preferred compared to an overt subject. However, in (96b), the possible antecedent of the null subject is non-specific (possibly the SE pronoun). Having a non-specific antecedent, the null subject is preferred, even if it is not in a Maximally Contrastive utterance. Therefore, the constraints that affect the use and interpretation of null and pronominal anaphoric subjects are multiple and probably act at the same time, making one or another form more likely to be used.

The hypothesis that the choice among anaphoric options and their interpretation are multiply constrained phenomena is not unheard of in the literature about both anaphora resolution and functional grammar, as well as in more structure-driven theories in the Generative Grammar tradition. In this last perspective, Reuland (2003, 17), for instance, points out that the linguistic calculus involving anaphoric co-reference can be related to several mechanisms within a given language system. He mentions lexical properties, such as argument structure and variations on it, strictly

syntactic and structural properties, such as Case (in the Generative Grammar sense), conditions on chains, variable binding in logical syntax, value assignment interpretation in the Logical Form, and discourse factors, such as source of the speech act, center of consciousness, degree of informativeness or speaker and audience's expectations. From a psycholinguistic point of view, Almor (2000, 343) also observes that multiple factors can interact when anaphora resolution comes to the scene. These factors include, but possibly are not restricted to, discourse topic (Givón 1976, 1983), informational status (new information vs. given information; Chafe 1976, Deemter 2008), grammatical function (Gordon et al. 1993), syntactic construction (Almor 1996, Carpenter & Just 1977, Cutler & Fodor 1979, Carminati 2002), order of mention (Gordon et al. 1993), the amount of intervening text since the most recent mention (Sandord & Garrod 1981, Givón 1983), and relevant world knowledge (Sandord & Garrod 1981). From a more functional perspective, Ariel (1990), for example, also suggests that there are a number of factors that influence the level of Accessibility of a referent, among which Saliency (comprising Topicality, Subjecthood, etc.), Competition, Distance and Unity (Discourse Cohesion, etc.). What is at stake in the present dissertation is how to understand and to model these multiple constraints and their interaction, and possibly which of them could be brought together in a more general principle.

In this sense, some research engages in the effort of isolating some properties that would be independently relevant. Carminati (2002) show that, in intra-sentential contexts in Italian, the structural position of the antecedent is decisive in the interpretation (and possibly in the use) of null and overt subjects. Following Carminati (2002), but based on data from Catalan, Mayol (2010) argues that cross-linguistically null subjects are more sensitive to syntax than overt subjects (which would be subject to pragmatic factors). However, de la Fuente (2016), based on data from French and mainly Spanish, shows that preferences in anaphora resolution are language specific and result from an interaction of language specific properties and general cognitive principles. Rather than pursuing one or another perspective and trying to understand whether the BP data presented here favor one or another theory, the present dissertation attempts to make sense of both claims, based on data presented here and elsewhere in the literature. In this vein, the fact that a given factor is not significant for anaphora resolution in a given language is not seen as evidence against a general linguistic principle. As shown in section 2.8.2, rather than not being present in a given language, a semantic or pragmatic principle can map into a different structural form by a different interface

constraint in such a language. One well known example in the literature is the mapping between discourse-semantic focus and its realization in different languages. The existence of discourse-semantic focus is probably universally attested (since every propositional system seems to have an articulation between the asserted and the backgrounded content, which turns an idea – *i. e.*, a propositional content – into a material realization whose content can be checked against the world – *i. e.*, an utterance). However, cross-linguistically the realization of semantic-discourse focus into a formal marking is quite variable: in English, it is marked by sentential stress (Selkirk 1984, Rochemont 1986, Truckenbrodt 1995, *inter alia*); in Catalan and Hungarian, it seems to be realized by (relative) position (see Vallduví 1990, É. Kiss 2008, among others); in French, it appears to be related to specific syntactic constructions (see Lambrecht 1994, *inter alia*); and in languages such as Chickasaw, a (Western) Muskogean language, it seems to be morphologically marked (Munro & Willmond 1994, Gordon in press; see Büring 2009 for a more complete typology). The use and interpretation of different anaphoric options seems to be a similar phenomenon: although the discourse and semantic notions of prominence may possibly be cross-linguitically relevant, the mapping of referents to anaphoric forms is likely to be language-specific. This idea will be advanced in the next section, but, before going over it, the convergence of the different discourse and semantic ideas of prominence presented so far must be spelled out in more detail.

Although Ariel (1990)'s idea of Salience relies on multiple factors, one observation seems crucial for the choice of the co-referential expression of the subject in BP: the most accessible antecedents are those which are easily recoverable from the linguistic context. In the present dissertation, four independent factors have been taken to be relevant to the use of null subjects in contrast with the choices of overt pronominal subjects: Specificity, Animacy, Contrast and Obviousness. The idea here is that these factors are the relevant constraints in BP and realizations of a cross-linguistic observation: null subjects are more likely to retrieve discourse prominent antecedents, because such antecedents are obviously part of the shared common ground. This claim captures the observation that null subjects are restricted by language internal properties of Salience (such as Subjecthood, Distance from the antecedent, Topicality, C-command, At-issueness, Contrast, etc) while overt pronominal subjects are more free in terms of co-reference and biased by extra-discourse factors. This also explain why 1st and 2nd persons are more likely to be retrieved by overt subjects in BP, since they deictically refer to entities in the world. Additionally, the same

principle accounts for the "Nobody antecedents" and "Sentential Antecedents": these entities only acquire reference within the discourse, so they are discursively salient; given this property, they are more likely to be retrieved by null subjects (and to look like a bound variable). Finally, this is the principle behind the other two factors identified in this dissertation: Contrast and Obviousness in the definition of Maximal Contrast are exclusively discourse properties. In this sense, a new formulation of the Maximal Contrast Null Argument Hypothesis is given in (97) below, along with a definition of Discursiveness in (98).

(97)   Maximal Contrast Null Argument Hypothesis [MaxCoNAH] (final version)

Null arguments are more likely to be used co-referentially when they are in utterances that are Discursively Maximally Contrastive (that is, within the discourse – linguistic context); co-referential overt subjects are more likely otherwise.

(98)   Discursiveness definition

A given property is discursive iff it refers exclusively to the shared linguistic common ground which the speaker knows the addressee is aware of.

The MaxCoNAH states that co-referential null argument is more likely to be used in Discursively Maximally Contrastive utterances.[40] According to the definition of Discursiveness in (98), a property is discursively defined if and only if this property is established within the linguistic common ground about which the speaker acknowledges the addressee's awareness. However, the MaxCoNAH in (97) does not address example (96b), where the sentence in which the null subject appears is no longer Maximally Contrastive. To account for examples such as this, a second hypothesis (the Non-Maximally Contrastive Null Arguments Interpretation Hypothesis) is proposed in (99) below, along with a definition of "Intra-Discursively Defined" in (100).

---

[40]   In this definition, the likelihood of only overt subjects is stated with respect to null arguments. Null objects seem to be some degrees lower in the scale of contrast in BP. Overt and null objects will be addressed in future work.

(99)   Non-Maximally Contrastive Null Arguments Interpretation Hypothesis [Non-MaxCoNAIH]
(first version)

If a null argument is in a non-Maximally Contrastive utterance, its interpretation is intra-discursively defined.

(100)   Intra-Discursively Defined

An interpretation is intra-discursively defined iff it defines a hypothetical set of possible worlds without recourse to the common ground, to the physical world or to world knowledge.

In (96b), the interpretation of the null subject does not come from the common ground, from the physical context or from world knowledge. In fact, the interpretation attributed to the null subject is exclusively established via discourse: the group of people who offer different activities is undefined, but is in a hypothetical set of possible worlds, where, if they are in "nowadays" times, it's true that they offer different activities to children. This interpretation is non-specific and, in Chapter 3, will be further explored as regards the generic interpretation of the null subject in BP. So far, the discussion is kept to co-referential subjects and the mapping constraints to their formal realization either as null or as overt.

2.8.2   Scales and Probabilistic Mapping

As discussed in section 1.8, Animacy and Specificity are not taken to be binary features. It seems that multiple degrees of Animacy and Specificity are possible in a language, and in BP this seems to influence the degree of acceptability of null and overt subjects and their bias toward different antecedents. The properties discussed in this chapter – Contrast and Obviousness – also seem to be scalar rather than binary. This is evident, for example, in the definitions of Maximal Contrast and At-Issueness, where the former is a restricted version of the latter. In this vein, a scale of Informativeness is proposed in (101) below.

(101)   Informativeness Hierarchy

| Inferentially | Partially | Completely | Maximally |
| Related | Non-at-Issue | At-Issue | Contrastive |

likelihood of null arguments

——>

-related to the current QUD   +related to the current QUD

[-inf] <——————————————————————————————————> [+inf]
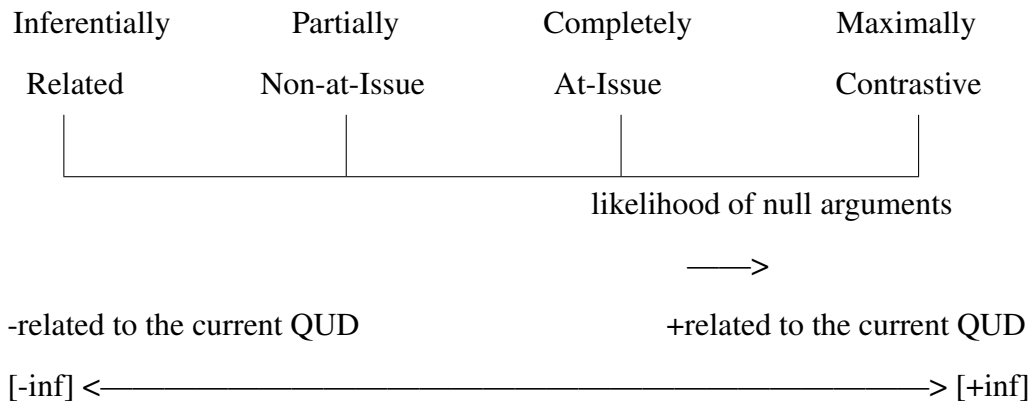
In the Informativeness Hierarchy (101), different degrees of informativeness as regards the current QUD are summarized. Maximally Contrastive is the highest degree. If addressing the QUD, a given content can be still Completely At-Issue without being Maximally Contrastive. One degree lower, a given content can be Partially Non-at-Issue, but still related to the current QUD (such as the answers with a Non-at-Issue evaluative adjunct in Experiment 4). As the lowest degree of the scale, a content can be only inferentially related to the current QUD (also the lowest degree according to Urbano et al. 1993). As in the Referential Hierarchy in Chapter 1, two scales are collapsed in this hierarchy: Contrast and Obviousness. In intuitive terms, the Informativeness Hierarchy in (101) represents the idea developed throughout this chapter. However, as for the mapping from semantics and discourse to anaphoric options, each of the factors seems to individually exert relevant effects on the choice between null and pronominal subjects in BP. For this reason, it is proposed here that there are at least four independent scales for the mapping between the discourse-semantic properties of a given referent and its realization as an anaphoric option. This mapping is summarized in the Figure 2.8 below.

Figure 2.8 – A Mapping Constraint for Null Subjects in A Multiply Constrained Probabilistic Scalar Grammar

"Obvious/Unique"   Semantic-pragmatic Scales   "Maximally Contrastive"

Accessibility   Contrast

Animacy   Specificity

Probabilistic Interface Constraints

"Sentential"

"Nobody"
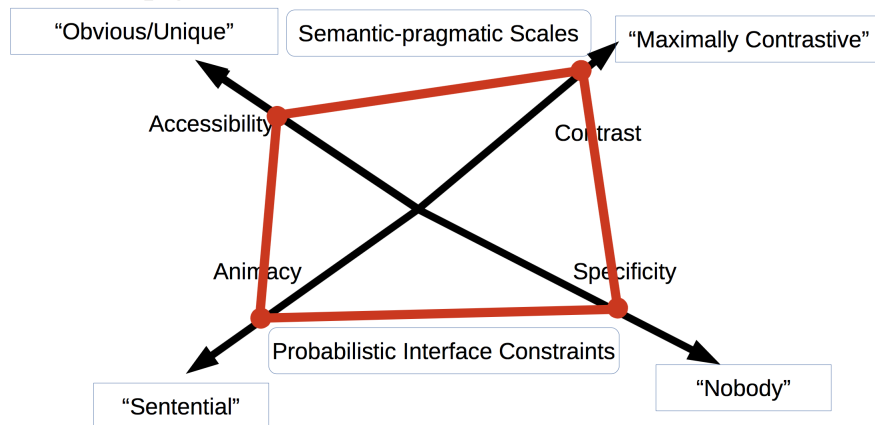
In Figure 2.8 above, each scale is one axis. At a given point on a scale, a referent has a probability of being mapped into a null subject. This probability increases as the referent is more exocentric on the scales and decreases as the referent is more endocentric. Taking the most exocentric point of each scale, the referent is in the optimal context to be mapped into a null subject and an overt subject would be considered completely unacceptable. At the most endocentric point of each scale, a null subject is absolutely ruled out and an overt subject is strongly preferred. All other points over the scale are intermediate cases, in which the probability of a referent mapping into a null subject is between 0.01 and 0.99. Since the ultimate mapping is given by the interaction of factors, representatively the likelihood of null subject is given by the area of the red square in the Figure 2.8.

2.8.3   How does this grammar predict?

A Multiply Constrained Scalar Grammar was proposed in subsection 2.8.2 above. However, being far more flexible than a symbolic system based on simple binary rules (or constraints), the explanatory and predictive power can be called into question. In what follows, it is detailed and explained how a grammar with this design predicts null and pronominal subject data in BP. Moreover, in subsection 2.8.2, the most endocentric point was mentioned, but not explicitly addressed; the most exocentric point was not discussed either. These cases will be discussed in what follows as the cases in which unacceptability is predicted by the grammar design presented here. After this

discussion, intermediate cases are analyzed and an attempt to explain divergent judgments about the data is presented, based on the probabilistic approach. Specifically, the question of gradient acceptability was discussed and implemented in several different ways, such as the Optimality Theory approach of Keller (2000), Hayes (2000), the Generative Grammar approach of Featherston (2005), and the probabilistic approach of Bresnan et al. (2007).[41] The design of the subject anaphoric grammar proposed here is closely related to the last of these, proposed and explored by Bresnan and colleagues in a series of papers about the dative alternation in English (Bresnan 2006, Bresnan et al. 2007, Bresnan & Nikitina 2009). Different ways of computing probabilities are also proposed in the literature (compare for instance the proposal made by Bresnan et al. 2007 with that proposed by Kehler & Rohde 2013*a*,*b*). Details of the mathematical implementation will not be explored in the present dissertation, since the interest here is primarily the linguistic phenomenon.

The most endocentric point of the multiple scales would be (i) the most animate and specific referent, which here is taken to be the 1st person singular, in (ii) the least obvious and contrastive context. A clause with these characteristics is presented in (102) below, in which a null subject is absolutely ruled out, despite being the only pragmatically possible subject of the verb (because it requires an animate antecedent).

---

[41] As pointed out to me by Sergio Menuzzi, most of the differences in mean acceptability judgments found in the experiments reported here, despite being statistically significant, are around 1 point ouf of a scale of 10. This result is expected in the present approach, since each experiment aimed at working out the relevance of one single factor (Animacy, Specificity, Contrast, e. g.), keeping the other factors constant and as good as possible. The statistical significance by itself shows that the judgments are consistent and that a given factor really exert an effect on the acceptability of the relevant sentences, but it by itself does not mean that the sentence turns out to be ungrammatical when violating only one mapping constraint. The combination of multiple violations produce less acceptable sentences, as detailed in what follows.

(102)　　a.　A – o　que　você$_1$ leva　　　　　em conta　　pra escolher　　o lugar$_2$?... ou _$_2$
　　　　　　　　the what you　take.PRS.3SG in　account to　choose.INF a place　　　or
　　　　　　　já　　　é　　　　　　um lugar fixo?
　　　　　　　already be.PRS.3SG a　place fixed
　　　　　　　"A – What do you take into account to choose a place? or is it a fixed location?"

　　　　b.　B – não... _$_2$ normalmente é　　　　　　um lugar que　　eu$_1$/*_$_1$ não
　　　　　　　No...　　normally　　BE.PRS.3SG a　place which I　　　not
　　　　　　　conheça　　　　　　　ainda...
　　　　　　　know.PRS.SUBJ.3SG yet...
　　　　　　　"B – No, it's usually a place that I don't know yet..."

<div align="right">(NURC-RJ, "Inquiry_158")</div>

In example (102), the subject of the relative clause is in the 1st person singular; it is thus at the top of the scale of Animacy and Specificity. The current QUD can be either question made by A in (102a). In terms of Contrast, answer (103b) is only inferentially related to the first question in (102a), not being part of the set of assertive propositions produced by the QUD. Moreover, the subject is in a relative clause, which usually is not used to directly address a QUD (see section 4.5 in Chapter 4). Finally, the antecedent of the subject is distant (to the extent that 1st person singular can have an antecedent), there is a possible intervining referent (*o lugar$_2$*, which is inanimate and non-specific) and the verb is in a tense in which the agreement marking is multiply ambiguous (it could be used for 2nd and 3rd person singular as well as 1st person plural). With so many endocentric properties, a null subject is absolutely unlikely in such a context, and really seems unacceptable.

In statistical terms, when 1st person singular subjects are not preceded by another 1st person singular pronoun and when they are in relative clauses whose verb is ambiguously marked, 98,5% of them are overt (data from the corpus research reported in the present dissertation – NURC-RJ). As the grammar predicts, this is a context in which the mapping to a null argument would be unlikely and dispreferred. For this reason, the null subject in (102b) seems to be unacceptable.

Turning now toward the other extreme of the scales, the most exocentric point would be an inanimate (preferably sentential or eventive) and non-specific (for example, *tudo/nada* "everything/nothing") in a Maximally Contrastive context (e.g. a vYNA or a Unary Wh-answer). Such an example was found in the corpus and is reported in (103) below.

(103)   a.   A – Você$_1$ acha   que _$_2$ foi boato, você$_1$ acha   que não
             You   think.PRS.3SG that it was rumour you   think.PRS.3SG that not

        ocorreu   nada$_3$ assim?
        happen.PST.3SG nothing like.this

        "A – Do you$_1$ think that it$_2$ was a rumour? Do you$_1$ think that nothing$_3$ like this has

        happened?"

   b.   B – Não, _$_3$/ *Ele$_3$/*Ela$_3$/Isso$_2$/$_3$* deve   ter   ocorrido,   né?
             No   he/she/that   must.PRS.3SG have.INF happen.PTCP [DM]

        "B – No, *[something]$_3$/that$_2$ must have happened, right?"

(Adapted from NURC-RJ, "Inquiry_ac_02")

In (103), the antecedent for the null subject in the original example is *nada* "nothing", but since the answer is a positive disagrement, it turns to be interpreted as "something". In context (103), an overt subject with this interpretation is absolutely unacceptable, since the antecedent is at the bottom of the scale of Specificity, Animacy and Obviousness, and the anaphoric subject at the top of the scale of Contrast. The only possibility of using a pronoun in the answer in (103b) is to refer to the previous (less obvious) referent in the first indirect question in (103a) (in the context, it refers to a kind of mass robbery) with a demonstrative pronoun. This possibility also corroborates the design proposed here, because the use of an overt subject retrieves the less obvious antecedent (demonstratives will not be further addressed here, because they seem to have special properties). Statistically, 100% of the cases found in similar contexts were null in the corpus research reported here (9 clauses).

Finally, the issue of acceptability and grammaticality in the non-extreme cases must be addressed. Much research has shown that acceptability is not a matter of binary judgments on data. The literature in psycholinguistics is extensive in this regard (see Schütze & Sprouse 2013, Schütze 2016, and the introductory section in the present dissertation). The variation in acceptability of a given example comes from several sources: dialectal variation, higher or lower sensitivity to certain constraints (such as crossover effects), possible divergent semantic and pragmatic interpretations, frequency of a given construction or combination, which sentences are taken to contrast, individual variation, among many other factors. Here, the factors that are being tested come from an extensive corpus investigation, and their relevance is further tested in psycholinguistic experiments. The results found using both methodologies are evidence for the relevance of such factors, but they

do not constitute binary judgments on the data in terms of acceptability or grammaticality. In this vein, they provide tendencies and indicate the likelihood of whether a given semantic feature, syntactic construction or discourse context exerts effects on the choice of the expression of anaphoric subject in BP. They are thus taken here to represent a probabilistic knowledge that BP speakers have about their options to express a given meaning and how they prefer to use and interpret one or another form. In this sense, the knowledge of the speaker is not limited to *in presentia* elements in the sentence, but also extends to the set of possible alternative formulations in the paradigm, which are excluded by each new word that is being computed during language comprehension and production (as in Baumann et al. 2014, Colonna et al. 2015, de la Fuente et al. 2016). More details and evidence regarding this idea are given in the next chapters.

2.9 Interim Summary

In this chapter, the simplest discourse context in which null and overt subjects appear in BP was studied: namely verbal yes-no answers. It was argued that they are a high relevant context for the theory of acquisition as well as the theory of the use of null and overt subjects. The literature about these answers was presented. A new proposal, which explains the use of null subjects in these answers, was put forward. In this proposal, the fundamental hypothesis was that null subjects are preferred in Maximally Contrastive discourse contexts. So, a corpus investigation of context-types was presented. In this corpus research, however, no context-type that decisively favors the proposed analysis was found. So, an experiment was carried out in order to test (i) whether Maximally Contrastive contexts favor null subjects regardless of possible underlying specific syntactic derivations and (ii) whether null subjects in two different Maximally Contrastive contexts pattern alike, even if one of them is not related to polarity questions. As shown by the results of this experiment, null subjects in verbal yes-no answers are similar to those in answers to verbal, unary wh-questions. In the account proposed here, these two cases were expected to pattern alike, since both are in Maximally Contrastive contexts. And indeed they do.

After the experimental data provided in favor of the Maximal Contrastive Hypothesis, a comparison of this account with previous discourse-based anaphora resolution literature was proposed. A review of the extensive literature on the matter was presented. In this review, some important

factors to be checked in corpora were highlighted. It was shown that the significant factors in the corpus research are tightly related to how Obvious the antecedent of the sentence is. A proposal to integrate all relevant factors found here was also sketched. Fundamentally, the design of the grammar must be multi-layered, scalar and probabilistic. It was shown how such a grammar can account for and predict the use and interpretation of null and overt co-referential subjects. The cases not found in the corpus and intuitively judged to be unacceptable are correctly ruled out, null subjects with lower degrees of acceptability correspond to cases with lower likelihood of null subjects in the grammar sketched, and the contexts with more frequent null subjects correspond to the optimal conditions for null subjects.

In Chapter 3 below, this proposal is extended to a crucial case, which apparently contradicts the proposal put forward in this chapter: generic null subjects. At the end of the chapter, however, it is shown that generic null subjects conform to the general predictions of the approach proposed in this dissertation if a comprehensive semantic description of generic sentences is assumed.

Chapter 3

Generic Null Subjects as a By-Product of Non-optimal Conditions for Referentiality

3.1 Introduction

The aim of this chapter is to discuss the constraints on the interpretation of generic 3rd person singular null subjects in BP. Proposals presented in the literature about BP and also proposals for other languages are presented and analyzed based on corpus examples and experimental data. In this chapter, the focus will be on the contrast between the co-referential interpretation, discussed in previous chapters, and the generic interpretation, which will be presented in what follows.

As shown so far, in most clauses in BP, 3rd person singular null subjects are used to co-refer with another previously mentioned 3rd person singular referent, as in (104) below, instead of an overt pronoun, such as *he* in the English translation provided in the gloss.

(104) … ele₁ ficou   sem   carro, então _₁ desceu   comigo …
          he   get.PST.3SG without car,   so   come.down.PST.3SG with.me
       "he₁ ended up without a car, so he₁ left [the building] with me [to take the bus]"

<div align="right">(NURC-RJ, "Inquiry_114")</div>

On the other hand, in other contexts, such as in (105) below, 3rd person null subjects in BP are used to indicate a generic non-specified referent, which can be broadly translated as the pronoun *one* in English.

(105) Lá   _gen não pode   fazer   isso não.
       There   not can.PST.3SG do.INF that not.
       "There [=in Brasilia] one cannot do that [=construct buildings in a different way]."

<div align="right">(NURC-RJ, "Inquiry_114")</div>

This chapter addresses the descriptions, proposals and theories that deal with sentences in which the generic interpretation of null subjects in BP is allowed and preferential or obligatory, as in (105). This interpretation is particularly challenging for two reasons: (a) this is a very particular feature of present day BP (since generic null subjects are not allowed in any other Romance language nor found in historical data from previous stages of BP) and (ii) it could be seen as counter-evidence to the claim proposed here that null subjects in BP are essentially likely to refer to discourse entities, since generics could be thought to refer to world entities. The starting point here is a literature review on both null subjects and genericity. Consequently, the first goals are

(i) to establish what a generic null subject in BP is, clarifying both the terminology and the concept itself, and (ii) to provide some tests to identify them and compare them with other generic constructions; (iii) to examine the relevance of these data to the general theory of null arguments and to the proposal made in the present dissertation; (iv) to present the analyses in the literature about these phenomena. In the second part of the chapter, the goals are (v) to discuss the proposals made in previous research and check their accuracy on the empirical ground and (vi) to show that the proposal presented in this dissertation accounts for generic null subjects BP straightforwardly based on a general analysis of generic constructions.

Outline of Chapter 3

This chapter is organized as follows. In section 3.2, some terminological discrepancies in the literature that deals with generic arguments are discussed and summarized. In section 3.2.2, some empirical diagnoses for generic null subjects in BP are provided. In section 3.3, some reasons for the study and account of generic null subjects as regards the theory of pro-drop are presented. In section 3.4, the main proposals put forth in the literature on generic null subjects are summarized. At the end of this section, the questions raised in the literature review are presented as regards three main points: (a) the nature of the generic null subject, (b) the constructions in which generic null subjects can appear and (c) the role of generic null subjects as evidence for so-called "partial pro-drop". Since question (a) depends on many intra-theoretical assumptions that are beyond the aims of this chapter, in section 3.5 the question (b) is explored in detail and in section 3.8 question (c) is commented on. The claims presented in the literature are confronted to data from NURC-RJ and experimental evidence in sections 3.5 and 3.7. Between the two empirical sections, in section 3.6, the previous accounts for generic null subjects are summarized. In section 3.8, it is shown how the approach presented here deals straightforwardly with all the observations made in the previous literature and those based on the corpus investigation and on the experimental data.

3.2   Data Under Consideration

In this section, the term "generic" null subject, also called "arbitrary" subject, is terminologically and conceptually clarified, and the related term "impersonal subject" is explained. In the second

part, some tests to identify and delimit the empirical coverage of generic null subjects are provided. These tests show that generic null subjects behave like any other generic construction.

3.2.1 Conceptual and Terminological Considerations

"Generic" subjects, in the sense in which this term is used here, have been taken by the literature to be a sort of impersonal referent. First of all, BP has at least two different types of "impersonal" 3rd person singular null subjects. If the term "impersonal" is taken in a broad sense (more or less as non-co-referential and non-deictic), the subjects that would fit this criterion can be distributed into two coarse-grained categories.[1] As for the first category, it is taken to include those null subjects that have been called non-argumental, or expletive, whose defining feature is the absence of semantic interpretation, as in (106) below.

---

[1]   Differently from previous literature, two other constructions will not be taken into consideration here. The first of these are sentences with post verbal subjects with unaccusative or unergative verbs, such as in (1) and (2) respectively. The other is constituted of sentences with clause extraposition, as in (3).

(1)   Na   casa   da   Maria _ chegou   algumas cartas.
       In.the house of.the Maria   arrive.PST.3SG some   letters

       "In Mary's house some letters arrived."

(2)   Naquele quarto   _ dormiu   várias pessoas.
       In.that   bedroom   sleep.PST.3SG many   people

       "In that bedroom many people slept."

                                                           (Avelar & Cyrino 2008, 68)

(3)   _ Parece   que as crianças comeram   o   bolo.
       seem.PRS.3SG that the children eat.PST.3PL the cake.

       "It seems that the children ate the cake."

                                                              (Duarte 2003, 2)

There is no controversy regarding the fact that the post-verbal NPs and the embedded clause are semantically the subject of the verbs *chegou* "arrive.PST.3SG" and *dormiu* "sleep.PRS.3SG" and of the verb *parece* "seem.PRS.3SG". The issue is whether they are syntactic subjects in any relevant sense of the term, and whether there is an empty subject position. In some theories, these types of sentences are also assumed to have syntactic null subjects. In those theories, the general issue of null subjects would require that they be treated here. However, for the purposes of the present study, they are not relevant, so they are left aside.

(106)   ... da   sala também _ tem   uma porta que   vai   para cozinha
         from.the room also   have.PRS.3SG a   door which go.PRS.3PL to   kitchen
         "... from the room also there is a door which leads to the kitchen ..."

(NURC-RJ, "Inquiry_048")

Note that the subject of the verb *tem* "have.PRS.3SG" is absent and has no interpretation at all, satisfying thus the criterion for being an expletive non-argumental null subject.

The second sort of impersonal 3rd person singular null subjects are those with a generic (or arbitrary) interpretation, which can broadly be described as an entity or group of whom "a general property" is true (following Chierchia 1995). In a sentence such as (107) below, the null subject is interpreted arbitrarily. The meaning of such a sentence could be broadly described as *a predicate of people that is in general true, given the appropriate conditions.*

(107)   compro   uma quantidade grande, a não ser fruta e   legume   que   _gen
         buy.PRS.1SG a   amount   big,   to not be fruit and vegetable which
         tem   que comprar semanalmente
         have.PRS.3SG that buy.INF weekly
         "I buy a big amount, except for fruits and vegetables which one has to buy weekly"

(NURC-RJ, "Inquiry_272")

The subject of the verb *tem* "have.PRS.3SG" in (107), which in this case is a deontic modal, is a generic referent, which does not restrict the property described by the sentence further than *to be true of people*. So the interpretation of sentence (107) can be expressed in quasi-logical terms as in (108) below.[2]

---

[2]   This proposed interpretation is only intended to capture the fact that, when using a generic null subject, the speaker must suppose that there exists at least one individual x such that the property P is true of x and for all or most of x P(x) is equally true. For instance, sentence (1) below would be false if there were no inhabitants in China (for the other quantifiers, it could be vacuously true) or if there are a few x such that x is in China and does not eat scorpions (see the discussion on the semantics of the generic interpretation in Heim 1982, Krifka et al. 1995, Chierchia 1995, Müller 2001, 2002, *inter alia*). This interpretation could also be explained in pragmatic terms: using this sentence in cases in which the denotation of the generic subject would violate principles of relevance (or quantity), and could be claimed to be false.

(1)   Na   China, _gen come   escorpião.
       In.the China   eat.PRS.3SG scorpion.
       "In China, one eats scorpions."

(108)   "it is true that there exists at least one individual x in the domain and for every/most of x,
         if x fulfills its obligations, x buys fruits and vegetables weekly."

The interpretation proposed in (108) resembles the semantics proposed for the "generic oper-
ator" by Krifka et al. (1995), based on Dahl (1975), Nunberg & Pan (1975), Heim (1982), Del-
grande (1987, 1988). The only caveat made here with respect to their proposal is that, at least
for the generic null subject in BP, the domain to which the predicate is applied cannot be empty.
Intuitively, it seems natural that speakers construct generic sentences on the basis of a number of
examples, and not from possible unattested examples. It may indeed be the case that this property
is derived not from the semantics *stricto sensu*, but via pragmatics (e.g. relevance), see note 2
above. This first hunch could be refined with further investigation, but this definition is enough for
the purposes of the present study.

The restriction on the predicate being applied to humans/people is not necessary in order to
correctly interpret example (107), because the predicate itself is suitable only for them (an animal
or a thing cannot buy fruits and vegetables, unless it is humanized). However, one can assume
that such a constraint is inserted in the generic constituent, as for instance Chierchia (1995) does.
However, this does not seem desirable and it is preferable to leave this question open for now.[3] In
fact, in section 3.5 of this chapter, it is shown that there are examples in which the generic subject
can be interpreted as inanimate, although they are not frequent.

This is broadly speaking the semantics assumed here for the generic null subject in BP. As in
other research, it has a sort of dual behavior between an operator and an argument which is able
to saturate a predicate (by closing it existentially) and also to quantify over the entity to which the
predicate can apply (which ultimately restricts the range of such an argument).

In this respect, this description is aligned with Chierchia (1995)'s proposal to deal with *si* in

---

    The above description does not mean that it is assumed here that generic sentences are universally quantified
    (see, for instance, Crone & Frank 2016, in this regard). This has been shown to be false since the XVII cen-
    tury (see Arnauld 1662, part 2, chap. 13). Also, no other quantifier can represent the semantics of generics (as
    shown specially by Carlson 1977). The approximation provided in (108) is good enough for the purposes of the
    present dissertation.
3   Technically, Chierchia (1995) attributes to the assignment function from internal language variables into values
    of the appropriate type, the function of correctly leading to the generic interpretation. According to him, if the
    index of a variable is *arb* ("arbitrary", which broadly means "generic"), the value of the assignment function for
    such a variable belongs to the set of groups of humans (a subset of the domain of individuals).

Italian (which is also a sort of impersonal subject), although there are slight differences between this pronoun and the null generic subject in BP.[4] The technical way the generic interpretation is obtained from the sentences in each theory is a matter that will indeed be discussed throughout this chapter, since it has theoretical and empirical consequences, which need careful considerations.

3.2.2 Empirical Diagnoses for Generic Null Subjects

The literature is extensive regarding tests for genericity. In this subsection, it is argued that generic null subjects are not different from generic NPs, since they pass the same tests. This subsection is based on Krifka et al. (1995), and Müller (2001, 2002), who attempts to apply Krifka et al. (1995)'s tests to BP generic sentences. These authors sum up the possible diagnoses by proposing some clear-cut distinctions to pick out genericity.

a. *The first test*. Although, as in English, there is no explicit marking typically assigned to generic sentences in BP, it is possible to highlight that a sentence is generic by combining it with an adverb such as *frequentemente* "usually" or *tipicamente* "typically", such as in (109) below, adapted from (107) in the previous section.

(109)   compro   uma quantidade grande, a não ser fruta e   legume   que
        buy.PRS.1SG a   amount   big,   to not be fruit and vegetable which
        frequentemente/tipicamente $_{gen}$ tem   que comprar toda   semana
        usually/typically   have.PRS.3SG that buy.INF every week
        "I buy a big amount, but fruits and vegetables which one usually/typically has to buy every week"

                                                                    (NURC-RJ, "Inquiry_272")

In (109), there is a slight difference in meaning as regards the original example, in (107). However, this is expected and shows that indeed no word can explicitly encode the meaning of a generic sentence itself. What *frequentemente* "usually" does, for instance, is to convey the idea that there may be exceptions to the situation expressed by the sentence, but there are instantiations

---

[4]   Chierchia (1995) also mentions the possibility that the subject *pro* in Italian could be analyzed exactly in the same line he proposes for *si*. The differences between them were not considered. Tackling these differences is beyond the scope of this chapter as well.

of such a sentence, and they are frequent.

b. *The second test*. While some predicates can easily be read as generic, others may appear to be generic, but turn out not to be on closer examination. This is the case of "kind predicates" (Krifka et al. 1995, 10; Müller 2001, 289; Müller 2002, section 1), such as in (110) below.

(110)   O   dodo está   extinto.
        the dodo be.PRS.3SG extinct
        "The dodo is extinct."

(Müller 2001, 289)

(110) clearly evokes a taxonomic reading. It is not expected that a generic null subject may be acceptable with a predicate such as "be extinct". In fact, in the corpus study carried out here, no null subject presenting such a taxonomic reading was found. Suppose then that generic null subjects cannot be kind-referring expressions. Observe the contrast that further confirms this claim in (111) below (based on the discussion and examples proposed by Krifka et al. 1995):

(111)   a. No   Brasil, quando pequeno, peixe de rio   é   protegido por lei.
           In.the Brazil when   small   fish of river be.PRS.3SG protected by law
           "In Brazil, when a river fish is small, it is protected by law."

        b. ??No   Brasil, quando pequeno, _gen é   protegido por lei.
           In.the Brazil when   small   be.PRS.3SG protected by law
           "In Brazil, when one is small, one is protected by law."

        c. No   Brasil, quando pequeno, _gen não pode   trabalhar.
           In.the Brazil when   small   not can.PRS.3SG work.
           "In Brazil, when one is small, one cannot work."

(111a) is a clear case of kind-referring expression – *peixe de rio* "river fish" –, which is used with the predicate "to be procted by law", which according to Krifka et al. (1995), is taken to instigate a taxonomic reading. In (111b), the same predicate preceded by a generic null subject produces an unacceptable sentence, which contrasts with an acceptable use of the generic null subject in (111c), a standard example of a generic sentence.

c. *The third test*. This test is used to go further in the distinction between kind-referring expressions and actual generic predicates. According to Krifka et al. (1995) and Müller (2001, 2002), kind-referring expressions must refer to a "well-establish kind". It is possible to observe in

(111a) above that *peixe de rio* "river fish" is a well establish kind, while the generic null subject can refer to any human being in for example (111c). For this reason, it is taken for granted that generic null subjects cannot be kind-expressions.

d. *The forth test*. Having established the difference between kind-referring and generic expressions, it is worth positing a set of criteria to contrast generic and particular sentences. Generic sentences express events that are regular (or were regular in the past), and do not refer to a specific event. According to Krifka et al. (1995) and Müller (2002), this opposition resembles the linguistic distinction between stative and non-stative (also called dynamic) sentences. Müller (2002) also argues that generic sentences are not episodic. Here the less specific description proposed by Krifka et al. (1995) is assumed, who show that the progressive aspect for instance leads a generic sentence to be interpreted as particular, as in (112) below.

(112)   a. Pensando em sair,   etc. etc. aquele negócio, _gen compra   apartamento, _gen
           thinking in leave.INF etc etc that   deal   buy.PRS.3SG apartment
           não compra   apartamento...
           not buy.PRS.3SG apartment
           "Thinking of going out, etc. etc. that deal, one buys an apartment or one doesn't buy an apartment..."

        b. ??Pensando em sair,   etc. etc. aquele negócio, _gen (es)tá   comprando
           thinking   in leave.INF etc etc that   deal   Ø be.PRS.3SG buying
           apartamento, _gen não (es)tá   comprando apartamento...
           apartment   not be.PRS.3SG buying   apartment
           "Thinking of going out, etc. etc. that deal, one is buying an apartment or one isn't buying an apartment..."

        c. Pensando em sair,   etc. etc. aquele negócio, a gente (es)tá   comprando
           thinking in leave.INF etc etc that   deal   the people be.PRS.3SG buying
           apartamento, a gente não (es)tá   comprando apartamento...
           apartment   the people not be.PRS.3SG buying   apartment
           "Thinking of going out, etc. etc. that deal, we are buying an apartment or we aren't buying an apartment..."

(adapted from NURC-RJ, "Inquiry_084")

In (112a), there is a typical context for a null generic subject: a sentence in the present tense – followed by a bare direct object – whose event is supposed to be regular and true of most peo-

ple. When the verbal tense is manipulated, such as in (112b), the generic interpretation for the null subject becomes unacceptable.[5] Finally, in (112c), the subject *a gente* "the people", whose preferential interpretation is the 1st person plural (but is ambiguous between specific and generic reference), with the progressive aspect is clearly interpreted as a particular sentence, and not as generic.

Extending this proposal, Egerland (2003) (based on Dahl 1975, Cinque 1988, 544, Krifka et al. 1995, 49ff and Kratzer 1995, 129-130) points out that to "suspend" perfective aspect is enough to lead to a generic reading. For this reason, modal environments favours such an interpretation: embedded interrogatives, conditionals as well as *quando* "when" with a conditional value, etc. are likely to trigger a generic interpretation. This proposal will be incorporated into the present account as a failure to "ground" the sentence (see Crone & Frank 2016).

e. *The fifth test*. Generic sentences are usually used to refer to non-accidental properties of individuals. Dahl (1975) indeed points out that these predicates are not merely descriptive, but rather they seem "normative". This explains (i) the reason why it is easy to find generic null subjects with modal verbs, present and imperfect tensed verbs, rules and normative sentences and finally conditionals; and (ii) the widespread uses of these constructions to make statements about human beings, since they are more likely to follow norms than things are. However, contrary to what is generally claimed, generic null subjects can be used with "inherent" properties. For instance, sciences-law-like sentences can have generic null subjects in BP, such as in (113) below:

(113)   pra cada prata que sair   de junção... (...) $_{gen}$ tem   que
         for each silver that leave.SUBJ.FUT.3SG of joint   have.PRS.3SG that
         arranjar um junto   com ela, um cloro   pra formar outro   cloreto de prata... OK?
         arrange one together with it   a   chlorine to   form   another chloride of silver Okay?
         "for every silver that comes out of the joint ... (...) it's necessary to get one with it, one

         chloride to form another silver chloride ... Okay?"

                                                              (adapted from NURC-RJ, "Inquiry_251")

The predicate in (113) seems descriptive, rather than normative. In terms of Krifka et al.

---

[5]   An interpretation in which the null subject is co-referential is acceptable, but the sentence turns out to be partic-
       ular and not generic, which reinforces the argument pursued here.

(1995), it would be considered as an "essential/inherent" property, and consequently not classified as generic. However, not being "essential/inherent" is absolutely not a required feature of generic sentences. Often, sentences such as *Water boils at one-hundred degrees.* are interpreted as generic, even if this is a inherent property of the entity. For this reason, the fifth test is set aside, assuming only four as relevant for the purposes of this dissertation.

In order to finish the conceptual characterization of generic null subjects and clearly circumscribe the relevant empirical data, all interpretations that obligatorily exclude the speaker and the addressee, *i. e.* the so-called exclusive reading, are set apart from the subjects under consideration here. This is an important clarification since BP allows the 3rd person plural to have a partially generic, exclusive interpretation which is not under consideration in this investigation, since it is a common construction in virtually all pro-drop languages (see note 10 and Jaeggli 1986, Cabredo-Hofherr 2006, for Spanish; Barbosa 2011, for BP and EP; Malamud 2012, for Italian, Spyropoulos 2002, for Modern Greek; Dalmi 2014, for Hungarian, *inter alia*).

3.3   Importance of Generic Null Subjects for the (Non-)Pro-Drop Character of Brazilian Portuguese

There are several reasons why generic null subjects should be included in the study of pro-drop in BP. In the current literature about the pro-drop parameter, it is generally claimed that languages like BP, which allows null subjects in some but not in all contexts, allow generic null subjects with an inclusive reading, differently from standard pro-drop languages. This is taken to be a distinctive feature of this group of languages that includes BP, the so-called "partial pro-drop languages", to which Finnish, Russian, Modern Hebrew, among others also belong (see, for example, Roberts & Holmberg 2010, Biberauer et al. 2010). Moreover, the generic interpretation could be taken as a counterexample for the generalization proposed in the Chapter 2. However, a closer inspection shows that it perfectly conforms to the theory and analysis proposed here. In this section, historical considerations about the so-called pro-drop parameter, the theory of pro-drop and the role that BP has played in this story are briefly recapitulated.

3.3.1 Partial (Non-)Pro-drop in BP

As discussed in the preceding chapters, there are two general claims about null subjects in BP. The first is that BP is changing regarding the "pro-drop parameter" (Rizzi 1980) and the "Avoid Pronoun" Principle (Chomsky 1981). This parameter was first said to be binary: languages such as English and French do not allow null subjects, and in other languages, such as Italian and Spanish, null subjects are permitted and preferential in topical co-referential positions.[6] The latter allow sentences with unpronounced subjects, because their morphologically rich agreement markers support and license them (see Chapter 1 in this dissertation).[7] Having taken this parameter to be binary, according to Tarallo (1983), Duarte (1993, 1995) and other authors mentioned in the Chapter 1, BP is in the process of changing from a pro-drop language to a non-pro-drop language. This explains why there is a decrease both in the frequency of null pronouns in general and in the number of different contexts in which they are used (Duarte 1993, 1995, Barbosa et al. 2005, Cyrino et al. 2000, and references in Chapter 1). As discussed in Chapter 1, this claim is supported by diachronic data provided by much previous research (see references there). In this context, many authors in the 80s and early 90s proposed analyses in the Government and Binding framework, sharing the idea that a little-*pro* (as in other standard pro-drop languages) is a vestige of previous stages of BP grammar (Duarte 1993, 1995, Galves 1987, *inter alia*).

However, there are many contexts in which null subjects are allowed and some in which they are obligatory (see Figueiredo-Silva 1994, 2000, Holmberg 2005, Muller 2005, Negrão & Müller 1996, *inter alia*). Moreover, they are not so uncommon, being preferential in many contexts in which a standard pro-drop language also accepts a null pronoun. In this perspective, generic null subject sentences are a surprising and challenging case, because they are contexts in which both standard pro-drop languages and non-pro-drop languages resort to an overt strategy to get the

---

[6]   Rizzi (1986), however, proposes that languages can vary regarding the type of pro-drop they allow, by proposing three different types of pro-drop: (i) null pronouns which bears a full theta-role; (ii) null subjects to which a quasi-theta-role is attributed; and (iii) null subjects without theta-role at all. See a comparison of this typology with use of generic null subjects in several languages in Cabredo-Hofherr (2006).

[7]   As mentioned before, in the 1980s, Huang (1982, 1984) already shows that there are other types of languages, which do not fit the binary description, such as Chinese Mandarin. In such languages, the subject is dropped regardless of the morphological complexity of the verb, but purely for discourse-syntactic reasons, as discussed in Chapter 2, when co-indexed with a topic syntactic position. This kind of language also allows generic interpretations for null subjects, given contexts similar to those discussed in section 4.6 in Chapter 4.

generic interpretation (see Roberts & Holmberg 2010, *inter alia*). From where on this diachronic path, thus, has the generic null subject come?

Furthermore, the theory of *pro* has become unacceptable under the framework of Minimalist Program (Chomsky 1995), specially after its own further developments (Chomsky 2000, 2001) and the emergence of Distributed Morphology (Halle & Marantz 1993, 1994, Bonet 1991, Harley & Noyer 1999, Embick & Noyer 2007, *inter alia*).[8] Having assumed post-syntactic insertion or uninterpretability of agreement features, *pro*-licensing is hardly determined by "richness" of the morphology of the verb, which in fact proves to be a suitable conclusion given the data in Chapter 1. Mainly, the relevant features would not be visible at the point of the derivation at which they are required.[9]

In this vein, the analyses of the pro-drop parameter have changed after the introduction of the Minimalist Program, as briefly summarized in Chapter 1. Mainly, null subject sentence licensing is taken to be carried out by specific functional projections, as in the following analyses: (i) the little-*pro* (revised) hypothesis: Holmberg (2003, 2005, 2007), Holmberg et al. (2009) – a kind of functional underspecified category [ϕP] checks the features of Agr and T; (ii) the Agr-feature-checking hypothesis: Barbosa et al. (2005) – the AgrP is itself a kind of pronoun that checks EPP at some level in standard (also called consistent) pro-drop languages. Holmberg (2003, 2005, 2007) fully rejects this second hypothesis, by showing that it would lead to a number of incorrect predictions about partial pro-drop languages (mainly Finnish, but also BP, Hebrew, Marathi, among others). Some researchers then propose a controlled or anaphor-logophoric PRO as a null subject in BP (Kato 1999, Barbosa et al. 2005, *inter alia*) and still others try to deal with it as an instance of null NP anaphora anaphorically related to topic position (Barbosa 2011); (iii) finally,

---

[8] The theoretical problems are that (i) "agreement" is assumed to be an uninterpretable feature, which is unable to license a null constituent, and cannot be interpreted at LF; and (ii) with post-syntactic insertion proposed by DM, in both pro-drop and non-pro-drop languages the fully specified set of syntactic features can be the same, varying only in the form in which they are morphophonologically realized (Halle & Marantz 1993, 1994, Bobaljik 2008, etc).

[9] There are two ways of obtaining reasonable theoretical results for the theory of *pro* in the Minimalist Program: (i) licensing of *pro* is a syntactic phenomenon but independent of the richness of morphological inventories (Grimshaw & Samek-Lodovici 1998, OT syntax in general); and (ii) licensing of *pro* is a post-syntactic (PF) phenomenon (Perlmutter 1971, Adger 2003, Holmberg 2005). I will not be able to discuss this point in this dissertation, but it is worth observing that the implications of the generic null subjects in partial pro-drop languages for these proposals are highly relevant.

the movement-hyperraising-based hypothesis: Ferreira (2000, 2004), Rodrigues (2004) – based upon the analysis of "control as movement", they claim that null subjects are traces of the main clause subjects, derived by movement. Each of these approaches for co-referential null subjects have their respective explanation for the generic interpretation of null subjects, cf. summarized in section 3.4 below.

Most of the recent research claims that BP is a "partial null subject language". (Rodrigues 2004, Holmberg 2005, Barbosa 2011, Roberts & Holmberg 2010, Biberauer et al. 2010, *inter alia*). As proposed in many different articles (Holmberg et al. 2009, Roberts & Holmberg 2010, Barbosa 2011, etc), null-subject generic constructions are a standard feature of partial null subject languages. Cabredo-Hofherr (2006) also underlines the fact that this kind of pro-drop languages (Modern Hebrew, Finnish, Icelandic and Russian) can differ regarding the possibility of having co-referential null subjects, although all of them allow null generic subjects and null subjects with meteorological predicates (differing from non-pro-drop languages).[10]

It is noteworthy that the generic construction is innovative in BP. It emerges at the same time as the impersonal existential construction with the verb *ter* "have" appears (from XIX century on, according to Nunes 1990, Callou & Avelar 2003), when the number of null subjects strongly declines (Duarte 1993, 1995, 2000, 2012). Furthermore, this is not a construction found in previous stages of BP nor a possible construction in current EP either.

The theoretical and empirical problem that generic null subjects pose for all the above-mentioned proposals is the following: since null subjects in embedded sentences are mostly co-referential (the generalization that such analyses attempt to capture), embedded generic constructions require a specific analysis. Compare, for instance, (114) and (115):

---

[10]   There are some differences concerning the languages that are analyzed by Cabredo-Hofherr (2006), which can be relevant. The first noticeable contrast is that some of them requires 3rd person singular verbal marking and others 3rd plural. The second is that the interpretation of null generic subjects can include or not the speaker and the addressee (inclusive vs. exclusive reading) according to the possible interpretations of the generic null subject (see also section 3.4.2 and Holmberg & Phimsawat 2015, 2016). BP has both possibilities and seems to split the interpretations according to the construction. That is, 3rd person singular is used for the inclusive reading, which can include at least the speaker; 3rd person plural triggers the exclusive reading, and does not allow the inclusive reading. Only the first construction is in the scope of the present investigation.

(114)   a   gente$_1$ achou   que _$_1$ tinha   que fazer alguma coisa...
            the people think.PST.3SG that   have.PST.IMP.3SG that do   some   thing
            "we$_1$ though that we$_1$ had to do something..."

<div align="right">(NURC-RJ, "Inquiry_158")</div>

(115)   Aí ela disse   que "mesmo nos   momentos mais sombrios... _ não pode
            So she say.PRS.3SG that even   in.the moments   more dark   not can.PST.3SG
            deixar de acreditar na   luz...
            leave of believe   in.the light
            "So she said that even in the darkest moments one cannot stop believing in the light."

<div align="right">(http://sitedobernardo.blogspot.fr/2011/05/pena-so-de-gente-ignorante.html)</div>

In (114), the null subject of the embedded sentence is interpreted as co-referring to the main clause subject *a gente* "the people[=us]", which is syntactically 3rd person singular. In (115) the embedded clause subject is not co-referential, although it could be if the adverbial phrase *mesmo nos momentos mais sombrios...* "even in the darkest moments" was not present. Rather, it is interpreted as a generic subject, since it passes all four tests in section 3.2.2: (i) adding *frequentemente* "usually" to the sentence would highlight the generic interpretation; (ii) it is not a kind-referring predicate; (iii) it is not a well established type of being and (iv) if the embedded clause verb inflected in present tense is transformed into an episodic inflected verb (for instance, *tinha podido* "had been able to"), the null subject is preferentially interpreted as co-referential and the sentence becomes episodic. For the sake of the analysis, it is taken for granted that the subject of sentence (115) is inclusive, appealing to the interpretation provided by native speakers of BP. The contrast between the preferential interpretations in (114)-(115) raises a problem for analyses whose generalization is that the null subject of embedded clauses is obligatorily co-referential with the subject of the main clause. This is true of (114), but it is not the case of (115), which is clearly a shortcoming for such analyses.

3.3.2   Elsewhere Generics - Strategies of Impersonalization in Typologically Different Languages

It is surprising in the parametric theory that pro-drop and non-pro-drop languages both resort to overt strategies for achieving the generic subject interpretation. It is also surprising that the so-

<div align="center">217</div>

called "partial" pro-drop languages can have the generic interpretation for null subjects (Biberauer et al. 2010, Holmberg et al. 2009, Holmberg & Phimsawat 2015, 2016). In what follows, some previous studies about generic subjects in typologically different languages are presented. Especial attention is given to the analysis that provide data which are different from what is observed in BP. Unfortunately, however, experimental studies on this topic are scarce. Therefore, although the next four experiments reported in this dissertation open a new path of investigation in psycholinguistics and anaphora resolution, there is unfortunately not much literature to compare their results with.

Possibly every language reported in previous research has more than one construction that expresses the generic interpretation for subjects. English, for instance, resorts to overt strategies in finite sentences, such as *one* in (116) and *you* in (117), but in infinitive clauses can express the generic meaning by a null pronoun, as in (118b).

(116)   One can see the picture from the entrance.

(Moltmann 2010)

(117)   On a day like this it's important [that you get enough water].

(Holmberg & Phimsawat 2016)

(118)   a. It is nice when one/*_ is walking in the park.

b. It is nice _/*one to walk in the park.

(adapted from Moltmann 2006)

The contrast between (118a) and (118b) obeys general constraints on the distribution of overt and null subjects in English. As a non-null-subject language, in English tensed clauses, overt strategies are used to achieve this interpretation. Other non-pro-drop Germanic languages present similar strategies, especially in finite clause environments, such as *man/du* "one/you" in German (Linthe 2002), *men/jo/do* in Frisian (Hoekstra 2010), *man/du* in Danish (Jensen 2009), among others.

Surprisingly, pro-drop languages also resort to some sort of overt strategy for the generic interpretation. In Spanish, beyond the pronouns in Spanish equivalent to those in English, *tu* "you" and *uno* "one", SE-construction is also available. The choice of the form can strongly differ across

varieties or groups of speakers. Ramírez (2007) shows that, in a corpus study of interviews with English-Spanish bilinguals living in New York and monolingual Spanish native-speakers living in Colombia, the distribution of forms used for the generic interpretation heavily diverges. Her results are summarized in Table 3.1 below (similar results in Caribbean Spanish monolingual speakers and bilingual Puerto-Rican Spanish-English bilinguals are reported by Morales 1995).

Table 3.1 – General distribution of impersonal forms according to group of speakers

Generic Pronoun	Bilinguals	Monolinguals
uno "one"	150 (64.7%)	28 (32.6%)
se SE	71 (30.6%)	58 (67.4%)
tu "you"	11 (4.7%)	0
total	232 (100%)	86 (100%)

(adapted from Ramírez 2007)

Other Romance languages also have overt strategies for the generic interpretation, such as *si* and *uomo* (literally "man") in Italian (Malamud 2012).

Languages that alternate between overt and null strategies to achieve the generic interpretation are also reported in the literature. Icelandic has both an overt generic pronoun, *madur* "one", and a generic null-subject with 3rd person singular agreement on the verb:

(119)   a.   Fyrst beygir   madur til haegri.
             first   turns.3SG one   to right
             "First, one turns to the right."

        b.   Fyrst verdur   _ ad beygja til haegri.
             first   must.3SG   to turn   to right
             "First, one must turn to the right."

(Sigurdsson & Egerland 2009)

Some other languages are unclear about the generic interpretation, although they allow only null subjects to be generic (and 2nd person singular). Russian (and possibly Modern Greek) is such a case (see, for instance, Gast & van der Auwera 2000, Spyropoulos 2002, Rudolf 2014). Although, when the subject required by the construction is nominative, the preferential way to achieve a generic interpretation is the null subject plus 3rd person plural verbal agreement as in

(120a), in dative quirky-subject constructions, the dropped subject is clearly interpreted as generic, as in (120b).

(120)   a. Teper' _ starajutsja prepodavat' anglijskij jazyk   v mladšix klassax.
           now   _ try.3PL   teach.INF   English   language in younger classes
           "Now they're trying to teach English in the lower grades."

        b. Etogo _ ne sleduet   delat'.
           this   not should.IMPS.PRS do.INF
           "One shouldn't do that."

                                                              (Gast & van der Auwera 2000)

BP is different from other Romance languages in the sense that both overt and null generic subjects are allowed. In BP *se* "SE" and *você* "you" alternate with the null generic subject in many contexts, although none of the strategies completely overlap, as briefly mentioned in the next sections. Before going over the facts and analyses of generic subjects in BP, however, one psycholinguistic study about the generic and co-referential interpretation is worth mentioning.

Haendler (2017) studies the effects of referential properties of subjects on object relative clause [ORC] processing in Modern Hebrew (a language taken to be "partial" pro-drop, like BP) and German. In the study on Hebrew, three types of sentences were compared, as shown in (121) below.

(121) Materials for Experiment on Processing Relative Clauses in Hebrew

<div align="center">Material</div>
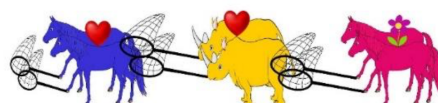
Type of embedded DP	Sentence
	Ma ha-ceva shel ha-susim... what (is) the-color of the-horses
FULL DP	... she-ha-karnafim tofsim (otam)? that-**the-rhinos** catch (them)
NON-REF-PRO	... she- tofsim otam? that-**pro**-catch them
HEM	... she-hem tofsim? that-**they** catch

<div align="right">(Haendler 2017, 14)</div>

The participants read a question with an ORC whose subject was (i) in the first condition, a referential full noun; (ii) in the second condition, a non-referential generic null subject; and (iii), in the third condition, a co-referential pronoun. His off-line results in terms of processing of the relative clause are summarized in Table 3.2 below.

<div align="center">Table 3.2 – General distribution of impersonal forms according to group of speakers</div>

Type of embedded subject in the ORC	Proportion of correct responses (and SD)
FULL DP	97% (7%)
Impersonal pronoun (NON-REF-PRO)	100%
Referential pronoun (HEM)	53% (14%)

<div align="right">Haendler (2017,18)</div>

As shown in Table 3.2, the accuracy of the interpretation in the off-line task suggests that the non-referential null pronoun is more similar to a full referential noun than to a co-referential subject. Haendler (2017) suggests that this can be explained by the fact that the impersonal null subject has less demanding referential properties. Since his sentences were tested without context,

he also claims that, in the absence of an explicit referent, *hem* (co-referential) is more "error-prone" than the impersonal subject.

In order to test the second hypothesis mentioned above, Haendler (2017) carried out a follow-up study on German (a non-pro-drop language). In this study, a context was provided, and the crucial manipulation concerned an alternation between the generic *man* "one" and the co-referential pronoun *er* "he". The materials were like those in (122) below.

(122)   a. Manuela traf Paul auf einer Hochzeit von Freunden.

        "Manuela met Paul at a friends' wedding."

    b. Paul mag bekannterweise schön gekleidete Männer.

        "Paul is known to like nicely dressed men."

    c. Manuela tanzte gerne   mit dem Bräutigam, den *er/man* sehr skeptisch
        Manuela danced willingly with the groom   that he/one very skeptically
        betrachtete, weil   der Ärmel seines Anzugs beschmutzt war.
        observed   because the sleeve of-his suit   dirty   was
        "Manuela danced willingly with the groom that he/someone looked at very skeptically, because the sleeve of his suit was dirty."

                                                         (Haendler 2017)

The experiment was an eye-tracking while reading task, in which the participants had their eye-movements tracked while the sentences were presented on a screen. The generic subject was read significantly faster than the co-referential subject in German. The results suggest that, once again, the generic subject is less complex to process in the ORC environment, than the co-referential subject. These results will be incorporated in the proposal presented in this dissertation in section 3.8 and in Chapter 4. Before going over this comparison, a more detailed study of the generic subject in BP is required.

As mentioned before, some of research that investigates null subjects involve specific proposals concerning the generic interpretation. The gist of these proposals is that the generic interpretation is the result of a specific syntactic construction. That is, sentences in which a generic null subject is used are the outcome of different syntactic combinations whose structure is not the same as that of other null (co-referential) subject constructions. In section 3.4 below, some of these proposals are briefly presented.

3.4 A Brief Review of the Literature

This section presents the previous analyses and observations in the literature about BP generic null subjects (in subsection 3.4.1) and about other languages (in section 3.4.2). This review is not supposed to be exhaustive, although most of the research about this structure in BP is covered in this section.

3.4.1 Accounts of Generic Null Subjects Focusing on Brazilian Portuguese

The discussion on the generic null subject in BP comes from Galves (1987, 1993), who claims that the impoverishment of agreement marking has structurally and semantically weakened verbal markings, null subjects and subject-verb agreement relation, so that 3rd person singular can be interpreted as generic. Cf. Galves (1987), in finite sentences in BP, such as in (123) below, the 3rd person singular [+ person] feature is defective, so that the empty category turns out to be interpreted as undetermined.

(123) Nos   dias de hoje, não usa   mais saia.
       In.the days of today not use.PRS.3SG plus skirt.
       "Nowadays one does not use skirt anymore."

(Galves 1987, 36-7)

Nunes (1990) argues that this construction comes either from the SE impersonal construction, as in (124), or from the 3rd plural null subject, as in (125). According to him, the first possibility would be a result of the general disappearance of clitics in BP, while the second would be related to the impoverishment of agreement markings, as in *ele/eles fala* "he/they speak.PRS.3SG". Duarte (1993, 1995) also mentions this construction, but she does not offer a fully detailed account of it. Indeed this research is primarily concerned with other phenomena related only indirectly with the generic null subject.

(124) Nos nossos dias não se usa   mais saia.
       In   our   days not SE use.PRS.3SG plus skirt.
       "Nowadays one does not use skirt anymore."

(Nunes 1990, 99)

(125)   Nos nossos dias não usam   mais saia.
        In   our   days not use.PRS.3PL plus skirt.
        "Nowadays they do not use skirt anymore."

<div align="right">(Nunes 1990, 99)</div>

Kato (1999), Cyrino et al. (2000) claim that the generic null subject is doubled by a *PRO*. According to them, the standard analysis for SE constructions has a *PRO* as subject, and *se* absorbs case. With the gradual disappearance of *se* (cf. Nunes 1990), the structure eventually comes to be a *pro* doubled by a *PRO*, such as in the structure sketched in (126b) below. That is, for a sentence such as (127a), in which the use of *se* is still possible, the structure in (126a) is on the way to be fixed as (126b). They also speculate that, as is the case for PRO in other syntactic environments, the generic PRO is always human and, like other strong pronouns, it can be modified in a left fronted small clause, such as in (127b).

(126)   A Syntactic Structure for Generic Null Subjects (*SE* and null)

   a. $[_{IP}PRO \; [_{I'}V + se \; [\; t_v \; \text{à direita}]]$
   b. $[_{IP}PRO_i \; [_{IP}pro_i \; [_{I'} V \; [\; t_v \; \text{à direita}]]$

<div align="right">(Cyrino et al. 2000, 64)</div>

(127)   a. No Brasil circula à direita.

        b. $[_{SC}$ PRO No   Brasil] circula   à   direita.
                        In.the Brazil circulate.PRS.3SG to.the right.
            "In Brazil one drives [cars] to the right."

<div align="right">(Cyrino et al. 2000, 64)</div>

Based on this previous research, analyzing diachronic data taken from Nunes (1990), Gonçalves (2000, 2002), as well as Duarte et al. (2001), argues that the null generic subject comes from indefinite SE. Duarte et al. (2001) suggest that this construction is a result of the on going decline in the number of clitics in BP and of the fact that the 3rd person singular is "pronominal", which can license referentially featureless null subjects. However, Gonçalves (2000, 2002) points out that there are several dissimilarities between the SE construction and null subject generic constructions, despite some similarities. He suggests the idea that if in fact both the constructions are compatible

with generic aspect-temporal sentences, the generic null subject construction requires a semantic external argument, that is, such as that of the transitive or intransitive verbs (mainly transitive, in fact), as in (128) below. However, his informants' judgments are not in all-or-nothing fashion. He assumes that there are several discrepancies among informants and items, in for example (128)-(130).[11]

(128)   a. Aqui _ vende   sapato.
            Here   sell.PRS.3SG shoe
            "Here one sells shoes."

        b. Nesse lugar _ usa   brinco   no   umbigo.
            In.this place   use.PRS.3SG piercing in.the navel
            "In this place one uses navel piercings."

        c. No   shopping,   _ almoça   antes   do   meio-dia.
            In.the shopping.mall   lunch.PRS.3SG before of.the midday
            "In the shopping mall one has lunch before midday."

        d. Nessa biblioteca _ lê   muito livro.
            In.this library   read.PRS.3SG many book
            "In this library one reads many books."

                                                                    (Gonçalves 2002, 37)

(129)   a. ??Aqui _ telefona   a   toda hora.
            Here   telephone.PRS.3SG the all   hour
            "Here one telephones every time."

        b. ?Aqui _ trabalha   até   tarde.
            Here   work.PRS.3SG until late
            "Here one works till all hours."

        c. *Aqui _ tosse   muito no   inverno.
            Here   cough.PRS.3SG in.the winter
            "Here one coughs in the winter."

---

[11] As pointed out to me by Patricia Cabredo-Hofherr, all of the examples with "to sell/repair shoes here" mentioned in previous literature look like an exclusive generic reading (and not the inclusive one). Intuitively, including the speaker in the generic interpretation of this sentences require a specific pragmatic context (as a speaker who is a shoe-repairer or a shopkeeper). This can be the reason for having divergent judgments about the data.

(130) a. *Aqui _ chega   tarde no   serviço.
         Here   arrive.PRS.3SG late   in.the job

         "Here one arrives late in the job."

    b. ?Aqui _ morre   de enfarto depois do   cinqüenta.
         Here   die.PRS.3SG of infarct after   of.the fifty

         "Here one dies from heart attack after being fifty years old."

    c. ??Aqui _ é   bem-vindo.
         Here   be.PRS.3SG welcome

         "Here one is welcome."

    d. *Aqui _ é   preso   por assassinato.
         Here   be.PRS.3SG arrested by   murdering

         "Here one is arrested by murdering."

(Gonçalves 2002, 38)

Paradigms (128)-(130) show that interpreting a null subject as a generic reference is completely context sensitive, and subject to some variability among speakers.[12] This point will be addressed later.

Gonçalves (2000, 2002) establishes a difference between generic and indefinite subjects by building this dissimilarity on four main characteristics: (i) generic subjects are (quasi-) universally quantified, indefinites are existentially quantified; (ii) generic sentences are incompatible with a specific time frame and are rather better accepted in contexts in which the temporal reference is suspended (modal contexts, for example), while indefinites can be temporally specific; (iii) indefinite subjects must be present in all the steps of the syntactic derivation, but generics do not have to be (they could be a last resort at LF, for example); and finally (iv) the speaker can be included in the reference of generic subjects, but not in the reference of indefinite subjects. He uses this set of criteria to differentiate the possible interpretations of SE and the interpretations of the generic null subject and those which overlap. Most of the criteria do indeed coincide with the set of tests proposed in section 3.2, apart from those that are theory-internal.

---

[12]   The judgments on paradigms (128)-(130) are those provided by Gonçalves (2002). Generic null subjects are delicate in terms of acceptability and interpretation as discussed in sections 3.7 and 3.8.

Gonçalves (2000, 2002) also observes that generic interpretation requires a range restriction scoping on the subject of the sentence. Gonçalves (2002) cites Pontes (1987) in order to claim that this restriction is somehow related to the fact that subjects and topics in BP are not distinctive. However, how topics, subjects and the range restriction are related to each other is unclear in his proposal, since they are different categories that can overlap in generic sentences only if the analysis assumes specifically intra-theoretical assumptions, such as those that any category can be a subject, the subject is not a null pronoun, etc.

Formally, Gonçalves (2000, 2002) proposes that the 3rd person is a zero morpheme, whose person features are morphologically and, thus, semantically defective (following Kato 1999's analysis, which is based on Cardinaletti & Starke 1994). Such a morpheme becomes co-referential through a bounding relation between a PRO in a higher than TP spec position (namely ΣP) (following Martins 1996's and Kato 1999's proposals, which are based on Kuroda 1979),[13] cf. the structure in (131) below.

(131)   A Syntactic Structure for Generic and Co-referential Null Subjects in BP

$[_{\Sigma P}$ PRO$_i$ $[_\Sigma[$-specific $]_{TP}$ -$\emptyset_{i[+pronominal]}$ $]]$

(Gonçalves 2000, 77)

(132)   a. A   Maria disse   que _ usa   saia.
             The Maria say.PST.3SG that   use.PRS.3SG skirt
         b. A Maria$_1$ disse que [PRO$_1$ [usa-$\emptyset_1$ saia]]

             "Maria$_1$ said that she$_1$ uses skirts."

(adapted from Gonçalves 2000, 78)

(133)   a. Aqui _ usa   saia.
             Here   use.PRS.3SG skirt

---

13   Kuroda (1979) analyses -*wa* marked DPs in Japanese as part of categorical sentences and -*ga* marked DPs as appearing in thetic sentences. Following him, Martins (1996) and Kato (1999) argues for a similar analysis for EP and BP, whose subject DPs of categorical sentences are interpreted at LF in [spec,ΣP] position, that is, the specifier of a polarity phrase. Kato (1999) goes further by arguing that any strong pronoun or topic is in this position in BP, while weak pronouns are in [spec,TP]. Lambrecht (1994) reinterprets Kuroda (1979)'s proposal in terms of information structure, by arguing for different pragmatic arrangements as regards each structure.

227

b. PRO$_{arb}$/$_1$ [Aqui usa-Ø$_1$ saia].

"Here one uses skirts."

(adapted from Gonçalves 2000, 78)

In case there is no available antecedent (that is, a c-commanding DP), the phi-features of PRO are not specified and, thus, it ends up to be interpreted as generic. Following Cyrino et al. (2000), Gonçalves (2002) assumes that the position of the fronted adverbial phrase is a result of the necessity of obtaining the correct syntactic geometry for scoping such phrase over PRO, as shown in (134) below.

(134)   a. No   Brasil _ circula   à   direita.
            In.the Brazil   circulate.PRS.3SG to.the right
       b. [$_{SC}$ PRO No Brasil] circula à direita.

"In Brazil one drives [cars] to the right."

(adapted from Gonçalves 2000, 81)

His proposal is, however, unable to tackle the different levels of acceptability in (129)-(130), which is claimed to be evidence to suggest an on-going change and normative pressure regarding generic null subject construction.

Rodrigues (2004, 78) assumes that the general analysis for co-referential null subjects is a case of control out of case domains, see Hornstein (1999, 2001, 2002). Discussing Kato (1999)'s proposal, she points out (i) that the presence of a controller does not necessarily imply that PRO is controlled, such as in (135) below;[14] (ii) that the absence of phi-feature specification in the agreement marker does not necessarily lead to the emergence of another phi-feature deficient element, namely PRO; and (iii) that it is unclear whether Agr can satisfy the requirements of TP, since in Kato (1999)'s proposal weak pronouns are in the specifier of TP position in sentences such as (136) below, cf. Modesto (2000*b*): it remains doubtful why 1st person Agr is not able to check TP (thus

---

[14]   I have corrected one typo and inserted the adverbial phrase *aqui* "here" in the example, based on the provided glossa. It is worthy clarifying this is Rodrigues (2004)'s interpretation of this example. My first intuition is that this sentence is co-referential, at least if the prosodic accent of the sentence is on the adverbial phrase in the end.

requiring a weak pronoun), while in generic null subject sentences, such as (135), 3rd person Agr can satisfy the feature checking in TP. This criticism can be extended to all the previous approaches based on Kato (1999)'s theory of co-indexing PRO with Agr.

(135) a. [O Pedro$_1$ disse que [$_{SP}$ PRO$_2$ [$_{TP}$ conserta-Agr$_2$ sapato aqui.]]]
the Pedro say.PST.3SG that repair.PRS.3SG shoe here
"Pedro said that one repairs shoes here."

(adapted from Rodrigues 2004, 101)

(136) a. Eu, eu (phonetically [o]) sinto demais isso, neh.
I, I feel.PRS.1SG too.much this, right
"I feel this too much, right."

(Kato 1999, 14 apud Rodrigues 2004, 102)

As for the analysis of generic null subjects, Rodrigues (2004) takes the existence of a syntactic null subject to be blurred, since it would be expected that they were able to (i) bind syntactically anaphoric elements; (ii) control infinitive clauses and (iii) take a secondary subject-oriented predicate. However, sentences with generic null subjects do not exhibit such properties, since according to her judgments sentences with such properties are unacceptable, as shown in (137) below.

(137) a. Aqui $_1$ não vende nada (*de si mesmo$_1$)
here not sell.PRS.3SG nothing of SE self
"Here one$_1$ sells nothing from oneself$_1$."

b. Aqui vende sapato caro (??para $_1$ evitar cliente pobre)
here sell.PRS.3SG shoe expensive to avoid.INF customer poor
"Here one$_1$ sells expensive shoes in order to $_1$ avoid having poor customers."

c. Aqui _ conserta sapato (*bêbado/irritado)
here repair.PRS.3SG shoe drunk/irritated
"Here one repairs shoes drunk/irritated."

(adapted from Rodrigues 2004, 71)

According to Rodrigues (2004), (137a) turns out to be ungrammatical if a subject-bound anaphor is added. For her, (137b) is an unacceptable sentence when it is followed by an infinitive clause with a controlled subject. And finally a secondary predication referring to the generic

null subject cannot be used in sentence (137c). However, Rodrigues (2004, 72) claims that using a modal verb improves the acceptability of sentences like (137b)-(137c), as shown in (138a)-(138b) below:

(138) a. Nesse hotel _ pode   entrar   na   piscina   sem   _1 tirar
           in.this hotel can.PRS.3SG enter.INF in.the swimming.pool without   take.off.INF
           a   roupa.
           the cloth
           "In this hotel one can enter into the swimming pool without taking his or her clothes
           off."

       b. Nesse hotel _ não pode   entrar   na   piscina   bêbado
           in.this hotel not can.PRS.3SG enter.INF in.the swimming.pool drunk
           "In this hotel one can enter into the swimming pool drunk."

                                                              (adapted from Rodrigues 2004, 72)

Rodrigues (2004, 72) suggests the modal verb selects an infinitival clause. This structure licenses the subject of the root clause as a null expletive, as shown in (139) below.

(139)   A Syntactic Structure for Generic Null Subjects in BP

[$_{AdvLoc}$ [$_{Infl}$ Pro$_{expl}$ [$_{VP}$ Verb$_{Modal}$ [$_{Infl}$ PRO$_1$ [$_{VP}$ t$_1$ V-inf.]]]]]

                                                                        (Rodrigues 2004, 72)

There is in fact no doubt that modals favor the generic interpretation of null subjects, as discussed in sections 3.2.2 and 3.8. In fact, this structure could be a representation of (138), but it cannot be the syntatic structure of (137) without the phrases within parentheses that make the sentences less acceptable for her. Following Guimarães & Rodrigues (2002), she also mentions that adverbial locatives (or other referential expressions) must be fronted for the null subject to be interpreted as generic: the AdvP would then be the sentential subject of generic clauses. For this reason, Rodrigues (2004, ch. 4), claims that in BP fronted locative adverbs block the extraction of subjects from the embedded clause. According to her proposal, such adverbs and other referential expressions, such as direct objects, prevent A-movement out of an embedded finite clause. She proposes that these elements are in FP (Focus Phrase). If CP is a phase, following Chomsky (2000,

2001), the matrix DP subject cannot come from the embedded clause, because the "escape hatch" from its CP is occupied by such fronted elements. Furthermore, Rodrigues (2004, 143), claims that only referential expressions can create this blocking effect, since other fronted constituents, such as temporal adverbial phrases, do not prevent the co-referential interpretation, as in (140) below:

(140) a. O   João$_1$ me contou   que todos os dias _$_1$ vende   cachorro quente
           the João me tell.PST.3SG that every the days   sell.PRS.3SG dog   hot
           na   praia.
           at.the beach

        "João$_1$ told me that he$_1$ sells hot-dogs at the beach every day."

(adapted from Rodrigues 2004, 143)

The reasons why only certain kinds of expressions are able to block subject movement remain unexplained in her dissertation. It is also not certain that the only possible interpretation for the subject in (140) is the co-referential. There is no doubt that the present tense is less likely to be interpreted as a generic predicate in the embedded clauses than, for instance, modal verbs. Using a modal verb in (140) would render the generic interpretation much clearer. This is however a general constraint on the semantics of generic NPs. Furthermore, the embedded clause in (140) is not straightforwardly acceptable out of the main clause, even if the subject is interpreted as generic. The reason for which the subject of the embedded clause in (140) is not at first sight interpreted as generic is probably related to the semantics of generic sentences in general, and not related to the position or type of adverbial phrase.

Avelar & Cyrino (2008) further develop the idea of the AdvP as a subject of generic sentences by arguing that generic null subject sentences are an instance of locative inversion.[15] According to them, the adverbial phrase moves to the specifier position of TP and thus satisfies EPP. An element in the specifier position of vP to be interpreted as an agent is no longer obligatory (as in passive constructions). They take the structure sketched in (141) to represent generic null subject sentences.

---

[15]   Avelar & Cyrino (2008) also argue for the rejection of the analyses of generic null subjects as a simple loss of the SE impersonal construction in present day spoken BP. They show that these two constructions are not subject to the same constraints, since fronted adverbial phrases, for instance, are optional in SE impersonal sentences.

(141) A Syntactic Structure for Generic Null Subjects

Naquela loja vende livros.
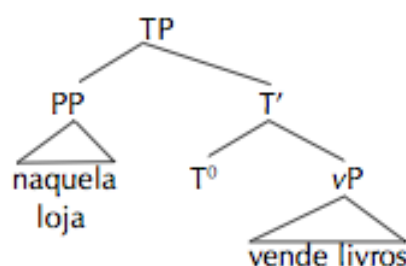
[TP [PP In.that store] [T' [vP sell.PRS.3SG books]]]

"In that store one sells books."

(Avelar & Cyrino 2008)

In the structure in (141), the locative adverbial phrase *Naquela loja* "In that store" occupies the position of specifier of TP. In this position, it is able to be the grammatical subject of the sentence and to satisfy EPP. According to Avelar & Cyrino (2008), the generic interpretation in embedded clauses is due to this syntactic representation.

Holmberg (2005, 2007), Holmberg et al. (2009) have also a similar proposal concerning null generic subjects. However, this proposal comes from a different treatment for co-referential null subjects, which is based on the idea that there is a null unpronounced pronoun in the subject position. As briefly mentioned in Chapter 1, he takes this constituent to be a D[efinite]-less ϕP, that is, determiner phrases [DP] – without a substructure required for a definite referential category – which is deleted. According to him, there are several kinds of syntactically projected null subjects in Finnish. However, he explicitly states that generic null subjects and embedded co-referential null subjects (what he calls "logophoric subject pronouns") are the same category, that is, ϕP (based on Déchaine & Wiltschko 2002).[16] Following Longobardi (2003), he proposes that such a category cannot be independently, deictically or co-referentially bound to a referring DP, but can be a bound variable referring to a quantifier phrase [QP] or "logophorically linked" to a DP in a higher clause,

---

[16]   According to Holmberg (2005), subjectless verbal answers to yes-no questions, for instance, are a different type of null subject constructions (see Holmberg 2001, 2003, 2007, 2016, and see Chapter 2 for another point of view).

as shown by the contrast between (142a)-(142b) below.

(142) a. Ele/*_ ganhou   na   loto.
           he   win.PST.3SG on.the lottery
           "He won the lottery."

       b. O   Pedro₁ disse   que ele₁/ele₂ /_₁/*_₂ ganhou   na loto.
           The Pedro   say.PST.3SG that he   win.PST.3SG the lottery.
           "Pedro₁ said he₁/*₂ won the lottery."

(adapted from Holmberg 2005, 32)

For Holmberg (2005, 2007) and Holmberg et al. (2009), the generic interpretation of the null subject comes from a last resort: the adverbial phrase occupies the [Spec,TP] position and checks EPP, so that ϕP needs to stay in its base-generated position [Spec,vP]; in this position, it is not accessible to be bound by the main clause subject, since it would be so only outside vP, and lacks of a definite interpretation. As a last resort, it is interpreted as generic having been bound by an abstract generic operator (as in Krifka et al. 1995), such as in the example provided in (143) below. According to Holmberg (2005), this operator is located no higher than I(nflectional), so that if the null subject is moved to a higher position, the generic interpretation is not available.

(143) a. Aqui _ não pode   nadar.
           here _ not can.PRS.3SG swim.INF
           "One can't swim here."

(adapted from Holmberg 2005, 32)

However, as Holmberg (2005) mentions, there is no independent evidence to show that generics are confined to the IP domain. Rather, non-pronominal generic subjects can be outside that domain. This is probably a flaw in his proposal: as pointed out in section 3.2, generic null subjects are not semantically different from other generic referents. In Holmberg (2007), he points out that the null subject cannot move out of vP because it is not referential and, thus, unable to check EPP. He thus claims that a generic null pronoun cannot satisfy EPP because it is lexically null (differing for example from 1st and 2nd person null pronouns, which are not lexically null but do not have a lexical form and whose features are deleted at PF). Since it cannot be lexically realized nor thus spelled out, EPP does not attract such an element, so that the null pronoun stands still. According

to him, the generic null subject is locked in its firstly merged position, since it is invisible to the EPP of T and C.

According to Holmberg (2007), furthermore, the generic subject can only be interpreted as human. First he puts forward the idea that this constraint on the interpretation of generic subjects is based on the null pronoun itself (following Egerland 2003), or is assigned by default (lexically underspecified subjects always refer to humans). Later in the same chapter, he explicitly claims that the generic null pronoun has the feature [+ human].

On the other hand, as Holmberg (2005, 2007) and Holmberg & Phimsawat (2015, 2016) propose for classical pro-drop languages, valuing category ϕP with uD features prevents the generic interpretation from emerging. According to this proposal, standard pro-drop languages have a richer agreement marking system than those of partial pro-drop languages. In these languages, in the agree relation, ϕP ends up fully copying the features from T, which includes a definiteness feature. In Holmberg (2005, 2007), he proposes that copying the definiteness feature the ϕP is no longer available to be bound by the generic operator (which as traditionally assumed scopes over indetermined DPs). In Holmberg & Phimsawat (2015), the explanation is slightly different: having copied the feature of T, the pronoun is deleted or not spelled out. According to them, an indefinite or generic 3rd person pronoun cannot be null. For this reason, in standard pro-drop languages, the generic interpretation must be achieved by some sort of overt strategy; in European Portuguese, for instance, as well as in Italian and Spanish, the pronoun SE is used – such as in (144) below – in the contexts BP would allow a null to be generic, such as in (143) above.

(144)   a. Aqui não *(se) pode   nadar.
            here not SE   can.PRS.3SG swim.INF
            "One can't swim here."

(adapted from Holmberg 2005, 34)

Based upon Holmberg's proposal, Barbosa (2011) claims that the null generic subject stays *in situ* (as in Holmberg 2005, 2007, Holmberg et al. 2009), while co-referential null subjects raise to a higher position, in which they get a definite reading. In the case of the generic interpretation, another element can check EPP in the specifier position of TP. The null generic subject inside the VP is thus interpreted under Existential Closure bound by a generic operator, since the topic is occupied by an AdvP. When raised to a higher position, the definite interpretation is available.

3.4.2 Accounts for Generic Null Subject Looking at Elsewhere Data

Cabredo-Hofherr (2006) collects data from different languages which allow generic null subjects – which she calls arbitrary – and anaphoric null subjects in some specific contexts. She argues that co-referential and generic null subjects are subject to two different types of agreement. She mainly proposes that the feature [number] plays a crucial role. Having assumed that 3rd person is standard for the generic interpretation, the distinction between singular and plural is thought to be parallel to the difference between uncountable and countable nouns, so that the former is mass/inanimate and the later is count/human. This difference is encoded formally by the idea that the 3rd person singular has no number at all, but the 3rd person plural has an underspecified number. That is, according to Cabredo-Hofherr (2006), the number feature in singular is absent and never has any value. She underlines that this agreement leads to a mass interpretation that eventually excludes the speaker and the addressee since mass interpretation is inanimate by default.

Cabredo-Hofherr (2006) points out two interesting properties of null subject constructions (very similar to those proposed by Cinque 1988 for Italian *si* construction and assumed also by Egerland 2003 for Swedish *man*): (i) the first factor is the ambiguity of pronouns, which can either denote a unique maximal group, such as a definite plural, or introduce a variable into the semantic representation, as an indefinite; and (ii) the generics that are interpreted as a plural have a mechanism of content identification, and those that are a variable are interpreted through existential closure. However, as said by Cabredo-Hofherr (2006), her account cannot explain data as in (145) from Modern Hebrew.

(145) a. ba-maxane ha-ze, _ 'ovdim   harbe (MH) _ kamim be-SeS ba-boker,   _
          in-camp   the-this,   work.3PL much   get.up at-6   in.the.morning
          matxilim la'avod be-Seva   ve- _ lo   mafsikim 'ad Seva _ ba-'erev
          start   working at-7.o'clock and   only finish   at 7   in.the.evening
          ve-be-sof   ha-xodeS   _ mekablim saxar   'aluv
          and-at.the-end of.the.month   earn   miserable wage
          "One gets up at 6 in the morning, one starts working at 7 o'clock and at the end of the month one earns a miserable wage."

<div align="right">(adapted from Cabredo-Hofherr 2006, 255)</div>

Notice that this example, despite being 3rd person plural, has the same reading as in (112a)

<div align="center">235</div>

in BP. That is, in Modern Hebrew, as well as in BP, if two or more null subjects are co-indexed without being co-referent with some other entity, the generic reading is acceptable.

Holmberg & Phimsawat (2015) propose a detailed description of generic (null) subjects in different types of pro-drop languages (following the typology proposed in Biberauer et al. 2010). According to them, in radical pro-drop languages (Thai is their example and focus), the generic null subject, as delimited in section 3.2 of this chapter, has to be null, such as in (146). It is also in partial pro-drop languages, such as, according to them, Finnish, as for example in (147). However, as mentioned in the previous section, consistent pro-drop languages, such as, in their example in (148), Italian and in the previous section EP, require some sort of overt strategy to bring about the same interpretation. The analysis of this strategy, which can be a SE/SI pronoun in consistent pro-drop Romance languages, but also can be a reflexive in consistent pro-drop Slavic languages, remains controversial (see Cinque 1988, Chierchia 1995, Dobrovie-Sorin 1998, D'Alessandro 2007, *inter alia*, for Romance languages, and Krzek 2013*a*,*b*, *inter alia* for Slavic languages). The fact is that the contrast between consistent pro-drop languages and the other types (including partial pro-drop languages) is clear, as attested when comparing EP with BP, for example.

(146) a. díawníi ŋaan hǎa yâak   mâak thâa _ mây cob trii.
          nowadays job   seek difficult very   if   NEG finish BA
          "Nowadays to seek a job is difficult if one hasn't finished a BA."

                                               (adapted from Holmberg & Phimsawat 2015, 56)

(147) a. Tässä _ istuu   mukavasti.
          here   sit.PRS.3SG comfortably
          "One can sit comfortably here."

                                               (adapted from Holmberg & Phimsawat 2015, 57)

(148) a. Si lavora   sempre troppo.
          SI work.PRS.3SG always too.much
          "One always works too hard."

                                               (adapted from Holmberg & Phimsawat 2015, 56)

After having analyzed these three kinds of null generic subjects in such different languages, Holmberg & Phimsawat (2015) propose that these subjects have no phi-features at all. According

to them, from this idea it follows that these subjects are "inclusive" (following Egerland 2003; see also the discussion in Barbosa 2011 and section 3.2 in this chapter), that is, they include the speaker, or the speaker and his/her associates, or the addressee, or a possible female person who is not the speaker or the addressee, even if they are 3rd person singular. They then mention that many researchers come up with the hypotheses that 3rd singular is the minimal phi-feature specification, either a minus-valued entity [– plural, – participant] or the absence of number and person (see Harley & Ritter 2002, Nevins 2007). However, they claim the partial pro-drop languages are cases in which generic null subjects must have person and number feature values, because in their theory the agreement features of T must have a value assigned to them. So they conclude that in partial pro-drop languages 3rd person and singular values are favored because they are the minimally specified.

In Thai, as described by Holmberg & Phimsawat (2015), there are no agreement markers on verbs. For this reason, they take Thai to have real featureless generic subjects, so that the subject has to be null (since there is nothing to be spelled out) and the reference is maximally inclusive. So, taking Phimsawat (2011)'s idea, Holmberg & Phimsawat (2015) suggest that the null subject has the following structure:

(149)   A Syntactic Structure of Generic Null Subjects

```
        R
       / \
      R   N
```
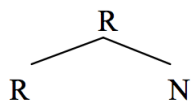
(Holmberg & Phimsawat 2015)

Cf. Holmberg & Phimsawat (2015), in (149), R is a referential feature, which is a property of every argument. As for referential arguments, the value of R is a referential index; as for generic arguments, R is bound by a generic operator, the adverbial operator GENERALLYx.

Holmberg & Phimsawat (2015) also claim that null generic subjects must have some feature specification, since mostly it is claimed to be applied only to humans. However, they verify that this claim is not completely true: Thai, as well as Chinese (Mandarin and Taiwanese), Korean, Japanese, Sinhala, Vietnamese can include plants along with humans and animals, when a generic sentence with a predicate like *to grow* is used. And, according to them, in Brazilian Portuguese,

Finnish, Hebrew, Icelandic, Polish and Thami, this is not possible, although, in Holmberg & Phimsawat (2016), the authors take a step back regarding BP, providing evidence that BP also allow generic null subjects to be inanimate. Holmberg & Phimsawat (2015) also point out that the former set of languages has no subject-verb agreement, while the latter does. Therefore, they propose a different structure for languages like BP, Finish, etc, that is, the partial pro-drop languages:

(150)   A Syntactic Structure of Generic Null Subjects in Partial Pro-Drop Languages
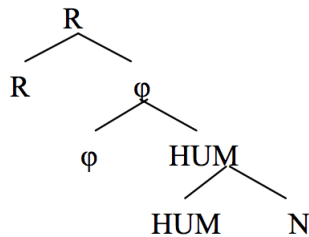
(Holmberg & Phimsawat 2015)

According to Holmberg & Phimsawat (2015), this structure is the null subject in partial pro-drop languages. Besides the referential feature, the structure also has a phi-feature and a HUM-feature.

3.4.3   Summary and Further Steps

Although there are several analyses of generic null subjects in BP and elsewhere, some points seem to be shared among them. Some considerations and proposals are briefly recapitulated and highlighted, because they are at the core of many analyses proposed in the literature. Six questions are put forth below, however only some of them will be further addressed.

(a) Is there a syntactic constituent that is the null generic subject?

(b) If so what is the nature of this constituent? Is it composed of features or an atomic entity?

(c) If they are composed, which features are specified and which ones are underspecified?

(d) In what kind of constructions can a generic null subject appear? What is the role of the fronted element in generic sentences? Which elements can they be?

(e) What is the relation between generic and co-referential null subjects?

(f) What is the role of generic null subjects in the classification of language as regards pro-drop?

All these questions were discussed in the previous literature summarized in section 3.4. In section 3.2, question (f) has already been briefly discussed, but certainly it must be further studied. From the point of view proposed in this dissertation, questions (a), (b) and (c) must be addressed in such a way that they are analyzed as regards other contexts of null subjects, such as co-referential null subjects, generic bound null subjects, etc., and maybe with other null constituents in general, such as null objects, null complement anaphora, etc. Furthermore, these questions seem to be subject to many intra-theoretical assumptions that have been changing since the early 1980s. These assumptions have to be well-established and made sense of in the theory of generic null subjects. From an empirically based point of view, it is intrinsically difficult to attest the syntactic existence of a null element. Even if in a given language null constituents alternate with pronounced pronouns and nouns, speakers can be aware of the paradigm and infer the semantic content of the subject by the absence of the subject with a verb. If this entity is syntactically present, evidence provided in this dissertation suggest that it is decomposable and probably underspecified. This could possibly explain the bias toward non-specified referents in Chapter 1, for example. Being less specified, null subjects can simply substitute for any pronoun and refer to absolutely any entity (from non-specific neuter "nobody" antecedents to extremely specific antecedents, such as the Speaker, from extremely inanimate entities, such as propositional content, to animate antecedents). Possible counter-evidence to this hypothesis would be the fact that generic null subjects refer only to humans. In section 3.5 below, some examples are discussed as regards the nature of the null subject. However no formal proposal for accounting for questions (a), (b) and (c) will be put forward in the present dissertation.

In sections (3.5) and (3.7), questions (d) and (e) are focused on, as they are empirically more accessible. In what follows, how the proposal put forward in Chapter 2 deals with the generic interpretation will be briefly outlined, showing that it does so straightforwardly without further assumptions.

In Chapter 2, it was argued that null subjects in Maximally Contrastive contexts tend to be co-referential with discourse Obvious antecedents. When they are in non-Maximally Contrastive contexts, they are likely to be interpreted intra-discursively. The hypothesis of the interpretation of null subjects in non-Maximally Contrastive contexts is repeated in (151) below, along with the definition of "Intra-discursiveness" in (152).

(151) Non-Maximally Contrastive Null Arguments Interpretation Hypothesis [Non-MaxCoNAIH] (first version)

   If a null argument is in a non-Maximally Contrastive utterance, its interpretation is intra-discursively defined.

(152) Intra-Discursively Defined

   An interpretation is intra-discursively defined iff it defines a hypothetical set of possible worlds without recourse to the common ground, to the physical world or to world knowledge.

There two points to be made regarding the generic interpretation in order to account for it without any further assumptions. The first is that the generic interpretation of null subjects comes up in non-Maximally Contrastive contexts in BP. This can happen when the antecedent is not obvious in the common ground or in the current QUD and when the sentence in which the null subject appears is not in the highest contrastive level; that is, the assertion does not coincide with the whole overt material. This is an empirical question which will be pursued in the next section. The second point is that the generic interpretation is not established through recourse to the common ground, to the physical world or to world knowledge. This is a theoretical assumption that is assumed here based on research reported elsewhere. In this sense, an idea of genericity close to that proposed by Crone & Frank (2016) and Tessler & Goodman (to appear) is assumed, according to which genericity depends on a failure to ground expressions and on subjective beliefs. Following these authors (see also Moltmann 2006, 2010), the generic interpretation fundamentally depends on failing to refer to specific entities and situations; failing to achieve a common-groundness, a sentence can be interpreted intra-discursively as an imprecise proposition that the speaker is putting in the discourse to be accepted or not by the audience when it is an assertion. Therefore, no access to the common ground, to the physical world or to world knowledge is necessary to interpret a generic sentence. Rather the interpretation follows from the acceptance that there is a predictable prevalence of a given belief. It is not necessary to go into the details of genericity, since the main focus are questions (d) and (e). What is important and will be pursued in the next sections is how the hypotheses proposed in Chapter 2 are able to deal with the generic interpretation and the contexts in which it emerges.

3.5 Some Discussion Based on Corpus Evidence

In this section, the role of the fronted element in triggering the generic interpretation for the null subject is studied, as well as some other previously mentioned constructions, and the relation of generic and co-referential null subjects – questions (d) and (e) in the previous section. All generic subjects presented in this section were tested for the four properties presented in section 3.2, and they thus fit the characterization of generic null subjects proposed there.

It is worth remembering that (i) as stressed by Avelar & Cyrino (2008), an adjunct is obligatory, but can be at the end of the sentence, as in (153); and in embedded clauses the fronting element facilitates the generic reading, blocking the co-referential interpretation as in (114)-(115), repeated here as (154)-(155).

(153)   a.   _ Vende   muitas calças naquela loja.
              Sell.PRS.3SG many   paints in.that   store
              "One sells many paints in that store."

                                              (adapted from Avelar & Cyrino 2008, 62)

(154)   a.   a   gente$_1$ achou   que _$_1$ tinha   que fazer alguma coisa...
              the people think.PST.3SG that   have.PST.IMP.3SG that do   some   thing
              "we$_1$ though that we1 had to do something..."

                                                            (NURC-RJ, Inquiry 158)

(155)   a.   Aí ela disse   que "mesmo nos   momentos mais sombrios... _ não
              So she say.PST.3SG that even   in.the moments   more dark   not
              pode   deixar de acreditar na   luz...".
              can.PRS.3SG leave   of believe   in.the light
              "So she said that even in the darkest moments one cannot stop believing in the light."

                              (http://sitedobernardo.blogspot.fr/2011/05/pena-so-de-gente-ignorante.html)

First of all, fronting a constituent is not necessary for the generic interpretation to emerge, as shown in (153), nor sufficient, since there are contexts in which the interpretation of the null subject is co-referential even if there is a fronted element, such as in (156) below:

(156) a. O, o quarto ela₁ atapetou de amarelo. É, _ chama ouro₂, né,
The the bedroom she carpet-pp-3s of yellow [dm] call.PRS.3SG gold [dm]
quase a mesma coisa, porque não, depois _₂ desbota e _₂
almost the same thing why not after discolor.PRS.3SG and
fica amarelo mesmo, né? _₁ Botou uma estante correndo numa.
get.PRS.3SG yellow really [dm] Put.PST.3SG a shelf running in.one
[AdvPNa janela que dá pra rua e, e encostado na, na,
In.the window which give.PRS.3SG to.the street and, and drawed.up in.the in.the
nessa estante] _₁ botou duas camas modernas sem, sem pés e com
in.this shelf put.PST.3SG two beds modern without without legs and with
dois gavetões.
two drawers.

"The, the bedroom she₁ carpeted with yellow. Yeah, it₂ is called 'gold', you know, al-most the same thing, why not? After it₂ discolors and ends up actually yellow, doesn't it? She₁ put a shelf along one of them [=the walls]. Next the window which faces to-ward the street and, and right next the, the, this shelf she₁ put two modern beds without legs but with two drawers."

<div align="right">(NURC-RJ, "Inquiry_011")</div>

(156) leads one to conclude that any analysis that claims that the fronted AdvP is the actual subject of the sentence or that it blocks any possibility of co-reference is overgeneralizing. There are fronted constructions with co-referential null subjects in BP, which for me seem less accept-able. However, this decline in acceptability does not make the sentence ungrammatical, but rather, as predicted by the hypothesis put forward in Chapter 2, is caused by the fact that it is a less optimal context for null co-referential subjects. In (156), the sentence does not fail completely to be grounded: being episodic (because of the verbal tense), the null subject is forced to be co-referential, since no other possible interpretation is available. A manipulation in verbal tense in (156) would lead to a generic interpretation if a modal or an imperfective verb were used. In fact, the contrast in embedded sentences is a matter of preferential interpretations: in (154), without a fronted constituent, the null subject tends to be interpreted as co-referential; in (155), the AdvP *mesmo nos momentos mais sombrios* "even in the darkest moments" dislocated to the left biases the null subject to a generic interpretation rather than a co-referential one.

Another shortcoming for many proposals presented in the previous section is the nature of the fronted element. It is claimed by Rodrigues (2004) and Avelar & Cyrino (2008) that only refer-

ential AdvPs, such as locative adverbials, are able to trigger the generic interpretation. However, other AdvPs can lead to such a reading of the null subject, such as the temporal AdvP in (157) below. According to the hypothesis in Chapter 2, any fronted element can trigger the generic interpretation, by producing non-Maximal Contrastive utterances.

(157) a. quando chove   _ tem   que fechar   tudo...   porque _ não
           when   rain.PRS.3SG   have.PRS.3SG that close.INF everything because   not
           tem   beiral
           have.PRS.3SG eaves
           "When it rains, one has to close everything, because there aren't eaves."

<div align="right">(NURC-RJ, "Inquiry_011")</div>

Temporal (and virtually all other semantic types) of AdvPs can bias the null subject toward a generic interpretation. However, not only AdvPs facilitate the generic interpretation. Other kinds of fronted constituents also favor it, as in (158) below:

(158) a. O que que _ mistura   primeiro?
           the what that   mix.PRS.3SG first
           "What does one mix first?"

<div align="right">(adapted from Duarte 1995, 62, from NURC-RJ)</div>

In (158) the WH-phrase *O que (que)* "What" leads to the null subject being interpreted as generic. In fact, Rodrigues (2004) says that WH-phrases are one of the contexts that blocks co-reference. This is not completely correct, as shown in Chapter 4 below. However, the point is that even if the WH was not fronted, the generic interpretation would be preferential, as in _ *Mistura primeiro o quê?* "one mixes first what?". So, it is possible to conclude that, whatever the fronted element may be, it is not in a subject position satisfying EPP.

Furthermore, clauses without conjunctions, linked to the main clause by juxtaposition, can also trigger the generic interpretation. In (159) below, the first clause, in which the verb *chega* "arrive.PRS.3SG" has an impersonal, athematic subject, triggers the generic interpretation in the second clause, in which the modal verb *tem* "have.PRS.3SG" has a deontic value.

(159) a. então _ chega   assim sete ho... horas da   noite,   oito horas, _
           so   _ arrive.PRS.3SG so   seven   hours of.the evening, eight hours _
           tem   que dar jantar apesar   de ter   babá.
           have.PRS.3SG that give dinner even.though of have baby.sitter
           "So, when it is around seven or eight pm, one has to prepare the dinner, even though

           one has a baby sitter."

It is worth noticing that these juxtaposed clauses can be of several semantic types. In (159), the clause has a temporal value. However, one of the standard contexts for generic interpretation is conditional-like sentences. In such syntactic-semantic environments, BP allows generic null subjects, bound to the main clause such as in (160) below, or not bound such as in (161):

(160) a. _ descia,   _ ia   dar no   quintal.
           down.IMP.3SG   go.IMP.3SG give in.the garden
           "If one left [the house], one would arrive to the garden."

(161) a. Aí _ chega   na   hora do   jantar. _ bota   lenha,   fica tudo
           So   arrive.PRS.3SG in.the hour of.the dinner,   put.PRS.3SG firewood, get all
           pretinho   outra vez.
           black.DIM other time
           "So, it is time to have dinner. If one puts the firewood, all gets black again."

In some sentences, the generic interpretation is possible even thought there is no left dislocation at all. However, it seems that a kind of restrictor must be at least implicit in the context, such as in (162), or the clause must have a conditional value, such as in (163).

(162) a. _ sabe   essas coisas que   _ têm   né? _ não pode
           know.PRS.3SG these things which   have.PRS.3SG [dm]   not can.PRS.3SG
           botar   na   cozinha e...
           put.INF in.the kitchen and
           "Do you know these things that exist, don't you? One cannot put [them] in the kitchen

           and..."

244

(163) a. _ Precisa   de uma coisa, compras de casa   sou   eu que faço,
          Need.PRS.3SG of a   thing shopping of house be.PRS.1SG I   that do.PRS.3SG
      "If one needs something, it's me that do the house shopping"

In (162), *essas coisas* "these things" is a topic that is highlighted by the interlocutor in the first sentence. In the second sentence, this constituent is active and salient, as a topic to be talked about. In this context, it can be taken as a restrictor to the scope of the generic sentence. In (163), however, there is nothing implicit at the beginning of the sentence, but the first clause needs a posterior link with a second clause; this second clause has a fronted element, *compras da casa* "house shopping", that can be interpreted as a restrictor as well.

To conclude, the restrictor is semantic and independent of syntactic realization or position. However, as proposed by Krifka et al. (1995), Diesing (1992), among many others, generic sentences must have a conditional-like semantic content, that is, a restrictor and a predicate. As regards syntactic position, it seems to be completely orthogonal to the possibility of interpreting a subject as generic. The difference in position that favors one or other interpretation is possibly a by-product of other constraints on the co-referentiality of null subjects, as proposed in Chapter 2. In section 3.7, this question is directly addressed by an experiment. The result suggests that the generic interpretation is not the preferential one in optimal contexts for co-reference; but, given non-optimal conditions for co-reference, the generic interpretation emerges straightforwardly, and not as a last resort, as previously proposed. The explanation for the generic interpretation will be fully spelled out in section 3.8, based on the idea that genericity comes from a failure to ground the sentences referentially.

Last but not least, the semantic restriction on the interpretation of the referent of the generic null subject, which would be constrained to human(ized) entities, cannot be applied to BP, since this language allows generic inanimate null subjects, cf. (164) below.

(164) a. Paredex é   uma tinta que   você$_1$ põe   dentro de um balde, $_1$
          Paredex be.PRS.3SG a   ink which you   put.PRS.3SG inside of a   bucket
          põe, $_1$ joga   um pouco de água, $_1$ mistura,   quer
          put.PRS.3SG   throw.PRS.3SG a   few   of water   mix.PRS.3SG want.PRS.3SG

245

dizer, qualquer um, inclusive você, pode   pegar   uma broxa   e
say   any   one including you   can.PRS.3SG get.INF a   paint-brush and
pintar   e,   e $*_{-1}/_{-2}$ sai   bem pintado.
paint.INF and and   leave.PRS.3SG well painted

"'Paredex' is a kind of ink which you$_1$ put into a bucket, you$_1$ do, you$_1$ throw some water into it, you1 mix them, I mean, anyone, including you, can get a paint-brush and paint and *you$_1$/anything$_{gen}$ ends up well painted."

<div align="right">(NURC-RJ, "Inquiry_144")</div>

In (164), although there are multiple good antecedents for the last null subject, the subject of the verb *sai* "leave.PRS.3SG", it is interpreted as generic, that is, the subject is roughly equivalent to "anything/everything" (against Chierchia 1995, Kato 1999, Cyrino et al. 2000, Holmberg & Phimsawat 2015).

3.6   Summary about Generic Null Subjects

Many points mentioned so far remain under-explored. Although it is not the intention of the present dissertation to be exhaustive, broaching the discussion and stressing what is at issue when dealing with generic null subjects and their implications for the theory of pro-drop are the main goals of this chapter. Throughout this chapter it is highlighted that the analysis of generic null subjects is far from being undisputed and needs to be better understood within the general theory of null subjects and other types of null co-referential entities. Furthermore, the data provided in this dissertation is taken from real attested examples and, as much as possible, is from a well established and controlled corpus, which has been already studied by many researchers (as in Duarte 1995, for example). The lack of a quantitative approach to generic null subjects in corpora is due to the fact that it is not clear with which overt strategy they alternate. For instance, null generic subjects are taken to alternate with the 2nd singular pronoun in Holmberg & Phimsawat (2016), with SE construction in Kato (1999) and with 3rd plural subjects in Nunes (1990). Not having clear criteria to decide with which of these elements they should be compared, a quantitative analysis of generic null subjects is left for future research.

Generic null subjects are, nonetheless, a fundamental piece of evidence for the theory proposed

in the present dissertation. In section 3.5, previous approaches were confronted and challenged by applying them to data from NURC-RJ and by showing that these analyses cannot account for the generic interpretation in the general theory of pro-drop in BP in a satisfactory way. Most of them have interesting empirical observations, but when their claims are confronted to real corpus-based examples, they fail to predict some patterns and overgeneralize for others. At the same time, generic null subjects could be challenging for the theory of pro-drop in BP presented in Chapter 2. At first sight, the reference of generic null subjects could be taken to be deictically established with respect to the world or encyclopaedic knowledge. In fact, not having a clear antecedent in the discourse is a requisite for the null subject to be interpreted as generic. However, in sections 3.4.3, the account for generic null subjects based on the theory proposed in Chapter 2 was explained. In section 3.8, this account is further explored based on experimental data provided (Experiment 5) in section 3.7. In this experiment, the most challenging context for many approaches was tested: the generic interpretation within embedded clauses. The results support the approach advanced in the present dissertation.
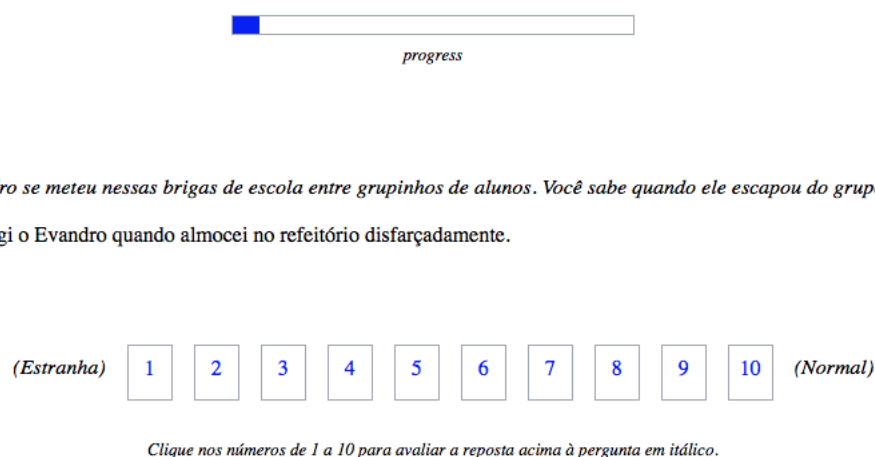
3.7   Experimental Data

Given the observations from the corpus research described in the previous sections and the theories that could possibly account for the interpretation of null subjects, an experiment was carried out in order to check whether manipulations in the position of adjuncts can trigger different interpretation and whether one of these interpretations is more acceptable than the other one. In this experiment, two features were manipulated to try to produce two different interpretations for the relevant null subjects: the position of the adjunct in the subordinate clause and the gender of the secondary predication; additionally, by the analysis of a judgment task, both possible interpretations of the sentences are evaluated in terms of mean acceptability. Two main findings arose from this experiment: (i) the generic interpretation seems to be restricted to certain speakers (there are clearly speakers who lack this interpretation altogether, interpreting all the items as co-referential regardless of the position of the adjunct) and (ii) for the group who has the generic interpretation, no significant difference in acceptability between the generic and the co-referential interpretation was found (differently from the forced co-reference with a non-salient antecedent, which was judged

significantly less acceptable), while for the group which lacks the generic interpretation at all, a significant interaction of factors between the position of the adjunct (Factor Adjunct) and gender of the secondary predication (Factor Gender) as well as each main factor was at least marginally significant.

3.7.1   Methodological Considerations

In this experiment, as in the experiments reported in Chapters 1 and 2, participants read a sequence of two turns, that is, a short dialogue. So they were asked about the acceptability of the answer in the relevant context on a scale from 1 to 10 in a judgment task, cf. Figure 3.1. They were told to use the full scale according how natural "Normal" or strange "Estranha" the answer sounded in the context of the question. After judging them, the participants were asked about the interpretation of the relevant subject in a closed yes-no question task, cf. Figure 3.2.

Figure 3.1 – Screen sample – Judgment Task

progress

O Evandro se meteu nessas brigas de escola entre grupinhos de alunos. Você sabe quando ele escapou do grupo rival?

Eu protegi o Evandro quando almocei no refeitório disfarçadamente.

(Estranha) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | *(Normal)*

Clique nos números de 1 a 10 para avaliar a reposta acima à pergunta em itálico.

Figure 3.2 – Screen Sample – Closed Question Task
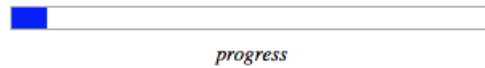
All participants voluntarily participate in the experiments by the IbexFarm platform (http://spellout.net/ibexfarm). They gave their written consent in having the data produced by them used for scientific purposes in a anonymous fashion. They filled in a basic information form and had 4 sentences to practice before starting the experiment, which took them around 25 minutes to be completed.

3.7.2 Experiment 5 – Generic vs Co-referential Subjects and Adjunct Position

This experiment was designed to find out whether BP speakers tend to interpret the null subject differently depending on where an adjunct is placed in the sentence. An introductory context presented two possible referents, a masculine and a feminine proper name. In the second context sentence, a question about what the masculine referent told the feminine one was presented. The answer to this question is a reported speech, whose referent of the relevant subject could be either the (masculine) subject of the main clause or the feminine referent present, but not salient, in the discourse; otherwise, a generic interpretation could emerge. Two Factors were manipulated in order to trigger each of these interpretations: Fronted vs. *In Situ* Adjunct (Factor "Adjunct") and Masculine vs. Feminine Secondary Predication (Factor "Gender"). As in the previous experiments reported in sections 1.7 and 2.5, highly acceptable control sentences were randomly presented among the experimental sentences, and completely unacceptable sentences were presented in the very end of the experiment.

The hypotheses were the following: (i) if the genericity is triggered by structural means (that is, the position of the Adjunct in [Spec, IP, TP or TopicP]), the generic interpretation should come up in the Fronted Adjunct Condition with masculine gender, but it should be less acceptable than the co-referential interpretation, since it would be produced by a last-resort strategy; with feminine gender marking, where the generic interpretation is excluded, again the fronted adjunct should be less acceptable than the *In Situ* adjunct, since it occupies the relevant spec position, and so produces an absolutely unacceptable derivation, because of a weak crossover effect (if, for example, movement is assumed for the co-referential null subject); (ii) if the participants interpret the null subject as generic in non-optimal contexts for co-reference, the generic interpretation would be triggered by the fronted Adjunct, and there would be no significant difference between the *In Situ* and Fronted Adjunct Condition regarding acceptability, as predicted by the Non-MaxCoNAIH; and (iii) if the participants do not interpret the Fronted Adjunct Condition sentence as generic, there would be a significant decrease in acceptability in this condition, and the mean acceptability rates should be closer to those of the subject referring to non-salient feminine referent in the context.

Material Design

Twenty items were created for this experiment, based on previous interpretations and judgments reported in the literature. As in the previous experiment, the relevant antecedent is introduced in the preceding context, to be available as a topic of the current conversation. Differently from the previous experiment, however, the referent of subject was not always the most salient in the context (Salience here is taken to be the closest antecedent, which is also the natural topic of a Maximal Contrastive answer for the current QUD). The context question, cf. (165a), was provided in order to make the sequence sound as natural as possible. This context question was composed of an introductory context and an indirect reported speech question asking *o que* "What" the masculine subject told the feminine referent (a Prepositional Phrase). The answer to the question displayed two Factors in a two by two design (two Factors with two levels). The relevant manipulations were carried out in the subordinate clause introduced by *que* "that", directly addressing the relevant question, cf. (165b-i)-(165c-ii) below. All three possible interpretations (the masculine referent, the feminine referent and the generic interpretation) were plausible in the pragmatic context, as an explanation for the act performed according to the first sentence in the dialogue.

The two by two design was the following: in items (b), such as (165b-i) and (165b-ii), the secondary predication was a masculine adjective; in items (c), such as (165c-i) and (165c-ii), the secondary predication was feminine (Factor Gender); in items (i), such as (165b-i) and (165c-i) the adjunct was at the beginning of the subordinate clause, while, in items (ii), such as (165b-ii) and (165c-ii), the adjunct was *in situ* (Factor Adjunct).

Afterwards, the participants had to answers to an interpretation question indirectly questioning whether the masculine referent recommended/suggested/told the described action to the feminine referent. The expected answer was *Sim.* "Yes" when the secondary predication was feminine and when the null subject is interpreted as generic; otherwise, when the null subject is interpreted as co-referential with the subject of the main clause, the expected answer was *Não.* "No", cf. (166).

(165)   a. A – O João$_1$ e a Maria$_2$ foram embora cedo. Você sabe o que o João$_1$ disse para a
            Maria$_2$?

            "João$_1$ and Maria$_2$ went away early. Do you know what João$_1$ told to Maria$_2$?"

    b. i. B – O   João$_1$ disse   que no   teste de motorista $_{\text{gen/1}}$ não pode
            The João tell.PST.3SG that in.the test of driver   not can.PRS.3SG
            ir   cansado.
            go.INF tired.MASC

            "João$_1$ told that one/he$_1$ could not go to the driving test tired."

        ii. B – O   João$_1$ disse   que $_{\text{gen/1}}$ não pode   ir   cansado
            The João tell.PST.3SG that   not can.PRS.3SG go.INF tired.MASC
            no   teste de motorista.
            in.the test   of driver

            "João$_1$ told that one/he$_1$ could not go to the driving test tired."

    c. i. B – O   João$_1$ disse   que no   teste de motorista $_2$ não pode
            The João tell.PST.3SG that in.the test of driver   not can.PRS.3SG
            ir   cansada.
            go.INF tired.FEM

            "João$_1$ told that she$_2$ could not go to the driving test tired."

        ii. B – O   João$_1$ disse   que $_2$ não pode   ir   cansada   no
            The João tell.PST.3SG that   not can.PRS.3SG go.INF tired.FEM in.the
            teste de motorista.
            test   of driver

            "João$_1$ told that she$_2$ could not go to the driving test tired."

251

(166) Então, o João₁ recomendou que a Maria₂ não deveria ir no teste cansada?

"So did João₁ recommend that Maria₂ should not go to the driving test tired?"

a. Sim.

"Yes."

b. Não.

"No."

Following the results from the corpus study and the previously reported literature, the empirical predictions are the following: the items such as (165b-i) would be interpreted as generic, while items such as (165b-ii) would be interpreted as co-referential, but no significant difference as regards acceptability is expected if the generic interpretation is not a last resort. If the generic interpretation is a last resort to save the derivation, the acceptability of sentences like (165b-i) should be lower than the co-referential sentence (165b-ii).[17] The sentences with a secondary predication with feminine agreement, such as (165c-i) and (165c-ii), should be considered less acceptable, because the antecedent for the null subject would be not the most obvious, and so the interpreter would have to use a last resort to achieve the correct meaning of the sentence. However, they would not be as bad as the control sentences, which are unacceptable for strong syntactic and pragmatic violations.

Participants

One-hundred-fifty-two participants took part in this experiment, all of them highly educated people (minimally undergraduation students) living in the south region of Brazil (Rio Grande do Sul and Santa Catarina). They were invited to participate via email and Facebook, and gave their written consent to use the data. Their age averaged 38,5 years old (more or less 6 years). They work or study in several different fields (all the 4 big areas of knowledge described by the Brazilian government were included). In the result analysis, eight participants were discarded because they systematically judged ungrammatical control sentences above eight, and twenty-three participants

---

[17]   Other sorts of "last resort" strategies (such as, resumptive pronouns in English) are shown to be less acceptable than "non-last-resort" sentences, althought they are significantly slightly more aceptable than sentence with syntactic contraint violations ECP (the higher difficult to extract from subject position than from object position) and Island Effects (see Keffala 2013).

who answered all the questions in the interpretation task positively were also set apart.[18]

Results

The results show two divergent tendencies across groups of participants. The first group consistently had both the co-referential and the generic interpretation across items (no particular item had only one interpretation of answers), answering "yes" to the interpretative task when a generic interpretation was expected. This group was constituted of sixty-eight participants, mainly older than 30 year old, with high volume of weekly electronic reading (also the participants in this group mostly work in the fields Math, Engineering, Life and Earth Sciences and Linguistics). The second group consistently interpreted the null subject as co-referential with the masculine referent in Condition (b) (masculine marked adjective) or the feminine referent in Condition (c) (feminine marked adjective). This group was constituted of fifty-three participants whose weekly reading time is lower than that of the previous group, who are younger (the group is mostly constituted by people from 20 to 30 years old), and mainly working in Humanities (see Table A.1 in Appendix A – Experiment 5 for the distribution of participants' answers, which motivated the division into two groups). Figure 3.3 shows the distribution of generic interpretations across age groups and daily time spent reading.

---

[18] It is possible that the participants who answered all questions of this experiment positively really interpreted Condition (165b-ii) (*in situ* adjuncts and masculine marked secondary predication) as generic. This could be explained by different hypotheses: (i) since the materials were written, they may have interpreted the focused constituent differently from what was expected (rightmost focal accent), as a non-Maximally Contrastive context; or (ii) their interpretations are more pragmatically permissive, extending to other possible people in the context a rule that the referent of the main clause imposes to himself (in this case, including the feminine referent under the scope of the predicate). In any case, it is impossible to be sure that they are not simply answering positively across the board. Because there was no reliable interpretation of these results, they were excluded from this analysis.

Figure 3.3 – Distribution of Participants across extra-linguistic factors
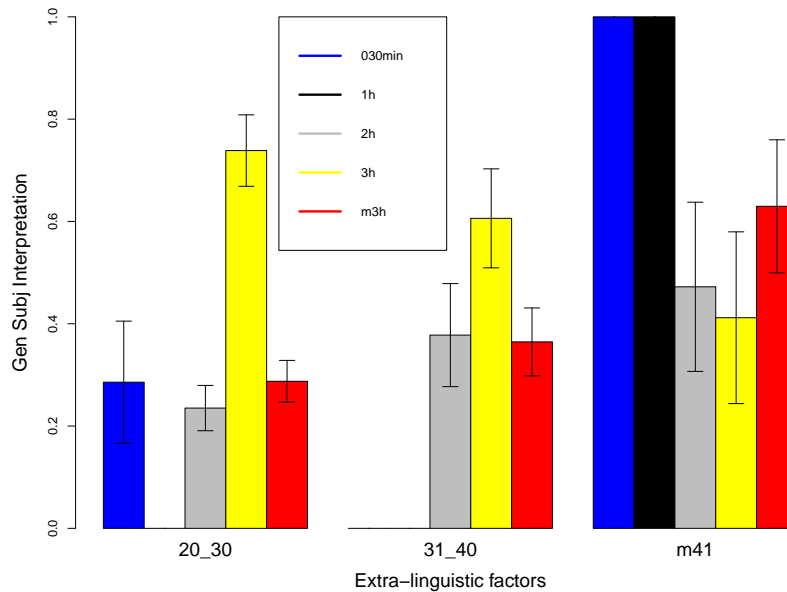
As shown in Figures 3.4 and 3.5 below, the experimental results strongly diverge according to the group of participants: while the group which interpreted null subjects in Condition Masculine:Fronted as generic judged the sentence almost as acceptable as the co-referential, the group which interpreted Condition Masculine:Fronted as co-referential judged it much less acceptable, almost at the same level as the feminine referent with a fronted adjunct (Condition Feminine:Fronted). Being an optimal context for the co-reference with the matrix clause subject, Condition Feminine:*In Situ* was considered the worst for both the groups of participants (although it is noticeable that there is a divergence when the adjunct is fronted).

Figure 3.4 – Mean Acceptability Judgments according to Adjunct and Gender Factors for Participants with Generic Interpretations
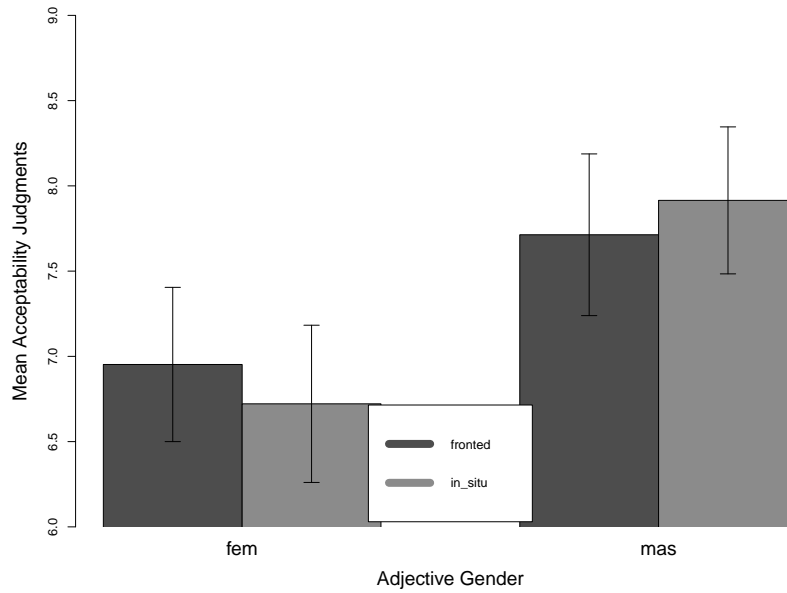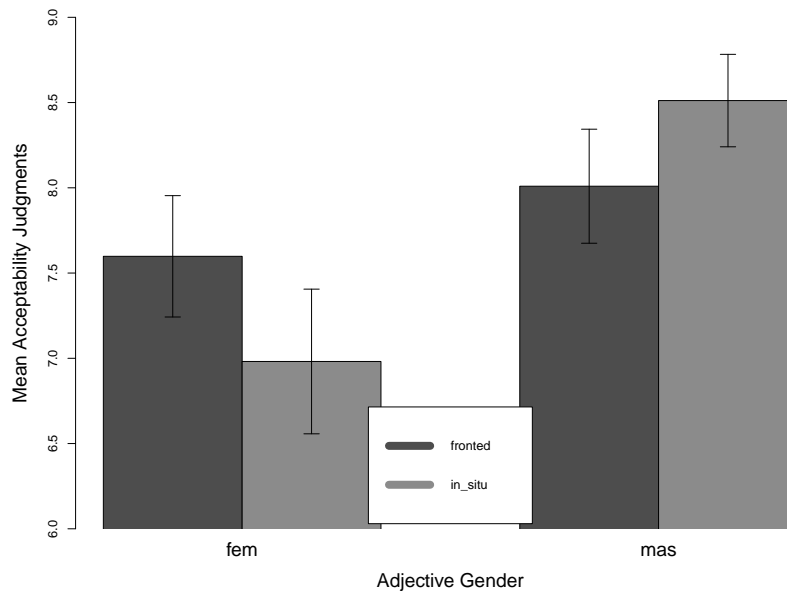

Figure 3.5 – Mean Acceptability Judgments according to Adjunct and Gender Factors for Participants with only Co-referential Interpretations

For the inferential statistical analysis, mean acceptability judgments were entered into a log-linear mixed-effects model analysis containing two main Factors (Adjunct and Gender) with two

levels and random effects (Participants and Items). As for the first group of participants (those who interpret Condition Masculine:Fronted as generic), no main effect or interaction was significant, except Gender, which clearly affects the acceptability of the sentences, cf. Table 3.3. In the second group of participants (those who do not interpret the subject as generic in any condition), Factor Adjunct approached significance, Factor Gender and the interaction between both Factors were significant, as summarized in Table 3.4.

Condition	Estimate	Std. Error	T-value	Pr(>|t|)
(Intercept)	6.9125	0.3353	20.616	< 2e-16 ***
Adjunct	-0.2525	0.3484	-0.725	0.47693
Gender	0.9186	0.2972	3.090	0.00331 **
Adjunct:Gender	0.3402	0.4540	0.749	0.45805

Table 3.3 – Log-linear mixed-effects model for Experiment 5 (Participants with both the interpretations)

Condition	Estimate	Std. Error	T-value	Pr(>|t|)
(Intercept)	7.6047	0.3236	23.497	<2e-16 ***
Adjunct	-0.5488	0.2933	-1.871	0.0772 .
Gender	0.4672	0.2233	2.092	0.0482 *
Adjunct:Gender	0.9322	0.3986	2.339	0.0295*

Table 3.4 – Log-linear mixed-effects model for Experiment 5 (Participants with only the co-referential interpretation)

Discussion

The result of Experiment 5 show variation across BP speakers, but not across items. That is, BP speakers consistently interpreted Condition Masculine:Fronted either only as generic or only as co-referential, but Condition Masculine:In_Situ as co-referential. Interestingly, as the interpretations diverge, the acceptability judgments do as well. The group of participants who interpret Masculine:Fronted as generic judged generic and co-referential subjects (that is, the fronted and the *In Situ* Adjunct Conditions) equally acceptable, while the forced co-reference with a less salient feminine antecedent was judged less acceptable. As for the group which has only the co-referential interpretation, the optimal context for co-reference (*In Situ* Adjunct with Masculine Secondary Predication) was judged the best, but the acceptability of the Fronted Adjunct Condition is much lower and closer to the acceptability of the non-obvious feminine antecedent. Not surprisingly

the context that would be optimal for co-reference is judged the worst when the interpretation of the null subject is changed at the end of the sentence by the feminine agreement, reversing the incremental expectation. These results strongly support the approach proposed here since the non-optimal context for null co-referential subjects is judged worse than the optimal when the interpretation is co-referential, but at the same level when the interpretation is generic. These findings show that the generic interpretation does not emerge as a last resort to save the sentence, but rather is triggered by the non-optimal context for co-reference (in case the context allows and the speaker has this interpretation).

3.8   General Discussion

Semantics of Generic Sentences   As shown in section 3.2.2, sentences with generic null subjects in BP pass all the standard tests used to identify genericity in other generic constructions either in BP or in other languages. It has been noted in the literature about generics that some terms are ambiguous between a generic and non-generic reading, and the sentences in which they appear are similarly ambiguous between the two readings. Carlson (1977) observed this phenomenon in English, cf. (167) below.

(167)   Dinosaurs ate kelp.

(Carlson 1982: p. 163)

(167) is clearly ambiguous between a generic reading (similar to "in general dinosaurs were kelp-eaters") and an episodic reading (equivalent to "there were some dinosaurs that ate some kelp"). It is not surprising that interpretations can vary across speakers (some of them prefer the episodic interpretation while some of them prefer the generic). It has also been observed by Holmberg & Phimsawat (2016) in a very recent paper that not all speaker accept the generic interpretation for the null subject in BP, but some prefer the 2nd person singular overt pronoun *você*, for conveying the inclusive generic reading. While the extra-linguistic factors that group some speakers into one or another set of interpreters must be further studied, the intra-linguistic factors that trigger the generic interpretation (for those who interpret the null subject this way) seem to be accounted for by the theory presented here. Similarly, for those who do not interpret the null

subject as generic, the theory also predicts the lower acceptability of the relevant conditions.

In the present dissertation, standard semantic representations for generic sentences are assumed (Krifka et al. 1995, Diesing 1992). In this vein, no especial syntactic or semantic structure is proposed for such constructions (contra the literature presented in section 3.4). As shown by Crone & Frank (2016) in a very recent series of psycholinguistic experiments about generic and non-generic sentences in English, such as example (167) above, the generic interpretation emerges from a failure to ground expressions as referring to specific entities or events. They found two main factors in their study that seem to trigger the generic interpretation: (i) tense of the sentence and (ii) definiteness of the sentence's subject. They point out that the effects of such factors can be accounted for by the proposal according to which generic interpretations are the result of listeners' failure to attribute expressions' references to particular entities or events. However, as observed also in Experiment 5, judgments regarding generic or non-generic meanings seem to be graded inferences about speakers' intentions. The factors that they identify make the sentences more likely to be interpreted as generic or not. Here, a similar path is taken. The modal verb and the position of the adverb are taken to be factors that favor the generic interpretation for the null subject because they seem to indicate a failure in linking the sentence to the common ground.

Failure in Grounding Utterances   Let me recapitulate how the present theory is able to deal with both patterns found in the Experimental Evidence. For this, the assumption that the generic semantic structure is the same in the null subject constructions and in other generic statements is taken for granted. The next step is, thus, to explain how the null subject fails to be grounded. In what follows, an explanation for the failure in grounding the reference of null subject is given. As shown below, it is a by-product of the strict conditions under which the co-referential null subjects are optimal in present day BP. No additional assumption is required.

In (168) below, the hypothesis about the co-referential null subjects is repeated, followed by the important definition of "Discursiveness" in (169).

(168)   Maximal Contrast Null Argument Hypothesis (final version)

Null arguments are more likely to be used co-referentially when they are in utterances that are Discursively Maximally Contrastive (that is, within the discourse – linguistic context); co-referential overt subjects are more likely otherwise.

(169) Discursiveness definition

A given property is discursive iff it refers exclusively to the shared linguistic common ground which the speaker knows the addressee is aware of.

The hypothesis in (168) predicts that co-referential null subjects would be preferred in Maximally Contrastive contexts in which the antecedent is discursively retrieved. The likelihood and the acceptability of co-referential null subjects are lower otherwise. This explains the decline in acceptability of the experimental data in the group of speaker who do not interpret null subjects as generics. There is an underlying assumption that must be properly spelled out: why is the Maximally Contrastive context with a forced co-reference to a non-expected antecedent the least acceptable sentence? As proposed previously in Colonna et al. (2010, 2012, 2015, 2014), Baumann et al. (2014), de la Fuente et al. (2016), speakers of any language process the sentences in a parallel fashion, taking into consideration not only the *in presentia* elements, but also the possible alternative formulations that are not used. They compute each word excluding alternative formulations. Since the co-referential null subject in the non-optimal conditions is less likely, the possibility that its antecedent is in the previous adjacent clause is also less expected. In the optimal context for the co-referential null subject, there is a strong violation of the expectation, so that an overt subject would be almost obligatory (using a "division of labor" strategy). In the less likely context, the null subject is less expected and so is the alternative formulation with an overt pronoun to co-refer to the less salient antecedent.

Taking into consideration that in the non-Maximally Contrastive context, the co-referential null subject is not discursively linked to the common ground, the generic interpretation should be triggered. To make clear the proposal about the generic interpretation, the Non-Maximally Contrastive Null Arguments Interpretation Hypothesis is repeated once more here in (170), along with the definition of Intra-Discursively Defined in (171).

(170) Non-Maximally Contrastive Null Arguments Interpretation Hypothesis (final version)

If a null argument is in a non-Maximally Contrastive utterance, its interpretation is more likely to be intra-discursively defined.

(171) Intra-Discursively Defined

Semantic content is intra-discursively defined iff it defines a hypothetical set of possible

worlds without recourse to the common ground, to the physical world or to world knowledge.

When the generic interpretation is achieved in non-Maximally Contrastive Contexts, the interpretation of the null subject is defined intra-discursively by the semantics of generic sentences. As observed in the literature and in the corpus research presented in section 3.5, some semantic conditions facilitate the generic interpretation, such as modal verbs, conditional environments, present and imperfect tenses (and imperfect aspect). These semantic triggers are not exclusive for generic null subjects and, for the most part, all types of sentences are easier to interpret as generic given such conditions. The point stressed by Holmberg et al. (2009), according to which the main problem in his model is the lack of general evidence that any generic NP must stay inside an IP or vP, is accounted for in an absolutely straightforward way in the present proposal. Given the semantic principles governing generic sentences, the only especial feature in null subject generic sentences is that the co-referential interpretation must be less likely for the generic interpretation to emerge.

Null vs. Overt Strategies An important point highlighted throughout this chapter is that different languages achieve the generic interpretation for the null subject by means of different strategies. Both so-called non-pro-drop languages and standard pro-drop languages usually resort to some kind of overt strategy to produce the generic interpretation. However, while non-pro-drop languages seem to have a lexicalized way of expressing this meaning, such as *one* in English, *on* in French and *man*/*men* in other languages such as German, Danish, Frisian, *inter alia*, standard pro-drop Romance languages seem to recycle the multiply functional SE construction for this interpretation, just as standard pro-drop Slavic Languages recycle the reflexive *sebja*. Although the details of such languages are beyond the scope of the present dissertation, a point that must be stressed is that the so-called partial pro-drop languages seem have this interpretation for null subject because of the tight restrictions on the use of co-referential null subjects. In BP, Maximal Contrast seems to constrain the use of co-referential null subjects to very specific contexts. Out of these optimal conditions, the null subject is not grounded in the discourse. Other languages, such as Russian, seem to achieve this interpretation by the fact that quirky null subjects are almost never co-referential (even in "control" environments) (see section 3.3.2). Because of this, they are not grounded in discourse, as in BP. With a more flexible co-referential null subject system, such as in

"standard" pro-drop languages, this interpretation is rarely available, which leads speakers to mark the generic interpretation overtly, usually with the least specified form available in the language.

3.9   Interim Summary

In this chapter, apparent counter-evidence to the approach proposed in the present dissertation was analyzed: the generic interpretation of null subjects in BP. This is taken as a fundamental piece of evidence in the general theory of null and overt subjects, because it is claimed in the literature that languages like BP allow these interpretation while many other typologically different languages resort to some overt strategy to achieve the generic interpretation. Throughout this chapter, it was argued that generic null subjects conform to the standard semantic structure of generic sentences and that the generic interpretation is achieved when the sentence fails to be discursively grounded either in the previous discourse or in the world. Consequently, the generic interpretation is intra-discursively established.

Contrary to previous approaches to generic null subjects in BP or elsewhere, no especial syntactic derivation is proposed. The constructions that favor the generic interpretation of null subjects are exactly those which favor generic interpretations of other sorts of NPs. Some of these constructions are frequent in the corpus, such as modal verbs, conditionals and conditional-like juxtapositions. However, even in these environments, the interpretation of null subjects is often ambiguous, and the generic interpretation depends on the low likelihood of the co-referential interpretation, along with some extra-linguistic factors. In Experiment 5, it was shown that the non-optimal context for co-referential null subjects either triggers the generic interpretation or significantly decreases the acceptability of the relevant tested materials. The interpretation seems to be modulated by extra-linguistic factors such as age and time spent reading on-line (which could possibly be related to the participant's degree of literacy). Modulo these extra-linguistic factors, the approach proposed in the present dissertation deals with the generic interpretation straightforwardly, without additional assumptions.

Chapter 4

Co-referential and Generic Null Subjects and the Theory of the Optimal Interpretation – On-line Evidence on Semantic and Discourse Processing

4.1 Introduction

This chapter aims to provide on-line evidence for the intra-discourse behavior of null subjects in BP, as well as addressing questions that remained open in Chapters 2 and 3. As argued in previous chapters, null subjects are likely to co-refer with highly obvious, less specific, less animate antecedents, especially when such anaphoric subjects are in highly contrastive contexts. Otherwise, acceptability rates decrease or, in appropriate semantic contexts, the generic interpretation emerges. These constraints are thought to be a way of optimizing interpretation in the sense of making the contribution to the common ground as informative as possible, in the most economical way with respect to form. That is, null subjects are constrained to optimal contexts of discourse co-referentiality, while overt subjects are more informative and thus more likely to refer to either discursively non-obvious or deictically established antecedents. Since the generic interpretation seems to be produced locally and intra-discursively, null subjects can be generic when no obvious antecedent is available and when the context is not Maximally Contrastive (in other words, when the context is not optimal for co-reference) by failing to be grounded. Moreover, the generic interpretation is preferred over non-obvious co-reference for being more informative and constraining the possible worlds more than the co-referential interpretation. In section 2.8 of Chapter 2, evidence reported in Chapters 1 and 2 is put together and discussed in order to model the knowledge of speakers about the use and interpretation of null and pronominal subjects in present day BP. This proposal is extended in Chapter 3 to account for generic null subjects. In the present chapter, the predictions and extensions made in the previous chapters are tested and compared to other previously mentioned accounts. Differently from previous accounts, here constraints on null subjects are not taken to generate ungrammatical or grammatical sentences, but rather to increase the difficulty of interpreting the null subject according to specific contexts.

All results discussed so far in Chapters 1, 2 and 3 constitute final interpretation preferences and acceptability judgments on the whole sentence, apart from the data from corpora which are essentially constituted of production data. However, in this chapter, three eye-tracking while reading experiments are reported in order to further study some properties of co-referential and generic null subjects in BP in an on-line fashion. The crucial measures were reading times [RT] in crit-

ical regions [CR] and post-critical regions [post-CR].[1] Here the basic premise is the "eye-mind assumption" (Carpenter & Just 1977, Just & Carpenter 1980), according to which readers maintain fixation on a word or region until processing is completed and this includes processes such as, for example, word recognition, syntactic parsing, semantic integration, referential integration, among other cognitive tasks involved in reading and comprehension. Following Rayner (1998), three ways of analyzing the data were taken into consideration in the present research: the first-pass total RT (the initial reading of a region starting from its first encounter until the region is exited to the left or to the right); the regression path RT (first pass RTs plus rereading of earlier regions until the a fixation lands to the right of the critical region) and total RTs (the sum of all fixations on the region). The methodology used in these experiments is discussed in detail in section 4.3 below. In what follows in this section, a brief discussion of measures and expectations is summarized and the three experiments and the findings are briefly outlined (beyond the literature cited below, see also Sturt 2003, who investigate different measures in the context of reflexive biding).

First Pass RTs. One of the main issues in eye-tracking while reading experiments is how to analyze the eye movement record to understand language processing. This issue concerns how to measure the processing demands associated with a given region of a sentence. According to Rayner (1998), when the region is longer than one word, the total first-pass RTs is generally used as the primary measure of interest. Boland (2004), however, suggests that when an effect observed in the total RTs is also found in the first fixation RTs, it especially shows that the relevant factor affects initial processing of a certain region. According to her, some issues may be observed when the various dependent measures do not all exhibit the same behavior: (i) do some possible constraints influence structural ambiguity resolution, but not the initial production of parsing alternatives? (ii) which factors guide the structure(s) that is/are initially interpreted/built and how difficult is this? and (iii) are phrase structure, morpho-syntactic, and semantic operations sequentially ordered or

---

[1]   CRs are taken to be the verbal region in Experiments 6 and 7, since, only when the participants come to the verb, they can integrate the overt subject into the derivation or undertand that the verb does not have an overt subject, in which case they have to attribute a proper semantic interpretation to the null subject. As much as possible, expectations were cancelled out by other sentences (in Experiment 7, for example, by subject relative clauses that were not analyzed here). Finally in Experiment 8, the critical region is the disambiguation area, so that no concern about the integration of overt subjects are at stake.

simultaneous? Maia et al. (2016) summarize two positions about these questions: (i) syntactic processing has priority over other factors that can potentially influence the parsing of a given structure (Frazier 1987, Frazier & Flores d'Arcais 1989, Maia 2014, Clifton & Frazier 1989); and (ii) all factors exert their influence at the same time on the parsing of a given sentence (e. g. McDonald et al. 1994, McRae et al. 1998). It is common in reading studies to contrast early (first pass RT) effects with later (regression path RT) effects. In this vein, researchers aim to use the multiple measures offered by the eye-tracking experiment to disentangle early structure-building processes from later processes that make use of information incorporated late in the process.[2] Boland & Blodgett (2001), for example, claim that lexical and syntactic constraints influence the initial structure parsing of a sentence, as each new word is integrated into the developing sentence syntactic analysis. They argue that constraints from higher levels of representation (e. g. semantic or discourse structure) can influence syntactic ambiguity resolution at a later point when multiple alternative structures are generated.

Regression Path RTs. Given that processing of a CR can spill over onto a post-CR, it can also be worth analyzing the regressions back into a CR (*i. e.* after having first exited the CR to the right). It is commonly assumed that regressions into a region reflect delayed processing and they are therefore considered a late measure, while first-pass RTs are associated with early effects (Pickering et al. 2004). Traxler et al. (1996), for example, found that animacy has an effect in attenuating the difficulty of reanalysis after a garden path in object relative clauses. This pattern was considered a tendency in the first pass RTs, and was stronger in the other measures (e. g., the likelihood of a regression and total RTs). Studies and results like these suggest that increased processing demands are reflected in longer fixations, secondary fixations in the difficult region, and regressions to earlier segments. Frazier & Rayner (1982) found also that regressions can reflect the fact that a reader went through a word that shows that a previous interpretation was not correct: readers often make a regression as they encounter disambiguating information. The analysis of regression RTs could

---

[2]   Importantly, researchers often analyze contrasts in different measures as reflexes of divergent cognitive properties, depending on their theoretical assumptions. Some researchers have, for instance, claimed that lexical information is not part of the initial parsing of a sentence (e. g. Mitchell 1987); while others argued for lexical information building the basis of the initial structure process (e. g. McDonald et al. 1994).

be taken to be uncontroversial. However, there is some disagreement about how to compute RTs in a given region when regressions are at stake (see the debate between Altmann 1994, Rayner & Sereno 1994*a*,*b*,*c*, for example). For example, Rayner & Sereno (1994*c*) noted that, when readers enter a region and then quickly make a regression out of that region, the first-pass time is very short in comparison to when the reader does not regress. How best to analyze such regressions is uncertain. It is common to simply record the first-pass RT and then examine the pattern of regressions out of the CR (in the post-CR, for example). Konieczny et al. (1994), however, argued for regression-path duration analysis (see also Liversedge et al. 1994, for further discussion of these issues). In the present dissertation, a position closer to Frazier & Rayner (1982), Traxler et al. (1996) is taken: regression-path RTs are taken to be strongly associated with disambiguating regions and lexical and semantic information that can help the reader to solve the possibly wrong parsing, *i. e.* garden path effects.

Total RTs. It is generally accepted that total RTs are affected by all sorts of extra cognitive demands associated to reading a given area, from lexical frequency to high level pragmatic inferential knowledge. Keeping factors controlled would allow researchers to test specific effects that should come up at least in the Total RTs, if they really reflect an extra cognitive effort for processing. One caveat about this idea is, however, necessary. Maia et al. (2016) argue in their study of the effect of islands on wh-extraction that semantic information about the plausibility of a sentence is processed after syntactic constraints. Their hypothesis was that the readers would not show the filled-gap effect nor the effect of the semantic implausibility when they first visualize the island structural context. The predictions were that, since wh-extraction is not allowed in island-contexts (see section 4.5 below), the reader would process this structural information faster and the gap would not be postulated. In the same vein, the processor would not deal with the plausibility of the sentence in such a context, which would be previously blocked by the syntactic analysis. These predictions are indeed realized in their experiments, reproducing previous experiments about island effects in other languages (Stowe 1986, in English, Bourdages 1992, in French and Maia 2014 in a self-pace reading experiment in BP, but contra Traxler & Pickering 1996 in English). This sort of experiments shows that some effects can precede other effects, which are in fact sequentially computed by the reader. Given the expected effect of increase in RTs, it is assumed here that such

primary effects can impede more complex effects of emerging even in the total RTs.

Outline of Chapter 4

In section 4.2, the account proposed in the present dissertation is outlined and the open questions from the previous chapters are recapitulated. In section 4.3, methodological considerations are briefly presented. In section 4.4, the first remaining issue from the previous chapters is addressed: a Maximal Contrastive context in which the null subject should potentially be unacceptable. In section 4.5, the puzzling case of relative clause islands is summarized and analyzed. In section 4.6, one more puzzling case is addressed: subject free relatives are compared to relative clauses in BP and to free relatives in another language. Finally, general assumptions about the proposal presented in this dissertation and its comparison with previous approaches with respect to data in this chapter are discussed in section 4.7.
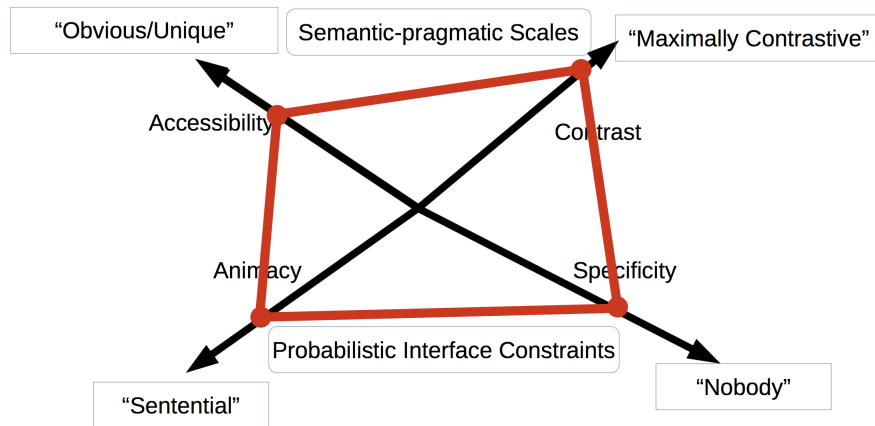
4.2 Summing up the Proposal and Getting Back to Open Questions

The results of the experiments, along with the corpus research, presented in Chapters 1 and 2 constitute evidence in favor of the predictions of the account proposed in the present dissertation. Also, the extension of the proposal, which was presented in Chapter 3, gives further evidence in favor of the account presented here. In what follows, the main claims of the approach proposed here are recapitulated, and the remaining problems examined in this chapter are presented.

The main assumptions in the present approach are the following: (i) typologically, null subjects are licensed in BP, but more restricted regarding their interpretation than in standard pro-drop languages (such as Italian, Spanish and European Portuguese); (ii) the constraints on the co-referential interpretation of null subjects are multiple and the form-meaning mapping is probabilistic; (iii) the relevant semantic and discursive constraints are scalar, rather than binary, and at each point of the scale there is a relative probability of using a null subject; and (iv) the final probability of a given co-referential null subject is a result of each mapping scalar probabilistic mapping constraint, being sketched and depicted as the area of the polygon in the Figure 2.8, repeated here as Figure 4.1

below.

Figure 4.1 – A Mapping Constraint for Null Subjects in A Multiply Constrained Probabilistic Scalar Grammar

"Obvious/Unique"   Semantic-pragmatic Scales   "Maximally Contrastive"

Accessibility   Contrast

Animacy   Specificity

Probabilistic Interface Constraints

"Sentential"   "Nobody"

The most exocentric point of the scheme in Figure 4.1 above is the optimal context for a co-referential null subject. Such a context was presented in section 2.8.2: when a null subject is in a Maximal Contrastive context and the antecedent is obvious, non-specific and inanimate. By changing these conditions, going toward the endocentric points of the scales, the likelihood and the acceptability of the use of a null subject decreases and other possible interpretations can come up given proper conditions. One of these possible interpretations is the generic inclusive one. Generic null subjects are possible in BP exactly as in other contexts in which generic interpretations are achieved for NPs (non-episodic contexts, for instance). So, the existence of such an interpretation in BP is not a specificity of null subjects themselves. For this reason, in Chapter 3, it was claimed that it is a by-product of non-optimal conditions for co-reference along with general properties of the generic interpretation, mainly when the reference fails to be grounded. Given that generic interpretation is a kind of intra-discursive property, it is not surprising that such interpretation emerges in non-optimal contexts for co-reference as a product of the "division of labour" between null and overt subjects. As proposed by Baumann et al. (2014), Colonna et al. (2014), de la Fuente et al. (2016), the computation of interpretation resorts to several mechanisms, in which the interpreter takes into consideration not only *in presentia* elements but also all sorts of possible alternative formulations and interpretations that are not uttered but could be. In this vein, given the preference of null subjects to refer to intra-discursive referents and the semantic conditions for genericity, the generic interpretation is the natural preferential interpretation when the "division of

labor" comes into play (overt subjects tend to refer to less obvious more specific antecedents).

Given the approach proposed here, the questions that remain unanswered are the following: (i) Is the open-polarity context a requirement for the optimality of the co-referential null subject? And why would null subjects be ruled out in a Maximally Contrastive context in which the polarity of the preceding sentence is not open? In such a context, does generic interpretation emerge as a last resort to save the sentence as a possible reanalysis? (ii) Are object relative clauses a context in which co-referential null subjects are ruled out and only the generic interpretation is available? If so, why is this context ruled out: under a bound variable analysis (Modesto 2000*a*,*b*, 2008*a*,*b*), it is predicted to be fine, since there can be no constraint preventing the antecedent from raising to a topic A-bar position, where it can scope over and bind the null subject in the relative clause; under a movement-based analysis Ferreira (2000, 2004), Rodrigues (2004), Sheehan (2006), the derivation would crash, since the CP is taken to be a phase and thus unable to be an input for the next required steps of the movement derivation; under the approach proposed here, the null co-referential subject is not in the optimal context to be interpreted, so that either it increases processing load or the generic interpretation emerges as a *division of labour* product; and (iii) are free-relatives similar to object relative clauses as regards the unacceptability of null co-referential subjects? As previously observed in the literature about topic drop languages, such as Chinese, free-relatives are a context in which the relative pronoun can be interpreted only as the direct object and the null subject can be interpreted only as generic. If so, when the relative pronoun is forced to be interpreted as the subject of the clause, these sentences must be harder to process; additionally, when the remaining gap (the null subject) resulting from the interpretation of the relative pronoun as an object is forced to be co-referential, and not generic, an increase in processing effort is also expect; on the other hand, the literature and the corpus research presented here are abundant in natural examples of object free relatives with co-referential null subjects. How can this difference be accounted for in a model for null subject use and interpretation?

In order to provide an answer to the questions above, in the present chapter, the results and findings of Experiments 4 and 5 are replicated through a slightly different psycholinguistic technique that is extremely suitable in the study of moment-to-moment pronoun interpretation preferences and sentential acceptability: eye-tracking while reading experiments. Section 4.3 below begins by providing a brief description of this methodology followed by the description of the equipment

used and the participants that took part in the experiments.

4.3   Methodological Considerations

When a text is being read, readers move their eyes over the lines through an alternating pattern of fixations (the stationary points where a given word or phrase is focused) and saccades (points at which the eyes are moving from one segment to the other). As mentioned in section 4.1, a very basic assumption of the eye-tracking methods literature is that higher processing demands are linked to increases in fixation time and in the pattern of fixations (in both fixations and saccades). These increases generally reflect difficulties in the processing of a given part of a sentence or text. According to Rayner et al. (2001), it has been demonstrated that certain types of experimental manipulations primarily influence the length of a saccade (the decision about where to look next) while other manipulations primarily influence the duration of a fixation (the decision about when to move the eyes)(see also Rayner & Pollatsek 1981). In particular, "low level" (*i.e.*, non-linguistic) variables such as word length and the spaces between words are the primary sources of information used to determine *where* to fixate next, while the difficulty associated with processing the fixated word primarily influences *when* to move the eyes.

Eye-tracking while reading experiments are taken to be one of the most precise methods to study the cognitive processes involved in sentence comprehension. They allow researchers to measure moment-by-moment (online) processing demands during sentence interpretation. Cognitive processing demands are reflected in several aspects of eye movement behavior, such as fixation duration, number of fixations, and number of regressions, duration of the regressions, etc, as mentioned before. Changes in local difficulty also affects several measures, such as reading times for a given critical region, the probability of fixating words, and likelihood of making regressions to specific regions. The first main advantage of eye-tracking experiments over acceptability judgments is that the latter do not provide detailed measures of processing and interpretation performance in specific regions, specifically, the regions where the sentences vary across conditions. The second reason for using this methodology is that eye movements are a natural part of reading; therefore, no additional task demands are placed on a reader (that is, no metalinguistic task is involved in the outcomes of the experiments). Third, there are a number of measures which may be analyzed (e.g.

fixation duration (RTs), regression frequency, etc), so that different elements of the processing and interpretation process may be taken into consideration. Finally, eye movements directly reflect processing demands given by features of the stimuli being read. For example, eye movements vary as a function of word frequency, word length, lexical ambiguity, contextual constraints and acceptability of the utterance, making up an accessible way of measuring cognitive complexity and difficulties.

Equipment

Participants' eye movements were recorded using an EyeLink® II head-mounted eye-tracker at a sampling rate of 500 Hz on a 21" screen (1024 x 768 px. image resolution). Viewing was monocular, that is, only the participant's dominant eye was tracked (pupil-only tracking).

Each trial was started by a fixation prompt in the left side of the screen that the participant had to fixate, which triggered the drift correction and allowed the presentation of the following stimulus. Each stimulus presented a sentence or sequence of two sentences and, in some cases, was followed by a question about the interpretation of the relevant subject. The interpretation questions were always closed questions, whose answer could be either *sim* "yes" or *não* "no". After choosing one answer (with the left or right button of the joystick), a new trial started.

Participants

Forty Brazilian Portuguese native speakers took part in the following experiments in two sessions. They were thirty-five female and five male, averaging 30 years old (ranging from 24.5 to 36.5). The completion of the experiment, which included a linguistic background questionnaire took from 45 minutes to 1 hour in the first session and from 30 to 45 minutes in the second. Participants were recruited through email lists and Facebook. They were paid eight euros per session in exchange for their participation. Two participants were excluded from subsequent analyses due to the fact that their fixations fell for the most part outside the interest areas. A table summarizing participants' information is given in Appendix C. Eye movement recording was done of the participant's dominant eye only, which was determined by means of a Miles test (Miles 1930). Procedural recommendations related to preparing the participant and setting up and calibrating the equipment

271

were given in the very beginning of the sessions. All the standard procedures regarding ethics, equipment adjustments and environment protocols were followed (Raney et al. 2014). Finally, participants that consistently did not interpret the subject as generic in the relevant conditions in each experiment were excluded in the final analysis (no more than five participants were left out in each experiment).[3]

Data and Inferential Statistical Analysis

In this chapter, the main outcome considered was RTs. Other measurements can be taken into consideration for further analysis, but for the present dissertation the relevant data are the RTs with respect to increasing difficulty of processing given certain conditions. Only the significant effects in RTs will be presented in detail. The main focus of the analysis were the CRs and post-CRs. For Experiments 6 and 7, they were the post-subject verbal area. As for Experiment 8, the CR was the disambiguation area, since a garden path effect was involved.

The regression models presented in this chapter are those in which the residuals were closest to a Gaussian distribution. Mainly, statistical regressions were carried out with the raw data (RTs), with the logs of the raw data (logs of RTs) and with the residuals of a model considering only the Length of the relevant region as a Factor. With the exception of the results of Experiment 7, all the models contained only random intercepts for Participants and Items (Baayen et al. 2008).[4] Before starting the analysis, outlier RTs below 100ms and out of the normal distribution of the data were excluded. Data points that were three standard deviations above the measure mean of the relevant region were also removed from the data set.

---

[3] A fifth of the experimental sentences were followed by a closed-question task, whose demanded information was the referent of the relevant subject (see Appendix B). The questions were mainly included to encourage attentive reading and to check whether participants interpreted some sentences as generic. They are not analyzed any further.

[4] As in previous chapters, the statistical analyses in this chapter were linear mixed models (carried out with the lmer function) including Items and Participants as random slopes. In some experiments the addition of Items and Participants as random slopes did not converge, probably due to the number of conditions, and, thus, were removed. These models are thus intercept models, not including maximal random structure.

4.4 Addressing (Open) Polarity and The Role of (Fronted) Adjuncts

In this section, the issue of open polarity is addressed along with the position of adjuncts and discourse contrast. Especially, the context of non-open polarity is addressed with respect to the Maximally Contrastive Null Argument Hypothesis [MasCoNAH], which was proposed and reformulated throughout Chapter 2. Since the main factors at stake here are the discourse and syntactic structures, other factors, namely Animacy, Specificity and Obviousness, are controlled (all of the possible referents were animate and specific proper names and they were also the closest antecedent to the (anaphoric) subject). The main controversial context is example (49) reported in section 2.3.3 in Chapter 2, repeated below as (172).

(172) a. Dizem   que o   João$_1$ não fala   francês, mas *_$_1$/ele$_1$
         Say.PRS.3PL that the João not speak.PRS.3SG French, but he
         fala.
         speak.PRS.3SG
         "They[indet] say that John$_1$ doesn't speak French, but he$_1$ does."

(adapted from Holmberg 2007, 215, ex. 4d)

According to Holmberg (2007), example (172) cannot be generated by IP-ellipsis, since the polarity is not open in the antecedent clause. This would lead to the unacceptability of (172) without an overt subject, given that in this context there is no other way of licensing the null subject according his approach. The post-comma sentence starting with an overt subject would be produced by standard VP-ellipsis. Indeed, the MaxCoNAH predicts the second clause to be acceptable without an overt subject, given a proper context for this sentence to be uttered.[5] In such a specific example, however, there are a number of factors that can intervene in the judgment of the sentence as a whole, given the fact that the first subject is a non-specific exclusive generic one, the negative sentence is embedded in a reported speech context, the second clause is ambiguous

---

[5]   In fact, in all partial or not pro-drop languages that I am aware of (BP, Spanish, Russian and Latvian), my informants and I judge this sentence odd when a null subject is used. Therefore, there seems to be a block on null subjects in this context, even in more permissive languages, such as Spanish and Latvian. This makes a point against the IP-ellipsis analysis, since this languages allow pro-drop out of IP-ellipsis contexts.

between an arbitrary interpretation of the null object (meaning "but he speaks too much") and a VP-ellipsis interpretation and, finally, the present tense is used to make two general contradictory statements about the same referent. Adding, for example, *eu acho que* "I think that" between the coordinate conjunction and the clause would increase the acceptability of the sentence by making it pragmatically more acceptable and making the polarity opposition clearer. As for the fact that judgments on this example are not clear, an eye-tracking while reading experiment can provide further data about similar contexts and how they must be addressed by the theory of null subjects in BP.

In Experiment 6 below, a slightly different context is proposed. First, for controlling the possible interference of the non-specific subject in the main clause, only proper names are used, avoiding the imbalance of specificity that could trigger a preference for this antecedent. Second, an adjunct phrase is added to the second clause in order (i) to make the polarity opposition pragmatically more plausible and (ii) to assure that IP-ellipsis is not a possible mechanism to produce these sentences. The pro-drop approach proposed by Holmberg (2005, 2007) is thus tested and contrasted with the MaxCoNAH proposed in Chapter 2. None of the sentences have open polarity, so IP-ellipsis cannot be a possible explanation for the difficulties in processing any condition. Before going over the materials, one important prediction of the MaxCoNAH must be stressed and spelled out: the role of fronting in the acceptability and interpretation of null subjects. This prediction is crucial, since it is the point where many other approaches and the approach proposed here converge and diverge at the same time as explained in the next subsection.

4.4.1 Fronting Again

In Chapter 2, it is argued that null subjects are more likely in Maximally Contrastive contexts. Non-Maximally Contrastive contexts, but still Completely At-Issue utterances, can appear in a variety of forms. One of them is sentences with fronted adjuncts that are at issue in the current QUD, but are backgrounded, as in (173) below.

(173)   a.   Em que   época do   ano   a   senhora esteve   na   Espanha?
            in   which time   of.the year the ms.   be.PST.3SG in.the Spain
            "In which time of the year have you been to Spain?"

b. Na   Espanha, eu estive   em setembro   ...
   in.the Spain,   I   be.PST.1SG in   September
   "In Spain, I've been in September."

(NURC-RJ, "Inquiry_267")

In (173), the fronted locative phrase is completely at-issue in the current QUD. However, the repetition of the adjunct as a strategy to restrict the scope of the focused material produces further inferences that the sentence is not Maximally Contrastive (either by recovering the topic of the conversation or, in the case at stake, by contrasting the topic with other possible locations included in the shared background, which were in QUDs in previous turns). Notice that fronting the adjunct phrase in (173) seems to force the use of an overt subject, even if it is obvious and the topic of the current QUD. In Chapter 3, it is argued that the fronting of an adjunct either triggers the generic interpretation or decreases the acceptability of the sentence in which the null co-referential subject occurs, by turning the clause into a non-optimal context for co-reference. In discourse-logical terms, it implicates a set of assertive alternatives whose backgrounded content is not the same; it rather produces a set with different backgrounded structures, whose adjunct specifies the chosen one (as in Büring 1999, 2003's QUD trees, etc).

In Experiment 6, a fronted or *in situ* adjunct is once more taken into consideration for testing the MaxCoNAH. By fronting the adjunct, it is expected that either a generic interpretation comes up (in cases in which the semantics of the sentence allows the generic interpretation) or a difficulty in interpreting the co-referential null subject in a non-optimal context would be found. By leaving the adjunct *in situ*, a Maximally Contrastive sentence is produced, and thus the null subject perfectly matches the optimal context for co-reference. Overt subjects are used as a baseline, since they would probably not be less acceptable in any condition, even if they might produce a slightly different pragmatic effect.

4.4.2   Experimental Evidence – Experiment 6

In this experiment, the semantic contribution of the adjunct is addressed, as well as the interpretation and use of null and pronominal subjects in a closed polarity context. First, in all contexts, the clause in which the null or overt subject appears has a different polarity from the clause where the

potential antecedent is referred to, but it is never open (in the sense of Holmberg 2001, 2007, 2016). Second, the adjunct is supposed to be part of the focus, when it is *in situ*, but part of the background when fronted. This difference can be formally explained in terms of D-trees, as in Büring (1999, 2003), and here is formally described as a non-Maximally Contrastive context. Third, as proposed in Chapter 3, modal verbs suspend the presupposition of existence and so their subjects can acquire reference intra-discursively in the contexts where the co-referential interpretation is not likely. Finally, null and overt subjects are tested in similar semantic-discursive conditions in order to compare the complexity of processing of each of them. The three way interaction of Factors approached significance and suggests that in non-optimal conditions the co-reference requires an extra effort from the interpreter.

Material Design

Twenty-four items were created for this experiment, cf. the sample in (174). As mentioned before, only proper names were used (half of them masculine), to avoid any imbalance that could possibly bias the interpretation of the anaphoric subject. As in Holmberg (2007)'s original example, the sentences started with a main clause that introduced an embedded clause with the potential antecedent of the anaphoric subject. The second sentence started with an opposition conjunction (*mas* "but") and finished with a reassertion of the reporting event (*ele/ela disse/falou* "(s)he told/said") in order to have a post-CR. The manipulation of Factors conformed to the following distribution: in sentences (a), the event is episodic and the same as the main verb of the antecedent clause, while in (b) the verb is modal and different from that described by the verb of the main clause (in a possibly intransitive use of the verb) (Factor "Verb"); in (i) the adjunct is *in situ* at the end of the clause, and in (ii) the adjunct is fronted right after the conjunction (Factor "Adjunct"). Within the manipulation of both Factors, the subjects of the second clause were either null or overt (Factor "Subject").

(174)   a.   i.   A   Gisele falou   que o Felipe$_1$ utiliza   tango$_2$ em performances
                   The Gisele say.PST.3SG that the Felipe use.PRS.3SG tango in   performances
                   em geral,   mas _$_1$/ele$_1$ |*não tinha*   *utilizado*| _$_2$ nesse show de
                   in   general, but he   not have.PST.3SG use.PTCP   in.this show of
                   talentos, ela disse.
                   talents   she tell.PST.3SG

ii. A Gisele falou   que o Felipe₁ utiliza   tango₂ em performances
The Gisele say.PST.3SG that the Felipe use.PRS.3SG tango in performances
em geral,   mas nesse show de talentos _₁/ele₁ |*não tinha   utilizado*|
in general, but in.this show of talents he   not have.PST.3SG use.PTCP
_₂, ela disse.
    she tell.PST.3SG

"Gisele said that Felipe₁ uses tango in general performances, but in this talent show

he₁ had not used it, she said."

b. i. A Gisele falou   que o Felipe₁ utiliza   tango₂ em performances
The Gisele say.PST.3SG that the Felipe use.PRS.3SG tango in performances
em geral,   mas _₁/ele₁ |*não pode   dançar*| _ nesse show de
in general, but he   not can.PRS.3SG dance.PTCP   in.this show of
talentos, ela disse.
talents   she tell.PST.3SG

ii. A Gisele falou   que o Felipe₁ utiliza   tango₂ em performances
The Gisele say.PST.3SG that the Felipe use.PRS.3SG tango in performances
em geral,   mas nesse show de talentos _gen/ele₁ |*não pode   dançar*|
in general, but in.this show of talents he   not can.PRS.3SG dance.PTCP
_, ela disse.
    she tell.PST.3SG

"Gisele said that Felipe₁ uses tango in general performances, but in this talent show

one/he₁ cannot dance, she said."

The CR in all the conditions is the italicized verbal sequence in the second conjunct (see 174 above), introduced by *mas* "but" (followed or not by an overt subject). In what follows, RTs refer to times measured only in this region. The empirical predictions were the following: based on the standard assumption that RTs (in terms of fixation duration) increase as the processing load, the non-optimal context for co-referential null subjects should present higher RTs than the comparison baseline with the overt pronoun (Condition Episodic Verb:Fronted Adjunct 174a-ii). On the other hand, a co-referential null subject preceding the same verb type with an *in situ* adjunct should be read as fast as or faster than the pronominal version (Condition Episodic Verb:In Situ Adjunct 174a-i). No effect of the interaction between Adjunct position and type of Subject is expected in the Modal Verb Condition, since it would trigger the generic interpretation in the non optimal context, which is as easy to process as the co-referential overt subject, following the results and proposal presented in Chapter 3.

Results

In Figures 4.2 and 4.3, the total RTs in the CR are plotted (respectively, the RTs for the episodic Verb and the RTs for the Modal Verb Condition. As shown by the comparison between the graphs, the CR of episodic sentences were generally read faster than the CR of modal sentences, except for Condition Null Subject plus Fronted Adjunct. Null and pronominal subjects were read almost at the same pace except for Condition Episodic Verb plus Fronted Adjunct. RTs were shorter for the sentences with the *In Situ* Adjunct than with the Fronted Adjunct.

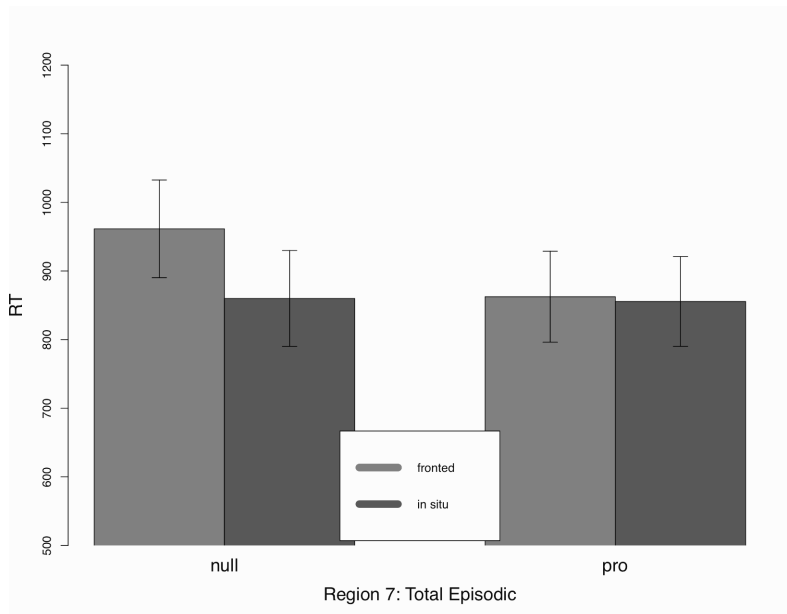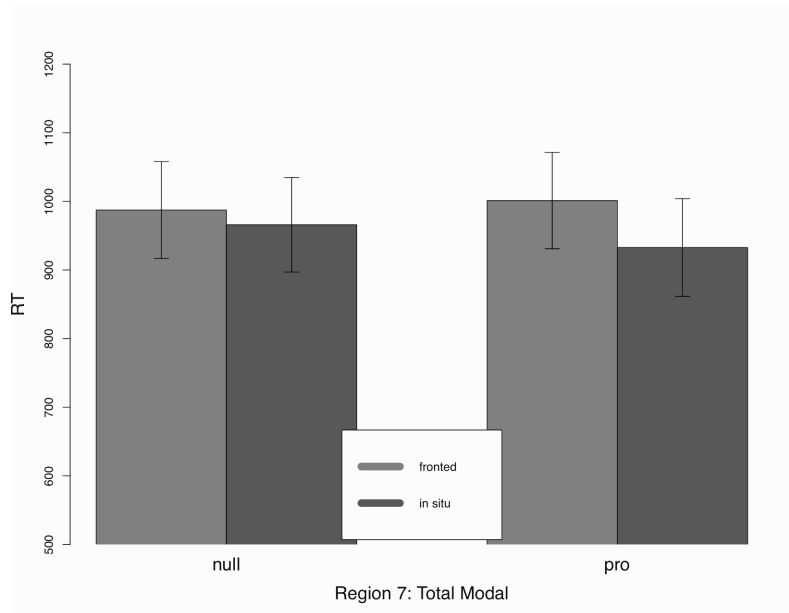Figure 4.2 – Total RTs for the CR ((negation+)verb) in Episodic Verbs as in 174a

Figure 4.3 – Total RTs for the CR ((negation+)verb) in Modal Verbs as in 174b

As for the inferential statistical analysis, the logarithm of the total RT of the CR were regressed in a mixed effect model. Main Factors were Adjunct, Subject and Verb (with two levels each) and random Factors were Participants and Items. Additionally, three possible intervening factors were regressed to ensure that factors observed do not depend on the disposition of characters or participant awareness of the structure: Length (number of characters including blank spaces), Line_Break (some items fell on the margins of the screen, causing a line break) and Session (given that the experiment had eight conditions, the stimuli were split into two sessions). Table 4.1 sums up the results of the mixed effect model.

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	9.071e+00	1.903e-01	47.662	<2e-16 ***
Adjunct	-1.751e-01	6.717e-02	-2.607	0.00925 **
Subject	-1.349e-01	6.599e-02	-2.044	0.04114 *
Verb	1.009e-01	6.565e-02	1.537	0.12442
Length	4.269e-02	9.650e-03	4.424	5.30e-05 ***
Line_Break	2.973e-03	4.038e-02	0.074	0.94131
Session	-1.688e-01	3.297e-02	-5.121	3.49e-07 ***
Subject:Verb	1.464e-01	9.354e-02	1.565	0.11775
Subject:Adjunct	1.441e-01	9.255e-02	1.557	0.11968
Verb:Adjunct	1.989e-01	9.206e-02	2.161	0.03091 *
Subject:Verb:Adjunct	-2.405e-01	1.314e-01	-1.831	0.06736 .

Table 4.1 – Log-linear mixed-effects model for Logs of Total RTs in CR in Experiment 6

First, two intervening Factors were significant: Length and Session. Two main Factors were also individually significant: Subject and Adjunct. Pronominal subjects were generally read faster, as indicated by the negative estimate, as were so *in situ* adjuncts. Adjunct seems to be modulated by the interaction with Verb, which also reaches significance. Finally, the three way interaction approached significance, which can also modulate both the main effects and the interaction between Verb and Adjunct. This is probably due to the higher RTs in the condition in which the null subject is co-referential and the adjunct is fronted.

Discussion

Apart from the intervening Factors that were not surprisingly significant (length naturally affects the RTs, as well as an effect of priming in the second session), Main Factors came up significant. Especially the three way interaction seems to be the main finding in this experiment, since, as seen in the Figures 4.2 and 4.3, RTs are lower in co-referential contexts when the adjunct is *in situ*, in the case in which a null subject is used, or when a pronoun is used. No significant effect was found in the first pass RTs. In the regression path RTs, only the effect of the Adverb is significant, and the effect of the Verb is marginally significant (see Appendix D), which can possibly be due to the unexpected position in the beginning of the sentence causing a regression to previous area to reread the sentence. In terms of total RTs, the general findings are according to the empirical predictions made beforehand, since the highest total RTs were expected to be the episodic co-referential null

subject in the Fronted Adjunct Condition. The fact that overall total RTs are a little higher in the Modal Verb Condition may be due to the interaction with the position of the adjunct, as shown by the significant interaction between Factors Verb and Adjunct. This could also be the case because the verb is repeated in the Episodic Condition, since repeated lexical items are generally read faster. The interaction, however, seems to be modulated by the use of null and overt subjects in the Episodic Condition, a modulation that does not appear in the Modal Verb Condition. Since the mean of RTs are close to each other across conditions when the verb is modal, the difficult to interpret such sentences comparing to episodic sentences is probably due to the different verbs. It is important to point out that, in the modal context, no significant difference between null and overt subjects are found across Fronted and *In Situ* Adjunct Conditions.

4.4.3 Accounting for Fronting

In Chapter 3, the findings of Experiment 5 show that BP speakers tend to interpret null subjects as generics in non-optimal contexts for the co-referential interpretation (modulo speaker to speaker variation). Also, when the generic interpretation is not achieved (due to either intra-linguistic or extra-linguistic factors), the acceptability of the co-referential null subject in non-optimal contexts is lower than in the optimal one (fronted vs. *in situ* adjunct phrases). In Chapters 2 and 3, it was proposed that the optimal context for the co-referential null subject is Maximally Contrastive. By fronting the adjunct, the sentence is no longer Maximally Contrastive and the null subject is not as good as with an *in situ* adjunct. These findings were reproduced in Experiment 6 with an additional caveat: regardless of the closed polarity of the antecedent, the sentences with and without overt subjects were read equally fast apart from the context in which it was not Maximally Contrastive. No such an effect was found, however, when the verb was modal, triggering the generic interpretation in the non-optimal context. The outcomes of Experiment 6 suggest that the MaxCoNAH correctly predicts when the co-referential null subject is more likely to appear (and is thus easier to process). Additionally, the Non-MaxCoNAIH predicts the generic interpretation in the non-optimal conditions for co-reference, in such a way the RTs do not diverge relevantly from those of overt subjects.

In the next section, another Non-Maximally Contrastive context is tested: object relative clauses.

In section 2.7 in Chapter 2, a corpus analysis found that null subjects in relative clauses were rare, and the effect of this syntatic environment showed a significant effect in the logistic regression. If the same hypotheses hold over the whole system of BP, it is expected that once more the interaction between the type of Verb (modal or episodic) and the type of Subject (null or overt) should be significant: the episodic verb with a null subject should be read more slowly than the three other conditions (modal verb with null and overt subjects and episodic verb with an overt subject). However, if the hypothesis presented in this dissertation is correct, the effects of such interaction should be found in the total RTs, which are more likely to be affected by interpretation preferences. Otherwise, first pass and regression RTs, which would suggest either the ungrammaticality of the sentence or a garden path related to multiple gaps, would be significantly different.

4.5   Addressing Relative Clauses

In much literature, BP relative clauses are taken to be a context in which co-referential null subjects are categorically ruled out (Figueiredo-Silva 1994, 2000, Modesto 2000*a*,*b*, 2008*a*,*b*, Ferreira 2000, 2004, Rodrigues 2004, Sheehan 2006). Other researchers have reported that null subjects in this context are extremely infrequent Duarte (1995), Magalhães (2007). The low frequency of co-referential null subjects in relative clauses is also confirmed in the corpus research reported in section 2.7 in Chapter 2. The fact that null subjects are less frequent in object relative clauses, however, does not entails that they are not licensed in this structures. In example (94) in section 2.6.2 in Chapter 2, a null subject is co-referential with the discourse salient non-specific antecedent presented in the question (the relevant turns are repeated in 175 below).

(175)   a.   A – E o professor$_1$?

"A – And what about the teachers$_1$?"

b.   B – $\_1$ Foi   despedido ... $\_1$ Recebe   a   indenização a   que $\_1$
           be.PST.3SG fired   receive.PRS.3SG the restitution   to which
tiver   direito e   pronto ...   e   o   problema social$_2$
have.FUT.SUBJ.3SG right   and ready   and the problem social   be.PRS.3SG
é   criado ...
created

"B – They$_1$ are fired... they$_1$ receive the restitution which they$_1$ have the right to and

that is it... and the social problem$_2$ is created..."

<div align="right">(NURC-RJ, "Inquiry_164")</div>

In (175), the discourse-salient antecedent spans over the three clauses of the answer in (175b). The third clause is a relative clause whose antecedent of the relative pronoun is *indenização* "restitution". Following the relative pronoun, a co-referential null subject starts the clause. This is an absolutely non-expected case for most of the theories about null subjects in BP. In this section, Experiment 7 is designed to check for the effects of relativization on the use and interpretation of null and overt pronouns in cases similar to (175b). The findings in this chapter suggest once more that non-optimal contexts for null co-referential subjects either trigger the generic interpretation or increase the difficulty of the interpretation in case the generic interpretation is not available, but do not make the sentence ungrammatical.

4.5.1 The Theory of Relativization – Island Effects and the Discourse Structure of Relative Clauses

As already mentioned, relative clauses are a context in which co-referential null subjects are dispreferred. Many researchers have claimed that co-referential null subjects are sensitive to the same constraints as fronted elements, such as Wh-elements or topicalized NPs (Sheehan 2006, Rodrigues 2004, Ferreira 2000, 2004), preventing co-indexing from inside the relative clause. This is claimed to be an effect of strong islands in Relative Clauses.[6] In these approaches, co-referential null subjects are supposed to be sensitive to Island Effects (also called, strong cross-over effects) for different reasons, but the gist of these proposals is that some kind of intervention prevents the null subject in this position from being syntactically licensed. Figueiredo-Silva (1994, 2000) observed that object relative clauses are a context where null subjects seems to be categorically ruled out:

---

[6]   The so-called island effect has also been a matter of debate in the recent literature both in theoretical linguistics (Rizzi 1982*b*) and in experimental psycholinguistics (Sprouse et al. 2016, Abeillé et al. in prep).

(176) *A Maria₁ achou   um carro₂ que₂ _₁ tem   grana para comprar _₂.
       the Maria find.PST.3SG a   car   that have.PRS.3SG money to   buy
       "Maria found a car that she has money to buy."

<div align="right">(Figueiredo-Silva 2000, 139)</div>

In the very same fashion, topicalization in BP seems to be unacceptable when the topicalized NP comes out of an object relative clause, as in (177).

(177) *A Maria₁, o   João achou   um carro₂ que₂ _₁ tem   grana para
       the Maria   the João find.PST.3SG a   car   that have.PRS.3SG money to
       comprar _₂.
       buy
       "Maria₁, João found a car that [she₁] has money to buy."

<div align="right">(Figueiredo-Silva 2000, 136)</div>

Moreover, Negrão & Viotti (2000) observe that Wh-extraction from an object relative is also impossible in BP, as in (178).

(178) *Que animais₁ a   televisão mostrou   as criancas₂ que₂ _₁ atacaram   _₂?
       which animals   the TV   show.PST.3SG the children   that   attack.PRS.3SG
       "Which animals₁ does the TV showed the children₂ that₂ they₁ attacked _₂."

<div align="right">(Negrão & Viotti 2000, 121)</div>

Cross-linguistically Object Island constraints are generally taken to lead to ungrammaticality (Uriagereka 1999, Sabel 2002, Johnson 2002, Truswell 2007). This ungrammaticality has been explained in terms of structural constraints on movement or co-indexicality between the antecedent and the gap. Other approaches propose that Fronted-Wh elements and topicalization are restricted by the information structure of the sentences (see Erteschik-Shir 1997, Polinsky et al. 1995, Takami 1989, Van Valin 1995, Goldberg 2013). These two proposals for dealing with such a potential constraint are discussed in section 4.5.3.

Before moving on to the theoretical account of these facts, Experiment 7 is reported in the next section. In this experiment, three hypotheses are tested: (i) whether strong island ungrammaticality effects are observed when a co-referential null subject is used in a ORC; (ii) whether a non-specific

topical antecedent can span over the object relative clause; and (iii) whether a generic interpretation can emerge in the object relative clause. Ultimately, once more, the effects in CR are found only in the total RTs, but not in the first pass RTs nor in the regression path RTs, which is claimed to be explained by the difficulty in interpreting the null subject, but not in licensing it, as predicted by the MaxCoNAH. As in Experiment 6, overt subjects are taken as a baseline for comparison with null subjects. However, if the context is really optimal for null subjects, the overt subjects are expected to be less acceptable since they tend to be dispreferred when retrieving a non-specific antecedent.

4.5.2 Experimental Evidence – Experiment 7

In this experiment, the context of object relative clauses is addressed, as well as co-reference with non-specific antecedents in this context. First, all sentences started with a universally quantified subject, which is the only possible topic of the sentence (as explained in the next subsection). Second, expectations about filler-gap relations are counter-balanced with four other sentences in which the relative pronoun was obligatorily interpreted as the subject of the relative clause, and a null object was used (these conditions are not analyzed in the present dissertation and will be addressed in future work). Third, as in the former experiment, modal verbs in the relative clause are used to possibly facilitate a generic interpretation for the null subject where co-reference is not likely. Last, the main manipulation was the comparison between overt and null subjects in the CR (post-subject verbal region).

Material Design

Twenty-four items were created for this experiment. All of them started with a universally quantified non-specific referent followed by an existential quantified referent and a restrictive object relative clause. A half of the sentences were positive and the other half negative, as in (179) and (180) below respectively.

(179)   a.   i.   Todo jogador$_1$ usufrui   de um esporte que $\_{gen}$ |*pode*   *jogar*|
             Every player   enjoy.PRS.3SG of a   sport   that   can.PRS.3SG play.INF
             com frequência.
             with frequency

ii. Todo jogador₁ usufrui   de um esporte que ele₁ |*pode*   *jogar*|
Every player   enjoy.PRS.3SG of a   sport   that he   can.PRS.3SG play.INF
com frequência.
with frequency

"Every player₁ enjoys a sport that one/they₁ can play frequently."

b. i. Todo jogador₁ usufruiu   de um esporte que _₁ |*tinha*   *praticado*|
Every player   enjoy.PST.3SG of a   sport   that   have.PST.3SG play.PTCP
com frequência.
with frequency

ii. Todo jogador₁ usufruiu   de um esporte que ele₁ |*tinha*   *praticado*|
Every player   enjoy.PST.3SG of a   sport   that he   have.PST.3SG play.PTCP
com frequência.
with frequency

"Every player₁ enjoyed a sport that they₁ had played frequently."

(180) a. i. Nenhum lojista₁   se aborrece   com um cheque que _gen |*não*
No   shopkeeper SE annoy.PRS.3SG with a   check   that   not
pode   *cobrar*| em antecipação.
can.PRS.3SG charge.INF in   anticipation

ii. Nenhum lojista₁   se aborrece   com um cheque que ele₁ |*não*
No   shopkeeper SE annoy.PRS.3SG with a   check   that he   not
pode   *cobrar*| em antecipação.
can.PRS.3SG charge.INF in   anticipation

"No shopkeeper₁ gets annoyed by a check that one/he₁ cannot charge in advance."

b. i. Nenhum lojista₁   se aborreceu   com um cheque que _₁ |*não*
No   shopkeeper SE annoy.PST.3SG with a   check   that   not
tinha   *recebido*| em antecipação.
have.PST.3SG charge.PTCP in   anticipation

ii. Nenhum lojista₁   se aborreceu   com um cheque que ele₁ |*não*
No   shopkeeper SE annoy.PST.3SG with a   check   that he   not
tinha   *recebido*| em antecipação.
have.PST.3SG charge.PTCP in   anticipation

"No shopkeeper₁ got annoyed by a check that he₁ had not charged in advance."

Quantified elements are traditionally taken to be ambiguous in sentences in which one referent can scope over the other, such as in *Todo jogador usufruiu de um esporte.* "Every player enjoyed a sport." (see Sennet 2016, for an overview). There are two possible interpretations for this sentences according to which quantifier (either the universal or the existential quantifier) scopes over the

other. Thus, this sentence can refer either to multiple sports which each one of the players enjoys or to one single sport that the group of players in the context enjoys. Since Chomsky (1976) and May (1977, 1985), these two interpretations are taken to be generated by some sort of movement mechanism either in syntax or in semantics [also known as, Logical Form] (see also Chapter 1, especially the sections 1.7.4).

Further studies show that the information structure status of quantifiers influences their scope. Sæbø (1997) (along with many others Sgall et al. 1986, Kuno 1991, Kuno et al. 1999, Van Valin & LaPolla 1997, Krifka 2016) observes that topical quantifiers have wide scope over focused quantifiers.

(181)   a.   i.   A – What did the candidates attend?

        ii.   B – [Several candidates]$_{Top}$ attended [every MEETING.]$_F$.

         'For several candidates it holds that they attended every meeting.'

   b.   i.   A – Who attended the meetings?

        ii.   B – [Several CANDIDATES]$_F$ attended [every meeting.]$_{Top}$.

         'For every meeting it holds that several candidates attended it.'

(Sæbø 1997, p. 4-5)

In more recent approaches, the scope of quantifiers is mostly captured structurally within a framework that assumes that topics occupy syntactic positions that c-command foci as in the "cartographic" theory (e.g. Rizzi 1997), as mentioned by Krifka (2016), based on facts observed by Szabolcsi (1997) about Hungarian and by Frey (2004) and Ebert (2009) about German.

However, in many psycholinguistic studies, the so-called "surface" interpretation in (181a) is considered predominant (Kurtzman & MacDonald 1993 have found this preference in continuation acceptability tasks, Dwivedi et al. 2010, in an off-line norming pre-test, and Raffray 2010, in a picture priming study). Thus, the preference for the surface interpretation is indeed a robust empirical finding. Also, Sæbø (1997), Krifka (1998a) and van de Koot & Neeleman (2012) observe that the "inverse scope" reading requires particular operations on the information structure, so that the element located more to the right must become a possible topic. Chomsky (1995, 377) also points out that Quantifier Raising (QR) raises the issue of over-generation. Unless the operation is

constrained, it creates readings that are not intuitively available, in particular regarding the inverse scope reading. Finally, the inverse scope interpretation does not seem to be a plausible interpretation for the materials tested in this experiment, either from a theoretical point of view or based on empirical grounds, so that it is almost undisputed that the quantified element has a topical status in these sentences.

A crucial property of the topical quantifiers is observed by Ebert (2009) and Cresti (1995). They suggest that topical quantifiers have the additional property that their anaphoric potential is larger than that of other elements. The wide-scope effects of some indefinites and quantified elements come about as a result of their topicality.

By choosing these quantified NPs, two properties that favor the use null subjects are obtained: (i) regardless of the theory assumed, the referents of these NPs are the topic of the utterance (either at a syntactic level or at a discourse level); (ii) the topical quantified phrases are lower in scale of specificity, being higher than possibly only *nobody* and *everybody*. These two Factors are taken to be relevant in Chapters 1 and 2 and elsewhere in the literature. Also, all the referents were animate, so that no effect of this Factor can be considered relevant. Having controlled these Factors, neither the discourse prominence nor the inherent salience of the antecedent can play any relevant role in the outcome of this experiment. What is at issue in the materials is only the relative clause environment, the interpretation that can be associated with the null and overt subjects and their interaction.

Keeping the controlled facts in mind, the experiment followed a two by two design (two binary Factors): in (179) and (180) above, sentences (a) had a modal verb in the relative clause and sentences (b) had a past perfect episodic verb (Factor Verb); and in sentences (i), the subject of the relative clause was null, while in sentences (ii), it was overt (Factor Subject). Considering this design, the empirical predictions were the following: (i) the RTs in the CR in the Episodic plus Null Condition should be the highest, while with an overt subject the RTs should be close to the Modal Condition; (ii) regardless of Factor Subject, the RTs in the Modal Condition should be similar to the Episodic Verb plus Overt Subject Condition. No main Factor is expected to be significant, but the interaction between both Factors is expected to reach significance.

Another crucial issue is in which measures these effects will be found. The predictions regarding this issue are summarized as follows. If in the object relative clause context null subjects

are not *licensed*, the effects of ungrammaticality should come up in the first pass (Frazier 1987, Frazier & Flores d'Arcais 1989, Maia 2014, Clifton & Frazier 1989). If the main effect is a garden path problem regarding the interpretation of the relative pronoun, given the two gaps, a significant difference in the reading times of the first pass in the Overt Pronoun Condition compared to the Null Subject Condition and a possible difference in the regression path RTs should come up (Frazier 1987, Frazier & Flores d'Arcais 1989, Clifton & Frazier 1989, Boland 2004, Maia 2014, Maia et al. 2016). Specifically, if the theoretical and analytical claims about the effects of ORC islands on null subjects in BP (proposed in most of the literature) hold, it is expected that similar results to those found by Maia et al. (2016) would be found in this experiment; that is, no effects of semantics or pragmatics should appear in the RTs across conditions, since all possible interpretations are blocked when they first get to the island constraint area. However, according to the findings presented in the next section, none of these predictions hold. Since the effects were found only in the total RTs, it is likely that the constraints on one or another condition are due to increased complexity of interpretation of the relevant null subject, rather than licensing. Indeed, the main effect found in this experiment was a significant interaction between the type of verb (either modal or episodic) and the subject (either null or overt) as regards the total RTs of the critical region. These results suggest that the effect of the relative context environment over the use of null subjects is not a structural problem, but rather an interpretation problem. The most likely possibility of using a null subject in a relative clause is to interpret it as a generic.

Results

As shown in Figure 4.4, total RTs of the CR do not diverge too much regarding the use of a null or an overt subject when a modal verb is used. However, when an episodic verb is in the object relative clause, the total RTs of the CR were higher when a null subject was used than when the subject was overt. Also, the mean RTs in the CR in the Overt Pronoun plus Episodic Verb Condition do not diverge from those found in the Modal Verb Condition with either null or overt subjects.

Figure 4.4 – Total RTs for the CR ((negation+)verb) according to Factors Subject and Verb

Region 1: Total RT
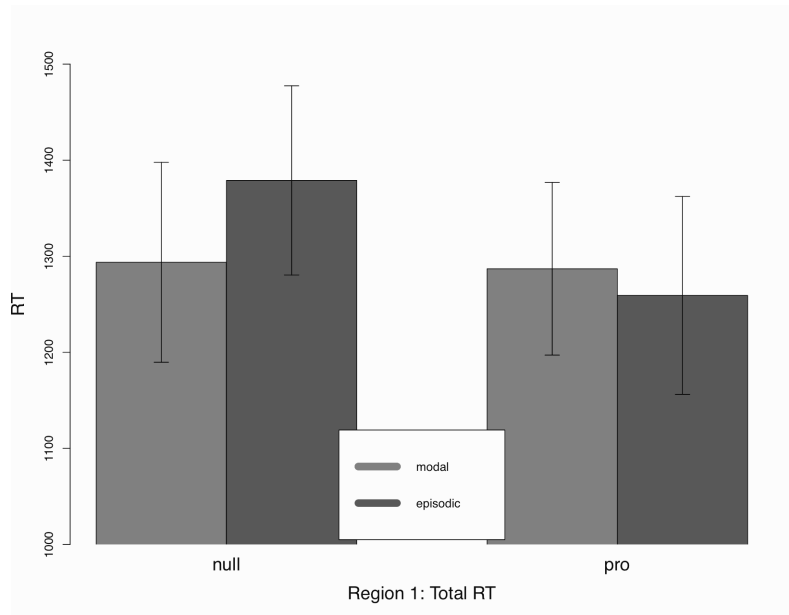
The total RTs in the CR were thus entered in a mixed effects linear model whose main Factors were Subject and Verb and the random Factors were Participants and Items, with maximal random structure (Barr et al. 2013). Additionally, Length and Session were also entered as possible intervening Factors. The inferential statistical analysis is summarized in Table 4.2 below.

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	9.145838	0.270908	33.760	<2e-16 ***
Subject	0.008365	0.077863	0.107	0.9150
Verb	0.078972	0.074309	1.063	0.2949
Length	0.058161	0.013655	4.259	0.0001 ***
Session	-0.159553	0.064781	-2.463	0.0188 *
Subject:Verb	-0.184967	0.086289	-2.144	0.0348 *

Table 4.2 – Log-linear mixed-effects model for Logs of Total RTs in CR in Experiment 7

As presented in Table 4.2, intervening Factors Length and Session were significant (Length increases the RTs, and in the second Session they were decreased). No main Factor reached significance, but the interaction between both main Factors is significant.

In Figures 4.5 and 4.6 below, the residuals of a model whose predictors were intervening Factors (Length: β:87.685, SD:1.794, T-value:48.877, p-value:< 2e-16 *** ; Session: β:-170.948, SD:24.753, T-value:-6.906, p-value:5.7e-12 ***) are plotted. As shown in these plots below, the

RTs diverge beyond the standard deviation (error bars) in the CR only when a episodic verb is in the object relative clause.

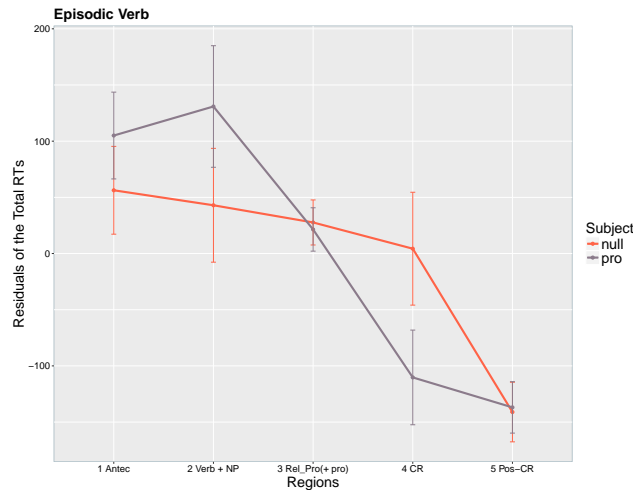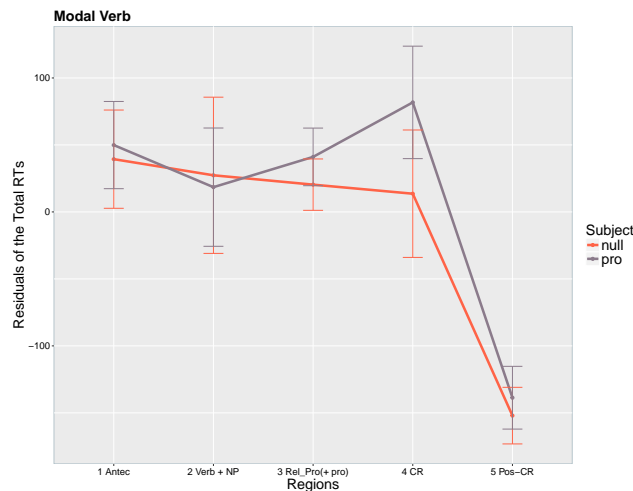Figure 4.5 – Residuals of the total RTs for Episodic Verbs as in (179b) and (180a)

Figure 4.6 – Residuals of the total RTs for Modal Verbs as in (179a) and (180a)

Discussion

The effects of the different conditions were found only in the total RTs of the CR, which suggests, as in Experiment 6, an increased difficulty in interpreting the relevant condition rather the ungrammaticality of these sentences. The main finding is the significant interaction between both conditions (Subject and Verb), which was predicted by the initial empirical hypotheses. As observed in the residuals of the intervening Factor model, there is an inversion in the RT patterns

across conditions, in which null subjects are read faster with a modal verb, while overt subjects are read faster with an episodic verb. These findings are a direct mirror of the relative frequency of null and overt subjects in relative clauses (although it is not possible to insure that frequency plays a role in the increased RTs or the RTs show why overt subjects are more frequent in such constructions). Finally, a difference between the co-referential non-specific interpretation and the generic interpretation is necessary, since they are significantly different in terms of how easy they are processed.

4.5.3 Accounting for Relativization

As mentioned in section 4.5.1, object relative clauses are taken to be islands for extraction. In previous literature, they are considered to be islands either for the movement of the subject to a higher c-commanding position in the clause Ferreira (2000, 2004), Rodrigues (2004), Sheehan (2006), so that it cannot be further deleted, or they are taken to have an intervening operator that does not allow the co-indexing of the antecedent and the null bound variable Figueiredo-Silva (1994, 2000), Modesto (2000*a*,*b*, 2008*a*,*b*). Here, based on the findings of Experiment 7 and the facts and hypotheses presented so far, another account is proposed. Based on the study of the semantics and discourse structure of relative clauses and the information structure-based accounts for islands, it is proposed that such constructions conform to the MaxCoNAH exactly the same way as the other non-Maximal Contrastive contexts studied before.

According to Goldberg (2013), some discourse-level phenomena are sensitive to islands, such as relative clauses. In a context in which the relevant fact to answer the question in (182) is that Shira is wearing a hat, answer (182b-i) seems to be infelicitous.

(182)   a. A – Why was Shira so happy?

        b.   i. B – The woman who thought she was wearing a new hat lives next door.

            ii. B – The woman who lives next door thought she was wearing a new hat.

(Goldberg 2013)

Morgan (1975) claimed that reply (182b-ii) is a felicitous answer to the question in (182) while (182b-i) does not seem to be adequate (see also James 1972 for related observations and the

experimental research by Roland et al. 2012). For accounting for the effect of islands in extraction and related phenomena, Goldberg (2013) claims that backgrounded constituents are islands and cannot be "extracted from", since "extracting" a constituent positions it in a discourse-prominent slot, and it is anomalous to give an element a backgrounded and discourse-prominent position at the same time.

The fact that topicalized and extracted constituents in BP are sensitive to ORC islands could possibly be explained by Goldberg's discursive account of extraction. The sensitivity of co-referential null subjects to a similar constraint is partially a coincidence, because indeed the explanations diverge although the source of the oddness is the same: the discourse status of ORCs. The main claim here is that ORCs are backgrounded parts of the sentences in which they occur by the simple fact that they are not asserted (see, among many others, Abbott 1995, which treats "presupposed content" as not asserted, or Lambrecht 1994, who also proposes similar ideas). Being backgrounded, they have no hope of being Maximally Contrastive. In this context, the MaxCoNAH predicts an overt subject rather than a null subject as the best co-referential option, thus straightforwardly explaining the higher RTs in the Episodic Verb plus Null Subject Condition. There is one point, however, that remains unclear as regards the approach proposed in the present dissertation: how non-specific referents mentioned in the discourse differ from generics which are not previously mentioned? In what follows in this section, this remaining issue is addressed.

The Non-MaxCoNAIH was formulated to address the complementary distribution in the interpretations of null subjects in Chapter 3. However, the interpretation and use of overt and null subjects are not in a binary complementary distribution. In real discourse, many possible antecedents (or more generally possible interpretations) are available for each expression. The fact that the generic interpretation emerges when the contexts does not favor the co-referential interpretation has to do with multiple factors (as proposed in Chapter 2, section 2.8). Taking all the multiple factors into consideration, it is not surprising that there are multiple "imbalances" between possible interpretations. One of them is the fact the generic interpretation is established more "locally" than co-reference with a previously mentioned non-specific antecedent. In such cases, regardless of the use of a structural operator, as in Krifka et al. (1995), or the establishment of the interpretation via semantic-pragmatic evaluation of alternatives, as in Baumann et al. (2014), the generic interpretation is preferred because of the "division of labor" between forms, according to which the most

293

underspecified form is used for the most obvious and contrastive interpretation.

The intervention of multiple factors as regards preferential interpretations is not unheard of in the literature about non-specific antecedents and null and overt subject distribution in Romance languages. Luján (1999) already observed that the generalization proposed by Montalbetti (1984), according to which null subjects pattern like bound variables when referring to quantified antecedents does not hold regardless of discourse context, as shown in (183) below.

(183)   a. Todo estudiante piensa   que él es   inteligente y   los demás no.
              Every student   think.PRS.3SG that he be.PRS.3SG intelligent and the others not
              "Every student thinks that he is intelligent and the others are not."

                                                                    (Luján 1999)

As observed in (183), manipulations in the discourse structure can easily change the interpretation preferences and the choice of anaphoric expressions, in a way that no hard constraints can explain. In Experiment 7, the overt pronouns are easily processed even if co-refering with quantified antecedents because of the discourse environment: being in the background, null subjects are not the optimal option. For the present account, this is captured by the probabilistic mapping proposed in section 2.8 in Chapter 2.

For the purposes of the analysis of the generic null subject, it is taken to be a by-product of the combination of the general properties of generic sentences, along with the unlikelihood of the co-referential interpretation, and finally the computation of parallel options which take into consideration not only the *in presentia* elements but also all other alternative paradigmatic formulations that could be used (as proposed by Baumann et al. 2014, Colonna et al. 2014, *inter alia*). When the semantics of the sentence allows the generic interpretation (as in any other generic sentence), this is one more competing option, which is preferred in the contexts in which other competing options are unlikely. In the next experiment, this prediction is tested once more, in another slightly different context: free relative clauses.

4.6 Addressing Free Relative Clauses

As shown in the previous section, RTs of object relative clauses with null co-referential subjects are higher than object relative clauses with generic null subjects or overt co-referential subjects. A similar effect would be expected in a similar structure (possibly a quite similar type of syntactic structure), such as free relatives. The aim of this section is to discuss this puzzling sentence structure from the point of view of the theory proposed in this dissertation.

The most interesting case is observed by Modesto (2000a), who points out that free relatives, against expectations, do not pattern like object relative clause islands, as show in (184) below. In the corpus research reported here, similar cases where also found, as shown in (185) (also compare to the very similar example in 186).[7]

(184) A Maria$_1$ não sabe quem$_2$ _$_1$ entrevistou _$_2$.
the Maria not know.PST.3SG who interview.PST.3SG
"Maria$_1$ doesn't know whom$_2$ she$_1$ interviewed _$_2$."

(Modesto 2000a, 100)

(185) todo mundo$_1$ que entra [no curso] sabe o que$_2$ _$_1$
all world that enter.PRS.3SG in.the course know.PRS.3SG what _
quer fazer _$_2$.
want.PRS.3SG do.INF
"Everybody$_1$ who enrolls for [the course] knows what$_2$ they$_1$ wanna do _$_2$."

(NURC-RJ, "Inquiry_ac_1")

---

[7] Anne Abeillé suggested that examples 184 and 185 might be analyzed as embedded interrogatives. At the syntactic level, it is defended here that they pattern alike. At the semantic and discourse level, on the other hand, they do not seem quite similar. As proposed in the section 4.6.3, they are not "backgrounded". Caponigro (2003) makes a comparison between questions and free relatives and finds some differences. In fact, it seems that there is a difference which comes from the fact that the focus of the interrogatives is the questioned constituent ("Wh-phrase", for example), while in free relatives the WH-phrase denotes the maximal set of individuals to which the predicate is attributed, and the focus of the object free relative is assigned "by default" on the last word of the sentence. Therefore, it is proposed here that there is a difference, and this is the reason why there is less subject null in the interrogatives than in the object free relatives.

(186)   Ninguém$_1$ revela realmente tudo que $\_1$ sente $\_2$.

   "Nobody$_1$ really reveals everything$_2$ they$_1$ feel $\_2$."

<div align="right">(NURC-RJ, "Inquiry_71")</div>

The similarity between the free relative in (185) and the object relative clause in (186) is evident. However, given the results in the previous section, why are these sentences acceptable and claimed to be a counter-evidence for syntactic-based analyses of null subjects?

4.6.1   If it is free, one can be null, but (s)he cannot

Huang (1984) observed that the gaps in subject free relatives in Chinese can only be interpreted in one single way: the relative pronoun must be the object and the subject must be generic, as in (187) below.

(187) $\_{gen}$ mai $\_2$ de gen $\_{gen}$ zu $\_2$ dou hao.
         buy   DE and   rent   all good
        "What one buys and what one rents is both good."

<div align="right">(Huang 1984, 545)</div>

This observation sharply contrasts with the data presented in the previous section about BP. However, many theories about BP null subjects rely on the initial proposal made by Huang (1984), according to which a null argument may be a bound variable linked to a covert operator in the topic position of the sentence (Modesto 2000*b*,*a*, 2008*a*,*b*). If this hypothesis can somehow be implemented, subject and object free-relatives should not be different in terms of acceptability and interpretation. Since object free relatives with co-referential null subjects are both straightforwardly considered acceptable in the previous literature and easily found in the corpus research, this structure does not need to be tested. Any native speaker of BP accepts sentences (184) and (185), for instance, as natural. The point is to test whether the predictions made by Huang (1984) also hold for BP. If they do, a proposal that is able to differentiate object and subject free-relatives is evidently needed.

Finally, assuming the semantic description proposed by Caponigro (2003), it is shown that once more the MaxCoNAH and the Non-MaxCoNAIH account for the data straightforwardly. The

main assumption is that pre-verbal subject free relatives are not Maximally Contrastive, differing from the object free relatives reported in the literature and found in the corpus. Essentially, the informational focus is not in the domain of the free relative, since the assertion tends to be in the verbal domain outside the subject free relative. The subject free relative is, thus, a non-Maximally Contrastive context.

4.6.2 Experimental Evidence – Experiment 8

In this section, an experiment to test whether subject free relatives behave similarly to object free relatives in BP or like subject free relatives in Chinese is reported. The results suggest that BP subject free relatives are similar to the cases observed by Huang (1984), but at the same time differ in two fundamental aspects: (i) the relative pronoun can be interpreted as the subject of the verb in the free relative; and (ii) having assumed that the two gaps can equally be bound by the relative pronoun, some possible effect of the topicality/subjechood of the antecedent would be expected. No such effect was found. Possible explanations for the facts observed in this Experiment rely on the fact that pre-verbal subject free relatives are taken to be backgrounded. As backgrounded content, they cannot be Maximally Contrastive and thus they are not an optimal context for co-referential null subjects. Also, the outcome of the present experiment suggests that background content possibly has no internal articulation in terms of information structure, so that being a topic in the discourse probably has little or no effect on the likelihood of a null subject when it is within the backgrounded part of the sentence. Consequently, as predicted by the definition of Maximal Contrast and by the MaxCoNAH, topichood does not show significant effects outside of its relation with focus in the use of null subjects in BP.

Material Design

In order to test for on-line effects of the context of subject free-relatives on the interpretation of null subjects, sentences in which the relative pronoun is disambiguated in the main clause verbal area were proposed. The hypothesis was that without a cue about which constituent is the relative pronoun and which one must be interpreted arbitrarily, the reader would disambiguate the syntactic parsing only in the CR, getting garden-pathed by the effects of the frequency of rela-

297

tive constructions (subject relatives are reported to be cross-linguistically easier to process in the psycholinguistics literature); moreover this effect was counter-balanced by the frequency of the interpretation of the relative pronoun *O que* "what or the one who/those who" as explained below.

Based on this idea, the free-relatives should start with the pronoun *o que*, which in BP can be interpreted as "what" (the most commom interpretation) and "those who/which" or "the one who/which" (the less common interpretation), as in example (188).

(188)   a.   A – quando... quantas... eh... quantos adolescentes assim... pessoal jovem₁ que mora na zona sul... que por... força das circunstâncias... tem que estudar numa faculdade lá em Deus me livre... na Piedade... né?

"When... how many... hmmm... how many teenagers like this... young folks who live in the south area... who because of the circumstances... have to study in a faculty there in a way out in the sticks... in Piedade... right?"

   b.   B – exatamente é... é... porque _₁ foi aprovado.

"exactly yes... yes... because they₁ were accepted."

   c.   A – ou vice-versa também... o   que   mora   na   zona norte...
          or vice-versa also   that which live.PRS.3SG in.the zone north
        vai   estudar   na   PUC... né?
        go.PRS.3SG study.INF in.the PUC   [DM]

"or vice-versa as well... those who live in the north area... go to study in the PUC [Pontifical Catholic University]... right?"

Since the "What" interpretation is far more frequent than "those who/which" or "the one who/which", this may cancel out any underlying preference for a subject relative clause. Moreover, the "those who" Condition is taken to be the baseline to attest whether the relative pronoun can be interpreted as the subject of the sentence and to produce equal expectations on the filler-gap relation.

Twenty-four items were created for this experiment. The eight conditions are exemplified in (189) and (190) below. The experiment was conceived in a two-by-two-by-two design, that is, three binary Factors. In (189), the potential antecedent of the null subject inside the free relative is the subject of the preceding sentence, while in (190) the potential antecedent is in a by-phrase (Factor Antecedent: Subject vs. non-Subject Conditions). In sequences (a) (189a and 190a), the

verbs of the relative clauses are episodic; while, in sequences (b) (189b and 190b), the verbs of the relative clauses are modal and potentially intransitive (Factor Verb: Episodic vs. Modal Conditions). Finally, in sentences (i), the relative pronoun is interpreted as "What" and so it is the object of the free relative; while in sentences (ii), the relative pronoun is interpreted as "Those who" or "The one who" and so it is the subject of the free relative (Factor Rel_Pron: What vs. Those_Who Conditions).

(189)   a. O bom jogador da nossa equipe preferiu o exercício esportivo extenuante.

"The good player(s) in our team preferred the strenuous workout(s)."

i. O   que   tinha   realizado   com frequência |*causou*   *ganhos*
that which have.PST.3SG realize.PRTC with frequency causa.PST.3SG gains
bem rápidos|, disse   o   treinador da   equipe.|
very fast   say.PST.3SG the coach   of.the team

"What (s)he/they had performed frequently caused very fast gains, said the coach of the team."

ii. O   que   tinha   realizado   com frequência |*teve*   *ganhos*
that which have.PST.3SG realize.PRTC with frequency have.PST.3SG gains
mais rápidos|, disse   o   treinador da   equipe.|
more fast   say.PST.3SG the coach   of.the team

"Those who had performed it/them frequently had faster gains, said the coach of the team."

b. O bom jogador da nossa equipe prefere o exercício esportivo extenuante.

"The good player(s) in our team prefer(s) the strenuous workout(s)."

i. O   que   tem   de treinar   com frequência |*causa*   *ganhos*
that which have.PRS.3SG of train.INF with frequency causa.PRS.3SG gains
bem rápidos|, disse   o   treinador da   equipe.|
very fast   say.PST.3SG the coach   of.the team

"What (s)he/they/one practice frequently causes very fast gains, said the coach of the team."

ii. O   que   tem   de treinar   com frequência |*tem*   *ganhos*
that which have.PRS.3SG of train.INF with frequency have.PRS.3SG gains
mais rápidos|, disse   o   treinador da   equipe.|
more fast   say.PST.3SG the coach   of.the team

"Those who have to work out frequently have faster gains, said the coach of the team.

(190) a. O exercício esportivo extenuante foi preferido pelo bom jogador da nossa equipe.

"The strenous workout(s) was/were preferred by the good player(s) in our team."

i. O   que   tinha   realizado   com frequência |*causou*   *ganhos*
that which have.PST.3SG realize.PRTC with frequency causa.PST.3SG gains
bem rápidos|, disse   o   treinador da   equipe.|
very fast   say.PST.3SG the coach   of.the team

"What (s)he/they had performed frequently caused very fast gains, said the coach of the team."

ii. O   que   tinha   realizado   com frequência |*teve*   *ganhos*
that which have.PST.3SG realize.PRTC with frequency have.PST.3SG gains
mais rápidos|, disse   o   treinador da   equipe.|
more fast   say.PST.3SG the coach   of.the team

"Those who had performed it/them frequently had faster gains, said the coach of the team."

b. O exercício esportivo extenuante é preferido pelo bom jogador da nossa equipe.

"The strenous workout(s) is/are preferred by the good player(s) in our team."

i. O   que   tem   de treinar   com frequência |*causa*   *ganhos*
that which have.PRS.3SG of train.INF with frequency causa.PRS.3SG gains
bem rápidos|, disse   o   treinador da   equipe.|
very fast   say.PST.3SG the coach   of.the team

"What (s)he/they/one practice frequently causes very fast gains, said the coach of the team."

ii. O   que   tem   de treinar   com frequência |*tem*   *ganhos*
that which have.PRS.3SG of train.INF with frequency have.PRS.3SG gains
mais rápidos|, disse   o   treinador da   equipe.|
more fast   say.PST.3SG the coach   of.the team

"Those who have to work out frequently have faster gains, said the coach of the team.

The empirical predictions were the following: overall, the disambiguation region (CR) should be read faster when the free relatives have a modal verb than when they have an episodic verb if the context facilitates the generic interpretation but the co-referential interpretation is dispreferred.

If the relative pronoun can be interpreted only as object, the total RTs in the CR should he higher when it is interpreted as "Those_Who" than when it is interpreted as "What". A garden path effect is expected in the disambiguation region (CR), as a consequence of the Factor Rel_Pronoun. Finally, no increase in total RTs of the position of the antecedent is expected if the topic or syntactic position of the antecedent does not exert effects on the null subject in such contexts, but it can modulate the garden path effect in the First Pass RTs in the CR and the regression path RTs in the CR and post-CR. Condition Those_Who with a modal verb is taken as a baseline for comparison, since the verbs in the free relatives have only one gap to be filled in that case.

Results

Due to the complexity of the effects found in this experiment, the analysis was split into three short subsections: in **First Pass RTs**, the first pass RTs in the CR, in the post-CR and in the sum of the CR and Post-CR are analyzed, *i. e.* the time that the reader spent from the beginning to the end of the regions until (s)he left the area for the first time; in **Regression Path RTs**, that is, the regression path RTs in the CR, in the Post-CR and in the sum of the CR and post-CR are analyzed, that is, RTs including first pass reading times and fixations from regressions to previous regions until the region is left to the right; finally, in **Total RTs**, the total RTs spent by the readers in the CR, in the Post-CR and in the sum of the CR and the Post-CR are analyzed.[8]

First Pass RTs. The most relevant analysis is shown in Figures 4.7 and 4.8. In the first pass RTs in the CR, many effects seem to play a relevant role at the same time. The first observable effect is main Factor Verb in the What Rel_Pronoun Condition. Clearly, in this combination of Factors, the CR is read faster when a modal verb is in the free relative than when an episodic verb is in the free relative. In the Those_Who Condition, the same effects are not found. Mainly, it seems that readers have more difficulties in understanding the correct interpretation of the relative pronoun when a modal verb is used and the antecedent is not the topical subject of the preceding sentence. Other conditions seem to follow the same pattern as in the What Rel_Pronoun Condition. Overall,

---

[8] As mentioned in section 4.3, only the models whose residuals follow a Gaussian distribution are presented in this chapter. In the present section, the models of either the logs of the raw data or the length-residuals are presented.

contrary to expectations, the What Rel_Pronoun Condition (which is far more frequent than "Those Who") is read slower than Those_Who.

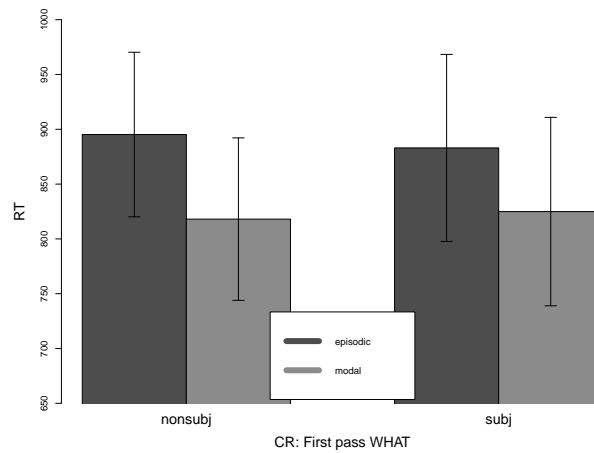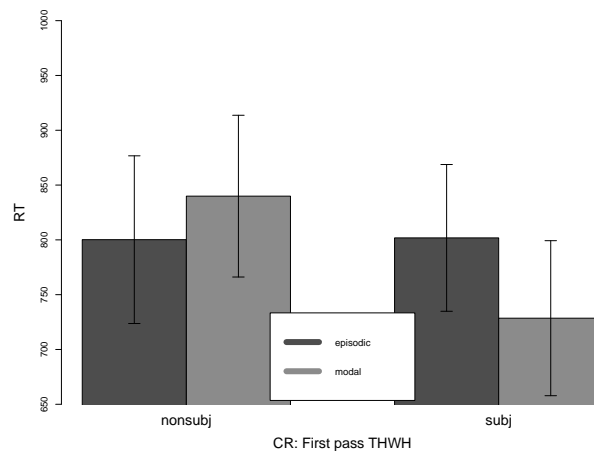Figure 4.7 – First Pass RTs for What Free Relatives according to Antecedent and Verb Factors

Figure 4.8 – First Pass RTs for Those Who Free Relatives according to Antecedent and Verb Factors

As for the inferential statistical analysis, the residuals of a model considering only the Length of the CR were entered in a linear mixed model. Main Factors were Rel_Pro, Antecedent and Verb (and their possible interactions), as well as the Session in which the participant took part, along with random Factors Participants and Items.

As shown in Table 4.3, main Factors Verb and Rel_Pro showed marginal effects, as well as the interactions Antecedent:Rel_Pro and Verb:Rel and the three-Factor interaction. The inter-

action Antecedent:Verb is significant, possibly due to the unexpected high RTs in the Those_-
Who:nonSubject:Modal Condition.

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	-18.636	43.071	-0.433	0.6662
Rel_Pro	76.079	45.484	1.673	0.0947 .
Antecedent	4.320	44.913	0.096	0.9234
Verb	81.128	46.461	1.746	0.0810 .
Session	-1.947	22.907	-0.085	0.9323
Antecedent:Verb	-137.594	63.484	-2.167	0.0304 *
Antecedent:Rel_Pro	-121.106	64.468	-1.879	0.0606 .
Verb:Rel_Pro	-121.106	64.468	-1.879	0.0606 .
Antecedent:Verb:Rel_Pro	166.628	92.721	1.797	0.0726 .

Table 4.3 – Log-linear mixed-effects model for Residuals of First Pass RTs in the CR in the model considering Length in Experiment 8

The only main Factor that was found to be marginally significant in the Post-CR with respect to the First Pass was Verb (β = -64.030; SD = 35.010; T-value = -1.829; p-value = 0.0676 .). The interaction between the three main Factors was also marginally significant (β = -64.030; SD = 35.010; T-value = -1.829; p-value = 0.0676 .). In the sum of the RTs of the CR and the post-CR again the only Factor that showing a marginal effect is Verb (β = 89.3645; SD = 49.4994; T-value = 1.805; p-value = 0.0712 .) (see Appendix D).

Regression Path RTs. In the Regression Path RTs, two relevant effects were found. First, in the CR, summarized in Figures 4.9 and 4.10, the effects found in the First Pass RTs are again observed: overall the RTs in the What Rel_Pronoun are higher than the RTs in the Those_Who Rel_Pronoun Condition; in the first case, the RTs of the Modal Verb Condition are lower than the RTs of the Episodic Verb Condition. However, the RTs in the Those_Who Rel_Pronoun Condition differ from those found in the First Pass and in the What Rel_Pronoun: they were higher in the Modal Verb Condition than in the Episodic Verb Condition. No effect of the Antecedent is observable in the graphs below.

Figure 4.9 – Regression Path RTs in the CR for What Free Relatives according to Antecedent and Verb Factors
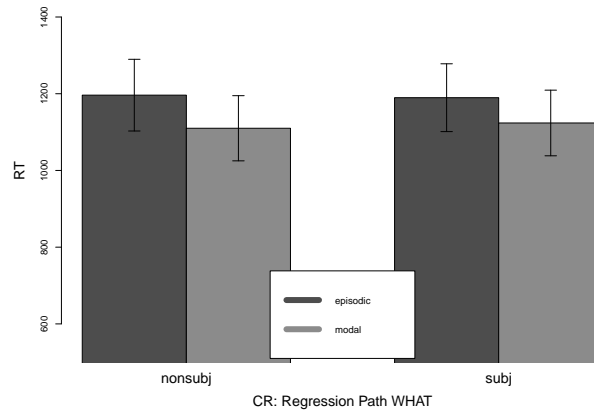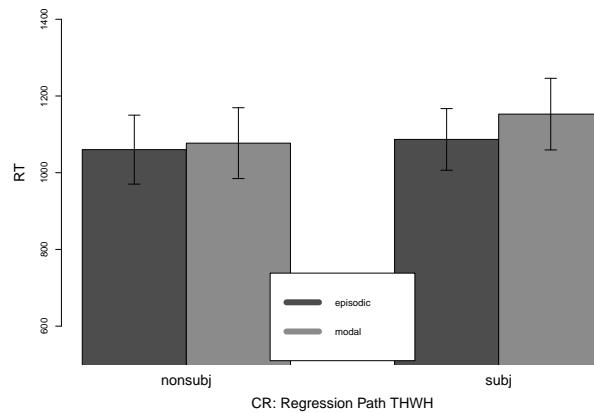

Figure 4.10 – Regression Path RTs in the CR for Those Who Free Relatives according to Antecedent and Verb Factors

As for the statistical analysis, the logs of the Regression Path RTs were entered in a linear mixed model. All main Factors, additional Factors and random Factors were the same as in the regression of the First Pass RTs, along with additional Factor Length. As summarized in Table 4.4, only significant main linguistic Factor was Rel_Pro. No other linguistic Factor reached significance. Length and Session were significant.

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	8.99072	0.31793	28.279	< 2e-16 ***
Rel_Pro	0.15243	0.06028	2.529	0.01158 *
Antecedent	0.02985	0.05878	0.508	0.61169
Verb	0.06160	0.06263	0.984	0.32548
Length	0.04059	0.01313	3.090	0.00243 **
Session	-0.06406	0.02993	-2.140	0.03254 *
Antecedent:Verb	0.04037	0.08302	0.486	0.62689
Antecedent:Rel_Pro	-0.01094	0.08437	-0.130	0.89687
Verb:Rel_Pro	-0.10970	0.08440	-1.300	0.19396
Antecedent:Verb:Rel_Pro	0.01842	0.12098	0.152	0.87900

Table 4.4 – Log-linear mixed-effects model for Logs of Regression Path RTs in CR in Experiment 6

The second relevant effect was found in the post-CR, as shown in Figures 4.11 and 4.12. A consistent effect of Factor Verb is observed across all other Conditions.

Figure 4.11 – Regression Path RTs in the Post-CR for What Free Relatives according to Antecedent and Verb Factors
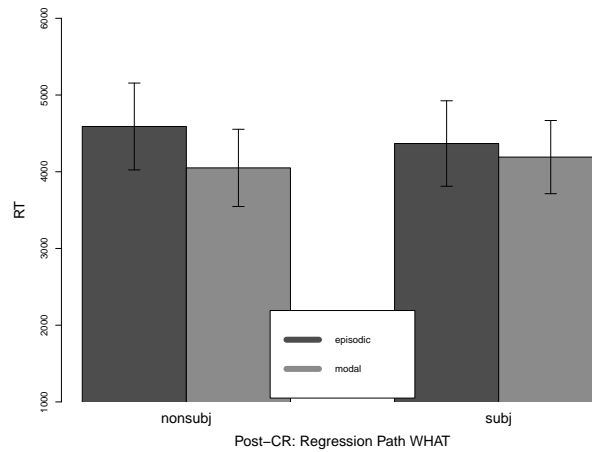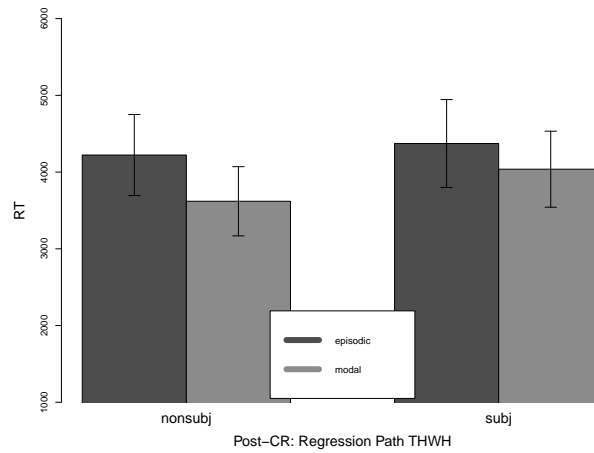
Figure 4.12 – Regression Path RTs in the Post-CR for Those_Who Free Relatives according to Antecedent and Verb Factors

As for the statistical analysis, as in the First Pass RTs inferential analysis, the residuals of the regression path RTs in the post-CR from a model considering only Length were entered into a linear mixed model. Main Factors were Rel_Pro, Antecedent and Verb (and their possible interactions), as well as the Session in which the participant took part, along with the random Factors Participants and Items.

As shown in Table 4.5, the only significant linguistic main Factor was Verb. The non-linguistic Factor Session also reached significance.

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	463.60	225.78	2.053	0.0404 *
Rel_Pro	307.02	295.27	1.040	0.2986
Antecedent	85.51	292.90	0.292	0.7704
Verb	-647.36	288.07	-2.247	0.0248 *
Session	-717.98	144.12	-4.982	7.07e-07 ***
Antecedent:Verb	298.16	405.73	0.735	0.4625
Antecedent:Rel_Pro	-253.23	415.33	-0.610	0.5421
Verb:Rel_Pro	129.34	408.99	0.316	0.7519
Antecedent:Verb:Rel_Pro	37.23	575.10	0.065	0.9484

Table 4.5 – Log-linear mixed-effects model for Residuals of Regression Path RTs in the Post-CR in the model considering Length in Experiment 8

No significant linguistic effect was found in the sums of the CR and Post-CR as regards the regression RTs (see Appendix D).

306

Total RTs. In the CR and Post-CR, no significant linguistic effect was found. For this reason, in the Figures 4.13 and 4.14 below, only the total RTs of the *sum* of the CR and Post-CR are shown.

As presented in Figures 4.13 and 4.14, once again Factor Verb showed a marginal effect in the total RTs for the sum of the CR and Post-CR. The only case in which the same pattern does not seem to come up is the Those_Who:Subject:Episodic Condition. In this case, the relevant region is read slightly faster than the others (which looks similar to the Those_Who:Modal Conditions).

Figure 4.13 – Total RTs in the CR + Post-CR for What Free Relatives according to Antecedent and Verb Factors
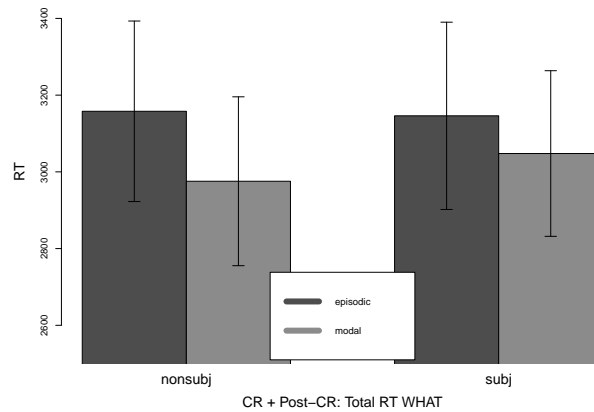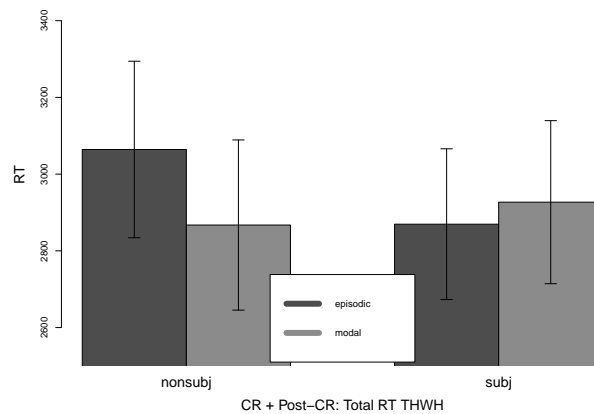

Figure 4.14 – Total RTs in the CR + Post-CR for Those_Who Free Relatives according to Antecedent and Verb Factors

As for the statistical inferential analysis, the logs of the Total RTs were entered in a linear

mixed model. All main Factors, additional Factors and random Factors were the same as in the model for the First Pass RTs and Regression Path RTs, along with additional Factor Length. As summarized in Table 4.6, the only marginally significant main linguistic Factor was Verb. No other linguistic Factor reached significance. Since Length is more or less constant across conditions with respect to this regions, the only extra-linguistic Factor which reached significance was Session.

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	1.148e+01	2.744e-01	41.825	< 2e-16 ***
Rel_Pro	5.273e-02	5.314e-028	0.992	0.3212
Antecedent	-5.899e-02	5.268e-02	-1.120	0.2630
Verb	-1.001e-01	5.336e-02	-1.876	0.0609 .
Length	-1.593e-04	1.077e-02	-0.015	0.9882
Session	-1.618e-01	2.611e-02	-6.196	7.45e-10 ***
Antecedent:Verb	8.857e-02	7.360e-02	1.203	0.2290
Antecedent:Rel_Pro	4.210e-02	7.450e-02	0.565	0.5721
Verb:Rel_Pro	1.643e-03	7.377e-02	0.022	0.9822
Antecedent:Verb:Rel_Pro	-1.136e-02	1.041e-01	-0.109	0.9131

Table 4.6 – Log-linear mixed-effects model for Logs of Total RTs in the sums of the CR and the Post-CR in Experiment 8

Discussion

As shown in section 4.6.2 above, subject free relatives pattern similarly to subject free relatives in Chinese Mandarin (as described by Huang 1984) and to object relative clauses in BP, as described in section 4.5.2. The main Factor which is persistently significant across relevant regions and measures is whether the verb allows a generic interpretation or not: modal verbs, which allow the generic interpretation, were read faster in the first pass in the CR, in the regression path in the post-CR and in the total RTs in the sums of both regions, in the What interpretation of the relative pronoun. In terms of total RTs, however, the Those_Who and What interpretations did not diverge significantly. The only measure in which they diverge significantly was in the Regression Path RTs of the CR, which suggests a garden path effect. This garden path effect is, however, modulated by the semantics of the verb, as observed in the Regression Path RTs of the post-CR. These findings are consistent with the initial hypothesis for this Experiment: generic interpretation are easier to process and modulate the garden path effect of the double gapped free-relative; however, no additional difficulty was expected in the interpretation of the relative pronoun either as subject or

object of the free relative. No relevant effect of the antecedent was found apart from its interaction with Factor Verb in the first pass RTs in the CR.

4.6.3   Accounting for Freedom

As shown in the Experiment 8, subject free relatives in BP pattern like object relative clauses in BP and subject free relatives in Chinese with respect to the interpretation of the null subject inside the free relative. However, the effect of the verb type shows that they differ from object relative clauses, since the latter can be co-referential with a topical subject. In the approach developed in the present dissertation, the main constraint which explain such different patterns is the informational focus of the sentences.

In section 4.5.1, it was argued that the apparently island effect of object relative clauses as regards the use of null subjects is indeed accounted for by a discourse-based approach, rather than a structural approach. In the same vein, instead of proposing additional syntactic machinery for accounting for the effects found in Experiment 8 (since the explanation proposed by Huang 1984 is ruled out by additional evidence in this experiment), the proposal here is to understand the semantic and discourse structure of free relatives and to check whether the same hypotheses proposed so far (namely, MaxCoNAH and Non-MaxCoNAIH) can also account for these cases straightforwardly. And indeed they do.

The first observation is that relative pronouns in free relatives (*quando* "when", *como* "as", *o que* "what", *quem* "who") are referentially empty (they have no referential content). Rather, as observed by Caponigro (2003), they are like variables which can denote the maximal content delimited by the predicate. In some sense, they act exactly like embedding conjunctions and not like standard relative pronouns. They do not count, for the sake of Maximal Contrast, as left dislocated elements because of their lack of semantic content. This proposal explains why object free relatives, as well as temporal PP free relatives and comparative PP free relatives, can easily and frequently present co-referential null subjects, even though there is no c-command relation between the antecedent and the null subject.

The second important observation is that focus tends to be on the rightmost constituent in BP. As in many other languages, when the focus is on this constituent it can spill over toward the

constituents located to the left. This way, the object free relatives can be analyzed as Maximally Contrastive. Pre-verbal subject free relatives, however, cannot. In the sentences in Experiment 8, none of the subject free relatives have a focused constituent; that is, no part of the free relative is asserted. As a consequence, pre-verbal free relatives are not likely to have co-referential null subjects in BP, since they have no hope of being Maximally Contrastive.

4.7   General Discussion

In this chapter, three syntactic-discursive contexts were analyzed with respect to the use and interpretation of null and overt subjects in BP. In the previous literature, these contexts were particularly challenging because some of them are expected to be ungrammatical. In the account presented here, however, they are supposed to contrast in terms of processing difficulty, mainly due to the restrictions on the interpretation of null subjects in BP.

The main aim of this chapter was to provide online evidence for the observations made throughout the present dissertation based on off-line data. As pointed out in the Introduction to this chapter, eye-tracking while reading experiments provide a window into the comprehension and processing of sentences in a very natural way. That is, in such experiments, no extra effort or metalinguistic task is demanded of the participant.

The lines that connect the three experiments reported in this chapter are (i) the methodology, (ii) the testing of non-optimal contexts for null subjects and (iii) the generic interpretation in such contexts. This chapter attempts to fill a gap in the literature as regards the lack of experimental data about null subjects and the generic interpretation in general. One of the few researchers which previously provided some insights into this area was reported in Chapter 3: Haendler (2017). In the present chapter, results quite similar to those reported in that paper are found: in Modern Hebrew and German, generic subjects facilitate the interpretation and processing of object relative clauses. Although Modern Hebrew is similar to BP in some aspects (being considered a "partial" pro-drop language), the indefinite null subject does not clearly present a generic inclusive interpretation (see Chapter 3). In German, the interpretation of the pronoun *man* seems to be similar to that of the generic null subject in BP. However, German is a language with obligatory expression of subjects, which makes the comparison between the results of the experiments harder.

The data provided in this chapter suggest that the MaxCoNAH and the NonMaxCoNAIH account straightforwardly for the three syntactic-discursive contexts experimentally tested. First, contrary to what is claimed by Holmberg (2007), null subjects in polarity contrastive environments do not need to be produced by IP-ellipsis (which requires open-polarity in the antecedent). In the proper Maximally Contrastive context, the null and the overt subject are read at the same pace, which suggests that the null subject is not ruled out in such contexts nor even more complex to process. What actually exerted effects on the RTs of the relevant area was the position of the adjunct phrase and the choice of verb (whether it was modal or not). As suggested previously (in Chapters 2 and 3), the effect of the position of the adjunct is explained by the general hypothesis that the optimal context for the use of null subjects is Maximally Contrastive. Fronting the adjunct produces a sentence that is no longer Maximally Contrastive, and consequently the null subject is not preferential. In such a context, the null subject fails to be grounded (see section 3.8 in Chapter 3). Given this analysis, the generic interpretation emerges as a by-product of the standard semantic structure of generic sentences without linking the referent of the subject to the common ground, but rather in a purely intra-discursive way. This account explains the results found in Experiment 6.

Surely, however, the account proposed here is not the only one that can explain the findings of Experiment 6. In general, accounts that are based on "movement out of case domains" (such as, for instance, Rodrigues 2004, Ferreira 2000, 2004, Sheehan 2006) could also deal with the data provided here. However, it is not clear what their predictions about each measure would be: would a model in which null subjects are produced in a way similar to Wh-movement predict a significant difference only in the total RTs? If syntactic violations are processed faster or first, should the effect of a possibly ungrammatical sentence not be observed in the first pass RTs? If the prediction of such models are that the syntactic structure is processed first, they would fail to explain the data found here. Accounts based on some kind of operator in the TopP position of the sentence could also account for the data presented here (Modesto 2000*a*, Ferreira 2000, Modesto 2008*a*,*b*, Barbosa 2011). However, they also seem to predict stronger difficulties in processing these sentences than the increase in processing time found here. Finally, accounts based on the syntactic anaphoric behavior of null subjects have little or nothing to say about the results found in Experiment 6 (Figueiredo-Silva 1994, 2000, Branco 2007, Menuzzi & Lobo 2016). These accounts

do not predict any interference of the fronted adjunct as long as syntactic c-command is maintained.

The findings in Experiment 7 are also hard to account for in TopP-based accounts or bound-variable-based accounts. As Sheehan (2006) points out, most of these approaches propose an *ad hoc* resource to rule out co-referential null subjects in object relative clauses. Movement-based analyses, on the other hand, categorically rule out the use of co-referential null subjects in such constructions. However, none of these predictions hold completely. First, although dispreferred, in corpora it is possible to find grammatical co-referential null subjects in object relative clauses. In Experiment 7, however, co-referential null subjects showed a significant increase in total RTs in the object relative environment. Furthermore, given the results obtained by Maia et al. (2016) about relative clause islands in BP and the results obtained in Experiment 7, it is unlikely that null subjects are similar to traces of topicalization or Wh-movement. In the proposal presented in this dissertation, this effect is due to the fact that the object relative clause is not an optimal context for co-referential null subjects, since they are not Maximally Contrastive. As proposed in previous chapters, the increase in processing is rather an interpretation problem than a lack of licensing properties (which would lead to strong unacceptability). Following the reasoning presented in Chapter 2 and the proposed discourse structure analysis by Goldberg (2013), the main difficulty found in object relative clauses is that they are usually backgrounded information. With a lack of focused material, these clauses have no hope of being Maximally Contrastive. Consequently, co-referential null subjects are not in their optimal context. In such contexts, once again failing to be grounded, null subjects are interpreted as generic when the sentence semantic structure allows it. When the generic interpretation is achieved, the sentences are read faster, reproducing the results found by Haendler (2017) for Modern Hebrew and German (notwithstanding the differences mentioned above).

Finally, in Experiment 8, the most puzzling case for syntactic-based accounts of co-referential null subjects was studied: free-relative clauses. Movement-based analyses cannot explain the uses of co-referential null subjects in object free-relative clauses mentioned in the relevant literature and found in corpus. At the same time, TopP-based accounts appeal to *ad hoc* mechanisms to account for free-relatives in BP. Based on Huang (1982, 1984, 1989), such accounts expect BP to be similar to Chinese Mandarin. And, indeed, it is. Subject free relatives facilitate the generic interpretation, which is read faster and reduce the garden path effect on the filler-gap relation. Co-referential null

subjects are overall read slower and increase the garden-path filler-gap effect. However, differently from Chinese Mandarin, the relative pronoun can be interpreted as object or as subject of the verb in the free-relative without significant increase in processing demands (in some measures they are even read faster). This finding suggests that the filler-gap relation in BP is different from Chinese Mandarin, and the null operator approach cannot account for BP without further assumptions.

Following the proposal presented in this dissertation, both the corpus research and the experimental data are easily accounted for. The only restriction regarding the use of co-referential null subjects in subject free-relatives is due to the fact that they are in the backgrounded area of the sentence. In this area, they have no hope of being Maximally Contrastive (because the whole free-relative is backgrounded). When a part of the free relative is focused, as in object free relative reported in section 4.5, co-referential null subjects are as natural as in other Maximally Contrastive contexts. Being in a non-optimal context for co-referentiality, the null subject fails to be grounded. So, its reference is established intra-discursively, *i. e.* the generic interpretation emerges in the proper context. In the context in which co-referentiality is forced, the garden-path effect of the interpretation of the relative pronoun is persistent and affects the RTs in the CRs.

Overall, the sentences studied in this chapter show consistent effects which clearly follow from the theory developed in the present dissertation. As previously discussed, null co-referential subjects seem to be semantically and discursively restricted by constraints on interpretation. These constraints were spelled out in Chapter 2 under the form of a scalar probabilistic mapping. In this chapter, online evidence for this approach was provided.

4.8   Summary

In this chapter, three syntactic-discourse contexts were analyzed in on-line experiments. The main hypotheses were that the difficulties in processing sentences would increase the cognitive effort demanded (and thus the RTs) in the conditions in which the co-referential null subjects are not likely, following the approach proposed in Chapter 2. In these contexts, however, when the semantic conditions for genericity are achieved, the generic interpretation emerges and the sentences would be read significantly faster.

In this chapter, on-line evidence suggests that the difficulties in processing co-referential null

subjects in some syntactic environments, rather than a licensing failure, seems to be due to the constraints on the identification and interpretation of these null subjects. The main findings, which were relevant in the three experiments, were the differences between modal verbs (which allows null generic subjects) and episodic verbs (which are not interpreted as generic). This effect is claimed to emerge from optimal vs. non-optimal contexts for co-referentiality and the failure to ground the referent of the subject in the common shared set of propositions (common ground). Constraints on opposite polarity sentences, relative clauses and subject free-relative clauses, elsewhere thought to be syntactic, are attributed to the scale of Contrast, according to which null co-referential subjects are preferred in Maximally Contrastive discourse structures. Along with other factors identified in Chapters 1 and 2 they seem to indicate where null subjects seem to be used and how they tend to be interpreted.

Notice, however, that differently from what is expected in other approaches based on hard constraints on the licensing of null subjects, here the effects were felt consistently only in the total RTs. These results indicate that the constraints in the use of co-referential null subjects can be accounted for in a model that focuses on interpretation, such as the anaphora resolution approach. In future research, once the use of null subjects is broadly described in terms of general anaphoric properties, syntactic and/or morphological idiosyncrasies should be addressed and accounted for. However, if a comprehensive model of anaphora resolution in broader contexts accounts for these idiosyncrasies, as proposed throughout this chapter, syntactic constraints can be considered a by-product of this approach at least in languages such as BP, which does not seem to conform to either a binary formulation or a multiply typed pro-drop theory.

Conclusion

The present dissertation shows the importance of setting the theoretical research on empirical grounds, combining corpus research along with experimental data. Much of the previous literature has been based on intuitions, contrasts and general descriptions that can be challenged through empirically based research. The results presented here aim to both shed light on the theoretical questions addressed and to set up a methodology to further investigate these questions on an empirical basis. In what concerns specifically the topic of the present dissertation, the use and interpretation of null and overt subjects in BP were studied, reanalyzed and accounted for. In this endeavor, a design for the grammar was sketched. This grammar has fundamental features that enables it to deal with the empirical data presented throughout this dissertation: (i) it is multilayered, (ii) it is probabilistic and (iii) it is scalar. This design fits the empirical findings in anaphora resolution (and in other interface phenomena) in many languages, although they can be challenging to accommodate in level-centric theories with binary constraints.

Even though many of the observations pointed out throughout the present dissertation could be attributed to performance effects, due to the fact that real data are analyzed, the facts discussed here are part of the core grammar of BP and have been used for linguists to provide general statements about the nature of grammatical knowledge for many years. If the effects and relevant factors described, analyzed and proposed here are due exclusively to performance, they should not constitute a reliable data base for such theories. Nonetheless, it is claimed throughout this dissertation that these factors are part of the probabilistic knowledge that speakers have about their language: they know that one meaning is more likely to be expressed and understood in one form rather than an alternative form that could lead to miscomprehension. Similar approaches have been proposed in many recent papers about anaphora resolution (Colonna et al. 2010, 2012, 2014, Baumann et al. 2014, Colonna et al. 2015), argument realization (Bresnan 2006, Bresnan et al. 2007, Bresnan &

Nikitina 2009), and filler-gap constructions (Gibson 2010).

Two theoretical traditions were brought together in this dissertation. The Generative Grammar Theories were presented and studied along with the anaphora resolution approach, which mainly combines psycholinguistics, functionalism and also Generative Grammar perspectives into a convergent theory of language comprehension and production. The main findings in the present dissertation suggest, however, that the anaphora resolution approach allows one to deal more comprehensively with the set of data presented here. This absolutely does not exclude the possibility of combining both theories and finding contexts in which a structural approach can explain the data, as proposed by Carminati (2002). In this vein, it is possible to find effects in a controlled environment that do not come up in a broader analysis of the phenomenon at stake. In section 2.7.3 in Chapter 2, for example, the embedded clause environment was significant statistically only when antecedents in the immediately preceding clause were taken into consideration. It is possible that the structural relation of c-command plays a role in the use of null subjects in BP, but discourse and semantic constraints that overcome this structural relation, such as Distance, Animacy, Specificity and Contrast, must be controlled for. As proposed throughout the present dissertation, other factors that were not studied here may also contribute to the choice of the anaphoric expression and its interpretation. Potentially any factor that can produce an imbalance among two or more antecedents would affect the interpretation and the preferred anaphoric option in a given context when other factors are in equilibrium. In some sense, this is the "division of labor" that Ariel (1990) proposes in her papers: given two (or more) anaphoric options in a language, multiple factors can affect the Salience of the potential antecedents, and consequently the choice and interpretation of anaphoric options can be differently biased. In the present dissertation, this proposal was incorporated and essentially claimed to be a by-product of the semantic and discourse structure of sentences.

In Chapter 1, after analyzing the widespread claim that impoverished inflectional morphology is directly related to availability of null referential subjects, it was claimed that the proposal based on the impoverishment of verbal morphology does not account for the distribution of null subjects in the system of present day BP. An alternative hypothesis was considered: inherent features of antecedents ([–animate] and [–specific]) could favor the use of null subjects and bias their interpretation. In a corpus investigation and in experimental data, these features showed significant effects, along with a weaker functional effect of the ambiguity of the verbal paradigm. These effects are

not accounted for in many morphosyntactic approaches previously proposed in the literature, but are possibly related to proposals based on the Referential Hierarchy and the specialization of the null subject in BP. However, at first sight, the results found there are surprising for anaphora resolution approaches, which claim that more salient referents are retrieved by less specified forms. At closer inspection, these data do not contradict the anaphora resolution inverse mapping relation: inherent features of antecedents are claimed to be different from discourse Salience of the antecedent, which is explored in Chapter 2.

In Chapter 2, the discourse properties of the relation between null and overt anaphoric subjects and their antecedents are studied. The chapter starts by an investigation of the most controlled discourse environment, namely verbal yes-no answers. These constructions are analyzed within the general theory of null subjects and they are claimed to be one more case in which the null subject is productive in BP. Differently from previous approaches based on structural assumptions, null subjects in verbal answers are taken to be constrained in the same way as null subjects in other sentences and utterances. The particular case of verbal answers is thus explained by a general discourse-based hypothesis: co-referential null subjects in BP are likely to be used in Maximally Contrastive utterances. This hypothesis is compared to the hypothesis of the reverse mapping of discourse Salience of antecedents in the anaphora resolution tradition. The data presented suggest that both hypotheses converge into a generalization that null subjects are likely to retrieve discursively predictable antecedents in Maximally Contrastive contexts, because in these contexts the background is highly obvious and the asserted information is optimally informative. A grammar design that can account for this generalization straightforwardly was proposed, along with the a proposal incorporating the constraints observed in Chapter 1. Finally, the predictive power of this theory and the way how multiple degrees of acceptability, frequency and processing difficulties are accounted for in such a grammar were explained.

In Chapter 3, a hypothesis for the generic interpretation of null subjects in BP was put forth. This case is particular interesting, since BP is the only Romance language in which this interpretation of null subjects is achieved and this interpretation could be thought to be a counterexample for the generalization proposed in Chapter 2, according to which null subjects are likely to retrieve discursively salient antecedents. However, as pointed out in Chapter 3, this case does not in fact provide counter-evidence for the present proposal but actually favors the hypothesis sketched

in this dissertation. Basically, generic null subjects conform to standard semantic restrictions on genericity, failing to be grounded in the discourse or world. The only particular feature of null generic subjects is that they do not appear in optimal contexts for the co-referential interpretation.

Finally, in Chapter 4, on-line evidence for the approach developed in the previous chapters is provided. In three different syntactic constructions, it is shown that the effect of the non-optimal discourse context for co-referential null subjects comes up when comparing them either with overt pronominal subjects or with the generic interpretation in the same context. All three constructions investigated in Chapter 4 are matters of controversy in the literature, since they are categorically taken not to license null subjects in BP. If different eye-tracking measures tap into different steps of processing in reading as suggested by Frazier (1987), Frazier & Flores d'Arcais (1989), Maia (2014), Clifton & Frazier (1989), among others, this would mean that, in the experiments reported here, the problem that shows up is one of pragmatic-semantic identification rather than of structural licensing, since the main effect was found only in the total RTs.[9] It was suggested that null subjects in BP are thus more discursively restricted than in other languages in which null subjects are allowed. Null subjects in BP can consequently be accounted for in the general theory of anaphora resolution, although some BP data suggest that some caveats with respect to this theory are necessary. The two main improvements put forward here are that (i) the notion of Salience should be replaced by Discourse Obviousness and (ii) the mapping hypothesis should somehow take into account principles of efficiency related to the discourse structure of the sentence in which the anaphoric form appears (such as Maximally Contrastive contexts).

Notational Variants or Different Theories?

Many different ways of formalizing the empirical observations presented in this dissertation can be proposed. However, some points of the proposal sketched here strongly diverge from those

---

[9]   As pointed out to me by Jeffrey Runner, if effects do not appear earlier on one or another experiment, it is never sure if there is a power problem (see critical writing by Vasishth & Nicenboim 2016, Nicenboim & Vasishth 2016). That is, a lack of effect does not make sure that there is no effect. However, the fact that, with the amount of participants in the Experiments 6, 7 and 8, effects are significant in later measures and not significant in earlier measures suggests that whatever is influencing the readers is related to late processing in these experiments; there is, Experiments 6, 7 and 8 failed to produce an effect on earlier measures, but did produce significant effects on late measures. This at least calls into question the approaches that would predict earlier (ungrammaticality) effects. Further studies can corroborate the null hypothesis in the relevant environments.

proposed elsewhere in the literature and must be highlighted.

First, the account proposed here is based on two intuitive observations about languages in general: (i) languages are intended to be meaningful and (ii) they (language systems not linguistic theories) are driven toward efficiency. Second, the design of the grammar proposed here is multi-layered, probabilistic and scalar. This design is not unheard of in the literature about computational linguistics, corpus linguistics and psycholinguistics. However, there seems to be a gap between these fields and some areas of theoretical linguistics. This dissertation puts all of them together and contrasts their predictions. Since the theory pursued here is empirically driven, as other which propose similar grammar designs, it was argued and shown that this style of grammar provides a better account for modeling speakers' linguistic knowledge about the phenomenon at stake, rather than a traditional single-layered binary grammar design.

Methodological Protocols

Although not all the experiments carried out focusing on the use and interpretation of null and pronominal subjects were reported in this dissertation (which included different methodology), those which have been described here aimed at providing not only contributions to the theoretical discussion about null and overt subjects in BP, but also to set up some important protocols for future research.

Corpus Research   First, BP lacks some important tools for corpus research. For example, annotated BP corpora in which missing arguments are somehow marked do not exist. Even the best automatic parsers of BP are not reliable, although reported precision is higher than 90%. However, for the kind of corpus research reported here, systematic annotation and pre-processing of texts is fundamental. The advantage of the use of VISL, reported in section 1.6 in Chapter 1, was that it provided a visually accessible way of searching for patterns. In future reseach, however, improved parsers are necessary.

Second, the establishment of criteria that are clearly accessible, and not theory-internal, is necessary. Moreover, these criteria must be graded (or scalar) when dealing with fuzzy semantic categories. This can lead to more fine-grained results than those obtained here.

Experimental Research   One of the main findings in the experimental research carried out here is the combination of three different tasks to support the conclusions about the data. Most of the experiments reported in this dissertation combined an acceptability judgment task along with an interpretation task, which highlighted that differences in the interpretation can lead to different outcomes in the acceptability of the sentences. The crucial results that show the importance of combining these tasks are in Experiment 5, where two different groups of interpreters were split, leading to different outcomes in terms of acceptability. Without this analysis, the selection of participants for the follow-up experiments would probably not be possible.

Another relevant methodological finding was the importance of the context sentence in some experiments, since, in many cases, manipulations of the context are not taken into consideration in experimental research. In the present dissertation, discourse structure is taken to play a central role in the phenomenon analyzed. However, in much previous research, the discourse structure was simply neglected. Incorporating this factor, or at least, controlling for this factor is necessary when dealing with almost any syntactic and interpretation issue.

4.8.1   Future Research

Many theoretical points remain under-explored in the present dissertation due to methodological and theoretical decisions that were taken in the beginning of the project. First, despite arguing for a scalar grammar, the degrees of Specificity and Animacy were studied in a binary fashion, and so were also Distance and Contrast. This was a decision taken before the first results came up. However, it is clear throughout this dissertation that some referents have a much lower degree of, for instance, Specificity, such as "Nobody" antecedents, which increases gradually, as for quantified antecedents, indefinite NPs, definite NPs, proper names and, finally, deictically established discourse persons. The lowest degree is much more likely to be retrieved by a null subject than medium low degrees, medium degrees and so on. Testing these degrees in both corpus and experimental research is a topic for future research.

Another point that can be further investigated is the effect of comparative structures on the use of missing arguments. It seems that comparative structures offer little to no resistance to dropping any part of the sentence that can be inferable from the previous context, almost regardless of

syntactic structure. As comparisons seem to be a context of Maximal Contrast, it would be worth studying these omissions in the light the proposal put forth here.

Bibliography

Abbott, B. (1995), 'Some Remarks on Specificity', *Linguistic Inquiry* **26**(2), 341–347.

Abeillé, A., Hemforth, B., Liu, Y., Winckel, E. & Gibson, E. (in prep), 'Some linguistic materials of which certain aspects may surprise some linguists: When subjects are easier to extract from'.

Adams, M. (1987*a*), 'From Old French to the Theory of Pro-Drop', *Natural Language and Linguistic Theory* **5**, 1–32.

Adams, M. (1987*b*), Old French, Null Subjects and Verb-Second Phenomena, PhD thesis, UCLA, Los Angeles, CA, USA.

Adger, D. (2003), *Core Syntax*, Oxford University Press, London, UK.

Alexiadou, A. & Anagnostopoulou, E. (1998), 'Parametrizing AGR: Word Order, V-Movement and EPP-Checking', *Natural Language and Linguistics Theory* **16**, 491–539.

Almor, A. (1996), Noun-phrase Anaphora and Focus: The Informational Load Hypothesis, PhD thesis, Brown University, Los Angeles, CA, USA.

Almor, A. (1999), 'Noun-phrase Anaphora and Focus: The Informational Load Hypothesis', *Psychological Review* **106**(4), 748–765.

Almor, A. (2000), Constraints and mechanisms in theories of anaphor processing, *in* M. W. Crocker, M. Pickering & C. Clifton, eds, 'Architectures and Mechanisms for Language Processing', Cambridge University Press, pp. 341–354.

Almor, A., de Carvalho Maia, J., Cunha Lima, M., Vernice, M. & Gelormini-Lezama, C. (2017), 'Language processing, acceptability, and statistical distribution: a study of null and overt subjects in Brazilian Portuguese', *Journal of Memory and Language* **92**, 98–113.

Alonso-Ovalle, L., Fernandes-Solera, S., Frazier, L. & Clifton, C. (2002), 'Null vs. overt pronouns and the Topic-Focus articulation in Spanish', *Rivista di Linguistica* **14**(2), 1–19.

Altmann, G. T. M. (1994), 'Regression-contingent analyses of eye movements during sentence processing : A reply to Rayner and Sereno', *Memory and Cognition* **22**, 286–290.

Amaral, P. M. & Schwenter, S. A. (2003), Contrast and the (Non-)Occurrence of Subject Pronouns, *in* D. Eddington, ed., 'Selected Proceedings of the 7th Hispanic Linguistics Symposium', Cascadilla Proceedings Project, Somerville, MA, USA, pp. 116–127.

Ariel, M. (1990), *Accessing Noun-Phrase Antecedents*, Routledge, London, UK.

Ariel, M. (1994), 'Interpreting anaphoric expressions: A cognitive versus a pragmatic approach', *Journal of Linguistics* **30**, 3–42.

Ariel, M. (2001), Accessibility Theory: An Overview, *in* T. Sanders, J. Schliperoord & W. Spooren, eds, 'Text representation', Springer Verlag, pp. 29–87.

Armstrong, M. E. (2008), Pragmatic Restrictions on Affirmative Response Choice in Brazilian Portuguese, *in* J. Bruhn de Garavito & E. Valenzuela, eds, 'Selected Proceedings of the 10th Hispanic Linguistics Symposium 2008', Colibri, Somerville, MA, USA, pp. 173–188.

Arnauld, A. (1662), *La logique ou L'Art du Penser*, critical edition by Pierre Clair and François Girbal, J. Vrin, Paris, France, 1981.

Arnold, J. E. (1998), Reference Form and Discourse Patterns, PhD thesis, Stanford University, Stanford, CA, USA.

Arnold, J. E. (1999), 'Marking salience: The similarity of topic and focus', Unpublished manuscript, University of Pennsylvania.

Avelar, J. & Cyrino, S. (2008), 'Locativos preposicionados em posição de sujeito: uma possível contribuição das línguas Bantu à sintaxe do português brasileiro', *Revista de Estudos Linguísticos da Universidade do Porto* **3**, 55–75.

Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008), 'Mixed-effects modeling with crossed random effects for subjects and items', *Journal of Memory and Language* **59**, 390–412.

Barbosa, M. d. P. P. (2009), 'Two kinds of subject pro', *Studia Linguistica* **63**(1), 2–58.

Barbosa, M. d. P. P. (2011), 'Partial pro drop as Null NP Anaphora', retrieved from http://ling.auf.net/lingbuzz/001717.

Barbosa, M. d. P. P., Duarte, M. E. L. & Kato, M. A. (2005), 'Null Subjects in European and Brazilian Portuguese', *Journal of Portuguese Linguistics* **4**(2), 11–52.

Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. (2013), 'Random effects structure for confirmatory hypothesis testing: Keep it maximal', *Journal of Memory and Language* **68**(3), 255–278.

Bates, D. & Maechler, M. (2009), 'Lme4: linear mixed-effects models using S4 classes, howpublished = R package version 0.999375-27, comment = '.

Bates, D., Maechler, M. & Bolker, B. (2011), 'Lme4: Linear mixed-effects models using S4 classes', R package version 0.999375-41. http://CRAN.R-Project.org/package=lme4.

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015), 'Fitting linear mixed-effects models using lme4', *Journal of Statistical Software* **67**(1), 1–48.

Baumann, P., Konieczny, L. & Hemforth, B. (2014), Conversational Implicatures in Anaphora Resolution: Alternative Constructions and Referring Expressions, *in* B. Hemforth, B. Mertins & C. Fabricius-Hansen, eds, 'Psycholinguistic Approaches to Meaning and Understanding Across Languages', Springer Verlag, pp. 197–212.

Belletti, A. & Rizzi, L. (1988), 'Psych-verbs and Theta-theory', *Natural Language and Linguistic Theory* **6**, 291–352.

Bennett, C. E. (1895), *New Latin Grammar*, ebook, retrieved from http://www.gutenberg.org/ebooks/15665?msg=welcome$_s$tranger.

Berlinck, R. d. A., Duarte, M. E. L. & de Oliveira, M. (2009), Predicação, *in* M. A. Kato & M. Nascimento, eds, 'Gramática do Portugues Culto Falado no Brasil: A Construção da Sentença', Vol. 3, Editora da UNICAMP, Campinas, SP, Brazil.

Berman, R. A. (1980), 'The case of an SVO language: Subjectless constructions in Modern Hebrew', *Language* **56**(3), 759–776.

Bianchi, V. & Figueiredo Silva, M. C. (1994), On Some Properties of Agreement-Object in Italian and Brazilian Portuguese, *in* M. L. Mazzola, ed., 'Issues and Theory in Romance Linguistics. Selected papers from the Linguistic Symposium on Romance Languages XXIII', Vol. 3, Georgetown University Press, Washington, DC, USA, pp. 181–197.

Biberauer, T., Holmberg, A., Roberts, I. & Sheehan, M. (2010), *Parametric Variation: Null Subjects in Minimalist Theory*, Cambridge University Press, Cambridge, MA, USA.

Bick, E. (2000), *The Parsing System "Palavras": Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*, University of Oklahoma Press, Aarhus, Denmark.

Bittner, D. (2007), Influence of animacy and grammatical role on production and comprehension of intersentential pronouns in German L1-acquisition, *in* D. Bittner & N. Gagarina, eds, 'Intersentential pronominal reference in child and adult language : Proceedings of the Conference on Intersentential Pronominal Reference in Child and Adult Language (ZAS Papers in Linguistics 48)', ZAS, Berlin , Germany, pp. 57–85.

Bobaljik, J. D. (2008), Where's phi? Agreement as a post-syntactic operation, *in* D. A. Daniel Harbour & S. Béjar, eds, 'Phi-Theory: Phi features across interfaces and modules', Vol. 4410, Oxford University Press, Oxford, UK, pp. 295–328.

Boland, J. E. (2004), Linking eye movements to sentence comprehension in reading and listening, *in* M. Carreiras & C. Clifton Jr, eds, 'The on-line study of sentence comprehension', Psychology Press, New York, NJ, USA, pp. 51–76.

Boland, J. E. & Blodgett, A. (2001), 'Understanding the constraints on syntactic generation: Lexical bias and discourse congruency effects on eye movements', *Journal of Memory and Language* **45**(3), 391–411.

Bonet, E. (1991), Morphology after Syntax: Pronominal Clitics in Romance, PhD thesis, Massachussetts Institute of Technology, Cambridge, MA, USA.

Bourdages, J. S. (1992), Parsing complex NPs in French, *in* H. Goodlick & M. S. Rochemond, eds, 'Island constraints: Theory, acquisition and processing', Kluwer, Dordrecht, The Netherlands, pp. 61–87.

Branco, A. (2007), Null Subjects Are Reflexives, Not Pronouns, *in* A. Branco, ed., 'Anaphora: Analysis, Algorithms and Applications. DAARC 2007. Lecture Notes in Computer Science', Vol. 4410, Springer, Berlin, Heidelberg, Germany.

Bresnan, J. (2006), Is syntactic knowledge probabilistic? experiments with the english dative alternation, *in* 'Linguistics in search of its evidential base, Studies in Generative Grammar. Mouton de Gruyter'.

Bresnan, J., Cueni, A., Nikitina, T. & Baayen, H. R. (2007), Predicting the dative alternation, *in* G. Boume, I. Kraemer & J. Zwarts, eds, 'Cognitive foundations of interpretation', Royal Netherlands Academy of Science, Amsterdam, The Netherlands, pp. 69–94.

Bresnan, J. & Nikitina, T. (2009), The gradience of the dative alternation, *in* L. Uyechi & L. H. Wee, eds, 'Reality exploration and discovery: Pattern interaction in language and life', CSLI Publications, Stanford, CA, USA, pp. 161–184.

Burzio, L. (1991), 'The Morphological Basis of Anaphora', *Journal of Linguistics* **27**(1), 81–105.

Burzio, L. (1996), The role of the antecedent in anaphoric relations, *in* R. Freidin, ed., 'Current issues in comparative grammar', Kluwer, Dordrecht, The Netherlands, pp. 1–45.

Burzio, L. (1998), Anaphora and soft constraints, *in* P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, eds, 'Is the best good enough?', Oxford University Press, pp. 93– 113.

Büring, D. (1999), Topic, *in* P. Bosch & R. van der Sandt, eds, 'Focus - Linguistic, Cognitive, and Computational Perspectives', Cambridge University Press, pp. 142–165.

Büring, D. (2003), 'On D-Trees, Beans, and B-Accents', *Linguistics and Philosophy* **26**(5), 511– 545.

Büring, D. (2005), *Binding theory. Cambridge Textbooks in Linguistics*, Cambridge University Press, Cambridge, MA, USA.

Büring, D. (2009), Towards a Typology of Focus Realization, *in* M. Zimmermann & C. Féry, eds, 'Information Structure', Oxford University Press, Oxford, pp. 177–205.

Cabredo-Hofherr, P. (2006), 'Arbitrary' pro and the theory of pro-drop, *in* P. Ackema, P. Brandt, M. Schoorlemmer & F. Weerman, eds, 'Agreement and arguments', Oxford University Press, Oxford, UK and New York, NJ, USA, pp. 230–257.

Callou, D. & Avelar, J. (2003), Estruturas com "ter" e "haver" em anúncios do século XIX, *in* T. Alkmim, ed., 'Para a história do português brasileiro', Vol. III, Humanitas-USP, São Paulo, SP, Brazil, pp. 47–68.

Caponigro, I. (2003), Free Not to Ask: On the Semantics of Free Relatives and Wh-Words Cross-linguistically, PhD thesis, University of California, California, LA, USA.

Caramazza, A., Grober, E., Garvey, C. & Yates, J. (1977), 'Comprehension of anaphoric pronouns', *Journal of Verbal Learning and Verbal Behavior* **16**, 601–609.

Cardinaletti, A. (1997), Subjects and Clause Structure, *in* L. Haegeman, ed., 'The New Comparative Syntax', Longman, Londres, UK.

Cardinaletti, A. & Starke, M. (1994), 'The Typology of Structural Deficiency on the Three Grammatical Classes', *Working Papers in Linguistics* **9**, 41–109.

Carlson, G. N. (1977), References to Kinds in English, PhD thesis, Massachussetts Institute of Technology, Cambridge, MA, USA, retrieved from http://semanticsarchive.net/Archive/jk3NzRlY/.

Carlson, L. (1983), *Dialogue Games: An approach to discourse analysis*, Reidel, Dordrecht, The Netherlands.

Carminati, M. N. (2002), The processing of Italian subject pronouns, PhD thesis, University of Massachussets, Cambridge, MA, USA, retrieved from http://scholarworks.umass.edu/dissertations/AAI3039345.

Carpenter, P. A. & Just, M. A. (1977), Eye fixations and comprehension, *in* 'Cognitive Processes in Comprehension', Lawrence Erlbaum Associates, Hillsdale, NJ, USA, pp. 109–139.

Chafe, W. L. (1976), Givenness, contrastiveness, definiteness, subjects, topics and points of view, *in* C. Li, ed., 'Subject and Topic', New York Academy press, New York, NJ, USA, pp. 26–55.

Chierchia, G. (1995), *Dynamics of Meaning. Anaphora, Presupposition and the Theory of Grammar, publisher=University of Chicago Press, address=Chicago, CH, USA, pages=270.*

Chomsky, N. (1976), On the nature of language, *in* S. R. Harnad, H. D. Steklis & J. Lancaster, eds, 'Origins and Evolution of Language and Speech', Vol. 280, Annals of the New York Academy of Sciences, pp. 46–57.

Chomsky, N. (1981), *Lectures in Government and Binding. Studies in generative grammar 9*, Foris, Dordrecht: The Netherlands.

Chomsky, N. (1986), *Knowledge of Language: Its Nature, Origin and Use*, Praeger, New York: USA.

Chomsky, N. (1995), *The Minimalist Program*, MIT Press, Cambridge, MA, USA.

Chomsky, N. (2000), Minimalist inquiries: the framework, *in* R. Martin, D. Michaels & J. Uriagereka, eds, 'Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik', MIT Press, Cambridge, MA, USA, pp. 89–155.

Chomsky, N. (2001), Derivation by phase, *in* M. Kenstowicz, ed., 'Ken Hale: A life in language', MIT Press, Cambridge, MA, USA, pp. 1–52.

Chung, S. (2000), *The Design of Agreement: Evidence from Chamorro*, University of Chicago Press, Chicago and London, USA and UK.

Cinque, G. (1988), 'On si constructions and the theory of arb', *Linguistic Inquiry* **19**, 521–581.

Clark, H. H. & Haviland, S. E. (1977), Comprehension and the given-new contract, *in* R. Freedle, ed., 'Discourse production and comprehension', Erlbaum, Hillsdale, NJ, USA, pp. 1–40.

Clifton, C. & Frazier, L. (1989), Comprehending sentences with long-distance dependencies, *in* M. Tanenhaus & G. Carlson, eds, 'Linguistic structure in language Processing', Kluwer Academic Press, Dordrecht, The Netherlands.

Colonna, S., Schimke, S. & Hemforth, B. (2010), Le rôle de la structure informationnelle dans l'interprétation d'une anaphore pronominale inter-phrastique, *in* F. Neveu, V. Muni Toke, J. Du-

rand, T. Klingler, L. Mondada & S. Prévost, eds, 'Actes du Congrès Mondial de Linguistique Française', Colibri, Paris, France, pp. 1489–1499.

Colonna, S., Schimke, S. & Hemforth, B. (2012), 'Information structure effects on anaphora resolution in German and French: A cross-linguistic study of pronoun resolution', *Linguistics* **50**(5), 991– 1013.

Colonna, S., Schimke, S. & Hemforth, B. (2014), Information structure and pronoun resolution in German and French: Evidence from the visual-world paradigm, *in* B. Hemforth, B. Schmiedtová & C. Fabricius-Hansen, eds, 'Psycholinguistic approaches to meaning and understanding across languages', Academic Press, pp. 175–195.

Colonna, S., Schimke, S. & Hemforth, B. (2015), 'Different effects of focus in intra- and intersentential pronoun resolution in German', *Language, Cognition and Neuroscience* .

Contreras, H. (1991), On the Position of Subjects, *in* S. Rothstein, ed., 'Perspectives on Phrase Structure: Heads and Licensing, Syntax and Semantics', Vol. 25, Academic Press, San Diego, CA, USA.

Corbett, A. & Chang, F. R. (2005), 'Pronoun disambiguation: Accessing potential antecedents', *Memory and Cognition* **11**, 283–294.

Correa Soares, E. (2016), Grammatical change and hindcast model statistics - A comparison between Medieval French and Brazilian Portuguese, *in* A. Botinis, ed., 'Proceedings of 7th Tutorial and Research Workshop on Experimental Linguistics', CSLI Publications, Saint Petersburg, Russia, pp. 171–175.

Corrêa, L. M. S. (1998), 'Acessibilidade e paralelismo na interpretação do pronome sujeito e o contraste pro/pronome em português', *DELTA* **14**(2), 295–329.

Costa, A. (2003), *Processamento de Frases em Portugues Europeu: Aspectos Cognitivos e Linguísticos Implicados na Compreensão da Língua Escrita*, Fundação Calouste Gulbenkian, Lisbon, Portugal.

Costa, A., Faria, I. & Kail, M. (2004), Semantic and Syntactic Cues' Interaction on Pronoun Resolution in European Portuguese, *in* A. Branco, McEnery & R. Mitkov, eds, '5th Discourse Anaphora Resolution Colloquium', Colibri, Lisbon, Portugal, pp. 45–50.

Costa, A., Faria, I. & Matos, G. (1998), Ambiguidade referencial na identificação do sujeito em estruturas coordenadas, *in* 'Textos seleccionados do XIII Encontro Nacional da Associação Portuguesa de Linguística', Colibri, Lisboa, Portugal, pp. 173–188.

Cowan, J. R. (1980), The significance of parallel function in the assignment of anaphora, *in* K. J. & A. E. Ojeda, eds, 'Papers from the parasession on pronouns and anaphora', Chicago Linguistic Society, Chicago, IL, USA.

Crain, S. & Fodor, J. D. (1985), How can grammars help parsers?, *in* D. Dowty, L. Kartunnen & A. Zwicky, eds, 'Natural language parsing: psycholinguistic, computational, and theoretical perspectives', Cambridge University Press, Cambridge, UK, pp. 94–128.

Crain, S. & Fodor, J. D. (1993), Competence and performance in child language, *in* E. Dromi, ed., 'Language and Cognition: A Developmental Perspective', Ablex, Norwood, NJ, USA, pp. 141–171.

Crawley, R., Stevenson, R. J. & Kleinman, D. (1990), 'Grammatical determinants of ambiguous pronoun resolution', *Journal of Psycholinguistic Research* **19**, 245–264.

Cresti, D. (1995), Indefinite Topics, PhD thesis, Massachussets Institute of Technology, Cambridge, MA, USA.

Creus, S. & Menuzzi, S. d. M. (2004), 'O papel do gênero na alternância entre objeto nulo e pronome pleno em português brasileiro', *Revista da ABRALIN* **3**(1-2).

Crone, P. & Frank, M. C. (2016), Inferring generic meaning from pragmatic reference failure, *in* 'Proceedings of the 37th annual meeting of the cognitive science society'.

Cunnings, I., Patterson, C. & Felser, C. (2014), 'Variable binding and coreference in sentence comprehension: evidence from eye movements', *Journal of Memory and Language* **71**(1), 39–56.

Cutler, A. & Fodor, J. A. (1979), 'Semantic focus and sentence comprehension', *Cognition* **7**, 49–59.

Cyrino, S. M. L., Duarte, M. E. L. & Kato, M. A. (2000), Visible subjects and invisible clitics in Brazilian Portuguese, *in* M. A. Kato & E. V. Negrão, eds, 'Brazilian Portuguese and the Null Subject Parameter', Vervuert-Iberoamericana, Frankfurt, Germany, pp. 55–104.

Dahl, O. (1975), On generics, *in* E. Keenan, ed., 'Formal Semantics of Natural Language', Cambridge University Press, Cambridge, MA, USA, pp. 99–111.

Dahl, O. & Fraurud, K. (1996), Animacy in Grammar and Discourse, *in* T. Fretheim & J. K. Gundel, eds, 'Reference and Referent Accessibility', Cambridge University Press, Pragmatics and Beyond New Series 38, pp. 47–64.

Dalmi, G. (2014), All-in-One: Generic Inclusive Null Subjects in Hungarian, *in* R. E. Santana-LaBarge, ed., 'Proceedings of the 31st West Coast Conference on Formal Linguistics', Mouton de Gruyter, Somerville, MA, USA, pp. 115–123.

Davidson, D. (2001), *Essays on Actions and Events*, Clarendon Press, Oxford, UK.

de la Fuente, I. (2016), Putting pronoun resolution in context: The role of syntax, semantics and pragmatics in pronoun interpretation, PhD thesis, Université Paris Diderot, Paris, France.

de la Fuente, I. & Hemforth, B. (2013), Topicalization and focusing effects on subject and object pronoun resolution in Spanish, *in* J. Cabrelli Amaro, G. Lord, A. de Prada Pérez & J. E. Aaron, eds, 'Selected Proceedings of the Hispanic Linguistics Symposium 2012', Cascadilla Press, Somerville, MA, USA, pp. 57–85.

de la Fuente, I., Hemforth, B., Colonna, S. & Schimke, S. (2016), The role of syntax, semantics, and pragmatics in pronoun resolution: A cross-linguistic overview, *in* A. Holler & K. Suckow, eds, 'Empirical Perspectives on Anaphora Resolution: Information Structural Evidence in the Race for Salience', Mouton de Gruyter.

de Oliveira, M. (1996), Respostas Assertivas e sua Variação nas Línguas Românicas: o seu papel na aquisição, PhD thesis, UNICAMP, Campinas, SP, Brazil.

Deemter, K. v. (2008), 'What's new? A semantic perspective on sentence accent', *Journal of Semantics* pp. 1–31.

Delgrande, J. P. (1987), 'A first-order conditional logic for prototypical properties', *Artificial Intelligence* **33**, 105–130.

Delgrande, J. P. (1988), 'An approach to default reasoning based on a first-order conditional logic: Revised report', *Artificial Intelligence* **36**, 63–90.

Destrual, E., Velleman, D., Onea, E., Bumford, D., Xue, J. & Beaver, D. (2013), A cross-linguistic study of the non-at-issueness of exhaustive inferences, *in* F. Schwarz, ed., 'Under Contract. Experimental Perspectives on Presuppositions', Springer's Studies in Theoretical Psycholinguistics Series.

Diesing, M. (1992), *Indefinites*, MIT Press, Cambridge, MA, USA.

Discole, A. (1997), *De la construction (syntaxe)*, Vol. 1, Librairie philosophique J. VRIN, Paris, France.

Dobrovie-Sorin, C. (1998), 'Impersonal se constructions in Romance and the passivization of unergatives', *Linguistic Inquiry* **29**(3), 399–437.

Dowty, D. & Jacobson, P. (1989), Agreement as a semantic phenomenon, *in* J. Powers & K. de Jong, eds, 'Proceedings of Fifth Eastern States Conference on Linguistics (ECSOL '88)', Ohio State University, Columbus, Ohio, USA, pp. 95–101.

Dryer, M. S. (2013), Expression of Pronominal Subjects, *in* M. S. Dryer & M. Haspelmath, eds, 'The World Atlas of Language Structures Online', Leipzig: Max Planck Institute for Evolutionary Anthropology, Available online at http://wals.info/chapter/101, Accessed on 2017-08-06, pp. 69–82.

Duarte, M. E. L. (1993), Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil, *in* I. Roberts & M. A. Kato, eds, 'Português brasileiro: Uma viagem diacrônica. Homenagem a Fernando Tarallo', Editora da Unicamp, Campinas, Brazil, pp. 107–129.

Duarte, M. E. L. (1995), A Perda do Princípio "Evite Pronome" no Português Brasileiro, PhD thesis, Universidade Estadual de Campinas, Campinas, SP, Brazil.

Duarte, M. E. L. (2000), The loss of the 'avoid pronoun' principle in Brazilian Portuguese, *in* M. A. Kato & E. V. Negrão, eds, 'Brazilian Portuguese and the Null Subject Parameter', Vervuert-Iberoamericana, Frankfurt, Germany, pp. 17–36.

Duarte, M. E. L. (2003), O sujeito expletivo e as construções existenciais, *in* J. Roncarati, C.; Abracado, ed., 'Português brasileiro: contato linguístico e história', 7 Letras, Rio de Janeiro, RJ, Brazil, pp. 123–35.

Duarte, M. E. L. (2012), *O sujeito em peças de teatro (1833-1992): Estudos diacrônicos*, Parábola Editorial, São Paulo, SP, Brazil.

Duarte, M. E. L. (2015), 'Avanço no estudo da mudança sintática associando a teoria da variação e mudança e a teoria de Princípios e Parâmetros', *Cadernos de Estudos Lingüísticos* **57**(1).

Duarte, M. E. L., Barbosa, M. d. P. P. & Kato, M. A. (2001), Sujeitos indeterminados em PE e PB, *in* 'Anais do II Congresso Internacional da ABRALIN', Vol. I, Colibri, pp. 405–409.

Duarte, M. E. L., Mourão, G. & Guimarães, L. (2012), A retomada dos sujeitos proposicionais: categoria vazia ou demonstrativo?, *in* M. E. L. Duarte, ed., 'O sujeito em peças de teatro (1833-1992): Estudos diacrônicos', Parábola Editorial, São Paulo, SP, Brazil, pp. 69–82.

Duarte, M. E. L., Mourão, G. & Santos, H. (2012), Os sujeitos de terceira pessoa: revisitando Duarte 1993, *in* M. E. L. Duarte, ed., 'O sujeito em peças de teatro (1833-1992): Estudos diacrônicos', Parábola Editorial, São Paulo, SP, Brazil, pp. 21–44.

Dwivedi, V., Phillips, N. A., Einagel, S. & Baum, S. R. (2010), 'The neural underpinnings of semantic ambiguity and anaphora', *Brain Res* **1311**, 93–109.

Déchaine, R.-M. & Wiltschko, M. (2002), 'Decomposing pronouns', *Linguistic Inquiry* **33**, 409–442.

D'Alessandro, R. (2007), *Impersonal si constructions*, Mouton de Gruyter, Berlin, Germany/New York, NJ, USA.

Ebert, C. (2009), *Quantificational topics: A scopal treatment of exceptional wide scope phenomena*, Springer, New York, NJ, USA.

Egerland, V. (2003), 'Impersonal pronouns in Scandinavian and Romance', *Working papers in Scandinavian Syntax* **71**, 75–102.

Embick, D. & Noyer, R. (2007), Distributed Morphology and the Syntax/Morphology Interface, *in* A. Holler & K. Suckow, eds, 'The Oxford Handbook of Linguistic Interfaces', Oxford University Press, Oxford, UK, pp. 289–324.

Enç, M. (1991), 'The semantics of specificity', *Linguistic Inquiry* **22**(1), 1–25.

Erteschik-Shir, N. (1997), *The Dynamics of Focus Structure*, Springer, Cambridge, MA, USA.

Ewert, A. (1943), *The French Language*, Faber and Faber, London, UK.

Falco, M. (2002), 'Specificity: the syntax/semantics mapping - A research project', ms. Center for Mind/Brain Sciences (CIMeC), University of Trento.

Falk, S. (2014), 'On the notion of salience in spoken discourse - prominence cues shaping discourse structure and comprehension', *TIPA. Travaux interdisciplinaires sur la parole et le langage* **30**.

Featherston, S. (2005), 'Magnitude estimation and what it can do for your syntax: Some wh-constraints in German', *Lingua* **115**, 1525–1550.

Felser, C., Clahsen, H. & Münte, T. F. (2003), 'Storage and Integration in the Processing of Filler-Gap Dependencies: An ERP Study of Topicalization and Wh-Movement in German', *Brain Lang* **87**(3), 345–354.

Fernandes, E., Luegi, P., Correa Soares, E., De la Fuente, I. & Hemforth, B. (2016), 'How to turn Brazilians into Europeans: Global and local exposure effects on co-reference in European and Brazilian Portuguese', Poster presented in Architectures and Mechanisms for Language Processing (AMLaP) XXII, Bilbao, Spain.

Fernandes, E., Luegi, P., Correa Soares, E., De la Fuente, I. & Hemforth, B. (to appear), 'Adaptation in Pronoun Resolution: Evidence from Brazilian and European Portuguese', manuscript

accepted for publication in Journal of Experimental Psychology: Learning, Memory, and Cognition.

Fernandez-Soriano, O. (1989), Strong pronouns in null subject languages and the Avoid Pronoun Principle, *in* 'MIT Working Papers in Linguistics,', Vol. 11, MIT, Cambridge, MA, USA, pp. 228–240.

Ferreira, F. (1994), 'Choice of passive voice is affected by verb type and animacy', *Journal of Memory and Language* **33**, 715–736.

Ferreira, M. B. (2000), Argumentos Nulos em Português Brasileiro, Master's thesis, Universidade de São Paulo, São Paulo, SP, Brazil.

Ferreira, M. B. (2004), Hyperraising and Null Subjects in Brazilian Portuguese, *in* 'MIT Working Papers in Linguistics 47: Collected Papers on Romance Syntax', Colibri, Somerville, MA, USA, pp. 57–85.

Figueiredo-Silva, M. C. (1994), La position sujet en portugais brésilien, PhD thesis, Université de Genève, Published in Portuguese translation as Figueiredo-Silva, C. (1996) A Posição Sujeito no Português Brasileiro : Frases finitas e infinitivas. Campinas : Editora da Unicamp.

Figueiredo-Silva, M. C. (2000), Main and embedded null subjects in Brazilian Portuguese, *in* M. A. Kato & E. V. Negrão, eds, 'Brazilian Portuguese and the Null Subject Parameter', Vervuert-Iberoamericana, Frankfurt, Germany, pp. 17–36.

Filiaci, F. (2010), Null and overt subject biases in Spanish and Italian: a cross-linguistic comparison, *in* C. Borgonovo, M. Español-Echevarría & P. Prévost, eds, 'Selected Proceedings of the 12th Hispanic Linguistic Symposium', Cascadilla Proceedings Project, Somerville, MA, USA, pp. 171–182.

Filiaci, F., Sorace, A. & Carreiras, M. (2013), 'Anaphoric biases of null and overt subjects in Italian and Spanish: a cross-linguistic comparison', *Language, Cognition and Neuroscience* **29**(7), 1–19.

Fillmore, C. (1982), Frame semantics, *in* 'Linguistics in the Morning Calm', Soeul, South Korea, pp. 111–137.

Frazier, L. (1987), 'Syntactic processing: Evidence from Dutch', *Natural Language and Linguistic Theory* **5**, 519–560.

Frazier, L. & Flores d'Arcais, G. (1989), 'Filler-driven parsing: a study of gap filling in Dutch', *Journal of Memory and Language* **28**, 331–344.

Frazier, L. & Rayner, K. (1982), 'Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences', *Cognitive Psychology* **14**, 178–210.

Frege, G. (1989), 'Über sinn und bedeutung', *Zeitschrift für Philosophie und philosophische Kritik* **100**, 25–50.

Frey, W. (2004), 'A medial topic position for German', *Linguistische Berichte* **198**, 153–190.

Gagarina, N. (2007), The hare hugs the rabbit. He is white ... Who is white? Pronominal anaphora in Russian, *in* D. Bittner & N. Gagarina, eds, 'Intersentential pronominal reference in child and adult language : Proceedings of the Conference on Intersentential Pronominal Reference in Child and Adult Language (ZAS Papers in Linguistics 48)', ZAS, Berlin , Germany, pp. 57–85.

Galves, C. (1987), 'A sintaxe do português brasileiro', *Ensaios Lingüísticos* **13**, 31–50.

Galves, C. (1993), O enfraquecimento da concordância no PB, *in* I. Roberts & M. A. Kato, eds, 'Português brasileiro: Uma viagem diacrônica. Homenagem a Fernando Tarallo', Editora da Unicamp, Campinas, Brazil, pp. 387–408.

Galves, C. (1997), La Syntaxe Pronominale du Portugais Brésilien et la Typologie des Pronoms, *in* A. Zribi-Hertz, ed., 'Les Pronoms: morphologie, syntaxe et typologie', Presses Universitaires de Vincennes (PUV), Saint-Denis, France, pp. 11–34.

Galves, C. (1998), 'La grammaire du portugais brésilien', *Língua e Instrumentos Linguísticos* **1**, 79–96.

Garnham, A. (2001), *Mental Models and the Interpretation of Anaphora. Essays in Cognitive Psychology*, Psychology Press, Hove, East Sussex, UK.

Gast, V. & van der Auwera, J. (2000), Towards a distributional typology of human impersonal pronouns, based on data from European languages, *in* D. Bakker & M. Haspelmath, eds, 'Languages across Boundaries: Studies in the Memory of Anna Siewierska', Mouton de Gruyter, Berlin, Germany, pp. 119–158.

Gathercole, S. E. (2006), 'Nonword repetition and word learning: The nature of the relationship', *Applied Psycholinguistics* **27**(4), 513–543.

Gernsbacher, M. A. (1989), 'Mechanisms that improve referential access', *Cognition* **32**, 99–156.

Gernsbacher, M. A. & Hargreaves, D. J. (1988), 'Accessing sentence participants: The advantage of first mention', *Journal of Memory and Language* **27**, 699–717.

Gibson, E. (1998), 'Linguistic complexity: Locality of syntactic dependencies', *Cognition* **68**, 1–76.

Gibson, E. (2010), Dependency locality theory: A distance-based theory of sentence processing difficulty, *in* A. Marantz, Y. Miyashita & W. O'Neil, eds, 'Image, Language, Brain: Papers from the first Mind Articulation Project Symposium', MIT Press, Cambridge, MA, USA, pp. 95–126.

Gildersleeve, B. L. & Lodge, G. (1905), *Latin Grammar*, Univerity Publishing Company, New York, New Orleans, Boston, USA.

Gilligan, G. M. (1987), A cross-linguistic approach to the pro-drop parameter, PhD thesis, University of Southern California, San Diego, CA, USA.

Ginzburg, J. (1996), Dynamics and the semantics of dialogue, *in* J. Seligman, ed., 'Language, Logic and Computation 1', Vol. 6, CSLI Lecture Notes, Standford, CA, USA, pp. 221–237.

Ginzburg, J. (2012), *The Interactive Stance: Meaning for Conversation*, Oxford University Press, Oxford,UK.

Ginzburg, J. (to appear), The Semantics of Dialogue, *in* M. Aloni & P. Dekker, eds, 'Cambridge Handbook of Formal Semantics', Cambridge University Press.

Givón, T. (1976), Topic, pronoun and grammatical agreement, *in* C. N. Li, ed., 'Subject and Topic', Vol. 6, Academic Press, New York, USA.

Givón, T. (1983), *Topic Continuity in Discourse - A Quantitative Cross Language Study*, John Benjamins, Amsterdam, The Netherlands, Philadelphia, USA.

Givón, T. (1985), Iconicity, isomorphism, and non-arbitrary coding in syntax, *in* J. Haiman, ed., 'Iconicity in Syntax: Proceedings of a symposium on iconicity in syntax', Vol. 6, Typological Studies in Language, Standford, CA, USA, pp. 187–202.

Givón, T. (1991), 'Isomorphism in the Grammatical Code: Cognitive and Biological Considerations', *Studies in Language* **15**(1), 85–114.

Goldberg, A. E. (2013), Backgrounded constituents cannot be "extracted", *in* J. Sprouse & N. Hornstein, eds, 'Experimental Syntax and islands effects', Cambridge University Press, Cambridge, MA, USA, pp. 221–238.

Gonçalves, A. (2000), 'Uma análise de sujeitos genéricos nulos de terceira pessoa do singular em sentenças finitas raízes no Português Brasileiro'.

Gonçalves, A. (2002), Uma análise de sujeitos genéricos nulos de terceira pessoa do singular em sentenças finitas raízes no Português Brasileiro, *in* 'Working Papers em Linguística', Vol. 6, UFSC, Florianópolis, SC, Brazil.

Gordon, M. (in press), The intonational realization of contrastive focus in chickasaw, *in* C. Lee, M. Gordon & D. Buring, eds, 'Intonational and Meaning from a cross-linguistic perspective', Kluwer, Dordrecht, The Netherlands.

Gordon, P. C., Grosz, B. J. & Gilliom, L. A. (1993), 'Pronouns, names, and the centering of attention in discourse', *Cognitive Science* **17**(3), 311–348.

Greenough, J. B. & Allen, J. H. (1888/1903), *New Latin Grammar*, Ginn and Company, The Athenaeum Press, Boston, MA, USA and London, UK.

Grice, H. P. (1975), Logic and Conversation, *in* P. Cole & J. Morgan, eds, 'Syntax and Semantics 3: Speech Acts', Vol. 6, CSLI Lecture Notes, London, UK, pp. 41–58.

Grimshaw, J. & Samek-Lodovici, V. (1998), Optimal subjects and subject universals. Is the best good enough?, *in* P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, eds, 'Image,

Language, Brain: Papers from the first Mind Articulation Project Symposium', MIT Press and MIT Working Papers in Linguistics, Cambridge, MA, USA, pp. 193–219.

Grober, E., Beardsley, W. & Caramazza, A. (1990), 'Parallel function strategy in pronoun assignment', *Cognition* **6**, 117–133.

Grodzinsky, Y. & Reinhart, T. (1993), 'The Innateness of Binding and Coreference', *Linguistic Inquiry* **24**(1), 69–101.

Groenendijk, J. & Stokhof, M. (1984), *Studies in the Semantics of Questions and the Pragmatics of Answers*, Akademish Proefschrift, Amsterdam, The Netherlands.

Grosz, B. J. & Snider, C. (1986), 'Attention, intentions, and the structure of discourse', *Computational Linguistics* **12**(3), 175–204.

Grosz, B. J., Weinstein, S. & Joshi, A. K. (1995), 'Centering: A framework for modeling the local coherence of discourse', *Computational Linguistics* **21**, 203–225.

Guesser, S. (2007), 'Il soggetto nullo incassato in Portoghese Brasiliano: problemi legati al movimento a posizioni tematiche', *Rivista di Grammatica Generativa* pp. 39–64.

Guesser, S. (2008), 'Embedded Null Subjects in Brazilian Portuguese', *Nanzan Linguistics: Special Issue 5* pp. 55–78.

Guimarães, M. & Rodrigues, C. (2002), 'Locative Determiner and Argumental Pseudo-Adverbs', ms. University of Maryland, College Park.

Gundel, J. (2010), 'Reference and Accessibility from a Givenness Hierarchy Perspective', *International Review of Pragmatics* **2**(2), 148–168.

Gundel, J., Hedberg, N. & Zacharski, R. (1993), 'Cognitive status and the form of referring expressions in discourse', *Language and Cognitive Processes* **69**, 274–307.

Gundel, J. K. (1988), Universals of topic-comment structure, *in* M. Hammond, E. Moravcsik & J. Worth, eds, 'Studies in Syntactic Typology', John Benjamins, Amsterdam, The Netherlands, pp. 209–239.

Haendler, Y. (2017), 'Effects of pronouns' referential properties on relative clause processing', Talk given at LingLunch Seminar, Université Paris Diderot, March 23rd 2017.

Halle, M. & Marantz, A. (1993), Distributed Morphology and the Pieces of Inflection, *in* A. Marantz, Y. Miyashita & W. O'Neil, eds, 'The View from Building 20', MIT Press, Cambridge, MA, USA, pp. 111–176.

Halle, M. & Marantz, A. (1994), Some Key Features of Distributed Morphology, *in* 'MIT Working Papers in Linguistics - Papers on Phonology and Morphology', MIT Press, Cambridge, MA, USA, pp. 275–288.

Halliday, M. A. K. (1967), 'Notes on transitivity and theme in English, Part 2', *Journal of Linguistics* **2**, 199–244.

Halliday, M. & Hasan, R. (1976), *Cohesion in English*, Longman, London, UK.

Hamblin, C. L. (1973), 'Questions in Montague English', *Foundations of Language: International Journal of Language and Philosophy* **10**(1), 41–53.

Harley, H. & Noyer, R. (1999), 'State-of-the-Article: Distributed Morphology', *GLOT International* **4-4**, 3–9.

Harley, H. & Ritter, E. (2002), 'Person and number in pronouns: a feature-geometric analysis', *Language* **78**, 482–526.

Hawkins, J. A. (1978), *Definiteness and Indefiniteness*, Croom Helm, London, UK.

Hawkins, J. A. (1991), 'On (in)definite articles: implicatures and (un)grammaticality prediction', *Journal of Linguistics* **27**, 405–442.

Hawkins, J. A. (1994), *A performance theory of order and constituency*, Cambridge University Press, Cambridge, MA, USA.

Hawkins, J. A. (2004), *Efficiency and complexity in grammars*, Oxford University Press, Oxford, UK.

Hawkins, J. A. (2014), *Cross-linguistic variation and efficiency*, Oxford University Press, Oxford, UK.

Hayes, B. (2000), Gradient well-formedness in Optimality Theory, *in* J. Dekkers, ed., '5th Discourse Anaphora Resolution Colloquium', Oxford University Press, Oxford, UK, pp. 88–120.

Heim, I. (1998), Anaphora and Semantic Interpretation: A Reinterpre- tation of Reinhart's Approach, *in* U. Sauerland & O. Percus, eds, 'The Interpretive Tract, MIT Working Papers in Linguistics 25', SfS Report 07-93, Universität Tübingen, Tübingen, Germany, pp. 205–246.

Heim, I. R. (1982), The Semantics of Definite and Indefinite Noun Phrases, PhD thesis, University of Massachusetts, MA, USA.

Hoekstra, J. (2010), 'On the impersonal pronoun men in Modern West Frisian', *Journal of Comparative Germanic Linguistics* **13**.

Holmberg, A. (2001), 'The syntax of yes and no in Finnish', *Studia Linguistica* **55**, 141–174.

Holmberg, A. (2003), Yes/No Questions and the Relation between Tense and Polarity in English and Finnish, *in* P. Pica, ed., 'Linguistic Variation Yearbook 3', pp. 43–68.

Holmberg, A. (2005), 'Is There a Little Pro? Evidence from Finnish', *Linguistic Inquiry* **36**, 533–564.

Holmberg, A. (2007), 'Null subjects and polarity focus', *Studia Linguistica* **61**, 212–236.

Holmberg, A. (2013), 'The syntax of answers to polar questions in English and Swedish', *Lingua* **128**, 31–50.

Holmberg, A. (2014), 'Answering Yes/No-Questions in English and Other Languages', *The EFL Journal 5* **2**, 23–42.

Holmberg, A. (2016), *The Syntax of Yes and No*, Oxford University Press, Oxford, UK.

Holmberg, A. (to appear), 'The syntax of negative questions and their answers', ms. Newcastle University.

Holmberg, A., Nayudu, A. & Sheehan, M. (2009), 'Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish, and Marathi', *Studia Linguistica* pp. 59–97.

Holmberg, A. & Nikanne, U. (2002), Expletives, subjects, and topics in Finnish, *in* P. Svenonius, ed., 'Subjects, expletives, and the EPP', Oxford University Press, New York and Oxford, USA and UK, pp. 71–106.

Holmberg, A., Nikanne, U., Oraviita, I., Reime, H. & Trosterud, T. (1993), The structure of INFL and the finite clause in Finnish, *in* A. Holmberg & U. Nikanne, eds, 'Case and other functional categories in Finnish syntax', Mouton de Gruyter, Berlin, Germany, pp. 176–206.

Holmberg, A. & Phimsawat, O.-U. (2015), Generic pronouns and phi-features: Evidence from Thai, *in* 'Newcastle and Northumbria Working Papers in Linguistics 21', retrieved from http://www.ncl.ac.uk/linguistics/research/workingpapers/Volume21.1.htm.

Holmberg, A. & Phimsawat, O.-U. (2016), 'Minimal Subjects', Ms. University of Newcastle, retrieved from http://ling.auf.net/lingbuzz/002964.

Hornstein, N. (1999), 'Movement and Control', *Linguistic Inquiry* **30**, 69–96.

Hornstein, N. (2001), *Move! A Minimalist Theory for Construal*, Blackwell, Oxford, UK.

Hornstein, N. (2002), On Control, *in* R. Hendriks, ed., 'Contemporary Grammatical Theory', Blackwell, Oxford, UK.

Huang, C.-T. J. (1982), Logical relations in Chinese and the theory of grammar, PhD thesis, MIT, Massachussets, MA, USA.

Huang, C.-T. J. (1984), 'On the distribution and reference of empty pronouns', *Linguistic Inquiry* **15**, 531–574.

Huang, C.-T. J. (1989), Pro-drop in Chinese, *in* O. Jaeggli & K. Safir, eds, 'The Null Subject Parameter', Kluwer, Dordrecht, The Netherlands, pp. 11–34.

Huang, Y. (1994), *The Syntax and Pragmatics of Anaphora*, Cambridge University Press, Cambridge, UK.

Huang, Y. (2000), *A Cross-linguistic study*, Oxford University Press, Oxford, UK.

Hyams, N., Mateu, V., Ortfitelli, R., Putnam, M., Rothman, J. & Sánchez, L. (2015), Parameters in language acquisition and contact, *in* F. A., M. J. & P. M., eds, 'Contemporary Handbook of Linguistic Parameters', Bloomsbury, London, UK, pp. 353–375.

Jackendoff, R. (1972), *Semantic Interpretation in Generative Grammar*, MIT Press, Cambridge, MA, USA.

Jackendoff, R. (1990), *Semantic Structures*, MIT Press, Cambridge, MA, USA.

Jaeger, T. F. (2008), 'Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models', *Journal of Memory and Language* pp. 434–446.

Jaeggli, O. (1986), 'Arbitrary plural pronominals', *Natural Language and Linguistic Theory* **4**, 43–76.

Jaeggli, O. & Safir, K. (1989), *The Null Subject Parameter*, Kluwer, Dordrecht, The Netherlands.

James, D. (1972), 'Some aspects of the syntax and semantics of interjections', University of Massachusetts at Amherst, unpublished manuscript, Amherst, MA, USA.

Jensen, T. J. (2009), 'Generic variation? Developments in use of generic pronouns in late 20th century spoken Danish', *Acta Linguistica Hafniensia* **41**, 83–115.

Jespersen, O. (1922), *Language: its nature, development and origin*, G. Allen and Unwin, London, UK.

Johnson, K. (2002), 'Towards an etiology of adjunct islands', University of Massachusetts at Amherst, unpublished manuscript, Amherst, MA, USA.

Just, M. A. & Carpenter, P. A. (1980), 'A theory of reading: From eye fixations to comprehension', *Psychological Review* **87**, 329–354.

Kadmon, N. & Roberts, C. (1986), Prosody and scope: the role of discourse structure, *in* P. T. Farley, A. M. Farley & K.-E. McCullough, eds, 'Chicago Linguistics Society (CLS) 22(2): papers from the Parasession on Pragmatics and Grammatical Theory', University of Chicago, Chicago, IL, USA, pp. 16–28.

Kagan, O. (2006), Specificity as speaker identifiability, *in* B. Gyuris, L. Kálmán, C. Piñón & K. Varasdi, eds, 'Proceedings of the ninth symposium on Logic and Language', Eötvös Loránd University, Besenyőtelek, Hungary.

Kaiser, E. (2011), 'Focusing on pronouns: Consequences of subjecthood, pronominalisation, and contrastive focus', *Language and Cognitive Processes* **26**, 1625–1666.

Kaiser, G. A. (2006), Sobre a (alegada) perda do sujeito nulo no português brasileiro, *in* T. e. a. Lobo, ed., 'Para a história do português brasileiro. Vol. VI: Novos dados, novas análises', Editora da Universidade Federal da Bahia, Salvador, BA, Brazil, pp. 11–42.

Kaiser, G. A. (2009), Losing the null subject. A contrastive study of (Brazilian) Portuguese and (Medieval) French, *in* G. A. Kaiser & E.-M. Remberger, eds, 'Proceedings of the Workshop "Null-Subjects, Expletives, and Locative in Romance"', Universität Konstanz, Konstanz, Germany, pp. 72–80.

Kamp, H. (1981), A theory of truth and semantic representation, *in* J. Groenendijk, T. Janssen & M. Stokhof, eds, 'Formal Methods in the Study of Language', Mathematical Centre Tracts, Amsterdam, The Netherlands, pp. 483–512.

Kamp, H. & Reyle, U. (1993), *From Discourse to Logic*, Kluwer, Dordrecht, The Netherlands.

Kathol, A. (1999), Agreement and the syntax-morphology interface in HPSG, *in* R. Levine & G. Green, eds, 'Studies in Current Phrase Structure Grammar', Cambridge University Press, Dordrecht, The Netherlands, pp. 223–274.

Kato, M. A. (1999), 'Strong and weak pronominals in the null subject Parameter', *Probus* **11**(1), 1–37.

Kato, M. A. (2000), The partial pro-drop nature and the restricted VS order in Brazilian Portuguese, *in* M. . E. N. Kato, ed., 'Brazilian Portuguese and the Null Subject Parameter', Vervuert-Iberoamericana, Frankfurt, Germany, pp. 223–258.

Kato, M. A. (2013), 'Polar positive answers in Brazilian Portuguese', Ms. Universidade Estadual de Campinas.

Kato, M. A. & Duarte, M. E. L. (2014), 'Restrições na distribuição de sujeitos nulos no Português Brasileiro', *Veredas* **18**(1), 1–24.

Kato, M. A., Duarte, M. E. L., Cyrino, S. M. L. & Berlinck, R. d. A. (2006), Português brasileiro no fim do século XIX e na virada do milênio, *in* J. M. e. R. V. M. e. S. Suzana Cardoso, ed., 'Quinhentos anos de história lingüística no Brasil', Empresa Gráfica da Bahia/Funcultura/Governo da Bahia, Salvador, Bahia, Brazil, pp. 413–438.

Kato, M. A. & Negrão, E. V. (2000), *Brazilian Portuguese and the Null Subject Parameter*, Vervuert-Iberoamericana, Frankfurt, Germany.

Kato, M. & Tarallo, F. (1993), Sim: respondendo afirmativamente em português, *in* M. S. Z. Paschoal & M. A. A. Celani, eds, 'Lingüística aplicada: da aplicação à linguística transdisciplinar', EDUC, São Paulo, Brazil, pp. 457– 489.

Kayne, R. S. (1994), 'The Antisymmetry of Syntax', *Linguistic Inquiry* **25**, ??

Keenan, E. & Comrie, B. (1977), 'Noun Phrase Accessibility and Universal Grammar', *Linguistic Inquiry* **8**, 63–99.

Keffala, B. (2013), Resumption and gaps in English relative clauses: Relative acceptability creates an illusion of 'saving', *in* G. F. S. K. C. S. S. Chundra Cathcart, I-Hsuan Chen & E. Stickles, eds, 'Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society', Colibri, Berkeley, CA, USA, pp. 140–154.

Kehler, A. & Rohde, H. (2013*a*), 'A Probabilistic Reconciliation of Coherence-Driven and Centering-Driven Theories of Pronoun Interpretation', *Theoretical Linguistics* **39**, 1–37.

Kehler, A. & Rohde, H. (2013*b*), 'Aspects of a theory of pronoun interpretation', *Theoretical Linguistics* **39**(3-4), 295–309.

Keller, F. (2000), Gradience in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality, PhD thesis, University of Edinburgh, Edinburgh, Scotland.

Konieczny, L., Hemforth, B., Scheepers, C. & Strube, G. (1994), 'The role of lexical heads in parsing: Evidence from German', *Language and Cognitive Processes* **12**, 307–348.

Kratzer, A. (1991), The Representation of Focus, *in* A. von Stechow & D. Wunderlich, eds, 'Semantik: an International Handbook of Contemporary Research', Walter de Gruyter, Berlin, Germany, pp. 825–834.

Kratzer, A. (1995), Stage-level and Individual-level Predicates, *in* G. N. Carlson & F. J. Pelletier, eds, 'The Generic Book', University of Chicago Press, Chicago, USA, pp. 125–175.

Krifka, M. (1998*a*), 'Scope-inversion under the rise-fall contour in German', *Linguistic Inquiry* **29**, 75–112.

Krifka, M. (1998*b*), The origins of telicity, *in* S. Rothstein, ed., 'Lingüística aplicada: da aplicação à linguística transdisciplinar', Kluwer, Dordrecht, The Netherlands, pp. 197–235.

Krifka, M. (2007), Basic notions of information structure, *in* C. Fery & M. Krifka, eds, 'Interdisciplinary Studies on Information Structure 6', Potsdam Universitätsverlag, Potsdam, Germany, pp. 13–56.

Krifka, M. (2008), Information structure: Notional distinctions, ways of expression, *in* P. van Sterkenburg, ed., 'Unity and diversity of languages', John Benjamins, Amsterdam, The Netherlands, pp. 123–136.

Krifka, M. (2016), Quantification and information structure, *in* C. Féry & S. Ishihara, eds, 'The Oxford Handbook of Information Structure', Oxford University Press, Oxford, UK, pp. 251–269.

Krifka, M., Pelletier, F., Carlson, G. N., Chierchia, G., Link, G. & ter Meulen, A. (1995), *Introduction to Genericity*, University of Chicago Press., Chicago, USA.

Kroch, A. (1989), 'Reflexes of Grammar in Patterns of Language Change', *Language Variation and Change* **1**, 199–244.

Krzek, M. (2013*a*), Generic subjects and Voice in Polish impersonal constructions, *in* C. S. Rhys, P. Iosad & A. Henry, eds, 'Microvariation, Minority Languages, Minimalism and Meaning: Proceedings of the Irish Network in Formal Linguistics', Cambridge Scholars Publishing, Cambridge, UK, pp. 186–206.

Krzek, M. (2013*b*), Interpretation and voice in Polish SIĘ and -NO/-TO constructions, *in* I. K. Chahine, ed., 'Contemporary Studies in Slavic Linguistics', John Benjamins, Amsterdam, The Netherlands, pp. 185–198.

Kuno, S. (1972), 'Functional sentence perspective', *Linguistic Inquiry* **3**, 269–320.

Kuno, S. (1973), *The structure of the Japanese language*, MIT Press, Cambridge, Mass, USA.

Kuno, S. (1978), Generative discourse analysis in America, *in* W. Dressler, ed., 'Current Trends in Text Linguistics', Mouton de Gruyter, Berlin, Germany and New York, USA, pp. 275–295.

Kuno, S. (1991), Remarks on quantifier scope, *in* H. Nakajima, ed., 'Current English linguistics in Japan', Mouton de Gruyter, Berlin, GE and New York, USA, pp. 261–287.

Kuno, S., Takami, K. & Wu., Y. (1999), 'Quantifier scope in English, Chinese and Japanese', *Language* **75**, 63–111.

Kuroda, S.-Y. (1979), 'The semantics of Japanese topic marker wa', *Linguisticae Investigationes* **3**, 75–86.

Kurtzman, H. S. & MacDonald, M. C. (1993), 'Resolution of quantifier scope ambiguities', *Cognition* **48**, 75–86.

Lambrecht, K. (1994), *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents*, Cambridge Studies in Linguistics 71, Cambridge University Press, Cambridge, MA, USA.

Landau, I. (2000), *Elements of Control. Structure and Meaning in Infinitival Constructions*, University of Oklahoma Press, Norman, OK, USA.

Laperuta, M. (2002), A realização do sujeito pronominal: um estudo sociolingüístico paramétrico para a cidade de Londrina - Norte do Paraná, PhD thesis, Universidade Estadual Paulista, Araraquara, SP, Brazil.

Laperuta, M. (2004), 'Sujeito nulo na aquição: um parâmetro em mudança - sujeito preenchido na aprendizagem: a eterna tentativa de mudança', *Revista do Gelne* **6**(141-152), 535–561.

Legendre, G., Culbertson, J., Barriere, I., Nazzi, T. & Goyet, L. (2010), Experimental and empirical evidence for the status and acquisition of subject clitics and agreement marking in adult and child spoken French, *in* V. Torrens, L. Escobar, A. Gavarro & J. G. Mangado, eds, 'Movement and Clitics', Academic Press, Berlin, Heidelberg, Germany, pp. 333–360.

Levin, B. & Rappaport-Hovav, M. (2005), *Argument Realization*, Cambridge University Press, Cambridge, UK.

Levinson, S. C. (1987), 'Pragmatics and the grammar of anaphora: A partial pragmatic reduction of binding and control phenomena', *Journal of Linguistics* **23**(2), 379–434.

Levinson, S. C. (1991), 'Pragmatic reduction of the Binding Conditions revisited', *Journal of Linguistics* **27**, 107–161.

Lewis, D. (1970), 'General Semantics', *Synthese* **22**, 18–67.

Lewis, D. (1979), Scorekeeping in a language game, *in* R. Bäurle, U. Egli & A. von Stechow, eds, 'Semantics from Different Point of View', Springer, Berlin, Heidelberg, Germany, New York, USA, pp. 172–187.

Li, C. N. & Thompson, S. (1976), Subject and topic: a new typology of language, *in* C. N. Li, ed., 'Subject and Topic', Academic Press, Austin, USA, pp. 457– 489.

Link, G. (1983), The logical analysis of plurals and mass terms, *in* R. Bäuerle, C. Schwarze & A. von Stechow, eds, 'Meaning, use, and interpretation of language', Academic Press, Berlin, Germany - New York, USA, pp. 302–323.

Linthe, A. (2002), Exploring the function and distribution of generic pronouns. The example of German man and du, PhD thesis, University of Sheffield, Sheffield, UK.

Lira, S. d. A. (1982), Nominal, Pronominal and Zero Subject in Brazilian Portuguese, PhD thesis, University of Pennsylvania.

Lison, P. & Tiedemann, J. (2016), OpenSubtitles2016: Extracting Large Parallel Corpora from Movie and TV Subtitles, *in* 'Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)'.

Liversedge, S. E., Paterson, K. B. & Pickering, M. J. (1994), Eye movements and measures of reading time, *in* G. Underwood, ed., 'Eye movements in reading and scene perception', Elsevier, Oxford, UK, pp. 55–76.

Longobardi, G. (2003), 'Reference and Proper Names: a Theory of N-Movement in Syntax and Logical Form', *Linguistic Inquiry* **25**(4), 609–665.

Lopes, C. S. L. (1999), A inserção de a gente no quadro pronominal do português, PhD thesis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil.

Lopes, R. E. V. (2003), 'The Production of Subject in Brazilian Portuguese by a Young Child', *Probus* **15**, 123–146.

Luegi, P. (2012), Processamento de sujeitos pronominais em Português: efeito da posição estrutural dos antecedentes, PhD thesis, Universidade de Lisboa, retrieved from http://repositorio.ul.pt/bitstream/10451/7531/1/ulsd063349$_t$d$_P$aula$_R$ibeiro.pdf.

Luegi, P., Costa, A. & Maia, M. (2014), 'Processamento e interpretação de sujeitos nulos e plenos em português Europeu e português do Brasil', *Cadernos de Letras da UFF* **49**, 67–88.

Luján, M. (1985), 'Binding properties of overt pronouns in null pronominal languages', *Chicago Linguistic Society* **21**, 424–438.

Luján, M. (1986), 'Stress and binding of pronouns', *Papers from the Parasession on Pragmatics and Grammatical Theory. Chicago Linguistic Society* **22**(2), 248–262.

Luján, M. (1999), Expresión y Omisión del Pronombre Personal, *in* I. Bosque & V. Demonte, eds, 'Gramática Descriptiva de la Lengua Española. Real Academia Española: Colección Nebrija and Bello, Fundación Ortega y Gasset', Vol. 2, Espasa-Calpe, Madrid, Spain, pp. 1276–1315.

Lust, B. (1986), *Studies in the Acquisition of Anaphora, vol. I, Defining the Constraints*, Reidel, Dordrecht, The Netherlands.

Lyons, C. (1999), *Definiteness*, Cambridge University Press, Cambridge, MA, USA.

Lyons, J. (1977), *Semantics. 2 vols*, Cambridge University Press, Cambridge, MA, USA.

Madeira, A., Xavier, M. & Crispim, M. (2010), Interpretação semântica e/ou pragmático-discursiva de sujeitos na aquisição de português L2, *in* 'Textos seleccionados do XXV Encontro Nacional da Associação Portuguesa de Linguística', Colibri, Lisbon, Portugal, pp. 513–529.

Magalhães, T. M. V. (2000), Aprendendo o Sujeito Nulo na Escola, PhD thesis, UNICAMP, Campinas, Brazil.

Magalhães, T. M. V. (2003), 'O Traço de Pessoa no PE e no PB: um estudo a partir de dados de aquisição', ms. Comunicação apresentada no 6º ENAL, PUCRS, Porto Alegre, outubro de 2003.

Magalhães, T. M. V. (2007), 'A aquisição de pronomes sujeitos no PB e no PE', *Letras de Hoje* **42**(1), 97–112.

Magalhães, T. M. V. & Santos, A. L. (2006), 'As respostas verbais e a freqüência de sujeito nulo na aquisição do Português Brasileiro e Português Europeu', *Letras de Hoje* **41**(1), 179–193.

Maia, M. (2014), 'Efeito da lacuna preenchida e plausibilidade semântica no processamento de frases em português brasileiro', *Cadernos de Letras da UFF* **49**, 23–46.

Maia, M., Moura, A. & de Souza, M. (2016), 'Ilhas sintáticas e plausibilidade semântica - um estudo de rastreamento ocular de frases com lacunas preenchidas em português brasileiro', *SCRIPTA* **20-38**, 293–311.

Malamud, S. (2012), 'Impersonal indexicals: one, you, man and du', *Journal of Comparative Germanic Linguistics* **15**, 1–48.

Martins, A. M. (1994*a*), Clíticos na História do Português, PhD thesis, Universidade de Lisboa, Lisboa, Porto Alegre.

Martins, A. M. (1994*b*), 'Enclisis, VP-Ellipsis and the Nature of Sigma', *Probus* **6**, 173–205.

Martins, A. M. (1996), A colocação dos clíticos e a relevância da categoria Sigma, *in* I. Duarte & I. Leiria, eds, 'Congresso Internacional sobre o Português. Actas', Vol. 2, Associação Portuguesa de Linguística and Colibri, Lisboa, Portugal, pp. 95–134.

Martins, A. M. (2004), Emphatic Affirmation and Polarity: Contrasting European Portuguese With Brazilian Portuguese, Spanish, Catalan and Galician, *in* J. Doetjes & P. Gonzalez, eds, 'Romance Languages and Linguistic Theory 2004: Selected Papers from Going Romance', John Benjamins, Amsterdam, The Netherlands, pp. 197–223.

Martins, A. M. (2012), 'The Portuguese answering system: affirmation, negation and denial', ms. available at http://alfclul.clul.ul.pt/wochwel/documents/Martins

Marušič, F., Marvin, T. & Žaucer, R. (2003), Secondary Predication in Slovenian, *in* W. Browne, ed., 'FASL 11 (The Amherst Meeting, 2002)', Ann Arbor: Michigan Slavic Publications, Michigan, USA, pp. 49–76.

Masur, E. F. (1995), "Infants' early verbal imitation and their later lexical development', *Merrill-Palmer Quarterly* **41**(3), 286–306.

Matos, G. (1992), Construções de Elipse do Predicado em Português: SV Nulo e Despojamento, school=Universidade de Lisboa, address=Lisboa, Porto Alegre, PhD thesis.

May, R. (1977), The Grammar of Quantification, PhD thesis, Massachusetts Institute of Technology, facsimile edition of 1977 MIT Ph.D. Dissertation with 1991 preface published by Garland, New York, NJ, USA.

May, R. (1985), *Logical Form: Its Structure and Derivation*, MIT Press, Cambridge, MA, USA.

Mayol, L. (2010), Redefining salience and the Position of Antecedent Hypothesis: a study of Catalan pronouns, *in* 'University of Pennsylvania Working Papers in Linguistics', University of Pennsylvania, Pennsylvania, USA.

McDonald, M., Pearlmutter, N. & Seidenberg, M. (1994), 'The lexical nature of syntactic ambiguity resolution', *Psychological Review* **101**, 676–703.

McEnery, T. (2000), *Corpus-based and computational approaches to discourse anaphora*, John Benjamins, Cambridge, MA, USA.

McRae, K., Spivey-Knowlton, M. J. & Tanenhaus, M. K. (1998), 'Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension', *Journal of Memory and Language* **38**, 283–312.

Melo, M. & Maia, M. (2005), 'O processamento da co-referência do sujeito pronominal em sentencas formadas por verbos de comunicação linguística no português do Brasil', *Linguística* **1**, 177–206.

Menuzzi, S. d. M. (2000), First Person Plural Anaphora in Brazilian Portuguese: chains and constraint interaction in binding, *in* J. Costa, ed., 'Portuguese Syntax: New Comparative Studies', Oxford University Press, Oxford, UK, pp. 191–240.

Menuzzi, S. d. M. (2002), 'Sobre as Opções Anafóricas dos Antecedentes Genéricos', ms. available at https://sites.google.com/site/smenuzzi/artigos-hand-outs-etc.

Menuzzi, S. d. M. & Lobo, M. (2016), Binding and Pronominal Forms in Portuguese, *in* W. L. Wetzels, S. Menuzzi & J. Costa, eds, 'The Handbook of Portuguese Linguistics', Wiley Blackwell, Oxford, UK, pp. 338–355.

Merchant, J. (2001), *The syntax of silence: Sluicing, islands, and the theory of ellipsis*, Oxford University Press, Oxford, UK.

Miles, W. R. (1930), 'Ocular dominance in human adults', *Journal of General Psychology* **3**, 412–429.

Miltsakaki, E. (2002), 'Toward an aposynthesis of topic continuity and intrasentential anaphora', *Computational Linguistics* **28**(3), 319–355.

Mitchell, D. C. (1987), 'Lexical guidance in human parsing - locus and processing characteristics', *Attention and Performance* **12**, 601–618.

Mitkov, R. (1999), 'Anaphora Resolution: The State of the Art', *Paper based on the COLING'98/ACL'98 tutorial on anaphora resolution; University of Wolverhampton* .

Mitkov, R. (2002), *Anaphora Resolution*, Pearson Education, Upper Saddle River, NJ, USA.

Modesto, M. (2000*a*), Null Subjects without "Rich" Agreement, *in* M. A. Kato & E. V. Negrão, eds, 'Brazilian Portuguese and the Null Subject Parameter', Vervuert-Iberoamericana, Frankfurt, Germany, pp. 147–174.

Modesto, M. (2000*b*), On the Identification of Null Arguments, PhD thesis, University of South California, Los Angeles, CA, USA.

Modesto, M. (2008*a*), Null Subjects in Brazilian Portuguese and Finnish: They are not derived by movement, *in* W. D. Davies & S. Dubinsky, eds, 'New Horizons in The Analysis of Control and Raising, Studies in Natural Language and Linguistic Theory', Vol. 71, Springer, Amsterdam, The Netherlands, pp. 231–248.

Modesto, M. (2008*b*), Topic Prominence and Null Subjects, *in* T. Biberauer, ed., 'The Limits of Syntactic Variation, Linguistics Today', Vol. 132, John Benjamins, Amsterdam, The Netherlands, pp. 375–409.

Moeschler, J. (2003), 'Answers to questions about questions and answers', *Journal of Pragmatics* **10**, 227–253.

Moltmann, F. (2006), 'Generic one, arbitrary PRO, and the first person', *Natural Language Semantics* **14**, 257–281.

Moltmann, F. (2010), 'Relative Truth and the First Person', *Philosophical Studies* **150**(2), 187–220.

Montalbetti, M. (1984), After Binding Properties. On the Interpretation of Pronouns, PhD thesis, MIT, Massachusetts, MA, USA.

Morales, A. (1995), The Loss of the Spanish Impersonal Particle se Among Bilinguals: A Descriptive Profile, *in* C. Silva-Corvalan, ed., 'Spanish in Four Continents. Studies in Language Contact and Bilingualism', Georgetown University Press, Washington, D.C., USA, pp. 72–80.

Moreira da Silva, S. (1984), Etudes sur la symétrie et l'asymétrie Sujet/Objet dans le Portugais du Brésil, PhD thesis, Université de Paris VIII, Paris, France.

Morgado, S. (2012), Processamento da co-referência pronominal - informação sintáctica e semântica, PhD thesis, Universidade de Lisboa, Lisbon, Portugal.

Morgan, J. L. (1975), Some interactions of syntax and pragmatics, *in* P. Gole & J. L. Morgan, eds, 'Syntax and Semantics: Speech Acts, vol. 3', Academic Press, New York, NJ, USA, pp. 221–238.

Muller, G. (2005), 'Pro-drop and impoverishment', ms. University of Leipzig.

Munro, P. & Willmond, C. (1994), *Chickasaw: an Analytical Dictionary*, University of Oklahoma Press, Norman, OK, USA.

Myers, J. (n.d.), 'Acceptability judgments'.
URL: *http://linguistics.oxfordre.com/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-333*

Müller, A. L. d. P. (2001), Genericity and the Denotation of Common Nouns in Brazilian Portuguese, *in* 'UMOP 25: Proceedings of SULA (Semantics of Under-Represented Languages in the Americas)', GLSA, University of Massachusetts, Amherst, USA, pp. 72–80.

Müller, A. L. d. P. (2002), 'The semantics of generic quantification in Brazilian Portuguese', *Probus* **14:2**, 279–298.

Müller, A. L. d. P. (2003), 'Anaphora and Genericity in Brazilian Portuguese', *Letras de Hoje* **38**, 109–124.

Negrão, E. V. (1990), 'A distribuição e a interpretação de pronomes na fala de crianças da escola pública', (Ms.) University of São Paulo, São Paulo.

Negrão, E. V. (1997), Asymmetries in the distribution of overt and empty categories in Brazilian Portuguese, *in* J. B. . V. Motapanyase, ed., 'Clitics Pronouns and Movement,Current Issues in Linguistic Theory', Vol. 14, John Benjamins, Amsterdam, The Netherlands, Philadelphia, USA, pp. 217–235.

Negrão, E. V. & Müller, A. L. d. P. (1996), 'As Mudanças no Sistema Pronominal do Português do Brasil: Substituição ou Especialização de Formas?', *DELTA* **12**, 125–152.

Negrão, E. V. & Viotti, E. d. C. (2000), Brazilian Portuguese as a Discourse-Oriented Language, *in* M. A. Kato & E. V. Negrão, eds, 'The Null Subject Parameter in Brazilian Portuguese', Vervuert-Iberoamericana, Frankfurt, Germany, pp. 105–126.

Nevins, A. (2007), 'The representation of third person and its consequences for person-case effects', *Natural Language and Linguistic Theory* **25**(2), 273–313.

Newmeyer, F. J. (2004), 'Against a parameter-setting approach to language variation', *Linguistic Variation Yearbook* **4**, 181–234.

Newmeyer, F. J. (2005), *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*, Oxford University Press, Oxford, UK.

Newmeyer, F. J. (2006), 'A rejoinder to 'On the role of parameters in Universal Grammar: a reply to Newmeyer' by Ian Roberts and Anders Holmberg', ms., revised version, February 2006.

Newmeyer, F. J. (to appear), Where, if Anywhere, are Parameters? A Critical Historical Overview of Parametric Theory, *in* C. Bowern, L. Horn & R. Zanuttini, eds, 'On looking into words (and beyond)', Language Science Press, Berlin, Germany, pp. 547–569.

Nicenboim, B. & Vasishth, S. (2016), 'Statistical methods for linguistic research: Foundational ideas - part ii', pp. 479–508.

Nunberg, G. & Pan, C. (1975), Inferring quantification in generic sentences, *in* 'Papers from the Eleventh Regional Meeting of the Chicago Linguistic Society', Chicago University Press.

Nunes, J. M. (1990), O Famigerado SE: uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador, Master's thesis, UNICAMP, Campinas, SP, Brazil.

Otheguy, R., Zentella, A. C. & Livert, D. (2007), 'Language contact in Spanish in New York toward the formation of a speech community', *Language* **83**(4), 770–802.

Paul, W. (2016), Null subject, null topics and topic prominence in Mandarin Chinese and beyond, *in* S. Fuzhen, ed., 'Studies on Syntactic Cartography Volume 1: Proceedings of the international workshop on syntactic cartography IWSC2015, Beijing Language and Culture University, December 6-7, 2015', Chinese Social Sciences Academy Press, Beijing, China, pp. 391–420.

Perlmutter, D. (1971), *Deep and Surface Structure Constraints in Syntax*, Holt, Reinhart and Winston, New York, NJ, USA.

Phillips, C., Kazanina, N. & Abada, S. H. (2005), 'ERP effects of the processing of syntactic long-distance dependencies', *Cognitive Brain Research* **22**(3), 407–428.

Phimsawat, O. (2011), The syntax of pro-drop in Thai, PhD thesis, Newcastle University, Newcastle, UK.

Picallo, C. M. (1984), 'The Infl Node and the Null Subject Parameter', *Linguistic Inquiry* **15**, 75–102.

Pickering, M. J., Frisson, S., McElree, B. & Traxler, M. J. (2004), Eye movements and semantic composition, *in* 'The On-line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond', Psychology Press, pp. 33–50.

Polinsky, M., Gómez Gallo, C., Graff, P., Kravtchenko, E., Milton, A. & Sturgeon, A. (1995), Subject islands are different, *in* J. Sprouse & N. Hornstein, eds, 'Experimental Syntax and islands effects', Cambridge University Press, Cambridge, MA, USA, pp. 29–60.

Pollard, C. J. & Sag, I. A. (1994), *Head-Driven Phrase Structure Grammar*, University of Chicago Press, Chicago, USA.

Pontes, E. (1987), *O Tópico no Português do Brasil*, Pontes, Campinas, SP, Brazil.

Potts, C. (2015), Presupposition and implicature, *in* S. Lappin & C. Fox, eds, 'The Handbook of Contemporary Semantic Theory', Wiley-Blackwell, Oxford, UK, pp. 168–202.

Prince, E. F. (1981), Toward a taxonomy of given-new information, *in* P. Cole, ed., 'Radical Pragmatics', The Academic Press, New York: USA, pp. 223–256.

Prince, E. F. (1986), On the syntactic marking of presupposed open propositions, *in* A. M. Farley, P. Farley & K.-E. McCullough, eds, 'Proceedings of CLS 22: Papers from the parasession on pragmatics and grammatical theory', IL, Chicago: USA, pp. 208–222.

Raffray, C. N. (2010), 'How do people construct logical form during language comprehension?', *Psychol Sci* **21**, 1090–1097.

Ramírez, D. B. (2007), Selected Proceedings of the Third Workshop on Spanish Sociolinguistics, *in* J. Holmquist, ed., 'Proceedings of CLS 22: Papers from the parasession on pragmatics and grammatical theory', Cascadilla Proceedings Project, Somerville, MA: USA, pp. 122–128.

Raney, G. E., J., C. S. & C., B. J. (2014), 'Movements to Evaluate the Cognitive Processes Involved in Text Comprehension', *Journal of Visualized Experiments* **83**.

Rayner, K. (1998), 'Eye Movements in Reading and Information Processing: 20 Years of Research', *Psychological Bulletin* **124**, 372–422.

Rayner, K., Binder, K., Ashby, J. & Pollatsek, A. (2001), 'Eye movement control in reading: word predictability has little influence on initial landing positions in words', *Vision Res* **41**(7), 943–954.

Rayner, K. & Pollatsek, A. (1981), 'Eye movement control during reading: evidence for direct control', *Quarterly Journal of Experimental Psychology* **33A**, 351–373.

Rayner, K. & Sereno, S. C. (1994*a*), Eye movements in reading: Psycholinguistic studies, *in* M. Gernsbacher, ed., 'Handbook of psycholinguistics', Academic Press, New York, NJ, USA, pp. 57–82.

Rayner, K. & Sereno, S. C. (1994*b*), 'Regression-contingent analyses: A reply to Altmann', *Memory and Cognition* **22**, 291–292.

Rayner, K. & Sereno, S. C. (1994*c*), 'Regressive eye movements and sentence parsing: On the use of regression-contingent analyses', *Memory and Cognition* **22**, 281–285.

Reinhart, T. (1983), *Anaphora and semantic interpretation*, The University of Chicago Press, Chicago: USA.

Reinhart, T. (l981), 'Pragmatics and linguistics: an analysis of sentence topics', *Philosophica* **27**(1), 53–94.

Reuland, E. (2001), 'Primitives of Binding', *Linguistic Inquiry* **32**(2), 439–492.

Reuland, E. (2003), 'Anaphoric dependencies: A window into the architecture of the language system 7', pp. 3–25.

Ritter, E. (l995), 'On the syntactic category of pronouns and agreement', *Natural Language and Linguistic Theory* **13**(3), 405–443.

Rizzi, L. (1980), 'Wh-movement, negation and the pro-drop parameter', (Ms.) Scuola Normale Superiore, Pisa, Italy.

Rizzi, L. (1982*a*), *Issues in Italian Syntax*, Foris, Dordrecht: The Netherlands.

Rizzi, L. (1982*b*), Violations of the wh-island constraint and the subjacency condition, *in* L. Rizzi, ed., 'Issues in Italian syntax', Foris, Dordrecht, The Netherlands, pp. 49–76.

Rizzi, L. (1986), 'Null Objects in Italian and the Theory of pro', *Linguistic Inquiry* **17**, 501–557.

Rizzi, L. (1997), The fine structure of the left periphery, *in* L. Haegeman, ed., 'Elements of Grammar. Handbook of Generative Syntax', Foris, Dordrecht: Kluwer, pp. 281–337.

Rizzi, L. (2002), 'On the Grammatical Basis of Language Development: A Case Study', Ms., Siena University, Siena, Italy.

Roberts, C. (1996), Information structure in discourse: Towards an integrated formal theory of pragmatics, *in* J. Yoon & A. Kathol, eds, 'OSUWPL Volume 49: Papers in Semantics', The Ohio State University Department of Linguistics, Ohio, USA, pp. 1–69.

Roberts, C. (2012), 'Information structure in discourse: Towards an integrated formal theory of pragmatics', *Semantics and Pragmatics* **5**, 1–69.

Roberts, I. (1993), *Verbs and Diachronic Syntax*, Kluwer, Dordrecht, The Netherlands.

Roberts, I. (2010), A Deletion Analysis of Null Subjects, *in* T. Biberauer, A. Holmberg, I. Roberts & M. Sheehan, eds, 'Parametric Variation: Null Subjects in Minimalist Theory', Cambridge University Press, Cambridge, MA, USA, pp. 58–88.

Roberts, I. (2014), Taraldsen's Generalization and Language Change: Two Ways to Lose Null Subjects, *in* P. Svenonius, ed., 'Functional Structure from Top to Toe: The Cartography of Syntactic Structures', Vol. 9, Oxford UniversityPress, Oxford: UK, p. ???

Roberts, I. (2016), 'The Null Subject Parameter in the 21st century', Ms. Downing College, University of Cambridge, Cambridge, UK.

Roberts, I. & Holmberg, A. (2010), Introduction, *in* T. Biberauer, A. Holmberg, I. Roberts & M. Sheehan, eds, 'Parametric Variation: Null Subjects in Minimalist Theory', Cambridge University Press, Cambridge: MA, USA, pp. 1–56.

Roberts, I. & Kato, M. A. (1993), *Português Brasileiro: uma viagem diacrônica*, Editora da Unicamp, Campinas, SP, Brazil.

Rochemont, M. S. (1986), *Focus in generative grammar*, John Benjamins, Amsterdam: The Netherlands.

Rodrigues, C. (2004), Impoverished Morphology and A-Movement out of Case Domains, PhD thesis, University of Maryland at College Park, Cambridge, MA, USA.

Rohrbacher, B. (1994), The Germanic VO Languages and the Full Paradigm: A Theory of V to I Raising, PhD thesis, University of Massachusetts, Cambridge, MA, USA.

Roland, D., Mauner, G., O'Meara, C. & Yun, H. (2012), 'Discourse expectations and relative clause processing', **66**, 479–508.

Rooth, M. (1985), Association with Focus, PhD thesis, University of Massachusetts at Amherst, Amherst, MA, USA.

Rooth, M. (1992), 'A theory of focus interpretation', *Natural Language Semantics* **1**, 75–116.

Rudolf, O. (2014), A Corpus-Based Study of Human Impersonal Constructions in Russian, *in* J. Witkos & S. Jaworski, eds, 'New Insights into Slavic Linguistics – Series Title Sprachund Kulturkontakte in Europas Mitte', Peter Lang AG, Bern: Switzerland, pp. 329–343.

Runner, J. T. & Ibarra, A. (2016), Information structure effects on null and overt subject comprehension in Spanish, *in* A. Holler & K. Suckow, eds, 'Empirical perspectives on anaphora resolution: Information structural evidence in the race for salience', de Gruyter, Berlin, Germany, pp. 1–69.

Russell, B. (1905), 'On Denoting', *Mind* **14**.

Rusten, K. (2010), A study of empty referential pronominal subjects in Old English, Master's thesis, University of Bergen, Bergen, Germany.

Sabel, J. (2002), 'A Minimalist Analysis of Syntactic Islands', *The Linguistic Review* **19**, 271–308.

Sadock, J. & Zwicky, A. M. (1985), Speech act distinctions in syntax, *in* T. Shopen, ed., 'Language Typology and Syntactic Description, Volume I, Clause structure', Cambridge University Press, Cambridge, UK, pp. 155–196.

Safir, K. (2004), *The syntax of anaphora*, Oxford University Press, Oxford, UK.

Sandord, A. J. & Garrod, S. C. (1981), *Understanding Written Language*, John Wiley and Sons, Chichester, USA.

Saussure, F. d. (1983), *Course in General Linguistics*, Open Court, La Salle, IL, USA.

Schwenter, S. A. (2005), 'The pragmatics of negation in Brazilian Portuguese', *Lingua* **115**, 1427–1456.

Schütze, C. T. (2016), *The empirical base of linguistics: Grammaticality judgments and linguistic methodology. Classics in Linguistics 2*, Language Science Press, Berlin, Germany.

Schütze, C. T. & Sprouse, J. (2013), Judgment data, *in* R. J. Podesva & D. Sharma, eds, 'Research methods in linguistics', Cambridge University Press, New York, NJ, USA, pp. 179–208.

Selkirk, E. (1984), *Phonology and Syntax: The Relation between Sound and Structure*, MIT Press, Cambridge, MA, USA.

Sells, P. (1987), 'Aspects of logophoricity', *Linguistic Inquiry* **18**, 445–480.

Sennet, A. (2016), 'Ambiguity', The Stanford Encyclopedia of Philosophy (Spring 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2016/entries/ambiguity/>.

Sgall, P., Hajičová, E. & Panevová, J. (1986), *The meaning of the sentence in its semantic and pragmatic aspects*, D. Reidel, Dordrecht, Germany and Boston, MA, USA.

Sheehan, M. (2006), The EPP and null subjects in Romance, PhD thesis, Newcastle University, Newcastle, UK.

Sheehan, M. (to appear), Subjects, null-subjects and expletives in Romance, *in* 'Manuals of Romance Linguistics (MRL): Grammatical Interfaces', http://ling.auf.net/lingbuzz/002466.

Sheldon, A. (1974), 'The role of parallel function in the acquisition of relative clauses in English', *Journal of Verbal Learning and Verbal Behavior* **13**, 278–281.

Shlonsky, U. (2009), 'Hebrew as a partial null-subject language', *Studia Linguistica* **63**(1), 133–157.

Sidnell, J. (2010), *Conversation analysis: An introduction*, Wiley and Blackwell, Chichester, UK.

Sigurdsson, H. A. (2004), 'The syntax of person, tense and speech features', *Italian Journal of Linguistics* **16**, 219–251.

Sigurdsson, H. A. (2012), 'About Pronouns', http://ling.auf.net/lingbuzz/001593.

Sigurdsson, H. A. & Egerland, V. (2009), 'Impersonal null-subjects in Icelandic and elsewhere', *Studia Linguistica* **63**, 158–185.

Sigurdsson, H. & Holmberg, A. (2008), Icelandic dative intervention, *in* R. D'Alessandro, S. Fischer & G. H. Hrafnbjargarson, eds, 'Agreement Restrictions', Mouton de Gruyter, Berlin, Germany, pp. 251–279.

Silverstein, M. (1976), Hierarchy of features and ergativity, *in* R. M. W. Dixon, ed., 'Grammatical categories in Australian languages', UMOP, Australian Institute of Aboriginal Studies, Canberra, AUS, pp. 112–171.

Simonenko, A., Crabbé, B. & Prévost, S. (2015), Morphological triggers of syntactic changes: Treebank-based Information Theoretic approach, *in* M. Dickinson, E. Hinrichs, A. Patejuk & A. Przepiórkowski, eds, 'Proceedings of the Fourteenth International Treebanks and Linguistic Theories (TLT14)', Warsaw, Poland, retrieved from https://hal.inria.fr/hal-01240245/document.

Simonenko, A., Crabbé, B. & Prévost, S. (2017), Taraldsen's Generalization in Diachrony: Evidence from a Diachronic Corpus, *in* A. e. a. Kaplan, ed., 'Proceedings of the 34th West Coast Conference on Formal Linguistics', CSLI Publications, Somerville, MA, USA, pp. 462–469.

Simons, M., Tonhauser, J., Beaver, D. & Roberts, C. (2010), What projects and why, *in* 'SALT', Vol. 20, CLC Publications, Ithaca, NY, USA, pp. 309–327.

Simões, L. J. (1997), Sujeito Nulo na Aquisição do Português do Brasileiro: um Estudo de Caso, PhD thesis, PUCRS, Porto Alegre, RS, Brazil.

Smyth, R. (1974), 'Grammatical determinants of ambiguous pronoun resolution', *Journal of Psycholinguistic Research* **23**(3), 197–229.

Sorjonen, M.-L. (2001), *Responding in Conversation: A study of response particles in Finnish*, John Benjamins, Amsterdam, The Netherlands.

Speas, M. (1994), Null Arguments in a Theory of Economy of Projection, *in* E. Benedicto & J. Runner, eds, 'Functional Projections', UMOP, Campinas, Brazil, pp. 179–208.

Sperber, D. & Wilson, D. (1986), *Relevance: communication and cognition (The Language and Thought Series)*, Harvard University Press, Cambridge, MA, USA.

Sprouse, J., Caponigro, I., Greco, C. & Cecchetto, C. (2016), 'Aspects of logophoricity', *Experimental syntax and the variation of island effects in English and Italian* **34**(1), 307–344.

Spyropoulos, V. (2002), A note on arbitrary null-subjects, *in* 'Reading Working Papers in Linguistics', Vol. 6, pp. 85–100.

Stalnaker, R. (1978), Assertion, *in* P. Cole, ed., 'Syntax and Semantics 9: Pragmatics', Academic Press, New York, USA, pp. 315–332.

Stalnaker, R. (2002), 'Common Ground', *Linguistics and Philosophy* **16**(5-6), 701–721.

Stowe, L. A. (1986), 'Parsing WH-constructions: Evidence for on-line gap location', *Language and Cognitive Processes* **1**, 227–245.

Stubbs, M. (1983), *Discourse Analysis. The Sociolinguistic Analysis of Natural Languages*, Basil Blackwell, Oxford, UK.

Sturt, P. (2003), 'The time-course of the application of binding constraints in reference resolution', *Journal of Memory and Language* **48**, 542–562.

Szabolcsi, A. (1997), Strategies for scope taking, *in* A. Szabolcsi, ed., 'Ways of scope taking', Kluwer, Dordrecht, The Netherlands, pp. 109–154.

Sæbø, K.-J. (1997), Topic, focus and quantifier raising, *in* 'Proceedings of the Eleventh Amsterdam Colloquium', ILLC/Department of Philosophy,University of Amsterdam, Amsterdam, The Netherlands, pp. 67–72.

Takami, K. (1989), 'Prepositional stranding: arguments against syntactic analyses and an alternative functional explanation', *Lingua* **76**, 299–335.

Taraldsen, K. T. (1980), 'On the NIC, vacuous application and the that-trace filter', Indiana University Linguistics Club.

Tarallo, F. (1983), Relativization Strategies in Brazilian Portuguese, PhD thesis, University of Pennsylvania.

Taylor, M. (2009), On the pronominal status of Brazilian Portuguese a gente, *in* 'NYU working papers in linguistics', New York, USA, pp. 1–36.

Tessler, M. H. & Goodman, N. D. (to appear), 'A pragmatic theory of generic language', (manuscript) Stanford University.

Traugott, E. (1992), Syntax, *in* R. Hogg, ed., 'The Cambridge history of the English language I: Old English', Cambridge University Press, Cambridge, MA, USA, pp. 168–289.

Traxler, M. J., Morris, R. K. & Seely, R. E. (1996), 'Processing subject and object relative clauses: Evidence from eye movements', *Journal of Memory and Language* **47**, 69–90.

Traxler, M. J. & Pickering, M. J. (1996), 'Plausibility and the processing of unbounded dependencies: An eyetracking study', *Journal of Memory and Language* **35**, 454–475.

Truckenbrodt, H. (1995), Phonological Phrases: Their Relation to Syntax, Focus, and Prominence, PhD thesis, MIT, Cambridge, MA, USA.

Truswell, R. (2007), 'Extraction from adjuncts and the structure of events', *Lingua* **117**, 1355–1377.

Urbano, H., Favero, L. L., Andrade, M. L. & Aquino, Z. G. (1993), Perguntas e respostas na conversação, *in* A. T. de Castilho, ed., 'Gramática do português falado', Vol. 2, Editora da Unicamp, Campinas, Brazil, pp. 75–98.

Uriagereka, J. (1999), Multiple spell-out, *in* S. Epstein & N. Hornstein, eds, 'Working Minimalism', MIT Press, Cambridge, MA, USA, pp. 251–282.

Vainikka, A. & Levy, Y. (1999), 'Empty subjects in Finnish and Hebrew', *Natural Language and Linguistic Theory* **17**, 613–671.

Vallduví, E. (1990), The Informational Component, PhD thesis, University of Pennsylvania.

Vallduví, E. & Engdahl, E. (1996), 'The linguistic realization of information packaging', *Linguistics* **34**.

van de Koot, H. & Neeleman, A. (2012), Towards a Unified Encoding of Contrast and Scope, *in* A. Neeleman & R. Vermeulen, eds, 'The Syntax of Topic, Focus and Contrast: An Interface-Based Approach', Vol. 14, Mouton de Gruyter, Amsterdam, The Netherlands, Philadelphia, USA, pp. 217–235.

van Gelderen, E. (2013), 'Null Subjects in Old English', *Linguistic Inquiry* **44**(2), 271–285.

Van Valin, R. (1995), Toward a functionalist account of so-called 'extraction constraints', *in* B. D. et al, ed., 'Complex Structures: A Functionalist Perspective', Mouton de Gruyter, Berlin, Germany, pp. 29–60.

Van Valin, R. D. & LaPolla, R. J. (1997), *Syntax: structure, meaning and function*, Cambridge University Press, Cambridge, MA, USA.

Vance, B. S. (1997), *Syntactic Change in Medieval French. Verb-Second and Null Subjects. Studies in Natural Language and Linguistic Theory*, Kluwer, Dordrecht, The Netherlands.

Vasishth, S. & Nicenboim, B. (2016), 'Statistical methods for linguistic research: Foundational ideas - part i', pp. 479–508.

Vennemann, T. (1993), An explanation of drift, *in* C. N. Li, ed., 'Word Order and Word Order Change', Vol. 2, University of Texas Press, Austin, TX, USA, pp. 269–305.

Visser, F. T. (1963), *An historical syntax of the English language I*, Brill, Leiden, The Netherlands.

Von Eckardt, C. & Potter, M. (1985), 'Clauses and the semantic representation of words', *Memory and Cognition* **13**, 371–376.

von Fintel, K. (1994), Restrictions on Quantifier Domains, PhD thesis, University of Massachusetts, Amherst, Massachusetts, USA.

von Heusinger, K. (2002), 'Specificity and definiteness in sentence and discourse structure', *Journal of Semantics* **19**, 254–274.

von Heusinger, K. (2011), Specificity, *in* K. von Heusinger, C. Maienborn & P. Portner, eds, 'Semantics: An International Handbook of Natural Language Meaning', Vol. 2, Mouton de Gruyter, Berlin, Germany, pp. 11–34.

von Stechow, A. (1991), Current Issues in the Theory of Focus, *in* A. von Stechow & D. Wunderlich, eds, 'Semantik: an International Handbook of Contemporary Research', Walter de Gruyter, Berlin, Germany, pp. 825–834.

Walker, M. & Prince, E. F. (1996), A bilateral approach to Givenness: A hearer-status algorithm and a centering algorithm, *in* T. Freitheim & J. Gundel, eds, 'Reference and referent accessibility', John Benjamins, Amsterdan, The Netherlands, pp. 291–306.

Wasow, T. (1986), Reflections on anaphora, *in* B. Lust, ed., 'Studies in the Acquisition of Anaphora, vol. I, Defining the Constraints', Reidel, Dordrecht, The Netherlands, pp. 107–123.

Wechsler, S. (2005), What is right and wrong about little v, *in* N. Oslo, ed., 'Grammar and beyond - Essays in honour of Lars Hellan', Novus Press, Oxford, UK, pp. 179–195.

Wechsler, S. & Zlatič, L. (2001), 'A theory of agreement and its application to serbo-croatian', *Language* pp. 391–423.

Williams, T. & Scheutz, M. (2017), Reference in Robotics: a Givenness Hierarchy Theoretic Approach, *in* B. Abbott & J. Gundel, eds, 'The Oxford Handbook of Reference', Oxford University Press, Oxford, UK.

Wilson, D. & Sperber, D. (1979), Ordered entailments: analternative to presuppositional theories, *in* C. Oh & D. Dinneen, eds, 'Syntax and Semantics 11 : Presuppositions', Academic Press, New York, NJ, USA, pp. 299–323.

Wong, J. (2000), 'Repetition in Conversation: A Look at "First and Second Sayings"', *Research on Language and Social Interaction* **33**(4), 407–424.

É. Kiss, K. (2008), 'Identificational focus vs. information focus', *Language* **74**, 245–273.

Appendices

Appendix A

Experimental Materials – Off-line Experiments

A.0.1 Experiment 1

Judgment Task

(191) A Maria estava muito nervosa. Você sabe quando ela ficou mais calma?

" Maria was very nervous. Do you know when she calmed down? "

  a. Eu tranquilizei a Maria quando (eu) divulguei os resultados do exame.

    " I calmed down Maria when I released the exam results. "

  b. Eu tranquilizei a Maria quando (eu) ia divulgar os resultados do exame.

    " I calmed down Maria when I was going to release the exam results. "

(192) O Pedro estava treinando para a competição. Você sabe quando ele começou a preparação?

" Pedro was training for the competition. Do you know when he started the preparation? "

  a. Eu treinei o Pedro quando (eu) desisti do campeonato regional.

    " I trained Pedro when I gave up the regional championship. "

  b. Eu treinei o Pedro quando (eu) ia desistir do campeonato regional.

    " I trained Pedro when I was going to give up the regional championship. "

(193) A Ana foi uma das testemunhas do caso julgado recentemente. Você sabe quando ela se sentiu instigada no tribunal?

" Ana was one of the witnesses in the recent judgment. Do you know when she felt instigated in court? "

a. Eu provoquei a Ana quando (eu) corrigi o depoimento sobre o crime.

" I provoked Ana when I corrected the testimony about the crime. "

b. Eu provoquei a Ana quando (eu) tinha corrigido o depoimento sobre o crime.

" I provoked Ana when I had corrected the testimony about the crime. "

(194) O Leandro recindiu o contrato com a empresa para construção do prédio. Você sabe quando ele se ofendeu com o parecer?

" Leandro retracted the contract with the company to build the building. Do you know when he took offense at the report? "

a. Eu critiquei o Leandro quando (eu) indiquei os defeitos da construção.

" I criticized Leandro when I pointed out the defects of the building. "

b. Eu critiquei o Leandro quando (eu) tinha indicado os defeitos da construção.

" I criticized Leandro when I had indicated the defects of the building. "

(195) A Clara se sentiu injustiçada com as palavras contra ela. Você sabe quando ela ficou ofendida?

" Clara felt injustice at the words against her. Do you know when she was offended? "

a. Eu insultei a Clara quando (eu) revelei as imagens do roubo.

" I insulted Clara when I revealed the images of the robbery. "

b. Eu insultei a Clara quando (eu) ia revelar as imagens do roubo.

" I insulted Clara when I was about to reveal the images of the robbery. "

(196) O Douglas tinha desistido do projeto do automóvel. Você sabe quando ele se motivou novamente?

" Douglas had given up the project of the car. Do you know when he got motivated again? "

a. Eu aconselhei o Douglas quando (eu) desmontei o protótipo do veículo.

" I advised Douglas when I dismantled the prototype of the vehicle. "

b. Eu aconselhei o Douglas quando (eu) ia desmontar o protótipo do veículo.

" I advised Douglas when I was going to dismantle the prototype of the vehicle. "

(197) A Diana é uma noiva muito complicada. Você sabe quando ela ficou furiosa no prova do vestido?

" Diana is a very complicated bride. Do you know when she was furious in the dress try-out? "

a. Eu ignorei a Diana quando (eu) atendi o celular durante o ajuste.

" I ignored Diana when I answered the phone during the adjustment. "

b. Eu ignorei a Diana quando (eu) estava atendendo o celular durante o ajuste.

" I ignored Diana when I was answering the cell phone during the adjustment. "

(198) O Lucas é o melhor médico especialista em contusões. Você sabe quando ele recomeçou a atender clientes?

" Lucas is the best specialist in bruises. Do you know when he has started to serve customers again? "

  a. Eu consultei o Lucas quando (eu) regressei da viagem para a Itália.

    " I consulted Lucas when I returned from the trip to Italy. "

  b. Eu consultei o Lucas quando (eu) tinha regressado da viagem para a Itália.

    " I consulted Lucas when I had returned from the trip to Italy. '

(199) A Marcela ficou emocionada com as últimas notícias da família. Você sabe quando ela ficou tão chorosa e emocionada?

" Marcela was thrilled by the latest news from the family. Do you know when she got so crying and excited? "

  a. Eu abracei a Marcela quando (eu) anunciei o noivado da filha dela.

    " I hugged Marcela when I announced the engagement of her daughter. "

  b. Eu abracei a Marcela quando (eu) estava anunciando o noivado da filha dela.

    " I hugged Marcela when I was announcing the engagement of her daughter. "

(200) O Michel não vem nas reuniões do grupo há semanas. Você sabe quando ele pretende aparecer por aqui de novo?

" Michel has not been in group meetings for weeks. Do you know when he plans to show up here again? "

  a. Eu encontrei o Michel quando (eu) adoeci com caxumba nas férias.

    " I met Michel when I got sick with mumps on vacation. "

  b. Eu encontrei o Michel quando (eu) tinha adoecido com caxumba nas férias.

    " I met Michel when I had been sick with mumps on vacation. "

(201) A Cláudia tem treinado mais do que todas as outras atletas do nosso time. Você sabe quando ela se informou sobre as rotinas de exercícios?

" Claudia has been training more than all other athletes in our team. Do you know when she learned about the exercise routines? "

  a. Eu orientei a Cláudia quando (eu) adquiri os acessórios do treino.

    " I guided Claudia when I got the training accessories. "

  b. Eu orientei a Cláudia quando (eu) ia adquirir os acessórios do treino.

    " I guided Claudia when I was going to get the training accessories. "

(202) O Cláudio andava estressado demais. Você sabe quando ele ficou mais tranquilo?

" Claudio was too stressed. Do you know when he got calmer? "

  a. Eu acalmei o Cláudio quando (eu) repeti as notícias da rádio.

    " I calmed down Claudio when I repeated the radio news. "

b. Eu acalmei o Cláudio quando (eu) ia repetir as notícias da rádio.

"I calmed down Claudio when I was going to repeat the radio news."

(203) A Luísa foi a última pessoa a mexer no aparelho desaparecido ontem. Você sabe quando ela foi perguntada sobre o aparelho?

"Luisa was the last person to touch the missing device yesterday. Do you know when she was asked about the appliance?"

a. Eu questionei a Luísa quando (eu) arrumei os materiais no armário.

"I questioned Luisa when I packed the materials in the closet."

b. Eu questionei a Luísa quando (eu) estava arrumando os materiais no armário.

"I questioned Luisa when I was packing the materials in the closet."

(204) A Fernanda ajuda muito nas reportagens sobre tragédias. Você sabe quando ela recebeu o prêmio da redação?

"Fernanda is very helpful in the news on tragedies. Do you know when she received the editorial board prize?"

a. Eu premiei a Fernanda quando (eu) terminei a reportagem sobre as cheias.

"I rewarded Fernanda when I finished the report on the floods."

b. Eu premiei a Fernanda quando (eu) tinha terminado a reportagem sobre as cheias.

"I rewarded Fernanda when I had finished the report on the floods."

(205) O Fernando trabalhou muito em uma pesquisa infrutífera. Você sabe quando ele desistiu da publicação do artigo?

"Fernando worked hard on fruitless research. Do you know when he gave up the paper publication?''

a. Eu contrariei o Fernando quando (eu) tinha confirmado a evidência científica.

"I countered Fernando when I had confirmed the scientific evidence."

b. Eu contrariei o Fernando quando (eu) tinha confirmado a evidência científica.

"I countered Fernando when I had confirmed the scientific evidence."

(206) A Eduarda não parecia feliz com a apresentação de ontem. Você sabe quando ela ficou chateada com a performance?

"Eduarda did not seem happy with yesterday's presentation. Do you know when she was upset about the performance?"

a. Eu atrapalhei a Eduarda quando (eu) apareci no palco secundário.

"I messed up Eduarda when I appeared on the secondary stage."

b. Eu atrapalhei a Eduarda quando (eu) tinha aparecido no palco secundário.

"I messed up Eduarda when I had appeared on the secondary stage."

(207) O Eugênio fez um depoimento abolutamente contraditório. Você sabe quando ele mudou a versão dos fatos?

" Eugenio made an absolutely contradictory testimony. Do you know when he changed the version of the facts? "

  a. Eu ameacei o Eugênio quando (eu) retornei do tribunal Constitucional.

    " I threatened Eugenio when I returned from the Constitutional Court. "

  b. Eu ameacei o Eugênio quando (eu) estava retornando do tribunal Constitucional.

    " I threatened Eugenio when I was returning from the Constitutional Court. "

(208) A Bianca estava se sentindo muito mal nos treinos recentemente. Você sabe quando ela teve o diagnóstico do problema?

" Bianca was feeling very bad in training recently. Do you know when she had the diagnosis of the problem? "

  a. Eu examinei a Bianca quando (eu) percebi as contusões do esforço.

    " I examined Bianca when I realized the impingements of the effort. "

  b. Eu examinei a Bianca quando (eu) estava percebendo as contusões do esforço.

    " I examined Bianca when I was noticing the impingements of the effort. "

(209) O Evandro se meteu nessas brigas de escola entre grupinhos de alunos. Você sabe quando ele escapou do grupo rival?

" Evandro got involved into these school fights among groups of students. Do you know when he escaped from the rival group? "

  a. Eu protegi o Evandro quando (eu) almocei no refeitório disfarçadamente.

    " I protected Evandro when I had lunch disguised in the cafeteria. "

  b. Eu protegi o Evandro quando (eu) estava almoçando no refeitório disfarçadamente.

    " I protected Evandro when I was having lunch disguised in the cafeteria. "

(210) A Gisele estava enganando todo mundo com a performance dela nos aparelhos. Você sabe quando ela ficou envergonhada com os resultados publicados?

" Gisele was fooling everyone with her performance on the sets. Do you know when she was embarrassed by the published results? "

  a. Eu desmenti a Gisele quando (eu) escrevi o relatório do treino.

    " I proved Gisele to be a liar when I wrote the training report. "

  b. Eu desmenti a Gisele quando (eu) ia escrever o relatório do treino.

    " I proved Gisele to be a liar when I was going to write the training report. "

(211) A Paula estava escondendo a jóia da mãe. Você sabe quando ela foi pega na mentira?

" Paula was hiding her mother's jewelry. Do you know when she got caught up in the lie?"

   a. Eu surpreendi a Paula quando (eu) encontrei o diamante escondido.

     " I surprised Paula when I found the diamond hidden. "

   b. Eu surpreendi a Paula quando (eu) tinha encontrado o diamante escondido.

     " I surprised Paula when I had found the diamond hidden. "

(212) O Paulo estava sendo acusado de ter causado um grande problema para a empresa. Você sabe quando ele se livrou da culpa pelo erro?

" Paul was being accused of causing a major problem for the company. Do you know when he got rid of the blame for the mistake? "

   a. Eu defendi o Paulo quando (eu) assumi o equívoco durante a reunião.

     " I defended Paul when I assumed the mistake during the meeting. "

   b. Eu defendi o Paulo quando (eu) tinha assumido o equívoco durante a reunião.

     " I defended Paul when I had assumed the mistake during the meeting. "

(213) A Natália é uma excelente funcionária nessa companhia. Você sabe quando ela foi admitida na empresa?

" Natalia is an excellent employee in this company. Do you know when she was admitted to the company? "

   a. Eu contratei a Natália quando (eu) alterei o projeto de investimento.

     " I hired Natália when I changed the investment project. "

   b. Eu contratei a Natália quando (eu) estava alterando o projeto de investimento.

     " I hired Natália when I was changing the investment project. "

(214) A Helena foi demitida justamente por causa do relatório escrito. Você sabe quando ela foi considerada culpada pelos erros no relatório?

" Helena was fired just because of the written report. Do you know when she was found guilty of the errors in the report? "

   a. Eu acusei a Helena quando (eu) relatei os problemas ocorridos.

     " I accused Helena when I reported the problems. "

   b. Eu acusei a Helena quando (eu) tinha relatado os problemas ocorridos.

     " I accused Helena when I had reported the problems. "

Interpretation Task

(215) Era a Maria quem tinha divulgado os resultados do exame?

" Was it Maria who had divulged the results of the exam? "

(216) Era o Pedro quem tinha desistido do campeonato regional?

" Was it Pedro who had given up the regional championship? "

(217) Era eu quem tinha corrigido o depoimento sobre o crime?

" Was I who corrected the testimony about the crime? "

(218) Era eu quem tinha indicado os defeitos da construção?

" Was I the one who had indicated the defects of the building? "

(219) Era a Clara quem tinha revelado as imagens do roubo?

" Was it Clara who had revealed the images of the robbery? "

(220) Era o Douglas quem tinha desmontado o protótipo do veículo?

" Was it Douglas who had dismantled the prototype of the vehicle? "

(221) Era eu quem tinha atendido o celular durante o ajuste?

" Was I the one who had answered the phone during the adjustment? "

(222) Era eu quem tinha regressado da viagem para a Itália?

" Was I the one who had returned from the trip to Italy? "

(223) Era a Marcela quem tinha anunciado o noivado da filha?

" Was it Marcela who had announced her daughter's engagement? "

(224) Era o Michel quem tinha adoecido com caxumba nas férias?

" Was it Michel who had got sick with mumps on vacation? "

(225) Era eu quem tinha adquirido os acessórios do treino?

" Was I the one who had bought the training accessories? "

(226) Era o Cláudio quem tinha repetido as notícias da rádio?

" Was it Claudio who had repeated the radio news? "

(227) Era a Luísa quem tinha arrumado os materiais no armário?

" Was it Louise who had packed the materials in the closet? "

(228) Era a Fernanda quem tinha terminado a reportagem sobre as cheias?

" Was it Fernanda who had finished the report on the floods? "

(229) Era o Fernando quem tinha confirmado a evidência científica?

" Was it Fernando who had confirmed the scientific evidence? '

(230) Era eu quem tinha aparecido no palco secundário?

" Was I the one who had appeared on the secondary stage? "

(231) Era eu quem tinha retornado do tribunal Constitucional?

" Was I the one who had returned from the Constitutional Court? "

(232) Era a Bianca quem tinha percebido as contusões do esforço?

" Was it Bianca who had noticed the bruises of the effort? "

(233) Era o Evandro quem tinha almoçado no refeitório disfarçadamente?

" Was it Evandro who had eaten lunch in the dining room? "

(234) Era eu quem tinha escrito o relatório do treino?

" Was I the one who wrote the training report? "

(235) Era eu quem tinha encontrado o diamante escondido?

" Was I the one who had found the diamond hidden? "

(236) Era eu quem tinha assumido o equívoco durante a reunião?

" Was I the one who had confessed the mistake during the meeting? "

(237) Era a Natália quem tinha alterado o projeto de investimento?

" Was Natalia who had altered the investment project? "

(238) Era eu quem tinha relatado os problemas ocorridos?

" Was I the one who had reported the problems? "

A.0.2   Experiment 2

Judgment Task

(239) a. A Maria estava muito irritada depois da reforma no apartamento. Você sabe o que aconteceu com a televisão dela lá?

" Maria was very annoyed after the renovation in the apartment. Do you know what happened to her television there? "

b. A Maria estava muito irritada depois da reforma no apartamento. Você sabe o que aconteceu com a colega de quarto dela lá?

" Maria was very annoyed after the renovation in the apartment. Do you know what happened to her roommate there? '

(Ela) caiu da bancada.

" She/It fell off the bench. '

(240) a. O Leandro estava rindo muito depois do mergulho no lago. Você sabe o que aconteceu com o equipamento dele lá?

" Leandro was laughing long after diving in the lake. Do you know what happened to his equipment there? "

b. O Leandro estava rindo muito depois do mergulho no lago. Você sabe o que aconteceu com o instrutor dele lá?

" Leandro was laughing long after diving in the lake. Do you know what happened to his instructor there? "

(Ele) emergiu de repente.

" He/It emerged suddenly. "

(241) a. A Clara ficou muito emocionada na sua participação no programa de televisão. Você sabe o que aconteceu com a apresentação dela lá?

" Clara was very excited about her participation in the television program. Do you know what happened to her presentation there? "

b. A Clara ficou muito emocionada na sua participação no programa de televisão. Você sabe o que aconteceu com a cantora preferida dela lá?

" Clara was very excited about her participation in the television program. Do you know what happened to her favorite singer there? "

(Ela) surgiu de surpresa.

" She/It came by surprise. "

(242) a. O Douglas estava desesperado depois do roubo da loja. Você sabe o que aconteceu com o material dele lá?

" Douglas was desperate after the theft of the store. Do you know what happened to his stuff there? "

b. O Douglas estava desesperado depois do roubo da loja. Você sabe o que aconteceu com o funcionário dele lá?

" Douglas was desperate after the theft of the store. Do you know what happened to his employee there? "

(Ele) desapareceu por alguns dias.

" He/It disappeared for a few days. '

(243) a. A Marcela ficou desolada depois do sequestro no shopping center. Você sabe o que aconteceu com a bolsa dela lá?

" Marcela was devastated after the kidnapping at the mall. Do you know what happened to her purse there? "

b. A Marcela ficou desolada depois do sequestro no shopping center. Você sabe o que aconteceu com a mãe dela lá?

" Marcela was heartbroken after the kidnapping at the mall. Do you know what happened to her mother there? "

(Ela) sumiu por mais de duas horas.

" She/It disappeared for more than two hours. "

(244) a. O Michel nunca mais foi visto depois da demissão naquela empresa. Você sabe o que aconteceu com o entusiasmo dele lá?

" Michel has never seen again after his resignation. Do you know what happened to his enthusiasm there? "

       b. O Michel nunca mais foi visto depois da demissão naquela empresa. Você sabe o que aconteceu com o chefe dele lá?

" Michel was never seen again after his resignation. Do you know what happened to his boss there? "

(Ele) morreu inexplicavelmente.

" He/It died inexplicably. '

(245) a. O Luís teve uma performance estranha no ensaio no palco ontem. Você sabe o que aconteceu com o braço dele lá?

" Luis had a strange performance in the stage rehearsal yesterday. Do you know what happened to his arm there? "

       b. O Luís teve uma performance estranha no ensaio no palco ontem. Você sabe o que aconteceu com o par dele lá?

" Luis had a strange performance in the stage rehearsal yesterday. Do you know what happened to his partner there? "

(Ele) adormeceu sem mais nem menos.

" He/It fell asleep/got dormant without reason. "

(246) a. O Cláudio perdeu a partida de volêi ontem na quadra poliesportiva. Você sabe o que aconteceu com o corpo dele lá?

" Claudio lost the volleyball match yesterday on the sports court. Do you know what happened to his body there? "

       b. O Cláudio perdeu a partida de volêi ontem na quadra poliesportiva. Você sabe o que aconteceu com o companheiro dele lá?

" Claudio lost the volleyball match yesterday on the sports court. Do you know what happened to his mate there? "

(Ele) esfriou depois do aquecimento.

" He/It cooled after heating. '

(247) a. A Luísa teve problemas na viagem para a Rússia por causa do frio. Você sabe o que aconteceu com a bota dela lá?

" Luisa had problems on her trip to Russia because of the cold. Do you know what happened to her boot there? "

    b. A Luísa teve problemas na viagem para a Rússia por causa do frio. Você sabe o que aconteceu com a amiga dela lá?

" Luisa had problems on her trip to Russia because of the cold. Do you know what happened to her friend there? "

(Ela) congelou na subida das montanhas.

" She/It froze on the rise of the mountains. "

(248)   a. A Fernanda tinha que providenciar alguns documentos depois de anos na França. Você sabe o que aconteceu com a certidão dela lá?

" Fernanda had to arrange some documents after years in France. Do you know what happened to her certificate there? "

    b. A Fernanda tinha que providenciar alguns documentos depois de anos na França. Você sabe o que aconteceu com a avó dela lá?

" Fernanda had to arrange some documents after years in France. Do you know what happened to her grandmother there? "

(Ela) caducou depois de tanto tempo.

" She/It passed after/expired so much time. "

(249)   a. O Evandro era considerado um dos principais profissionais da empresa. Você sabe o que aconteceu com o funcionário dele lá?

" Evandro was considered one of the main professionals of the company. Do you know what happened to his employee there? "

    b. O Evandro era considerado um dos principais profissionais da empresa. Você sabe o que aconteceu com o contrato dele lá?

" Evandro was considered one of the main professionals of the company. Do you know what happened to his contract there? "

(Ele) durou muito mais do que o esperado.

" He/It lasted much longer than expected. "

(250)   a. O Cláudio fez uma grande contribuição dando aulas na Holanda. Você sabe o que aconteceu com o potencial dele lá?

" Claudio made a great contribution by giving classes in Holland. Do you know what happened to his potential there? "

    b. O Cláudio fez uma grande contribuição dando aulas na Holanda. Você sabe o que aconteceu com o

orientando dele lá?

" Claudio made a great contribution teaching in Holland. Do you know what happened to his guiding there? "

(Ele) amadureceu bastante.

" He/It matured a lot. "

(251)   a. O Paulo inscreveu um grupo de estudantes para a vaga aberta na universidade. Você sabe o que aconteceu com o arquivo dele lá?

" Paulo enrolled a group of students for the open position at the university. Do you know what happened to his file there? "

       b. O Paulo inscreveu um grupo de estudantes para a vaga aberta na universidade. Você sabe o que aconteceu com o estudante dele lá?

" Paulo enrolled a group of students for the open space at the university. Do you know what happened to his student there? "

(Ele) constou como registrado.

" He/It was recorded. "

(252)   a. A Natália está dando péssimas aulas ultimamente no grupo de dança folclórica. Você sabe o que aconteceu com a performance dela lá?

" Natalia has been doing terrible classes lately in the folk dance group. Do you know what happened to her performance there? "

       b. A Natália está dando péssimas aulas ultimamente no grupo de dança folclórica. Você sabe o que aconteceu com a dançarina dela lá?

" Natalia has been doing terrible classes lately in the folk dance group. Do you know what happened to her dancer there? "

(Ela) enferrujou depois das férias.

" She/It rusted after the holidays. '

(253)   a. A Helena era uma candidata política à frente do seu tempo. Você sabe o que aconteceu com a ideia dela lá?

" Helena was a political candidate ahead of her time. Do you know what happened to her idea there? "

       b. A Helena era uma candidata política à frente do seu tempo. Você sabe o que aconteceu com a secretária dela lá?

" Helena was a political candidate ahead of her time. Do you know what happened to her secretary there? "

(Ela) sobreviveu aos ataques dos adversários.

" She/It survived her opponents' attacks. "

(254) a. A Mônica teve problemas com a entrega dos produtos no novo apartamento. Você sabe o que aconteceu com a compra dela lá?

" Monica had problems with the delivery of the products in the new apartment. Do you know what happened to buying her there? "

b. A Mônica teve problemas com a entrega dos produtos no novo apartamento. Você sabe o que aconteceu com a colega de apartamento dela lá?

" A Mônica teve problemas com a entrega dos produtos no novo apartamento. Você sabe o que aconteceu com a colega de apartamento dela lá? "

(Ela) veio de volta.

" She/It came back. "

(255) a. O Flávio não encontrava a última encomenda feita pelo site da empresa. Você sabe o que aconteceu com o pedido dele lá?

" Flavio did not find the last order made by the company's website. Do you know what happened to his request there? "

b. O Flávio não encontrava a última encomenda feita pelo site da empresa. Você sabe o que aconteceu com o entregador dele lá?

" Flavio did not find the last order made by the company's website. Do you know what happened to his deliveryman there? "

(Ele) foi para o lugar errado.

" He/It went to the wrong place. "

(256) a. A Bruna parecia muito triste depois da apresentação de dança. Você sabe o que aconteceu com a sapatilha dela lá?

" Bruna looked very sad after the dance performance. Do you know what happened to her shoe there? "

b. A Bruna parecia muito triste depois da apresentação de dança. Você sabe o que aconteceu com a dançarina dela lá?

" Bruna looked very sad after the dance performance. Do you know what happened to her dancer there? "

(Ela) escorregou no palco.

" She/It slipped on the stage. "

(257) a. O André teve uma péssima experiência na empresa. Você sabe o que aconteceu com o currículo dele lá?

" André had a terrible experience in the company. Do you know what happened to his curriculum there? "

b. O André teve uma péssima experiência na empresa. Você sabe o que aconteceu com o diretor dele lá?

" André had a terrible experience in the company. Do you know what happened to his director there? "

380

(Ele) retornou ao departamento pessoal.

" He/It returned to the personal department. "

(258)   a. A Júlia se desencontrou com os materias para a competição. Você sabe o que aconteceu com a encomenda dela lá?

      " Júlia got in the way of the competition. Do you know what happened to her order there? "

   b. A Júlia se desencontrou com os materias para a competição. Você sabe o que aconteceu com a treinadora dela lá?

      " Júlia disagreed with the subjects for the competition. Do you know what happened to her coach there? "

(Ela) partiu antes do planejado.

" She/It left earlier than planned. "

(259)   a. O Mateus teve uma excelente surpresa voltando para casa. Você sabe o que aconteceu com o pacote dele lá?

      " Matthew had a great surprise coming home. Do you know what happened to his package there? "

   b. O Mateus teve uma excelente surpresa voltando para casa. Você sabe o que aconteceu com o irmão dele lá?

      " Mateus had a great surprise coming home. Do you know what happened to his brother there? "

(Ele) chegou antes do esperado.

" He/It arrived earlier than expected. "

(260)   a. A Rebeca viveu uma tragédia na última competição. Você sabe o que aconteceu com a panturrilha dela lá?

      " Rebecca experienced a tragedy in the last competition. Do you know what happened to her calf there? "

   b. A Rebeca viveu uma tragédia na última competição. Você sabe o que aconteceu com a dupla dela lá?

      " Rebecca experienced a tragedy in the last competition. Do you know what happened to her duo there? "

(Ela) inchou repentinamente.

" She/It swelled suddenly. "

(261)   a. O Ricardo teve um fim trágico na competição de rali. Você sabe o que aconteceu com o carro dele lá?

      " Ricardo had a tragic end to the rally competition. Do you know what happened to his car there? "

   b. O Ricardo teve um fim trágico na competição de rali. Você sabe o que aconteceu com o copiloto dele lá?

      " Ricardo had a tragic end to the rally competition. Do you know what happened to his co-driver there? "

(Ele) despencou da colina.

" He/It fell off the hill. "

(262)   a. A Viviane ficou muito feliz depois da surpresa no jantar na casa dos pais na sexta. Você sabe o que aconteceu com a encomenda dela lá?

" Viviane was very happy after the surprise dinner at her parents' house on Friday. Do you know what happened to her order there? "

b. A Viviane ficou muito feliz depois da surpresa no jantar na casa dos pais na sexta. Você sabe o que aconteceu com a irmã dela lá?

" Viviane was very happy after the surprise dinner at her parents' house on Friday. Do you know what happened to her sister there? "

(Ela) apareceu de surpresa.

" She/It came by surprise. "

Interpretation Task

(263) Então, era uma pessoa que tinha caído da bancada?

" So was it a person who had fallen off the bench? "

(264) Então, era um objeto que tinha emergido de repente?

" So was it an object that had suddenly emerged? "

(265) Então, era uma pessoa que tinha surgido de surpresa?

" So was it a person who had come by surprise? "

(266) Então, era um objeto que tinha desaparecido por alguns dias?

" So was it an object that had disappeared for a few days? "

(267) Então, era um objeto que tinha sumido por mais de duas horas?

" So was it an object that had disappeared for more than two hours? "

(268) Então, era uma pessoa que tinha morrido inexplicavelmente?

" So was it a person who had died inexplicably? "

(269) Então, era uma pessoa que tinha adormecido sem mais nem menos?

" So was it a person who had just fallen asleep? "

(270) Então, era uma pessoa que tinha esfriado depois do aquecimento?

" So was it a person who had cooled after the warm up? "

(271) Então, era um objeto que tinha congelado na subida das montanhas?

" So was it an object that had frozen in the ascent of the mountains? "

(272) Então, era uma pessoa que tinha caducado depois de tanto tempo?

" So was it a person who had expired after so long? "

(273) Então, era uma pessoa que tinha durado muito mais do que o esperado?

" So was it a person who had lasted much longer than expected? "

(274) Então, era uma pessoa que tinha amadurecido bastante?

" So was it a person who had matured enough? "

(275) Então, era um objeto que tinha constado como registrado?

" So was it an object that had been recorded as recorded? "

(276) Então, era uma pessoa que tinha enferrujado depois das férias?

" So was it a person who had rusted after the vacation? "

(277) Então, era uma pessoa que tinha sobrevivido aos ataques dos adversários?

" So was it a person who had survived the opponents' attacks? "

(278) Então, era um objeto que tinha vindo de volta?

" So was it an object that had come back? "

(279) Então, era um objeto que tinha ido para o lugar errado?

" So was it an object that had gone to the wrong place? "

(280) Então, era um objeto que tinha escorregado no palco?

" So was it an object that had gone to the wrong place? "

(281) Então, era um objeto que tinha retornado ao departamento pessoal?

" So was it an object that had returned to the personal department? "

(282) Então, era um objeto que tinha partido antes do planejado?

" So was it an object that had left before planned? "

(283) Então, era um objeto que tinha chegado antes do esperado?

" So was it an object that had arrived earlier than expected? "

(284) Então, era uma pessoa que tinha inchado repentinamente?

" So was it a person who had swollen suddenly? "

(285) Então, era um objeto que tinha despencado da colina?

" So was it an object that had fallen off the hill? "

(286) Então, era um objeto que tinha aparecido de surpresa?

" So was it an object that had come by surprise? "

A.0.3 Experiment 3a

Judgment Task

(287) O João é um dos principais advogados da firma de contabilidade. Você sabe se o João trabalha na empresa até as 20h?

" John is one of the accounting firm's top lawyers. Do you know if John works at the company until 8:00 p.m.? "

   a. A Maria diz que trabalha (lá sempre trancado).

     " Maria says that (s)he works there always locked in."

   b. Todo advogado diz que trabalha (lá sempre trancado).

     " Every lawyer says they/he work(s) there always locked in."

(288) O Pedro é um dos melhores jodadores da equipe de vôlei da escola. Você sabe se o Pedro treina na quadra poliesportiva sábado de manhã?

" Pedro is one of the best players on the school's volleyball team. Do you know if Pedro practices on the multi-sport court on Saturday mornings?"

   a. A Joana diz que treina lá mesmo cansado.

     " Joana says that (s)he practice there even tired."

   b. Todo jogador diz que treina lá mesmo cansado.

     " Every player says they/he practice(s) there even tired."

(289) A Caroline é uma das enfermeiras mais pontuais do hospital. Você sabe se a Caroline se prepara na enfermaria antes das 8h?

" Caroline is one of the most puny nurses in the hospital. Do you know if Caroline gets ready at the infirmary before 8:00?"

   a. O Marcelo fala que se prepara lá toda atenciosa.

     " Marcelo says that (s)he gets ready there very attentively."

   b. Toda enfermeira fala que se prepara lá toda atenciosa.

     " Every nurse says that they/she get(s) ready there very attentively."

(290) O Henrique é um dos escultores mais talentosos da oficina de artesanato. Você sabe se o Henrique pratica na oficina mais de 8h por dia?

" Henry is one of the most talented sculptors in the craft shop. Do you know if Henrique practices in the workshop more than 8 hours a day?"

   a. A Cláudia fala que pratica lá muito atento.

     " Claudia says that (s)he practices very attentively."

    b. Todo escultor fala que pratica lá muito atento.

       " Every sculptor says that they/he practice(s) there very attentively."

(291) A Inês é uma das médicas mais trabalhadoras da clínica. Você sabe se a Inês vai para o consultório depois das 6h da tarde?

" Inês is one of the clinic's most hardworking doctors. Do you know if Inês goes to the office after 6:00 p.m.?"

    a. O Felipe fala que vai lá mesmo exausta.

       " Felipe says (s)he goes there even exhausted."

    b. Toda médica fala que vai lá mesmo exausta.

       " Every doctor says they/she go(es) there even exhausted."

(292) A Ana é uma das melhores dentistas do consultório. Você sabe se a Ana fica na sala depois das 6h da tarde?

" Ana is one of the best dentists in the office. Do you know if Ana stays in the room after 6:00 p.m.?"

    a. O Cláudio diz fica lá sempre super disposta.

       " Claudio says that (s)he stays there always super willing."

    b. Toda dentista diz fica lá sempre super disposta.

       " Every dentist says that they/she stay(s) there always super willing."

(293) A Eduarda é uma das bailarinas mais bonitas da companhia de dança. Você sabe se a Eduarda dança nos fins de semana no teatro?

" Eduarda is one of the most beautiful ballerinas in the dance company. Do you know if Eduarda dances on weekends at the theater?"

    a. O Gustavo fala que dança lá muito bem vestida.

       " Gustavo says that (s)he dances there very well dressed."

    b. Toda bailarina fala que dança lá muito bem vestida.

       " Every ballerina says that they/she dances there very well dressed."

(294) O Sérgio é um dos mais pintores mais requisitados na exposição. Você sabe se o Sérgio pinta no ateliê durante a semana?

" Sergio is one of the most requested painters in the exhibition. Do you know if Sergio paints in the studio during the week?"

    a. A Cecília diz que pinta lá mesmo às vezes sonolento.

       " Cecilia says that (s)he paints there even sometimes drowsy."

    b. Todo pintor diz que pinta lá mesmo às vezes sonolento.

       " Every painter says that they/he paint(s) there even sometimes drowsy."

(295) O Lúcio é um dos professores mais produtivos da nossa universidade. Você sabe se o Lúcio estuda no escritório pela manhã?

" Lucio is one of the most productive teachers of our university. Do you know if Lucio studies in the office this morning?"

  a. A Mariana diz que estuda lá mesmo distraído.

    " Mariana says that (s)he studies there even distracted."

  b. Todo professor diz estuda lá mesmo distraído.

    " Every teacher says that they/he study/ies even distracted."

(296) A Sônia é uma das veterinárias mais conhecidas desta cidade. Você sabe se a Sônia atende na clínica no sábado?

" Sônia is one of the best known veterinarians of this city. Do you know if Sonia works in the clinic on Saturdays?"

  a. O Eduardo fala que atende lá sempre muito cuidadosa.

    " Eduardo says that (s)he always works there very carefully."

  b. Toda veterinária fala que atende lá sempre muito cuidadosa.

    " Every veterinarian says that they/she work(s) very carefully."

(297) A Érica é uma das jornalistas mais lidas deste jornal. Você sabe se a Érica permanece na redação até de madrugada?

" Erica is one of the newspaper's most read journalists. Do you know if Erica stays in the newsroom until dawn?"

  a. O Júlio diz que permanece lá mesmo contrariada.

    " Julius says that (s)he remains there even upset."

  b. Toda jornalista diz que permanece lá mesmo contrariada.

    " Every journalist says that they/she remain(s) there even upset."

(298) A Lúcia é uma das chefes de cozinha mais apreciadas do nosso restaurante. Você sabe se a Lúcia se concentra na cozinha antes das 10h da manhã?

" Lucia is one of the most appreciated chefs of our restaurant. Do you know if Lucia gets concentrated in the kitchen before 10 AM?"

  a. O Rafael diz que se concentra lá sempre animada.

    " Rafael says that (s)he get(s) concentrated there always excited."

  b. Toda cozinheira diz que se concentra lá sempre animada.

    " Every cook says that they/she get(s) concentrated there always excited."

(299) A Vera é uma das motoristas mais contratadas na nossa empresa. Você sabe se a Vera dirije na rota São Paulo-Goiás somente durante o dia?

" Vera is one of the most hired drivers in our company. Do you know if Vera drives on the São Paulo-Goiás route only during the day?"

  a. O Daniel fala que dirige lá sempre sozinha.

    " Daniel says that (s)he drives there alone."

  b. Toda motorista fala que dirige lá sempre sozinha.

    " Every driver says that they/she drive(s) there alone."

(300) A Helena é uma das juízas mais corretas desta cidade. Você sabe se a Helena atua no tribunal todas as tardes?

" Helena is one of the most correct judges in this city. Do you know if Helena works in the courtroom every afternoon?"

  a. O Luís diz que atua lá rigorosa como sempre.

    " Luis says that (s)he acts there as strictly as ever."

  b. Toda juíza diz que atua lá rigorosa como sempre.

    " Every judge says that they/she act(s) there as rigorously as ever."

(301) A Pâmela é uma das garçonetes que recebe mais gorjetas neste bar. Você sabe se a Pâmela sorri enquanto serve as mesas?

" Pâmela is one of the waitresses who receives more tips at this bar. Do you know if Pâmela smiles while serving the tables?"

  a. O Paulo fala que sorri lá super simpática.

    " Paulo says that (s)he smiles there super nice."

  b. Toda garçonete fala que sorri lá super simpática.

    " Every waitress says that they/she smile(s) there super friendly."

(302) O Cristian é um dos ginastas mais fortes da academia. Você sabe se o Cristian malha na academia no sábado?

" Cristian is one of the strongest gymnasts of the academy. Do you know if Cristian knits at the academy on Saturday? '

  a. A Giovana diz que malha lá muito concentrado.

    " Giovana says that (s)he works out there very concentrated."

  b. Toda ginasta diz que malha lá muito concentrado.

    " Every gymnast says that they/she work(s) out there very concentrated."

(303) O Gabriel é um dos arquitetos mais competentes do nosso escritório. Você sabe se o Gabriel chega na construção antes das 9h?

" Gabriel is one of the most competent architects in our office. Do you know if Gabriel gets to the building before 9am?"

   a. A Francine fala que chega lá envolvido no trabalho.

      " Francine says that (s)he gets there involved in work."

   b. Todo arquiteto fala que chega lá envolvido no trabalho.

      " Every architect says that they/he get(s) there involved in work."

(304) O Fabiano é um dos melhores triatletas no nado nas competições regionais. Você sabe se o Fabiano nada na piscina da academina nos fins de semana?

" Fabiano is one of the best triathletes in swimming in regional competitions. Do you know if Fabiano swims in the academy pool on the weekends?"

   a. A Natália diz que nada lá muito entusiasmado.

      " Natalia says that (s)he swims there very enthusiastic."

   b. Todo triatleta diz que nada lá muito entusiasmado.

      " Every triathlete says that they/he swim(s) there very enthusiastic."

(305) O Jean é um dos vendedores com as maiores metas na nossa loja. Você sabe se o Jean grita na porta da loja durante o dia?

" Jean is one of the salespeople with the highest goals in our store. Do you know if Jean shouts at the door of the store during the day?"

   a. A Júlia diz que grita lá mesmo rouco.

      " Julia says that (s)he shouts there even in a hoarse voice."

   b. Todo vendedor diz que grita lá mesmo rouco.

      " Every salesperson says that they/he shout(s) there even in a hoarse voice."

(306) O Bruno é um dos modelos mais chamados desta agência. Você sabe se o Bruno desfila para aquela marca famosa na fashion week?

" Bruno is one of the most called models of this agency. Do you know if Bruno cat-walks for that famous brand in the fashion week?"

   a. A Aline diz que desfila lá muito seguro.

      " Aline says that (s)he cat-walks there very confident."

   b. Todo modelo diz que desfila lá muito seguro.

      " Every model says that they/he cat-walk(s) there very confident."

Interpretation Task

(307)   a.   A Maria afirma que o João trabalha na empresa até as 20h?

          " Does Maria say that João works at the company until 8:00 p.m.? "

       b.   Todo advogado afirma que o João trabalha na empresa até as 20h?

          " Does every lawyer say that João works at the company until 8:00 p.m.? "

(308)   a.   A Joana confirma que o Pedro treina na quadra poliesportiva?

          " Does Joana confirm that Pedro practices on the multi-sport court? "

       b.   Todo jogador confirma que o Pedro treina na quadra poliesportiva?

          " Does every player confirm that Pedro practices on the multi-sport court? "

(309)   a.   O Marcelo garante que a Caroline se prepara na enfermeira antes das 8h?

          " Does Marcelo ensure that Caroline gets ready for the nurse before 8am? "

       b.   Toda enfermeira garante que a Caroline se prepara na enfermeira antes das 8h?

          " Does every nurse ensure that Caroline gets ready for the nurse before 8:00? "

(310)   a.   A Cláudia afirma que o Henrique pratica na oficina mais de 8h por dia?

          " Does Claudia say that Henrique practices in the workshop more than 8 hours a day? "

       b.   Todo escultor afirma que o Henrique pratica na oficina mais de 8h por dia?

          " Does every sculptor say that Henrique practices in the workshop more than 8 hours a day? "

(311)   a.   O Felipe confirma que a Inês vai para o consultório depois das 6h da tarde?

          " Does Felipe confirm that Ines goes to the office after 6:00 p.m.? "

       b.   Toda médica confirma que a Inês vai para o consultório depois das 6h da tarde?

          " Does every doctor confirm that Ines goes to the office after 6:00 p.m.? "

(312)   a.   O Cláudio garante que a Ana fica na sala depois das 6h da tarde?

          " Does Claudio ensure that Ana stays in the room after 6:00 p.m.? "

       b.   Toda dentista garante que a Ana fica na sala depois das 6h da tarde?

          " Does Claudio ensure that Ana stays in the room after 6:00 p.m.? "

(313)   a.   O Gustavo afirma que a Eduarda dança nos fins de semana no teatro?

          " Gustavo say that Eduarda dances on weekends in the theater? "

       b.   Toda bailarina afirma que a Eduarda dança nos fins de semana no teatro?

          " Does every ballerina say that Eduarda dances on weekends in the theater? "

(314)   a.   A Cecília confirma que o Sérgio pinta no ateliê durante a semana?

          " Does Cecilia confirm that Sergio paints in the studio during the week? "

389

b. Todo pintor confirma que o Sérgio pinta no ateliê durante a semana?

" Does every painter confirm that Sergio paints in the studio during the week? "

(315) a. A Mariana garante que o Lúcio estuda no escritório pela manhã?

" Does Mariana ensure that Lucio studies in the office in the morning? "

b. Todo professor garante que o Lúcio estuda no escritório pela manhã?

" Does every teacher ensure that Lucio studies in the office in the morning? "

(316) a. O Eduardo afirma que a Sônia atende na clínica no sábado?

" Does Eduardo say that Sonia works the clinic on Saturday? "

b. Toda veterinário afirma que a Sônia atende na clínica no sábado?

" Does every veterinarian say that Sonia works the clinic on Saturday? "

(317) a. O Júlio confirma que a Érica permanece na redação até de madrugada?

" Does Julio confirm that Erica stays in the newsroom until dawn? "

b. Toda jornalista confirma que a Érica permanece na redação até de madrugada?

" Does every journalist confirm that Erica stays in the newsroom until dawn? "

(318) a. O Rafael garante que a Lúcia se concentra na cozinha antes das 10h da manhã?

" Does Rafael ensure Lucy gets concentrated the kitchen before 10:00 in the morning? "

b. Toda cozinheira garante que a Lúcia se concentra na cozinha antes das 10h da manhã?

" Does every cook ensure that Lucy gets concentrated the kitchen before 10 a.m. in the morning? "

(319) a. O Daniel afirma que a Vera dirige na rota São Paulo-Goiás durante o dia?

" Does Daniel say that Vera drives on the São Paulo-Goiás route during the day? "

b. Toda motorista afirma que a Vera dirige na rota São Paulo-Goiás durante o dia?

" Does every driver say that Vera drives on the São Paulo-Goiás route during the day? "

(320) a. O Luís confirma que a Helena atua no tribunal todas as tardes?

" Does Luis confirm that Helena works in court every afternoon? "

b. Toda juíza confirma que a Helena atua no tribunal todas as tardes?

" Does every judge confirm that Helena works in court every afternoon? "

(321) a. O Paulo garante que a Pâmela sorri enquanto serve as mesas?

" Does Paulo ensure that Pâmela smiles while serving the tables? "

b. Toda garçonete garante que a Pâmela sorri enquanto serve as mesas?

" Does every waitress ensure that Pâmela smiles while serving the tables? "

(322) a. A Giovana afirma que o Cristian malha na academia no sábado?

" Does Giovana say that Cristian works out in the gym on Saturday? "

b. Todo ginasta afirma que o Cristian malha na academia no sábado?

" Does every gymnast say that Cristian works out in the gym on Saturday? "

(323) a. A Francine confirma que o Gabriel chega na construção antes das 9h?

" Francine confirms that Gabriel gets to the building before 9am? "

b. Todo arquiteto confirma que o Gabriel chega na construção antes das 9h?

" Does every architect confirm that Gabriel gets to the building before 9am? "

(324) a. A Natália garante que o Fabiano nada na piscina da academia nos fins de semana?

" Does Natalia ensure that Fabiano swims in the gym pool on weekends? "

b. Todo triatleta garante que o Fabiano nada na piscina da academia nos fins de semana?

" Does every triathlete ensure that Fabiano swims in the gym pool on weekends? "

(325) a. A Júlia afirma que o Jean grita na porta da loja durante o dia?

" Does Julia say that Jean shout out at the door of the store during the day? "

b. Todo vendedor afirma que o Jean grita na porta da loja durante o dia?

" Does every salesman say that Jean shout out at the store door during the day? "

(326) a. A Aline confirma que o Bruno desfila para aquela marca famosa na fashion week?

" Aline confirms that Bruno cat-walks for that famous brand in fashion week? "

b. Todo modelo confirma que o Bruno desfila para aquela marca famosa na fashion week?

" Every model confirms that Bruno cat-walks for that famous brand in fashion week? "

A.0.4 Experiment 3b

Judgment Task

(327) a. O João é um dos principais advogados da firma de contabilidade. Você sabe se o João trabalha na empresa até as 20h?

"John is one of the accounting firm's top lawyers. Do you know if John works at the company until 8:00 p.m.?"

b. Todo advogado diz que (ele) trabalha (lá sempre trancado).

" Every lawyer says that he/they work(s) there always locked in. "

(328) a. O Pedro é um dos melhores jodadores da equipe de vôlei da escola. Você sabe se o Pedro treina na quadra poliesportiva sábado de manhã?

"Pedro is one of the best players on the school's volleyball team. Do you know if Pedro is training on the multi-sport court Saturday morning?"

b. Todo jogador diz que (ele) treina (lá mesmo cansado).

" Every player says they/he practice(s) there even tired. "

(329) a. A Caroline é uma das enfermeiras mais pontuais do hospital. Você sabe se a Caroline se prepara na enfermaria antes das 8h?

"Caroline is one of the most puny nurses in the hospital. Do you know if Caroline gets ready at the infirmary before 8:00?"

b. Toda enfermeira fala que (ela) se prepara (lá toda atenciosa).

" Every nurse says that they/she get(s) ready there very attentively. "

(330) a. O Henrique é um dos escultores mais talentosos da oficina de artesanato. Você sabe se o Henrique pratica na oficina mais de 8h por dia?

"Henry is one of the most talented sculptors in the craft shop. Do you know if Henrique practices in the workshop more than 8 hours a day? "

b. Todo escultor fala que (ele) pratica (lá muito atento).

"Every sculptor says that they/he practice(s) there very attentively."

(331) a. A Inês é uma das médicas mais trabalhadoras da clínica. Você sabe se a Inês vai para o consultório depois das 6h da tarde?

"Inês is one of the clinic's most hardworking doctors. Do you know if Agnes goes to the office after 6:00 p.m.?"

b. Toda médica fala que (ela) vai (lá mesmo exausta).

" Every doctor says they/she go(es) there even exhausted. "

(332) a. A Ana é uma das melhores dentistas do consultório. Você sabe se a Ana fica na sala depois das 6h da tarde?

"Ana is one of the best dentists in the office. Do you know if Ana stays in the room after 6:00 p.m.?"

b. Toda dentista diz que (ela) fica (lá sempre super disposta).

" Every dentist says that they/she stay(s) there always super willing. "

(333) a. A Eduarda é uma das bailarinas mais bonitas da companhia de dança. Você sabe se a Eduarda dança nos fins de semana no teatro?

"Eduarda is one of the most beautiful ballerinas in the dance company. Do you know if Eduarda dances on weekends at the theater?"

b. Toda bailarina fala que (ela) dança (lá muito bem vestida).

" Every ballerina says that they/she dances there very well dressed. "

(334) a. O Sérgio é um dos mais pintores mais requisitados na exposição. Você sabe se o Sérgio pinta no ateliê durante a semana?

"Sergio is one of the most requested painters in the exhibition. Do you know if Sergio paints in the studio during the week?"

b. Todo pintor diz que (ele) pinta (lá mesmo às vezes sonolento).

" Every painter says that they/he paint(s) there even sometimes drowsy. "

(335) a. O Lúcio é um dos professores mais produtivos da nossa universidade. Você sabe se o Lúcio estuda no escritório pela manhã?

"Lucio is one of the most productive teachers of our university. Do you know if Lucius is studying in the office this morning?"

b. Todo professor diz (ele) estuda (lá mesmo distraído).

" Every teacher says that they/he study/ies even distracted. "

(336) a. A Sônia é uma das veterinárias mais conhecidas desta cidade. Você sabe se a Sônia atende na clínica no sábado?

"Sonia is one of the best known veterinarians of this city. Do you know if Sonia attends the clinic on Saturday?"

b. Toda veterinária fala que (ela) atende (lá sempre muito cuidadosa).

" Every veterinarian says that she always goes there very carefully. "

(337) a. A Érica é uma das jornalistas mais lidas deste jornal. Você sabe se a Érica permanece na redação até de madrugada?

"Erica is one of the newspaper's most read journalists. Do you know if Erica stays in the newsroom until dawn?"

b. Toda jornalista diz que (ela) permanece (lá mesmo contrariada).

" Every journalist says that they/she remain(s) there even upset. "

(338) a. A Lúcia é uma das chefes de cozinha mais apreciadas do nosso restaurante. Você sabe se a Lúcia se concentra na cozinha antes das 10h da manhã?

"Lúcia is one of the most appreciated chefs of our restaurant. Do you know if Lucy focuses on the kitchen before 10 AM?"

(339) a. Toda chefe de cozinha diz que (ela) se concentra (lá sempre disposta).

" Every cook says that they/she get(s) concentrated there always excited. "

(340) a. A Vera é uma das motoristas mais contratadas na nossa empresa. Você sabe se a Vera dirije na rota São Paulo-Goiás somente durante o dia?

"Vera is one of the most hired drivers in our company. Do you know if Vera drives on the São Paulo-Goiás route only during the day?"

b. Toda motorista fala que (ela) dirige (lá sempre sozinha).

"Every driver says that they/she drive(s) there alone."

(341) a. A Helena é uma das juízas mais corretas desta cidade. Você sabe se a Helena atua no tribunal todas as tardes?

"Helena is one of the most correct judges in this city. Do you know if Helena works in the courtroom every afternoon?"

b. Toda juíza diz que (ela) atua (lá rigorosa como sempre).

"Every judge says that they/she act(s) there as rigorously as ever."

(342) a. A Pâmela é uma das garçonetes que recebe mais gorjetas neste bar. Você sabe se a Pâmela sorri enquanto serve as mesas?

"Pâmela is one of the waitresses who receives more tips at this bar. Do you know if Pâmela smiles while serving the tables?"

b. Toda garçonete fala que (ela) sorri (lá super simpática).

"Every waitress says that they/she smile(s) there super friendly."

(343) a. O Cristian é um dos ginastas mais fortes da academia. Você sabe se o Cristian malha na academia no sábado?

"Cristian is one of the strongest gymnasts of the academy. Do you know if Cristian knits at the academy on Saturday?"

b. Todo ginasta diz que (ele) malha (lá muito concentrado).

"Every gymnast says that they/she work(s) out there very concentrated."

(344) a. O Gabriel é um dos arquitetos mais competentes do nosso escritório. Você sabe se o Gabriel chega na construção antes das 9h?

"Gabriel is one of the most competent architects in our office. Do you know if Gabriel arrives at the building before 9am?"

b. Todo arquiteto fala que (ele) chega (lá envolvido no trabalho).

"Every architect says that they/he get(s) there involved in work."

(345) a. O Fabiano é um dos melhores triatletas no nado nas competições regionais. Você sabe se o Fabiano nada na piscina da academina nos fins de semana?

"Fabiano is one of the best triathletes in swimming in regional competitions. Do you know if Fabiano swims in the academy pool on the weekends?"

b. Todo triatleta diz que (ele) nada (lá muito entusiasmado).

"Every triathlete says that they/he swim(s) there very enthusiastic."

(346)   a. O Jean é um dos vendedores com as maiores metas na nossa loja. Você sabe se o Jean grita na porta da loja durante o dia?

"Jean is one of the salespeople with the highest goals in our store. Do you know if Jean screams at the door of the store during the day?"

b. Todo vendedor diz que (ele) grita (lá mesmo rouco).

"Every salesperson says that they/he shout(s) there even in a hoarse voice."

(347)   a. O Bruno é um dos modelos mais chamados desta agência. Você sabe se o Bruno desfila para aquela marca famosa na fashion week?

"Bruno is one of the most called models of this agency. Do you know if Bruno parades for that famous brand in fashion week?"

b. Todo modelo diz que (ele) desfila (lá muito seguro).

"Every model says that they/he cat-walk(s) there very confident."

Interpetation Task

(348)   Todo advogado afirma que o João trabalha na empresa até as 20h?

" Does every lawyer say that João works at the company until 8:00 p.m.? "

(349)   Todo jogador confirma que _ treina na quadra poliesportiva?

" Does every player confirm that they practice on the multi-sport court? "

(350)   Toda enfermeira diz que a Caroline se prepara na enfermaria antes das 8h?

" Does every nurse say that Caroline gets ready at the infirmary before 8:00? "

(351)   Todo escultor afirma que _ pratica na oficina mais de 8h por dia?

" Does every sculptor say that they practice in the workshop more than 8 hours a day? "

(352)   Toda médica confirma que a Inês vai para o consultório depois das 6h da tarde?

" Does every doctor confirm that Inês goes to the office after 6:00 p.m.? "

(353)   Toda dentista fala que _ fica na sala depois das 6h da tarde?

" Does every dentist say they stay in the room after 6:00 in the afternoon? "

(354)   Toda bailarina afirma que a Eduarda dança nos fins de semana no teatro?

" Does every dancer say that Eduarda dances on weekends in the theater? "

(355) Todo pintor fala que _ pinta no ateliê durante a semana?

" Every painter says that they paint in the studio during the week? "

(356) Todo professor fala que o Lúcio estuda no escritório pela manhã?

" Does every teacher say that Lucio studies in the office in the morning? "

(357) Toda veterinária afirma que _ atende na clínica no sábado?

" Does every veterinarian say that they work at the clinic on Saturday? "

(358) Toda jornalista confirma que a Érica permanece na redação até de madrugada?

" Does every journalist confirm that Erica stays in the newsroom until dawn? "

(359) Toda chefe de cozinha fala que _ se concentra na cozinha antes das 10h da manhã?

" Does every chef say they get concentrated in the kitchen before 10 a.m. in the morning? "

(360) Toda motorista afirma que a Vera dirige na rota São Paulo-Goiás somente de dia?

" Does every driver claim that Vera drives on the São Paulo-Goiás route only during the day? "

(361) Toda juíza confirma que _ atua no tribunal todas as tardes?

" Does every judge confirm that they acts in the court every afternoon? "

(362) Toda garçonete diz que a Pâmela sorri enquanto serve as mesas?

" Every waitress says that Pâmela smiles while serving the tables? "

(363) Todo ginasta afirma que _ malha na academia no sábado?

" Does every gymnast says that they work out at the gym on Saturday? "

(364) Todo arquiteto confirma que o Gabriel chega na construção antes das 9h?

" Does every architect confirm that Gabriel arrives at the building before 9am? "

(365) Todo triatleta fala que _ nada na piscina da academina nos fins de semana?

" Does every triathlete say they swim in the gym pool on the weekends? "

(366) Todo vendedor afirma que o Jean grita na porta da loja durante o dia?

" Does every salesman says that Jean shouts at the store door during the day? "

(367) Todo modelo confirma que _ desfila para aquela marca famosa na fashion week?.

" Every model confirms that they cat-walk for that famous brand in fashion week ?. "

A.0.5   Experiment 4

Judgment Task

(368)   a. Só o João estava comendo os sanduíches hoje à tarde.

" Only João was eating the sandwiches this afternoon. "

b. i. O João comeu aqueles com maionese?

" Has João eat those with mayonnaise? "

ii. O que o João fez com aqueles com maionese?

" What has João done to those with mayonnaise? "

c. (Infelizmente) (ele) comeu.

" Yes, (he has)./He's eaten them./Yes, (he has) unfortunately (done so)./Unfortunately, he's eaten them. "

(369) a. Só o Felipe fez os exercícios com a material esportivo hoje de manhã.

" Only Felipe worked out with the sport equipment this morning. "

b. i. O Felipe guardou as bolas de basquete?

" Has Felipe stored the basketballs? "

ii. O Felipe o João fez com as bolas de basquete?

" What has Felipe done to the basketballs? "

c. (Sem dúvida) (ele) guardou.

" Yes, (he has)./He's stored them./Yes, (he has) no doubt (done so)./He's no doubt stored them. "

(370) a. Só o Pedro fez pesquisa na biblioteca hoje à tarde.

" Only Peter researched at the library this afternoon. "

b. i. O Pedro arrumou os livros depois?

" Has Peter organized the books then? "

ii. O que o Pedro fez com livros depois?

" What has Peter done to the books then? "

c. (Aparentemente) (ele) arrumou.

" Yes, (he has)./He's organized them./Yes, (he has) apparently (done so)./He's organized them. "

(371) a. Só o Paulo tinha feito cópia dos contratos que foram enviados para o governo.

" Only Paulo had made a copy of the contracts that were sent to the government. "

b. i. O Paulo pegou as cópias dos documentos?

" Has Paulo taken the copies of the documents? "

ii. O que o Paulo fez com as cópias dos documentos?

" What has Paulo done with the copies of the documents? "

c. (Felizmente) (ele) pegou.

" Yes, (he has)./He's taken them./Yes, (he has) fortunately (done so)./He's fortunately taken them. "

(372) a. Só o Mário revisou os textos que a Maria escreveu.

"Only Mario reviewed the texts that Maria's written."

b. i. O Mário jogou os contos da Maria fora?

"Has Mario thrown Maria's shortstories away?"

ii. O que o Mário fez com os contos da Maria?

"What has Mario done with Maria's shortstories?"

c. (Francamente) (ele) jogou.

"Yes, (he has)./He's thrown them away./Yes, (he has) frankly (done so)./He's frankly thrown them away.
"

(373) a. Só o Jânio foi selecionado para a vaga na empresa que exige dedicação exclusiva.

"Only Jânio was selected for the job in the company that requires exclusive dedication."

b. i. O Jânio largou os outros empregos?

"Has Jânio quitted the other jobs?"

ii. O que o Jânio fez com os outros empregos?

"What has Jânio done with the other jobs?"

c. (Na realidade), (ele) largou.

"Yes, (he has)./He's quitted them./Yes, (he has) in fact (done so)./He's in fact quitted them."

(374) a. Só o Henrique foi ao bistrô renomado que abriu recentemente.

"Only Henrique went to the renowned bistro that has recently opened."

b. i. O Henrique rejeitou as sobremesas do restaurante ?

"Did Henrique refuse the restaurant's desserts?"

ii. O que o Henrique fez com as sobremesas do restaurante?

"What did Henry do to the restaurant's desserts?"

c. (Na real) (ele) rejeitou.

"Yes, (he did)./He threw them away./Yes, (he) actually (did so)./He actually refused them."

(375) a. Só o Luís trabalhou com as pastas que tinham os dados da pesquisa.

"Only Luís worked with the folders containing the survey data files."

b. i. O Luís arquivou a pasta no armário?

"Has Luis filed the folders in the closet?"

ii. O que o Luís fez com as pastas?

"What has Luís done to the folders?"

c. (De fato) (ele) arquivou.

" Yes, (he has)./He's filed them./Yes, (he has) indeed (done so)./He's indeed archived them. "

(376) a. Só o Carlos entrou no salão da festa de Ano Novo.

" Only Carlos entered the New Year's eve party hall. "

b.   i. O Carlos retirou as cadeiras de lá?

     " Has Carlos taken the chairs way from there? "

  ii. O que o Carlos fez com as cadeiras de lá?

     " What has Carlos done with the chairs from there? "

c. (Realmente) (ele) retirou.

" Yes, (he has)./He's taken them away (from there)./Yes, (he has) actually (done so)./He's actually taken them away (from there). "

(377) a. Só o Hugo estava cozinhando com as panelas do restaurante ontem.

" Only Hugo was cooking with the retaurant pots yesterday. "

b.   i. O Hugo lavou as panelas sujas?

     " Has Hugo washed the dirty pots? "

  ii. O que o Hugo fez com as panelas sujas?

     " What has Hugo done to the dirty pots? "

c. (Certamente) (ele) lavou.

" Yes, (he has)./He's washed them./Yes, (he has) certainly (done so)./He's certainly washed them. "

(378) a. Só a Maria tem acesso aos arquivos de alta segurança da empresa.

" Only Maria has access to high-security files of the company. "

b.   i. A Maria apagou os arquivos durante a reunião?

     " Has Maria deleted the files during the meeting? "

  ii. O que a Maria fez com esses arquivos?

     " What has Mary done with these files? "

c. (Supreendentemente) (ela) apagou.

" Yes, (she has)./She's deleted them./Yes, (she has) surprisingly (done so)./She's surprisingly deleted them. "

(379) a. Só a Helena viu as contas que estavam em cima da mesa.

" Only Helena's taken a look at the accounts that were on the table. "

b.   i. A Helena pagou as contas de luz e de água?

" Has Helena paid the electric and water bill? "

    ii. O que a Helena fez com as contas de luz e de água?

        " What has Helen done to the electric and water bill? "

  c. (Na verdade) (ela) pagou.

    " Yes, (she has)./She's paid them./Yes, (she has) in truth (done so)./She's in truth paid them. "

(380) a. Só a Júlia tinha ficado até mais tarde para terminar a decoração da escola.

    " Only Júlia had stayed late to finish school decoration. "

  b. i. A Júlia fixou os cartazes do festival?

        " Has Júlia set the festival posters? "

    ii. O que a Júlia fez com os cartazes do festival?

        " What has Júlia done with the festival posters? "

  c. (Com certeza) (ela) fixou.

    " Yes, (she has)./She's set them./Yes, (she has) for sure (done so)./She's for sure set them. "

(381) a. Só a Bruna trabalhou até mais tarde na realização do balanço da firma.

    " Only Bruna worked late to make the balance sheet of the firm. "

  b. i. A Bruna terminou aqueles relatórios?

        " Has Bruna finished those reports? "

    ii. O que a Bruna fez com aqueles relatórios?

        " What has Bruna done to those reports? "

  c. (Efetivamente) (ela) terminou.

    " Yes, (she has)./She's finished them./Yes, (she has) effectively (done so)./She's effectively finished them. "

(382) a. Só a Núbia organizou os trajes da peça de teatro ontem à tarde.

    " Only Nubia organized the theater play costumes last afternoon. "

  b. i. A Núbia escondeu as roupas que estavam manchadas?

        " Has Nubia hidden the clothes that were stained? "

    ii. O que a Núbia fez com as roupas que estavam manchadas?

        " What has Nubia done to the clothes that were stained? "

  c. (Claramente) (ela) escondeu.

    " She hid./Yes, (she has)./She's hidden them./Yes, (she has) clearly (done so)./She's clearly hidden them. '

(383) a. Só a Vânia permaneceu no tribunal onde estava o material do processo.

   " Only Vânia remained in the court in which the material of the case was. "

   b. i. A Vânia eliminou as gravações suspeitas?

      " Has Vânia eliminated suspicions recordings? "

      ii. O que a Vânia fez com as gravações suspeitas?

      " What has Vânia done to the suspicious recordings? "

   c. (Obviamente) (ela) eliminou.

   " Yes, (she has)./She's eliminated them. /Yes, (she has) obviously (done so)./She's obviously eliminated
   them. "

(384) a. Só a Tânia recebeu os presentes que foram adquiridos por seu marido no tour pela Europa.

   " Only Tânia's received the gifts that were purchased by her husband on the tour in Europe. "

   b. i. A Tânia destruiu as estátuas compradas na viagem?

      " Has Tânia destroyed the statues bought on the trip? "

      ii. O que o Tânia fez com as estátuas compradas na viagem?

      " What has Tânia done to the statues bought on the trip? "

   c. (Lamentavelmente) (ela) destruiu.

   " Yes, (she has)./She's destroyed them./Yes, (she has) unfortunately (done so)./She's unfortunately de-
   stroyed them. "

(385) a. Só a Luísa estava autorizada a mexer no estoque onde estavam os abajures da nova coleção.

   " Only Luiza had permission to handle the stock in which the lamps of the new collection were. "

   b. i. A Luísa vendeu aqueles com múltiplas cores?

      " Has Luísa sold those with multiple colors? "

      ii. O que a Luísa fez com aqueles com múltiplas cores?

      " What has Luísa done to those with multiple colors? "

   c. (Evidentemente) (ela) vendeu.

   " Yes, (she has)./She's sold them./Yes, (she has) evidently (done so)./She's evidently sold them. "

(386) a. Só a Fernanda elaborou as sequências de pratos servidas no restaurante chique da cidade.

   " Only Fernanda's created the sequences of dishes served at the fancy restaurant in the town. "

   b. i. A Fernanda alterou aqueles com ervas aromáticas?

      " Has Fernanda changed those with herbs? "

      ii. O que a Fernanda fez com aqueles com ervas aromáticas?

401

“ What has Fernanda done to those with those herbs? ”

c. (Estranhamente) (el)a alterou.

“ Yes, (she has)./She's changed them./Yes, (she has) strangely (done so)./She's strangely sold them. ”

(387) a. Só a Sílvia podia entrar na sala onde os livros foram depositados.

“ Only Sílvia could enter the room where the books were deposited. ”

b. i. A Sílvia organizou os livros da enciclopédia?

“ Has Sílvia organized the encyclopedic books? ”

ii. O que a Sílvia fez com os livros da enciclopédia?

“ What has Sílvia done to the encyclopedic books? ”

c. (Incrivelmente) (ela) organizou.

“ Yes, (she has)./She's organized them./Yes, (she has) incredibly (done so)./She's incredibly organized them. ”

Interpretation Task

(388) Alguém mais além do João comeu os sanduíches com maionese?

“ Has anyone else besides João eaten the sandwiches with mayonnaise? ”

(389) Alguém mais além do Felipe guardou as bolas de basquete?

“ Has anyone else besides Felipe stored the basketballs? ”

(390) Alguém mais além do João arrumou os livros?

“ Has anyone else besides João organized the books? ”

(391) Alguém mais além do Paulo pegou os documentos ?

“ Has anyone else besides Paulo taken the documents? ”

(392) Alguém mais além do Mário odiou os romances da Maria?

“ Has anyone else besides Mario thrown Maria's shortstories away? ”

(393) Alguém mais além do Jânio largou os empregos?

“ Has anyone else besides John quitted the jobs? ”

(394) Alguém mais além do Henrique rejeitou as sobremesas do restaurante?

“ Did anyone else besides Henrique refuse the restaurant's desserts? ”

(395) Alguém mais além do Luís arquivou a pasta?

“ Has anyone else besides Luís filed the folders? ”

(396) Alguém mais além do Carlos retirou as cadeiras do salão?

"Has anyone else besides Carlos taken the chairs from the hall?"

(397) Alguém mais além do Hugo lavou as panelas sujas?

"Has anyone else besides Hugo washed the dirty pots?"

(398) Alguém mais além da Maria apagou os arquivos da empresa?

"Has anyone else besides Maria deleted company files?"

(399) Alguém mais além da Helena pagou as contas de luz e de água?

"Has anyone else besides Helena paid the electric and water bill?"

(400) Alguém mais além do Julia fixou os cartazes do festival?

"Has anyone else besides Julia set the festival posters?"

(401) Alguém mais além da Bruna terminou os relatórios?

"Has anyone else besides Bruna finished the reports?"

(402) Alguém mais além da Núbia escondeu as roupas manchadas?

"Has anyone else besides Nubia hidden the stained clothes?"

(403) Alguém mais além da Vânia eliminou as gravações suspeitas?

"Has anyone else besides Vânia eliminated suspicions recordings?"

(404) Alguém mais além da Tânia destruiu as estátuas compradas na viagem?

"Has anyone else besides Tânia destroyed the statues bought on the trip?"

(405) Alguém mais além da Luísa vendeu os abajures com múltiplas cores?

"Has anyone else besides Luísa sold the lamps with multiple colors?"

(406) Alguém mais além da Fernanda alterou os pratos com ervas aromáticas?

"Has anyone else besides Fernanda changed the dishes with herbs?"

(407) Alguém mais além da Sílvia organizou os livros da enciclopédia?

"Has anyone else besides Sílvia organized the encyclopedic books?"

A.0.6   Experiment 5

Judgment Task

(408)   a. O João e a Maria foram embora cedo. Você sabe o que o João disse para a Maria?

"João and Maria went away early. Do you know what João told to Maria?"

b.　i.　O João disse que no teste de motorista _ não pode ir cansado/a.

　　　" João told that one could not go to the driving test tired. "

　　ii.　O João disse que _ não pode ir cansado/a no teste de motorista.

　　　" João told that he could not go to the driving test tired. "

(409)　a.　O Pedro e a Inês chegaram ao hospital para uma visita ao seu neto recém-nascido. Você sabe o que Pedro falou para a Inês?

　　　" Pedro and Inês got to the hospital in order to visit their newborn grandchild. Do you know what Pedro said to Inês? "

　　b.　i.　O Pedro falou que na sala da maternidade _ tem que ser cuidadoso/a.

　　　" Pedro said that in the maternity room one had to be careful. "

　　ii.　O Pedro falou que _ tem que ser cuidadoso/a na sala da maternidade.

　　　" Pedro said that he had to be careful in the maternity room. "

(410)　a.　O Pedro e a Ana estavam escolhendo onde iam comprar os trajes pro casamento da Maria. Você entendeu o que o Pedro disse para a Ana?

　　　" Pedro and Ana were choosing where they would buy wedding party costumes to go to Maria's marriage. Did you understand what Pedro told to Ana? "

　　b.　i.　O Pedro disse que para essa ocasião _ precisa sair bem vestido/a.

　　　" Pedro told that for this occasion one needed to go out dressed up. "

　　ii.　O Pedro disse que _ precisa sair bem vestido/a para essa ocasião.

　　　" Pedro told that he needed to go out dressed up for this occasion. "

(411)　a.　O Marcelo e a Caroline vieram aqui ontem com os filhos que eles tiveram recentemente. Você sabe o que o Marcelo falou para a Caroline?

　　　" Marcelo and Caroline came here yesterday with their recently-born children. Do you know what Marcelo said to Caroline ? "

　　b.　i.　O Marcelo falou que com filho pequeno _ tem que ser atencioso/a.

　　　" Marcelo said that with recently-born children one had to be attentive. "

　　ii.　O Marcelo falou que _ tem que ser atencioso/a com filho pequeno.

　　　" Marcelo said that he had to be attentive with recently-born children. "

(412)　a.　O Henrique e a Cláudia estavam conversando sobre as aulas de música. Você sabe o que o Henrique falou para a Cláudia?

　　　" Henrique and Claudia were talking about the music classes. Do you know what Henrique said to Claudia? "

b. i. O Henrique falou que na aula de piano _ não pode ir desatento/a.

" Henrique said that to the piano lessons one could not go inattentively. "

ii. O Henrique falou que _ não pode ir desatento/a na aula de piano.

" Henrique said that he could not go inattentively to the piano lessons. "

(413) a. O João e a Maria estavam conversando sobre o tempo que eles trabalharam juntos. Você sabe o que o João disse para a Maria?

" João and Maria were talking about the period they worked together. Do you know what João told to Maria? "

b. i. O João disse que naquela empresa _ podia trabalhar sentado/a o tempo todo.

" João told that in that company one could work seated down all work-day long. "

ii. O João disse que _ podia trabalhar sentado/a o tempo todo naquela empresa.

" João told that he could work seated down all work-day long in that company. "

(414) a. O Cláudio e a Denise estavam falando sobre a escola onde eles estudaram. Você sabe o que o Cláudio disse para a Denise?

" Cláudio and Denise were talking about the school they attended to. Do you know what Cláudio told to Denise? "

b. i. O Cláudio disse que naquela escola _ tinha que ser estudioso/a.

" Cláudio told that in that school one had to be studious. "

ii. O Cláudio disse que _ tinha que ser estudioso/a naquela escola.

" Cláudio told that he had to be studious in that school. "

(415) a. O Edson e a Vânia relembraram o tempo que trabalharam na fábrica de artesanato. Você sabe o que o Edson disse para a Vânia?

" Edson and Vânia remembered of the period they worked in the craftwork factory. Do you know what Edson told to Vânia? "

b. i. O Edson disse que na fábrica de artesanato _ não podia ser desajeitado/a.

" Edson told that in the craftwork factory one could not be clumsy. "

ii. O Edson disse que _ não podia ser desajeitado/a na fábrica de artesanato.

" Edson told that he could not be clumsy in the craftwork factory. "

(416) a. O Pedro e a Isadora relembraram saudosos a educação de antigamente. Você sabe o que o Pedro falou para a Isadora?

" Pedro and Isadora remembered wistfully the erstwhile education. Do you know what Pedrosaid to Isadora? "

b. i. O Pedro falou que na época de criança _ não podia responder aos pais mal-educado/a.

" Pedro said that in his childhood one could not answer to the parents ill-bredly. "

ii. O Pedro falou que _ não podia responder aos pais mal-educado/a na época de criança.

" Pedro said that he could not answer to the parents ill-bredly in his childhood. "

(417) a. O Fernando e a Helena estavam trocando ideias sobre as receitas do restaurante. Você sabe o que o Fernando disse para a Helena?

" Fernando and Helena were exchanging words about the restaurant recipes. Do you know what Fernando told to Helena? "

b. i. O Fernando disse que no feijão _ pode misturar os temperos despreocupado/a.

" Fernando told that in the bean stew one could mix the spices carelessly. "

ii. O Fernando disse que _ pode misturar os temperos despreocupado/a no feijão.

" Fernando told that he could mix the spices carelessly in the bean stew. "

(418) a. O Gabriel e a Nicole discordaram no debate sobre os jornais disponibilizados na internet hoje em dia. Você sabe o que o Gabriel disse para Nicole?

" Gabriel and Nicole disagreed in the debate about the online available newspapers. Do you know what Gabriel told to Nicole? "

b. i. O Gabriel disse que nos dias atuais _ precisa estar bem-informado/a.

" Gabriel told that nowadays one needed to be well informed. "

ii. O Gabriel disse que _ presisa estar bem-informado/a nos dias atuais.

" Gabriel told that he needed to be well informed nowadays. "

(419) a. O Igor e a Fernanda estavam conversando sobre os problemas que os filhos causaram quando jogavam futebol no condomínio. Você sabe o que o Igor disse para a Fernanda?

" Igor and Fernanda were talking about the problems their children caused when they played soccer on the condominium. Do you know what Igor told to Fernanda? '

b. i. O Igor disse que nessas situações _ tinha que ser compreensivo/a.

" Igor told that in those situations one had to be comprehensive. "

ii. O Igor disse que _ tinha que ser compreensivo/a nessas situações.

" Igor told that he had to be comprehensive in those situations. "

(420) a. O Pedro e a Isadora estavam falando sobre a empresa de tecnologia da informação onde eles trabalham. Você sabe o que o Pedro disse para a Isadora ?

" Pedro and Isadora were talking about the IT Company where they worked. Do you know what Pedro told to Isadora? "

b.   i. O Pedro disse que nessa empresa _ pode ir para a reunião desarrumado/a.

         " Pedro told that in this company one could go to the meetings untidily. "

    ii. O Pedro disse que _ pode ir para a reunião desarrumado/a nessa empresa.

         " Pedro told that he could go to the meetings untidily in this company. "

(421)   a. O João e a Cláudia estavam discutindo como iriam viajar para um congresso no exterior. Você sabe o que o João disse para a Cláudia?

      " João and Cláudia were discussing how to travel for a congress abroad. Do you know what João told to Cláudia? "

  b.   i. O João disse que em certas companhias aéreas _ pode viajar mais reservado/a.

         " João told that in certain air companies one could travel more quietly. "

    ii. O João disse que _ pode viajar mais reservado/a em certas companhias aéreas.

         " João told that he could travel more quietly in certain air companies. "

(422)   a. O Pedro e a Joana queriam participar uma competição esportiva amadora na sexta, depois de uma semana de trabalho. Você sabe o que o Pedro disse para a Joana?

      " Pedro and Joana wanted to take part in a amateur athletic competition on Friday, after a workweek. Do you know what Pedro told to Joana? "

  b.   i. O Pedro disse que nessa competição _ pode jogar mesmo exausto/a.

         " Pedro told that in this competition one could play even exhausted. "

    ii. O Pedro disse que _ pode jogar mesmo exausto/a nessa competição.

         " Pedro told that he could play even exhausted in this competition. "

(423)   a. O Antônio e a Tânia falaram sobre as aulas que eles estão dando. Você sabe o que o Antônio disse para a Tânia?

      " Antônio and Tânia have talked about the classes they are given. Do you know what Antônio told to Tânia? '

  b.   i. O Antônio disse que em sala de aula _ não pode parecer inseguro/a.

         " Antônio told that in the classroom one couldn't seem insecure. "

    ii. O Antônio disse que _ não pode parecer inseguro/a em sala de aula.

         " Antônio told that he couldn't seem insecure in the classroom. "

(424)   a. O Gustavo e a Tânia dicutiram por causa dos filhos. Você sabe o que o Gustavo falou para a Tânia ?

      " Gustavo and Tânia have argued because of their children. Do you know what Gustavo said to Tânia? "

  b.   i. O Gustavo falou que no castigo dos filhos _ tem que ser justo/a.

         " Gustavo said that in the children's punishment one had to be fair. "

    ii. O Gustavo falou que _ tem que ser justo/a no castigo dos filhos.

       " Gustavo said that he had to be fair in the children's punishment. "

(425) a. O Fernando e a Isabela relembraram o tempo que tiveram aulas na Companhia de Dança do Teatro Municipal. Você sabe o que o Fernando disse para a Isabela?

      " Fernando and Isabela remembered the period when they had lessons in the Company of Dance of the Municipal Theater. Do you know what Fernando told to Isabela? "

    b. i. O Fernando disse que nessa companhia _ tinha que estar sempre atento/a.

       " Fernando told that in this company one had to be attentive. "

      ii. O Fernando disse que _ tinha que estar sempre atento/a nessa companhia.

       " Fernando told that he had to be attentive in this company. "

(426) a. O Henrique e a Gabriela falaram sobre o trabalho como professores de educação física e árbitros esportivos. Você sabe o que o Henrique falou para a Gabriela?

      " Henrique and Gabriela spoke about the work as PE teachers and as referees. Do you know what Henrique said to Gabriela? "

    b. i. O Henrique falou que nos jogos profissionais _ tem que ser rápido/a.

       " Henrique said that in the professional matches one had to be fast. "

      ii. O Henrique falou que _ tem que ser rápido/a nos jogos profissionais.

       " Henrique said that he had to be fast in the professional matches. "

(427) a. O Cristiano e a Eduarda debateram sobre ensaio da peça de teatro na última reunião. Você sabe o que o Cristiano disse para a Eduarda?

      " Cristiano and Eduarda debated about the rehearsal of the play in the last meeting. Do you know what Cristiano told to Eduarda? "

    b. i. O Cristiano disse que no palco _ não pode ser tímido/a.

       " Cristiano told that on the stage one could not be shy. "

      ii. O Cristiano disse que _ não pode ser tímido/a no palco

       " Cristiano told that he could not be shy on the stage. "

Interpretation Task

(428) Então, o João recomendou que a Maria não deveria ir no teste cansada?

    " So did João recommend that Maria should not go to the driving test tired? "

(429) Então, o Pedro recomendou que a Inês deveria ser cuidadosa na sala da maternidade?

    " So did Pedro recommend that Inês should be careful in the maternity room? "

(430)   Então, o Pedro recomendou que a Ana deveria se vestir bem para o casamento da Maria?

" So did Pedro recommend that Ana should dress up to go to Maria's marriage? "

(431)   Então, o Marcelo recomendou que a Caroline seja atenciosa com os filhos?

" So did Marcelo recommend that Caroline should be attentive with their children? "

(432)   Então, o Henrique recomendou que a Claudia não deveria frequentar a aula de piano desatenta?

" So did Henrique recommend that Claudia should not frequent the piano lessons inattentively? "

(433)   Então, o João falou que a Maria podia trabalhar sentada naquela empresa?

" So did João say that Maria could work seated down in that company? "

(434)   Então, o Cláudio disse que a Denise tinha que ser estudiosa naquela escola?

" So did Cláudio tell that Denise had to be studious in that school? "

(435)   Então, o Edson disse que a Vânia não deveria ser desajeitada na fábrica de artesanato?

" So did Edson tell that Vânia should not be clumsy in the craftwork factory? "

(436)   Então, o Pedro falou que a Isadora não podia responder mal-educada na época de criança?

" So, did Pedro say that Isadora could not answer ill-bredly in the childhood. "

(437)   Então, o Fernando disse que a Helena pode misturar os temperos despreocupada no feijão?

" So, did Fernando tell that Helena could mix the spices carelessly in the bean stew? "

(438)   Então, Gabriel recomendou que a Nicole deveria estar bem-informada nos dias atuais?

" So did Gabriel recommend that Nicole should be well informed nowadays? "

(439)   Então, o Igor falou que a Fernanda tinha que ser compreensiva com os filhos nessas situações.

" So did Igor say that Fernanda had to be comprehensive in those situations? "

(440)   Então, o Pedro disse que a Isadora pode ir para a reunião desarrumada nessa empresa?

" So, did Pedro tell that Isadora could go to the meetings untidily in this company? "

(441)   Então, o João disse que a Cláudia poderia viajar mais reservada em certas companhias aéreas?

" So, did João tell that Cláudia could travel more quietly in certain air companies? "

(442)   Então, o Pedro disse que a Joana pode jogar mesmo exausta nessa competição?

" So, did Pedro tell that Joana could play even exhausted in this competition? "

(443)   Então, o Antônio recomendou que a Tânia não parecesse insegura?

" So, did Antônio tell that Tânia couldn't seem insecure in the classroom? "

(444)   Então, o Gustavo recomendou que a Tânia deveria ser justa no castigo dos filhos?

" So, did Gustavo recommend that Tânia should to be fair in the children's punishment? "

(445)   Então, o Fernando disse que a Isabela deveria estar sempre atenta nessa companhia?

"So, did Fernando tell that Isabela should be attentive in this company?"

(446)   Então, o Henrique falou que a Gabriela deveria ser rápida nos jogos profissionais?

"So, did Henrique tell that Gabriela should be fast in the professional matches?"

(447)   Então, o Cristiano recomendou que a Eduarda não deveria ser tímida no palco?

"So, did Cristiano recommend that Eduarda should not be shy on the stage?"

Participants' Expected Answers in the Interpretation Task

part_id	part_class	insitu_fem	insitu_mas	fronted_fem	fronted_mas	Total
part_001	both_int	1	1	1	0.8	0.75
part_002	both_int	1	1	1	1	1
part_003	both_int	1	1	1	1	1
part_004	both_int	1	1	1	1	1
part_005	both_int	1	1	1	1	1
part_006	both_int	1	1	1	1	1
part_007	both_int	1	1	1	1	1
part_008	more_coref	1	1	1	0	0.75
part_009	both_int	1	1	0.8	1	0.7
part_010	both_int	1	1	1	1	1
part_011	more_coref	0.8	0.8	1	0.2	0.7
part_012	both_int	1	1	1	1	1
part_013	more_coref	0.8	1	0.6	0.2	0.65
part_014	more_gen	0.8	0.2	1	0.8	0.7
part_015	more_coref	0.8	1	0.4	0.2	0.6
part_016	more_coref	1	0.8	0.8	0.2	0.75
part_017	more_coref	0.8	1	1	0.2	0.8
part_018	more_gen	1	0.2	0.8	1	0.75
part_019	more_coref	0.8	0.8	1	0.2	0.65
part_020	both_int	1	1	1	1	1
part_021	both_int	0.8	1	1	0.8	0.65
part_022	both_int	0.8	1	1	0.8	0.65
part_023	both_int	0.8	0.8	0.8	0.8	0.7
part_024	more_gen	1	0.2	0.8	0.8	0.65
part_025	more_coref	0.6	1	0.6	0.2	0.65
part_026	more_coref	0.8	0.8	1	0.2	0.65
part_027	more_gen	1	0	1	1	0.75
part_028	both_int	1	1	1	1	1
part_029	both_int	1	1	1	1	1
part_030	both_int	1	1	1	1	1
part_031	both_int	1	1	1	1	1
part_032	both_int	1	1	1	1	1
part_033	both_int	0.8	1	1	1	0.8
part_034	more_coref	1	0.8	1	0.2	0.75
part_035	more_coref	1	1	0.8	0.2	0.8
part_036	both_int	1	1	1	1	1
part_037	more_coref	0.8	0.8	0.8	0.2	0.65
part_038	more_coref	0.8	0.8	1	0.2	0.7
part_039	both_int	1	1	1	1	1
part_040	more_coref	0.6	0.8	0.8	0.2	0.6

part_id	part_class	insitu_fem	insitu_mas	fronted_fem	fronted_mas	Total
part_041	both_int	0.8	1	1	1	0.8
part_042	both_int	1	1	1	0.8	0.8
part_043	both_int	1	1	1	1	1
part_044	more_coref	1	0.8	1	0	0.7
part_045	more_coref	0.6	0.8	0.6	0.2	0.6
part_046	more_gen	1	0.2	1	1	0.8
part_047	both_int	1	1	1	1	1
part_048	both_int	0.8	0.8	1	0.8	0.68
part_049	more_coref	1	1	1	0	0.75
part_050	more_coref	1	1	1	0	0.75
part_051	more_gen	1	0	1	0.8	0.7
part_052	more_coref	1	1	1	0.2	0.8
part_053	both_int	1	1	1	1	1
part_054	more_coref	0.8	0.8	0.8	0	0.55
part_055	more_coref	1	1	1	0.2	0.8
part_056	more_coref	0.8	0.8	1	0.2	0.7
part_057	more_gen	1	0	1	0.8	0.7
part_058	both_int	0.6	0.8	1	0.8	0.6
part_059	both_int	1	1	1	1	1
part_060	both_int	0.8	1	0.6	1	0.65
part_061	more_coref	1	1	1	0	0.75
part_062	more_coref	0.8	0.8	0.8	0.2	0.7
part_063	more_gen	1	0	0.8	0.8	0.65
part_064	more_coref	1	0.8	0.6	0.2	0.65
part_065	more_gen	1	0	1	1	0.82
part_066	more_coref	0.8	0.8	0.8	0.2	0.65
part_067	more_gen	1	0.2	1	0.8	0.7
part_068	both_int	1	1	1	1	1
part_069	more_coref	0.2	0.8	0.8	0.2	0.5
part_070	both_int	1	1	1	1	1
part_071	both_int	1	1	1	1	1
part_072	more_gen	1	0.2	1	1	0.8
part_073	more_gen	1	0	1	0.8	0.61
part_074	more_coref	1	0.8	1	0.2	0.7
part_075	both_int	1	0.8	0.4	0.8	0.65
part_076	both_int	1	0.8	1	1	0.85
part_077	more_coref	0.8	0.8	0.5	0	0.52
part_078	more_gen	1	0.2	1	1	0.8
part_079	more_coref	1	1	0.8	0.2	0.75
part_080	both_int	1	1	1	1	1

part_id	part_class	insitu_fem	insitu_mas	fronted_fem	fronted_mas	Total
part_081	both_int	1	1	1	1	1
part_082	more_coref	0.8	1	0.8	0.2	0.7
part_083	more_gen	1	0	1	1	0.75
part_084	both_int	0.8	1	0.8	0.8	0.65
part_085	both_int	0.8	1	0.4	0.8	0.55
part_086	both_int	1	1	1	1	1
part_087	both_int	1	1	1	1	1
part_088	more_coref	0.8	0.8	1	0.2	0.7
part_089	more_coref	0.4	1	0.4	0.2	0.5
part_090	more_coref	0.4	0.8	0.8	0	0.45
part_091	both_int	1	1	1	1	1
part_092	more_coref	0.8	1	0.4	0.2	0.65
part_093	more_coref	1	1	0.8	0	0.7
part_094	both_int	0.8	0.8	1	0.8	0.7
part_095	more_gen	1	0	0.8	0.8	0.65
part_096	more_gen	0.8	0	1	1	0.7
part_097	more_coref	1	1	1	0	0.75
part_098	more_coref	1	1	1	0.2	0.8
part_099	more_coref	0.6	0.8	0.6	0.2	0.5
part_100	both_int	1	0.8	0.6	1	0.7
part_101	both_int	1	1	1	1	1
part_102	both_int	0.8	0.8	0.4	0.8	0.6
part_103	both_int	1	1	1	1	1
part_104	both_int	0.8	0.8	1	0.8	0.7
part_105	more_gen	1	0	1	1	0.75
part_106	both_int	1	0.8	0.8	0.8	0.7
part_107	both_int	1	1	1	1	1
part_108	more_coref	0.8	0.8	0.6	0.2	0.65
part_109	more_gen	1	0	0.8	1	0.7
part_110	more_coref	1	0.8	1	0	0.7
part_111	both_int	1	1	1	1	1
part_112	both_int	1	0.8	0.8	0.8	0.7
part_113	both_int	1	1	1	1	1
part_114	more_gen	1	0	0.8	0.2	0.55
part_115	more_coref	1	1	1	0.2	0.8
part_116	both_int	1	1	1	1	1
part_117	both_int	0.8	0.8	0.6	0.8	0.55
part_118	both_int	1	1	1	0.8	0.75
part_119	more_coref	0.8	0.8	1	0.2	0.75
part_120	more_coref	0.8	0.8	1	0.2	0.7

part_id	part_class	insitu_fem	insitu_mas	fronted_fem	fronted_mas	Total
part_121	both_int	1	1	1	1	1
part_122	both_int	1	1	1	1	1
part_123	more_coref	0.4	0.8	0.6	0.2	0.5
part_124	more_coref	0.6	0.8	0.8	0.2	0.6
part_125	more_coref	1	0.8	0.8	0.2	0.75
part_126	both_int	1	1	1	0.8	0.75
part_127	more_coref	0.6	0.8	0.4	0	0.45
part_128	more_coref	0.8	0.8	0.4	0.2	0.55
part_129	both_int	0.8	0.8	1	0.8	0.65
part_130	more_coref	0.6	0.8	0.8	0	0.55
part_131	both_int	0.2	1	1	0.8	0.5
part_132	more_coref	0.6	0.8	0.6	0.2	0.5
part_133	both_int	1	1	1	1	1
part_134	more_coref	1	0.8	1	0.2	0.75
part_135	both_int	1	1	1	1	1
part_136	both_int	1	1	1	1	1
part_137	both_int	1	1	1	1	1
part_138	more_gen	1	0	1	1	0.75
part_139	more_coref	1	1	1	0	0.73
part_140	both_int	1	1	0.8	0.8	0.7
part_141	more_gen	1	0	1	0.8	0.7
part_142	more_gen	0.8	0	0.8	0.8	0.6
part_143	both_int	0.8	1	1	1	0.8
part_144	more_gen	1	0	1	0.8	0.7
Total		0.89	0.80	0.88	0.82	0.75

Table A.1 – Percentage of Expected Answers and Classification of Participants in Experiment 5

A.0.7 Unacceptable Control Sentences in Experiments 1, 2, 3, 4 and 5

(448) a. O Manoel sempre teve muito medo de passeios de barco. Você sabe quando ele perdeu o medo?

" Manoel has always been very afraid of boat trips. Do you know when he has lost his fear? "

b. Eu encorajei o Manoel quando eu recolhemos as amarras do navio.

" I encouraged Manoel when I picked up the moorings of the ship. "

(449) a. O Mário e a Juliana estavam conversando sobre as festas que eles frequentavam. Você sabe o que o Mario disse para a Juliana?

" Mário and Juliana were talking about the parties they attended. Do you know what Mario said to Juliana? "

b. O Mario queria ela ir na festa de carnaval.

" Mario wanted her to go to the carnival party."

(450) a. A Eduarda tava com muitas dificuldades de adaptação na sua mudança para a Europa. Você sabe o que aconteceu com ela lá?

“ Eduarda had many difficulties adapting to her move to Europe. Do you know what happened to her there? ”

b. Incrivelmente ela vai rachar por causa do frio.

“ Unbelievably she will crack because of the cold.”

(451)   a. As festas na casa do Cézar sempre têm umas sobremesas maravilhosas. O que você achou do que ele ofereceu?

“ The parties at Cézar's house always have wonderful desserts. What did you think of what he offered? ”

b. Isso mágico foi realmente.

“ This magical thing really was.”

Appendix B

Experimental Materials – On-line Experiments

B.0.1 Experiment 6

Eye-tracking while Reading Task

(452) a. O João disse que a Natália toma cerveja artesanal frequentemente, mas nessa festa (ela) não pode beber, ele falou.

"João said that Natalia often drinks craft beer, but at this party one/she cannot drink, he said."

b. O João disse que a Natália toma cerveja artesanal frequentemente, mas nessa festa (ela) não tinha tomado, ele falou.

"João said that Natalia often drinks craft beer but at this party she had not drunk it, he said."

c. O João disse que a Natália toma cerveja artesanal frequentemente, mas (ela) não pode beber nessa festa, ele falou.

"João said that Natalia often drinks beer but she can not drink at this party, he said."

d. O João disse que a Natália toma cerveja artesanal frequentemente, mas (ela) não tinha tomado nessa festa, ele falou.

"João said that Natalia often drinks beer but she had not drunk it at this party, he said."

(453) a. O Pedro disse que a Helena não tinha de negociar os livros rapidamente, mas nessa loja (ela) tinha de vender, ele falou.

"Pedro said that Helena did not have to trade books quickly but in that store one/she had to sell [much], he said."

b. O Pedro disse que a Helena não tinha de negociar os livros rapidamente, mas nessa loja (ela) tinha negociado, ele falou.

" Pedro said that Helena did not have to trade books quickly but in that store she had traded them, he said. "

c. O Pedro disse que a Helena não tinha de negociar os livros rapidamente, mas (ela) tinha de vender nessa loja, ele falou.

" Pedro said Helena did not have to trade books quickly but she had to sell [much] at this store, he said. "

d. O Pedro disse que a Helena não tinha de negociar os livros rapidamente, mas (ela) tinha negociado nessa loja, ele falou.

" Pedro said that Helena did not have to trade books quickly but she had traded them at that store, he said. "

(454) a. A Ana falou que o Pedrinho não pode redigir textos no computador, mas nessa escola (ele) pode escrever, ela disse.

" Ana said that Pedrinho cannot write texts on the computer, but at this school one/he can write [much], she said. "

b. A Ana falou que o Pedrinho não pode redigir textos no computador, mas nessa escola (ele) tinha redigido, ela disse.

" Ana said that Pedrinho cannot write texts on the computer but in that school he had written them, she said. "

c. A Ana falou que o Pedrinho não pode redigir textos no computador, mas (ele) pode escrever nessa escola, ela disse.

" Ana said that Pedrinho cannot write texts on the computer but he can write [much] in this school, she said. "

d. A Ana falou que o Pedrinho não pode redigir textos no computador, mas (ele) tinha redigido nessa escola, ela disse.

" Ana said that Pedrinho cannot write texts on the computer but he had written them in that school, she said. "

(455) a. O Leandro disse que a Mônica não compreendia textos complexos facilmente, mas no jardim de infância (ela) tinha de ler, ele falou.

" Leandro said that Monica did not understand complex texts but in kindergarten one/she had to read, he said. "

b. O Leandro disse que a Mônica não compreendia textos complexos facilmente, mas no jardim de infância (ela) compreendeu, ele falou.

" Leandro said that Monica did not understand complex texts but in kindergarten she understood them, he said. "

c. O Leandro disse que a Mônica não compreendia textos complexos facilmente, mas (ela) tinha de ler no jardim de infância, ele falou.

" Leandro said that Monica did not understand complex texts but she had to read in kindergarten, he said. "

d. O Leandro disse que a Mônica não compreendia textos complexos facilmente, mas (ela) compreendeu no jardim de infância, ele falou.

" Leandro said that Monica did not understand complex texts but she understood them in kindergarten, he said. "

(456) a. O Douglas disse que a Bruna pode questionar problemas na escola, mas na aula de física (ela) não pode discutir, ele falou.

" Douglas said that Bruna can question problems at school but in physics class one/she cannot argue, he said. "

b. O Douglas disse que a Bruna pode questionar problemas na escola, mas na aula de física (ela) não questionou, ele falou.

" Douglas said that Bruna can question problems at school but in physics she did not question them, he said. "

c. O Douglas disse que a Bruna pode questionar problemas na escola, mas (ela) não pode discutir na aula de física, ele falou.

" Douglas said that Bruna can question problems at school but she cannot argue in physics class, he said. "

d. O Douglas disse que a Bruna pode questionar problemas na escola, mas (ela) não questionou na aula de física, ele falou.

" Douglas said that Bruna could question problems at school but she did not question them in the physics class, he said. "

(457) a. A Diana falou que o André deve defender as teorias com fervor, mas em aula (ele) não pode polemizar, ela disse.

" Diana said that André should defend the classical theories with fervor but in class one/he cannot polemize, she said. "

b. A Diana falou que o André deve defender as teorias com fervor, mas em aula (ele) não tinha defendido, ela disse.

" Diana said that Andre should defend the classical theories with fervor but in class he had not defended them, she said. "

c. A Diana falou que o André deve defender as teorias com fervor, mas (ele) não pode polemizar em aula, ela disse.

" Diana said that André should defend the classical theories with fervor but he cannot polemize in class, she said. "

d. A Diana falou que o André deve defender as teorias com fervor, mas (ele) não tinha defendido em aula, ela disse.

" Diana said that André should defend the classical theories with fervor but he had not defended them in class, she said. "

(458) a. A Marcela falou que o Mateus pode botar dinheiro em casinos, mas na Flórida (ele) não pode apostar, ela disse.

" Marcela said that Matthew can use money at casinos but in Florida one/he cannot bet, she said. "

b. A Marcela falou que o Mateus pode botar dinheiro em casinos, mas na Flórida (ele) não tinha botado, ela disse.

" Marcela said that Matthew could use money at casinos but in Florida he had not used it, she said. "

c. A Marcela falou que o Mateus pode botar dinheiro em casinos, mas (ele) não pode apostar na Flórida, ela disse.

" Marcela said that Matthew can use money at casinos but he can not bet in Florida, she said. "

d. A Marcela falou que o Mateus pode botar dinheiro em casinos, mas (ele) não tinha botado na Flórida, ela disse.

" Marcela said Matthew could use money at casinos but she had not used it in Florida, she said. "

(459) a. O Michel disse que a Rebeca não joga futebol no ginásio diariamente, mas na academia (ela) tem de treinar, ele falou.

" Michel said that Rebeca does not play soccer in the gym every day but at the gym one/she has to work out, he said. "

b. O Michel disse que a Rebeca não joga futebol no ginásio diariamente, mas na academia (ela) tinha jogado, ele falou.

" Michel said that Rebeca does not play soccer in the gym every day but at the gym she had played, he said. "

c. O Michel disse que a Rebeca não joga futebol no ginásio diariamente, mas (ela) tem de treinar na academia, ele falou.

" Michel said that Rebeca does not play soccer in the gym every day but she has to work out in the gym, he said. "

d. O Michel disse que a Rebeca não joga futebol no ginásio diariamente, mas (ela) tinha jogado na academia, ele falou.

" Michel said that Rebecca does not play soccer in the gym every day but she had played it in the gym, he said. "

(460) a. O Luís disse que a Viviane pode receber muito dinheiro certamente, mas nessa profissão (ela) não ganha, ele falou.

" Luis said that Viviane can make a lot of money but in this profession one/she does not win, he said. "

b. O Luís disse que a Viviane pode receber muito dinheiro certamente, mas nessa profissão (ela) não recebeu, ele falou.

" Luis said that Viviane can make a lot of money but in this profession she didn't receive it, he said. "

c. O Luís disse que a Viviane pode receber muito dinheiro certamente, mas (ela) não ganha nessa profissão, ele falou.

" Luis said that Viviane can make a lot of money but she does not win in this profession, he said. "

d. O Luís disse que a Viviane pode receber muito dinheiro certamente, mas (ela) não recebeu nessa profissão, ele falou.

" Luis said that Viviane can make a lot of money but in this profession she didn't receive it, he said. "

(461) a. A Cláudia falou que o Vagner recebe muito dinheiro pela lavagem automotiva, mas nesse estacionamento (ele) não pode cobrar, ela disse.

" Claudia said that Vagner receives a lot of money for the car wash but in that parking lot one/he cannot charge, she said. "

b. A Cláudia falou que o Vagner recebe muito dinheiro pela lavagem automotiva, mas nesse estacionamento (ele) não tinha recebido, ela disse.

" Claudia said that Vagner receives a lot of money for the car wash but in that parking lot he had not received it, she said. "

c. A Cláudia falou que o Vagner recebe muito dinheiro pela lavagem automotiva, mas (ele) não pode cobrar nesse estacionamento, ela disse.

" Claudio said that Vagner receives a lot of money for the car wash but he cannot charge in that parking lot, she said. "

d. A Cláudia falou que o Vagner recebe muito dinheiro pela lavagem automotiva, mas (ele) não tinha recebido nesse estacionamento, ela disse.

" Claudia said that Vagner receives a lot of money for the car wash but he had not received it in that parking lot, she said. "

(462) a. A Luísa disse que o Paulo não explica cálculos complexos nos cursos, mas nessa escola (ele) tem de ensinar bem, ela falou.

" Luisa said that Paulo does not explain complex calculations in courses, but in this school one/he has to teach well, she said. "

b. A Luísa disse que o Paulo não explica cálculos complexos nos cursos, mas nessa escola (ele) tinha explicado bem, ela falou.

" Luisa said that Paulo did not explain complex calculations in the courses, but at this school he had explained them well, she said. "

c. A Luísa disse que o Paulo não explica cálculos complexos nos cursos, mas (ele) tem de ensinar bem nessa escola, ela falou.

" Luisa said that Paulo does not explain complex calculations in the courses, but he has to teach well in this school, she said. "

d. A Luísa disse que o Paulo não explica cálculos complexos nos cursos, mas (ele) tinha explicado bem nessa escola, ela falou.

" Luisa said that Paulo does not explain complex calculations in the courses, but (he) had explained well in this school, she said. "

(463) a. O Manoel disse que a Fabíola sabe manobrar moto de qualquer jeito, mas na frente do portão (ela) não pode estacionar, ele disse.

" Manoel said Fabíola knows how to maneuver a motorcycle anyway but in front of the gate one/she cannot park, he said. "

b. O Manoel disse que a Fabíola sabe manobrar moto de qualquer jeito, mas na frente do portão (ela) não tinha manobrado, ele disse.

" Manoel said Fabíola knows how to maneuver a motorcycle anyway but in front of the gate she had not maneuvered, he said. "

c. O Manoel disse que a Fabíola sabe manobrar moto de qualquer jeito, mas (ela) não pode estacionar na frente do portão, ele disse.

" Manoel said Fabíola knows how to maneuver a motorcycle anyway but she can not park in front of the gate, he said. "

d. O Manoel disse que a Fabíola sabe manobrar moto de qualquer jeito, mas (ela) não tinha manobrado*na frente do portão, ele disse.

" Manoel said that Fabíola knows how to maneuver a motorcycle anyway but she had not maneuvered in front of the gate, he said. "

(464) a. A Ivana falou que o Jorge quis disputar uma partida de vôlei no verão, mas em má condição física (ele) não pode jogar, ela falou.

421

" Ivana said that Jorge wanted to dispute a volleyball match but in poor physical condition one/he cannot play, she said. "

b. A Ivana falou que o Jorge quis disputar uma partida de vôlei no verão, mas em má condição física (ele) não tinha disputado, ela falou.

" Ivana said that Jorge wanted to dispute a volleyball game but in poor physical condition he had not disputed, she said. "

c. A Ivana falou que o Jorge quis disputar uma partida de vôlei no verão, mas (ele) não pode jogar em má condição física, ela falou.

" Ivana said that Jorge wanted to dispute a volleyball game but he can not play in poor physical condition, she said. "

d. A Ivana falou que o Jorge quis disputar uma partida de vôlei no verão, mas (ele) não tinha disputado em má condição física, ela falou.

" Ivana said that Jorge wanted to dispute a volleyball game but he had not played in poor physical condition, she said. "

(465) a. A Fernanda falou que o Gabriel pode cobrir paredes com verniz, mas na chuva (ele) não pode pincelar, ela falou.

" Fernanda said that Gabriel can cover the walls with varnish but in the rain one/he cannot paint, she said. "

b. A Fernanda falou que o Gabriel pode cobrir paredes com verniz, mas na chuva (ele) não tinha coberto, ela falou.

" Fernanda said that Gabriel can cover the walls with varnish but in the rain it had not covered, she said. "

c. A Fernanda falou que o Gabriel pode cobrir paredes com verniz, mas (ele) não pode pincelar na chuva, ela falou.

" Fernanda said that Gabriel can cover the walls with varnish but he cannot paint in the rain, she said. "

d. A Fernanda falou que o Gabriel pode cobrir paredes com verniz, mas (ele) não tinha coberto na chuva, ela falou.

" Fernanda said that Gabriel can cover the walls with varnish but he had not varnished in the rain, she said. "

(466) a. O Fernando disse que a Gabriela percebe todos os detalhes em músicas, mas nessa brincadeira (ela) não pode escutar, ele disse.

" Fernando said that Gabriela notices all the details of the songs but in this game one/she can not hear, he said. "

b. O Fernando disse que a Gabriela percebe todos os detalhes em músicas, mas nessa brincadeira (ela) não tinha percebido, ele disse.

" Fernando said that Gabriela notices all the details of the songs but in this game one/she had not noticed them, he said. "

c. O Fernando disse que a Gabriela percebe todos os detalhes em músicas, mas (ela) não pode escutar nessa brincadeira, ele disse.

" Fernando said that Gabriela notices all the details of the songs but she can not hear in this game, he said. "

d. O Fernando disse que a Gabriela percebe todos os detalhes em músicas, mas (ela) não tinha percebido nessa brincadeira, ele disse.

" Fernando said that Gabriela notices all the details of the songs but she had not noticed them in this game, he said. "

(467) a. A Eduarda falou que o Luquinhas não sabe nem rabiscar casas com chaminé em geral, mas na pré-escola (ele) tem de desenhar, ela falou.

" Eduarda said that Luquinhas does not know how to scribble even houses with chimney but in the preschool one/he has to draw, she said. "

b. A Eduarda falou que o Luquinhas não sabe nem rabiscar casas com chaminé em geral, mas na pré-escola (ele) tinha rabiscado, ela falou.

" Eduarda said that Luquinhas does not know how to scribble even houses with chimney but in the preschool he had scribbled them, she said. "

c. A Eduarda falou que o Luquinhas não sabe nem rabiscar casas com chaminé em geral, mas (ele) tem de desenhar na pré-escola, ela falou.

" Eduarda said that Luquinhas does not know how to scribble even houses with chimneys but he has to draw in preschool, she said. "

d. A Eduarda falou que o Luquinhas não sabe nem rabiscar casas com chaminé em geral, mas (ele) tinha rabiscado na pré-escola, ela falou.

" Eduarda said that Luquinhas does not know how to scribble even houses with chimneys but he had scribbled them in preschool, she said. "

(468) a. O Eugênio disse que a Andreia não sabe retratar animais selvagens no papel, mas na escolinha (ela) tem de pintar, ele disse,

" The Eugenio said that Andreia does not know how to depict wild animals on a sheet but in the school one/she has to paint, he said. "

b. O Eugênio disse que a Andreia não sabe retratar animais selvagens no papel, mas na escolinha (ela) tinha retratado, ele disse,

" Eugenio said Andreia does not know how to depict wild animals on a sheet but in the school she had depicted them, he said. "

c. O Eugênio disse que a Andreia não sabe retratar animais selvagens no papel, mas (ela) tem de pintar na escolinha, ele disse.

" Eugenio said Andreia does not know how to depict wild animals on a sheet but she has to paint in the school, he said. "

d. O Eugênio disse que a Andreia não sabe retratar animais selvagens no papel, mas (ela) tinha retratado na escolinha, ele disse.

" Eugenio said that Andreia does not know how to depict wild animals on a sheet but she had depicted them in the school, he said. "

(469) a. O Evandro disse que a Malu não toca músicas populares normalmente, mas nesse concurso (ela) só podia cantar, ele disse.

" Evandro said that Malu does not play popular songs generally but in this contest one/she only could sing, he said. "

b. O Evandro disse que a Malu não toca músicas populares normalmente, mas nesse concurso (ela) só tinha tocado, ele disse.

" Evandro said that Malu does not play popular songs generally but in this contest she only had played them, he said. "

c. O Evandro disse que a Malu não toca músicas populares normalmente, mas (ela) só só podia cantar nesse concurso, ele disse.

" Evandro said that Malu does not play popular songs generally but she could sing in that contest, he said. "

d. O Evandro disse que a Malu não toca músicas populares normalmente, mas (ela) só tinha tocado nesse concurso, ele disse.

" Evandro said that Malu does not play popular songs generally but she only had played them in that contest, he said. "

(470) a. A Gisele falou que o Felipe utiliza tango em performances em geral, mas nesse show de talentos (ele) não pode dançar, ela falou.

" Gisele said that Felipe uses tango in general performances but in this show of talents one/he cannot dance, she said. "

b. A Gisele falou que o Felipe utiliza tango em performances em geral, mas nesse show de talentos (ele) não tinha utilizado, ela falou.

" Gisele said that Felipe uses tango in performances in general but in this talent show he had not used it, she said. "

c. A Gisele falou que o Felipe utiliza tango em performances em geral, mas (ele) não pode dançar nesse show de talentos, ela falou.

" Gisele said that Felipe uses tango in performances in general but he cannot dance in this talent show, she said. "

d. A Gisele falou que o Felipe utiliza tango em performances em geral, mas (ele) não tinha utilizado nesse show de talentos, ela falou.

" Gisele said that Felipe uses tango in performances in general but he had not used it in this talent show, she said. "

(471) a. O Cézar disse que a Rafaela não domina as exatas muito bem, mas para as provas (ela) tem de estudar, ele disse.

" The Cézar said that Rafaela does not handle the hard sciences very well but for the tests one/she has to study, he said. "

b. O Cézar disse que a Rafaela não domina as exatas muito bem, mas para as provas (ela) tinha dominado, ele disse.

" The Cézar said that Rafaela does not handle the hard sciences very well but for the tests she had handled them, he said. "

c. O Cézar disse que a Rafaela não domina as exatas muito bem, mas (ela) tem de estudar para as provas, ele disse.

" The Cézar said that Rafaela does not handle the hard sciences very well but she has to study for the tests, he said. "

d. O Cézar disse que a Rafaela não domina as exatas muito bem, mas (ela) tinha dominado para as provas, ele disse.

" Cézar said that Rafaela does not handle the hard sciences very well but she had handled them for the tests, he said. "

(472) a. A Paula falou que o Renam faz questões provocativamente, mas para o professor (ele) não pode responder, ela falou.

" Paula said that Renam provocatively makes questions but with the teacher one/he cannot argue, she said. "

b. A Paula falou que o Renam faz questões provocativamente, mas para o professor (ele) não tinha feito, ela falou.

" Paula said that Renam provocatively makes questions but for the teacher he had not made it, she said. "

c. A Paula falou que o Renam faz questões provocativamente, mas (ele) não pode responder para o professor, ela falou.

" Paula said that Renam provocatively makes questions but he cannot argue with the teacher, she said. "

d. A Paula falou que o Renam faz questões provocativamente, mas (ele) não tinha feito para o professor, ela falou.

" Paula said that Renam provocatively makes questions but he had not made them for the teacher, she said. "

(473) a. A Manuela falou que o Humberto retrata certos personagens muito bem, mas no show de improviso (ele) não pode representar, ela falou.

" Manuela said that Humberto depicts very well certain characters but in the impromptu show one/he cannot play, she said. "

b. A Manuela falou que o Humberto retrata certos personagens muito bem, mas no show de improviso (ele) não tinha retratado, ela falou.

" Manuela said that Humberto depicts very well certain characters but at the impromptu show he had not depicted them, she said. "

c. A Manuela falou que o Humberto retrata certos personagens muito bem, mas (ele) não pode representar no show de improviso, ela falou.

" Manuela said that Humberto depicts very well certain characters but he cannot play in the impromptu show, she said. "

d. A Manuela falou que o Humberto retrata certos personagens muito bem, mas (ele) não tinha retratado no show de improviso, ela falou.

" Manuela said that Humberto depicts very well certain characters but he had not depicted them in the impromptu show, she said. "

(474) a. O Flávio disse que a Luana pode notar vultos vendada, mas nesse jogo (ela) não pode ver, ele disse.

" Flavio said that Luana may notice figures blindfolded but in this game one/she cannot see, he said. "

b. O Flávio disse que a Luana pode notar vultos vendada, mas nesse jogo (ela) não tinha notado, ele disse.

" Flavio said that Luana may notice figures blindfolded but in that game she had not noticed them, he said. "

c. O Flávio disse que a Luana pode notar vultos vendada, mas (ela) não pode ver nesse jogo, ele disse.

" Flavio said that Luana may notice figures blindfolded but she cannot see in that game, he said. "

d. O Flávio disse que a Luana pode notar vultos vendada, mas (ela) não tinha notado nesse jogo, ele disse.

" Flavio said that Luana may notice figures blindfolded but she had not noticed them in that game, he said. "

(475) a. A Flávia falou que o Mário não sabe usar o italiano em público, mas nas palestras (ele) tem de falar bem, ela falou.

" Flavia said Mario does not know how to speak Italian in public but in the lectures he/one has to speak well, she said. "

b. A Flávia falou que o Mário não sabe usar o italiano em público, mas nas palestras (ele) tinha usado bem, ela falou.

" Flavia said Mario does not know how to speak Italian in public but in the lectures he had used it well, she said. "

c. A Flávia falou que o Mário não sabe usar o italiano em público, mas (ele) tem de falar bem nas palestras, ela falou.

" Flavia said Mario does not know how to speak Italian in public but he has to speak well in the talks, she said. "

d. A Flávia falou que o Mário não sabe usar o italiano em público, mas (ele) tinha usado bem nas palestras, ela falou.

" Flavia said that Mário does not speak Italian in public but he had used it well in the talks, she said. "

Question Task

(476) Então, a Rafaela tem de estudar as exatas para as provas?

" So, does Rafaela have to study the hard-sciences for the tests? "

(477) Então, é proibido beber qualquer coisa nessa festa?

" So, is it forbidden to drink anything at this party? "

(478) Então, é proibido fazer apostas na Flórida?

" So, is it forbidden to make bets in Florida? "

(479) Então, é proibido tomar qualquer coisa nessa festa?

" So, is it forbidden to take anything at this party? "

(480) Então, era necessário saber compreender textos complexos no jardim de infância?

" So was it necessary to know how to understand complex texts in kindergarten? "

(481) Então, era necessário saber ler no jardim de infância?

" So was it necessary to know how to read in kindergarten? "

(482) Então, eram permitidas somente perfomances de canto nesse concurso?

" So, were they allowed only singing performances in this contest? "

(483) Então, o Gabriel não pode pincelar paredes na chuva?

" So cannot Gabriel brush walls under the rain? "

(484) Então, o Humberto não pode representar certos personagens no show de improviso?

" So, cannot Humberto represent certain characters in the impromptu show? "

(485) Então, o Humberto não tinha retratado certos personagens no show de improviso?

"So, had not Humberto represented certain characters on the impromptu show?"

(486) Então, o Luquinhas tem de desenhar casas com chaminé na pré-escola?

"So, does Luquinhas have to draw houses with chimney in the preschool?"

(487) Então, o Luquinhas tinha rabiscado casas com chaminé na pré-escola?

"So, had Luquinhas scrawled chimney houses in preschool?"

(488) Então, o Paulo tem de ensinar cálculos complexos bem nessa escola?

"So, does Paul have to teach complex calculations well at this school?"

B.0.2   Experiment 7

Eye-tracking while Reading Task

(489) a. Todo cliente se deleita com um prato que (ele) pode comer à vontade.

"Every customer delights in a dish that one/they can eat in an 'all you can eat'."

b. Todo cliente se deleitou com um prato que (ele) tinha provado à vontade.

"Every customer delighted in a dish that they had tasted in an 'all you can eat'."

(490) a. Toda consumidora quer uma champagne que (ela) pode beber sem restrição.

"Every consumer wants a champagne that one/they can drink without restriction."

b. Toda consumidora quis uma champagne que (ela) tinha provado semrestrição.

"Every consumer wanted a champagne that they had tasted without restriction."

(491) a. Nenhuma aluna gosta de uma redação que (ela) pode escrever sem tema definido.

"No student likes an essay that one/they can write without a defined topic."

b. Nenhuma aluna gostou de uma redação que (ela) tinha redigido sem tema definido.

"No student likes an essay that they had written without a defined topic."

(492) a. Nenhum aluno adoraria um roteiro que (ele) tinha de ler em voz alta.

"No student would love a script that one/they had to read out loud."

b. Nenhum aluno adorou um roteiro que ele tinha interpretado em voz alta."

"No student loved a script they had interpreted out loud.

(493) a. Toda paciente ignora uma iluminação que (ela) não pode enxergar de longe.

"Every patient ignores an illumination that one/they cannot see from a distance."

b. Toda paciente ignorou uma iluminação que (ela) não tinha percebido de longe.

" Every patient ignored an illumination she had not perceived from a distance."

(494) a. Todo estudante receia um problema que (ele) não pode discutir em aula.

" Every student fears a problem that one/they cannot discuss in class. "

b. Todo estudante receou um problema que (ele) não tinha avaliado em aula.

" Every student feared a problem he had not assessed in class. "

(495) a. Toda estudante cisma com uma teoria que (ela) não pode polemizar em público.

" Every student wonders about a theory that one/they have to polemicize in public."

b. Toda estudante cismou com uma teoria que (ela) tinha defendido em público.

" Every student wondered about a theory they had espoused in public."

(496) a. Todo motorista detesta um caminhão que (ele) não pode dirigir na estrada.

" Every driver hates a truck that one/they cannot drive on the road."

b. Todo motorista detestou um caminhão que (ele) não tinha colocado na estrada.

" Every driver hated a truck they had not put on the road."

(497) a. Todo investidor aprova um ganho que (ele) pode apostar na bolsa.

" Every investor approves a gain that one/they can bet on the stock market."

b. Todo investidor aprovou um ganho que (ele) tinha utilizado na bolsa.

" Every investor approved a gain that they had used on the stock exchange.

(498) a. Toda jogadora simpatiza com uma prática que (ela) tem de treinar com frequência.

" Every player sympathizes with a practice one/they has to practice frequently.

b. Toda jogadora simpatizou com uma prática que (ela) tinha realizado com frequência.

" Every player sympathized with a practice they had performed frequently.

(499) a. Nenhuma bailarina prefere uma valsa que (ela) não pode dançar na competição.

" No ballerina prefers a waltz that one/they cannot dance in the competition."

b. Nenhuma bailarina preferiu uma valsa que (ela) tinha utilizado na competição.

" No ballerina preferred a waltz that they had used in competition."

(500) a. Nenhum lojista se aborrece com um cheque que (ele) não pode cobrar em antecipação.

" No shopkeeper gets annoyed with a check which one/they cannot charge in anticipation."

b. Nenhum lojista se aborreceu com um cheque que (ele) não tinha recebido em antecipação.

" No shopkeeper got annoyed by a check that they had not received in anticipation."

(501) a. Toda professora aprecia uma matéria que ele tem de ensinar bem para todos.

" Every teacher appreciates a subject that one/they has/have to teach well for everyone."

b. Toda professora apreciou uma matéria que ele tinha explicado bem para todos.

" Every teacher appreciated a subject they had explained well for everyone."

(502) a. Todo motorista se irrita com um veículo que (ele) não pode estacionar com facilidade.

" Every driver gets irritated with a vehicle that one/they cannot park with ease."

b. Todo motorista se irritou com um veículo que (ele) não tinha comprado com facilidade.

" Every driver was irritated by a vehicle that they had not bought with ease."

(503) a. Todo jogador usufrui de um esporte que (ele) pode jogar com frequência.

" Every player enjoys a sport that one/they can play frequently."

b. Todo jogador usufruiu de um esporte que (ele) tinha praticado com frequência.

" Every player enjoyed a sport that they had played frequently."

(504) a. Nenhuma estudante teme uma figura que (ela) tem de desenhar com regularidade.

" No student fears a figure that one/they has/have to draw regularly."

b. Nenhuma estudante temeu uma figura que (ela) tinha reproduzido com regularidade.

" No student feared a figure that the had scribbled regularly."

(505) a.

b. Nenhum colegial se interessa por um brasão que (ele) tem de pintar na escolinha.

" No schoolboy is interested in a coat of arms that one/they has to paint in his school."

c. Nenhum colegial se interessou por um brasão que ele tinha retratado na escolinha.

" No schoolboy is interested in a coat of arms that they had portrayed in the school."

(506) a. Nenhuma costureira seleciona uma camisa que ela tem de costurar na vertical.

" No seamstress selects a shirt that one/they has/have to sew upright."

b. Nenhuma costureira selecionou uma camisa que ela tinha cortado na vertical.

" No seamstress selected a shirt they had cut vertically."

(507) a. Nenhum artista reconhecia um bolero que (ele) podia cantar no festival.

" No artist knew a bolero that one/they could sing at the festival."

b. Nenhum artista reconheceu um bolero que (ele) tinha escolhido no festival.

" No artist knew a bolero that they had chosen in the festival."

(508) a. Nenhum universitário se surpreendeu com um tópico que (ele) tinha dominado com atenção.

" No college student was surprised by a topic that one/they had mastered with attention."

b. Nenhum universitário se surpreendeu com um tópico que (ele) tem de estudar com atenção.

" No college student was surprised by a topic that they have to study with attention."

(509) a. Nenhuma educadora odeia uma mal-criação que (ela) não pode responder de pronto.

" No educator hates an ill-breeding that one/they cannot respond immediately."

b. Nenhuma educadora odiou uma mal-criação que (ela) não tinha punido de pronto.

" No educator hated an ill-breeding that they had punished immediately."

(510) a. Toda candidata execra uma estampa que (ela) tem de colorir com destreza.

" Every candidate dislikes a stamp that one/they has/have to color with dexterity."

b. Toda candidata execrou uma estampa que (ela) tinha tingido com destreza.

" Every candidate disliked a stamp that they dyed with dexterity."

(511) a. Nenhum ator descarta um papel que (ele) não pode representar com naturalidade.

" No actor discards a role that one/they cannot represent naturally."

b. Nenhum ator descartou um papel que (ele) não tinha reproduzido com naturalidade.

" No actor discarded a role that they had not reproduced naturally."

(512) a. Nenhum paciente reconhece um aparelho que (ele) não pode ver no escuro.

" No patient recognizes a device that one/they cannot see in the dark."

b. Nenhum paciente reconheceu um aparelho que (ele) não tinha notado no escuro.

" No patient recognized a device that they had not noticed in the dark."

Question Task

(513) É a estampa que a candidata tem de colorir com destreza?

" Is the stamp that the candidate has to color with dexterity? "

(514) É a estampa que a candidata tinha tingido com destreza?

" Is the stamp that the candidate dyed with dexterity? "

(515) É a teoria que a estudante não pode polemizar em público?

" Is it the theory that the student can not quarrel in public? "

(516) É a teoria que a estudante tinha defendido em público?

" Is it the theory that the student had defended in public? "

(517) É o cliente que pode comer o prato à vontade?

" Is it the client who can eat the dish as much as he want? "

(518) É o cliente que tinha provado o prato à vontade?

"Is it the client who had tasted the dish as much as he want? " "

(519) É o colegial que tem de pintar um brasão na escolinha?

"Is the schoolboy who has to paint a coat of arms in the school? "

(520) É o colegial que tinha retratado um brasão na escolinha?

"Is the schoolboy who had a coat of arms on his school? "

(521) É o logista que não pode cobrar um cheque em antecipação?

"Is it the shopkeeper who can not cash a check in advance? "

(522) É o logista que não tinha recebido um cheque em antecipação?

"Is the shopkeeper who had not received a check in advance? "

(523) É um roteiro que o estudante tinha de ler em voz alta?

"Is it a script that the student had to read aloud? "

(524) É um roteiro que o estudante tinha interpretado em voz alta?

"Is it a script that the student had interpreted out loud? "

B.0.3 Experiment 8

Eye-tracking while Reading Task

(525) a. O nosso cliente apreciou muito o vinho especial do nosso restaurante. O que tinha tomado de graça ficou contente demais, disse o graçom do restaurante.

"Our customer greatly appreciated the especial wine from our restaurant. The one who had drunk it for free got too happy, said the restaurant waiter. "

b. O nosso cliente aprecia muito o vinho especial do nosso restaurante. O que pode beber de graça fica contente demais, disse o garçom do restaurante.

"Our customer greatly appreciates the especial wine from our restaurant. The one who can drink for free gets too happy, said the restaurant waiter. "

c. O vinho especial do nosso restaurante foi muito apreciado pelo nosso cliente. O que tinha tomado de graça foi logo desprezado, disse o garçom do restaurante.

"The especial wine from our restaurant was much appreciated by our client. What he had drunk for free was soon despised, said the restaurant waiter. "

d. O vinho especial do nosso restaurante é muito apreciado pelo nosso cliente. O que pode beber de graça é logo desprezado, disse o garçom do restaurante.

" The especial wine of our restaurant is much appreciated by our client. What one can drink for free is soon despised, said the restaurant waiter. "

e. O vinho especial do nosso restaurante foi muito apreciado pelo nosso cliente. O que tinha tomado de graça ficou contente demais, disse o garçom do restaurante.

" The especial wine from our restaurant was much appreciated by our client. The one who had drunk it for free got too happy, said the restaurant waiter. "

f. O vinho especial do nosso restaurante é muito apreciado pelo nosso cliente. O que pode beber de graça fica contente demais, disse o garçom do restaurante.

" The especial wine of our restaurant is much appreciated by our client. The one who can drink for free gets too happy, said the restaurant waiter. "

g. O nosso cliente apreciou muito o vinho especial do nosso restaurante. O que tinha tomado de graça foi logo desprezado, disse o garçom do restaurante.

" Our customer greatly appreciated the especial wine from our restaurant. What he had drunk for free was soon despised, said the restaurant waiter. "

h. O nosso cliente aprecia muito o vinho especial do nosso restaurante. O que pode beber de graça é logo desprezado, disse o garçom do restaurante.

" Our customer greatly appreciates the especial wine from our restaurant. What one can drink for free is soon despised, said the restaurant waiter. "

(526) a. O frequentador do nosso restaurante escolheu em geral o buffet sofisticado oferecido aos sábados. O que tinha provado sem restrição se sentiu prestigiado, disse o chefe de cozinha.

" Our restaurant patron generally chose the sophisticated buffet offered on Saturdays. The one who had tasted it unreservedly felt prestigious, said the head chef. "

b. O frequentador do nosso restaurante escolhe em geral o buffet sofisticado oferecido aos sábados. O que pode comer sem restrição se sente prestigiado, disse o chefe de cozinha.

" Our restaurant patron usually chooses the sophisticated buffet offered on Saturdays. The one who can eat without restriction feels prestigious, said the head chef. "

c. O buffet sofisticado oferecido aos sábados foi em geral escolhido pelo frequentador do nosso restaurante. O que tinha provado sem restrição não sobrou no buffet, disse o chefe de cozinha.

" The sophisticated buffet offered on Saturdays was generally chosen by our restaurant patron. What he had tasted without restriction was not left on the buffet, said the head chef. "

d. O buffet sofisticado oferecido aos sábados é em geral escolhido pelo frequentador do nosso restaurante. O que pode comer sem restrição não sobra no buffet, disse o chefe de cozinha.

" The sophisticated buffet offered on Saturdays is usually chosen by the patron of our restaurant. What one can eat without restriction is not left in the buffet, said the head chef. "

e. O buffet sofisticado oferecido aos sábados foi em geral escolhido pelo frequentador do nosso restaurante. O que tinha provado sem restrição se sentiu prestigiado, disse o chefe de cozinha.

" The sophisticated buffet offered on Saturdays was generally chosen by our restaurant patron. The one who had tasted it unreservedly felt prestigious, said the head chef. "

f. O buffet sofisticado oferecido aos sábados é em geral escolhido pelo frequentador do nosso restaurante. O que pode comer sem restrição se sente prestigiado, disse o chefe de cozinha.

" The sophisticated buffet offered on Saturdays is usually chosen by our restaurant patron. The one who can eat without restriction feels prestigious, said the head chef. "

g. O frequentador do nosso restaurante escolheu em geral o buffet sofisticado oferecido aos sábados. O que tinha provado sem restrição não sobrou no buffet, , disse o chefe de cozinha.

" Our restaurant-goer generally chose the sophisticated buffet offered on Saturdays. What he had tasted without restriction was not left on the buffet, said the head chef. "

h. O frequentador do nosso restaurante escolhe em geral o buffet sofisticado oferecido aos sábados. O que pode comer sem restrição não sobra no buffet, disse o chefe de cozinha.

" Our restaurant usually chooses the sophisticated buffet offered on Saturdays. What one can eat without restriction is not left in the buffet, said the head chef. "

(527) a. O vendedor da nossa loja ofereceu em primeiro lugar o livro mais procurado. O que tinha negociado com rapidez recebeu boa comissão, disse o chefe do departamento.

" The seller of our store first offered the most sought-after book. The one who had negotiated it quickly received a good commission, said the head of the department. "

b. O vendedor da nossa loja oferece em primeiro lugar o livro mais procurado. O que pode vender com rapidez recebe boa comissão, disse o chefe do departamento.

" The seller of our store first offers the most sought-after book. The one who can sell quickly receives good commission, said the head of the department. "

c. O livro mais procurado foi oferecido em primeiro lugar pelo vendedor da nossa loja. O que tinha negociado com rapidez esgotou no lançamento, disse o chefe do departamento.

" The most sought-after book was offered first by the seller of our store. What he had traded quickly sold out at the launch, said the department head. "

d. O livro mais procurado é oferecido em primeiro lugar pelo vendedor da nossa loja. O que pode vender com rapidez esgota no lançamento, disse o chefe do departamento.

" The most sought-after book is offered first by the seller of our store. What one can sell quickly sells out at launch, said the department head. "

e. O livro mais procurado foi oferecido em primeiro lugar pelo vendedor da nossa loja. The one who had negotiated it quickly received a good commission, said the head of the department.

" The most sought-after book was offered first by the seller of our store. The one who had negotiated it quickly received a good commission, said the head of the department. "

f. O livro mais procurado é oferecido em primeiro lugar pelo vendedor da nossa loja. O que pode vender com rapidez recebe boa comissão, disse o chefe do departamento.

" The most sought-after book is offered first by the seller of our store. The one who can sell quickly receives good commission, said the head of the department. "

g. O vendedor da nossa loja ofereceu em primeiro lugar o livro mais procurado. O que tinha negociado com rapidez esgotou no lançamento, disse o chefe do departamento.

" The seller of our store first offered the most sought-after book. What he had traded quickly sold out at the launch, said the department head. "

h. O vendedor da nossa loja oferece em primeiro lugar o livro mais procurado. O que pode vender com rapidez esgota no lançamento, disse o chefe do departamento.

" The seller of our store first offers the most sought-after book. What one can sell quickly sells out at launch, said the department head. "

(528) a. O aluno do nosso curso selecionou o projeto mais interessante. O que tinha redigido sobre o tema recebeu muitos prêmios, disse o professor da turma.

" The student of our course selected the most interesting project. The one who had written it on the subject received many awards, said the class teacher. "

b. O aluno do nosso curso seleciona o projeto mais interessante. O que pode escrever sobre o tema recebe muitos prêmios, disse o professor da turma.

" The student of our course selects the most interesting project. The one who can write about the subject receives many awards, said the class teacher. "

c. O projeto mais interessante foi selecionado pelo aluno do nosso curso. O que tinha redigido sobre o tema favoreceu os estudantes, disse o professor da turma.

" The most interesting project was selected by the student of our course. What he had written on the subject favored the students, said the class teacher. "

d. O projeto mais interessante é selecionado pelo aluno do nosso curso. O que pode escrever sobre o tema favorece os estudantes, disse o professor da turma.

" The most interesting project is selected by the student of our course. What one can write about the subject favors the students, said the class teacher. "

e. O projeto mais interessante foi selecionado pelo aluno do nosso curso. O que tinha redigido sobre o tema recebeu muitos prêmios, disse o professor da turma.

" The most interesting project was selected by the student of our course. The one who had written it on the subject received many awards, said the class teacher. "

f. O projeto mais interessante é selecionado pelo aluno do nosso curso. O que pode escrever sobre o tema recebe muitos prêmios, disse o professor da turma.

" The most interesting project is selected by the student of our course. The one who can write about the subject receives many awards, said the class teacher. "

g. O aluno do nosso curso selecionou o projeto mais interessante. O que tinha redigido sobre o tema favoreceu os estudantes, disse o professor da turma.

" The student of our course selected the most interesting project. What he had written on the subject favored the students, said the class teacher. "

h. O aluno do nosso curso seleciona o projeto mais interessante. O que pode escrever sobre o tema favorece os estudantes, disse o professor da turma.

" The student of our course selects the most interesting project. What one can write about the subject favors the students, said the class teacher. "

(529) a. O estudante da nossa escola elegeu o trecho do texto mais interessante. O que tinha interpretado em voz alta ficou bem envergonhado, disse o professor da escola.

" The student at our school chose the most interesting piece of text. The one who had interpreted it out loud was very ashamed, said the school teacher. "

b. O estudante da nossa escola elege o trecho do texto mais interessante. O que tem de ler em voz alta fica bem envergonhado, disse o professor da escola.

" The student at our school chooses the most interesting piece of text. The one has to read out loud is very ashamed, said the school teacher. "

c. O trecho do texto mais interessante foi eleito pelo estudante da nossa escola. O que tinha interpretado em voz alta foi preferido na escolha, disse o professor da escola.

" The most interesting piece of text was chosen by our school student. What he had interpreted out loud was preferred in the choice, the school teacher said. "

d. O trecho do texto mais interessante é eleito pelo estudante da nossa escola. O que tem de ler em voz alta é preferido na escolha, disse o professor da escola.

" The most interesting piece of text is chosen by the student at our school. What one has to read out loud is preferred in the choice, said the school teacher. "

e. O trecho do texto mais interessante foi eleito pelo estudante da nossa escola. O que tinha interpretado em voz alta ficou bem envergonhado, disse o professor da escola.

" The most interesting piece of text was chosen by the student at our school. The one who had interpreted it out loud was very ashamed, said the school teacher. "

f. O trecho do texto mais interessante é eleito pelo estudante da nossa escola. O que tem de ler em voz alta fica bem envergohado, disse o professor da escola.

" The most interesting piece of text was chosen by the student at our school. The one has to read out loud is very ashamed, said the school teacher. "

g. O estudante da nossa escola elegeu o trecho do texto mais interessante. O que tinha interpretado em voz alta foi preferido na escolha, disse o professor da escola.

" The student at our school chose the most interesting piece of text. What he had interpreted out loud was preferred in the choice, the school teacher said. "

h. O estudante da nossa escola elege o trecho do texto mais interessante. O que tem de ler em voz alta é preferido na escolha, disse o professor da escola.

" The student at our school chooses the most interesting piece of text. What one has to read out loud is preferred in the choice, said the school teacher. "

(530) a. O paciente do nosso hospital odiou o clarão do corredor central. O que não tinha percebido na verdade conseguiu dormir melhor, disse a enfermeira do hospital.

" The patient in our hospital hated the flash of the central hallway. The one who had not noticed actually got better sleep, said the hospital nurse. "

b. O paciente do nosso hospital odeia o clarão do corredor central. O que não pode enxergar na verdade consegue dormir melhor, disse a enfermeira do hospital.

" The patient in our hospital hates the glare in the central hallway. The one who cannot see actually get better sleep, said the hospital nurse. "

c. O clarão do corredor central foi odiado pelo paciente do nosso hospital. O que não tinha percebido na verdade não atrapalhou o sono, disse a enfermeira do hospital.

" The flash of the central corridor was hated by the patient in our hospital. What he had not noticed actually did not disturb his sleep, said the hospital nurse. "

d. O clarão do corredor central é odiado pelo paciente do nosso hospital. O que não pode enxergar na verdade não atrapalhou o sono, disse a enfermeira do hospital.

" The flash of the central corridor is hated by the patient in our hospital. What one cannot see actually does not disturb their sleep, said the hospital nurse. "

e. O clarão do corredor central foi odiado pelo paciente do nosso hospital. O que não tinha percebido na verdade conseguiu dormir melhor, disse a enfermeira do hospital.

" The glimmer of the central corridor was hated by the patient in our hospital. The one who had not noticed actually got better sleep, said the hospital nurse. "

f. O clarão do corredor central é odiado pelo paciente do nosso hospital. O que não pode enxergar na verdade consegue dormir melhor, disse a enfermeira do hospital.

" The flash of the central corridor is hated by the patient in our hospital. The one who cannot see actually get better sleep, said the hospital nurse. "

g. O paciente do nosso hospital odiou o clarão do corredor central. O que não tinha percebido na verdade não atrapalhou o sono, disse a enfermeira do hospital.

" The patient in our hospital hated the flash of the central corridor. What he had not noticed actually did not disturb his sleep, said the hospital nurse. "

h. O paciente do nosso hospital odeia o clarão do corredor central. O que não pode enxergar na verdade não atrapalha o sono, disse a enfermeira do hospital.

" The patient in our hospital hates the flash of the central corridor. What one cannot see actually does not disturb their sleep, said the hospital nurse. "

(531) a. O professor em regimes não-democráticos não mencionou o problema político fundamental. O que não tinha levantado em aula sofreu psicologicamente, disse o cientista politico.

" The teacher in undemocratic regimes did not mention the fundamental political problem. The one who had not raised it in class suffered psychologically, said the political scientist. "

b. O professor em regimes não-democráticos não menciona o problema político fundamental. O que não pode discutir em aula sofre psicologicamente, disse o cientista politico.

" The teacher in undemocratic regimes does not mention the fundamental political problem. The one who cannot discuss in class suffers psychologically, said the political scientist. "

c. O problema político fundamental não foi mencionado pelo professor em regimes não-democráticos. O que não tinha levantado em aula foi omitido do curriculo, disse o cientista politico.

" The fundamental political problem was not mentioned by the teacher in non-democratic regimes. What he had not raised in class was omitted from the curriculum, said the political scientist. "

d. O problema político fundamental não é mencionado pelo professor em regimes não-democráticos. O que não pode discutir em aula é omitido do currículo, disse o cientista político.

" The fundamental political problem is not mentioned by the teacher in non-democratic regimes. What one cannot discuss in class is omitted from the curriculum, said the political scientist. "

e. O problema político fundamental não foi mencionado pelo professor em regimes não-democráticos. O que não tinha levantado em aula sofreu psicologicamente, disse o cientista politico.

" The fundamental political problem was not mentioned by the teacher in non-democratic regimes. The one who had not raised it in class suffered psychologically, said the political scientist. "

f. O problema político fundamental não é mencionado pelo professor em regimes não-democráticos. O que não pode discutir em aula sofre psicologicamente, disse o cientista politico.

" The fundamental political problem is not mentioned by the teacher in non-democratic regimes. The one who cannot discuss in class suffers psychologically, said the political scientist. "

g. O professor em regimes não-democráticos não mencionou o problema político fundamental. O que não tinha levantado em aula foi omitido do curriculo, disse o cientista politico.

" The teacher in undemocratic regimes did not mention the fundamental political problem. What he had not raised in class was omitted from the curriculum, said the political scientist. "

h. O professor em regimes não-democráticos não menciona o problema político fundamental. O que não pode discutir em aula é omitido do currículo, disse o cientista político.

" The teacher in undemocratic regimes does not mention the fundamental political problem. What one cannot discuss in class is omitted from the curriculum, said the political scientist. "

(532) a. O político nos dias atuais não enfatizou o projeto de governo. O que tinha defendido em público sofreu com preconceitos, disse o candidato do partido.

" The politician today did not emphasize the project of government. The one who had defended it in public suffered with prejudices, said the party's candidate. "

b. O político nos dias atuais não enfatiza o projeto de governo. O que tem de polemizar em público sofre com preconceitos, disse o candidato do partido.

" The politician in nowadays does not emphasize the project of government. The one who has to polemize in public suffers from prejudices, said the party candidate. "

c. O projeto de governo não foi enfatizado pelo político nos dias atuais. O que tinha defendido em público foi discutido na campanha, disse o candidato do partido.

" The government project was not emphasized by the politician in nowadays. What he had defended in public was discussed in the campaign, said the party's candidate. "

d. O projeto de governo não é enfatizado pelo político nos dias atuais. O que tem de polemizar em público é discutido na campanha, disse o candidato do partido.

" The government project is not emphasized by the politician in nowadays. What one has to polemize in public is discussed in the campaign, said the party candidate. "

e. O projeto de governo não foi enfatizado pelo político nos dias atuais. O que tinha defendido em público sofreu com preconceitos, disse o candidato do partido.

" The government project was not emphasized by the politician in nowadays. The one who had defended it in public suffered with prejudices, said the party's candidate. "

f. O projeto de governo não é enfatizado pelo político nos dias atuais. O que tem de polemizar em público sofre com preconceitos, disse o candidato do partido.

" The government project is not emphasized by the politician in nowadays. The one who has to polemize in public suffers from prejudices, said the party candidate. "

g. O político nos dias atuais não enfatizou o projeto de governo. O que tinha defendido em público foi discutido na campanha, disse o candidato do partido.

" The politician in nowadays did not emphasize the government project. What he had defended in public was discussed in the campaign, said the party's candidate. "

h. O político nos dias atuais não enfatiza o projeto de governo. O que tem de polemizar em público é discutido na campanha, disse o candidato do partido.

" The politician in nowadays does not emphasize the project of government. What one has to polemize in public is discussed in the campaign, said the party candidate. "

(533) a. O motorista de caminhão detestou o veículo muito pesado. O que não tinha testado em autoestradas se sentiu prejudicado, disse o presidente do sindicato.

" The truck driver hated the very heavy vehicle. The one who had not tested it on motorways felt wronged, said the union president. "

b. O motorista de caminhão detesta o veículo muito pesado. O que não pode dirigir em autoestradas se sente prejudicado, disse o presidemte do sindicato.

" The truck driver hates the very heavy vehicle The one who cannot drive on motorways feels wronged, said the union president. "

c. O veículo muito pesado foi destestado pelo motorista de caminhão. O que não tinha testado em autoestradas enferrujou muitas vezas, disse o presidente do sindicato.

" The very heavy vehicle was dislodged by the truck driver. What he had not tested on motorways rusted many times, the union president said. "

d. O veículo muito pesado é destestado pelo motorista de caminhão. O que não pode dirigir em autoestradas enferruja muitas vezas, disse o presidente do sindicato.

" The very heavy vehicle is driven by the truck driver. What one cannot drive on motorways rustles many times, the union president said. "

e. O veículo muito pesado foi destestado pelo motorista de caminhão. O que não tinha testado em autoestradas se sentiu prejudicado, disse o presidente do sindicato.

" The very heavy vehicle was dislodged by the truck driver. The one who had not tested it on motorways felt wronged, said the union president. "

f. O veículo muito pesado é destestado pelo motorista de caminhão. O que não pode dirigir em autoestradas se sente prejudicado, disse o presidente do sindicato.

" The very heavy vehicle is dislodged by the truck driver. The one who cannot drive on motorways feels wronged, said the union president. "

g. O motorista de caminhão detestou o veículo muito pesado. O que não tinha testado em autoestradas enferrujou muitas vezas, disse o presidente do sindicato.

" The truck driver hated the very heavy vehicle. What he had not tested on motorways rusted many times, the union president said. "

h. O motorista de caminhão detesta o veículo muito pesado. O que não pode dirigir em autoestradas enferruja muitas vezes, disse o presidente do sindicato.

" The truck driver hates the very heavy vehicle What one cannot drive on motorways rustles many times, said the union president. "

(534) a. O bom jogador da nossa equipe preferiu o exercício esportivo extenuante. O que tinha realizado com frequência teve ganhos mais rápidos, disse o treinador da equipe.

" The good player of our team preferred the strenuous sporting exercise. The one who had accomplished it often had faster gains, the team coach said. "

b. O bom jogador da nossa equipe prefere o exercício esportivo extenuante. O que tem de treinar com frequência tem ganhos mais rápidos, disse o treinador da equipe.

" The good player of our team prefers strenuous sporting exercise. The one who has to work out frequently has faster gains, the team coach said. "

c. O exercício esportivo extenuante foi preferido pelo bom jogador da nossa equipe. O que tinha realizado com frequência causou ganhos bem rápidos, disse o treinador da equipe.

" The strenuous sporting exercise was preferred by the good player of our team. What he had accomplished often caused very quick gains, the team coach said. "

d. O exercício esportivo extenuante é preferido pelo bom jogador da nossa equipe. O que tem de treinar com frequência causa ganhos bem rápidos, disse o treinador do equipe.

" The strenuous sporting exercise is preferred by the good player of our team. What one has to train often causes very fast gains, said the coach of the team. "

e. O exercício esportivo extenuante é preferido pelo bom jogador da nossa equipe. O que tem de treinar com frequência tem ganhos mais rápidos, disse o treinador da equipe.

" The strenuous sporting exercise is preferred by the good player of our team. The one who has to work out frequently has faster gains, the team coach said. "

f. O bom jogador da nossa equipe preferiu o exercício esportivo extenuante. O que tinha realizado com frequência teve ganhos mais rápidos, disse o treinador da equipe.

" The good player of our team preferred the strenuous sporting exercise. The one who had accomplished it often had faster gains, the team coach said. "

g. O exercício esportivo extenuante foi preferido pelo bom jogador da nossa equipe. O que tinha realizado com frequência teve ganhos mais rápidos, disse o treinador da equipe.

" The strenuous sporting exercise was preferred by the good player of our team. What he had accomplished often caused very quick gains, the team coach said. "

h. O bom jogador da nossa equipe prefere o exercício esportivo extenuante. O que tem de treinar com frequência causa ganhos bem rápidos, disse o treinador da equipe.

" The good player of our team prefers strenuous sporting exercise. What one has to train often causes very fast gains, said the coach of the team. "

(535)   a. O adolescente chato abominou até o evento legal. O que tinha curtido com tranquilidade gostou de todas diversões, disse o psicólogo da universidade.

" The annoying teenager abhorred even the cool event. The one who had enjoyed with ease liked all diversions, said the university psychologist. "

b. O adolescente chato abomina até o evento legal. O que pode aproveitar com tranquilidade gosta de todas diversões, disse o psicólogo da universidade.

" The boring teenager abhors even the cool event. The one who can enjoy with ease likes all diversions, said the university psychologist. "

c. O evento legal foi abominado pelo adolescente chato. O que tinha curtido com tranquilidade foi lembrado para sempre, disse o psicólogo da universidade.

" The cool event was abhorred by the annoying teenager What he had enjoyed with ease was remembered forever, said the university psychologist. "

d. O evento legal é abominado pelo adolescente chato. O que pode aproveitar com tranquilidade é lembrado para sempre, disse o psicólogo da universidade.

" The cool event is abhorred by the annoying teenager What one can enjoy with ease is remembered forever, said the university psychologist. "

e. O evento legal foi abominado pelo adolescente chato. O que tinha curtido com tranquilidade gostou de todas diversões, foi abominado pelo adolescente chato.

" The cool event was abhorred by the annoying teenager The one who had enjoyed with ease liked all diversions, said the university psychologist. "

f. O evento legal é abominado pelo adolescente chato. O que pode aproveitar com tranquilidade gosta de todas diversões, disse o psicólogo da universidade.

" The cool event is abhorred by the annoying teenager The one who can enjoy with ease likes all diversions, said the university psychologist. "

g. O adolescente chato abominou até o evento legal. O que tinha curtido com tranquilidade foi lembrado para sempre, disse o psicólogo da universidade.

" The boring teenager abhorred even the cool event. What he had enjoyed with ease was remembered forever, said the university psychologist. "

h. O adolescente chato abomina até o evento legal. O que pode aproveitar com tranquilidade é lembrado para sempre, disse o psicólogo da universidade.

" The boring teenager abhors even the cool event. What one can enjoy with ease is remembered forever, said the university psychologist. "

(536)   a. O lojista empreendedor negou o pagamento por cheque. O que não tinha recebido em antecipação esperou a compensação, disse o exonomista da revista.

" The entrepreneurial shopkeeper denied payment by check. The one who had not received it in anticipation waited for the compensation, said the exonomist of the magazine. "

b. O lojista empreendedor nega o pagamento por cheque. O que não pode cobrar em antecipação espera a compensação, disse o economista da revista.

" The entrepreneurial shopkeeper denies payment by check. The one who can not charge in anticipation waits for compensation, said the economist of the magazine. "

c. O pagamento por cheque foi negado pelo lojista empreendedor. O que não tinha recebido em antecipação se converteu em despesas, disse o economista da revista.

" The payment by check was denied by the entrepreneurial shopkeeper. What he had not received in anticipation became expenses, said the economist of the magazine. "

d. O pagamento por cheque é negado pelo lojista empreendedor. O que não pode cobrar em antecipação se converte em despesas, disse o economista da revista.

" The payment by check is denied by the entrepreneurial shopkeeper. What one cannot charge in anticipation becomes expenses, said the economist of the magazine. "

e. O pagamento por cheque foi negado pelo lojista empreendedor. O que não tinha recebido em antecipação esperou a compensação, disse o economista da revista.

" The payment by check was denied by the entrepreneurial shopkeeper. The one who had not received it in anticipation waited for the compensation, said the exonomist of the magazine. "

f. O pagamento por cheque é negado pelo lojista empreendedor. O que não pode cobrar em antecipação espera a compensação, disse o economista da revista.

" The payment by check was denied by the enterprising shopkeeper. The one who can not charge in anticipation waits for compensation, said the economist of the magazine. "

g. O lojista empreendedor negou o pagamento por cheque. O que não tinha recebido em antecipação se converteu em despesas, disse o economiste da revista.

" The entrepreneurial shopkeeper denied the payment by check. What he had not received in anticipation became an expense, said the economist of the magazine. "

h. O lojista empreendedor nega o pagamento por cheque. O que não pode cobrar em antecipação se converte em despesas, disse o economista da revista.

" The entrepreneurial shopkeeper denies payment by check. What one cannot charge in anticipation becomes expenses, said the economist of the magazine. "

(537) a. O professor interessado apreciou o tema desafiador. O que tinha explicado bem para todos lidou com a dificuldade, disse o diretor da escola.

" The interested teacher appreciated the challenging topic. The one who had explained it well to everyone dealt with the difficulty, said the principal of the school. "

b. O professor interessado aprecia o tema desafiador. O que tem de ensinar bem lida com a dificuldade, disse o diretor da escola.

" The interested teacher appreciates the challenging topic. The one who has to teach well deals with difficulty, said the principal of the school. "

c. O tema desafiador foi apreciado pelo professor interessado. O que tinha explicado bem para todos foi menos sofisticado, disse o diretor da escola.

" The challenging topic was appreciated by the interested teacher What he had explained well to everyone was less sophisticated, said the principal of the school. "

d. O tema desafiador é apreciado pelo professor interessado. O que tem de ensinar bem é menos sofisticado, disse o diretor da escola.

" The challenging topic is appreciated by the interested teacher. What one has to teach well is less sophisticated, "said the school principal. "

e. O tema desafiador foi apreciado pelo professor interessado. O que tinha explicado bem para todos lidou com a dificuldade, disse o diretor da escola.

" The challenging topic was appreciated by the interested teacher. The one who had explained it well to everyone dealt with the difficulty, said the principal of the school. "

f. O tema desafiador é apreciado pelo professor interessado. O que tem de ensinar bem lida com a dificuldade, disse o diretor da escola.

" The challenging topic is appreciated by the concerned teacher. The one who has to teach well deals with difficulty, said the principal of the school. "

g. O professor interessado apreciou o tema desafiador. O que tinha explicado bem para todos foi menos sofisticado, disse o diretor da escola.

" The interested teacher appreciated the challenging topic. What he had explained well to everyone was less sophisticated, said the principal of the school. "

h. O professor interessado aprecia o tema desafiador. O que tem de ensinar bem é menos sofisticado, disse o diretor da escola.

" The interested teacher appreciates the challenging topic. What one has to teach well is less sophisticated, "said the school principal. "

(538) a. O motorista nas metrópoles da Europa se irritou como veículo muito grande. O que não tinha botado na garagem perdeu a paciência fácil, disse o fiscal de trânsito.

" The driver in the metropolises of Europe got annoyed as a very large vehicle. The one who had not put it in the garage easily lost his patience, the traffic supervisor said. "

b. O veículo muito grande irritou o motorista nas metrópoles da Europa. O que não tinha botado na garagem não foi vendido nas lojas, disse o fiscal de trânsito.

444

" The very large vehicle irritated the driver in Europe's metropolises. What he had not put in the garage was not sold in stores, the traffic supervisor said. "

c. O veículo muito grande irrita o motorista nas metrópoles da Europa. O que não pode estacionar na garagem não é vendido nas lojas, disse o fiscal de trânsito.

" The very large vehicle irritates the driver in the European metropolises What one cannot park in the garage is not sold in stores, said the traffic supervisor. "

d. O veículo muito grande irritou o motorista nas metrópoles da Europa. O que não tinha botado na garagem perdeu a paciência fácil, disse o fiscal de trânsito.

" The very large vehicle irritated the driver in Europe's metropolises. The one who had not put it in the garage easily lost his patience, the traffic supervisor said. "

e. O veículo muito grande irrita o motorista nas metrópoles da Europa. O que não pode estacionar na garagem perde a paciência fácil, disse o fiscal de trânsito.

" The very large vehicle irritates the driver in the European metropolises. The one who cannot park in the garage easily loses his patience, said the traffic supervisor. "

f. O motorista nas metrópoles da Europa se irritou como veículo muito grande. O que não tinha botado na garagem não foi vendido nas lojas, disse o fiscal de trânsito.

" The driver in the metropolises of Europe got annoyed as a very large vehicle. What he had not put in the garage was not sold in the stores, said the traffic supervisor. "

g. O motorista nas metrópoles da Europa se irrita como veículo muito grande. O que não pode estacionar na garagem perde a paciência fácil, disse o fiscal de trânsito.

" The driver in the metropolises of Europe gets annoyed as a very large vehicle. The one who cannot park in the garage easy loses his patience, said the traffic supervisor. "

h. O motorista nas metrópoles da Europa se irrita como veículo muito grande. O que não pode estacionar na garagem não é vendido nas lojas, disse o fiscal de trânsito.

" The driver in the metropolises of Europe gets annoyed as a very large vehicle. What one cannot park in the garage is not sold in stores, said the traffic supervisor. "

(539) a. O estudante de arquitetura temeu o projeto complexo. O que tinha estruturado com regularidade não se incomodou tanto, disse o professor universitário.

" The student of architecture feared the complex project. The one who had structured it regularly did not bother so much, said the university professor. "

b. O estudante de arquitetura teme o projeto complexo. O que tem de desenhar com regularidade não se incomoda tanto, disse o professor universitário.

" The student of architecture fears the complex project. The one who has to draw with regularity does not bother so much, said the university professor. "

445

c. O projeto complexo foi temido pelo estudante de arquitetura. O que tinha estruturado com regularidade foi extremante simples, disse o professor universitário.

" The complex project was feared by the architecture student. What he had structured with regularity was extremely simple, said the university professor. "

d. O projeto complexo é temido pelo estudante de arquitetura. O que tem de desenhar com regularidade é extremante simples, disse o professor universitário.

" The complex project is feared by the student of architecture. What one has to draw with regularity is very simple, said the university professor. "

e. O projeto complexo foi temido pelo estudante de arquitetura. O que tinha estruturado com regularidade não se incomodou tanto, foi temido pelo estudante de arquitetura.

" The complex project was feared by the student of architecture. The one who had structured it regularly did not bother so much, said the university professor. "

f. O projeto complexo é temido pelo estudante de arquitetura. O que tem de desenhar com regularidade não se incomoda tanto, disse o professor universitário.

" The complex project is feared by the student of architecture. The one who has to draw with regularity does not bother so much, said the university professor. "

g. O estudante de arquitetura temeu o projeto complexo. O que tinha estruturado com regularidade foi extremante simples, disse o professor universitário.

" The student of architecture dreaded the complex project. What he had structured with regularity was extremely simple, said the university professor. "

h. O estudante de arquitetura teme o projeto complexo. O que tem de desenhar com regularidade é extremante simples, disse o professor universitário.

" The student of architecture is afraid of the complex project. What one has to draw with regularity is very simple, said the university professor. "

(540) a. O brasão de um país instigou o colegial primário. O que tinha retratado na escolinha soube o desenho de cor, disse a professora primária.

" The coat of arms of a country instigated the primary schoolboy. The one who had pictured it in the school knew the drawing by heart, said the primary teacher. "

b. O brasão de um país instiga o colegial primário. O que tem de pintar na escolinha sabe o desenho de cor, disse a professora primária.

" The coat of arms of a country instigates the primary schoolboy. The one who has to paint in the school knows the drawing by heart, said the primary teacher. "

c. O colegial primário foi instigado pelo brasão de um país. O que tinha retratado na escolinha foi muito mais fácil, disse a professora primária.

" The primary schoolboy was instigated by the coat of arms of a country What he had pictured in the school was much easier, said the primary teacher. "

d. O colegial primário foi instigado pelo brasão de um país. O que tinha retratado na escolinha soube o desenho de cor, disse a professora primária.

" The primary schoolboy was instigated by the coat of arms of a country. The one who had pictured it in the school knew the drawing by heart, said the primary teacher. "

e. O colegial primário é instigado pelo brasão de um país. O que tem de pintar na escolinha sabe o desenho de cor, disse a professora primária.

" The primary schoolboy is instigated by the coat of arms of a country. The one who has to paint in the school knows the drawing by heart, said the primary teacher. "

f. O colegial primário é instigado pelo brasão de um país. O que tem de pintar na escolinha é muito mais fácil, disse a professora primária.

" The primary schoolboy is instigated by the coat of arms of a country. What one has to paint in the school is much easier, said the primary teacher. "

g. O brasão de um país instigou o colegial primário. O que tinha retratado na escolinha foi muito mais fácil, disse a professora primária.

" The coat of arms of a country instigated the primary schoolboy. What he had pictured in the school was much easier, said the primary teacher. "

h. O brasão de um país instiga o colegial primário. O que tem de pintar na escolinha é muito mais fácil, disse a professora primária.

" The coat of arms of a country instigates the primary schoolboy. What one has to paint in the school is much easier, said the primary teacher. "

(541) a. O costureiro experiente separou o vestido mais complexo. O que tinha cortado na vertical ficou irritado demais, disse a presidente da cooperativa.

" The experienced dressmaker separated the more complex dress. The one who had cut it vertically was too angry, said the president of the cooperative. "

b. O costureiro experiente separa o vestido mais complexo. O que tem de costurar na vertical fica irritado demais, disse a presidente da cooperativa.

" The experienced dressmaker separates the more complex dress. The one who has to sew vertically gets too angry, said the president of the cooperative. "

c. O vestido mais complexo foi separado pelo costureiro experiente. O que tinha cortado na vertical foi colocado separado, disse a presidente da cooperativa.

" The most complex dress was separated by the experienced dressmaker. What he had cut vertically was placed separately, said the president of the cooperative. "

d. O vestido mais complexo é separado pelo costureiro experiente. O que tem de costurar na vertical fica irritado demais, disse a presidente da cooperativa.

" The more complex dress is separated by the experienced dressmaker. The one who has to sew vertically is too angry, said the president of the cooperative. "

e. O costureiro experiente separou o vestido mais complexo. O que tinha cortado na vertical foi colocado separado, dise a presidente da cooperativa.

" The experienced dressmaker separated the more complex dress. What he had cut vertically was placed separately, designed the president of the cooperative. "

f. O vestido mais complexo é separado pelo costureiro experiente. O que tem de costurar na vertical é colocado separado, disse a presidente da cooperativa.

" The most complex dress is separated by the experienced dressmaker. What one have to sew vertically is placed separately, said the president of the cooperative. "

g. O vestido mais complexo foi separado pelo costureiro experiente. O que tinha cortado na vertical ficou irritado demais, disse a presidente da cooperativa.

" The most complex dress was separated by the experienced dressmaker. The one who had cut it vertically was too angry, said the president of the cooperative. "

h. O costureiro experiente separa o vestido mais complexo. O que tem de costurar na vertical é colocado separado, disse a presidente da cooperativa.

" The experienced dressmaker separates the more complex dress. What one has to sew vertically is placed separately, said the president of the cooperative. "

(542) a. O cantor de barzinho reconheceu o bolero clássico. O que tinha executado em festivais deleitou a audiência, disse o apresentador de televisão.

" The bar singer recognized the classic bolero. The one who had played it in festivals delighted the audience, said the television presenter. "

b. O cantor de barzinho reconhece o bolero clássico. O que pode cantar em festivais deleita a audiência, disse o apresentador de televisão.

" The singer of bar recognizes the classic bolero. The one who can sing in festivals delights the audience, said the television presenter. "

c. O bolero clássico foi reconhecido pelo cantor de barzinho. O que tinha executado em festivais foi extremamente moderno, disse o apresentador de televisão.

" The classic bolero was recognized by the singer of barzinho What he had played in festivals was extremely modern, said the television presenter. "

d. O bolero clássico é reconhecido pelo cantor de barzinho. O que pode cantar em festivais é extremamente moderno, disse o apresentador de televisão.

448

" The classic bolero is recognized by the singer of barzinho. What one can sing in festivals is extremely modern, said the presenter of television. "

e. O bolero clássico foi reconhecido pelo cantor de barzinho. O que tinha executado em festivais deleitou a audiência, disse o apresentador de televisão.

" The classic bolero was recognized by the barzinho singer The one who had played it in festivals delighted the audience, said the television presenter. "

f. O bolero clásico é reconhecido pelo cantor de barzinho. O que pode cantar em festivais deleita a audiência, disse o apresentador de televisão.

" The classic bolero is recognized by the bar singer. The one who can sing in festivals delights the audience, said the television presenter. "

g. O cantor de barzinho reconheceu o bolero clássico. O que tinha executado em festivais foi extremamente moderno, disse o apresentador de televisão.

" The singer from bar recognized the classic bolero. What he had played in festivals was extremely modern, said the television presenter. "

h. O cantor de barzinho reconhece o bolero clássico. O que pode cantar em festivais é extremamente moderno, disse o apresentador de televisão.

" The singer of bar recognizes the classic bolero. What one can sing in festivals is extremely modern, said the presenter of television. "

(543)   a. O bailarino profissional reconheceu o balé clássico. O que não tinha utilizado na competição ficou muito desapontado, disse o organizador do evento.

" The professional dancer recognized the classic ballet. The one who had not used it in the competition got very disappointed, said the organizer of the event. "

b. O bailarino profissional reconhece o balé clássico. O que não pode dançar na competição fica muito desapontado, disse o organizador do evento.

" The professional dancer recognizes the classic ballet. The one who cannot dance in the competition gets very disappointed, said the organizer of the event. "

c. O balé clássico foi reconhecido pelo bailarino profissional. O que não tinha utilizado na competição teve muito apelo popular, disse o organizador do evento.

" The classic ballet was recognized by the professional dancer. What he had not used in the competition had much popular appeal, said the organizer of the event. "

d. O balé clássico é reconhecido pelo bailarino profissional. O que não pode dançar na competição tem muito apelo popular, disse o organizador do evento.

" The classic ballet is recognized by the professional dancer. What one cannot dance in the competition has much popular appeal, said the organizer of the event. "

e. O balé clássico foi reconhecido pelo bailarino profissional. O que não tinha utilizado na competição ficou muito desapontado, disse o organizador do evento.

" The classical ballet was recognized by the professional dancer. The one who had not used it in the competition got very disappointed, said the organizer of the event. "

f. O balé clássico é reconhecido pelo bailarino profissional. O que não pode dançar na competição fica muito desapontado, disse o organizador do evento.

" The classic ballet is recognized by the professional dancer. The one who cannot dance in the competition gets very disappointed, said the organizer of the event. "

g. O bailarino profissional reconheceu o balé clássico. O que não tinha utilizado na competição teve muito apelo popular, disse o organizador do evento.

" The professional dancer recognized the classic ballet. What he had not used in the competition had much popular appeal, said the organizer of the event. "

h. O bailarino profissional reconhece o balé clássico. O que não pode dançar na competição tem muito apelo popular, disse o organizador do evento.

" The professional dancer recognizes the classic ballet. What one cannot dance in the competition has much popular appeal, said the organizer of the event. "

(544) a. O universitário preocupado preparou o tópico da próxima aula. O que tinha dominado com dedicação não pensou duas vezes, disse o instrutor do curso.

" The worried college student prepared the topic for the next class. The one who had mastered it with dedication did not think twice, said the instructor of the course. "

b. O universitário preocupado prepara o tópico da próxima aula. O que tem de estudar com dedicação não pensa duas vezes, disse o instrutor do curso.

" The worried college student prepares the topic for the next class. The one who has to study with dedication does not think twice, said the instructor of the course. "

c. O tópico da próxima aula é preparado pelo universitário preocupado. O que tem de estudar com dedicação aterroriza a maioria, disse o instrutor do curso.

" The topic of the next class is prepared by the worried college student. What one has to study with dedication terrifies the majority, said the instructor of the course. "

d. O tópico da próxima aula foi preparado pelo universitário preocupado. O que tinha dominado com dedicação não pensou duas vezes, disse o instrutor do curso.

" The topic of the next class was prepared by the worried college student. The one who had mastered it with dedication did not think twice, said the instructor of the course. "

e. O tópico da próxima aula é preparado pelo universitário preocupado. O que tem de estudar com dedicação não pensou duas vezes, disse o instrutor do curso.

" The topic of the next class is prepared by the concerned university The one who has to study with dedication does not think twice, said the instructor of the course. "

f. O universitário preocupado preparou o tópico da próxima aula. O que tinha dominado com dedicação aterrorizou a maioria, disse o instrutor do curso.

" The worried college student prepared the topic for the next class. What he had mastered with dedication terrified most, said the instructor of the course. "

g. O universitário preocupado prepara o tópico da próxima aula. O que tem de estudar com dedicação aterroriza a maioria, disse o instrutor do curso.

" The worried college student prepares the topic for the next class. What one has to study with dedication terrifies the majority, said the instructor of the course. "

h. O tópico da próxima aula foi preparado pelo universitário preocupado. O que tem de estudar com dedicação aterroriza a maioria, disse o instrutor do curso.

" The topic of the next class was prepared by the concerned university What one has to study with dedication terrifies the majority, said the instructor of the course. "

(545) a. O educador experiente desconsiderou o mal comportamento. O que não tinha punido de imediato segurou todas as emoções, disse a pedagoga da escola.

" The experienced educator disregarded the bad behavior. The one who had not punished it immediately held all the emotions, said the school teacher. "

b. O educador experiente desconsidera o mal comportamento. O que não pode responder de imediato segura todas as emoções, disse a pedagoga da escola.

" The experienced educator disregards bad behavior. The one who cannot respond immediately holds all emotions, said the school teacher. "

c. O mal comportamento foi desconsiderado pelo educador experiente. O que não tinha punido de imediato não foi geralmente aceito, disse a pedagoga da escola.

" The bad behavior was disregarded by the experienced educator. What he had not punish immediately was not generally accepted, said the school teacher. "

d. O mal comportamento é desconsiderado pelo educador experiente. O que não pode responder de imediato não é geralmente aceito, disse a pedagoga da escola.

" Bad behavior is disregarded by the experienced educator What one cannot respond immediately is not generally accepted, said the school teacher. "

e. O mal comportamento foi desconsiderado pelo educador experiente. O que não tinha punido de imediato segurou todas as emoções, disse a pedagoga da escola.

" Bad behavior was disregarded by the experienced educator. The one who had not punished it immediately held all the emotions, said the school teacher. "

f. O mal comportamento é desconsiderado pelo educador experiente. O que não pode responder de imediato segura todas as emoções, disse a pedagoga da escola.

" Bad behavior is disregarded by the experienced educator. The one who cannot respond immediately holds all emotions, said the school teacher. "

g. O educador experiente desconsiderou o mal comportamento. O que não tinha punido de imediato não foi geralmente aceito, disse a pedagoga da escola.

" The experienced educator disregarded the bad behavior. What he had not punish immediately was not generally accepted, said the school teacher. "

h. O educador experiente desconsidera o mal comportamento. O que não pode responder de imediato não é geralmente aceito, disse a pedagoga da escola.

" The experienced educator disregards misbehavior. What one cannot respond immediately is not generally accepted, said the school teacher. "

(546) a. O candidato à vaga de tintureiro receou o tecido com muitas cores. O que tinha finalizado com rapidez ficou bastante nervoso, disse o avaliador da empresa.

" The candidate for the spot of the dyer feared the fabric in many colors. The one who had finished it quickly became quite nervous, said the company appraiser. "

b. O candidato à vaga de tintureiro receia o tecido com muitas cores. O que tem de colorir com rapidez fica bastante nervoso, disse o avaliador da empresa.

" The candidate for the job of dyer fears the fabric in many colors. The one who has to color quickly becomes quite nervous, said the company appraiser. "

c. O tecido com muitas cores foi receado pelo candidato à vaga de tintureiro. O que tinha finalizado com rapidez foi bastante complicado, disse o avaliador da empresa.

" The cloth with many colors was feared by the candidate for the job of dyer. What he had finished quickly was quite complicated, said the company appraiser. "

d. O tecido com muitas cores é receado pelo candidato à vaga de tintureiro. O que tem de colorir com rapidez é bastante complicado, disse o avaliador da empresa.

" The cloth with many colors is feared by the candidate for the job of dyer. What one has to color quickly is quite complicated, said the appraiser of the company. "

e. O tecido com muitas cores foi receado pelo candidato à vaga de tintureiro. O que tinha finalizado com rapidez ficou bastante nervoso, disse o avaliador da empresa.

" The cloth with many colors was feared by the candidate for the job of dyer. The one who had finished it quickly got quite nervous, said the company appraiser. "

f. O tecido com muitas cores é receado pelo candidato à vaga de tintureiro. O que tem de colorir com rapidez fica bastante nervoso, disse o avaliador da empresa.

" The fabric with many colors is feared by the candidate for the job of dyer. The one who has to color quickly becomes quite nervous, said the company appraiser. "

g. O candidato à vaga de tintureiro receou o tecido com muitas cores. O que tinha finalizado com rapidez foi bastante complicado, disse o avaliador da empresa.

" The candidate for the dyer job feared the fabric in many colors. What he had finished quickly was quite complicated, said the company's appraiser. "

h. O candidato à vaga de tintureiro receia o tecido com muitas cores. O que tem de colorir com rapidez é bastante complicado, disse o avaliador da empresa.

" The candidate for the job of dyer fears the fabric in many colors. What one has to color quickly is quite complicated, said the company appraiser. "

(547) a. O ator reconhecido não descartou o papel desafiador. O que não tinha reproduzido com naturalidade perdeu a oportunidade, disse o diretor do espetáculo.

" The recognized actor did not rule out the challenging role. The one who had not reproduced it naturally missed the opportunity, said the director of the show. "

b. O ator reconhecido não descarta o papel desafiador. O que não pode representar com naturalidade perde a oportunidadem disse o diretor do espetáculo.

" The recognized actor does not rule out the challenging role. The one who cannot play naturally misses the opportunity, said the director of the show. "

c. O papel desafiador não foi descartado pelo ator reconhecido. O que não tinha reproduzido com naturalidade não foi escolhido primeiro, disse o diretor do espetáculo.

" The challenging role was not dismissed by the recognized actor. What he had not reproduced naturally was not chosen first, said the director of the show. "

d. O papel desafiador não é descartado pelo ator reconhecido. O que não pode representar com naturalidade não é escolhido primeiro, disse o diretor do espetáculo.

" The challenging role is not discarded by the recognized actor. What one cannot play naturally is not chosen first, said the director of the show. "

e. O papel desafiador não foi descartado pelo ator reconhecido. O que não tinha reproduzido com naturalidade perdeu a oportunidade, disse o diretor do espetáculo.

" The challenging role was not dismissed by the recognized actor. The one who had not reproduced it naturally missed the opportunity, said the director of the show. "

f. O papel desafiador não é descartado pelo ator reconhecido. O que não pode representar com naturalidade perde a oportunidade,disse o diretor do espetáculo.

" The challenging role is not discarded by the recognized actor. The one who cannot play naturally misses the opportunity, said the director of the show. "

g. O ator reconhecido não descartou o papel desafiador. O que não tinha reproduzido com naturalidade não foi escolhido primeiro, disse o diretor do espetáculo.

" The recognized actor did not rule out the challenging role. What he had not reproduced naturally was not chosen first, said the director of the show. "

h. O ator reconhecido não descarta o papel desafiador. O que não pode representar com naturalidade não é escolhido primeiro, disse o diretor do espetáculo.

" The recognized actor does not rule out the challenging role. What one cannot play naturally is not chosen first, said the director of the show. "

(548) a. O paciente do setor oftalmológico reconheceu o vulto das pessoas. O que não tinha notado no escuro não distinguiu quase nada, disse o enfermeiro do hospital.

" The patient in the ophthalmology recognized the figure of people. The one who had not noticed it in the dark distiguished almost nothing, said the hospital nurse. "

b. O paciente do setor oftalmológico reconhece o vulto das pessoas. O que não pode ver no escuro não distingue quase nada, disse o enfermeiro do hospital.

" The patient in the ophthalmology recognizes the figure of people. The one who cannot see in the dark distinguishes almost nothing, said the hospital nurse. "

c. O vulto das pessoas foi reconhecido pelo paciente do setor oftalmológico. O que não tinha notado no escuro não foi muito incômodo, disse o enfermeiro do hospital.

" The figure of the people was recognized by the ophthalmological patient. What he had not noticed in the dark was not very uncomfortable, said the hospital nurse. "

d. O vulto das pessoas é reconhecido pelo paciente do setor oftalmológico. O que não pode ver no escuro não é muito incômodo, disse o enfermeiro do hospital.

" The figure of people is recognized by the ophthalmologist. What one cannot see in the dark is not very uncomfortable, said the hospital nurse. "

e. O vulto das pessoas foi reconhecido pelo paciente do setor oftalmológico. O que não tinha notado no escuro não distinguiu quase nada, disse o enfermeiro do hospital.

" The figure of the people was recognized by the ophthalmological patient. The one who had not noticed it in the dark distiguished almost nothing, said the hospital nurse. "

f. O vulto das pessoas é reconhecido pelo paciente do setor oftalmológico. O que não pode ver no escuro não distingue quase nada, disse o enfermeiro do hospital.

" The figure of people is recognized by the ophthalmological patient. The one who cannot see in the dark distinguishes almost nothing, said the hospital nurse. "

g. O paciente do setor oftalmológico reconheceu o vulto das pessoas. O que não tinha notado no escuro não foi muito incômodo, disse o enfermeiro do hospital.

" The patient in the eye care sector recognized the figure of the people. What he had not noticed in the dark was not very uncomfortable, said the hospital nurse. "

h. O paciente do setor oftalmológico reconhece o vulto das pessoas. O que não pode ver no escuro não é muito incômodo, disse o enfermeiro do hospital.

" The patient in the ophthalmology recognizes the figure of the people. What one cannot see in the dark is not very uncomfortable, said the hospital nurse. "

Question Task

(549)   Então, foi o vinho especial do nosso restaurante que foi tomado de graça?

" So, was it the especial wine from our restaurant that was taken for free? "

(550)   Então, é o bom jogador da nossa equipe que tem de treinar com frequência um exercício esportivo extenuante?

" So, it is the good player of our team that has to train frequently a strenuous work out? "

(551)   Então, é o clarão do corredor central que não se pode enxergar no hospital?

" So it's the flash of the central corridor that you can not see in the hospital? "

(552)   Então, é o paciente do setor oftalmológico que não pode ver no escuro o vulto das pessoas?

" So, is the ophthalmological patient who can not see the shape of people in the dark? "

(553)   Então, é o vinho especial do nosso restaurante que se pode beber de graça?

" So, is the especial wine in our restaurant that you can drink for free? "

(554)   Então, foi o bolero clássico que foi executado em festivais?

" So, was it the classic bolero that was run at festivals? "

(555)   Então, foi o bom jogador da nossa equipe que tinha realizado com frequência um exercício esportivo extenuante?

" So it was the good player of our team who had frequently performed a strenuous work out? "

(556)   Então, foi o clarão do corredor central que foi percebido pelo paciente do hospital?

" So it was the flash of the central corridor that was perceived by the hospital patient? "

(557)   Então, foi o paciente do setor oftalmológico que tinha notado no escuro o vulto das pessoas?

" So it was the ophthalmological patient who had noticed in the dark the figure of the people? "

Appendix C

Participants' Information – Experiments 6, 7 and 8

Total of Participants	Mean Age	Sex	Eye Recorded	Mean Foreign Language Level	Mean Time Abroad
40	31 (range: $\pm$ 6)	35 female	17 left	4.5 (out of 5)	3 years (range: $\pm$ 2,5)

Table C.1 – Participants' Information

Latest Completed Level of Education	High School	Under-Graduation	Graduation
Number of Participants	10	22	8

Table C.2 – Participants' Completed Level of Education

Figure C.1 – Participants according to Region
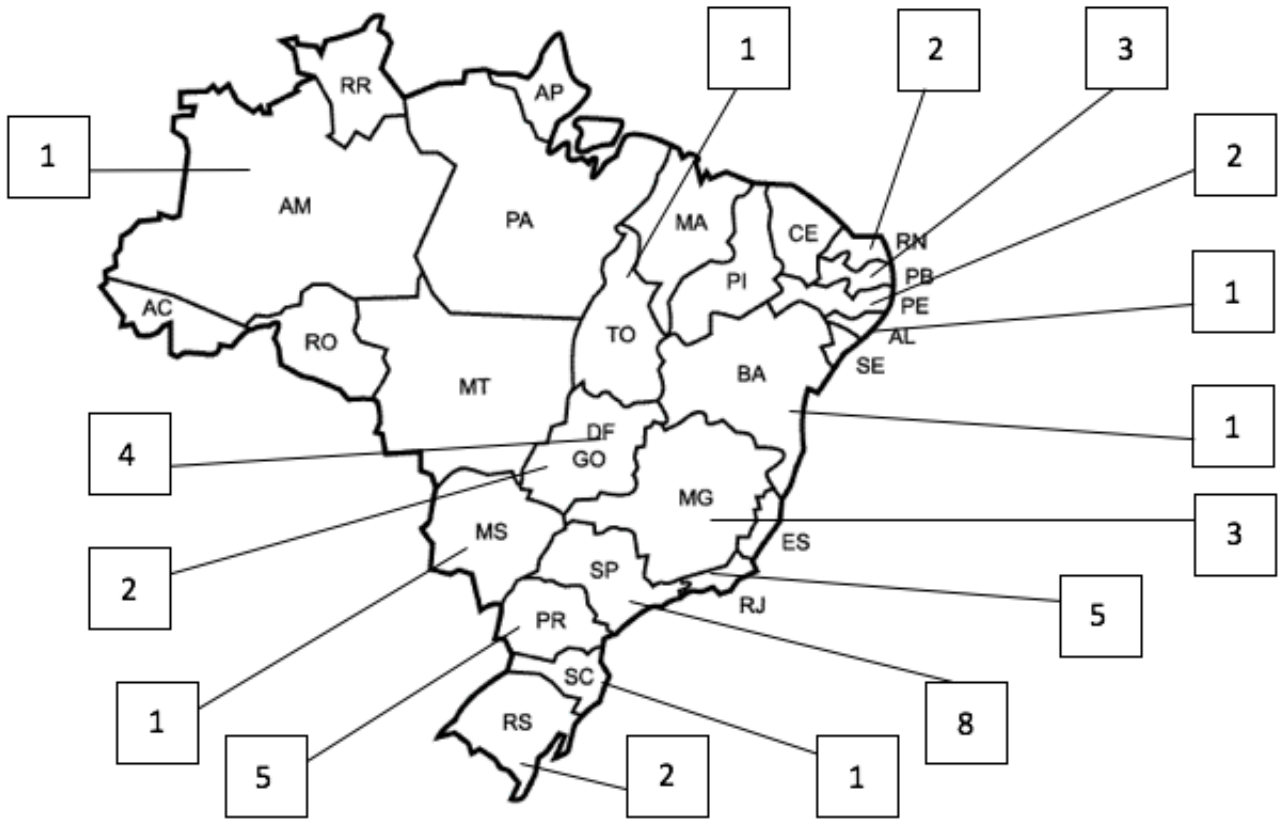
Appendix D

Extra Log-linear Mixed-Effects Models – Experiments 6, 7 and 8

In this appendix, models with the same measures and same structure used in the mixed models presented in Chapter 4 are shown. That is, when a maximal model of the logs of the total RTs in the CR was discussed in Chapter 4, here maximal models of the logs of the first pass and regression path RTs in the CR are presented.

D.1   Experiment 6

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	7.980e+00	1.762e-01	45.290	< 2e-16 ***
Adjunct	-5.350e-02	6.587e-02	-0.812	0.417
Subject	-9.136e-02	6.391e-02	-1.430	0.153
Verb	5.493e-02	6.396e-02	0.859	0.391
Length	6.345e-02	9.084e-03	6.984	7.98e-09 ***
Line_Break	2.317e-01	3.887e-02	5.960	3.31e-09 ***
Session	2.427e-03	3.199e-02	0.076	0.940
Subject:Verb	2.913e-02	9.084e-02	0.321	0.749
Subject:Adjunct	6.704e-02	8.988e-02	0.746	0.456
Verb:Adjunct	1.081e-01	8.983e-02	1.204	0.229
Subject:Verb:Adjunct	-1.009e-01	1.275e-01	-0.791	0.429

Table D.1 – Log-linear mixed-effects model for the Logs of First Pass RTs in CR in Experiment 6

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	8.569e+00	1.516e-01	56.509	< 2e-16 ***
Adjunct	-2.151e-01	5.989e-02	-3.591	0.000341 ***
Subject	-6.093e-02	5.912e-02	-1.031	0.302927
Verb	1.032e-01	5.876e-02	1.756	0.079372 .
Length	5.304e-02	7.387e-03	7.180	1.24e-08 ***
Line_Break	-1.643e-02	3.521e-02	-0.467	0.640869
Session	-3.550e-02	2.927e-02	-1.213	0.225425
Subject:Verb	9.844e-03	8.419e-02	0.117	0.906929
Subject:Adjunct	3.191e-02	8.227e-02	0.388	0.698179
Verb:Adjunct	4.920e-02	8.187e-02	0.601	0.547980
Subject:Verb:Adjunct	-3.653e-02	1.168e-01	-0.313	0.754435

Table D.2 – Log-linear mixed-effects model for the Logs of Regression Path RTs in CR in Experiment 6

D.2 Experiment 7

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	8.27688	0.20407	40.559	< 2e-16 ***
Subject	-0.02886	0.11205	-0.258	0.798395
Verb	-0.06006	0.08831	-0.680	0.503983
Length	0.04247	0.01033	4.110	0.000254 ***
Session	0.07230	0.06857	1.054	0.298753
Subject:Verb	-0.09497	0.12516	-0.759	0.455543

Table D.3 – Log-linear mixed-effects model for the Logs of First Pass RTs in CR in Experiment 7

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	8.85537	0.25293	35.011	< 2e-16 ***
Subject	-0.03352	0.07344	-0.456	0.65033
Verb	-0.04544	0.07545	-0.602	0.55049
Length	0.03683	0.01292	2.852	0.00601 **
Session	-0.04467	0.06089	-0.734	0.46769
Subject:Verb	-0.11533	0.10301	-1.120	0.26892

Table D.4 – Log-linear mixed-effects model for the Logs of Regression Path RTs in CR in Experiment 7

D.3   Experiment 8

D.3.1   Critical Region

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	9.95404	0.36943	26.944	< 2e-16 ***
Rel_Pro	0.11100	0.06893	1.610	0.1076
Antecedent	-0.07555	0.06636	-1.138	0.2552
Verb	-0.06882	0.07083	-0.972	0.3315
Length	0.02555	0.01521	1.680	0.0948 .
Session	-0.19485	0.03414	-5.707	1.45e-08 ***
Antecedent:Verb	0.05103	0.09404	0.543	0.5875
Antecedent:Rel_Pro	0.12803	0.09654	1.326	0.1850
Verb:Rel_Pro	-0.04384	0.09644	-0.455	0.6495
Antecedent:Verb:Rel_Pro	0.01352	0.13819	0.098	0.9221

Table D.5 – Log-linear mixed-effects model for Logs of Total RTs in CR in Experiment 8

D.3.2   Post-Critical Region

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	24.278	29.896	0.812	0.4176
Rel_Pro	-35.761	35.756	-1.000	0.3174
Antecedent	-7.685	35.187	-0.218	0.8272
Verb	-64.030	35.010	-1.829	0.0676 .
Session	3.218	17.596	0.183	0.8549
Antecedent:Verb	77.489	49.632	1.561	0.1187
Antecedent:Rel_Pro	38.962	50.016	0.779	0.4361
Verb:Rel_Pro	62.992	49.802	1.265	0.2061
Antecedent:Verb:Rel_Pro	-129.791	70.234	-1.848	0.0648 .

Table D.6 – Log-linear mixed-effects model for Residuals of First Pass RTs in the Post-CR in the model considering Length in Experiment 8

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	9.637e+00	2.983e-01	32.302	< 2e-16 ***
Rel_Pro	-3.323e-02	6.235e-02	-0.533	0.5941
Antecedent	-2.368e-02	6.134e-02	-0.386	0.6995
Verb	-6.733e-02	6.039e-02	-1.115	0.2651
Length	2.109e-02	9.618e-03	2.193	0.0392 *
Session	-1.448e-01	3.046e-02	-4.752	2.21e-06 ***
Antecedent:Verb	3.451e-03	8.568e-02	0.040	0.9679
Antecedent:Rel_Pro	7.044e-02	8.736e-02	0.806	0.4202
Verb:Rel_Pro	7.199e-02	8.620e-02	0.835	0.4038
Antecedent:Verb:Rel_Pro	-7.937e-02	1.217e-01	-0.652	0.5145

Table D.7 – Log-linear mixed-effects model for Logs of Total RTs in Post-CR in Experiment 8

D.3.3 Sums of Critical and Post-Critical Regions

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	-31.93	45.86	-0.696	0.48758
Rel_Pro	11.24	49.11	0.229	0.81894
Antecedent	16.47	49.13	0.335	0.73747
Verb	86.49	47.58	1.818	0.06931 .
Session	-36.61	24.14	-1.516	0.12961
Antecedent:Verb	-108.50	67.63	-1.604	0.10885
Antecedent:Rel_Pro	37.86	70.41	0.538	0.59090
Verb:Rel_Pro	-78.39	68.04	-1.152	0.24943
Antecedent:Verb:Rel_Pro	33.04	96.57	0.342	0.73231

Table D.8 – Log-linear mixed-effects model for Residuals of First Pass RTs in the sums of the CR and Post-CR in the model considering Length in Experiment 8

Condition	Estimate	Std. Error	T-value	Pr(>|z|)
(Intercept)	84.31	102.49	0.823	0.411350
Rel_Pro	75.84	127.26	0.596	0.551396
Antecedent	-31.96	124.67	-0.256	0.797747
Verb	16.06	119.47	0.134	0.893082
Session	-205.45	60.63	-3.388	0.000737 ***
Antecedent:Verb	77.489	49.632	1.561	0.1187
Antecedent:Rel_Pro	62.37	178.31	0.350	0.726574
Verb:Rel_Pro	63.56	168.18	0.378	0.705596
Antecedent:Verb:Rel_Pro	-70.21	240.92	-0.291	0.770800

Table D.9 – Log-linear mixed-effects model for Residuals of Regression Path RTs in the sums of the CR and Post-CR in the model considering Length in Experiment 8