UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE ESTUDOS DA LINGUAGEM

Alba Taboas Garcia

WH *IN SITU* AND TOPIC-SUBJECT CONSTRUCTIONS IN BRAZILIAN PORTUGUESE ACQUISITION AS A SECOND LANGUAGE

QU *IN SITU* E CONSTRUÇÕES DE TÓPICO-SUJEITO NA AQUISIÇÃO DE PORTUGUÊS BRASILEIRO COMO SEGUNDA LÍNGUA

Campinas

2020

ALBA TABOAS GARCIA

WH *IN SITU* AND TOPIC-SUBJECT CONSTRUCTIONS IN BRAZILIAN PORTUGUESE ACQUISITION AS A SECOND LANGUAGE

QU *IN SITU* E CONSTRUÇÕES DE TÓPICO-SUJEITO NA AQUISIÇÃO DE PORTUGUÊS BRASILEIRO COMO SEGUNDA LÍNGUA

Dissertation presented to the Institute of Language Studies of the University of Campinas in partial fulfillment of the requirements for the degree of Master in Linguistics

Dissertação de mestrado apresentada ao Instituto de Estudos da Linguagem da Universidade Estadual de Campinas para obtenção do título de Mestra em Linguística.

Advisor: Ruth Elisabeth Vasconcellos Lopes
Co-advisor: Juanito Ornelas de Avelar

Este arquivo digital corresponde à versão final da dissertação de mestrado defendida pela aluna Alba Taboas Garcia, orientada pela Profa. Dra. Ruth Elisabeth Vasconcellos Lopes e coorientada pelo Prof. Dr. Juanito Ornelas de Avelar

Campinas

2020

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Estudos da Linguagem
Leandro dos Santos Nascimento - CRB 8/8343

BANCA EXAMINADORA:

Ruth Elisabeth Vasconcellos Lopes

Charlotte Marie Chambelland Galves

Elaine Bicudo Grolla

Ata da defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós Graduação do IEL.

To my family,
the one I got and the one I created.

Acknowledgements

My gratitude goes first and foremost to my advisors. I would like to thank Ruth Lopes for her dedication and expertise, her constant support through my Linguistics career, and her sensibility to know when to challenge me and when to encourage me. To Juanito Avelar, who guided and motivated me throughout the process, I owe my fascination for syntax and the inspiration behind this project.

I am extremely grateful to Prof. Elaine Grolla for her valuable suggestions and contributions to this work, and to Prof. Charlotte for her inspiring lessons and discussions with me, which undoubtedly improved the quality of this dissertation.

I am indebted to all my professors at the Institute of Language Studies for giving me the tools to become a linguist and a researcher.

A special acknowledgement goes to Prof. Eleonora Cavalcante and my fellow students in her class for their emotional support and academic encouragement in times of need.

Thanks to my colleagues Rafael and Marcos, who offered me a helping hand in this endeavor.

My appreciation extends as well to all the people who took part in the study, making this research not only possible but also fun. A shoutout to Xabier, for finding many Spanish speakers and persuading them to help me.

To my closest friends, those who became my second family, thank you for taking care of me and letting me take care of you.

I am most sincerely thankful to my parents in law, Franca and Giuseppe, for their invaluable assistance despite the distance and for their kind words of encouragement always at the ready.

Above all, I am indebted to my parents, Antonia y Jose, for their unconditional love and support in all the possible meanings of the word. Thank you for being always on call.

Last, but not least, huge thanks to Lorenzo for giving me a new purpose and perspective, and my deepest gratitude to Simone for lovingly supporting all my decisions while taking me out of my comfort zone and constantly pushing me to go further.

Resumo

De uma perspectiva gerativista e adotando uma abordagem experimental, este trabalho se propõe a comparar duas estruturas sintáticas do Português Brasileiro, interrogativas in situ e construções de tópico-sujeito (CTS), na sua aquisição como segunda língua por falantes nativos de Espanhol. Os objetivos principais são confirmar que o QU in situ é adquirido muito antes do que as CTS, que podem até não chegar a ser incorporadas à gramática da interlíngua, e discutir as possíveis causas por trás desse contraste. O estudo experimental foi baseado em julgamentos de aceitabilidade, medindo os tempos de reação, e em uma tarefa de correção das sentenças consideradas inaceitáveis por cada participante, que fez as funções de um *follow-up*. Uma análise aprofundada dos resultados confirma a hipótese, mostrando que há efetivamente uma diferença significativa na aquisição das duas construções: o QU in situ não traz problemas aos aprendizes de PB, enquanto as CTS não são facilmente adquiridas, sendo, de fato, consistentemente rejeitadas à medida que o nível de proficiência do aprendiz aumenta. A explicação mais satisfatória para essa distinção será, no estudo, atribuída às diferenças sintáticas entre as duas estruturas. De acordo com os trabalhos de Avelar e Galves (2011, 2016), o tópico-sujeito nas CTS ocupa a posição de Spec-TP, o que é possível graças a phi-independência do EPP em T e a possibilidade de DPs sem traços de Caso integrar a derivação em PB. A aquisição das CTS requer, dessa perspectiva, a implementação de uma mudança substancial na gramática da L2 de falantes que têm o Espanhol como L1. Em contraste, assumindo uma análise sem movimento para as QU in situ do PB, como a de Grolla (2005), a aquisição desta estrutura implica o relaxamento da força de um traço formal compartilhado, de forte na L1 (Etxepare & Uribe-Etxebarria, 2005) a opcional na L2, o que constitui uma mudança muito menos drástica do que a outra. Como contribuição às discussões mais abrangentes sobre aquisição de segunda língua, o estudo aponta que os resultados referentes às interrogativas in situ podem ser considerados como evidência em favor das abordagens que advogam pelo acesso à GU (na linha de Hettiarachchi & Pires, 2016) e contra as que defendem um déficit representacional na aquisição do traço-QU (como em Hawkins & Chan, 1997; ou Hawkins & Hattori, 2006). Por outro lado, os resultados para as CTS sugerem que há uma interferência linguística tipológica da L1 (como defende Jung, 2004) e apoiam a ideia de que a psicotipologia é um fator influente na interferência linguística na aquisição de uma segunda língua. Finalmente, os resultados da pesquisa sugerem fortemente que as CTS constituem um fenômeno sintático em variação em PB como L1, o que leva à necessidade de ulteriores investigações sobre o assunto, para obter uma melhor compreensão do lugar que as CTS ocupam nas gramáticas nativas.

Palavras chave: gramática gerativa, aquisição de segunda língua, construções de tópico-sujeito, QU *in situ*

Abstract

From a Generative perspective and taking an experimental approach, this research aims to compare two Brazilian Portuguese syntactic structures, *in situ* interrogatives and topic-subject constructions (TSCs), in their second language acquisition (SLA) by Spanish native speakers. The main objectives are to confirm wh *in situ* is acquired much earlier than TSCs, which may even fail to be incorporated to the interlanguage grammar, and to explore the plausible causes for this contrast. The experimental study was based on an acceptability judgement task, for which reaction times were recorded, and a follow-up in the form of a correction task of the sentences considered unacceptable by each participant. A comprehensive analysis of the results verifies the hypothesis, revealing there is indeed a significative difference in the acquisition of the two constructions: wh *in situ* poses no problem for BP learners whereas TSCs are not easily acquired, being in fact increasingly rejected as the proficiency level rises. This distinction is best accounted for by attributing it to the syntactic differences between the two structures. According to Avelar and Galves' work (2011, 2016), the topic-subject constituent in TSCs is located in Spec-TP, something that is possible due to T's EPP phi-independence in BP and the possibility of DPs without Case features being inserted in the derivation in this language. Hence, under these assumptions, the acquisition of TSCs requires the implementation of a substantial change in the L2 grammar of Spanish native speakers. On the other hand, adopting an *in situ* approach like Grolla's (2005) for BP wh *in situ*, the acquisition of this structure involves the loosening of a shared formal feature's strength value, from strong in the L1 (Etxepare & Uribe-Etxebarria (2005), to optional in the L2, which constitutes a far less drastic change. As a contribution to wider discussions on general SLA theories, this work points out that the results regarding *in situ* interrogatives can be considered evidence in favor of Full Access approaches (in line with Hettiarachchi & Pires, 2016) and against Representational Deficit accounts of wh-feature acquisition (like those in Hawkins & Chan, 1997; or Hawkins & Hattori, 2006). In turn, TSCs results indicate there is typological transfer from the L1 (as defended by Jung, 2004) and provide support to the influence of psychotypology in L1 transfer in SLA. Finally, this research strongly suggests TSCs constitute a syntactic phenomenon under variation in BP as L1, which calls for further research on this topic so that a better understanding of how TSCs fit in native grammars can be achieved.

keywords: generative grammar, second language acquisition, topic-subject constructions, wh *in situ*

List of Abbreviations

ACC	Accusative Case
Adv	Adverb
AGR	Agreement
AgrP	Agreement Phrase
AJT	Acceptability Judgement Task
ASP	Aspect
BP	Brazilian Portuguese
CL	Clitic
CLASS	Classifier
CP	Complementizer Phrase
D	Determiner
DAT	Dative Case
DECL	Declarative
DEM	Demonstrative
DO	Direct Object
DP	Determiner Phrase
ec	Empty category
ECM	Exceptional Case Marking
EP	European Portuguese
EPP	Extended Projection Principle
Expl	Expletive
FI	Full Interpretation
FocP	Focus Phrase
FP	Finite Phrase
GJT	Grammaticality Judgement Task
IMPERF	Imperfective
IO	Indirect Object
IP	Inflection Phrase
L1	First Language
L2	Second Language
L3	Third Language
LBC	Left Branch Condition
LF	Logical Form
Ln	nth Language
LOC	Locative
LR	Last Resort
N	Noun
NEG	Negative particle
NegP	Negative Phrase
NOM	Nominative Case
NP	Noun Phrase
OP	Operator
P	Preposition
PAST	Past Tense

PERF	Perfective
Pers	Person
PF	Phonetic Form
PL	Plural
PP	Prepositional Phrase
Pres. Cont.	Present Continuous
Q	Interrogative particle
QSV	Interrogative element - Subject - Verb
QVS	Interrogative element - Verb - Subject
Q_{WH}	Wh interrogative particle
$Q_{Y/N}$	Yes/No interrogative particle
RP	Relator Phrase
RRC	Restrictive Relative Clause
SC	Small Clause
SD	Standard Deviation
SG	Singular
SLA	Second Language Acquisition
Spec	Specifier
TOP	Topic
Top	Topic
TP	Temporal Phrase
TSC	Topic-Subject Construction
UG	Universal Grammar
V	Verb
VAR	Variable
VP	Verbal Phrase

Table of contents

PART III Theoretical and Empirical Implications

Introduction

What, why and how

Taking a Generative view of second language acquisition (SLA), the main focus of this research is to analyze and compare the acquisition of two Brazilian Portuguese (BP) syntactic structures, *in situ* interrogatives and topic-subject constructions, by Spanish adult native speakers.

Interrogative sentences and crosslinguistic differences related to them have attracted syntacticians' interest from very early on and the theoretical discussion is far from over. Why do some languages necessarily move wh-phrases to the front of the sentence (like English) while others leave them *in situ*, i.e. in the position where they are semantically interpreted (like Chinese)? And more interestingly still, why do some languages allow both possibilities? The example below shows BP falls into the latter category, and, as we shall see in Chapter 1, accounting for this optionality constitutes a challenge for the theory.

(1)   a. ***O que*** *o Pedro comprou?*

          the what the Pedro bought

      b. *O Pedro comprou* ***o que****?*

          the Pedro bought what

      'What did Pedro buy?'

Topic-subject constructions are even more fascinating. These can be roughly defined as sentences in which a non-argumental constituent behaves like a subject, triggering verb agreement. The term was introduced by Pontes' pioneer work (1988) and it is widely accepted, but it has turned into an 'umbrella' term under which different structures can fit. The most classic examples, though, are those where a locative (2.a) or genitive (2.b) constituent becomes the sentence's grammatical subject. These are the two types of topic-subject constructions I will be studying here, and Chapter 2 is devoted to the relevant literature on this topic.

(2)   a. ***As ruas do centro*** *não estão passando carro.*

            the streets of.the downtown not are passing car

            'Cars are not passing through downtown streets.'

        b. ***As crianças*** *estão nascendo o dentinho.*

            the children are growing the tooth

            'Children's teeth are growing.'

What is most relevant for this research is that the chosen linguistic phenomena have been singled out as syntactic singularities of BP among other Romance languages. And this is precisely what makes both structures interesting phenomena to study from a second language acquisition perspective.

For example, Galves (1998) states that BP differs from the other Romance languages in that it allows topic-subject constructions. Locative and genitive constituents in these other languages occur in adverbial or adnominal position and are introduced by a preposition.

(3)   a. ***Por las calles del centro*** *no pas<u>a</u> <u>ningún coche</u>.*   (Spanish)

        b. ***Per le strade del centro*** *non pass<u>a</u> <u>nessuna machina</u>.*   (Italian)

        by the streets of.the centre not passes any car.

        'No car is passing through downtown streets.'

(4)   a. ***A mi hijo*** *le est<u>án</u> saliendo <u>los dientes</u>.*   (Spanish)

        b. ***A mio fliglio*** *gli sta<u>no</u> uscendo <u>i denti</u>.*   (Italian)

        to my son CL are coming.out the teeth

        'My son's teeth are coming out.'

But even if it can be argued that in very specific cases locatives can occupy the subject position (Fernandez-Soriano, 1999), they never trigger verb agreement, as can be seen in the following Spanish examples.

(5)   a. ***En esta clase*** *sobr<u>an</u> <u>estudiantes</u>.*

            in this classroom are.extra students

            'There are too many students in this classroom.'

        b. ***En Barcelona*** *h<u>a</u> ocurrido <u>un accidente</u>.*

            in Barcelona has happened an accident

            'There has been an accident in Barcelona.'

As for wh *in situ*, Ambar and Veloso (2001) point out that Portuguese stands out in showing *in situ* interrogatives that are true information-seeking questions and not just echo questions. And, although it is well known that French also allows *in situ* interrogatives (Cheng & Rooryck, 2000, among many others), in other Romance languages, like Spanish, on the other hand, these constructions are either only possible in echo questions or highly restricted to very specific contexts (highly presuppositional and maybe even involving a contrasting topic), as illustrated in this Spanish example from Etxepare & Uribe-Etxebarria (2005).

(6)   Speaker A: *Mi padre, mi madre y yo fuimos a comprar huevos, leche y café.*
                        Mi madre compró los huevos.
                        'My father, my mother and I went to buy eggs, milk and coffee.
                        My mother bought the eggs.'
        Speaker B: *¿Y tu padre compró qué?*
                        'And your father bought what?'

In fact, the initial idea for this research came from my own personal experience as a Spanish native speaker acquiring Brazilian Portuguese. Quite early in the process of acquisition, after only two months of prior formal training in Spain and another couple of months of living in Brazil, I found myself extensively producing wh *in situ* in Portuguese. Shortly after that, this syntactic structure began to cause attrition to my L1, something that was first noted by other Spanish native speakers who had no contact with BP and found the utterances strange enough to call my attention to them.

In contrast, after having been living in the country for over seven years, I do not think I have ever produced a topic-subject construction in a non-metalinguistic conversation. My first (conscious) encounter with these sentences happened after more than two years in the country and I have, since then, always found them to be quite odd.

Of course, being a Linguistics student, I do not believe that my experience as a BP learner is exactly the standard one. However, there was such a significant difference between the acquisition of those two syntactic structures in my case, that I could not help but think other Spanish native speakers must have had a somewhat similar experience. The deep strangeness I feel when I hear topic-subject constructions strongly contrasts with the ease with which I respond to, produce in BP and copy into Spanish *in situ* interrogatives. This led me to believe that there must be an underlying syntactic reason behind this difference and that a comparison between the acquisition of such structures was worth studying.

Research on second language acquisition (henceforth, SLA) can offer important insights about the human language capacity and it has helped move the field of linguistics forward. The different theories about SLA are briefly discussed in Chapter 3. As we shall see, this research's results can constitute evidence in favor of Full Access approaches, such as the Full Access Full Transfer Hypothesis introduced by Schwartz and Sprouse (1994); or in favor of Representational Deficit approaches, like the Interpretability Hypothesis proposed by Tsimpli (2003). This will be examined in Chapter 7, but, in any case, this study generated a fruitful discussion related to remarkably interesting SLA theoretical matters such as Universal Grammar access, linguistic transfer and feature strength resetting.

In addition to that, the two syntactic constructions are worthy of interest in themselves and have generated a regular research production from a purely theoretical perspective both in studies specifically about BP and in wider investigations about natural languages' syntactic properties. We shall see as well how the results of this study can contribute to the ongoing debates regarding these two structures.

But, recapitulating for a moment, the main question I would like this research to answer is whether there is a difference in the acquisition of *in situ* interrogatives and topic-subject constructions in BP as a second language. And, of course, if there is a contrast, I am interested to know why, what this difference is due to.

To answer the first question, I took an experimental approach. The core of the experiment designed for this purpose was an acceptability judgement task (AJT) of the BP syntactic constructions under study. The experiment was conducted on 24 Spanish native speakers, 26 BP native speakers and 4 people who declared to be bilingual. Each participant was shown 50 sentences, one at a time, and had to rate them on a five level Likert scale. The AJT offered off-line measurements which were complemented with on-line measurements by recording the reaction times. A follow-up task in which participants had to reformulate the rejected sentences was used to reveal the reason of rejection. The methodology outlined here is described in detail in Chapter 4, and the results, which confirmed our suspicions, are carefully analyzed and discussed in Chapters 5 and 7.

As for the second question, I explored the possible reasons behind the difference by taking the experiment results as the departure point of a theoretical discussion based on the existing proposals for the syntactic structures underlying the two constructions and on the relevant second language acquisition theories.

The rest of the text is structured as follows. In Chapter 1, I review some of the most relevant literature regarding wh BP interrogative paradigm with a special focus on the fact that both fronted and *in situ* questions are allowed in this language. Chapter 2 is devoted to topic-subject constructions. There, I discuss the theoretical proposals put forth to account for this singular BP syntactic structure. Chapter 3 offers a panoramic view on SLA. In this chapter, I briefly explain the most important hypotheses available in the field and I review some of them in a bit more detail through a selection of SLA research on wh fronting vs. *in situ* languages and on topic vs. subject prominent languages. In Chapter 4, I discuss the specific research questions of this study, formulate a hypothesis, and give a detailed description of the experimental design. Chapter 5 comprises an in-depth analysis of the results obtained from the experimental study. There, I examine the results quantitatively comparing the control and experimental groups and I provide a qualitative analysis of the individual results. In Chapter 6, I offer some empirical remarks underlining potential problems and suggesting improvements to the experimental design. Finally, Chapter 7 constitutes the conclusion of this work, in which I provide answers to the research questions posed in Chapter 4 and discuss the theoretical implications of the study results both for the field of SLA and for the pre-existing theoretical proposals concerning the syntactic structures under study.

PART I
Preliminaries

Chapter 1

Wh-movement and its optionality in Brazilian Portuguese

1.1. Introduction

Syntactic derivation of interrogative sentences is a topic profusely debated in generative grammar since the early stages of this framework. With the evolution of the theory, several different proposals have been put forward to explain the syntax behind these sentences and to try to account for cross-linguistic differences. In some languages, like Chinese, wh-phrases remain *in situ*, while in other languages, like English, wh-phrases are fronted. Compare (1) and (2) below.

(1)   *hufei mai-le **sheme***

         Hufei buy-ASP what

         'What did Hufey buy?'

(2)   ***What** did John buy?*

Languages that admit both fronted and *in situ* wh-phrases, like Brazilian Portuguese, constitute a challenge for this theory since wh-movement appears to be optional[1] in this case.

(3)   a. ***O que** o Pedro comprou?*

         the what the Pedro bought

      b. *O Pedro comprou **o que**?*

         the Pedro bought what

      'What did Pedro buy?'

In this chapter, I briefly review some of the generative literature on wh-movement in general, namely Cheng's *Clausal Typing Hypothesis* (1991) and Rizzi's *Q-Criterion* (1996). These proposals are at the core of the analyses for wh *in situ* in BP proposed in Mioto and Kato (2005),

---

[1] Here and throughout most of the chapter (unless otherwise specified), I will use a loose notion of optionality by which I only mean the possibility to either front wh-phrases or leave them *in situ* which some languages offer.

in Kato (2013), and in Grolla (2005), which constitute the focus of this chapter's discussion. The main goal of this review is to better understand wh-movement, but also to investigate how (or if) the existing proposals accommodate optionality in the particular context of interrogative sentences. I also look into optionality *per se*, exploring Biberauer and Richards's concept of *True Optionality* (2005) to see whether or not it applies to *in situ* interrogatives in BP. The chapter concludes arguing in favor of Grolla's proposal for BP by comparing it with Etxepare and Uribe-Etxebarria's analysis of wh *in situ* in Spanish (2005).

1.2. The foundations for the analyses in Brazilian Portuguese

After Chomsky's seminal work *On wh-movement* (1977), countless researchers took to the study of wh-movement and the many issues raised in that paper. The available literature on interrogatives (or even on *in situ* interrogatives alone) is so vast, that a comprehensive review goes way beyond our goals for this master's dissertation. Selecting the relevant proposals becomes a matter of knowing when to stop pulling the thread, quite a difficult task! Let me begin the chapter by focusing on the general analyses suggested in Cheng (1991) and in Rizzi (1996), because they were later used as the foundation for the research regarding BP that will be discussed in the following section.

These analyses generally assume the theoretical background leading up to Chomsky's *Minimalist Program* (1995). Minimalism postulates a computational system as economical as possible, where less is better. In this model, movement is necessarily triggered by feature checking, otherwise it cannot occur. And there are two types of features: strong or weak. Strong features are not phonologically acceptable and thus must be checked/deleted at PF, whereas weak features are acceptable at PF and, for economy reasons (do as less as possible), only need to be checked at LF. This contrast in feature strength gives rise to crosslinguistic differences.

So, in our case, what determines the need for the interrogative constituent to move to Spec-CP is the strength of the wh-feature (either on the wh-element or on C itself). If the feature is strong, then it needs to be immediately checked to grant the derivation's success and the movement will be overt. On the contrary, if the feature is weak, it does not block the derivation, so the feature can be checked in LF and thus the movement will be covert.

Languages with optional wh-movement are a challenge for the model, since, to accommodate them, the feature needs to be strong at times and weak at others. That is why many analyses, as

we shall see below, treat optionality as being only apparent, either considering wh *in situ* as the result of vacuous movement or considering fronted interrogatives as an instance of other syntactic structures in which the wh-phrase is base generated instead of having undergone movement.

1.2.1. *Cheng (1991) and Cheng and Rooryck (2000)*

In her doctoral thesis '*On the typology of wh questions*' (1991), Cheng argues that there are two types of languages: *in situ* languages, which never front wh-phrases, and languages with syntactic wh-movement, which never leave them *in situ*.

She begins by suggesting the following generalization:

(4)   *In situ* languages have special markings in yes-no questions and languages with special markings in yes-no questions are *in situ* languages.

The special marking is usually done by a particle, but it can also be done by another element or by a specific structure. It so happens that both options are available in Mandarin Chinese, as we can see in (5). There is a yes-no particle *ma* in (a), and a special structure that the author calls 'A-not-A' in (b).

(5)   a. *hufei hui lai ma*
          Hufei will come $Q_{Y/N}$
        b. *hufei hui-bu-hui lai*
          Hufei will-not-will come
        'Will Hufei come?'

Some of the languages that have yes-no particles also have wh-particles. In the case of Mandarin Chinese, the wh-particle is *ne* and it is optional.

(6)   *hufei mai-le na-yi-ben-shu (ne)*
        Hufei buy-ASP which-one-CL-book Q_{WH}
        'Which book did Hufei buy?'

Based on cross-linguistic evidence from several languages, as we can see in the table below, Cheng assumes there is a relation between yes-no particles and wh-particles. She claims that if a language has overt yes-no particles, then it also has wh-particles, although the latter can be overt or not.

(7)   Crosslinguistic evidence of yes/no and wh-particles

Language	Yes/No particle	Wh-particle
Hindi	kyaa	Ø
Palauan	special agreement	special agreement
Iraqi Arabic	hal	Ø
Egyptian Arabic	-ʃ / pronouns	Ø
Gulf Arabic	ʔidha / lo	Ø
Mandarin	ma / A-not-A	ne / Ø
Cantonese	A-not-A	a
Navajo	daʔ…(-ísh)	-lá / -sh
Papago	n-	Ø
Hopi	ya	ya
Japanese	no-(ka)	ka / (no)-ka
Korean	ci	ci
Indonesian	apa(kah)	Ø
Swahili	je	Ø
Amharic	wey	Ø
Lardil	kara	Ø
Turkish	-ma	Ø

Hence, the generalization in (4) can be reformulated as (8):

(8)   *In situ* languages have wh-particles. Languages with wh-particles are *in situ* languages.

In Cheng's proposal, wh-particles also serve to determine the interrogative force for certain wh-words. In some languages, wh-words have an ambiguous reading, they can have interrogative, existential or universal interpretations depending on the sentence. This is the case in Japanese. In (9), the wh-word *dare* ('who') can have an interrogative (a), a universal (b) or an existential reading (c).

(9)   a. *Dare-ga kimasu-ka*

           who-NOM come-Q$_{WH}$

           'Who's coming?'

     b. *Dare-ga kite-mo, boku-wa awa-nai*

       who-NOM come-Q I-NOM meet-not

       'Whoever comes, I will not meet (them).'

       ($\forall$ x, if x comes, I would not meet x)

     c. *Dare-kara-ka henna tegami-ga todoita*

       who-from-Q strange letter-NOM arrived

       'A strange letter came from god knows who.'

       ($\exists$ x : a strange letter came from x)

According to the author, languages like Japanese, in which wh-words are ambiguous, need to have overt wh-particles in order to resolve the ambiguity. Languages in which wh-words are non-ambiguous, that is, in which wh-words can only be interpreted as interrogatives, do not require the presence of overt wh-particles, so non-overt ones are allowed.

Exploring what could the particles' function be leads Cheng to adopt the *Clausal Typing Hypothesis* formulated by Chomsky and Lasnik (1977), which states that every clause must be typed at S structure. With the *Clausal Typing Hypothesis* there is no need to posit a Q-morpheme or [+wh] feature in C for languages like English, an assumption made in most of the previous work on wh-movement.

In particular, Cheng claims that, for wh-interrogative clauses, the typing may be done either by a particle in C^0 or by fronting a wh-phrase to Spec-CP. Since insertion of a particle is more economic than movement, a language with this option must always use it. Thus, wh-movement occurs as a last resort to type the sentence in languages for which the option of particles is not available.

To illustrate this, she compares a multiple question in English with a single one in Mandarin Chinese:

(10)   a. [$_{CP}$ Who$_i$ [$_{IP}$ t$_i$ bought what]]?

       b. qiaofong mai-le sheme ne

         Qiaofong buy-ASP what Q$_{WH}$

         'What did Qiaofong buy?'

In the English example, the wh-word *who* has moved to Spec-CP to type the sentence, thus blocking the movement of *what*. This is very similar to what happens in the Chinese example,

where the wh-word *sheme* ('what') stays *in situ* because the wh-particle *ne* has already typed the sentence.

Concerning acquisition, the author claims that a child acquiring a language with yes-no particles will assume there is also a wh-particle (whether overt or null), since these children never front wh-words. On the other hand, a child acquiring a wh-movement language has no evidence to assume the existence of such particles, which is supported by the early age in which these children show syntactic wh-movement (at around age 3).

Cheng (1991) leaves no room for optionality: a language that admits both fronted and *in situ* wh-phrases would constitute a counterexample for her proposal. She analyzes three languages that despite having a wh-particle also show fronted interrogatives: Egyptian Arabic, Bahasa Indonesia and Palauan. The author argues that apparent fronting of wh-phrases in these optional fronting languages does not involve wh-movement.

In her analysis, based on the form of the complementizer and on island violations and other syntactic phenomena, fronted interrogatives are nothing but reduced cleft sentences (when the wh-phrase is an argument) or an instance of topicalization (when the wh-phrase is an adjunct).

In order to illustrate this proposal, I will now try to summarize her argumentation for Egyptian Arabic. This language allows wh-phrases to remain *in situ* in both main and embedded clauses.

(11)   a. *Fariid hawil yi'mil <u>eeh</u>*
             Fariid tried to.do <u>what</u>
             'What did Fariid try to do?'
         b. *Mona 'irfit Ali haawil yisaafir <u>feen</u>*
             Mona knew Ali tried to.travel <u>where</u>
             'Mona knew where Ali tried to go.'

But it also shows fronted interrogatives. Therefore, this language is a potential counterexample for Cheng's proposal.

To show these fronted interrogatives are not the result of wh-movement, Cheng starts by calling the reader's attention to the similarities between this kind of interrogative sentences and relative clauses or clefts. The complementizer *illi* appears in all three types of sentences, however it is

not the typical complementizer used in embedded clauses[2].

(12)   a. *il-raagil **illi** Mona shaafit-uh*   Relative clause

          the-man that Mona saw-him

          'the man that Mona saw'

    b. *(dah) Muhamad **illi** gih*   Cleft

          this Mohammed that came

          'It is Mohammed who came.'

    c. *miin **illi** Mona darabit-uh*   Fronted interrogative

          who that Mona hit-him   (wh-argument)

          'Who did Mona hit?'

Notice, as well, that, just as it happens with relative clauses, fronted interrogatives require the presence of a resumptive pronoun without which the sentence becomes ungrammatical.

However, this only applies to argumental wh-phrases. When the wh-phrase is an adjunct, the author notices the sentence does not resemble relatives or clefts (there is no resumptive pronoun and no complementizer *illi*), but topicalization of a DP instead.

(13)   a. *fi-l-shari'dah, Mona kaanit bitdawwar'ala sha??ah.*   Topicalization

          on-the-street.DEM Mona was looking.for apartment

          'On that street, Mona was looking for an apartment.'

    b. *ma'a miin Mona raahit il-Qahirah*   Fronted interrogative

          with whom Mona went to-Cairo   (wh-adjunct)

          'With whom did Mona go to Cairo?'

Cheng proceeds with her argumentation claiming that fronted interrogatives behave like relative clauses in that they violate island constraints. However, her argument is not a very strong one, since, according to Cheng herself, the acceptability of sentences like that in (14.b) varies among speakers.

(14)   a. *dah il-beet illi baba ye'raf il-raagil illi bana-ah*

          this the-house that father knows the-man that built-it

          'This is the house that my father knows the man who built it.'

---

[2] The typical complementizer in Egyptian Arabic embedded clauses is *inn*, as we can see in (i).

    (i)   *Mona iftakarit **inn** Fariid shaafir.*

          'Mona thought that Fariid left.'

b. *anhi kitab illi Mona te'raf miin illi saraʔ-uh*

which book that Mona knows who that stole-it

'Which book does Mona know who stole?'

Then, she proposes an analysis for fronted interrogatives[3] in terms of a reduced cleft structure. 'Reduced' because there is no copula nor an expletive subject, but still a 'cleft' because there is a subject-predicate relationship between the wh-phrase *miin* 'who' and the rest of the sentence. This means the wh-phrase is base generated in this position, and thus there is no wh-movement here.

(15) $[_{CP}[_{DP}$ miin$_i]$ $[_{CP}$ OP$_i$ illi $[_{TP}$ Mona shaafit-uh$_i]]]$

who   that   Mona saw-him

'Who did Mona see?'

As we will see in the next section, Kato and Raposo (1994) propose to account for Brazilian Portuguese interrogatives in a very similar fashion, considering wh-fronting as deriving from reduced clefts. I also discuss how Grolla (2005), however, argues against this analysis and suggests a different approach partially based on Cheng's Ph.D. dissertation.

Continuing this line of research, Cheng and Rooryck (2000) examine French wh *in situ*. Their proposal is mainly based on the raising intonation of wh *in situ* questions and posits a root intonational Q-morpheme (similar to the overt complex Q-morpheme *est-ce que*) which is responsible for licensing wh *in situ*: it checks the Q-feature in C^0 thus blocking overt movement of the wh-phrase (much like the wh-particle in Chinese).

They suggest that Q-morphemes in general can be either specified as a wh-question morpheme or as a yes/no morpheme, or underspecified and thus compatible with both types of questions. This is the case of *est-ce que* and of the intonational Q-morpheme. But their underspecification needs to be resolved, which can be done at S-structure or at LF. Following the assumption that only features move at LF (Chomsky, 1995), Cheng and Rooryck propose then that underspecified morphemes trigger movement of the wh-feature if there is one present, otherwise they are set to [yes/no] by default.

Let me show an example to illustrate this proposal. Consider the following pairs of sentences:

---

[3] This analysis only applies to fronting of argumental wh-phrases. For adjuncts, the author does not supply a specific structure, though what she suggests is an analysis analogous to that of topicalization.

(16)   a. *Est-ce que Jean a acheté un livre?*

EST-CE QUE Jean has bought a book

'Did Jean buy a book?

b. *Quel livre est-ce que Jean a acheté?*

which book EST-CE QUE Jean has bought

'Which book did Jean buy?'

(17)   a. *Jean a acheté un livre?*   (rising intonation)

Jean has bought a book

'Did Jean buy a book?'

b. *Jean a acheté quoi?*   (rising intonation)

Jean has bought what

'What did Jean buy?'

In both cases there is an underspecified Q-morpheme. In (16.a), the morpheme *est-ce que* is set to [yes/no] by default, while in (16.b) it is set to [wh] by overt movement of the wh-phrase. Similarly, in (17.a) the intonational Q-morpheme is set to [yes/no] by default, while in (17.b) it is set to [wh] at LF by movement of the wh-feature in the wh-phrase (only features can move at LF). It is important for their argument to clarify that this movement is not to check the Q-feature in C^0 (this is done by the intonational morpheme itself) but to disambiguate the underspecified intonational Q-morpheme.

As we can see, there is a difference between *est-ce que* and the intonational Q-morpheme: The former triggers overt movement of wh-phrases while the latter does not. Cheng and Rooryck attribute this difference to their checking relation with C^0. They claim that the intonational Q-morpheme is merged as an adjoining head to a null C^0 (like the wh-particle in Chinese) thereby checking its Q-feature and blocking overt wh-movement. On the other hand, the morpheme *est-ce que* is an instantiation of the C^0 feature itself, therefore needing to be checked. This is done by overt movement of a wh-phrase to Spec-CP in (16.b), but their explicit proposal for (16.a) is not evident. From their argument, it appears that the feature is checked by merging a X^0 in C^0, although it remains unclear what would C^0 be in this case.

On the topic of optionality, the authors conclude that given their analysis, optional wh-movement in French is only apparent, attributing it to different numerations. If the intonational

Q-morpheme is in the numeration, then it is merged in C^0 and the wh-phrase cannot undergo movement; otherwise movement is necessary to check the Q-feature.

Cheng and Rooryck mention European Portuguese, but only to show that wh *in situ* in this language differs from that of French, since it is also allowed in indirect questions.

(18)   *O João perguntou se tu compraste o que?*

the João asked whether you bought the what

'Did João ask what you bought.' / 'What did Joao ask if you bought?'[4]

They compare *se* ('whether') with the intonational Q-morpheme since both are underspecified for [yes/no] or [wh] features. However, nothing is said about the analysis of matrix *in situ*.

In section 1.3.3., I discuss the analysis proposed in Kato (2013) for Brazilian Portuguese. She compares French and BP interrogatives to show that Cheng and Rooryck's (2000) analysis can also be applied to echo questions and yes/no questions but she argues that a different approach is needed for ordinary wh *in situ*.

1.2.2. Rizzi (1996) and the Q-Criterion

To explain T to C movement in interrogative sentences that languages like English and most modern Romance languages show (referred to as 'residual V2'), Rizzi (1996) formulates the *Q-Criterion*. This criterion is a general well-formedness condition on interrogatives which he claims is responsible for the S-structure distribution and the LF interpretation of Q-operators.

The criterion states the following:

(19)   *A Q-operator must be in a Spec-head configuration with a Q-head and, in turn, a Q-head must be in a Spec-head configuration with a Q-operator[5].*

Rizzi claims that the criterion can be applied either at S-structure or at LF depending on the language: for example, in English it applies at S-structure whereas in Chinese it applies in LF. Based on evidence from languages in which verbs have special morphology in interrogatives, he proposes that the Q-feature can be anchored in T.

---

[4] The original BP sentence is ambiguous and admits both readings. Note that the second reading is not a grammatical sentence in English beacause it involves extraction from a wh island.

[5] In his original paper, Rizzi (1996) states the *wh-criterion* in terms of wh-heads and wh-operators. I follow here a more modern formulation presented in Mioto and Kato (2005).

This accounts for English subject-aux inversion in interrogatives. Consider the following interrogative sentence and its ungrammatical versions:

(20)   a. Who has Mary seen?

    b. *Mary has seen who?

    c. *Who Mary has seen?

    d. *Has Mary seen who?

In (20.a), the Q-head *has* moved from T to C and the Q-operator *who* moved from the internal argument position to Spec-CP, so they are in a Spec-head relation, hence satisfying the Q-criterion. If none or only one of them moves, the Spec-head relation is not created, thus violating the criterion, which explains the ungrammaticality of the rest of the examples in (20). The proposed underlying structure is given in (21).

(21)   [$_{CP}$ who$_i$ has$_{j[+Q]}$ [$_{TP}$ Mary t$_j$ seen t$_i$]]

The Q-criterion also accounts for indirect questions.

(22)   a. I wonder who Mary has seen.

    b. *I wonder Mary has seen who.

    c. *I wonder has Mary seen who.

    d. *I wonder who has Mary seen.

In this case, T to C movement is not allowed because the C selected by the main verb is already filled by a phonetically null Q-head. For the Q-criterion to be satisfied, this head needs to be in a Spec-head relation with a Q-operator, thus triggering the movement of the wh-phrase to Spec-CP of the embedded clause.

(23)   I wonder [$_{CP}$ who$_i$ $\emptyset_Q$ [$_{TP}$ Mary has seen t$_i$]]

This analysis presents some problems when the questioned element of the interrogative sentence is the subject. T to C movement is not allowed in this case (24.b), but then, how can the Q-criterion be satisfied?

(24)   a. Who loves Mary?

    b. *Who does love Mary?

A possible explanation would be to assume that the subject remains in TP, where the Q-criterion would be easily satisfied. However, Rizzi points out some technical issues with this simpler

analysis, which I will not discuss here, and concludes the structure in (25) constitutes a better option.

(25) $[_{CP}$ who$_i$ C $[_{TP}$ t$_i$ T $[_{VP}$ t$_i$ love-s$_{[+Q]}$ Mary$]]]$

To explain how this does not result in a violation of the Q-criterion, the author extends its interpretation claiming that it is the chain of the relevant head that needs to have the feature [+Q] and not the position itself. In this case, C, T and the lower inflection containing the feature [+Q] form a chain, part of which is in a Spec-Head configuration with the Q-operator *who*, hence the criterion is satisfied at S structure.

Rizzi continues his paper by looking into wh *in situ* in English, which leads him to reconsider the definition of Q-operators. The Q-criterion straightforwardly accounts for the ungrammaticality of wh *in situ* in single questions. But again, it encounters some problems in multiple questions, where one of the operators has been moved to C and the others have remained *in situ*.

(26)   a. What did you give to whom?

b. $[_{CP}$ what$_i$ did$_{j[+Q]}$ $[_{TP}$ you t$_j$ give t$_i$ to **whom**$]]$

If the criterion applies in S-structure in English, the Q-operator that remains *in situ* violates the criterion (since it is not in a Spec-head configuration with a Q-head). Rizzi solves this by considering that a wh-phrase only qualifies as a Q-operator when it is in an A'-position[6]. This means the *in situ* wh-phrase in (26) is not a Q-operator and thus there is no need for it to satisfy the Q-criterion.

According to the author, the proposed solution is independently justified since it solves another problem as well, one that wh-phrases pose for Theta Theory. Consider the following question, with its D-structure (27.a) and its S-structure (27.b):

(27)   a. Mary saw whom

b. Who$_i$ did Mary see t$_i$

In (27.a), the verb *see* assigns an object theta role to the wh-phrase, hence it is an argument at D-structure. However, since the Theta Criterion applies at both levels, D-structure and S-structure, this theta role is assigned to its trace in (27.b), hence the wh-phrase is not an argument

---

[6] Rizzi then proposes an ulterior refinement by requiring the wh-phrase to be in a scope position, see the original paper for more details and a deeper discussion.

at S-structure. In other words, the status of the same element varies from one level to the other. But the refined notion of Q-operator solves this problem. Since the wh-phrase is in an A-position at D-structure, it does not qualify as an operator, it is an argument and receives the object theta role. At S-structure, the wh-phrase is in an A'-position so it qualifies as an operator, it is not an argument and the object theta role can be assigned to its trace.

Rizzi also explores the interrogative paradigm in French, which, as can be seen below, is much less restrictive for main clauses but just as restrictive for embedded clauses if compared to the English equivalent, see (20) and (22) respectively.

(28) a. *Elle a rencontré qui?*
        she has met who

      b. *Qui elle a rencontré?*
        who she has met

      c. **A-t-elle rencontré qui?*
        has-she met who

      d. *Qui a-t-elle rencontré?*
        who has-she met
      'Who did she meet with?'

(29) a. **Je ne sais pas elle a rencontré qui.*
        I NEG[7] know NEG she has met who

      b. *Je ne sais pas qui elle a rencontré.*
        I NEG know NEG who she has met

      c. **Je ne sais pas a-t-elle rencontré qui.*
        I NEG know NEG has-she met who

      d. **Je ne sais pas qui a-t-elle rencontré.*
        I NEG know NEG who has-she met
      'I don't know who she met with.'

Assuming the Q-criterion applies at LF in French would explain the facts regarding main questions, in (28), but not those regarding embedded questions, in (29). To deal with this, he proposes a "*dynamic agreement*" process (available for French in syntax as well as in LF) in

---

[7] Merely for ease of exposition, here I use NEG both for *ne* and *pas* in French. However, the reader should bear in mind that these two elements differ in several aspects.

which "*the specifier is able to endow the head with the relevant feature specification*" (Rizzi, 1996, p.76).

(30) $Q_{op} \, X \implies Q_{op} \; \dfrac{X}{[+Q]}$

So, in an *in situ* interrogative sentence, like (28.a), there is no violation of the Q-criterion at S-structure because there is no Q-head and the wh-element is in an A-position, so it is not an operator. At LF, the wh-element moves to Spec-CP for interpretive reasons, there it becomes a Q-operator[8] and by *dynamic agreement* endows C with the feature [+Q], thus verifying the Q-Criterion. A similar argument is used to account for the grammaticality of fronted interrogatives without T to C movement in French, like (28.b). There is no Q-head either but the moved wh-phrase at S-structure endows C with [+Q] feature through *dynamic agreement* hence ensuring the satisfaction of the criterion both at this level and at LF. According to Rizzi, none of these options is available in English due to lack of *dynamic agreement*.

The well formedness of (28.c) is analyzed as its equivalent in English. T is a Q-head and T to C movement grants satisfaction of the criterion at S-structure and at LF. As for the ungrammaticality of (28.c), it is due to a violation of the Q-criterion. In this case, T is a Q-head but there is no Q-operator that can enter in a Spec-head configuration with it.

Rizzi justifies the unavailability of *dynamic agreement* in the embedded paradigm (29) in the following manner. The embedded C is lexically selected by the main verb with its Q-feature set to [-Q] at D-structure, therefore, it cannot be changed later (due to the Projection Principle).

I find this analysis unsatisfying for several reasons. It does not explain what triggers some of the syntactic movements assumed to occur. The author does not say why the wh-phrase moves when there is no T to C movement in French. In addition to that, in fronted interrogatives with T to C movement and in main interrogative clauses in English, it is unclear which of the movements occurs first (T to C or wh-phrase to C) and why it occurs. It seems reasonable to assume T to C happens first and then this movement triggers wh-movement, but Rizzi does not make this explicit. In addition to that, the proposal does not seem to truly account for the French interrogative paradigm. All the different types of interrogative sentences satisfy the criterion, but the paper does not explain how they are generated and why there are three different options. Finally, and maybe most importantly, the concept of *dynamic agreement* is quite problematic. If specifiers can endow heads with the relevant feature specification, then there would be no

---

[8] Because Spec-CP is a scope-position (see footnote 5).

justification for any element to move to this position. It seems like a provable source of overgeneration.

As we are about to see in next section, Rizzi's proposal is at the core of Mioto (1994) and also Kato and Mioto (2005), hence these analyses for Brazilian Portuguese inherit the same problems I just discussed.

1.3. Some proposals for Brazilian Portuguese

Now that the basis on which the selected literature on BP stands are clearer, I can move on to reviewing and discussing some of the proposals put forth for this language's interrogative paradigm.

Since my interest resides on BP wh *in situ*, I chose to work with three different proposals that are representative of the research on this particular topic, considering the controversy on wh movement optionality in this language. On one hand, Mioto and Kato (2005), and Grolla (2005) both follow the *in situ* approach, that is, they support the idea that the wh constituent in *in situ* interrogatives really does remain *in situ*. Their proposals, as we are about to see, are however quite different from one another. On the other hand, Kato (2013) argues that BP wh *in situ* is only apparent, thus following the movement approach.

1.3.1.   *Mioto (1994) and Mioto and Kato (2005)*

Mioto (1994) focuses on Brazilian Portuguese interrogatives and applies a slightly earlier version of Rizzi's *Q-Criterion*. Then, Mioto and Kato (2005) extend this idea comparing Brazilian with European Portuguese, but their proposal for BP remains essentially the same. They argue that, with some supplementary considerations, the *Q-Criterion* accounts for the differences between European and Brazilian Portuguese with respect to their interrogative paradigms.

Let me start by stating these differences. In BP, there is no subject-verb inversion in fronted interrogatives, the subject precedes the finite verb yielding the order QSV. In EP, however,

there is subject-verb inversion yielding the order QVS. The authors claim these restrictions can be relaxed in embedded questions[9].

(31)   Main clause

a. *Como reagiu o Adriano Pinto?*   EP   *BP

how reacted the Adriano Pinto

b. *Como o Adriano Pinto reagiu?*   *EP   BP

how the Adriano Pinto reacted

'How did Adriano Pinto react?'

(32)   Embedded clause

a. *O João perguntou como reagiu o Adriano Pinto.*   EP   ?BP

the João asked how reacted the Adriano Pinto

b. *O João perguntou como o Adriano Pinto reagiu.*   EP   BP

the João asked how the Adriano Pinto reacted

'João asked how Adriano Pinto reacted.'

In addition to this, both languages allow clefts in main and embedded questions. But BP allows doubly filled Comp whereas EP does not.

(33)   a. *Onde é que você estava em 82?*   EP   BP

where is that you were in 82

b. *Onde que você estava em 82?*   *EP   BP

where is you were in 82

'Where is it that you were in 82?

Also, both languages have wh *in situ*, but its status is quite different. According to the authors, only in BP are *in situ* interrogatives true questions. The authors also claim that wh *in situ* is only allowed in main questions[10] and it does not show subject-verb inversion nor a realized complementizer.

---

[9] Notice that in this case the main verb needs to select an interrogative clause. Verbs that select declarative clauses like *achar* 'think/believe' do not allow these constructions.

[10] Notice that the sentence in (33.b) can become grammatical if the main verb selects a declarative clause, like *achar* 'think/believe'. However, in this case the sentence can only be interpreted as a main question:

    O João acha que você saiu de lá como?

    the João thinks that you got.out from there how

    'How does João think that you got out from there?'

(34)   a. *Você saiu de lá como?*   EP   BP

you got.out of there how

'How did you get out of there?'

b. **O João perguntou que você saiu de lá como.*   *EP   *BP

the João asked that you got.out of where how

'João asked how you got out of there.'

To explain all this, Mioto and Kato propose that in EP, the Q-feature anchored in T is specified as [+Q], whereas in BP, it is specified as [-Q].

This straightforwardly accounts for the word order QVS in EP main questions, see (31), and cleft questions, see (33.a), and it also explains the possibility of QSV in EP embedded clauses like (32.b): Considering that the main verb selects an interrogative clause, then C is already a Q-head (so T to C movement does not take place) and the Q-criterion is satisfied.

In EP, wh *in situ* like (34.a) should not be possible according to the author's assumption of T being [+Q]. They account for this fact by arguing that it is necessary to allow the Q-feature in T to be [-Q] in these cases. This optionality is justified saying that the discomfort it generates is reduced considering that EP wh *in situ* is not very frequent and it has "*an echo flavor*". The ungrammaticality of wh *in situ* in embedded questions (34.b) is due to a violation of the Q-criterion. The main verb selects a C with the feature [+Q] but there is no Q-operator (recall that a wh-phrase in an A-position does not qualify as such) to enter a Spec-head configuration with C.

As for BP, the assumption of T being [-Q] justifies the grammaticality of *in situ* interrogatives in this language, see (34.a). In S-structure, there is no Q-head and no Q-operator. However, to ensure the Q-criterion is satisfied also in LF, Mioto and Kato resort to *dynamic agreement*. At this level, the wh-phrase moves to Spec-CP, becomes an operator and, under *dynamic agreement*, endows C with the feature [+Q].

The QSV order in BP main questions, as seen in (30), also needs some clarification. Since T is assumed to be [-Q] in this language, it is not unreasonable to conjecture that C could be [+Q]. However, as Mioto and Kato argue, if we assume that this is done at the base (positing a [+Q] phonetically null C), it would constitute a problem for the analysis of wh *in situ*. So, they invoke *dynamic agreement* again (this time in S-structure) to ensure that the criterion is satisfied.

They also propose that in BP, C [+Q] can be lexically realized by the overt complementizer *que*. This explains the possibility of a sentence like (33.b) in which the complementizer is base generated and then triggers wh-movement to satisfy the Q-criterion. But to account for BP cleft interrogatives, as (33.a), they need again *dynamic agreement* for the criterion to be satisfied, because T is [-Q] and thus the verb *ser* ('to be') stays in T.

Finally, Mioto and Kato suggest that the possibility of the unexpected QVS order in both BP and EP embedded questions, as seen in (32.a), has nothing to do with the Q-criterion, it is a consequence of the focusing process. Hence, this optionality is treated as being only apparent.

This analysis raises some of the same problems that Rizzi (1996) did. It relies too much on *dynamic agreement*, which, as I have already mentioned, seems to be a dubious concept. If the moved wh-phrase endows a null C with the feature [+Q], what triggers the movement of this wh-phrase in the first place? Why doesn't it remain *in situ*? Or, if there is no need for a base-generated C to ensure wh-movement then, why doesn't the wh-phrase move in *in situ* interrogatives as well? The analysis, just like Rizzi's, does not really seem to accommodate this optionality. But nothing in these terms, about the option of moving the wh-phrase or leaving it *in situ*, is mentioned in Mioto (1994) or Mioto and Kato (2005).

1.3.2. *Grolla (2005)*

Grolla (2005) takes a different approach to the study of wh-movement in Brazilian Portuguese by bringing language acquisition to the center of the discussion. She argues that, in light of BP's behavior with regard to interrogative sentences, the general theory of feature strength should be loosened. Partially assuming Cheng (1991) and the Cue-based theory of acquisition introduced by Lightfoot (1999), she proposes to treat BP as a language with optional [+wh movement]. Her aim is to account for the late acquisition of wh *in situ*, when compared to fronting.

As a departure point, Grolla argues against Kato and Raposo (1994), who proposed an analysis very similar to Cheng's account of Egyptian Arabic (see section 1.2.1, examples (11 – 15)). They claim that fronted interrogatives in BP are derived from cleft sentences. First, consider a declarative cleft sentence and its underlying structure:

(35)   [FP [TP pro copula [α$_i$] [DP (o N)$_i$ [CP (que)$_i$ [TP … t$_i$ …]]]]]

   É Maria que me deu o CD.

   is Maria that CL gave the CD

   'It is Maria who gave me the CD.'

Then, the derivation process of a fronted interrogative from a cleft, as proposed by the authors, is the following:

(36) i. *É quem que me deu o CD?*   Cleft structure
          is who that CL gave the CD

      ii. *Quem é que me deu o CD?*   Overt raising to Spec-FP
          who is that CL gave the CD

      iii. *Quem que me deu o CD?*   Erasing of the copula
          who that CL gave the CD

      iv. *Quem me deu o CD?*   Erasing of the complementizer
          who CL gave the CD

Adopting this analysis for BP interrogatives and assuming Cheng (1991), if fronting is not a result of wh-movement, this suggests that BP should be classified as an in-situ language. But Grolla notes that this analysis is problematic because BP does not have a wh-particle, which is a necessary condition for being an in-situ language in Cheng's generalization (as we saw in (8), section 1.2.1).

Grolla also uses acquisition evidence to point out another problem for this analysis. Since the analysis assumes fronted interrogatives are derived from clefts, the latter should be acquired before the former, but that is not the case. Grolla shows exactly the opposite occurs in her data[11]. The child in her study presents a productive use of fronted interrogatives like *O que você quer?* ('what do you want?') as soon as age 2;3[12], whereas the first cleft of the type *O que é que você quer?* ('what is it that you want?'), from which the other one should derive, appears much later, at age 3;2. This is a fact that strongly weakens Kato and Raposo's analysis.

However, given the possibility of wh *in situ* and how common it is in adult speech, BP cannot be simply considered a fronting language either. Grolla then proposes that the [wh] feature in wh-phrases of this language can be both strong or weak. To back this idea, she presents more

---

[11] Longitudinal study (53 sessions) consisting of audio recordings of a child, from age 2 to age 4.
[12] Child's age: (years; months)

acquisition data showing that *in situ* interrogatives are acquired later than fronted ones. The sentences below were produced by the child in Grolla's study at the age in parentheses.

(37)   Fronted interrogatives

    a. ***Que*** *é isso aqui?*   (2;2)

       'What is this?'

    b. ***Dõndi*** *vai?*   (2;5)

       'Where do you go?'

    c. ***Porque*** *você tá fazendu assim?*   (2;6)

       'Why are you doing it like that?'

    d. ***Qual*** *que o pai fez?*   (2;11)

       'Which one did dad do?'

(38)   *In situ* interrogatives

    a. *Eu quero brincar* ***com quê?***   (3;11)

       'I want to play with what?'

    b. *Parece* ***quem*** *da novela?*   (3;11)

       'It looks like whom from the novel?'

    c. *Eu tô brincando sabe* ***com quem?***   (3;11)

       'I'm playing you know with whom?'

The author also brings the following table with adult statistics[13] compared to child statistics (from her study) to show how the status of wh *in situ* in child grammar is quantitatively different from that of adult grammar.

(39)   Adult vs. child quantitative production of interrogatives

Type		NURC[14]	TV	Child
Fronted	*O que você comprou?*	29%	30%	66%
Cleft	*O que é que você comprou?*	21%	19%	29.4%
Double comp.	*O que que você comprou?*	37.5%	18.6%	2.9%
In situ	*Você comprou o que?*	**12.5%**	**32.4%**	**1.7%**

---

[13] The adult data comes from Lopes-Rossi (1996).

[14] NURC stands for *Norma Urbana Linguística Culta* '*Cultured Linguistic Urban Norm', a Project that collected one of the biggest Brazilian Portuguese speech corpora.*

Grolla then gives a theoretical account for these acquisition facts, which further supports her optionality-based proposal. The author assumes the Cue-based theory of acquisition, first proposed by Lightfoot in the late nineties. That is, for every parameter to be set in accordance with the input, there is a *cue* that triggers this setting. *Cues* need to be unambiguous, robust and salient in the child's input. She also assumes that Cheng's generalization is somehow present in UG: the child will look for a *cue*, in this case an interrogative particle in C^0, that will tell her if the language she is acquiring is an *in situ* language. If she finds it, she will set the parameter as [-wh movement], if she does not, she will set it as [+wh movement].

A child acquiring BP will not find a C^0 filled with a particle, so she will first consider the language as being [+wh movement] and all the questions she produces will be fronted. Later on, though, children start producing wh *in situ.* Grolla proposes that the presence in the input of a special empty category such as the one in (40) functions as a *cue*.

(40)   a. *Quem que você gosta* [PP ec]*?*

'Who (is it) that you like?'

b. *O que você vai precisar* [PP ec] *para fazer o bolo?*

'What are you going to need to make the cake?'

c. *Esse livro eu vou precisar* [PP ec] *pra semana que vem.*

'That book I am going to need (it) next week.'

d. *O João eu gosto* [PP ec].

'(It is) João I like.'

She assumes that the wh-element and the topic in these sentences are base-generated, they have not undergone movement, otherwise they would have carried the preposition with them. So, when the child encounters this special empty category, she perceives there are structures without movement in BP, and this makes her question the previous marking of the parameter for wh movement.

According to the author, this *cue* will lead the child to reevaluate the input. Perceiving the possibility of A' dependencies without movement to the left periphery of the sentence, the child is able to analyze wh *in situ* as structures without movement. Her first decision is then reconsidered, and the parameter is set to [±wh movement], its value in the adult grammar. This explains why in situ interrogatives are acquired much later than fronted ones despite their high frequency in adult's speech, as we saw in table (39).

Maybe admitting this kind of loosening of the feature driven approach to movement could solve the problem of French interrogatives as well. I suggest that we would need to consider T as being [±Q] and wh-phrases as being [±wh], but a more detailed study would go beyond the limits and purpose of this dissertation.

While purely theoretical discussions about adult grammar can be extremely interesting and provide very valid analyses, they often neglect the important role that language acquisition can have in the developing of a theory of syntax. When assessing a formal proposal for a syntactic phenomenon, we should always keep in mind to ask ourselves how children acquire it. Grolla (2005) is one of the analyses for BP that takes acquisition into account. As she, herself, notes in the conclusion of her paper, her analysis could be strengthened if acquisition data from the three optional movement languages studied by Cheng (1991) were to show that in those cases children produce first wh *in situ*, then clefts and then apparent fronting.

Grolla's proposal is particularly appealing to me for several reasons. First, she works from a language acquisition point of view and I too want to look at wh *in situ* through that same glass, although in her case it was L1 acquisition and here it is L2 acquisition. Second, she takes the *in situ* approach, which means she considers there is no movement of the wh-word in *in situ* interrogatives. Her analysis is more interesting for this research because it can provide an explanation for the difference in L2 acquisition of wh *in situ* and topic-subject constructions. And third, but perhaps most important, Grolla (2005) establishes a possible connection between these two syntactic structures through her idea of the special empty category in (38) functioning as cue for wh *in situ*.

However, the main problem for this proposal is that, as I mentioned in the beginning, this optionality or loosening of feature strength, does not fit well within the principles of generative grammar. Assuming the possibility of a parameter [± wh movement] opens the door to extend this optional setting for other features, which could also be a source of overgeneration.

1.3.3. _Kato (2013)_

Kato (2013) avoids what she calls "*the uncomfortable problem of optionality*" proposing an integrated approach to both fronting and wh *in situ*. She claims that wh-movement in Brazilian Portuguese is not to the root CP/FP but to an internal FP position, at the edge of vP, in Belleti's

(2004) sense, as we will shortly see. In her proposal, echo-questions are the only ones in which the interrogative constituent really remains *in situ*.

Kato notes that, in BP, yes/no questions and echo-questions have rising intonation, while what she calls 'ordinary' questions (which I have been referring to as wh *in situ*) and 'fronted' questions have falling intonation. As we are about to see, this plays an important role in her argument.

The rising intonation is given by a silent operator Q in C, like the intonational Q-morpheme posited in Cheng and Rooryck (2000), which serves to type the clause. In the case of yes/no questions, she argues that the whole TP then moves to Spec-TP, as can be seen in the proposed structure (41.b) for the sentence in (41.a).

(41)   a. *O professor chegou?*   ↑   (rising intonation)
          the professor arrived
          'Has the professor arrived?'
        b. [$_{CP}$ [$_{TP}$ O professor chegou]$_i$ [$_{C'}$ Q [$_{TP}$ t$_i$]]]] ?

For echo-questions, Kato proposes that they are elliptical forms of explicit performative main clauses with embedded indirect speech:

(42)   a. *O professor viu quem?*   ↑   (rising intonation)
          'The professor saw who?'
        b. [$_{CP}$ [$_{TP}$ ~~você disse~~ [que o professor viu quem]]$_i$ Q [$_{TP}$ t$_i$]] ?
                ~~you said~~ that the professor saw who
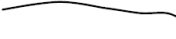
However, the author does not explain how the complementizer *que* ('that') is deleted to obtain the eco question in (42.a). This poses a problem for any analysis in terms of ellipsis. In addition to that, in BP, a phonetically realized complementizer makes wh-movement to Spec-CP obligatory.

In the remaining two types of interrogatives, the falling intonation is attributed to different reasons. For what she calls 'ordinary' questions (wh *in situ* to me), the author assumes the following (like Miyagawa (2001)): First, that the wh-phrase is associated both with wh and Q features and thus needs to end up in Spec-CP to satisfy C's EPP. And second, that the wh-phrase moves before to a Focus internal position at the edge of *v*P (following Belletti (2004)). The focused element requires nuclear stress and a falling prosody (given by Ø$_{+F}$), canceling the

effect of the operator Q that typed the sentence as a question. Kato shows the two steps of the derivation, as seen in (43.b) and (43.c).

(43)   a. *Você viu quem?*   ↓   (falling intonation)

          'You saw who?'

          b. [$_{TP}$ você viu [$_{FP}$ quem [$_{vP}$ t$_{você}$ t$_{viu}$ [$_{VP}$ t$_{viu}$ t$_{quem}$]]]] ?

          c. [$_{CP}$ [$_{TP}$ você viu [$_{FP}$ quem [Ø$_{+F}$]]$_i$ [$_{C'}$ Q [$_{TP}$ t$_i$]] ?

She then gives some arguments in favor of this analysis. In Belleti's proposal, the Focus internal position is the landing site of postposed subjects. Since Kato suggests it is also the landing site for the wh-element in *in situ* interrogatives, intonational similarities between the two structures are to be expected. The author provides phonetic evidence of these similarities, as can be seen in (44).

(44)   a.   b.

          Você viu quem? ↓ (falling int.)   *Chegou a Maria.* ↓   (falling int.)

          you saw who   arrived the Maria

Another piece of evidence in favor of her analysis is that *in situ* questions only admit postposed subjects when the verb is not transitive. See the contrast between (a) and (b, c) below.

(45)   a. *Telefonou quem?* ↓

          telephoned who

          'Who called?'

          b. **Compraram o que os meninos?*

          bought the what the boys

          'What did the boys buy?'

          c. **Comprou os CDs quem?*

          bought the CDs who

          'Who bought the CDs?'

As a final argument, Kato shows that, contrary to declaratives but just like in fronted interrogatives, when questioning multiple adverbial adjuncts, the *in situ* wh-phrases appear in coordination (suggesting that movement has taken place).

(46)   a. *A Maria nasceu na Bahia em 1980.*

          'Maria was born in Bahia in 1980.'

  b. *Onde e quando nasceu a Maria?*

   'When and where was Maria born?'

  c. *A Maria nasceu onde *(e) quando?*

   the Maria was.born where (and) where

   'When and where was Maria born?'

For 'fronted' interrogatives, Kato proposes an analysis similar to the proposal in Cheng (1991) for Egyptian Arabic and to the one in Kato and Raposo (1994) for BP. She argues that these sentences derive from cleft constructions. I will try to summarize her argumentation below.

First, consider the cleft question in (47.a):

(47) a. *Onde é que você vai?*

   where is that you go?

  b. [$_{ForceP}$ Q [$_{FP}$ Onde [$_{TP}$ é [$_{VP}$ t$_{é}$ [$_{CP}$ que [$_{TP}$ você vai t$_{onde}$]]]]]

According to the author, in (47.a), the wh-word moved to Spec-FP in the sentence periphery from the complement position of the embedded VP after the complementizer *que* and the copula were merged, yielding the structure we see in (47.b).

Now, consider the 'fronted' interrogative with *que* in (48.a). Kato claims it is derived from the canonic or *in situ* cleft in (48.b) by deleting the copula at PF. The proposed structure for this canonic cleft is (48.c), in which the focus (that is, the wh-element) is moved to the low VP periphery instead of being moved to the sentence periphery. The copula, in initial position, can then be phonetically deleted, yielding the sentence in (48.a)

(48) a. *Onde que você vai?*

   where that you go

  b. *É onde que você vai?*

   is where that you go

  c. [$_{ForceP}$ Q [$_{TP}$ é [$_{FP}$ onde [$_{VP}$ t$_{é}$ [$_{CP}$ que [$_{TP}$ você vai t$_{onde}$]]]]]

As an argument in favor of this analysis Kato presents acquisition evidence of a child producing the above-mentioned cleft construction at a very young age (2;3). This data, though, comes a specific dialect in Bahia and it may not be representative of how interrogatives are acquired in other BP dialects.

As further argument in favor of her analysis, she also shows independent evidence of copula loss in clefts, as we can see below.

(49)   *É inteligente o seu menino!*   >   *() Inteligente o seu menino!*
         is intelligent the your boy
         'Intelligent your boy is!'

Finally, Kato claims that the classic 'fronted' interrogative in (50) derives from (48.a) by deleting the complementizer as a result of a stylistic phonological process called *haplology*.

(50)   *Onde você vai?*
         where you go
         'Where do you go?'

The haplology process involves the loss of a syllable next to a phonetically identical (or similar) one, as can be seen in (51). The author claims that higher frequency of haplology in written texts constitutes evidence for this analysis, since it suggests that (48.a), and not (50), is the vernacular version.

(51)   a. *O que que ele quer?*   >   a'. *O que () ele quer?*
            the what that he wants   the what he wants
         b. *Quem que chegou?*   >   b'. *Quem () chegou?*
            who that arrived   who arrived

Yet, Kato does not explain why this also happens with wh-words that share no phonetical similarities with the complementizer (such as *onde* 'where'), or worse, with heavier wh-phrases (like *Que livro () você comprou?* 'Which book did you buy?'). Nor does she elaborate further on why the complementizer disappears on some occasions but not on others. From what I have noticed, preference for (a, b) in (50) vs. (a', b') may also be due to dialectal differences, but a confirmation of this suggestion would require further study.

Kato's analysis accounts for intonational facts, it also brings independent supporting evidence and seems to tie up many loose ends. However, it is not without its problems, as I have been discussing throughout this section.

What in my opinion seems to be more problematic is that she assumes fronted interrogatives derive from clefts, but although she gives acquisition evidence showing that children produce the canonic cleft construction, she does not show the real relevant data. To support her proposal, it is superlative that these constructions be acquired before fronted interrogatives, since the latter are supposed to derive from the former. So, she should show there is no fronting in

children speech until after they have acquired interrogative clefts of the kind in (48.b). This is the same argument that Grolla (2005) uses against the analysis in Kato and Raposo (1994).

1.4. True optionality

As I have mentioned several times before, optionality of the kind *move* vs. *do not move* is not really compatible with a theory in which movement must be triggered and triggered movement is obligatory. That is why most proposals for syntactic phenomena involving optionality consider it to be only apparent or attribute it to different numerations.

Biberauer and Richards (2006), though, seem to have found a different way of solving this matter. They also treat optionality as a consequence of different numerations, but only when the alternations give rise to different interpretations. Semantically vacuous optionality – *true optionality* – on the other hand, is conceived here to be a natural and expected result of considering *Move* as internal *Merge*. So, they argue that, just like it is possible to choose which elements will satisfy thematic selectional features, it should be equally possible to choose between several elements to satisfy non-thematic selectional features (EPP-features), as long as the relevant formal trigger is motivated. In this section, I will try to explain their proposal.

The authors start their paper introducing the problem. The main aim within the Minimalist Program (Chomsky, 1995; and subsequent work) is to reduce to the bare minimum the computational system underlying the Human Language Capacity. Thus, only three (and later on, just two) operations are postulated, namely Merge, Move (which then became External and Internal Merge), and Agree. In addition to the operations themselves, two fundamentally opposite economy principles govern syntactic operations. Last Resort (LR) requires that movement needs to be motivated ('don't do too much'), while Full Interpretation (FI) guarantees that, when there is a motivation, movement necessarily occurs ('don't do too little'). In other words, a constituent cannot move without the presence of the relevant feature, and such feature obligatorily triggers movement. *A priori*, this leaves optionality out of the picture.

However, many languages show word order alternations. The last versions of the model (Chomsky, 2000, 2001) propose to deal with optionality via the optional EPP-feature, which applies only when necessary to yield a new outcome. Therefore, structures with different interpretations but apparently the same lexical items are the result of different numerations (one

with the EPP-feature and one without it). This means, though, that optionality is only possible when it yields different interpretations.

But, then again, semantically vacuous optionality is present in natural languages. As an example, Biberauer and Richards show alternate placement of auxiliary verbs in Modern Spoken Afrikaans embedded clauses without any semantic consequences.

(52) a. *Ek weet dat sy dikwels Chopin gespeel **het***.

       I know that she often Chopin played has

       b. *Ek weet dat sy **het** dikwels Chopin gespeel*.

       I know that she has often Chopin played

       'I know that she has often played Chopin

The authors then claim that cost considerations play an important role in optionality. According to them, the computational system will be deterministic when it has an internal reason to be (based on LR and FI), but when two options are equally costly the system will not be able to choose. And what do they mean by 'equally costly'? They consider two operations are equally costly at a specific stage of the derivation when they both satisfy the formal imperative driving operations at that stage. LR and FI require that this formal imperative be satisfied right away, but how it is satisfied is not specified.

The formal imperatives in question are EPP-features. Biberauers and Richards claim that obligatory EPP comes for free in the numeration, whereas optional EPP comes at a cost and is detectable at the interface. Thus, obligatory EPP-features yield optional movement, and optional EPP-features yield obligatory movement. This means that obligatory EPP-features are the reason behind total semantic equivalence.

Biberauers and Richards affirm there are four different ways in which T's obligatory EPP can be satisfied depending on the *source* and *size* of the nominal category that moves to value T's phi-features. As for the source, some languages seek this category in Spec-vP while others (with rich-agreement paradigms) do it in the verbal head, more specifically, in the agreement morpheme. Regarding the size, in non piedpiping languages only the category bearing the agreeing phi-features moves, whereas in piedpiping languages this category carries with her the whole vP when it raises.

(53)   Typology of T's EPP satisfaction[15]

	Spec-seeking grammars	Head-seeking grammars
- piedpipe vP	Spec-rising (English)	Head-rising (Italian)
+ piedpipe vP	Spec-piedpiping (Afrikaans)	Head-piedpiping (German)

To illustrate their proposal, the authors analyze optional 'verb movement' in Afrikaans embedded clauses (like (52)). These alternations are treated as two computationally equivalent options to satisfy the obligatory EPP-feature in T. According to them, in Spec-piedpiping languages, the nominal category bearing the relevant phi-features is the DP in Spec-vP, which means it can either raise independently (54.a) or piedpipe the entire category that immediately contains it (54.b), and since both options satisfy the same EPP-feature, they are equally costly.

(54)   a. The spec-piedpiping option

---
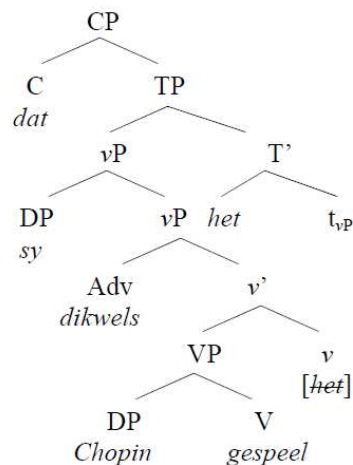
[15] According to the authors, this parametrization of EPP satisfaction can explain, for example, why nominative arguments may remain 'unraised' in German (and other languages in this family). Since this language is [+piedpipe vP], the whole vP is raised to Spec-TP.

(i)   ...daß dem Mann das Buch geschenkt wurde.
         that the.DAT man the.NOM book presented became
         '…that the book was given to the man.'
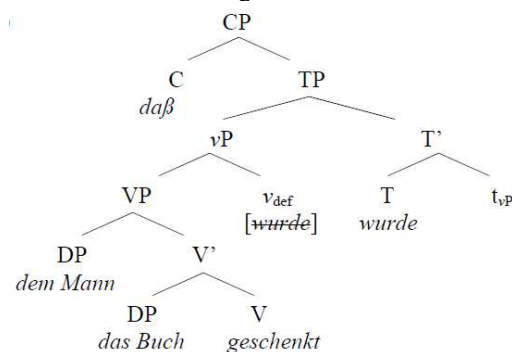
b. The spec-raising option

```
                    CP
                  /    \
                 C      TP
                dat    /   \
                     DP     T'
                     sy    /   \
                        het      vP
                               /    \
                            t_DP     vP
                                   /    \
                                 Adv     v'
                                dikwels /   \
                                      VP      v
                                     /  \   [het]
                                   DP    V
                                 Chopin gespeel
```
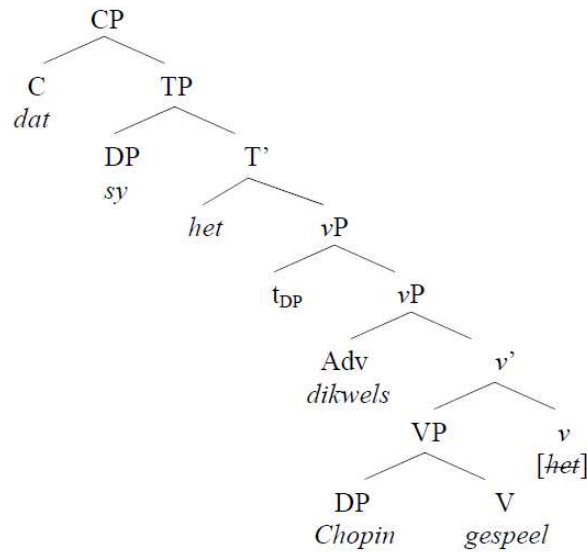
Biberauer and Richards offer a similar account for optional expletives in impersonal passives of spec-piedpiping languages like Afrikaans and Dutch.

(55)   a. *...dat (daar) gedans word.*   (Afrikaans)

b. *...dat (er) gedanst wordt.*   (Dutch)

   that (Expl) danced becomes

'... that there is dancing.'

In this case, what needs to be explained is how T's EPP-feature is satisfied when the expletive is omitted. The authors suggest that the passive participle of impersonal passives bears the appropriate phi-features. So, if there is an expletive in Spec-*v*P, its phi-features will agree with T's, otherwise T will look elsewhere and find the relevant features in the participle. In both cases the whole *v*P is piedpiped, the difference is that the former is an instance of spec-piedpiping and the latter of head-piedpiping. Again, since both options satisfy the same EPP-feature, they are equally costly.

The authors also give further supporting evidence from wh-interrogatives in Afrikaans, LBC violations in Russian and Ancient Greek, and quantifier stranding in Arabic, but I will not go into any more detail here.

Biberauer and Richards' idea of *true optionality* is quite an interesting one, because it solves, in a principled way, the incompatibility between optionality and a feature driven approach to movement. However, it only applies in a very specific situation: the alternative structures have

to be semantically equivalent and the result of two equally costly operations (in terms of computation) to satisfy the same obligatory EPP-feature.

This means that, unfortunately, *true optionality* cannot account for the alternation fronted vs. *in situ* in Brazilian Portuguese interrogative sentences. For one, it is not clear that the two options have exactly the same semantic characteristics. But more importantly, even if the alternation were semantically vacuous, the two operations should be equally costly and satisfy the same obligatory EPP-feature. The only possible candidate in this case is the EPP-feature in C, which we could assume to be obligatory for interrogative clauses (a peculiar notion of obligatoriness). However, this would still not account for BP fronted vs. *in situ* optionality, since there would be no way of satisfying the feature in the *in situ* option.

1.5. How Spanish can help us pick a side

Considering the problems in Mioto and Kato (2005) discussed earlier, we are left with two possible analyses for Brazilian Portuguese wh *in situ*: Grolla's (2005) loosened view of parameter setting in which the language is first set to be [+wh movement] and later reconsidered as being [±wh movement], and Kato's (2013) unified treatment of fronted and *in situ* interrogatives as instances of the same movement to an internal focus position. Both approaches have their own issues, as I have already noted in section 1.3. However, based on a comparison with wh *in situ* in Spanish, I would like to argue in favor of Grolla's proposal.

Spanish *in situ* interrogatives are much more restricted than those in BP. They are only felicitous when there is a heavily presupposed context (and probably also a contrastive topic) such as in this example from Etxepare and Uribe-Etxebarria (2005) (apud. Jiménez 1997):

(56)   Speaker A: *Mi padre, mi madre y yo fuimos a comprar huevos, leche y café.*
                       Mi madre compró los huevos.
                       'My father, my mother and I went to buy eggs, milk and coffee.
                       My mother bought the eggs.'
          Speaker B: *¿Y tu padre compró qué?*
                       'And your father bought what?'

Also, in this kind of interrogatives in Spanish, the wh-phrase must be sentence final, hence, giving rise to marked word orders. This does not happen in BP.

(57)　　a. *María le dio un libro a Pedro.*　　　　(natural, unmarked word order)

　　　　Maria CL gave a book to Pedro

　　　　b. **¿María le dio (el) qué a Pedro?*

　　　　Maria CL gave (the) what to Pedro?

　　　　c. *¿María le dio a Pedro (el) qué?*

　　　　Maria CL gave to Pedro (the) what?

Etxepare and Uribe-Etxebarria (2005) claim that in situ wh-phrases in Spanish are overtly moved to a focus position, the final word order being a result of subsequent remnant movement. This proposal explains the restricted distribution of wh *in situ*, the sentence final requirement and other characteristics that this structure shows in Spanish.

Recall that, according to Grolla (2005), what triggers the reevaluation of the input and thus the resetting of the parameter for wh-movement in BP is the presence of a special empty category like the one in the sentences in (40), repeated here as (58).

(58)　　a. *Quem que você gosta* [PP ec]*?*

　　　　'Who (is it) that you like?'

　　　　b. *O que você vai precisar* [PP ec] *para fazer o bolo?*

　　　　'What are you going to need to make the cake?'

　　　　c. *Esse livro eu vou precisar* [PP ec] *pra semana que vem.*

　　　　'That book I am going to need (it) next week.'

　　　　d. *O João eu gosto* [PP ec]*.*

　　　　'(It is) João I like.'

Spanish lacks this kind of empty category. In interrogative sentences with verbs that require a PP as a complement the preposition always needs to be present, as shown in (59), where we can see Spanish sentences equivalent to those in (58). And some authors (for example, Casielles-Suárez, 2003) argue that there is no topicalization in Spanish, only different kinds of left-dislocation[16].

(59)　　a. *¿ *(De) quién te has enamorado?*

　　　　'(With) whom have you fallen in love?

---

[16] In Ross' sense (1967):

　　(i)　　John, I saw him yesterday.　　　Left-dislocation
　　(ii)　　Julia I couldn't reach ().　　　Topicalization

b. *¿ *(De) qué te quejas?*

'(About) what are you complaining?'

c. Ese libro *(lo) voy a necesitar la semana que viene.

'That book I am going to need (it) next week.'

d. A Juan *(le) quiero muchísimo.

'Juan I love (him) so much.'

Following Grollas's reasoning, the Spanish child, then, does not find the relevant *cue* that allows her to analyze wh *in situ* as a structure without movement, since this kind of A' dependencies without movement do not occur in the input.

The differences between wh *in situ* in BP and in Spanish suggest that the syntactic structures underlying the construction in both languages are not the same. So, if we assume Etxepare and Uribe-Etxebarria's analysis for Spanish, Grolla's proposal for BP seems a better choice. Given how restricted this phenomenon is in Spanish, to assume Grolla's analysis would also be the most reasonable thing to do if this dissertation's results confirm the initial hypothesis of BP wh *in situ* early acquisition in Spanish native speakers.

Chapter 2
Topic-Subject Constructions in Brazilian Portuguese

2.1. Introduction

In Linguistics, the term 'topic' is used in many different contexts and with somewhat different meanings. Even reducing our scope to the sentence, the topic can be defined in several ways. For example, it can be considered as the element about which the sentence is about. With this definition of topic, subjects are usually (but not necessarily) topics. But this is not the definition that interests us here.

Trying to leave pragmatic and discursive notions aside, the syntactic concept of topic refers to a marked constituent, an NP or a PP in a sentence-peripherical position, about which a proposition is made through the rest of sentence. Languages differ in how they implement this notion. In English, for example, there are two kinds of topic constructions, topicalization ('*That ice-cream* I absolutely love') and left dislocation ('*That ice-cream*, people go crazy about it').

In addition to those two possibilities, BP has a special type of 'topic' constructions in which a non-argumental[17] constituent behaves like a subject, triggering verb agreement. Although it is not so clear that these constructions really involve the raising of a topic, they are called topic-subject constructions. This term was coined by Pontes (1987) and it is widely accepted in the literature. There are several kinds of sentences that could fit into this label, but the most typical ones are those in which either a locative (1) or a genitive (2) constituent is raised to subject position.

---

[17] Actually, as shall be discussed later in this chapter, some authors consider the raised elements as being argumental (at least in one of the types of topic-subject construction). In any case, the element raised to subject position is not the logical subject, which remains in post-verbal position.

(1)   *Essa casa bate bastante sol.*

        that house beat.3SG plenty sun

        'That house gets a lot of sunlight.'

(2)   *O meu carro furou o pneu.*

        the my car punctured.3SG the tire

        'My car got a flat tire.'

In this chapter, I will briefly review the most relevant literature on topic-subject constructions, from the first works that noticed these structures (Pontes, 1987; Galves, 1986) or the first that proposed an analysis within the minimalist frame of work (Galves, 1998), to the latest proposals put forward to try to account for their grammatical peculiarities (Munhoz & Naves, 2012; Andrade & Galves, 2014; Avelar & Galves, 2011, 2016).

I will try to show that, despite the technical differences between the various proposals, what they all have in common is that they postulate a syntactic position for BP, the one receiving the topic-subject constituent, that behaves quite unlike the topic or subject positions found in other Romance languages.

2.2. Breaking the ice: Pontes (1987)

Eunice Pontes was the first to study, in a systematic way, how topics behave in BP and to show that their behavior differs drastically from what happens in the European variety. She is considered the pioneer in the study of topics in BP, and, in particular, in the study of topic-subject constructions. Pontes was the first to call attention to sentences in which the topic, being in initial position, looks like a subject and also behaves as such, triggering verb agreement, as we saw in (1) and (2).

Pontes' book *'O tópico no Português do Brasil'* is a collection of papers and conference presentations exploring, from a functionalist perspective, the syntactic notion of topic in BP. The first and fourth chapters are the ones most relevant to this research because that is where she discusses topic-subject constructions and introduces what have now become the classic examples of these structures.

In the first chapter, which underlines the importance of topics in BP, the author tries to show how this language is situated within the typology proposed by Li & Thompson (1976). Their

proposal divided languages in four different types, depending on the role that the grammatical relations subject-predicate and topic-comment[18] play in the structure of sentences:

i)   Subject-prominent languages, in which the basic sentence structure favors a description in terms of the subject-predicate relation (English and most Indo-European languages)

ii) Topic-prominent languages, in which the basic sentence structure favors a description in terms of the topic-comment relation (Chinese)

iii) Languages both Subject and Topic-prominent, in which subject-predicate and topic-comment constructions are equally important sentence structures and different from one another (Japanese)

iv) Languages neither Subject nor Topic-prominent, in which subject and topic have blended and are impossible to distinguish. (Tagalog)

Pontes claims that sentences with a topic-comment structure are very common in BP and she brings many examples throughout the whole book. She points out that perhaps the most frequent type in this language is that in which the topic and the subject in the comment have the same reference, just like it happens in the following example:

(3)   *Os livros, eles estão em cima da mesa.*
      the books they be.3PL in top of.the table
      'As for the books$_i$, they$_i$ are on the table.'

However, sentences like the one in (4) below are more interesting, since they have the same structure as the most typical topic-comment construction (5) in a language like Mandarin Chinese, which is the prototypical Topic-prominent language. According to the author, these sentences are inaccurately referred to as 'double subject' constructions because they have two different sentence-initial DPs, the first one is the topic, which establishes a context for the comment, and the second one is the actual subject of the sentence that forms the comment.

(4)   *Essa bolsa as coisas somem aqui dentro.*
      that bag the things disappear.3PL. here inside
      'In that bag, things disappear.'

---

[18] The following English examples can help clarify these two grammatical relations.
   i)   Subject-predicate: Mary / kissed John.
   ii) Topic-comment: As for literature, / she prefers the classics.
Notice that the topic establishes a setting for the comment that follows, which is formed by a whole sentence with subject and predicate.

(5)   *Neì-xie shùmu shù-shén dà.*

those trees trunks big

'As for those trees, their trunks are big.'

Pontes shows, though, that these are not the only topic constructions in BP. It seems any DP can be a topic in this language, regardless of its syntactic function within the sentence, just like in Chinese. This is one of the properties of topics in Topic-prominent languages on which Li and Thomson based their typology. Pontes lists these properties and argues that topics in BP verify all of them (definiteness, discourse-related semantic role, sentence-initial position, lack of selectional restrictions, lack of verb agreement[19] and lack of involvement in syntactic processes).

She then proceeds to enumerate too the main characteristics of Topic-prominent languages pointed out by Li and Thompson (1976) and she claims BP shares almost all of them.

- Passives are rarely used.
- There are no expletive subjects.

(6)   a. *Chove*.

'(It) rains.'

b. *Tem um gato no jardim.*

'(There) is a cat in the garden.'

- 'Double subject' constructions of the type in (4) are quite typical.
- There are no restrictions on what can constitute a topic.

As mentioned before, any DP can be a topic in BP, regardless of its syntactic function.

(7)   a. *Essa competência ela é de natureza mental*.   Subject

'That skill, it is of a mental nature.'

b. *A Belina o Hélio levou para a oficina.*   Direct Object

'The car, Hélio took (it) to the mechanic.'

c. *A Joana, o Pedro não deu o presente*.[20]   Indirect Object

'Joana, Pedro did not give (her) the present.'

d. *Washington a neve é pouca*.   Adjunct (locative)

'(In) Washington, the snow is spare.'

---

[19] At this point in her discussion she is referring to topic constructions like that in (4). As we will shortly see, topic-subject constructions like those in (1) and (2) are a special case and, as such, have special properties, for example, verb agreement.

[20] Except for this example added by me, all the rest were taken from Pontes (1987).

e. *Semana que vem eu vou à praia.*   Adjunct (temporal)

'Next week, I am going to the beach.'

f. *Esse negócio o prazo acaba.*   Noun adjunct

'That thing, (its) deadline is approaching.'

- Co-reference is controlled by the topic instead of the subject.

  This is evident in examples like (7.b, 7.c) and (4).

- Topic constructions are not the result of any transformation of a more basic sentence structure[21].

  Pontes argues that if topic constructions were not basic sentence structures, they should not be able to appear in negative, interrogative or exclamative sentences, nor in subordinate clauses, but they do all that in BP. She also claims that sentences like (4) cannot be derived from more basic ones. However, this point is quite controversial, as we will see in the following sections of this chapter.

Pontes admits there are two characteristics of Topic-prominent languages that BP does not share, namely, verb in sentence-final position and morphological marking of the topic. She takes some time to explore the use of a co-referring pronoun in BP topic constructions stating the problems that a transformational analysis would ensue in this case and suggesting that the pronoun could function as a morphological marking for the topic.

Based on all the above, Pontes concludes then, that

> *[…] if Li and Thompson are on the right track with their proposed typology, then Portuguese is at least a language of the third type, in which both notions (subject and topic) are prominent.*

(Pontes, 1987, p. 39, my translation)

But before reaching this conclusion, she notices a special type of topic constructions, that would, from then on, be known as topic-subject constructions. In these sentences, the logic subject appears in post-verbal position and a sentence-initial topic looks like the grammatical subject, resulting in an apparent SVO word order. Examples (1) and (2) in this chapter's introductory section have become the most classic ones, but below are some other examples of these constructions.

---

[21] Here Pontes is using the original generative model or one of its reformulations (Standard Theory, Extended Standard Theory), in which every sentence is considered to be a transformation of a more basic sentence called kernel sentence.

(8)   a. *Esse carro cabe 60l de gasolina.*

          that car fit.3SG 60l of gasoline

          '60l of gasoline fit in that car.'

      b. *Essa torneira não sai água?*

          that tap not get.out.3SG water?

          'Is water not coming out of that tap?'

(9)   a. *A Sarinha tá nascendo dente.*

          the Sarinha be.3SG being.born tooth

          'Sarah's teeth are coming out.'

      b. *O jasmim amarelou as pontas.*

          the jasmine became.yellow.3SG the tips

          'The tips of the jasmine leaves became yellow.'

Notice that, the topic-subjects in (8) are locative constituents whereas in (9) they are genitive constituents. Pontes points out that each one of these sentences has a counterpart where the constituents are in 'direct order'. For example, (10.a) and (10.b) would be (8.a) and (9.a)'s counterparts.

(10)   a. *Nesse carro cabem 60l de gasolina.*

          in.that car fit.3SG 60l of gasoline

          '60l of gasoline fit in that car'

      b. *O dente da Sarinha está nascendo.*

          the tooth of.the Sarinha be.3SG being.born

          'Sarah's teeth are coming out.'

But she spends some time throughout the book arguing these versions are not semantically equivalent to the topic-subject versions. According to the author, the sentences in (10) have a more neutral reading than those in (8) and (9). She uses this difference in meaning as an argument against an analysis via transformations[22].

As Pontes says, these sentences raise an interesting problem for the theory. The first DP behaves like a topic. It has a discourse-related semantic role, it controls co-reference and it occupies the sentence-initial position. But it also behaves like a subject. It agrees with the verb, as we can

---

[22] In the version of the model Pontes is using, transformations were supposed to preserve meaning. Currently, though, this is no longer accepted. Movement to A'-positions is considered to have a semantic effect. (See for example Rizzi's proposal for *Criterial positions* (Rizzi, 2006))

see in (11), and, as (12) shows, it competes for the subject position when the second DP is placed before the verb, thus making the sentences ungrammatical.

(11)   a. *Essas casas batem bastante sol*.
   those houses beat.3PL plenty sun
   'Those houses get a lot of sunlight.'

   b. *Meus carros furaram o pneu*.
   My.PL cars punctured.3PL the tire
   'My cars got a flat tire.'

(12)   a. **Essa casa bastante sol bate*.
   that house plenty sun beat.3PL
   'That house gets a lot of sunlight.'

   b. **O meu carro 60l de gasolina cabe*
   the my car 60l of gasoline fit.3SG
   '60l of gasoline fit in my car.'

When there is no topic, though, the verb agrees with the postposed subject. However, later in the book, she claims that verb agreement is not a very reliable test for subject position in BP since it is a very variable aspect of this language's grammar. She also shows that topic-subject constructions cannot undergo passivization and she says this constitutes, at least, a challenge for an analysis considering the sentences' first DP as a subject.

On the other hand, Pontes claims that the first DP in some of these sentences shows a 'responsibility' for what the verb conveys, which according to Lakoff (1977) is one of the main characteristics of a subject[23]. But, then again, this is only so for topic-subject constructions of the locative type as in (1) and (8), those in which the topic-subject is a possessive constituent as in (2) and (9) do not show this 'responsibility'.

Pontes suggests that these sentences may constitute a middle ground between 'double subject' topic constructions of the kind in (4) and subject-predicate constructions. Topic and subject are intertwined, and so are comment and predicate. Finally, she also suggests that this behavior is

---

[23] The notion of 'responsibility' as a main semantic characteristic of the subject is quite problematic. See, for example, the sentence below. It is very hard to sustain the idea that 'the phone' has any responsibility whatsoever over the situation.
The phone hit the ground after the cat pushed it over the table.

similar to what happens in languages of the fourth type in Li and Thompson's typology, in which the notions of topic-comment and subject-predicate are no longer distinguishable.

2.3. Galves' earlier work

I want to single out now the work of Charlotte Galves, because of her continuous search for a wider vision and deeper understanding of topic-subject constructions. Both in her 1986 paper and in more recent papers by herself or with other authors (as we will see in this section and the next), one of her main concerns was showing that other syntactic phenomena, although apparently unrelated with topic-subject constructions, were in fact very much connected to these structures. She seemed (and still seems) to be looking for some characteristic of BP grammar that could explain all these phenomena at once.

2.3.1. *Galves (1986)*

In her 1986 paper, Galves gets on board with Pontes' idea about topics in BP and their importance as a distinguishing characteristic, arguing in favor of considering BP as a Topic-Comment kind of language. Galves (1986) claims that this similarity with Topic-Comment languages is exactly what constitutes the main difference between Brazilian and European Portuguese, but also between BP and other Romance languages. She carefully discusses a series of syntactic differences between the two Portuguese varieties and argues they can be accounted for on the basis of this distinction.

Galves shows how BP and EP differ as to the use of the tonic pronoun *ele* ('he') and the use of the multifunctional Romance pronoun *se*, and she relates these differences to the interpretation of empty categories in these languages. I will not discuss the details of her argumentation, but I will try to summarize her conclusions below.

She claims that third person verb inflection is no longer enough in BP to assign a determined reference to the null subject, hence the pronoun *ele* ('he') becomes necessary whenever there is no explicit DP in the sentence (13.a) and plays the role of verbal agreement when there is (13.b). To express indetermination, the impersonal pronoun *se* is no longer used. The null subject of a sentence with a verb in the third person is interpreted as undetermined whenever the tonic pronoun is not present and there is no plausible antecedent for it (14).

(13)   a. *O João é muito inteligente. Ele gosta de estudar.*

     'John is very intelligent. He likes to study.'

    b. *Essa competência, ela é de natureza mental.*

     'That competence, it is of a mental nature.'

(14)   *Nos nossos dias, não usa mais saia.*

    'In these days, skirts are no longer used.'

According to Galves, the weakening of BP's verbal inflection resulted in a new sentence structure where the topic plays a central role. Unlike in other Romance languages, the topic in BP can be the direct antecedent for null objects (15.a), thus making clitic pronouns unnecessary, and it can also be the antecedent for null subjects in infinitive sentences (15.b).

(15)   a. *Apanharam as maçãs e guardaram [ec] no porão.*

     'They took the apples and put (them) in the basement.'

    b. (Speaking about Mary's baby) *Eu queria acabar antes de [ec] nascer.*

     'I wanted to finish before (he) would be born.'

The author recalls that although the behavior in (15) is perfectly acceptable, the tonic pronoun *ele* can be used too, both in subject and in object position, in order to make the pronominal features more evident. In addition to that, in BP, any syntactic position can be co-indexed with the topic. These are two characteristics related to the syntax of Topic-Comment languages.

Besides the relevance of Galves (1986) for placing topic constructions within a bigger set of BP peculiarities and for trying to find some characteristic that could explain all of them, this paper is important for the study of topic-subject constructions because it introduced a new type of sentence within this category that had not been considered before. When discussing the interpretation of empty categories in object position, Galves relates the behavior in (15.a), a direct link between topic and object, to the lack of passives in BP and she gives some examples of sentences with passive interpretation but without the usual morphological modifications:

(16)   a. *A revista está xerocando.*   vs.   *está sendo xerocada.*

     the magazine be.3SG photocopying   be.3SG being photocopied

     'The magazine was being photocopied'

    b. *A linha dele tinha parado de fabricar.* vs. *ser fabricada.*

     the line of.his had.3SG stopped of to.produce   to.be produced

     'His line had stopped being produced'

c. *Aquela verba que liberou agora...*   vs.   *foi liberada...*

that money that released.3SG now...   was.3SG released

'The funds that were just released...'

The author says that it is precisely the link between the topic and the null object what renders the passive form unnecessary. She also says this is related to the possibility of interpreting the null subject in the finite sentence as undetermined. And she concludes that, since verbal inflection can no longer assign a specific reference to the null subject, the topic is reanalyzed as a subject, also because there are no phonological or syntactic marks that could distinguish one from the other.

2.3.2. *Galves (1998)*

In this work, Galves resumes the study of topic-subject constructions that she just hinted at in 1986. I believe it is an important paper because it constitutes the consolidation of this notion and it presents one of the first (if not the first) systematic analyses for these structures within the Minimalist Program.

Here, Galves begins by introducing the concept of topic-subject constructions and dividing it into two types. On one hand, we have sentences like the ones in (16) above, *'with a transitive verb accompanied only by its internal argument in pre-verbal position, but showing no inflectional marks that indicate a modification in the projection of its argumental structure'* (Galves, 1998, p.19, my translation). I will call them Type I sentences. On the other hand, there are the classic examples in (1) and (2), also in (8) and (9), defined by the author as *'pseudo-transitive constructions with an ergative verb in which the pre-verbal DP is not interpreted as agent or cause of the process expressed by the verb, but as a locative or a whole of which the post-verbal DP is a part'.* (idem). I will call these other type, Type II.

Galves brings some evidence in support of the idea that the topic is behaving like a subject. Regarding agreement, she shows that the pre-verbal DP agrees with the verb, as we saw in the previous section, ex. (11), and as can be seen again in (17); that this agreement is in complementary distribution with the presence of a resumptive pronoun, see (18); and that, on the contrary, agreement with the post-verbal DP is not possible unless there is a resumptive pronoun in the construction, compare (19.a) and (19.b).

(17)   *As balanças estão consertando.*

        the scales be.3SG fixing

        'The scales are being fixed.'

(18)   a. *Estas casas batem muito sol.*

        these houses beat.3SG plenty sun

      b. **Estas casas, batem muito sol nelas.*

        these houses beat.3PL plenty sun in.them

      'These houses get a lot of sunlight.'

(19)   a. ??*Este carro cabem muitas pessoas.*

        this car fit.3PL many people

        'Many people fit in this car.'

      b. *Este carro, cabem muitas pessoas nele*.

        this car fit.3PL many people

        'As for this car, many people fit in it.'

As further evidence for the subject qualities of the first DP, the author claims that the verb cannot project an external argument in topic-subject constructions. However, in the example she provides, the chosen verb (*pôr* – 'to put') does not fall into any of the categories that, according to her, license these structures, so its ungrammaticality could be due to that fact and not to the projection of the external argument. I provide here another example that I believe proves her point[24].

(20)   a. *A xícara, o João quebrou a asa dela*.

        the cup the João broke.3SG the handle of.it

        'As for the cup, João broke its handle.'

      b. ??*A xícara o João quebrou a asa*[25].

        the cup the João broke.3SG the handle

        'João broke the cup handle.'

---

[24] Notice that this is somewhat similar (but not exactly the same) to what Pontes (1987) says when she shows that the first and second DPs compete for the subject position when both are pre-posed (example 12).

[25] As Galves herself points out, the intonational structure plays a significant role in the acceptability of examples such as this or the one in (19.b). If there is a pause between the first DP and the rest of the sentence, it is interpreted as a topic-comment structure and the sentence can become acceptable.

She also points out that when the pre-posed DP is in a genitive relation with the postposed DP, this relation must have a whole/part semantic interpretation to license the topic-subject construction. Compare the two sentences in the following example.

(21)   a. *A mesa quebrou o pé.*

the table broke.3SG the foot

'The foot of the table broke.'

b. **A mesa quebrou o pote.*

the table broke.3SG the jar

'The jar on the table broke.'

To conclude her discussion about the properties of topic-subject constructions, Galves calls our attention to the ambiguity of a sentence like (22.a). Recall that this is the first of the two different types of topic-subject constructions considered by the author. This sentence can be interpreted as a passive (22.b) or as topicalization where there is an indetermination of the subject (22.c).

(22)   a. *A cueca de dinossauros do Calvin está lavando.*

the underwear of dinosaurs of Calvin be.3SG washing

b. *A cueca de dinossauros do Calvin está sendo lavada.*

the underwear of dinosaurs of Calvin be.3SG being washed

'Calvin's dinosaur underwear is being washed'

c. *A cueca de dinossauros do Calvin, está-se lavando.*

the underwear of dinosaurs of Calvin, be.3SG – SE[26] washed

'Someone is washing Calvin's dinosaur underwear'

Based on all of this, and also on the difference between BP constructions with resumptive pronouns and their analogous ones in other Romance languages[27], Galves proposes an analysis

---

[26] Romance pronoun *se* can function in several different ways:

(i)	Reflexive *se*	*A Maria se escovou os dentes.* 'Maria brushed her (own) teeth'
(ii)	Reciprocal *se*	*A Maria e o Pedro se escrevem cartas.* 'Maria and Pedro write letters to each other'
(iii)	Impersonal *se*	*Na festa só se falou de política.* 'At the party, (people) only talked about politics'
(iv)	Passive *se*	*Procura-se garçonete.* 'Waitress wanted'
(v)	Medium *se*	*Esse livro se lee muito rápido.* 'That book reads very easily'

However, the distinction between the last three types of *se* can be quite complicated, especially in BP, where *se* is in disuse. For example, in (22.c) it is not clear whether *se* is a passive marker or a particle to avoid making explicit the logical subject of the sentence.

[27] Without getting into all the details of her discussion here, she points out that BP tonic pronouns can be compared to EP clitic pronouns. She therefore proposes that, in sentences with resumptive pronouns, BP has invisible clitics created by a movement after Spell-Out of the tonic pronoun's phi-features.

for BP sentence structure, which, in particular, applies to topic-subject constructions. I will try to explain it below.

She adopts the following hypotheses:

i)   Agreement (AGR) is a formal feature. It is parametrically associated with functional categories. Categories with AGR project a Specifier.

ii)   There is a functional category between T and C: Person. This category Person has uninterpretable phi-features.

iii)   Pronouns are uninterpretable phi-features.

iv)   For BP, parametrization of functional categories T and Pers regarding features V and AGR goes as follows:

    a. Pers: -V, -AGR

    b. T: +V, +AGR

Galves claims this proposal accounts for some of BP's syntactic properties like 'short' verb movement and its relationship with weak verb morphology, it also reflects the idea of the subject being in Spec-TP. According to the author, the novelty here is considering there is a functional category Person with uninterpretable phi-features that need to be checked. And, since V and AGR are absent from this category, its phi-features cannot be checked by verb nor subject movement. This leaves invisible movement of pronominal phi-features as the only alternative.

To exemplify, a sentence like (7.a) and (13.b) would have the following underlying structure:

(23)   [PersP Essa competência [PersP ϕ_i [TP ela$_i$ é de natureza mental]]]

Notice that pronominal features moved after Spell-Out to check Person's phi-features. Notice too that the DP *Essa competência* is an adjunct to PersP. This analysis implies that, whenever there is no resumptive pronoun, there must be a null pronoun in subject position whose features check Person's phi-features. In this case, (24) is the structure proposed by Galves. In (25), I apply Galves' analysis to a simple sentence in BP.

(24)   [PersP DP [PersP ϕ_i [TP pro$_i$ ...]]]

(25)   [PersP A Maria [PersP ϕ_i [TP pro$_i$ dançou]]]   (BP)

In other words, every BP sentence is the result of dislocation. This is a consequence of phi-features and AGR being associated to different functional categories, unlike in the rest of Romance Languages, for which Person is +AGR.

As for the structure underlying topic-subject constructions, further explanations are in order. Recall that what licenses the pre-verbal DP is not the presence of a resumptive pronoun but verbal agreement, and also, recall that the verb's external argument (if there is one) cannot be present. Based on this, Galves proposes that Spec-TP is not projected, which means the verb's uninterpretable phi-features cannot be checked in a Spec-Head configuration. The only option, then, is that they move to Person and check themselves with this category's own phi-features. The author claims that with this movement, Person inherits T's AGR feature, which, in turn, is checked by the movement of a DP to Spec-PersP. The resulting structure is the following:

(26) $[_{PersP} DP_j [_{Pers'} \phi_i [_{TP} V_i t_j]]]$

Galves argues this structure can account for the characteristics of topic-subject constructions previously discussed. It explains the complementary distribution between resumptive pronoun and agreement with pre-verbal DP. If Spec-TP is projected, only pronominal features can check Person's phi-features; if it is not projected, the verb's features move and expand T's projection. The author claims her proposal explains as well the interpretive restrictions of these constructions (lack of external argument and whole-part relationship in genitive constructions). She argues that the pre-verbal DP is not external to the sentence (as in the constructions with resumptive pronoun), it occupies an A-position because it is in a Spec/Head configuration with verbal features (in accordance with Chomsky, 1995). As for the nature of the theta role it receives, the author herself admits it needs further research.

Let us now see how the proposed structure applies to specific examples built by me using Galves' analysis:

(27)   a. $[_{PersP} [As\ balanças]_j [_{Pers'} \phi_i [_{TP} estão\ consertando\ t_j]]]$ Type I
             b. $[_{PersP} [Essa\ casa]_j [_{Pers'} \phi_i [_{TP} bate\ [muito\ sol]\ t_j]]]$ Type II - Locative
             c. $[_{PersP} [O\ relógio]_j [_{Pers'} \phi_i [_{TP} quebrou\ [o\ ponteiro\ t_j]]]]$ Type II - Genitive

Galves refers to the movement of a DP to Spec-PersP to check T's AGR features inherited by Person, but she does not mention where this DP comes from or why the preposition is lost along the way. As we will see in next section, Munhoz and Naves (2012) make the origin of this movement their central concern; however, as we are about to see, their analysis has other problems.

2.4. State of the art: three different proposals

In this section, I sum up three of the most recent proposals for topic-subject constructions. Munhoz & Naves (2012) are interested in the argumental structure of the verbs that license topic-subject constructions and they propose there is a category between T and C that inherits C's phi-features and receives the topic-subject constituent. For Andrade & Galves (2014), the main concern is offering a unified analysis for both locative and genitive structures, which they do by suggesting the verbs involved instantiate a secondary predication or small clause. Avelar & Galves (2016), based on previous work (Avelar & Galves, 2011), place topic-subject constructions in a bigger frame and propose they are a consequence of a linguistic change involving two parameters, namely, phi-independence of T's EPP and the possibility of generating DPs without Case feature. The proposals are quite different from one another, but all three of them offer interesting insights about these constructions.

2.4.1. *Munhoz & Naves (2012)*

As I said, Munhoz and Naves place their research focus on the argumental structure of verbs licensing topic-subject constructions. They assume these verbs belong to the class of unaccusatives and they argue it is due to the heterogeneity of this class that there are also different types of topic-subject constructions.

According to the authors, in topic-subject constructions, the locative and genitive constituent have different status, the former modifies the verb whereas the latter modifies a post-verbal DP, the former is an argument whereas the latter is not. As evidence, they show locatives cannot be omitted without affecting the grammaticality of the sentence while genitive constituents can.

(28)　　　a. *(Essa casa) bate bastante sol. / Bastante sol bate *(nessa casa).*

　　　　　　that house beat.3SG plenty sun / plenty sun beat.3SG in.that house

　　　　b. *(O meu carro) furou o pneu. / O pneu (do meu carro) furou.*

　　　　　　the my car punctured.3SG the tire / the tire of.the my car punctured.3SG

The authors also apply constituency tests like hanging topics or cleft sentences in order to show how theme and genitive form a single constituent while theme and locative do not.

(29)　　　Hanging topics – 'As for X, it …'

　　　　　a. *[O pneu do meu carro]$_i$, ele$_i$ furou.*

　　　　　　the tire of.the my car he punctured.3SG

a'. *[O pneu]ᵢ, eleᵢ furou do meu carro.

the tire he punctured.3SG of.the my car

b. *[O sol nessa casa]ᵢ, eleᵢ bate.

the sun in.that house he beat.3SG

b'. [O sol]ᵢ, eleᵢ bate nessa casa.

the sun he beat.3SG in.that house

(30)   Cleft sentences – 'It is/was X that …'

a. *Foi [o pneu do meu carro] que furou.*

be. PAST.3SG the tire of.the my car that punctured.3SG

a'. *Foi [o pneu] que furou do meu carro.*

be. PAST.3SG the tire that punctured.3SG of.the my car

b. *É [sol nessa casa] que bate.*

be.3SG sun in.that house that beat.3SG

b'. *É [sol] que bate nessa casa.*

be.3SG sun that beat.3SG in.that house

Based on this, they propose there are two types of unaccusative verbs that license topic-subject constructions:

-   Monoargumental unaccusative verbs license genitive topic-subject constructions. In this case, a single argument is split and linearized in a discontinuous fashion.
-   Biargumental unaccusative verbs license locative topic-subject constructions. These verbs select two arguments, a theme and a locative, that are projected in different syntactic positions.

The authors point out, though, that theme and locative do not correspond to internal and external arguments. To support this claim, they show how their behavior differs from that of internal and external arguments of transitive verbs regarding passivization and cliticization.

(31)   Transitive verb – Admits both passivization (b) and cliticization (c)

a. *A Livraria do Chiquinho vende livros acadêmicos.*

'Chiquinho's bookstore sells academic books.'

b. *Livros acadêmicos são vendidos na/pela Livraria do Chiquinho.*

books academics be.3PL sold in.the/by.the bookstore of.the Chiquinho

'Academic books are sold at Chiquinho's bookstore.'

c. *A livraria do Chiquinho os vende / vende eles.*

the bookstore of.the Chiquinho CL.3PL sell.3SG / sell.3SG them

'Chiquinho's bookstore sells them.'

(32)   Biargumental unaccusative verb – Admits neither

    a. *Essa casa bate bastante sol.*

    'That house gets plenty of sunlight.'

    b. **Bastante sol é batido por/*nessa casa. / *Essa casa é batida por bastante sol.*

    plenty sun be.3SG beaten by/in.that house / that house be.3SG beaten by plenty sun

    c. **Essa casa o bate / *bate ele. / *O sol a bate / *bate ela.*

    that house CL.3SG beat.3SG / beat.3SG it / the sun CL.3SG beat.3SG / beat.3SG it

In order to further argue in favor of their proposal, Munhoz and Naves show locative and theme can both be extracted from a wh-island[28], as shown in (33), but genitive and theme can only be extracted together, as shown in (34).

(33)   *Esse caminhão cabe toda a mudança.*

    that truck fit.3SG all the furniture

    'All the furniture fits in that truck.'

    a. Theme

    O que_i o motorista não sabe [se o caminhão vai caber t_i]?

    the $what_i$ the driver not know.3SG [if the truck go.3SG fit t_i]?

    'What does the driver not know whether is going to fit in the truck?'

    b. Locative

    $Onde_i$/O que_i o motorista não sabe [se t_i vai caber toda a mudança]?

    $where_i$/the $what_i$ the driver not know.3SG [if t_i go.3SG fit all the furniture]

    'Where does the driver not know whether all the furniture is going to fit?'

(34)   *O carro furou o pneu.*

    'The car got a flat tire.'

    a. Theme

    *O que_i o mecânico não sabe [se o carro furou t_i]?

    the $what_i$ the mechanic not know.3SG [if the car punctured.3SG t_i]

    'What does the mechanic not know the car punctured?'

---

[28] Extraction from wh-islands is a classic test to know whether a constituent is an argument or not.

b. Genitive

*O que*ᵢ *o mecânico não sabe [se t*ᵢ *furou o pneu]?*

the whatᵢ the mechanic not know.3SG [if tᵢ punctured.3SG the tire]

'What does the mechanic not know punctured a tire?'

c. Theme and locative

*O que*ᵢ *o mecânico não sabe [se furou t*ᵢ*]?*

the whatᵢ the mechanic not know.3SG [if punctured.3SG tᵢ]

'What does the mechanic not know punctured?'

Regarding the preposition's presence or lack thereof, the authors say that prepositions are compulsory in post-verbal locatives, whereas pre-verbal locatives may appear with or without the preposition. Only in this last case they consider the sentence a topic-subject construction. However, as shown in Avelar & Cyrino (2008), a prepositioned constituent can appear in subject position when the verb presents the features of singular third person, see the following example.

(35)   *Na casa de Maria chegou umas cartas.*

in.the house of Maria arrived.3SG some letters

'Some letters arrived at Maria's house.'

Munhoz and Naves suggest there is a numeration with preposition and another one without it. When the preposition is present, it merges with its complement forming the locative PP that is then selected by the verb. After that, the second argument (the theme) merges with the rest of the structure forming the VP. When there is no preposition, the first *Merge* occurs between the verb and the theme, after which the locative is merged, thus remaining in a position from where raising is possible. They assume that, since these verbs do not project external argument, there is no *v*P.

As for their analysis of topic-subject constructions underlying structure, the authors take Pilati & Naves (2012) as a starting point and assume that, in Subject-prominent languages, C's phi-features are transferred to T and agreement is triggered by Case, whereas in Topic-prominent languages, they are transferred to α, a projection between C and T, and agreement is not triggered by Case. The authors do not explain exactly what this projection is or how it is semantically interpreted. Given the middle ground in which BP stands, they claim C's features can be inherited either by T or by α, depending on the syntactic construction. They propose that in topic-subject constructions, α inherits C's phi-features, making T defective and allowing

raising of a DP to Spec-αP. Since this category is not able to assign Case, any DP can be attracted there. This means that the proposed underlying structure is the following, both for locative and genitive topic-subject constructions.

(36)   a. [CP [αP Essa casa bate [TP ... [VP ~~essa casa bate~~ [DP sol]]]]]]

         b. [CP [αP O carro furou [TP ... [VP ~~furou~~ [DP [DP ~~o carro~~ [D' o [NP pneu]]]]]]]]]

Munhoz and Naves also claim that since α cannot assign Case and T is defective, both DPs in these constructions receive default Nominative Case in the morphological component (instead of doing so in the syntactic derivation).

The authors do not explain explicitly how the argumental DP is formed in genitive topic-subject constructions and why it splits in two DPs that occupy different positions in the final structure. Following their own reasoning for locative constructions, I guess there must be two different numerations as well, one with the preposition 'de' and one without it. When the preposition is in the numeration, it merges with its complement (the possessor) forming a PP that is then adjoined to the NP (the possessed). When there is no preposition, I suppose they assume an analysis analogous to that of the English genitive, in which the possessor is the specifier of the DP whose head is the genitive mark and that has the possessed as the complement, see (37). This leaves the possessor DP in a position from where it can raise to higher layers of the structure.

(37)   [DP Mary [D' 's [NP car]]]   English
         [DP [DP o carro] [D' o [NP pneu]]]   BP

Apart from this, as Andrade & Galves (2014) point out, there are two main problems in this analysis:

i)   *it does not account for the fact that both configurations* [locative and genitive] *allow the movement of some constituent to become the logical subject of the predication in BP;*

ii)   *it renders the correlation between the subject topic construction and the argument structure of the verbs involved completely opaque.*

(Andrade & Galves, 2014, p. 121-122)

Let us now see what these authors propose to solve the problems in Munhoz and Naves' analysis.

2.4.2. *Andrade & Galves (2014)*

Andrade and Galves notice there are two aspects of topic-subject constructions that should be explained, namely, why genitive and locative constructions are licensed by different types of verbs (what they call the distribution problem), and why these constituents trigger verb agreement, but only in the absence of a resumptive pronoun (what they call the agreement problem).

These problems are closely related to where topic-subjects come from and where do they end up in the syntactic structure. The paper focuses on the first issue, which the authors claim had been neglected in the literature. As for the second one, they adopt Avelar & Galves (2011) proposal and assume Spec-TP as the final position. I return shortly to their proposal in the next subsection, 3.4.3.

Andrade and Galves begin by stating that the distribution of both genitive and locative topic-subject constructions is not only a matter of verb class. To ensure nominative Case is available for the topic-subject constituent, the verb must be unaccusative. But this is not enough. There must be a semantic relation between the theme and the genitive or locative constituent, namely, part-whole or inalienable possession in the first case, content-container in the second. They propose these semantic relations are syntactically encoded as secondary predicates or small clauses (SC)[29]. So, the verb must also select a secondary predicate.

On one hand, this means the syntactic structure underlying inalienable possession is different from that of alienable possession and modification. This explains the ungrammaticality of genitive topic-subject constructions lacking this particular semantic relation, that is, constructions in which theme and genitive do not form a secondary predicate. See the following example:

---

[29] A small clause is a constituent consisting of a subject and a predicate but lacking tense inflection. See the following examples compared with their full clause counterparts. (Citko, 2011)

(i)	I consider [Mary smart]	vs.	I consider [Mary to be smart]
(ii)	I consider [Mary my best friend]	vs	I consider [Mary to be my best friend]
(iii)	I consider [Mary out of her mind]	vs	I consider [Mary to be out of her mind]

(38)   a. *A mesa quebrou o pé*.

the table broke.3SG the foot.

'The table leg broke.'
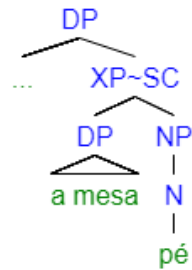
b. **O João pifou o carro*.[30]

the João broke.3SG the car

'João's car broke.'
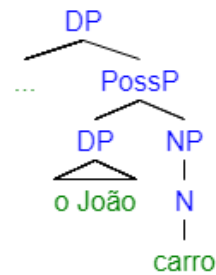
c. **Essa mesa quebrou o pote*.

this table broke.3SG the pot

'This table pot broke.'
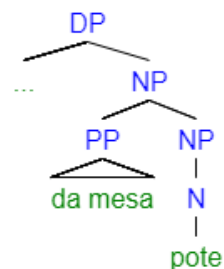
On the other hand, the authors show there is also a different structure underneath sentences with locative arguments. When the verb is unaccusative and selects a secondary predicate with a

---

[30] Andrade and Galves clarify this sentence can be grammatical with a slightly different meaning 'João broke his car', that is, if João is considered the agent. The verb *pifar* ('to break') is ergative and can admit both unaccusative and transitive constructions.

locative relation, the locative is attached inside the small clause (39.a). When the verb is unergative or transitive, the locative is attached at VP/RootP level (39.b).

(39)   a. *Caber muita gente no carro*

          'to fit many people in the car'
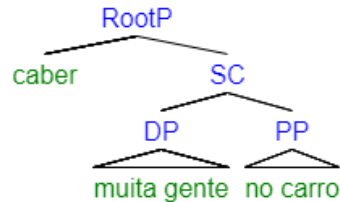
      b. *Colocar muita gente no carro*

          'to put many people in the car'
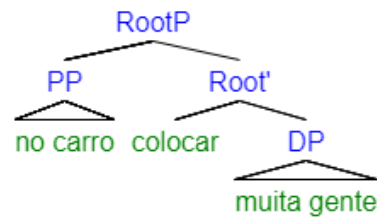
When the locative is an external modifier, it is interpreted as the endpoint of a dynamic event, resulting in a perfective reading. In topic-subject constructions, the locative is an internal modifier and cannot provide the event with internal dynamism, which is therefore imperfective. According to Andrade and Galves, this explains why topic-subject constructions sound more natural when the verb's aspect is imperfective, as shown below.

(40)   a. ?*Aquele consultório chegou um paciente*.

          that (doctor's) office arrived.3SG a patient

          'A patient arrived at that doctor's office.'

      b. *Aquele consultório chega paciente todos os dias*.

          that (doctor's) office arrive.3SG patient all the days

          'Patients arrive at that doctor's office every day.'

As evidence in favor of their proposal, the authors show there are, cross-linguistically, two different types of small clause, one expressing an integral relation (similar to part-whole) and the other one expressing a spatial relation (similar to content-container). This can be seen in existential predicates such as (41), which can be interpreted as (42.a) or (42.b) depending on the small clause underlying structure. The authors bring examples taken from (Hornstein, Rosen & Uriagereka, 2002).

(41)   There is a Ford T engine in my Saab.

(42)   a. My Saab has a Ford T engine.

            [$_{SC}$ My Saab [a Ford T engine]]

            b. (Located) in my Saab is a Ford T engine.

            [$_{SC}$ a Ford T engine [in my Saab]]

Andrade and Galves also present an empirical argument to support their analysis. Quantifier raising is not allowed in small clauses, so sentences containing a small clause and two quantifiers do not show the usual ambiguity due to scope. They show this is exactly what happens with topic-subject constructions, only the distributive reading is available, as can be seen in (43.b) and (44.b).

(43)   <u>Alienable vs. inalienable possession</u>

            a. *Todos os alunos quebraram duas garrafas.* ∀>2; 2>∀

            all the students broke.3PL 2 bottles

            'All the students broke two bottles.'

            b. *Todos os carros quebraram duas lanternas.* ∀>2; *2>∀

            all the cars broke.3PL 2 lanterns

            'All the cars broke two lanterns.'

(44)   <u>External vs. internal locatives</u>

            a. *Por todas as vans passaram seis alunos.* ∀>6; 6>∀

            by all the vans passed.3PL 6 students

            'Six students passed by all the vans.'

            b. *Todas as vans cabem seis alunos.* ∀>6; *6>∀

            all the vans fit.3PL 6 students

            'All the vans fit six students.'

Next, the authors claim that the movement restrictions (28)-(30) shown in Munhoz & Naves (2012) are only apparent counterexamples to the unified analysis they are proposing. They argue that the differences between locative and genitive constructions can be explained in terms of the internal structure of the small clause.

They adopt Den Dikken's proposal (2006) and assume the small clause corresponds to a Relator Phrase (RP), where RELATOR is a functional projection, and there is no fixed ordering for the

small clause predicate and its argument[31]. In the case of genitive and locative topics, they show a predicate-complement configuration and their underlying small clause base structures are the following:

(45)   a. <u>Part-whole relation</u> – *o pé da mesa* ('table leg')

RP
├── DP
│   └── a mesa
└── R'
    ├── R
    └── NP
        └── pé
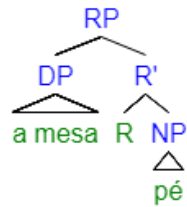
b. <u>Content-container relation</u> – *muita gente no carro* ('many people in the car')

RP
├── DP
│   └── muita gente
└── R'
    ├── R
    └── PP
        └── no carro
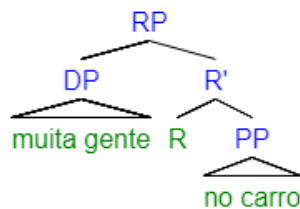
Then the authors proceed to show how these base structures are derived to form topic-subject constructions. In the case of genitive small clauses, a DP is projected on top of RP[32]. Due to locality restrictions, only the small clause subject, the DP in Spec-RP, is able to move to Spec-DP, a movement that is discourse-motivated. From there, at the edge of the DP, which constitutes a phase, it can further move to reach Spec-TP and become a topic-subject constituent. This option is not available for the small clause predicate, thus explaining some of the above-mentioned movement restrictions.
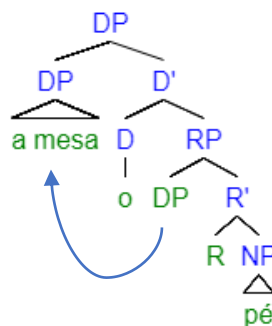
DP
├── DP
│   └── a mesa
└── D'
    ├── D
    │   └── o
    └── RP
        ├── DP
        └── R'
            ├── R
            └── NP
                └── pé

In the case of locative small clauses, no DP is projected on top of RP. This contrast with genitive small clauses is . Since RP is not a phase, both the subject and the predicate can move to reach Spec-TP, hence the possibility of sentences like *Muita gente cabe nesse carro* ('Many people fit in that car') and also *Nesse carro cabe muita gente* ('(In) that car fits many people'). But there is another possibility, that P is incorporated to R, for which they suppose the existence of a null P in the lexicon. In this case, the locative DP may raise to subject position and become a topic-subject constituent.
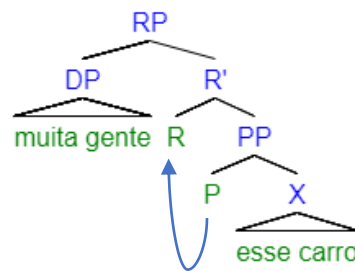
As implications of this analysis, Andrade and Galves mention that genitives lacking part-whole relation (those encoding alienable possession or modification) do not allow topic-subject constructions, as shown in (38). Similarly, locatives functioning as external modifiers are not part of a small clause, so they lack the RELATOR head. This makes P-incorporation not possible, as shown in (46) below.

(46)   a. **(D)essa casa eu não saio*.

         (of)that house I not leave.1SG

         'I don't leave from that house.'

     b. **(N)essa sopa eu coloquei sal demais*.

         (in)that soup I put.1SG salt too.much

         'I put too much salt in that soup.'

Finally, the authors claim that unaccusative verbs in EP have the same argument structure, so lack of topic-subject constructions in this language must be explained. They propose the absence of these constructions is due to dative Case valuation (by means of *a*-marked genitives and dative clitics) and to the inexistence of null Ps, which makes P-incorporation impossible.

However, this explanation has some problems. Genitives can show dative Case in EP and other Romance languages, but this is not always necessary, both options are available, especially if the whole DP is in pre-verbal position. See the following examples I propose:

(47)   a. *Apodreceu a raiz **às** árvores.*

           Rotted the root to.the trees

      b. *A raiz das árvores apodreceu.*

           the root of.the trees rotted

      'The trees roots were rotten.'

(48)   a. *Se les han podrido las raíces **a** los árboles.*

           SE CL.3PL had rotted the roots to the trees.

      b. *Se han podrido las raíces de los árboles.*

           SE had rotted the roots of the trees

      'The trees roots were rotten.'

In addition, datives in (47.a) and (48.a) suggest there may be a more complex argument structure, since part-whole relation does not intrinsically need *a*-marked genitives and dative clitics.

So, the only thing blocking topic-subject constructions would be the lack of null Ps. This can also be problematic, since locatives DP in EP can have no preposition and move to the left periphery of the sentence, just like in BP, but without triggering verb agreement.

(49)   a. *Esses carros, furou o pneu.*

           those cars punctured the tire

           'As for those cars, they got a flat tire.'

      b. *Essa árvore, caíram as folhas.*

           that tree fell.3PL the leaves

           'As for that tree, its leaves fell off.'

The main issue with their argument, hence, is that it does not explain why the topic agrees with the verb in BP, but it does not do so in EP or in other Romance languages. However, as I mentioned in this subsection's introduction, Andrade and Galves (2014) are only concerned with what they call the distribution problem, that is, the licensing conditions for topic-subject constructions, what verbs allow these constructions and how their argument structure looks like. They are and not really interested in the agreement problem and they are fully aware of this, as openly discussed in the beginning of their article.

As we are about to see, the focus of Avelar and Galves (2011, 2016), on the other hand, is the final position of the topic-subject constituent and especially why it triggers verb agreement, and they are not particularly interested in where this constituent comes from.

2.4.3. *Avelar & Galves (2011, 2016)*

Based on Chomsky (2008) and Holmberg (2010), Avelar and Galves explore the idea that topic-subject constructions and other BP syntactic peculiarities related to agreement and Case can be accounted for by assuming that the EPP feature in T is phi-independent in this language, and that DPs can be inserted in the derivation without a Case feature.

The authors adopt Chomsky's proposal for C-to-T phi-feature transfer, which means that T does not inherently possess phi-features, but it inherits them from C (which explains why T is defective when it is not projected as a complement of C, as, for example, in ECM constructions: *I want [$_{TP}$ you to be happy]*). The role these phi-features play in the creation of Spec-TP in a particular language determines whether T's EPP feature is phi-independent or not.

If the creation of Spec-TP is a necessary condition for the satisfaction of T's phi-features, then T's EPP is phi-dependent. In this case, Spec-TP can only be created after C has been merged and T inherits C's unvalued phi-features, which are then valued against those of the element in Spec-*v*P. This is what happens in English, EP and other Romance languages.
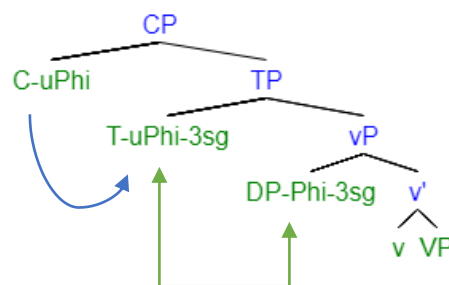
If, on the other hand, the creation of Spec-TP occurs independently of T's phi-features, then T's EPP is phi-independent. This means Spec-TP is created as soon as T is merged, which inherits C's phi-features after they have been already valued against those of the element in Spec-TP. According to the authors, this is the case of BP.
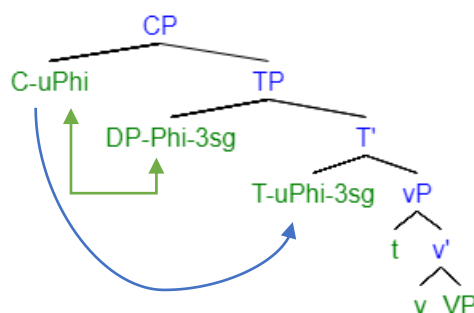
Therefore, in a language where T's EPP is phi-dependent, Spec-TP is created by movements triggered by phi-features and thus it is an A-position. In a language where T's EPP is phi-independent, however, Spec-TP is created by movements that were not triggered by phi-features (movements triggered by T's EPP itself), hence, it is an A'-position. This means in a language like BP, non-argumental constituents can occupy the subject position.

In addition to the fact that BP allows verb agreement with non-argumental DPs (or even with PPs), Avelar and Galves show hyper-raising sentences (50) and ambiguity in tough-constructions (52) constitute further evidence of T's EPP phi-independence in BP.

(50)   *Os carros parecem que o pneu não foi trocado.*
        the cars seem.3PL that the tire not was replaced
        'It seems this car's tire was not replaced.'

In this sentence, the DP *Os carros* moved from Spec-TP to Spec-TopP in the embedded clause, and from there to Spec-TP in the matrix clause, see (51). This is allowed in BP because, due to T's EPP phi-independence, Spec-TP is an A'-position and hence the movement is uniform.

(51) [$_{TP}$ [$_{DP}$ os carros]$_i$ [$_{T'}$ parecem … [$_{CP}$ que [$_{TopP}$ t$_i$ Top [$_{TP}$ [$_{DP}$ o pneu t$_i$] [$_{T'}$ não foi trocado]

For a similar reason, BP tough-constructions like (52) admit a second interpretation (the one in b) in which the subject of the matrix clause is interpreted as the external argument in the embedded clause.

(52)   *O João é difícil de agradar.*
        the João be.3SG tough of to.please'
    a. 'It is tough to please João.'
    b. 'It is tough for João to please somebody.'

Here, to be able to reach Spec-TP of the matrix clause, the DP *O João* needs to move first from Spec-*v*P to Spec-CP in the embedded clause, see (50).

(53)   [$_{CP}$ C [$_{TP}$ [$_{DP}$ o João]$_i$ [$_{T'}$ T ... [$_{CP}$ t$_i$ [$_{C'}$ de [$_{TP}$ [$_{vP}$ t$_i$ agradar]]]]]]]

Since this first movement is not triggered by phi-features but by C's edge-features, it constitutes an A'-movement. The second movement is allowed in BP because Spec-TP is an A'-position, which makes the movement uniform in this language.

But let us come back now to a topic-subject construction like (54), in which verb agreement with the topic-subject constituent is evident.

(54)   *Os carros furaram o pneu.*

the cars punctured.3PL the tire

'The cars got a flat tire.'

Avelar and Galves propose this sentence is derived as follows. Since the DP constitutes a phase, they begin by assuming the non-argumental topic *os carros* moves to the edge of the DP it modifies so as to be able to move further up the structure later. Right after T is merged, Spec-TP is created and occupied by the DP *os carros.* Then, C is merged and its phi-features are valued via agreement with the DP in Spec-TP, after which they are transferred to T. The derivation follows below.
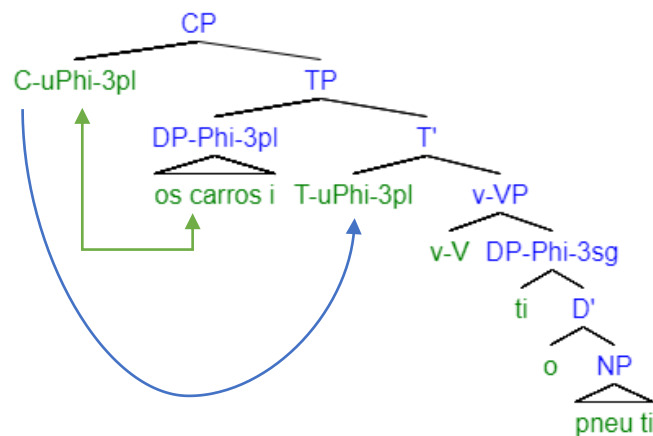
Notice here that, although they do not discuss the details of the argumental DP's original structure, their analysis is compatible with the proposal in Andrade & Galves (2014) that considers this argument as a small clause.

According to the authors, in languages for which T's EPP is phi-dependent, like EP and other romance languages, this derivation is not possible. Since in these languages Spec-TP is only created after C is merged, the DP *os carros* can no longer move to this position because the phi-features in the maximal projection of the DP that dominates it (*o pneu*) are detected before by T's unvalued phi-features, inherited when C was merged.

There is still a loose end that needs our attention. In topic-subject constructions like (54) there are two DPs but only one source of Case. T's phi-features assign Case to the pre-verbal DP, but since the verbs in these constructions are unaccusative, they cannot assign Case to the post-verbal DP.

To solve this problem, Avelar and Galves propose that in BP, DPs can be inserted into the derivation with or without a Case feature. They suggest that topic-subject constructions are licensed only when the complement DP lacks a Case feature. This means T's phi-features do not need to agree with it or assign Case to it, so they are free to do so with the non-argumental topic.

In support of this idea, the authors argue BP shows a morphological uniformity between pronouns in nominative and non-nominative positions. When they lack a Case feature, they appear in its nominative form by default, even if they occupy a non-nominative position, as can be seen in (55)[33].

(55)   a. *Você foi visto na escola*.

          you.NOM were seen in.the school

          'You were seen at school.'

        b. *A Maria viu você na escola*.

          the Maria saw you.NOM in.the school

          'Maria saw you at school.'

---

[33] BP pronouns can also appear in its accusative form, as can be seen below. In this case, they were inserted with a Case feature. Notice, however, that this form is only available for first person, both singular and plural, and for second person singular.

(52')   a. A Maria *me* viu na escola.   d. A Maria *nos* viu na escola.
        b. A Maria *te* viu una escola.   e. ?A Maria *vos* viu na escola.
        c. ?A Maria *o* viu una escola.   f. ?A Maria *os* viu na escola

The authors argue as well that variation in subject-verb agreement, such as that exemplified in (56), can also be accounted for in terms of their proposal. Insertion of the subject DP with an unvalued Case feature enables agreement with T's phi-features; insertion without a Case feature disallows this agreement; hence giving raise to the two options.

(56)   *As criança(s) brincavam/brincava na varanda.*
          the children played.3PL/played.3SG in.the veranda
          'The children played on the veranda.'

In Avelar & Galves (2016), the authors go on from here to analyze the differences between EP and BP discussed above from a diachronic point of view. They defend the idea that BP main peculiarities are due to linguistic contact with Bantu languages. And they propose the two parameters involved in the change that led to BP, namely T's EPP phi-independence and the possibility to generate DPs without a Case feature, are part of an Agreement/Case parametric tree in the line of those proposed by Roberts & Holmberg (2010) and Roberts (2012).

2.5. Considerations for the present study

As I mentioned in the introduction, despite their technical differences, the proposals discussed in the previous sections have something in common: They postulate a special position that receives topic-subject constituents and marks a contrast between BP and all the other Romance languages.

In Galves (1998), this position was Spec-PersP. The difference between BP and other Romance languages in this case resided in the parametrization of the proposed functional category Person, projected on top of T, that lacked both V and AGR formal features in this language. Therefore, Person did not project a specifier and its head could not receive the verb either, so its phi-features could only be checked by the movement of phi-features on their own. In the special case of topic-subject constructions, since the verb does not accept an external argument, T's phi-features were transferred to Person, which only then could project its specifier and thus receive the topic-subject constituent.

In Munhoz & Naves (2012), the position receiving this constituent was Spec-αP, the specifier of a functional category between C and T only present in Topic-prominent languages. The difference between BP and the rest of Romance languages here is that the latter, being Subject-

prominent languages, lack this projection, whereas the former, with its in-between behavior, can sometimes project it, in which case, C's phi-features are transferred to α instead of T.

Finally, in Avelar & Galves (2011, 2016) it was Spec-TP that received the topic-subject constituent. BP's peculiarity in this case is not the position itself but how and when it is created. Unlike in the other Romance languages, in BP this is done before C is connected to the structure and without the action of phi-features. Together with the assumption that DPs can be generated without a Case feature in BP, this allows verb agreement with non-argumental constituents, as it happens with topic-subject constructions.

As mentioned before, Andrade & Galves (2014) were not really concerned with the special position receiving the topic-subject DP, their focus was on the structure of the internal argument where the constituent originally comes from. But to fully understand BP topic-subject constructions and what their derivation is like, we need to understand both parts of the problem, we need to know what the starting point and the place of arrival are, and how and why the DP gets from one to the other. I believe that the combination of both proposals (Andrade & Galves, 2014; and Avelar & Galves, 2011, 2016) provides a complete explanation of the phenomenon.

Although there are some instances of locative inversion in Spanish, their occurrence is much more restricted than in BP. As discussed in Fernández-Soriano (1999), this construction is only licensed with two types of impersonal verbs, and despite the author argues that locatives occupy the subject position (more specifically, Spec-TP) in these sentences, there is never verb agreement with the locative constituent, as can be seen in the following example. Notice, also, that locatives need to keep its preposition, because DPs cannot be inserted without a Case feature in this language.

(57) a. *En esta casa faltan ventanas*.
          in this house miss.3PL windows
          'Windows are missing in this house.'
      b. *En Barcelona ha ocurrido un accidente*.
          in Barcelona have.3SG happened an accident
          'An accident has happened in Barcelona.'

What Fernández-Soriano proposes to account for these sentences in Spanish is that T's EPP feature and its phi-features agreement/nominative Case assignment can be fulfilled by different elements. She argues that a non-nominative constituent that does not trigger verb agreement

can satisfy T's need to project its specifier, and that nominative Case can be assigned to a posverbal DP with the appropriate features thus triggering verb agreement with this constituent.

In this dissertation, I will assume Fernández-Soriano's analysis for Spanish and both Andrade & Galves and Avelar & Galves' proposals for BP. Following Holmberg's reasoning for English (Holmberg, 2010, p.112), this phenomenon is too highly restricted in Spanish to suggest T's EPP phi-independence for this language. However, I will resume this discussion in Chapter 7 when reviewing the results of the experiment on which this study is based and their theoretical implications.

Chapter 3
On Second Language Acquisition

3.1. Introduction

Second language acquisition (from now on SLA) has long been a matter of interest for its educational and social relevance in multilingual societies. In western culture, for example, scholars had to reflect about this topic and conceptualize its nature as early as the sixth or seventh century, when the expansion of the Roman Empire turned Latin into a second language for almost half of Europe, the Northern coast of Africa and even parts of the Middle East. Certainly, other cultures have speculated too about the essence of second language learning since very early on. However, the origins of SLA systematic research are conventionally dated to the 1950s, when the field was identified as a scientific discipline. (Thomas, 2013)

At this early stage, the motivation behind SLA studies was the improvement of teaching methods. By the 1980s, though, the field had gained enough weight and theoretical relevance so that making references to pedagogical applications of research findings was no longer necessary, unless this was the specific aim of the study. As Cook (1986) pointed out, apart from the improvement of language teaching, the other three main reasons for research interest in this field were the investigation of the process of SLA in itself, contribution to wider issues in Linguistics and in the Linguistic Theory of Language Acquisition, and contribution to general issues in Psychology or, more precisely, in Psycholinguistics.

In fact, SLA is a very intricate and complex phenomenon that inevitably calls for different theoretical approaches and methodological tools depending on the particular dimension of the process that the researcher wants to address, but also on his/her views on the nature of language, on the acquisition process and on the learner's role in it. Not one approach alone has yet been able to capture the many facets of SLA. (Myles, 2013).

Sociocultural approaches focus on the social aspect of acquisition, viewing language as a tool for communication and thought in a social context; interactionist perspectives are interested on interactional patterns, on the role of input and output in the learning process and on the learners' negotiation strategies in the communication process; sociolinguistic approaches focus on the social and pragmatic aspects of language in use, from variation to identity construction. These research perspectives put context in the foreground. Differently from those, cognitive approaches are interested in the learner's mind, on processing skills and individual differences that influence the acquisition process. And yet another way of looking at SLA comes from formal or linguistic approaches, which focus on the linguistic system underlying the learner's performance, on its formal properties and their relationship with those of the native language, of the target language and with universal properties of human languages in general. Obviously, some of these approaches can sometimes overlap. (Myles, 2013).

Here, I will adopt a linguistic approach, in particular, a Universal Grammar approach. But this restriction of our theoretical perspective is not yet enough. Within the Generative framework there are still many different, and even contradicting, theories and hypotheses trying to account for the formal properties of the learners' linguistic systems and their development. These theories vary as to the dimension of the system on which the focus is placed (phonology, morphology, syntax, semantics, etc.) and to the assumptions made regarding the central issues of SLA from this perspective, such as the role of the native language and of UG, the influence of age of exposure, or the nature of the linguistic system(s) at different developmental stages.

In this chapter, I present an overview of some of the hypotheses put forth during the last couple of decades of SLA research, especially those most relevant for the syntactic domain. I will follow White's discussion of these theories regarding the advances made in the 1990s (White, 2003), and I will complement it with the more updated discussion provided by several authors in Herschensohn & Young-Scholten (2013).

Having provided a general outlook of SLA, I go on to review this field's most representative research on the syntactic phenomena that interest us here. Regarding acquisition of wh-movement (or of wh *in situ*), I devote section 3.4 to the discussion of three different approaches to this topic. Taking a Representational Deficit approach, Hawkins & Chan (1997) and Hawkins & Hattori (2006) are two of the more representative studies, respectively arguing in favor of the Failed Functional Features Hypothesis and the Interpretability Hypothesis. Hettiarachchi & Pires (2016) replicate Hawkins and Hattori's experiment and claim their results are better explained adopting a Full Access perspective. Finally, Choi (2016) studies the problem

reversing the acquisition direction (from a fronting L1 to an *in situ* L2) and interprets the results as evidence in favor of the Feature Reassembly Hypothesis.

As for topic-subject constructions, there is no SLA literature that focused on these structures, so section 3.5 is a very brief review of two representative articles providing two opposite accounts for the role of topic/subject prominence typology in SLA. On one side of the debate, Fuller and Gundel (1987) argue in favor of a universal early topic-comment stage and against L1 transfer. On the other side, Jung (2004) defends that early interlanguage grammars show typological transfer from the L1.

3.2. Some preliminary remarks

Before entering the discussion about the different hypotheses available in Generative SLA literature, I believe it is necessary to clarify some terminological choices made throughout this work which actually entail deeper theoretical assumptions.

3.2.1. *Second vs. third language acquisition*

When speaking about a second language (L2), it is quite common to use the term as referring to any language acquired in adulthood or subsequent to the native or first language (L1). This means that, chronologically, it could in fact be the third, fourth or nth language acquired. To this day, many researchers use the term in this way, and, for the present study, I am going to include myself in this group too.

Nowadays, English can be considered the predominant language for business and scientific exchange worldwide, so it is quite difficult to find 'true' L2 speakers of other languages, unless in contexts of multilingualism. This is the case of Brazilian Portuguese as a 'second' language, it is practically impossible for BP learners to not have had previous experiences with English or other language traditionally studied as L2 (like Spanish or French, for example).

The theoretical distinction between the acquisition of a second and a third language can be very important, though, and this subfield has seen an increase in rigorous studies over the last two decades.

For example, third language acquisition allows for interesting research about the initial state of the acquisition process and the role that all other previously acquired languages have in it. This

matter has called the attention of many researchers interested in L3 syntax, and different hypotheses or models have been advanced. The **L2 Status Factor Hypothesis** proposed by Bardel and Falk (2007) claims that, due to the similarities in L2 and L3 acquisition, L2 plays a major role in the initial state of L3 grammar, acting as a filter and blocking access to the L1 grammatical features. Other models suggest transfer may come from all previously acquired languages. The **Cumulative Enhancement Model** put forth by Flynn, Foley and Vinnitskaya (2004) predicts that transfer will come from any previously acquired language as long as it has a positive or facilitative effect. Rothman's **Typological Proximity Model** (2010) claims transfer is constrained by (either actual or perceived) typological proximity between grammars. These models can be tested against each other, as can be seen in Rothman (2010) and Ionin et. al (2015) for BP, for example.

Since the aim of this study does not fall into third language acquisition research agenda and given the fact that the participants' specific previously acquired languages do not differ in any of the relevant grammatical aspects, I will not make a distinction between L2 and L3/Ln and will refer to BP as the participants L2, without forgetting nonetheless that this is an oversimplification.

3.2.2. *Interlanguages vs. Multicompetence*

Somewhat connected to the L2/L3 distinction is another issue, the nature of the linguistic system or systems in the speaker's mind related to the knowledge of two or more languages.

According to White (2003) and Thomas (2013), the concept of interlanguage was proposed during the 1960s and early 1970s by several researchers, like Corder or Selinker, who actually coined the term in his 1972 article 'Interlanguage'. These researchers were concerned with the significance of the errors learners make when acquiring a second language, arguing they suggested a rule-governed systematic behavior similar to that of a L1, independently of the system's degree of approximation to the native-speakers grammars of that language (Corder, 1967). In Selinker's words, an *interlanguage* is a 'separate linguistic system based on the observable output which results from a learner's attempted production of a target language norm' (Selinker, 1972, p. 214).

After this initial interest in accounting for the systematicity found in L2 learners' output, researchers adopted this notion and, by the 1980s, they had taken to testing hypotheses on the

basis of the presence or absence of evidence for a determined theoretical constraint in learners' interlanguage (Thomas, 2013). The concept of interlanguage was particularly productive in Generative SLA research and it is still used to this day.

However, a very interesting new perspective surfaced in the 1990s with the notion of *multicompetence*. According to Rothman et al. (2013), the term was introduced by Cook in 1992 and it constituted a rupture with the traditional view and a reaction against the notion of *interlanguage*. Multicompetence sees the different languages in a speaker's mind as being closely interconnected and part of a macrosystem, rather than considering them separate entities. As a consequence, transfer or crosslinguistic influence is no longer seen as one-directional (from L1 to L2) and an acquired language is no longer considered a stable system, changes in one linguistic subsystem may affect all the others (including the native language). Studies about attrition have shown that any language may be subject to decline if it stops being used (Schmid, 2010; de Bot & Weltens, 1995) and that growth and decline are influenced by sociolinguistic and psycholinguistic factors (Clyne, 2003). (Rothman et. al, 2013).

Since I am going to review Generative SLA literature, the term *interlanguage* will appear very frequently throughout the chapter. For simplicity, I will also adopt it for the discussion of this study's results, keeping nonetheless an open mind about the notion of *multicompetence* in case it should become a more interesting approach to the theoretical discussion of our data.

3.3. SLA theories and hypotheses within the Generative framework

Generative SLA research exhibits a wide range of theoretical perspectives regarding the characteristics of the acquired grammar, mainly differing in the role assigned to UG within the underlying linguistic knowledge of L2 learners, the assumed initial state and/or the influence of the native language.

The notion of UG originated as an explanation for the *poverty of the stimulus problem* in native language acquisition, that is, the disparity between the linguistic input a child receives and the complexity of the system that results from the acquisition process, the grammar. It is commonly assumed that this problem is also present in SLA, with L2 speakers showing abstract and subtle properties of the L2 grammar that could not have been learned from the input or formal instruction alone.

Earlier Generative SLA research focused on whether access to UG was also a possible explanation for the *poverty of the stimulus problem* in SLA. Later, the focus of the discussion changed from the availability of UG to the nature of the interlanguage grammars and whether they behave like natural languages (i.e. whether they are constrained by UG) or not.

To start this section, I will try to summarize White's discussion (2003), in which the author presents many of the theories put forth in the field up to that point and provides evidence to support or refute each one. As a possible means for classification, we can consider two broad and opposite views that encompass different proposals and hypotheses. One assumes interlanguage grammars as being always constrained by UG, the other does not.

A representative proposal arguing against UG access is the **Fundamental Difference Hypothesis** (Bley-Vroman, 1990). This hypothesis considers first and second language acquisition as fundamentally different processes, the former being quick and successful, the latter being slow and variable among speakers. Within this perspective, it has been proposed that interlanguage grammars are only weakly UG-constrained by means of L1 reconstructed properties (Schachter, 1999), or that they show no evidence of UG constraints whatsoever, for example, that they are construction specific (Clahsen & Hong, 1995; Neeleman & Weerman, 1997) or resort to linear sequencing strategies (Meisel, 1997). This means each L2 construction must be learned separately, there are no parameters allowing the clustering of properties, and acquisition is achieved through *pattern matching*, i.e. focusing on form or surface properties. All these views presuppose a global impairment on interlanguage grammars.

Other perspectives assume a more local and/or temporary impairment. The **Minimal Trees Hypothesis,** put forth by Vainikka and Young-Scholten (1994), defends that interlanguage grammars completely lack functional categories at their early stages. According to this hypothesis, L1 lexical categories and some related properties are present in the initial state, but no functional category is (neither coming from L1 nor from UG itself). Triggered by the L2 input, though, functional categories gradually emerge in the interlanguage grammar, following a pre-determined bottom-up sequence in which structurally deeper categories are acquired before more peripheric ones, in a process called Structure Building. They claim the lexical stage VP constitutes the initial state, the next stage corresponds to a functional category FP - finite phrase (which does not exist in natural languages), in the following stage IP replaces FP, and, lastly, CP is acquired at the final stage.

Still assuming local impairment, Eubanks's **Valueless Features Hypothesis** (1993/1994) claims that both L1 lexical and functional categories are transferred to the interlanguage grammar in its initial state, but feature strength is not. He claims features are valueless in the beginning of the acquisition process, which allows, for example, for optional verb raising in L2 grammars of languages with strong or weak features, like French and English respectively. In Eubank's proposal, though, L2 feature strength is acquired once morphological paradigms are acquired too.

However, there is a more extreme position regarding feature strength. The **Local Impairment Hypothesis** argues that features remain permanently inert in interlanguage grammars (Beck, 1998), which means some parameters are never set. Contrary to Eubank's proposal, Beck claims there is no causal relation between the acquisition of inflectional morphology and the acquisition of feature strength. Even when L2 learners show accurate overt morphology, verb raising remains optional.

In any case, a consequence of both global and local impairment is that interlanguage grammars (whether permanently or only during early stages of the acquisition process) are not fully UG-constrained, so they may show a behavior that differs from that of natural languages.

On the opposite side of the debate, interlanguage grammars are considered to be unimpaired, they behave like natural languages and can be characterized in terms of parameter settings. The possibility of parameter setting or resetting sets apart three groups of proposals within this wider perspective, those assuming only L1 settings are available in interlanguage grammars, those for which L2 settings are directly achieved through L2 input, and those arguing in favor of a resetting process from the L1 values to the L2 values.

For the **No Parameter Resetting Hypothesis** the only available parameter settings are those of the L1, meaning that learners' representations for L2 data need to be accommodated within the L1 grammar with ad hoc local modifications (Liceras, 1997) or based on the L1 grammar but using UG constraints (Tsimpli & Roussou, 1991). Sharing this last proposal, Hawkings and Chan (1997) adopt the **Failed Functional Features Hypothesis**, which suggests it is impossible for L2 learners to acquire functional features not instantiated in the L1. Their study was focused on the [wh] feature of English and the properties of relative clauses associated to this feature, comparing Chinese and French native speakers. According to the authors, Chinese speakers analyze L2 data based on their L1 and following UG constraints. I will further discuss their proposal in the next section.

The **Full Access Hypothesis**, which White (2003) refers to as **Full Access Without Transfer**, was proposed by Epstein, Flynn and Martohardjono (1998). Contrary to the Fundamental Difference Hypothesis, this proposal considers there are strong similarities between the first and second language acquisition processes, and, based on that, it argues UG is fully accessible in SLA at all stages. According to White (2003), the authors explicitly reject L1 as being the initial state and she argues that, as a consequence of their reasoning, UG itself constitutes SLA's starting point (although they do not make this assumption explicit). This means parameters will be set to L2 values as the relevant L2 input is received.

Finally, Schwartz and Sprouse (1994) put forth the **Full Transfer Full Access Hypothesis**, which states that the L1 grammar (except specific lexical items) constitutes the initial state. According to this hypothesis, the abstract properties of the interlanguage grammar are initially equal to those of the L1. As the L2 input is received, and if the L1 grammar should fail to satisfactorily analyze it, new parameter settings, functional categories or feature values are incorporated into the interlanguage grammar, but always within the possibilities sanctioned by UG. These new settings provide a more appropriate analysis for the L2 data; however, they do not necessarily converge to those of the L2 grammar.

Many researchers have tested these hypotheses, finding evidence that constitutes arguments or counterarguments for each of them. In some cases, depending on how it is interpreted, the same data can support or refute the same hypothesis. Throughout her book, White (2003) advocates Full Transfer Full Access providing evidence in favor of this hypothesis from many different studies and re-interpreting the evidence provided by others to show problems in their arguments.

White (2003) does not discuss age effects in SLA, but it is a much-debated issue. The idea of a critical period for first language acquisition was introduced by Lenneberg in 1967 and generated profuse discussion, but it is widely accepted in the scientific community given the amount of evidence pointing in its favor. This idea was extended to SLA on the basis of the inverse correlation found by many studies between age of onset of acquisition and proficiency. Herschensohn (2013) sorts research on age effects in SLA into two different approaches. One considers adult and child SLA as categorically different (Bley-Vroman's Fundamental Difference Hypothesis, 1990; DeKeyser 2000). The other considers the differences are more gradient (Singleton & Ryan, 2004; Montrul, 2008; Herschensohn, 2009). She claims that true critical periods are strictly biological and linked to neural maturation, whereas a range of non-biological factors affect SLA. And although evidence points to a maturationally *sensitive* (as

opposed to *critical*) period for SLA, with offset starting as soon as age 4 and steeper decline occurring through adolescence, there is no definitive terminus.

The notion of transfer is another issue that, according to Foley and Flynn (2013), has recently aroused interest among researchers. Some studies, like Montrul's research on change-of-state verbs (2009), have considered the possibility of the L1 having a different role depending on the module of language knowledge, showing, for example, a deeper influence in overt verbal morphology than in argument structure. Whong-Barr (2006) suggested that derivational processes may also transfer and argued more explicit research on the nature of transfer is needed.

According to Vainikka and Young-Scholten (2013), SLA theories and hypothesis put forth within the Generative framework until the early 2000s have mainly neglected to study the role of development or stages in SLA process. Following the Structure Building reasoning proposed in their Minimal Trees Hypothesis in 1994, they have continued to research stage-like development in L1 and L2 acquisition and have come to develop their theory of **Organic Grammar**. This theory posits different stages of development corresponding to functional projections (NegP, TP, AgrP, CP, etc), which are predicted to follow the same order found in the adult syntactic tree starting from the deepest layer. According to Organic Grammar, the acquisition of a particular functional projection implies the acquisition of other syntactic and morphologic phenomena related to this projection. Due to economy constraints, the learner can only acquire projections for which there is overt evidence. And transfer from the L1 is only possible at the very beginning of the process, i.e. only the VP projection containing the verb and its arguments can be transferred. Within this view, inflectional morphology plays an extremely relevant role, being involved at each stage of the acquisition process. This is precisely what detractors of the theory use to argue against it, because L2 learners showing evidence of having acquired a particular functional projection may still fail to consistently produce target inflectional morphology.

Another difference between generative SLA research produced in the 1980s/1990s and more recent proposals is closely related to the introduction of the Minimalist Program (Chomsky, 1995) and the shift from parameters towards features that this program implied. As Ionin puts it, in this framework, '*parametric differences among languages are restricted to differences in formal features on functional items*', so '*learning a new parameter setting is now viewed as a matter of lexical learning of the corresponding functional items*' (Ionin, 2013, p.506).

As I did at the beginning of the section with the proposals discussed in White (2003), I will now present the more modern hypotheses discussed in Ionin (2013). All the following proposals adopt an impairment approach, considering that the non-target-like morphology showed by L2 learners corresponds to a representational impairment at some level or another.

Proposals suggesting impairment at the syntactic level consider that morphology errors are symptomatic of deeper syntactic problems. The Failed Functional Features Hypothesis previously discussed falls into this category, as does Hawkins' **Representational Deficit Hypothesis** (2003), which is often used as a label for all the proposals following the same line of reasoning. The main claim is that functional features absent from the L1 are impossible to acquire in a second language. These hypotheses were later refined by restricting the unavailability only to uninterpretable features, giving rise to the **Interpretability Hypothesis** proposed by Tsimpli (2003) and adopted by Hawkins and Hattori (2006), whose study focused on the uninterpretable wh-feature in C, which I will further discuss in the next section. The hypothesis was also explored by Tsimpli and Dimitrakopoulou (2007), and Tsimpli and Mastropavlou (2008).

Proposals arguing in favor of impairment at the morphological level include the **Missing Surface Inflection Hypothesis** (Haznedar & Schwartz, 1997, Lardiere 2000, Prévost & White, 2000), which attributes the L2 learners' non-target-like morphology to a mapping problem between syntax and morphology. Their suggestions are based on a much higher omission than misuse rate of inflectional morphology and on presence of syntactic properties related to inflectional functional categories. This hypothesis is fully compatible with Full Access models. McCarthy's **Morphological Underspecification Hypothesis** (2007, 2008) also considers the impairment occurs at the morphology level, but it views errors as a matter of competence rather than performance. Since a task comparison between production and comprehension did not show significant differences in target-like behavior, McCarthy proposed that features are underspecified at lower proficiency levels and are acquired later.

The **Prosodic Transfer Hypothesis** (Goad, White & Steele, 2003, Goad & White, 2006) suggests impairment occurs at the phonology level. Adopting a hierarchical view of prosodic constituents, this hypothesis considers morphosyntactic errors exhibited by L2 learners as a consequence of transfer of L1 prosodic properties, which causes difficulty to produce specific morphemes.

Recently, a proposal following a completely new direction has been put forth by Lardiere (2008, 2009). Under her **Feature Reassembly Hypothesis**, the acquisition of a L2 is viewed as a selection process of the relevant features for lexical items in the target language. Since languages organize their grammars choosing different features from the same universal inventory and combining them differently, L2 learners need to select new features, when they are not selected by the L1 too, and reassemble existing features, when they are present both in L1 and L2 but bundled together into lexical items and functional categories in different ways. This perspective explores features' fine-grained semantic properties and focuses on whether L2 learners are able to acquire them or not. I will tangentially discuss Lardiere's proposal in the next section by reviewing a Ph.D. dissertation on SLA of wh *in situ* advised by her.

3.4. The wh-feature in SLA research

Wh-movement and the crosslinguistic differences related to this syntactic phenomenon have been profusely studied in generative literature, because it provides a very good testing ground for theoretical hypotheses. In particular, this is also true for generative research on second language acquisition. The wh-movement parameter and later the wh-feature and its strength have been the focus of some of the theories that moved, and are still moving, the field forward.

However, languages showing both fronted and *in situ* interrogatives, like BP, have never been central in SLA research, at least, not to my knowledge. This is probably due to two main reasons. First, contrast between fully fronting and fully *in situ* languages provides a much clearer context for investigating the role of the wh-feature and its strength in how native speakers of a fronting language acquire an *in situ* language and vice-versa. And second, as I already discussed in chapter 1, this kind of optionality does not fit well within generative models and, more often than not, it is treated as being only apparent, with one of the possible structures being the result of some other syntactic operation.

In this section, I discuss four of the latest proposals related to the wh-feature in SLA and which constitute the literature most relevant to our topic of study. Two of them adopt a Representational Deficit approach (Hawkins & Chan, 1997; Hawkins & Hattori, 2006), one argues for Full Access (Hettiarachchi & Pires, 2016), and the last one follows the Feature Reassembly Hypothesis (Choi, 2016). Interestingly, contradicting findings and reinterpretation of similar data make possible for wh-movement to be used to argue in favor of completely opposite theories, as we will shortly see.

3.4.1. _Representational Deficit approach_

Hawkins and Chan (1997) examine the acquisition of English restrictive relative clauses (RRC from now on) by French and Cantonese native speakers to explore the availability of the wh-feature in SLA. The authors use the findings of this study to propose their Failed Functional Features Hypothesis. The original paper was published in 1997 and some of the terminology was a bit dated, so I will use here Hawkins's more modern revision of his original work (Hawkins, 2005).

English RRCs like the one in (1) involve movement of the wh-pronoun from the position where it is first merged to the specifier of CP.

(1)   _The patient [$_{CP}$ who [$_{TP}$ I visited ~~who~~]] was very sick._

This analysis is based on subjacency violations involving wh-islands (2.a) or complex NPs (2.b). If movement were not involved, the sentences in (2) should be grammatical.

(2)   a. _*This is the patient who Mary told me when she will visit._
           b. _*This is the patient who Mary heard the news that the doctor had sent to the hospital._

Hawkins and Chan make use of these and two other phenomena, namely ungrammaticality of resumptive pronouns in these constructions (3.a) and doubly filled complementizers in English (3.b), to build a acceptability judgement task.

(3)   a. _*The patient who I visited him was very sick._
           b. _*The patient who that I visited was very sick._

RRCs in Cantonese differ from their English counterparts in several respects[34]. First, they are head-final; second, they do not involve wh-movement, instead, a null topic is base generated in CP, it is coindexed with the RRC head and it binds a pronominal element in the embedded clause; in addition, this pronominal element may be null (_pro_, obligatorily in subject position, optionally in object position) or a resumptive pronoun; finally, the complementizer _ge_ is optional due to phonological conditions.

---

[34] Hawkins and Chan (1997) base their analysis for Cantonese RRCs on the analysis for Mandarin RRCs proposed in Xu and Langendoen (1985) and Xu (1986), claiming their structural properties appear to be identical with the only exception of the complementizer's optionality.

(4)   *[CP Topi [TP Ngo jungyi proi] (ge)] go go   neuijaii hou leng.*

       null topic I   like   pro   C   that CLASS girl   very pretty

       'The girl who I like is very pretty.'

The authors assume that English and Cantonese differ in respect to C's functional feature specifications. While the presence of the [wh] feature in the former forces operator movement to Spec-CP, the lack of this same feature in the latter results in lack of operators and hence of operator movement.

Hawkins and Chan's study compared three experimental Chinese groups divided by proficiency level in English with three French control groups also divided by proficiency level, and a control group of English native speakers. All of them were between 12 and 21 years old.

As mentioned before, the authors conducted a acceptability judgement task (GJT). Out of a total of 101 sentences, the task contained 59 relevant items of which 20 were grammatical RRCs and 39 involved some kind of ungrammaticality (wh-island violations, complex NP violations, misuse of resumptive pronouns and doubly filled Cs). The sentences had to be judged as definitely correct, probably correct, probably incorrect or definitely incorrect. Results in % of accurate judgements are shown in tables I and II.

(5)   Table I: Judgement accuracy (%) – grammatical RRCs, ungrammatical doubly filled C, ungrammatical resumptive pronouns (adapted from Hawkins, 2005)

	Grammatical RRCs	*Doubly filled C	*Resumptive pronouns
Elementary Chinese	56	50	38
Intermediate Chinese	67	68	55
Advanced Chinese	79	83	90
Elementary French	81	91	81
Intermediate French	88	95	90
Advanced French	92	98	96
Native Speakers	96	99	98

Based on the significative differences in accuracy between French and Chinese speakers, and between the three proficiency levels within each of those two groups, Hawkins and Chan claim the GJT is a valid indicator of syntactic competence and it is because the L1s are different that

the results between the two groups are too. Like English and unlike Cantonese, French also has a [wh] feature in C.

Since Table I results show accuracy increasing with proficiency both for Chinese and French speakers (despite differences in absolute value and range), the authors argue that the subjects in both groups are progressively acquiring C's main morphosyntactic properties in English. This may lead to conclude that Chinese and French speakers have adopted wh-movement as part of their interlanguage grammars regarding RRCs.

If this were true, and assuming interlanguage grammars are constrained by UG principles, both groups should be sensitive to Subjacency effects. However, the results in table II indicate that only the French speakers show a proficiency related increase in this sensitivity. Unlike the French group, Chinese speakers seem to significantly lose accuracy at detecting Subjacency violations with increasing proficiency level.

(6)   Table II: Judgement accuracy (%) – wh-islands, complex NP (adapted from Hawkins, 2005)

	*Wh-islands	*Complex NP
Elementary Chinese	63	71
Intermediate Chinese	54	61
Advanced Chinese	41	38
Elementary French	59	72
Intermediate French	66	79
Advanced French	85	90
Native Speakers	98	85

The authors provide the following explanation for these results. The rejection of sentences involving Subjacency violations by the less proficient Chinese subgroup is due to lack of resumptive pronouns in these sentences and not because of the violations related to wh-movement. In support of this claim, these speakers only reject English RRCs with resumptive pronouns 38% of the time, and grammatical RRCs are accepted at chance level (56%). Recall that there is no wh-movement in Cantonese RRCs but resumptive pronouns are allowed. Having this in mind, Hawkins and Chan suggest L1 properties have been transferred to these speakers' interlanguage grammars.

The advanced Chinese group shows the opposite behavior. They are able to detect the ungrammaticality of English RRCs with resumptive pronouns 90% of the time and they accept grammatical RRCs with 79% accuracy, but their judgement accuracy for violations related to wh-movement is around 40%. Hence, the authors claim Chinese speakers have not acquired wh-movement, their mental representations for RRCs involve a base generated wh-phrase that binds an obligatory null resumptive pronoun in the embedded clause:

(7)   *The patient [$_{CP}$ who$_i$ [$_{TP}$ I visited pro$_i$]] was very sick.*

Finally, Hawkins and Chan conclude the wh-feature must be subject to a critical period and so must happen too to any functional feature determining parametric differences. This constitutes the Failed Functional Features Hypothesis, which, as mentioned in the previous section, was later updated and became the Interpretability Hypothesis. Hawkins & Hattori (2006) argue in favor of this refined version of the hypothesis by studying the acquisition of superiority and subjacency effects in English multiple wh-questions by Japanese native speakers.

The general idea is more or less the same as in Hawkins and Chan (1997). Since Japanese interrogative C lacks the uninterpretable strong wh-feature present in English, if we assume Japanese learners of English have been able to acquire this feature then this implies their interlanguage grammars should respect subtle wh-movement properties such as superiority and subjacency effects.

Superiority effects appear in multiple wh-questions where the wh-phrase deepest in the structure cannot cross a higher situated wh-phrase. See the examples in (8):

(8)   a. *What did who buy ~~what~~?*
            b. *When did who leave ~~when~~?*
            c. *What did John buy ~~what~~ where?*

The syntactic structure of Japanese interrogatives is quite different than the English one. In Japanese, *v*P, TP and CP are head-final, interrogative C is realized by the particle *ka* and wh-phrases may remain *in situ* (8.a) or undergo scrambling (8.b).

(9)   a. *[$_{CP}$ [$_{TP}$ John-ga [$_{vP}$ kinou nani-o kaimashi] -ta] ka]?*
               John-NOM yesterday what-ACC buy -PAST Q

            b. *Nani-o John-ga kinou ~~nani-o~~ kaimashita ka?*
            'What did John buy yesterday?'

Japanese scrambling, however, is not the result of wh-movement because there is no superiority/subjacency effects in this language. Note the contrast with English in (10). Japanese scrambling is considered to involve movement to a Focus position, which is not subject to superiority effects because the Focus projection does not establish an operator-variable construction.

(10)   a. *What did John say [who ate ~~what~~]?

           b. *Nani-o John-wa   [dare-ga ~~nani-o~~   tabeta ka] itta no?*

           what-ACC John-TOP who-NOM what-ACC   ate   C say Q

Based on this difference, Hawkins and Hattori tested if Japanese speakers with a high level of proficiency in English were capable of detecting the ungrammaticality of multiple questions in biclausal sentences involving superiority or subjacency violations.

Multiple questions in biclausal sentences may have two possible interpretations depending on whether the matrix wh-word has scope over the matrix or the embedded clause:

(11)   *When did Sophie's brother warn [Sophie would phone who]?*

           a. Matrix scope – Reading A: *When did he make the warning?*

           b. Embedded scope – Reading B: *When would Sophie phone?*

However, this is not always the case. Reading B is no longer available when there is a superiority violation, as in (12), a subjacency violation, as in (13), or both at the same time, as in (14):

(12)   *Who did Sophie's brother warn [Sophie would phone when]?*

           a. Matrix scope – Reading A: *Who did he warn?*

           b. *Embedded scope – Reading B: *Who would Sophie phone?*

(13)   *When did Rupert discover who Nora had met?*

           a. Matrix scope – Reading A: *When did he discover it?*

           b. *Embedded scope – Reading B: *When had Nora met somebody?*

(14)   *Who did the weather office warn when the hurricane might strike?*

           a. Matrix scope – Reading A: *Who did the weather office warn?*

           b. *Embedded scope – Reading B: *Who might the hurricane strike?*

To test Japanese speakers' sensitivity to this subtle property of English wh-movement, Hawkins and Hattori used a truth value judgement task consisting of a short story, plus a question, plus three possible answers for each item. The 14 experimental items included questions like (11-

14) and also questions like (15), in which the only possible interpretation is the one corresponding to the embedded scope (with no syntactic violation). All three answers for the experimental items were pragmatically possible ('true') given the context provided by the story.

(15)   *Who did the headteacher suspect had taken what?*

          Only Embedded scope is possible – Reading: *Who had taken something?*

The authors included 8 items involving simple questions in biclausal sentences, some had unambiguous scope over the embedded clause, some had ambiguous scope over the matrix or the embedded clause. These items were used as a syntax test to ensure participants were able to interpret long-distance wh-movement. Those who answered three or more items incorrectly were not considered for analysis.

The results for the experimental items are shown in Table III below.

(16)   Table III: Mean choice of answers corresponding to the matrix wh-word (adapted from Hawkins & Hattori, 2006)

Question Type	Embedded scope		Matrix scope	
	Japanese speakers	English controls	Japanese speakers	English controls
ES – NV	0.96	0.97		
MS & ES – NV	0.78	0.75	0.92	0.91
MS & *ES – SprV	0.75 ⟷	0.33[35]	0.88	0.85
MS & *ES – SbjV	0.58 ⟷	0.21	0.93	1.00
MS & *ES – SSV	0.58 ⟷	0.00	0.95	0.91

Key: ES – Embedded scope / MS – Matrix scope / NV – No syntactic violation

SprV – Superiority violation / SbjV – Subjacency violation / SSV – Superiority and subjacency violation

The first two rows (corresponding to sentences like (15) and (11)) show that Japanese speakers and English controls had a very similar response to items that did not involve any syntactic violation. According to the authors, this shows the experimental group's ability to interpret long distance wh-movement.

As for the rest of the results (corresponding to sentences like (12-14)), they show Japanese speakers' acceptance of all three kinds of violations is significantly higher than that of the control group. In addition, their responses to the ungrammatical embedded readings involving

---

[35] Throughout the chapter, orange arrows represent a statistically significant difference.

a superiority violation were very similar to their responses to the grammatical embedded readings in ambiguous sentences like (11).

From these results, Hawkins and Hattori conclude that Japanese speakers with a high proficiency in English are not sensitive to superiority or subjacency, subtle constraints on wh-movement. The authors claim this implies English C's uninterpretable strong wh-feature is missing from their interlanguages as a consequence of the lack of this feature in their L1 and the effect of a critical period, which constitutes evidence for the Interpretability Hypothesis.

To explain how these Japanese speakers compensate the lack of the uninterpretable wh-feature, they suggest their mental representation for fronted interrogatives involves a Focus projection in the left periphery of interrogative sentences which has a strong uninterpretable feature. In addition to this, since wh-words constitute non-presupposed information, an interpretable focus feature is associated to all of them. Wh-words then move to the sentence's left periphery to value the uninterpretable focus feature in FocP. With this analysis, in multiple wh-questions, it does not matter which wh-word is fronted and which one stays in situ because the Focus projection does not establish an operator-variable construction, it only identifies non-presupposed information.

3.4.2. *Full Access approach*

Hettiarachchi and Pires (2016) recently replicated Hawkins and Hattori's study (2006) with native speakers of Sinhala that were acquiring English as a L2. The results they obtained are not in line with Hawkins and Hattori's, so Hettiarachchi and Pires claim they can be interpreted as additional evidence for Full Access to UG in adult SLA.

The language chosen by the authors, Sinhala, shares with Japanese the relevant structural properties regarding wh-questions:

i)   It is a wh *in situ* language, it has no overt wh-movement.

(17)   *siri mokak də kiyeww-e?*

Siri what Q read

'What did Siri read?'

ii)   Sinhala presents scrambling, but it is considered to be triggered by a topic or focus feature.

(18)   *mokak də siri kiyeww-e?*

            what Q Siri read

            'What is it that Siri read?'

iii)   Sinhala's scrambling is not subject to superiority or subjacency violations.

(19)   *siri [mokak də kau də kiwwe kiyəla] kalpəna-kəruwa.*

            Siri what   Q who Q said that   wonder

            'Siri wondered what who said.'

iv)   Wh-words are allowed within islands, a fact that further supports the claim that there is no wh-movement in this language.

(20)   *chitra Ranjith monəwa gatta kiənə katəkataawə də æhuwe?*

            Chitra Ranjith what   bought that rumor   Q heard

            'What did Chitra hear the rumor that Ranjith bought?'

The experiment conducted by Hettiarachchi and Pires was a slightly modified version of the Truth Value Judgement task in Hawkins and Hattori (2006). The main differences consisted in the elimination of one of the question types (the one in which only the embedded scope was possible but there was no syntactic violation, (15)) and one of the possible answers (the one in which the embedded wh-phrase was interpreted as having scope over the matrix clause).

Each test item consisted of a story providing some context, a multiple wh-question and two possible answers, both pragmatically acceptable. In some test items one of the answers was ungrammatical due to superiority violations, subjacency violations or both. Participants were asked to choose one or both answers according to their acceptability as an answer for the question given the context provided in the story.

(21)   **Story**

            James is making plans to go hike the Great Wall of China during the summer. Last Tuesday, James promised to call Lois the following day with the details of the trip, so that Lois can join him too.

            Question

            Who did James promise he would call when?

            Answers

            a) James promised that on Wednesday he would call Lois.

            b) James promised Lois that he would call on Wednesday.

The results of the experiment are summarized in Table IV.

(22)   Table IV: Mean choice of answers corresponding to the matrix wh-word (adapted from Hettiararchchi & Pires, 2016)

Question Type	Embedded scope			Matrix scope		
	Intermediate learners	Advanced learners	English controls	Intermediate learners	Advanced learners	English controls
MS & ES – NV	0.47	0.66	0.85	0.73	0.64	0.73
MS & *ES – SprV	0.56	0.71	0.54	0.8	0.66	0.92
MS & *ES – SbjV	0.14	0.2	0.31	0.92	0.91	0.91
MS & *ES – SSV	0.41	0.29	0.5	0.78	0.88	0.9

Key: ES – Embedded scope / MS – Matrix scope / NV – No syntactic violation

SprV – Superiority violation / SbjV – Subjacency violation / SSV – Superiority and subjacency violation

The results are not as conclusive as those obtained by Hawkins and Hattori. However, based on a statistical comparison between groups, readings and conditions, Hettiarachchi and Pires draw the following conclusions.

Despite the slight preference for the embedded reading shown by native speakers and a more significant preference for the matrix reading shown by the intermediate L2 group in the no violation condition, all three groups were sensitive to the ambiguity and were able to assign both readings when no syntactic violation is involved.

Regarding subjacency, all three groups showed a clear preference for the grammatical matrix reading in the subjacency violation condition. In addition, each group's ungrammatical embedded reading in this condition significantly differed from their own assignment of embedded readings in the no violation condition. Hence, both L2 groups show strong sensitivity to subjacency violations, since their behavior in this respect is not significantly different than that of native speakers. The authors interpret this as evidence of the acquisition of the uninterpretable wh-feature.

Regarding superiority, though, the results were not exactly as expected for the L2 groups. There was no significant difference between both groups' ungrammatical embedded reading in this condition and their own assignment of embedded readings in the no violation condition. In the condition involving both violations, only the advanced learners showed a significant difference between the embedded reading in this condition and in the no violation condition. Therefore, these results do not indicate sensitivity to superiority effects in L2 learners.

To account for this unexpected behavior, the authors point out that even native controls showed weaker sensitivity to superiority violations and they argue this could be due to the way the specific items were built. All items for this condition involved only argument over adjunct extractions, which are acceptable to many native speakers. Compare (23.a) and (23.b).

(23)　　　a. ?*What did you read where?*
　　　　　b. **What does Siri believe who said?*

This difference in grammaticality constitutes then an even subtler property of English wh-movement. Hettiarachchi and Pires claim that if learners were able to perceive this difference then this would constitute further evidence for their sensitivity to wh-movement violations and hence of their having acquired the relevant uninterpretable feature, contrary to what the Interpretability Hypothesis predicts. The authors state they conducted a second experiment including an acceptability judgement task to test learners' sensitivity to the subtle grammatical contrast in (23), and they claim their results confirm learners perceive this contrast. However, they do not report this results in the article.

3.4.3. *Feature Reassembly approach*

As mentioned in section 3.3, recent work by Lardiere (2008, 2009) has explored a new approach to SLA attacking the problem in terms of the selection and reorganization of relevant features associated with lexical items and functional categories in language-specific ways. This is a very interesting idea because, unlike parameter setting approaches, it can account for persistent morphological variability in interlanguages and divergence from the target grammar, without resorting to performance errors or assuming native-like grammars to be unachievable for L2 learners.

Choi's Ph.D. dissertation (2016) examines the interpretation of wh *in situ* elements in L2 Korean by English native speakers and argues in favor of a Feature Reassembly approach to SLA. The dissertation provides evidence of this theory's explanatory force by showing it constitutes the best way to account for the results obtained in an experimental study investigating the acquisition of Korean wh *in situ*.

Choi chose English and Korean as the L1/L2 pair for two reasons. First, acquisition of overt wh movement by native speakers of *in situ* languages has been profusely studied, but the opposite direction of the acquisition process had not. And second, the differences between the two

languages regarding wh expressions cannot be reduced to the wh-parameter. In addition, both languages select the same relevant features but assemble them in a very different manner, making this L1/L2 pair a perfect candidate for testing the Feature Reassembly hypothesis.

According to the Binding Approach assumed in Choi's work, wh-elements in *in situ* languages are variables that must be bound by a quantifier licensor, whereas in wh-movement languages, the operator and the variable are lexicalized in the wh-words, forcing overt movement to the scope position.

This analysis accounts for the multiple interpretations of Korean variables, which can have a universal, indefinite or wh-word reading depending on the context. Overt lexical items function as quantifiers (universal, existential and wh quantifiers) and license the variables' interpretation. The indefinite and wh-word readings are illustrated in (24). Note that matrix questions allow both the indefinite and wh-word reading of the variable depending on the intonation. A rising intonation determines the indefinite interpretation for the variable, thus making the sentence a yes/no question. A falling intonation determines the wh-word reading which guarantees the sentence is interpreted as a wh question.

(24)   a. *Mary-ka **mwues**-(i)-nka-(lul) sa-ss-**ta**.*
            Mary-NOM THING-∃-(ACC) buy-PAST-DECL
            'Mary bought something.'
         b. *Mary-ka **mwues**-(ul) sa-ss-**ni**?*
            Mary-NOM THING-ACC buy-PAST-Q
            (i)   Rising intonation: 'Did Mary buy something?'
            (ii)   Falling intonation: 'What did Mary buy?'

The sentential particles *-ta*, and *-ni* respectively type the clauses as declarative and interrogative. The particle *ni* can be considered a Q-morpheme with the features [+Q, ±wh]. A rising intonation sets the feature to [-wh] and the variable is bound by an implicit existential operator. A falling intonation sets the feature to [+wh] and the variable is bound by a covert Q-operator.

In embedded sentences, declaratives are again introduced by the particle *-ta* [-Q] attached to the embedded clause, whereas interrogatives are introduced by the particle *-nunci* [+Q]. In contrast to the underspecified particle *-ni* in matrix interrogatives, *-nunci* disallows the indeterminate reading for the variable element, as can be seen in (25). To obtain the indeterminate reading, an A-not-A construction (*nunci aninci*) needs to be used.

(25)   a. *Mary-nun [Tom-i **mwues**-ul sa-ss-**ta**-ko] an-ta.*

Mary-TOP Tom-NOM THING-(ACC) buy-PAST-DECL-C know-DECL

'Mary knows that Tom bought something.'

b. *Mary-nun Tom-i **mwues**-ul sa-ss-**nunci** an-ta?*

Mary-TOP Tom-NOM THING-ACC buy-PAST-Q know-DECL

(i)   Mary knows what Tom bought.'

(ii)   *Mary knows if Tom bought something.'

c. *Mary-nun Tom-i **mwues** -ul sa-ss-**nunci aninci** an-ta.*

Mary-TOP Tom-NOM THING-ACC buy-PAST-Q not-Q know-DECL

'Mary knows whether or not Tom bought something.'

In terms of features, the differences between English and Korean are summarized in Table V.

(26)   Table V: Feature organization and its (c)overt realizations (Choi, 2016; p. 45)

	Matrix Clause		Embedded Clause	
	English	Korean	English	Korean
Declarative [−Q, −wh]	Ø	*ta*	*that/ Ø*	*ta*
Yes/No Interrogative [+Q, −wh]	Ø	*ni (+rising)*	*if/whether*	*nunci(aninci)*
Wh-interrogative [+Q, +wh]	Ø	*ni (+falling)*	Ø	*nunci*

In addition to that, C selects an (uninterpretable) EPP feature in English, whereas in Korean it does not. And within the DP domain, the interpretable [Q-operator] and [Variable] features are conflated into wh-expressions in English, whereas in Korean, wh-lexical items are variables lacking quantificational force, so they need to be bound by an operator.

The experimental study conducted by Choi consisted of two translations tasks (Korean to English), one from listening and one from reading, and a truth value judgment that were administered to 47 adult learners of Korean at high intermediate (n=24) and advanced level (n=23) of proficiency.

The listening-translation task tested the learners' ability to disambiguate the reading of matrix questions, like (24.b), by making use of phonological cues. There were 12 distracters and 12 experimental items, half for each condition (falling vs. rising intonation). The number of syllables was balanced, the sentences were pre-recorded and played twice with a pause between items for writing the translated sentence.

The reading-translation task tested the learners' interpretation of variable expressions in embedded sentences like (25.a and b), which is morphologically determined by the sentential particle attached to the embedded clause (*ta* [-Q] vs. *nunci* [+Q]). Again, there were 12 distracters and 12 experimental items, half for each condition. All the experimental items involved variable expressions embedded by the matrix verb *kiekhata* 'remember' (which accepts either a declarative or an interrogative clause).

The truth value judgement task also tested learners' ability to use morphological cues to correctly interpret variable expressions in embedded sentences. This task's purpose was to corroborate the results of the reading-translation task. There were 12 test items. Each one consisted of a short story in English and four related sentences in Korean, two distracters and two biclausal sentences with indefinite and question readings (one each) for an embedded variable. Stories belonged to two groups (with six items each), one provided non-specific information that only allowed for the indefinite reading, the other provided specific information making both readings acceptable (although the expected response was the question reading). An example of each context is provided in (27) and (28).

(27)   Sample of a non-specific information context test item

Story

John and Mary are close co-workers working in the financial division. One day John saw a large and beautiful flower basket delivered to Mary. Mary was not there at the time. John was so curious about it, and then opened a card attached to the flower basket. A love message was written on the card. But there was no name of the sender on the card.

Related sentences

(i)   *John-un [__nwu__(kwu)-ka Mary-lul cohahan-__ta__-ko] an-ta.*
        John-TOP [PERSON-NOM Mary-ACC like-DECL-C] know- DECL
        'John knows that **somebody** likes Mary.'
        True ☑ False ☐ Don't Know ☐
(ii)   *John-un [__nwu__(kwu)-ka Mary-lul cohaha-__nunci__] an-ta.*
        John-TOP [PERSON-NOM Mary-ACC like-Q] know- DECL
        'John knows **who** likes Mary.'
        True ☐ False ☑ Don't Know ☐

Plus two distracters.

(28)   Sample of a specific information context test item

Story

Chelswu likes playing with his brother. He was looking for his brother to play with after he came back from school. But he could not find him. He asked his mother about his brother. His mother told him that his brother was at Yenghi's home. So, he went to Yenghi's home to find him.

Related sentences

(i)   *Chelswu-nun [tongsayng-i **eti**-ey ka-ss-**ta**-ko] an-ta.*
        Chelswu-TOP [brother-NOM PLACE-LOC go-PAST-DECL-C] know-DECL
        'Chelswu knows that his brother went **somewhere**.'
        True ☑ False ☐ Don't Know ☐

(ii)   *Chelswu-nun [tongsayng-i **eti**-ey ka-ss-**nunci**] an-ta.*
        Chelswu-TOP [brother-NOM PLACE-LOC go-PAST-Q] know- DECL
        'Chelswu knows **where** his brother went.'
        True ☑ False ☐ Don't Know ☐

Plus two distracters.

The group results of the two translating tasks are summarized in Table VI.

(29)   Table VI: Mean accuracy (Standard Deviation) of variable interpretations (adapted from Choi, 2016)

	Phonological condition		Morphological condition	
	Rising intonation (Indefinite reading)	Falling intonation (Question reading)	Declarative *-ta* (Indefinite reading)	Interrogative *-nunci* (Question reading)
High Intermediate	7% (.23) ⟷	97% (.09)	14% (.21) ⟷	85% (.23)
Advanced	68% (.28) ⟷	94% (.12)	60% (.36) ⟷	84% (.22)
Controls	100% (.00)	100% (.00)	100% (.00)	100% (.00)

Both intermediate and advanced learners show a significant asymmetry in their ability to correctly interpret variable expressions with indefinite and question readings. Their performance on question readings was very accurate in the prosodic licensing environment (almost native-like) and above chance level in the morphological licensing environment. However, the indefinite reading was more problematic for both groups. The intermediate learners' performance was very low in the two licensing environments, although the presence of sentential particles seemed to aid the variables' interpretation. The advanced group

performed better than the lower proficiency learners indicating an increase in target-like responses with increasing proficiency. But they still often failed to assign the indefinite reading to variable expressions, both in matrix and embedded sentences, with no clear sign of environment effects.

Individual results from both translating tasks confirm the asymmetry in performance between readings and between groups, and show some of the advanced learners have acquired target-like knowledge of the interpretive constraints on variable expressions.

For the data analysis of the truth value judgement task, each 'True' response was assigned a +1 score and each 'False' response a -1 score. This means negative mean judgments indicate a tendency to reject the test sentences whereas positive mean judgements indicate a tendency to accept them. Group results are summarized in Table VII.

(30) Table VII: Mean Judgements in the interpretation of variable expressions (adapted from Choi, 2016)

	Non-specific Information Context		Specific Information Context	
	Declarative *-ta* (Indefinite reading)	Interrogative *-nunci* (Question reading)	Declarative *-ta* (Indefinite reading)	Interrogative *-nunci* (Question reading)
High Intermediate	-0.75	-0.61	+0.63 ←→	+0.86
Advanced	+0.65 ←→	-0.76	+0.75	+0.74
Controls	+0.84 ←→	-0.98	+0.73	+0.91

The results from the truth value judgement task show higher divergence between the two groups, advanced learners displaying native-like behavior and lower proficiency learners making no significative distinction in the interpretation of variables licensed by the declarative and the interrogative particle.

Choi argues against a possible interpretation of the data in terms of a Full Access approach. Although learners seem to have reset the uninterpretable wh-feature in C from strong (L1) to weak (l2) (the *in situ* placement of wh lexical items did not pose problems in any of the test tasks), they still showed learning problems regarding the interpretation of variable expressions. The author conducted a corpus-based research on three types of variable expressions (*nwukwu* 'PERSON', *mwues* 'THING' and *eti* 'PLACE') and found out at least 30% of occurrences involving indefinite reading for each variable expression. Therefore, he concludes there is no poverty-of-the-stimulus problem for the acquisition of indefinite readings for variable

expressions, which under a Full Access approach would be the only possible source of non-target L2 grammars in this case.

In addition, English and Korean share the same interpretable features for the generation of wh-words, namely [+Q], [+wh] and [VAR]. However, learners displayed considerable difficulty acquiring wh-word interpretation in Korean and its licensing conditions. Choi argues this cannot be explained either by the Interpretability Hypothesis (a Representational Deficit approach), since the only candidate source for non-target L2 grammars under this proposal are uninterpretable features, which are not directly involved in the learner/native grammar divergences found in the study.

According to Choi, the data can be better accounted for by adopting a Feature Reassembly approach. Under this perspective, the L1 and L2 are compared on the basis of their differently assembled lexical items, and persistent non-target like behavior is addressed in terms of the organization of relevant features.

In the particular case under study, the author claims learners initially treat Korean wh-lexical items as having the same feature configuration of English wh-words, [OP + VAR], and until these features have not been disentangled, learners seem to ignore licensor environments. However, the disentanglement does not necessarily imply an abrupt change. According to the Feature Reassembly Hypothesis, variability in learner data is due to the co-existence in the interlanguage of L1-type and L2 type lexical items. This explains inconsistent target-like behavior in a principled way. For Choi's study, this means the wh constructions in interlanguage grammars can correspond to two competing configurations, [OP + VAR] and [OP … VAR], any of which can be inserted into the numeration respectively yielding a non-target like construction or a target-like one. The choice between co-existing items depends on the licensing environments, and this requires a gradual learning process.

Based on the reasoning above, Choi argues his test results constitute evidence in favor of the Feature Reassembly Hypothesis.

3.5. Topic prominence in SLA research

Topic-subject constructions have not yet been studied from an acquisition perspective, so the most relevant literature, given the circumstances, is related to topic prominence in second language acquisition. Below, I briefly discuss an older proposal arguing in favor of an initial

topic-comment stage in all SLA and a more recent one that brings evidence against this idea and argues in favor of typological transfer from the native language.

3.5.1.   *Early topic-comment stage*

The idea of an early topic comment stage in SLA comes from first language acquisition research and it was explored mainly in the 1980s. The comparative study in Fuller and Gundel (1987) is representative of this proposal.

Fuller and Gundel (1987) compared three sets of oral narratives: English narratives produced by native speakers of three highly topic prominent languages (Chinese, Japanese and Korean) and three less topic prominent languages (Arabic, Farsi and Spanish); native language narratives produced by the same speakers; and English narratives produced by English native speakers. The narratives were elicited as a response to a short film with no spoken language in it.

The authors chose the following six syntactic properties that had previously been proposed in the literature as a way of measuring how topic prominent or subject prominent a given language is.

(i) Coding of topic: All languages mark topics in some way or another, but only in topic prominent languages the marking is invariant. Topics in these languages stand out thanks to some kind of surface coding, usually a morphological marker, a specific syntactic position or both.

(ii) 'Double subject' constructions: Topics in topic prominent languages do not need to be selected by the verb nor be coindexed with a noun phrase inside the 'comment' part of the sentence. This allows constructions in which two initial DPs look like a double subject, as I mentioned in chapter 2.

(iii) Zero-NP anaphora[36]: In topic prominent languages, zero-NP anaphora's only constraint is that the missing NP is co-referent with the topic, whereas in subject prominent languages, zero-NP anaphora is syntactically constrained.

---

[36] Zero-NP anaphora refers to a gap in a clause whose interpretation depends on an antecedent that can come from the same sentence or from previous discourse. Null subjects and null objects are examples of zero-NP anaphora.

(iv) 'Subject-creating' constructions: Subject prominent languages have grammatical strategies to turn the topic into the subject, like passivization or subject raising. Topic prominent languages do not need these strategies.

(v) Expletive subjects: Topic prominent languages do not have expletive subjects.

(vi) Subject-verb agreement: Subject-verb agreement is characteristic only of non-topic prominent languages.

Fuller and Gundel used the properties listed above as criteria for analyzing and identifying relative topic prominence between the languages involved in their study, looking for evidence of each property in the L2 oral narratives produced by the participants.

On one hand, the authors found topics were consistently placed in sentence-initial position and left-dislocation was used as a coding strategy, but this could be due both to transfer from L1 or to an early topic-comment stage in the interlanguage. They found several examples of co-reference controlled by the topic in both groups of learners, with highly topic prominent L1s and with more subject prominent L1s. There were no passives or any other 'subject creating' constructions in the learners' English narratives, although this kind of construction was present both in the data produced by English native speakers and in the data produced in other native languages. All of this indicates the interlanguage is more topic prominent than both native English and the other non-topic-prominent languages.

On the other hand, they did not find any 'double subject' constructions, although these constructions did not appear either in native narratives of topic prominent languages, so the authors argue this cannot really be considered as counterevidence to topic prominence. They found some expletive subjects in English narratives produced by speakers of the two types of languages, and in this case, it could not be due to transfer from the L1 because none of the languages in the study admits expletive subjects. Finally, the authors found some instances of subject-verb agreement, but they claimed it was very inconsistent, so again they argue this cannot be considered as counter evidence to topic-prominence.

No difference was found in their English spoken production between native speakers of topic prominent languages and non-topic prominent languages. Fuller and Gundel conclude that their results provide empirical support for the hypothesis of an early universal topic-comment stage in SLA. However, their methodology did not meet the modern scientific standards and I believe their results could be re-interpreted to support the opposite idea.

3.5.2. *Typological transfer from L1*

Jung (2004) points out that most SLA research on topic prominence had focused on the acquisition of a subject prominent L2 by native speakers of a topic prominent L1, making hard to discern whether instances of topic-comment structures in L2 data were due to L1 transfer or to some other reason like the universal topic-comment stage postulated by Fuller and Gundel (1987). To fully understand the role of topic/subject prominence typology in SLA, Jung claims acquisition must be studied in the other direction, from a subject prominent L1 to a topic prominent L2.

That is what Jung does in her 2004 paper. The participants were English native speakers acquiring Korean as a L2 at three different levels of proficiency. Their results were compared with a group of Korean native speakers. The study consisted in an elicited written production task. After watching a silent short movie, the participants were asked to retell the film in writing from a particular scene around the middle to the end. The criteria used by Jung to measure topic prominence in the interlanguage was the presence of zero anaphora, topic markers, and 'double-subject' or double-nominative constructions.

The results show evidence of L1 transfer, which decreased with increasing proficiency level. The frequency of clauses making use of zero anaphora, topic markers and double-nominative constructions started out quite low and gradually increased with proficiency almost reaching native-like levels for zero anaphora.

(31)   Table VIII: Frequency of clauses with zero anaphora, topic markers and double-nominative constructions (adapted from Jung, 2004)

	Zero anaphora	Topic marker	Double nominative
Level 1	10.1%	4.2%	0%
Level 2	34.0%	12.9%	2.1%
Level 3	43.5%	19.6%	2.6%
Native	49.8%	33.8%	6.3%

The author studies these results in more detail, for example, noticing an asymmetry between subject and topic in all three groups, learners showed preference for subject rather than object

dropping. I will not discuss these details here, but all of Jung's finer grained observations are also consistent with L1 transfer.

3.6. Considerations for the present study

Since wh *in situ* in a language that also shows fronting, and topic-subject constructions both represent some kind of middle ground in the *in situ* vs. fronting and topic vs. subject prominence typology scales, it may seem *a priori* unclear how the acquisition of these structures by native speakers of a fronting, subject prominent language is going to fit within the current existing SLA theories and hypotheses. But, bear with me, the picture will become clearer once the results of this experimental study have been analyzed, and their implications for SLA will be discussed in detail in Chapter 7.

PART II

Experimental study

Chapter 4
Methodology

4.1. Introduction

In this chapter, I advance a hypothesis regarding the relation between BP *in situ* interrogatives and topic-subject constructions during the acquisition process of BP as L2 by Spanish native speakers and pose some research questions based on this hypothesis, this corresponds to section 4.2.

To test the hypothesis, an experimental study was designed to be conducted among an experimental group of Spanish native speakers acquiring BP with different levels of proficiency, and among a control group of BP native speakers. The participants characteristics are discussed in section 4.3.

The experimental study combined off-line with on-line measurements. Its backbone consisted of an acceptability judgement task of sentences containing the BP syntactic structures under study, during which, reaction times were recorded. This was followed by a correction task or follow-up in which each participant had to modify the sentences he/she had previously rejected, in order to make them fully acceptable. The participants also needed to complete a language background questionnaire and a BP proficiency test.

I carefully explain the ins and outs of the design process in section 4.4, where the acceptability judgement task and the follow-up task are described in detail, and where the use of the language background questionnaire and proficiency test in BP is accounted for.

4.2. Putting the pieces together: hypothesis and research questions

As I mentioned in the introduction, this research was inspired by my personal experience with BP as a second language. Very early on in the acquisition process, I started using the *in situ* strategy to form interrogative sentences in Portuguese in a rather productive way. So much so,

that this structure begun to cause attrition to my native language, Spanish, which (as previously discussed) does not allow *in situ* questions as freely as BP.

However, this never happened with topic-subject constructions. I lived in Brazil for almost seven years and still found these constructions very strange. Sometimes, I even struggled to understand them, finding myself pragmatically searching for a meaning I could assign to one of these sentences and failing to grasp the right one.

The difference was so bold that it made me think a syntactic reason must be behind it, and my scientific curiosity was awakened by this. So, the hypothesis I want to put forward is that this contrast in the acquisition of the two structures was not idiosyncratic but something that can be extended to all Spanish native speakers acquiring BP. In other words:

(1)   *Hypothesis: There is a significant difference in the acquisition of BP wh* in situ *and topic-subject constructions by Spanish native speakers. More specifically, the former is acquired earlier (and possibly even causes attrition to the speaker's L1) whereas the latter is acquired much later and may never be incorporated into the L2 grammar in its steady state.*

If the hypothesis proves to be true and such difference really exists, this opens the door to some interesting questions: Why is there a distinct contrast between the acquisition of the two structures? What is it due to or how can it be explained? Is there a syntactic reason behind it or is it just a matter of frequency of these constructions in the input?

My first intuition about this problem was to consider that the licensing of BP wh *in situ* requires a derivation with less syntactic operations than Spanish fronted interrogatives, whereas the derivation of topic-subject constructions involves more operations than that of its Spanish counterpart.

If we assume a non-movement analysis of BP wh *in situ*, then the interrogative constituent remains in the position where it is first merged in this language. This means that the final step of the derivation of *in situ* interrogatives in BP (2) does not entail internal merge of the wh-element and coincides with one of the steps of the derivation in the case of fronted interrogatives both in BP (3.i) and in Spanish (4.i) (or other fronting Romance language). So, to produce an *in situ* question in BP, a Spanish speaker only needs to stop the derivation some steps before where he/she would stop in Spanish (especially if we consider subject-verb inversion, which is necessary in Spanish).

(2)   *O Pedro comprou **o que**?*   (BP *in situ* interrogative)

         'Pedro bought what?'

(3)   i. O Pedro comprou **o que**

         ii. ***O que**$_i$ o Pedro comprou **t**$_i$?*   (BP fronted interrogative)

            'What did Pedro buy?'

(4)   i. Pedro compró **qué**

         ii. **Qué**$_i$ Pedro compró

         iii. *¿**Qué**$_i$ compró$_j$ Pedro t$_j$ **t**$_i$?*   (Spanish fronted interrogative)

            'What did Pedro buy?'

This could then be a factor that may contribute to facilitate the acquisition of *in situ* interrogatives and explain why they are so quickly incorporated into the L2 grammar.

In a language like Spanish, for which fronting is obligatory and *in situ* interrogatives are only apparent (according to Etxepare and Uribe-Etxebarria (2005)), the uninterpretable wh-feature in C is strong, assuming Chomsky (1995). So, speaking in terms of features and feature strength, for a Spanish native speaker, L2 acquisition of a language with true *in situ* interrogatives, such as BP (if we assume Grolla (2005)), would imply either a modification in the strength setting of this formal feature or a reanalysis of the syntactic structure that accommodates it to the L1 grammar.

The first option is in line with Full Access approaches (like Hettiarachchi & Pires, 2016) but can also be compatible with a Feature Reassembly approach (similar to that of Choi, 2016) if we extend our study to other features that may be involved or interconnected with the wh-feature (at the top of my head, verb inversion comes to mind, but other more subtle properties may also be relevant).

The second option, on the other hand, falls under a Representational Deficit approach (such as Hawkins and Chan's Failed Functional Features Hypothesis, 1997; or the Interpretability Hypothesis adopted in Hawkins & Hattori, 2006). From this perspective, the wh-feature is subject to a critical period and its strength cannot be modified. This constrains us to assume BP learners' mental representations of *in situ* interrogatives do not involve a true *in situ* structure, but some other analysis consistent with the L1 grammar.

What can be said then about the L2 acquisition of BP wh *in situ* if we adopt a Representational Deficit approach? Or, tackling the problem from the opposite direction, is it possible to look at

the results of this study as evidence of the resetting of feature strength, providing support to Full Access or Feature Reassembly Hypotheses?

But let me go back to the initial idea, which was that topic-subject constructions require a derivation with more syntactic operations than the semantically equivalent sentences in Spanish, while *in situ* interrogatives require less.

To start off on the right foot, I want to clarify again that, although 'subject-topic constructions' has become an umbrella term, the specific constructions that interest me here are those in which a locative or a genitive constituent in pre-verbal position triggers verbal agreement, such as (5) and (6) below.

(5)   *Essa casa bate bastante sol*.
        that house beat.3SG plenty sun
        'That house gets a lot of sunlight.'

(6)   *O meu carro furou o pneu*.
        the my car punctured.3SG the tire
        'My car got a flat tire.'

According to authors like Galves (1998), Munhoz & Naves (2012), Avelar & Galves (2011, 2016), and Nunes (2017), in the syntactic derivation of topic-subject constructions, a constituent that is not the logical subject of the sentence moves to a 'subject'-like position triggering verbal agreement (recall from Chapter 2 that the exact nature of this position varies with the proposals).

This movement is not allowed in Spanish, or in any other Romance languages. As previously discussed, Fernandez-Soriano (1999) argues there are two types of verbs in Spanish that license a locative constituent in subject position (Spec-TP), unaccusative stative verbs like *faltar* 'to miss/lack' or eventive verbs like *ocurrir* 'to happen' and meteorological verbs.

(7)   a. ***En esta casa** falt<u>an</u> <u>ventanas</u>*.
           in this house miss.3PL windows
           'Windows are missing in this house.'

        b. ***En Barcelona** h<u>a</u> ocurrido <u>un accidente</u>*.
           in Barcelona have.3SG happened an accident
           'An accident has happened in Barcelona.'

However, the locative always appears with its corresponding preposition and it never triggers verbal agreement (which in such constructions may only be done by a post-verbal NP with the

appropriate features). So, it can be assumed that movement of a non-argumental constituent to a position in which it triggers verbal agreement is not possible in Romance languages other than BP (and African varieties of Portuguese).

It appears to me now (after the deeper theoretical study in previous chapters) that compared complexity measured by number of syntactic operations involved in the derivation may not be a reasonable explanation for the expected difference in the acquisition of the two syntactic structures under consideration.

First of all, in Spanish versions of BP topic-subject constructions, locative and genitive constituents may either remain *in situ* or move to the sentence's left periphery.

(8)   a. *Pega mucho el sol **en esa casa***.

            hit.3SG a.lot the sun in that house

        b. ***En esa casa** pega mucho el sol*.

            in that house hit.3SG a.lot the sun

        'That house gets a lot of sunlight.'

(9)   a. *Se ha pinchado la rueda del coche*.

            SE have.3SG punctured the tire of.the car

        b. *La rueda del coche se ha pinchado*.

            the tire of.the car SE have.3SG punctured

        c. Al coche se le ha pinchado la rueda

            to the car SE have.3SG punctured the tire

        'The car got a flat tire.'

Thus, the problems Spanish native speakers may have acquiring BP topic-subject constructions cannot be accounted for in terms of a lack of movement when these structures are analyzed under L1 grammar properties. There can be movement but just not to the same position, it is a movement of a different nature and hence the constituent cannot trigger verb agreement nor loose its preposition.

So, an analysis based on the number of syntactic operations, comparing what needs to be done to derive BP *in situ* interrogatives and topic-subject constructions to what is done in their Spanish counterparts does not really apply here. Wh *in situ* requires less operations, but topic-subject constructions do not necessarily require more.

But most importantly, such a measure for complexity may be flawed on principle. The idea came from a notion of economy, if they share the same numeration and they converge, a derivation with less syntactic operations is more economical than one with more. However, I am not sure what a comparison in these terms can tell us about the early acquisition of a determined structure compared with another.

I still believe though, that if a formal feature has its strength value set to strong in the L1, when learning a L2 that shares the same formal feature but with its strength value set to weak, the acquisition of L2 structures involving this feature should happen quite quickly. But I will go back to this in Chapter 7.

Another possible explanation that could account for the difference in acquisition comes from Avelar and Galves' proposal (2011, 2016). As seen in Chapter 2, the authors argue that in BP, the EPP in T is phi-independent, they claim the position of Spec-TP is inserted before Comp and without Case features. According to Avelar and Galves, this explains a number of other BP singularities involving the subject position, agreement variation and pronouns. The example below shows that BP allows prepositional subjects (10.a), subject-verb agreement is variable (10.b), and there is morphological uniformity in nominative and non-nominative positions (10.c). Again, this was discussed in more detail in Chapter 2.

(10)   a. *Na minha escola aceita cartão de crédito.*
            in.the my school accept.3SG credit card
            'My school accepts credit cards.'
         b. *As criança(s) brincavam/brincava na varanda.*
            the.PL children played.3PL/played.3SG in.the veranda
            'The children played in the veranda.'
         c. *Você foi visto na escola. / A Maria viu você na escola.*
            you.NOM were seen in.the school / the Maria saw you.NOM in.the school
            'You were seen at school.' / 'Maria saw you at school.'

Neither Spanish nor any other Romance languages share any of these characteristics. Their EPP in T is not phi-independent and they do not allow DPs to be inserted without Case features. As we have seen, these two properties of BP grammar have strong implications for what this language allows regarding the subject position. Therefore, it can be argued that going from a phi-dependent EPP grammar to a phi-independent one constitutes a much more substantial change than merely modifying the strength value of a shared formal feature. And so, this looks

like a more plausible reason for the expected difference in the acquisition of topic-subject constructions and *in situ* interrogatives.

I will resume the theoretical discussion in Chapter 7, after the empirical results have been analyzed.

Objectives

To sum up, the main objective of this experimental study is to describe and analyze how L2 BP learners behave regarding *in situ* interrogatives and topic-subject constructions in order to verify or refute the hypothesis proposed in (1). In other words, this research aims to check whether the syntactic differences between these constructions also translate into differences in their acquisition by Spanish native speakers learning BP as a second language, more specifically, if topic-subject constructions are acquired later (if at all) than *in situ* interrogatives. And, taking this as a departure point, theorize about the reasons that may underlie this difference.

The results of this experimental study, however, could also contribute to the discussion of wider theoretical problems providing answers to the following questions:

- What does the difference in the acquisition of these two syntactic constructions (if this difference is actually confirmed by the experiment results) tell us about the general theories of L2 acquisition?
- Do the obtained results support the UG access hypothesis? Or are they more in line with Representational Deficit approaches? And, what information can they give us on the role of L1 linguistic knowledge in L2 acquisition?

Although my main interest resides in L2 acquisition, as far as I know, BP topic-subject constructions have not been studied experimentally. The results of this study can then offer new information about these structures' acceptability and perhaps give us some hint about the processing cost for native (and nonnative) speakers.

4.3. Participants

The experiment was conducted on a total of 54 participants.

The control group contained 26 Brazilian Portuguese native speakers, all college students, or graduates. This ensured some uniformity for the group, which is very important, since sociocultural variation can have a strong influence in the acceptability of *in situ* interrogatives

and especially of topic-subject constructions. Although geographic variation may also have an impact, it was not possible to restrict the group to a single region of origin due to a lack of time and volunteers. However, most of the participants came from the State of São Paulo.

After careful examination of the data, I found one member of the control group that had given the highest rate to almost all the sentences in the AJT and whose reaction times were much shorter than those of the other participants. This person was excluded from the study.

The experimental group contained 24 Spanish native speakers from Spain[37] who were living in Brazil at the moment of testing. All of them were acquiring BP, most through formal training but some in a more naturalistic manner. The results of the proficiency test for this group showed scores ranging from 37 to 50 (out of 50) with a mean score of 43.8. This shows that the proficiency level of the group, as a whole, can be considered to be advanced.

A non-parametric version of the one-way ANOVA (Kruskal-Wallis Test) showed a significant difference in the proficiency test scores between the experimental and the control group ($H_1=23.2$, $p<10^{-5}$), a difference which is visually apparent in the following graph.

(11) BP Proficiency test results by group
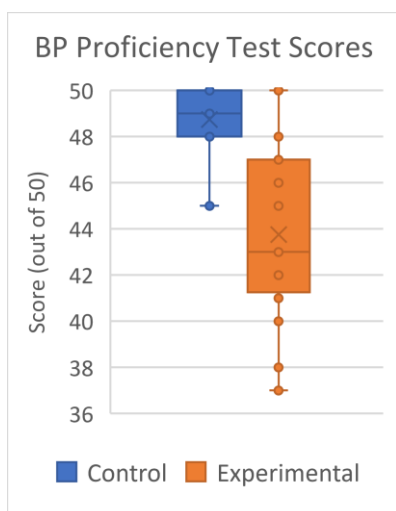
Considering the higher dispersion shown by their test results, the experimental group was divided into three subgroups according to their proficiency level. This division was based on the comparison of their test scores with those of the control group, on the scores of the experimental group in itself, and on the size of the subgroups. The purpose was to obtain a

---

[37] This restriction is due to the suspected differences between European Spanish and South American Spanish regarding the syntactic status of *in situ* interrogatives and the grammaticality of topic-subject constructions. According to Lipski (2012), along the Brazilian border, linguistic contact has resulted in occasional *in situ* questions in Spanish.

somewhat uniform distribution between the subgroups, which is essential for their statistical comparison.

Since 92% of the control participants scored 48 or higher on the proficiency test, the score range for the almost native-like group was established between 47 and 50. The lower limit was slightly relaxed (by one point) because keeping the limit at 48 left only 3 people in this group, a number too small to use statistic tools, whereas with 47 (which happened to be the third quartile) there were 8 participants in this subgroup. In addition to this, the fact that 2 people in the control group got a lower score than that was also considered.

The score ranges for the other two subgroups were determined using the first and third quartiles. The lower proficiency group was formed by the participants with scores below the first quartile ($Q_1 = 41.75$). And the participants with scores below the third quartile ($Q_3 = 47$) formed the middle group. Considering that the scores were in general quite high, I will refer to the subgroups as A – Intermediate, B – Advanced, and C – Almost native-like.

A minimum level of proficiency was required to understand the sentences that had to be judged in the task. Nonetheless, as mentioned above, all the participants in the experimental group scored at least 37 out of 50 on the proficiency test, so there was no need to exclude any of them for fear of lack of comprehension.

In this case, none of the participants in the group showed any kind of unwanted behavior during the experiment, nor did their results imply any erratic performances or misunderstandings of the tasks. Hence, there were no exclusions of this type.

The language background questionnaire provided more information relevant to the study, like how long the participants had been in the country, how much time they were daily exposed to BP and what kind of exposure it was, how they had acquired the language, or what other languages they spoke.

Some of the Spanish native speakers had been living in Brazil for two or three months to one year prior to their participation in the study, while others had been living in the country for several years. Although the test scores corresponding to the participants who had been living in Brazil for a short period of time were in the lower range, not all the long-term residents got high scores.

This was partially due to their daily exposure time to BP. Many of the participants were working in a mostly Spanish speaking environment, which reduced drastically the time spent interacting

with Brazilian Portuguese native speakers. Most of them reported a daily exposure time to BP between 10% and 30% of their usual day. They had acquired the language prior to their arrival in the country and mainly through formal training. This was not ideal, given the nature of the experimental study, but I will discuss this issue later (see chapter 7).

All members of the experimental group but one had studied English prior to having any contact with BP. Some had also studied French and German, and others came from the Basque Country and were bilingual in Spanish and in Euskera.

Participants that acquired a language with wh *in situ* like French prior to learning BP may show effects of transfer from that language to their BP grammar, which could interfere with the results. Exclusion of these speakers from the data analysis was considered. However, given the difficulty in recruiting participants and considering that, on the Language Background Questionnaire, they reported a low proficiency level in French and a very restricted presence of this language in their lives (in the past and especially in the present), I decided against their exclusion.

In addition to the experimental and control groups already discussed, there were also 4 other participants who did not fit into these two groups. During our initial conversation and in their Language Background Questionnaires, these participants declared themselves to be bilinguals. None of them considered any of the two languages, Spanish or Brazilian Portuguese, as being dominant, although one had trouble understanding me in Spanish. Since their backgrounds were quite different from one another and there were only 4 people in this category, it was not viable to treat them as an independent group. And, being bilinguals, they could not be part of the experimental nor the control group. Hence, I excluded their data from the group results analysis, but I included it in the individual results analysis.

The following table is a summary of the ages and proficiency test scores of the groups (after the pertinent exclusions).

Group	N	Proficiency test score (out of 50)		Age	
Control	25	Mean	**48.7**	Mean	**23.4**
		SD	(1.4)	SD	(4.1)
		Range	45 – 50	Range	19 – 35
Experimental	24	Mean	**43.8**	Mean	**29.4**
		SD	(3.8)	SD	(6.0)
		Range	37 – 50	Range	21 – 45
C – Almost native-like	8	Mean	**47.9**		
		SD	(1.4)		
		Range	47 – 50		
B – Advanced	10	Mean	**43.3**		
		SD	(1.7)		
		Range	42 – 46		
A – Intermediate	6	Mean	**39.0**		
		SD	(1.9)		
		Range	37 – 41		
Bilingual	4	Mean	**45.0**	Mean	**59.3**
		SD	(2.3)	SD	(8.3)
		Range	43 – 47	Range	50 – 70

Table 1 - Participant Characteristics

4.4. Procedure

4.4.1. *Acceptability Judgement Task (AJT)*

As I mentioned earlier, the main task of the experimental study was an acceptability judgement task. It is important to clarify that, although acceptability judgements were traditionally referred to as grammaticality judgements, these two concepts are related but are not the same. Roughly speaking, (un)acceptability is related to performance whereas (a)grammaticality has to do with competence. According to Guimarães (2017), the former constitutes raw empirical data, it is the speaker's reaction to a given sentence; the latter is a theoretical construct, it is an inference that a linguist makes based on acceptability judgements but also taking into account other theoretical, analytical and methodological aspects (such as the assumption that the speaker's reaction partially reflects his/her linguistic knowledge on a more abstract level).

This experimental paradigm has been severely criticized in the literature. Some authors are concerned with its metalinguistic nature, making the task artificial and unecological (Bresnan, 2007), others go as far as considering acceptability judgements as completely useless (Gethin, 1990). Its use in SLA research has also been the subject of scrutiny. Some authors question this

paradigm's reliability by pointing out inconsistencies between the results from acceptability judgement tasks and from other production-oriented tasks (like Ellis, 1991, among others), while others do exactly the opposite, defend the task's reliability comparing its results to those of a production task of some kind (see for example, Mandell, 1999).

Despite the controversy, acceptability judgement tasks continue to be a fundamental tool in the study of grammar. Being one of the most direct ways to obtain negative evidence, they offer an important empirical base for grammatical theories. Since this study aims to determine whether some syntactic constructions belong to the interlanguages of BP learners, i.e., if some syntactic constructions constitute negative evidence for the grammars of those interlanguages or not, this paradigm is the one best suited to our purposes.

Acceptability judgements of a given sentence, then, provide insight into its grammaticality within the interlanguage, an insight which will be used here as a way of determining whether the sentence's syntactic structure has been acquired by the learners. But the reader should bear in mind that this relation is not uncontroversial either. When can it be said that a specific syntactic structure has been acquired? How can this be quantitatively assessed? In the field of Language Acquisition, it is quite common to assume a threshold of 80% of acceptability. Intuitively, the threshold should be a somewhat high percentage, but any concrete number seems to be quite arbitrary and hard to justify. In this research, no such threshold was used. Instead, the learners' behavior towards the sentences under study was compared to that of native speakers to determine their acquisition status (statistically indistinguishable behavior being considered evidence of acquisition).

Finally, a more pragmatic advantage of this experimental paradigm is that it is one of the cheapest in terms of materials and equipment needed.

With the help of a computer, participants were asked to judge a series of 50 randomly ordered sentences in BP containing *in situ* interrogatives or topic-subject constructions. After reading some instructions, the sentences appeared on the screen one at a time and the participants had to rate each one using a 5-level Likert scale. Once the sentence was rated, the next one automatically appeared on the computer screen. To try to ensure reading, during the first 2 seconds, no response was allowed. The participant also had a maximum time of 15 seconds per sentence, after which the next sentence appeared on the screen (leaving the previous one unrated). The purpose of this time limit was to encourage instinctive answers and avoid overthinking. A reminder of the available ratings was always visible to facilitate the task.

1	2	3	4	5
'Horrível'	*'Estranha'*	*'Não sei'*	*'Aceitável'*	*'Perfeita'*
Horrible	Strange	Don't know	Acceptable	Perfect

Table 2 - Likert Scale

In addition to this score, the participants reaction times when rating the sentences were also recorded. This on-line measurement provides a concise way of comparing the processing costs of the two types of constructions under study and it can help distinguish two sentences that received the same score on the Likert scale. Some suspected unwanted behaviors such as automatic answers due to tiredness or disinterest can also be confirmed (or refuted) by this measurement.

The acceptability judgement task plus reaction times was built with the aid of the open software PsychoPy, version 1.90.1 for Windows (Peirce, 2007).

Four different lists of 50 sentences each were used in the experiment and distributed uniformly between the participants. The complete lists can be found in Appendix I. Below, I describe the structure of the lists and discuss how the sentences were selected.

Each list has 24 sentences intended to test topic-subject constructions. Half of those deal with locative agreement, as in (12.a) below, and the other half deal with genitive agreement, as in (13.a). Each of these two sets of 12 is divided in three smaller sets: 4 target sentences whose grammatical subject is a locative or genitive constituent (a), 4 sentences with the same verb as the target sentences but in a canonical structure (b), and 4 ungrammatical sentences copying the target structure but using unergative verbs, which do not admit[38] topic-subject constructions (c).

    (12)   <u>Sample sentences for **locative** agreement</u>

          a. *Essa casa bate bastante sol.*   (target)

              that house beat.3SG plenty sun

              'That house gets plenty of sunlight.'

          b. *Nessa praça não bate vento.*   (canonical)

              in.that square not beat.3SG wind

              'In that square, the wind does not blow.'

---

[38] This assertion is based on the existing literature on topic-subject constructions discussed in Chapter 2. The results of this study, however, may show a different picture of this structure's acceptability within native speakers, providing interesting new information about topic-subject constructions.

    c. *Aquela lagoa pesca muito pintado.*   (ungrammatical)

       that lagoon fish.3SG many catfish

       'In that lagoon, it is easy to catch catfish.'

(13)   Sample sentences for **genitive** agreement

    a. *Minha xícara quebrou a asa.*   (target)

       my cup broke the handle

       'My cup's handle broke.'

    b. *O teclado do computador quebrou.*   (canonical)

       the keyboard of.the computer broke

       'The computer keyboard broke.'

    c. *A Maria sempre chora o bebê à noite.*   (ungrammatical)

       the Maria always cry.3SG the baby at.the night

       'Maria's baby always cries at night.'

Both the number of the topic (locative/possessive constituent) and the number of the logical subject were controlled. The combination of these numbers is important because it is precisely when they do not match that verb agreement with the topic becomes evident. In each 4-sentence set (target, canonical and ungrammatical) there was one sentence for each combination: singular/singular (a), singular/plural (b), plural/singular (c), and plural/plural (d).

(14)   Number combination in target set for genitive agreement

    a. *O portão* (SG) *enferrujou a fechadura* (SG).

       the gate rusted.3SG the lock

       'The gate lock got rusty.'

    b. *O quarto* (SG) *está mofando as paredes* (PL).

       the room is molding the walls

       'The room walls are getting moldy.'

    c. *Os bateristas* (PL) *sangraram a mão* (SG).

       the drummers bleeded.3PL the hand

       'The hand of the drummers bled.'

    d. *Os armários* (PL) *arranharam as portas* (PL).

       the closets scratched the doors

       'The doors of the closets were scratched.'

The choice of verb was also controlled. As seen in the examples, each target sentence was built using an unaccusative verb. The same verb was used for the canonical version of the sentence,

for which the rest of the words were different. These pairs of sentences sharing the same verb (target and canonical) were placed in different lists so that no verb was repeated in any of the lists.

Since the ungrammatical versions of the sentences were built using completely different verbs (because they needed to be unergative to ensure their ungrammaticality), there was no need to have 4 different sets. So, lists 1 and 3, and lists 2 and 4 shared the same ungrammatical sentences.

As for the interrogative sentences, two different sets of 26 items were chosen. Again, lists 1 and 3, and lists 2 and 4 shared the same set. In each set, 14 sentences had argumental wh constituents: 4 subjects, 6 direct objects and 4 indirect objects. The 12 remaining sentences in each set had non-argumental wh constituents questioning when, where, how, what for, why and with whom (2 sentences each). For each of those subsets, half of the sentences were *in situ* interrogatives and the other half were fronted ones.

(15)   Examples of **argumental** wh constituents

    a. *A Maria queria o que?*   (DO)   (*in situ*)
        'Maria wanted what?'

    b. *O que que o Pedro vai comprar?*   (DO)   (fronted)
        'What (is) that Pedro is going to buy?'

    c. *A Maria emprestou a caneta para quem?*   (IO)   (*in situ*)
        'Maria lent the pen to whom?'

    d. *Para quem o Pedro enviou aquele pacote?*   (IO)   (fronted)
        'To whom did Pedro sent that package?'

(16)   Examples of **non argumental** wh constituents

    a. *A festa é quando?*   (Temporal)   (*in situ*)
        'The party is when?'

    b. *Quando seus pais vão embora?*   (Temporal)   (fronted)
        'When are your parents leaving?'

    c. *Os alunos vão se reunir onde?*   (Locative)   (*in situ*)
        'The students are going to meet where?'

    d. *Onde que fica a loja?*   (Locative)   (fronted)
        'Where (is) that the store is located?'

Some clarifications are in order. First, notice that two of the sentences in the examples ((15.b) and (16.d)) have a doubly filled Comp, that is, an 'extra' complementizer beside the interrogative pronoun. Actually, half of the fronted interrogatives in the lists were built with doubly filled Comp. This was due to the results of a previous pilot study, which showed that some participants rejected perfectly grammatical sentences simply because of this complementizer and others because of its lack thereof. Having the same number of sentences with and without a doubly filled Comp is a way of neutralizing this division of preferences. However, as we will discuss shortly, the follow-up task will reveal the reasoning behind any negative judgement and provide us with a way of dealing with this kind of situation.

More importantly, recall that each list contained 4 interrogative sentences that question the subject. *A priori*, considering the *in situ* versions here does not seem to make much sense because the apparent order of the constituents remains unchanged. However, there is a double intention behind this. On one hand, I want to compare the participant's reaction to interrogative sentences with unaccusative verbs and post-verbal subjects (which in this case are allowed, since they are nothing but an internal argument), such as (17.a), with the reaction to the same structure with unergative verbs, as in (17.b). On the other hand, if the control group rejects the unergative versions, this constitutes an argument against the analyses for BP *in situ* interrogatives that consider them only apparent (such as the one proposed in Kato, 2013).

(17)   Contrast in interrogative sentences questioning the subject

    a. *Chegou quem na aula?*　　　　　(unaccusative)
      arrived who in.the class
      'Who arrived in class?'

    b. *Tossiu quem na reunião?*　　　　(unergative)
      coughed who in.the meeting
      'Who coughed in the meeting?'

Note that apart from these sentences questioning the subject (which can be controversial as to their grammaticality), all the other interrogatives are grammatical. It is possible that this could induce a bias towards acceptability in our participants. However, this effect is an issue that affects mostly experimental studies with children, but it is not as relevant with adults.

Finally, given the diverse nature of the sentences designed to test the structures under study and also considering the lengthiness of the tasks, filler sentences were deemed unnecessary.

4.4.2. *Follow-up: Correction Task*

The acceptability judgement task was immediately followed by a 'correction' task that served as a follow-up. Here, again using a computer and after carefully reading some instructions, each participant was asked to reformulate the sentences to which he/she had assigned a score lower than 3 on the Likert scale, i.e., those that him/herself had previously considered to be unacceptable (either 'horrible' or 'strange').

The sentences were again randomly ordered and appeared on the screen one at a time. The participants were asked to turn them into acceptable versions of the original sentences, according to their own opinion. The instructions clearly stated and stressed to keep the modifications to a minimum trying to focus on the structure of the sentence and not on the choice of words or the spelling. There was no time limit for this task, the next sentence only appeared after the participant had pressed 'enter'.

The purpose of the correction task is to discover what was the specific reason behind the rejection of a sentence. This allows us to verify whether the cause of rejection was in fact the structures that interest us here and not unknown lexical items, hypercorrection of some other syntactic construction or any other external factors. This follow-up, together with the previously recorded reaction times, provide a very interesting and reliable source for a deeper qualitative analysis of the results obtained in the acceptability judgement task.

4.4.3. *Language Background Questionnaire and BP Proficiency Test*

In addition to the experiment in itself, participants had to perform two extra tasks supplying information about themselves and their relationship with BP: a language background questionnaire and a BP proficiency test.

The language background questionnaire was designed based on several similar questionnaires found in the literature, adding, removing and modifying some of the questions. It consists of a total of 8 questions regarding the languages that the participant speaks. The objective is to find out the participant's mother tongue, dominant language, time of exposure to each of the languages, first age of contact with them, most influential factors in the participant's learning experience, self-assessed proficiency level, and other information of this kind that can have a relevant impact on the main tasks results. The complete questionnaire can be found in Appendix II. This questionnaire was given only to the experimental group.

Proficiency level in the tested language is obviously an essential factor for organizing the data analysis (sorting the group in subgroups, for example). The proficiency test administered to the participants consisted of 50 multiple-choice items, 30 of which were intended to test grammar and the remaining 20 to test vocabulary. It was a cloze test, which means that each item contained a sentence in BP in which a word or phrase had been omitted and 4 possibilities to fill the gap.

It was the same test used in Ionin et al. (2015), where it was administered to Spanish and English native speakers. It was kindly provided to us by the authors, who had already used it in previous work. Ionin and her colleagues based the test on the Spanish proficiency test created by the Spanish Ministry of Education, known as DELE (*'Diploma de Español como Lengua Extranjera'*). The reliability of the test was confirmed by a reliability analysis conducted in Montrul et al. (2011). The complete test can be found in Appendix III.

Chapter 5
Results Analysis

5.1. Introduction

This chapter is devoted to the discussion of the study results. I will use both descriptive and inferential statistics for the quantitative analysis of the AJT results, and qualitative analysis for the follow-up task.

First, in section 5.2, I will consider the results by group, comparing some structures between the experimental and control groups. The results of the three experimental subgroups sorted by proficiency level will also be compared between each other and with the control group. Then, in section 5.3, I will look into the individual behavior of the participants.

Throughout the inferential analysis, I will use non-parametric versions of one and two-way ANOVAS, Kruskal Wallis and SHR tests, respectively. For post-hoc tests, I will use the Wilcoxon test with Bonferroni adjustment method, again a non-parametric version of the better-known Tukey test. The nature of our data requires non-parametric versions because it does not meet the requirements to apply the parametric tests (normality of the residues and homogeneity of variances). I will take a 5% significance level.

The qualitative part of the analysis was focused on the results of the follow-up or correction task and can be found in section 5.4. There, I will show and discuss many examples of how the participants rewrote the sentences they rejected on the AJT to understand better the reasons behind the rejections.

This qualitative analysis had a crucial methodological impact on the quantitative analysis. The data corresponding to sentences rejected by reasons other than those under study or rejected by unknown reasons were removed from the quantitative analysis in sections 5.2 and 5.3, see section 5.4.1 for a more detailed explanation.

5.2. Group Results

Let us start the discussion by studying the scores attributed by both groups of participants to each kind of sentence during the Acceptability Judgement Task. Recall that the scale went from 1 – 'horrible' to 5 – 'perfect', and that the experimental group was divided into three subgroups according to their proficiency level: group A – Intermediate, group B – Advanced, and group C – Native-like.

5.2.1. _Interrogative Sentences_

As we can see both in the table and in the boxplots below, neither the experimental nor the control group showed a preference for fronted interrogatives against _in situ_ ones. The experimental group gave slightly lower ratings in general and their data was a tad more disperse, but, looking at the graphs, the two groups seem to behave in a very similar way.

Group		Fronted		In situ
Experimental	Mean	**4.12**	Mean	**4.31**
	SD	(1.11)	SD	(0.95)
Control	Mean	**4.62**	Mean	**4.67**
	SD	(0.77)	SD	(0.64)

Table 3 – Interrogatives Mean Scores (AJT)

(1)   Boxplot graphs for Interrogative Scores in the AJT – Experimental vs. Control
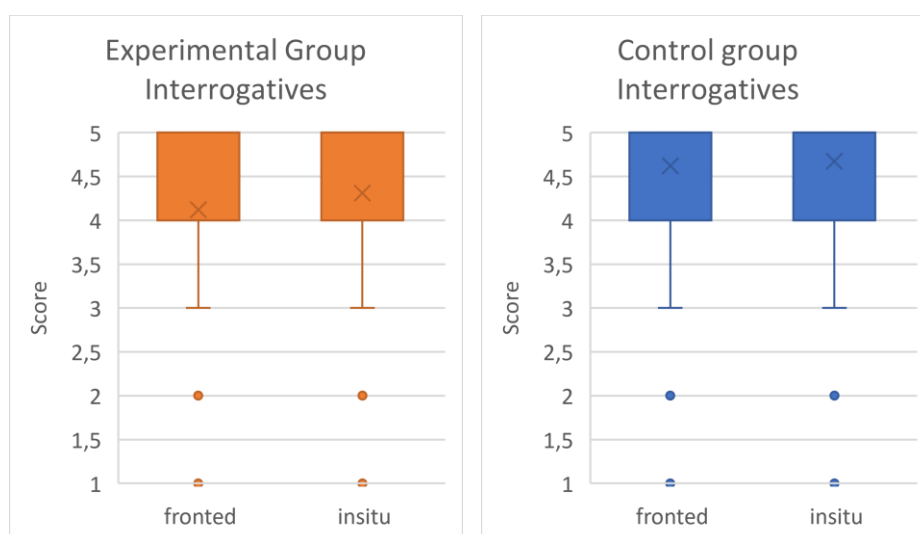
A non-parametric version of a two-way ANOVA showed there was a significant difference between groups ($H_{1,1047}=39.68$; p~0). Bear in mind that, whenever there is a considerable

amount of data, small differences may be statistically significant, but this does not necessarily imply they are relevant. In this case, the fact that the group was a significative factor indicates there is a statistically significative difference between experimental and control ratings for interrogative sentences considered as a whole, but mean scores were 4.22 and 4.65, respectively, which means both groups found interrogative sentences acceptable. Hence, this difference is not indeed relevant for us.

The statistical test showed as well that the type of interrogative, fronted or *in situ*, was not a significative factor ($H_{1,1047}=$; p=0.37), and that there was no interaction ($H_{1,1047}=$; p=0.50) between the two factors either, group and type of interrogative. The fact that there was no interaction means that if we consider the two sets of interrogatives divided by type, both groups behaved similarly in their ratings of each set's sentences.

All three subgroups of the experimental group showed exactly this behavior as well. None of them preferred a specific interrogative structure. However, the mean ratings for both kinds of interrogatives grew ever so slightly with the proficiency level, as can be seen in table 4. A non-parametric version of a two-way ANOVA confirmed there was no significant difference between the three subgroups ($H_{2,503}=4,22$; p=0.12), nor between the two types of interrogatives ($H_{1,503}=1.52$; p=0.22). There was also no interaction between those two factors ($H_{2,503}=3.84$; p=0.15).

Group		Fronted		In situ
A – Intermediate	Mean	**3.95**	Mean	**4.05**
	SD	(1.39)	SD	(1.09)
B – Advanced	Mean	**4.09**	Mean	**4.42**
	SD	(1.07)	SD	(0.84)
C – Native-like	Mean	**4.30**	Mean	**4.38**
	SD	(0.89)	SD	(0.95)

Table 4 – Interrogatives Mean Scores (AJT) by proficiency level

(2)   Experimental subgroups by proficiency level – Boxplot graphs for AJT Interrogative
        Scores
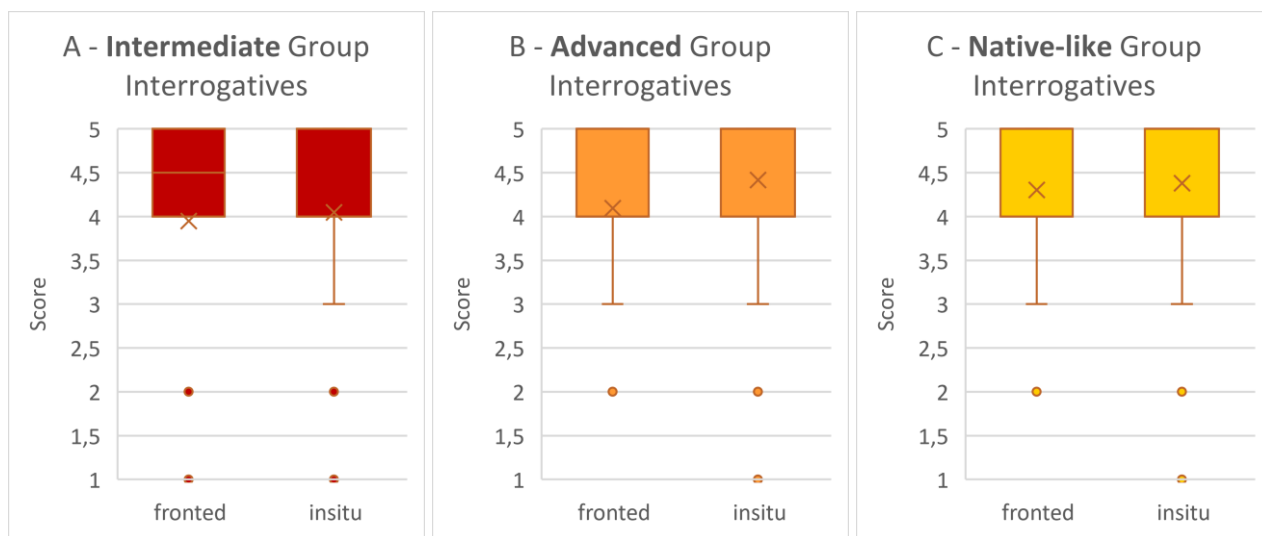
I excluded from this analysis the interrogatives questioning the subject of the sentence (those described in the previous chapter, example (17)) because they cannot be compared with the rest. Contrary to Kato's analysis (2013), I expected to find a higher rejection of postverbal subjects in questions with unergative verbs than in questions with unaccusative verbs, at least within the control group.

As we can see in the following graphs, this was not exactly the case. There was a small preference for postverbal subjects with unaccusative verbs in the control group (the mean, median and mode show this preference). In the experimental group the difference was even more subtle. The inferential tests, though, show these were not significant differences, there was no interaction between the type of verb and the group ($H_{1,93}$=1.29; p=0.26). However, the type of verb was a significant factor in itself ($H_{1,93}$=4.38; p=0.036). This means that, considering both groups together, the scores for postverbal subjects with inergative verbs differ significantly from those for postverbal subjects with unaccusative verbs. Although there was no interaction between the factors, I decided to run a Wilcoxon post-hoc test which showed that the difference between scores for sentences with unergative verbs and unaccusative ones within the control group almost reached significance (p=0.067).

Since this was not the main concern of our experimental study, the size of the available data is not big enough to reach a trustworthy conclusion from these results.

(3)　　Boxplot graphs for Interrogatives questioning post-verbal subjects (AJT Scores)
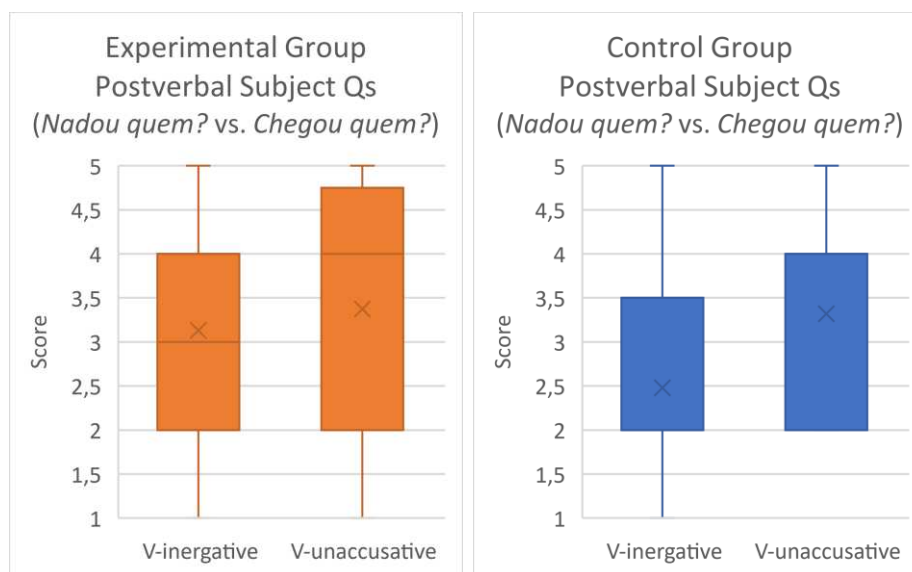
Note, however, that both groups showed lower scores for interrogatives questioning a postverbal subject than *in situ* interrogatives questioning other constituents (compare the graphs in (3) with the right boxes in the graphs in (1)). This is so even for the unaccusative verbs, which should allow a postverbal subject, since it is the verb's internal argument.

Since within the set of interrogative sentences some questioned argumental constituents and others non-argumental constituents, the results were studied considering this distinction to see whether it was an influential factor or not. The control group showed no difference in behavior regarding this distinction, as can be seen in the boxplot graphs. Statistic inferential tests (again, a non-parametric version of a two-way ANOVA) confirmed that neither the condition (argumental vs. non-argumental) nor the type of question (fronted vs. *in situ*) played a significant role ($H_{1,538}=0.01$; $p=0.91$ and $H_{1,538}=0.50$; $p=0.48$, respectively) and there was no interaction between them either ($H_{1,538}=0.02$; $p=0.89$).

(4)   Control Group – Boxplots for Argumental *vs*. Non-argumental Interrogatives (AJT Scores)
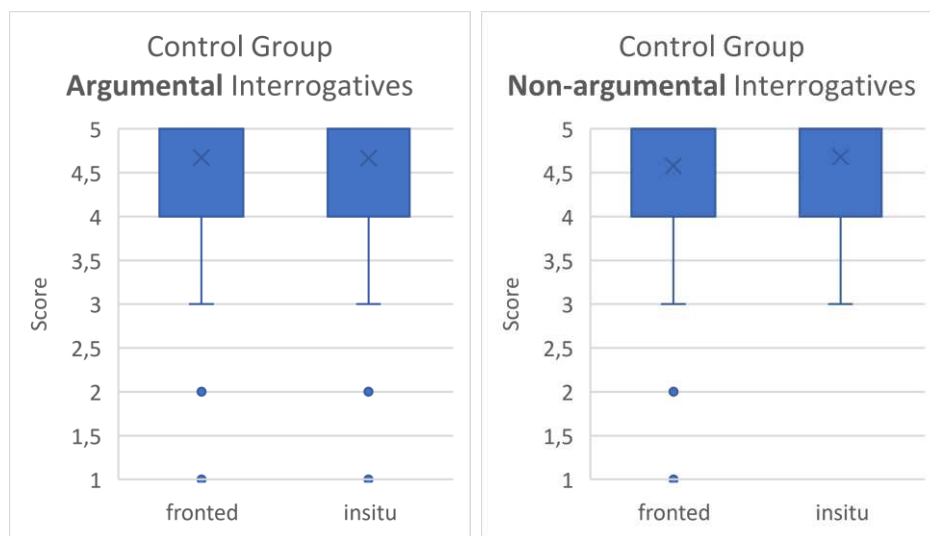
Like the control group, the experimental group did not show a preference for fronted or *in situ* interrogatives in any of the two conditions, argumental *vs*. non-argumental. The statistical tests confirmed that, within the experimental group, there was no difference between conditions ($H_{1,505}$=1.52; p=0.21), nor between type of question ($H_{1,505}$=1.58; p=0.22), nor interaction between the factors ($H_{1,505}$=0.99; p=0.32).

(5)   Experimental Group – Boxplots for Argumental *vs*. Non-argumental Interrogatives (AJT Scores)
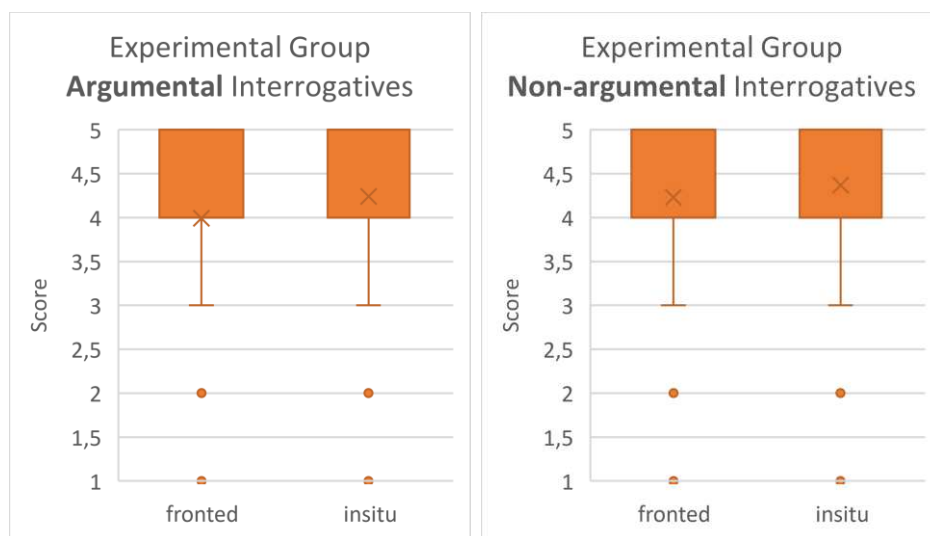
However, since the preliminary data showed this group had a slight preference for *in situ* interrogatives in the argumental condition, we decided to investigate further and looked at these results divided by proficiency group.

(6)   Experimental Subgroups by proficiency level – Boxplots for Argumental and Non-argumental Interrogatives (AJT Scores)
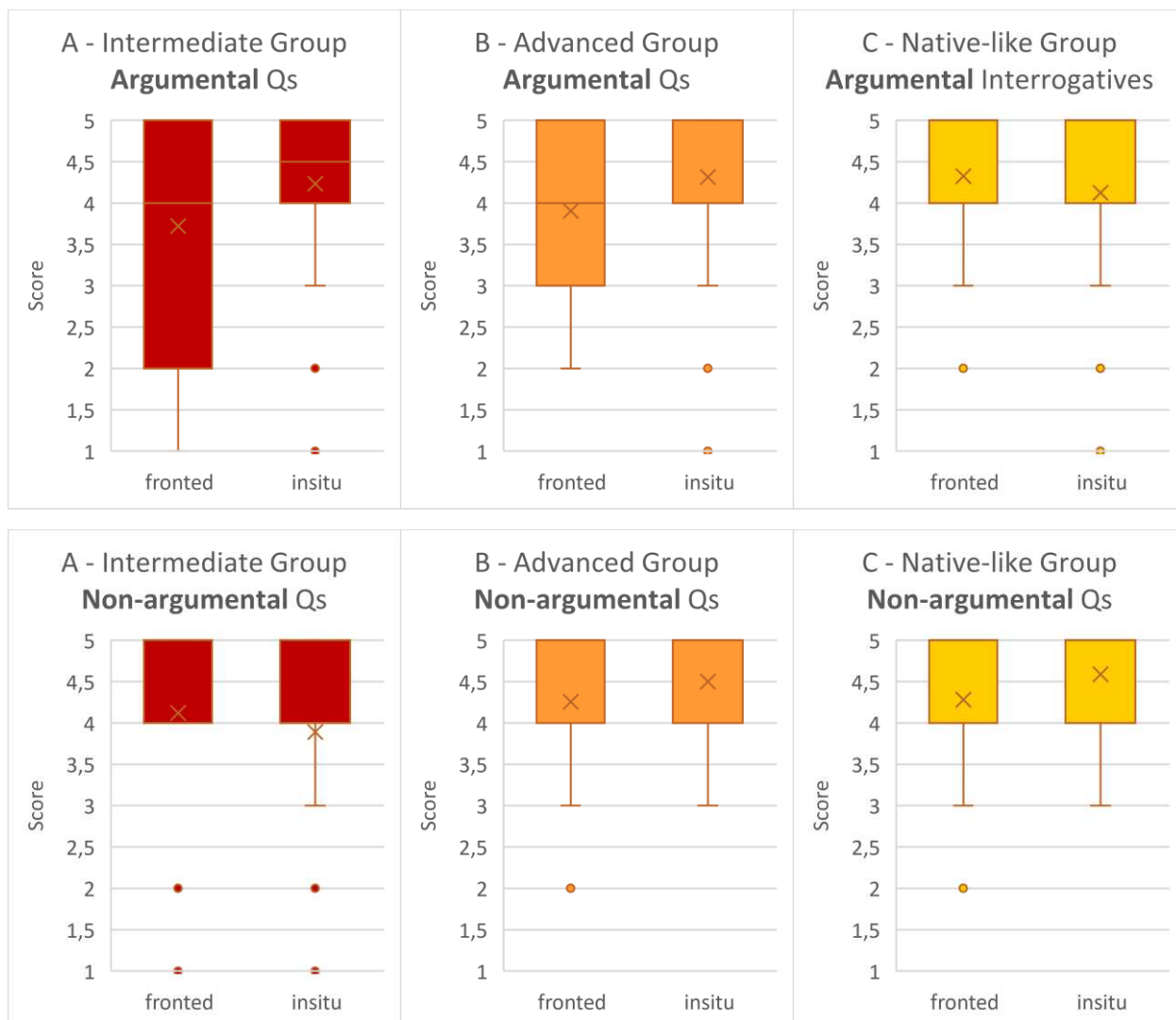
Do not let these graphs mislead you, though. What they show is only a bigger dispersion in the argumental condition for the intermediate and advanced groups. The differences that can be seen visually did not reach significance.

A non-parametric version of a two-way ANOVA test was applied to the argumental set of interrogatives and it showed that neither the type of interrogative (fronted or *in situ*), nor the experimental subgroup (A, B or C) were significant factors ($H_{1,224}$=2.31; p=0.13 and $H_{2,224}$=0.93; p=0.63, respectively), and there was no interaction either ($H_{2,224}$=2.69; p=0.26). The same was done to non-argumental interrogatives and the test confirmed neither the type of sentences nor the subgroup were significant factors ($H_{1,273}$=0.04; p=0.83 and $H_{2,273}$=3.75; p=0.15, respectively), but there was interaction between them ($H_{2,273}$=6.56; p=0.038). A Wilcoxon post-hoc test revealed that the scores given by subgroup A to *in situ* interrogatives

differed significatively from those by the other two subgroups to that kind of interrogatives (p=0.045 for subgroup B and p=0.021 for subgroup C).

Although the differences seen for fronted interrogatives were not significant, low scores for these sentences were not expected *a priori*. However, I suspected this was a possibility from a previous pilot study, the one responsible for the inclusion as experimental items of interrogatives with doubly filled Comp (i.e., a complementizer beside the interrogative pronoun), as discussed in section 4.4.1. In fact, the following graph shows the sentences with this 'extra' complementizer were those responsible for low scores received by fronted interrogatives. This was further confirmed by the qualitative analysis of the follow-up task, which can be seen in detail in section 5.6.2.

(7)   Experimental Group – Fronted Interrogatives with doubly filled Comp *vs*. without it
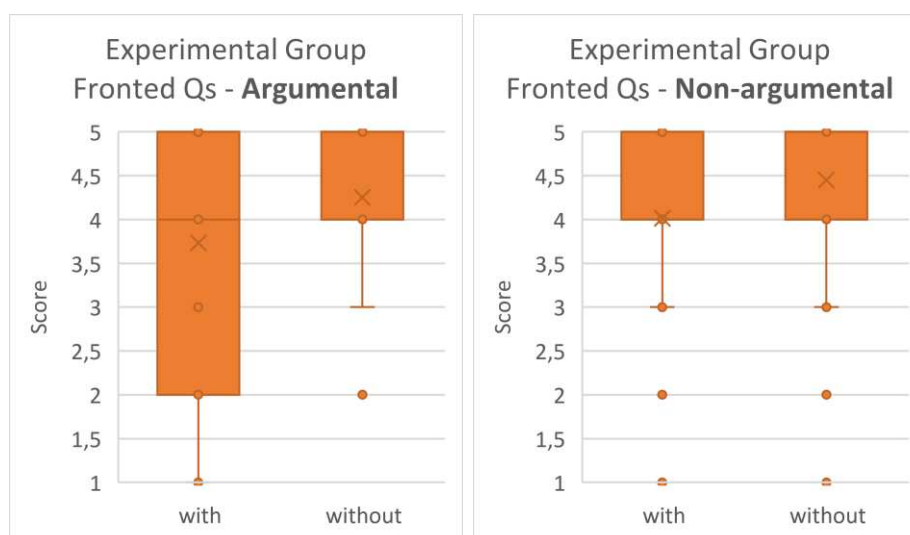
Again, BP's doubly filled Comp in interrogative sentences was not the main focus of this experimental study, so there is not enough data to extract concise conclusions. Nonetheless, these results suggest, at least, that this BP structure makes for an interesting research topic.

5.2.2. *Topic-Subject Constructions (TSC)*

As we can see below, in Table 5 and in the boxplot graphs in (8), the experimental and control groups behaved in a much more similar way than expected, regarding topic-subject constructions. The experimental group shows a slightly higher variation in their ratings of target sentences and lower mean scores (except for the ungrammatical versions). The differences, however, are smaller than expected.

Group	Canonical			Target		Ungrammatical
Experimental	Mean	**4.15**	Mean	**2.56**	Mean	**2.07**
	SD	(1.06)	SD	(1.32)	SD	(1.25)
Control	Mean	**4.49**	Mean	**3.23**	Mean	**1.86**
	SD	(0.88)	SD	(1.36)	SD	(1.03)

Table 5 – Topic-subject Constructions Mean Scores (AJT)

(8) Boxplot graphs for Topic-subject Constructions Scores in the AJT – Experimental *vs*. Control
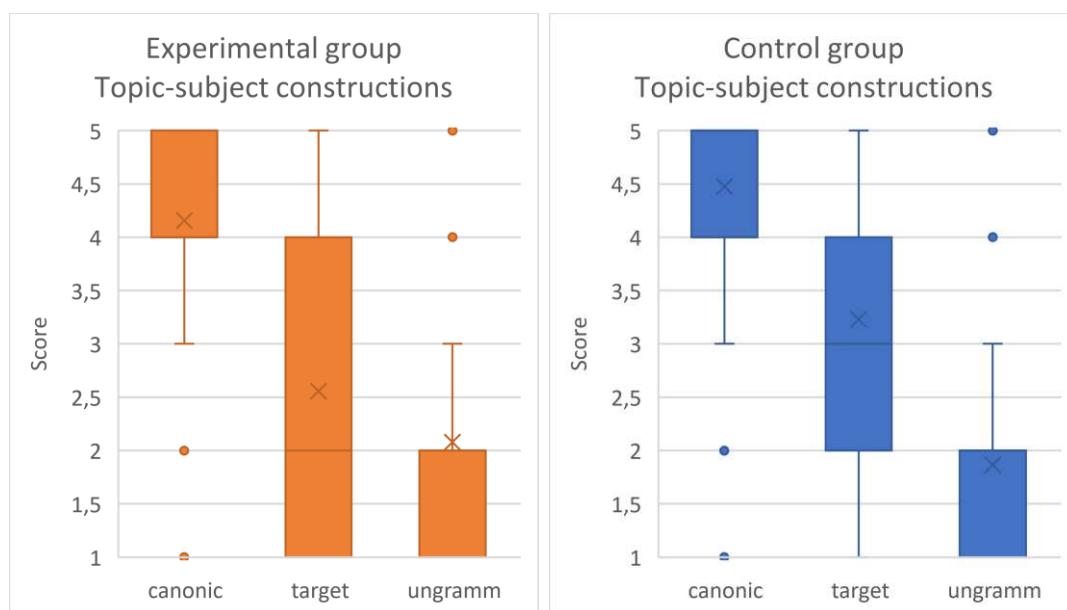
Contrary to what happened with the interrogative sentences, the type of structure was a very relevant factor here. Recall that the canonical sentences were those in which the locative or genitive appeared introduced by a preposition (and in adnominal instead of pre-verbal position, in the case of genitives), the target sentences were true topic-subject constructions, and the ungrammatical versions shared the target structure but used an unergative verb instead of an unaccusative one. See examples (12) and (13) in Chapter 4 for further clarification.

A non-parametric version of a two-way ANOVA confirmed that the type of structure was a significant factor influencing the scores ($H_{2,1134}$=417.31; p~0). Although the behavior of the two groups seems quite similar from the graphs (as I mentioned above), the statistical test showed this was also a significant factor ($H_{1,1134}$=7.34; p=0.007). In addition, there was interaction between the two factors, group and structure ($H_{2,1134}$=18.89; $p<10^{-4}$). A post-hoc Wilcoxon test revealed the differences were significative between all possible combinations but

one, experimental and control scores for ungrammatical TSCs were statistically undistinguishable.

In spite of reaching statistical significance, the difference between experimental and control ratings for the set of canonical sentences is not very relevant, since their mean scores are both over 4 (4.15 and 4.49 respectively), which means these sentences were considered acceptable by the two groups.

The fact that these three types of sentences were treated each in its own different way was somewhat expected, especially for the control group. However, I anticipated higher scores for target sentences in the control group, and lower scores for them in the experimental group. I also thought that the experimental group would have more trouble distinguishing between target and ungrammatical versions of TSCs. I will come back to discuss the possible reasons responsible for these unexpected results in Chapter 7.

So, as I was saying, these results were not exactly in line with the expectations. But looking at the experimental group by proficiency level provided a clearer picture. Note, in the following boxplot graphs, how the behavior regarding TSCs evolves with the participant's proficiency level.

Group A, with the lowest BP level, behaves more erratically, there is more dispersion in their data in general. In group B, there is a little less dispersion and the scores for canonical sentences slightly improve while the scores for ungrammatical sentences slightly decrease. The mean values change in that same direction, although the change is minor. The most interesting behavior, though, is that of the native-like group. In this case, data dispersion is considerably reduced, and, most importantly, the scores for canonical sentences improve a little, those for ungrammatical sentences continue to decrease, and the scores for target sentences (which did not change from group A to group B) change quite drastically in this group.

Group	Canonical			Target	Ungrammatical	
A – Intermediate	Mean	**3.86**	Mean	**2.81**	Mean	**2.36**
	SD	(1.27)	SD	(1.38)	SD	(1.26)
B – Advanced	Mean	**4.11**	Mean	**2.83**	Mean	**2.23**
	SD	(1.02)	SD	(1.28)	SD	(1.35)
C – Native-like	Mean	**4.44**	Mean	**1.95**	Mean	**1.59**
	SD	(0.86)	SD	(1.12)	SD	(0.92)

Table 6 – Topic-subject Constructions Mean Scores (AJT) by proficiency level

(9)   Experimental subgroups by proficiency level – Boxplot graphs for Topic-subject Constructions Scores in the AJT
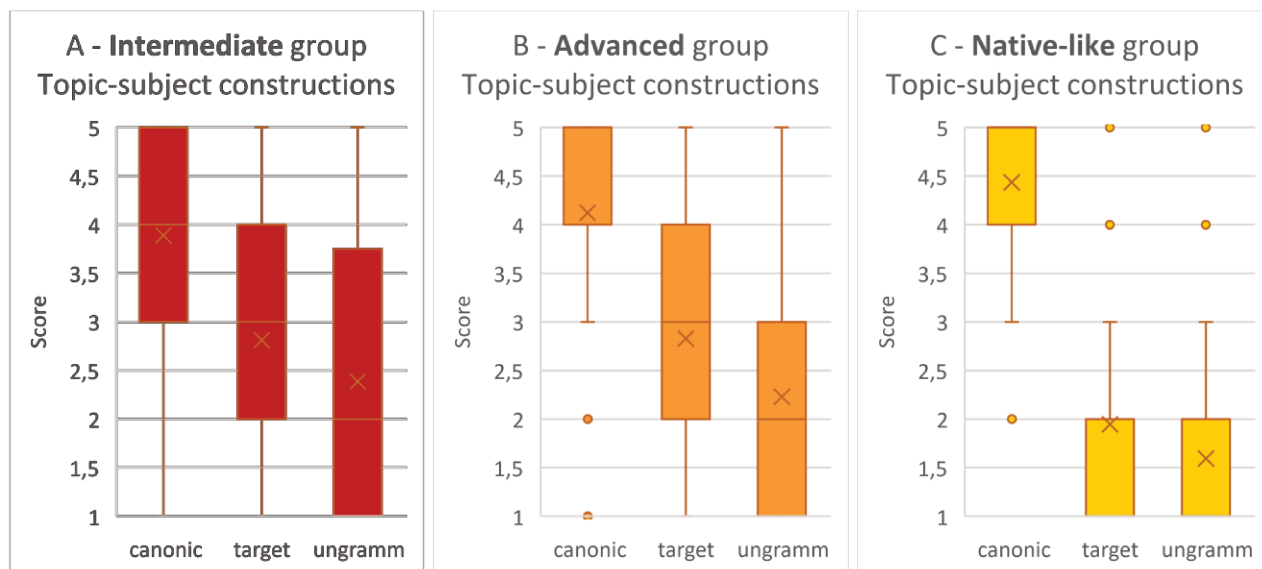
In fact, this analysis based on descriptive statistics was confirmed by inferential statistics. Both the type of structure (canonic *vs.* target *vs.* ungrammatical) and the subgroup by proficiency level were significant factors ($H_{2,543}$=168.80; p~0 and $H_{2,543}$=7.56; p=0.02, respectively), and there was also interaction between them ($H_{4,543}$=16.78; p=0.002).

A post-hoc Wilcoxon test revealed that, within each subgroup, canonic sentences received significatively different scores from the other two types, which did not differ between each other for any of the groups (although for the intermediate group significance was almost reached).

It also revealed that there was no significative difference between the three subgroups in their scores for canonical sentences, nor between group A and B in their scores for target and ungrammatical TSCs. But group C did differ in its ratings for target and ungrammatical TSCs from the other two groups. The native-like participants clearly rejected these kinds of sentences, and they did not differentiate between them.

Surprisingly, the other subgroups, A and B show a behavior regarding target TSCs remarkably similar to the control group. In fact, including the control group in the inferential test revealed that there was no significative difference in scores for target TSCs between these two subgroups and the control group, but there was such a difference between subgroup C and the control group (which did fall under our expectations). Although uneven ratings were expected for the

experimental groups, it was somewhat surprising to find this behavior was shared by the control group too.

In any case, it is crucial to investigate whether individual behavior is consistent or not. This will help us discern if the group behavior regarding TSCs is all over the place because each individual behaved in an erratic way, or because some participants accepted TSCs whereas others did not. This will be done in section 5.3.

Note that canonical sentences were grammatical BP sentences that were expected to receive high scores. Unexpected rejection of these sentences was related to the position of the locative constituent or to the addition of the pronoun *se* in genitive TSCs. I will go into this in more detail when discussing the follow-up task results in section 5.4.

To further investigate the results concerning topic-subject constructions, I considered two more factors that could play a role in the matter. The first one was the grammatical number combination of the topic-subject constituent and the sentence's logical subject. As discussed in the previous chapter (see example (14)), when these elements have different grammatical number, verb agreement with the topic-subject constituent becomes much more evident. This led me to believe that, at least for the experimental group, target sentences in this case may receive lower scores than in the opposite case, where both elements have the same grammatical number and agreement might go unnoticed. However, no such difference was present in the results.

What happened, though, is that the control group gave slightly higher ratings to target sentences in the equal number condition, and that the experimental group showed more dispersion for target and ungrammatical sentences in the different number condition.

(10)   Boxplot graphs for topic-subject constructions with the same *vs*. different grammatical number between topic-subject constituent and the logical subject (AJT Scores)
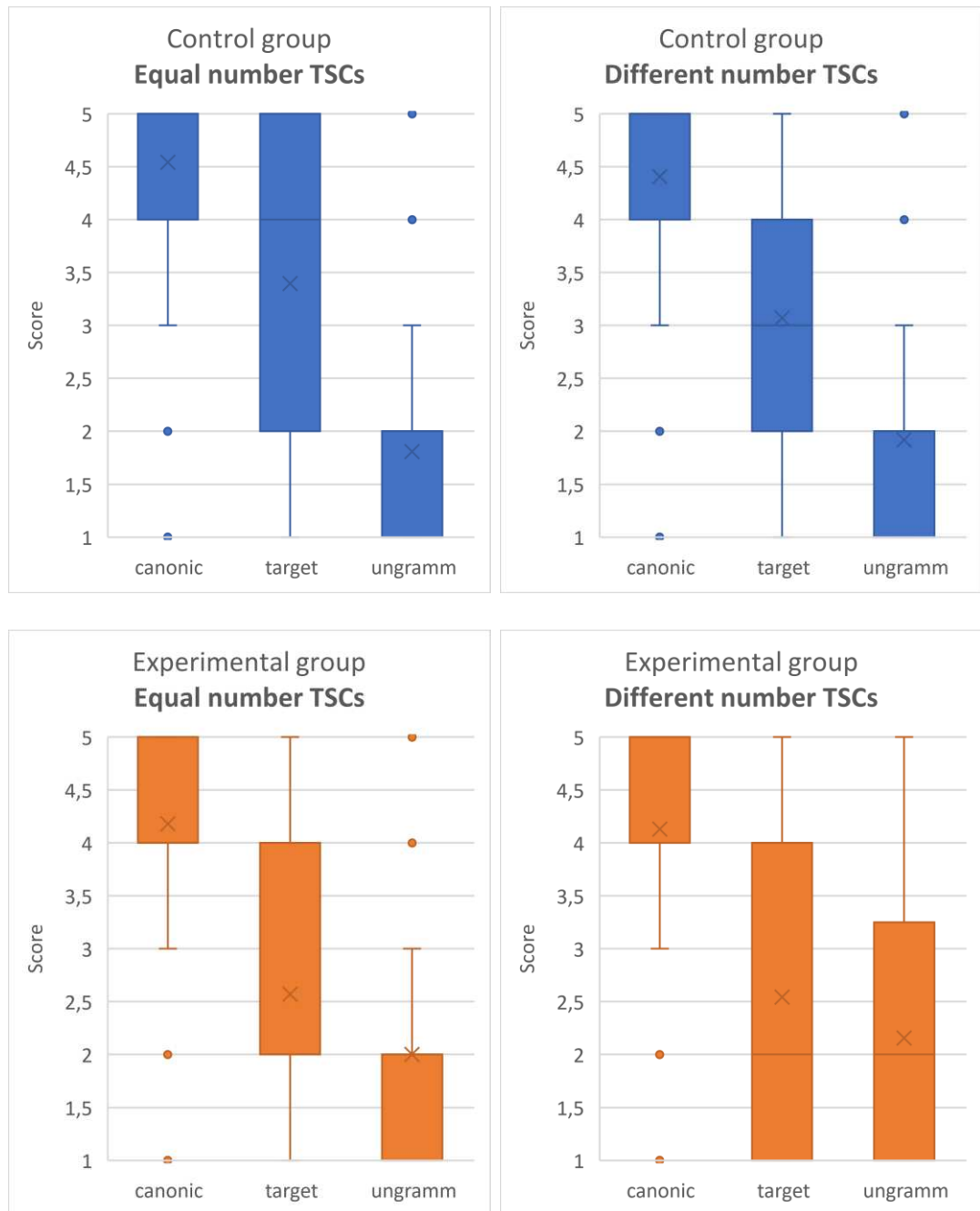
A non-parametric version of a two-way ANOVA test applied both to the experimental and the control group showed that number combination (equal *vs*. different) was not a significative factor for either of the two groups ($H_{1,546}=0.04$; $p=0.85$ and $H_{1,582}=0.916$; $p=0.34$ respectively) and there was no interaction with the structure of the sentence (canonical, target and ungrammatical) ($H_{2,546}=0.97$; $p=0.62$ and $H_{2,582}=2.305$; $p=0.32$, respectively).

In this case, even dividing the experimental group in their subgroups by proficiency level, grammatical number did not provide any further information, since it did not have a strong impact on the results.

The second factor considered for the investigation of topic-subject constructions was the type of construction (locative vs. genitive), which did not prove to be very relevant either. As can be seen from the graphs below, the control group was marginally more accepting of locative target TSCs. The experimental group, on the other hand, seems to be a bit more inclined to accept target and ungrammatical versions of genitive constructions rather than locative, which was again against my initial intuition and expectations.

(11)   Boxplot graphs for locative vs. genitive topic-subject constructions (AJT Scores)
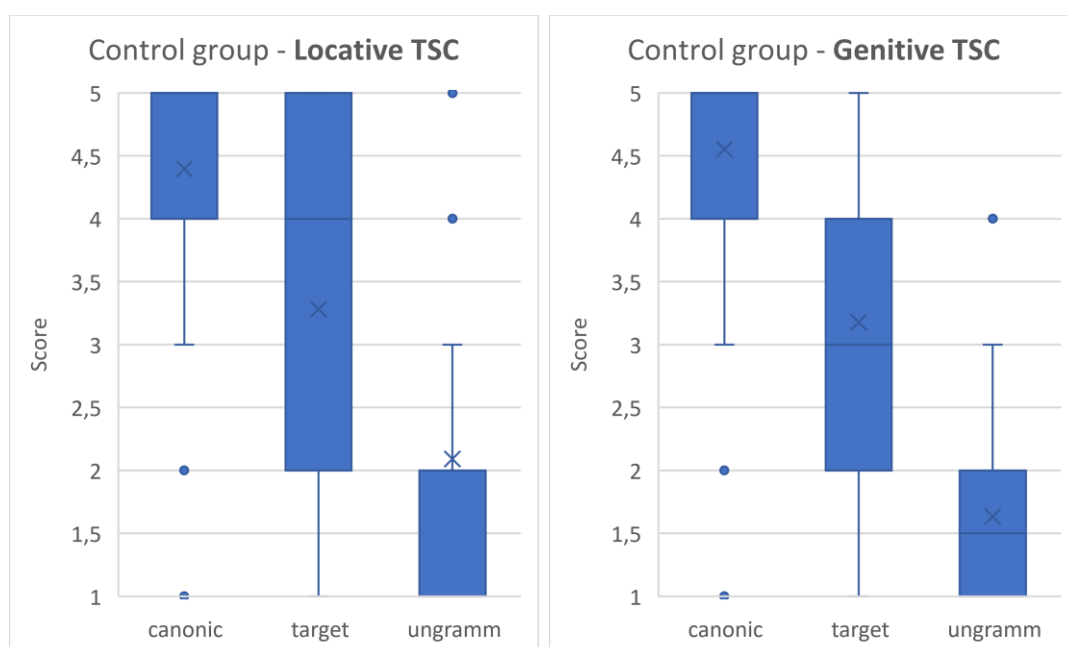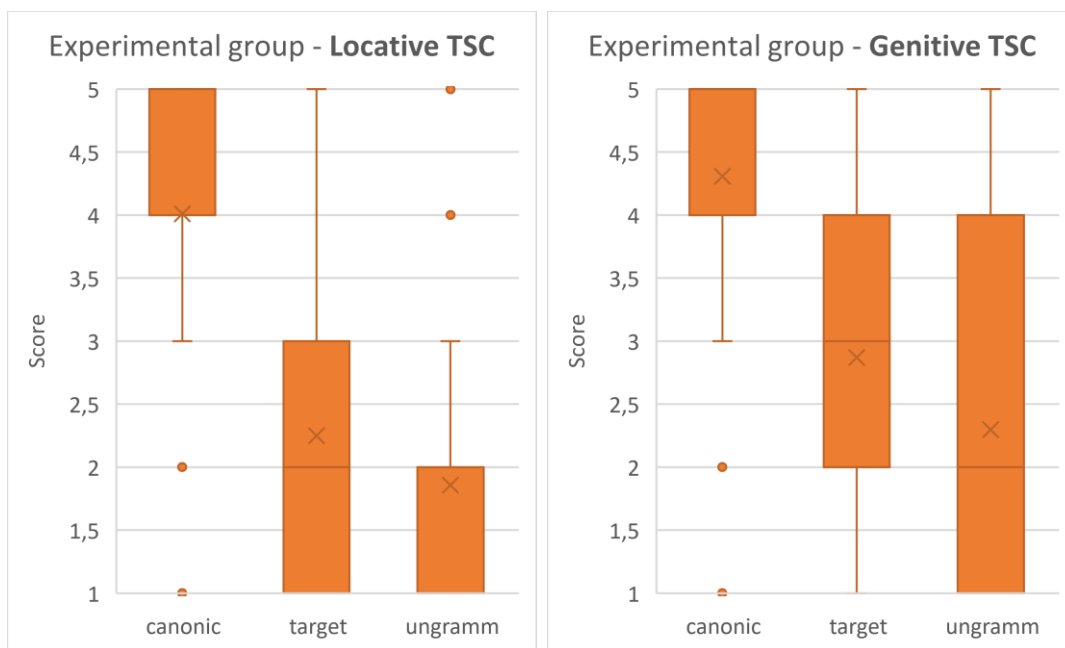
Comparing locative *vs*. genitive topic-subject constructions within groups, using non-parametric versions of two-way ANOVA tests, only the experimental group had a significatively different behavior in their rating of the two types ($H_{1,546}$=10.17; p=0.001). But again, this was only considering the scores as a whole because there was no interaction ($H_{2,546}$=1.50; p=0.47) between the type of topic-subject construction (locative *vs*. genitive) and the structure of the sentence (canonical, target and ungrammatical), which means the differences perceived graphically were only apparent. For the control group, neither the type of construction ($H_{1,582}$=0.729; p=0.39) nor the interaction between type and structure ($H_{2,582}$=4.452; p=0.11) were significative.

The study of this pair of conditions, locative vs. genitive TSCs, for the different proficiency subgroups did not provide any more information than that provided considering the experimental group as a whole. The three subgroups showed a slight preference for genitive TSCs in the boxplot graphs . Nonetheless, the inferential statistic tests confirmed these differences were not significant.

5.2.3. *Wh in situ* vs. *TSCs: a comparative discussion*

To summarize the results discussion so far, I want to directly compare the AJT scores given to target structures of topic-subject constructions and *in situ* interrogatives both by the experimental subgroups and by the control group. Note that this is not new information, it has been previously seen and discussed above, but here the scores for these two type of sentences

are isolated and presented directly one against the other to focus the attention towards this comparison.

I expected to find a different behavior for the experimental and the control groups with respect to the scores attributed to target topic-subject constructions, but not to *in situ* interrogatives. This is true for interrogatives, the experimental subgroups and control group behave very similarly, although the intermediate subgroup mean is lower. However, only the native-like experimental subgroup shows a different behavior regarding topic-subject constructions, which received much lower ratings from this group. The other two experimental subgroups rated these constructions almost like the control group. Their ratings were lower and more variable than those for *in situ* interrogatives, but just in the same way they were for the control group.

This can be seen in the graphs in (12) and it is confirmed by statistical tests. Both type of sentence and group are significative factors in this case ($H_{1,999}=215.46$; p~0 and $H_{3,999}=25.95$; $p<10^{-5}$), and there is interaction between the two factors ($H_{3,999}=11.29$; p=0.01). A post-hoc test corroborates that, for each of the groups, the two types of sentences are significatively different; that, for topic-subject sentences, only the native-like group (C) differs from the rest, and that, for *in situ* interrogatives, only the intermediate group (A) rated them differently.

(12)   Boxplots for scores in the AJT – Topic-subject constructions vs. *in situ* interrogatives
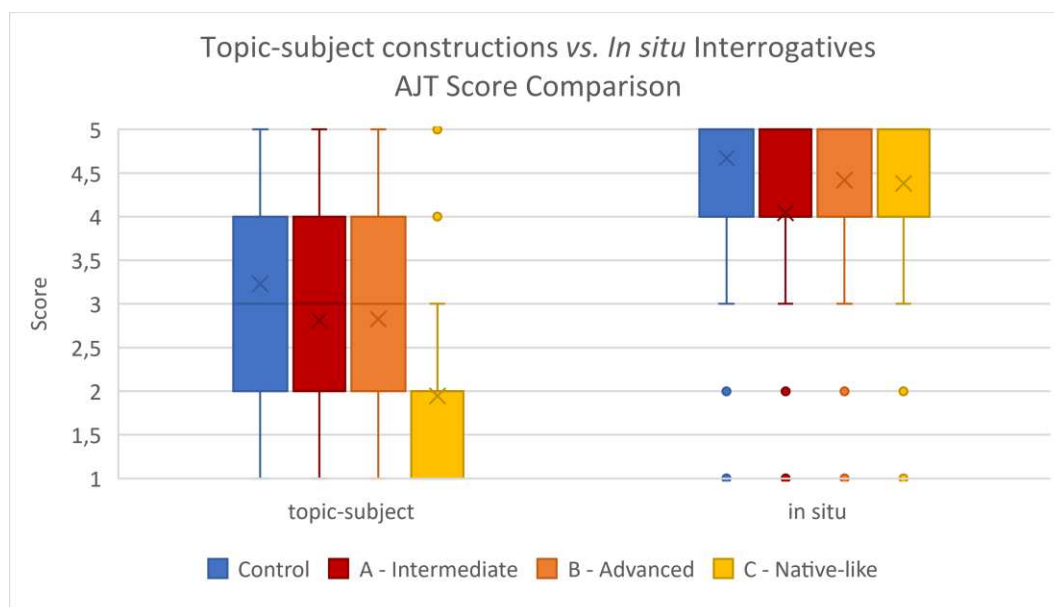
I will come back to the theoretical interpretation of this results in Chapter 7, but, although they do not meet exactly all my expectations, they point in the same direction as the hypothesis being tested (which was introduced in Chapter 4, (1)).

5.2.4. <u>Reaction times</u>

The analysis of reaction times shows that the experimental and control groups behaved quite similarly. Both the mean and standard deviation for each type of sentence are lower for the control group, as can be seen numerically in Table 7 and graphically in the boxplots in (13). Nonetheless, the differences type by type between the two groups did not prove to be significant.

Group	Topic-subject constructions			Interrogatives	
	Canonical	Target	Ungramm.	Fronted	In situ
Experimental	**5.35**	**5.56**	**5.24**	**4.02**	**3.67**
	(2.33)	(2.48)	(2.47)	(2.05)	(2.15)
Control	**4.04**	**4.24**	**3.93**	**2.83**	**2.73**
	(2.61)	(2.60)	(2.02)	(1.93)	(2.38)

Table 7 – Mean reaction times in seconds (Standard Deviation in parentheses)

Applying inferential statistics (again a non-parametric version of a two-way ANOVA), both the group and the type of sentence are significative factors for the reaction times ($H_{1,2373}=183.36$; $p\sim0$ and $H_{4,2373}=205.03$; $p\sim0$, respectively). So, on one hand, we can say that the experimental participants took significatively more time to respond, in general, than the controls. This is exactly as expected, non-native speakers take longer than native speakers.

On the other hand, what these results tell us is that there were some differences between sentence types considering the reaction times of both groups together. A post-hoc test revealed that the reaction times for topic-subject constructions (in any of its versions: canonical, target or ungrammatical) were significatively bigger than those for fronted and *in situ* interrogatives, which also differed significatively from one another.

The fact that fronted interrogatives took significatively more time to process than *in situ* ones was also expected, from what is discussed in psycholinguistic literature. (The wh element needs to be kept in the working memory until the place where it should be interpreted appears, it is known as a filler-gap dependency). What is most interesting to us, though, is the fact that topic-subject constructions required more processing time than interrogatives, in both groups. But such a comparison between conditions may be problematic, I will come back to this later.

However, there was no interaction between the two factors, group and sentence type ($H_{4,2373}$=0.70; p=0.95). This means that comparing type by type the experimental and the control groups behave in the same way. In other words, although the experimental group took more time to respond than the control group, their relative behavior regarding the different sentence structures was the same.

(13) Boxplot graphs for Reaction times sorted by group and type of sentence
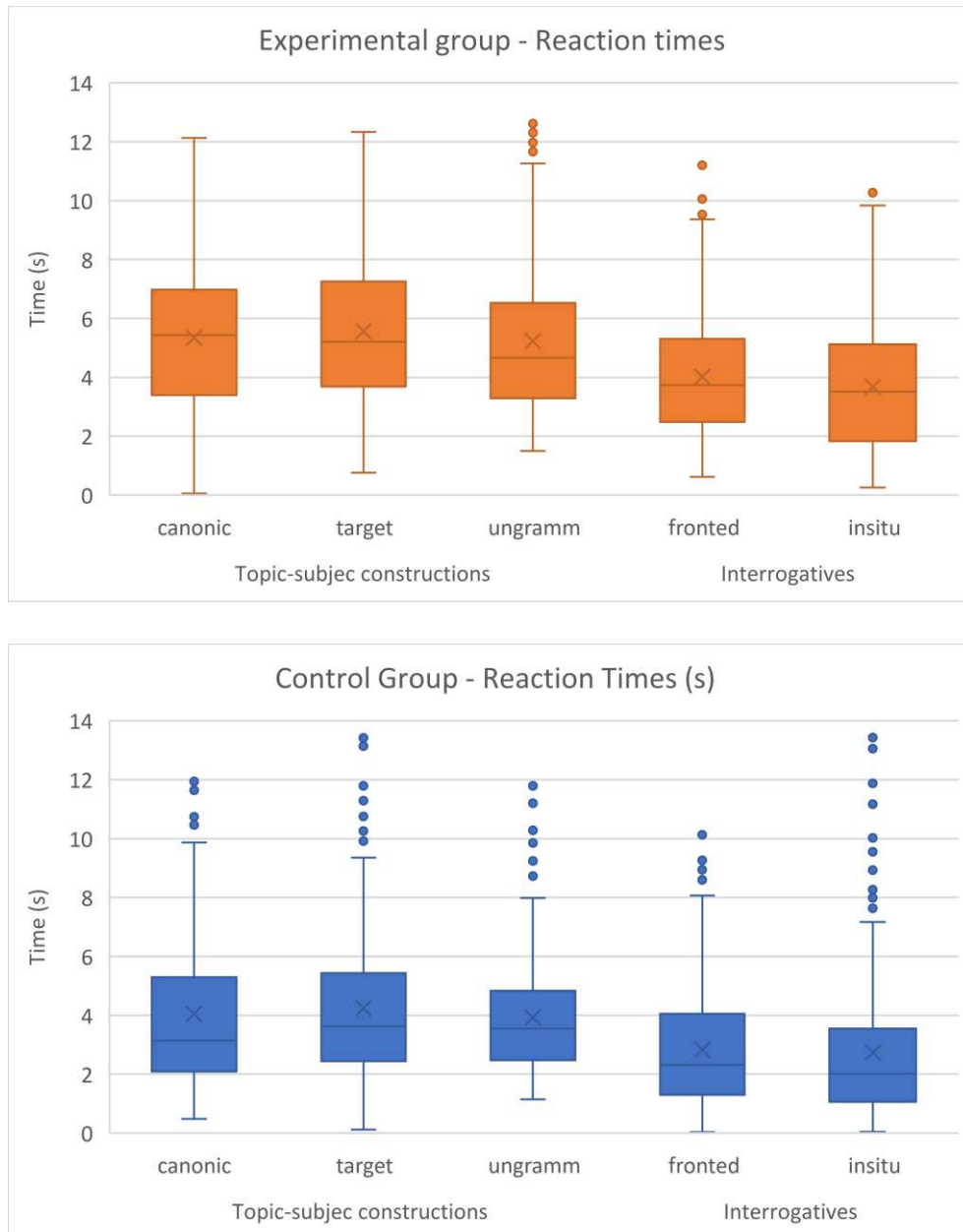
Again, division of the experimental group in its subgroups by proficiency level did not provide any additional information. Reaction times slightly decreased with proficiency, but the statistical tests showed the differences were not significant.

In Chapter 6, I will reflect upon the measurement of reaction times in this experimental study, from a purely empirical point of view. Whereas in chapter 7, I will resume the discussion on the possible theoretical implications of these results.

5.3. Individual Results

Doing experimental research, it is reasonable to assume some homogeneity within groups, however individuals always show idiosyncrasies. Let us now analyze if some of the individual behaviors may have had an impact on the overall results.

More specifically, I want to investigate whether the data dispersion seen in the previous section is due to individually erratic behavior of the participants or to different people (within the same group) behaving consistently as individuals but differently from one another.

5.3.1. *Consistency levels*

In order to have a panoramic view of the individual results, I designed a way of measuring (in a scale from 0 to 3) how consistent were the participants' responses for each of the five kind of sentence structures present in the experimental study. The responses were considered to be consistent when the same type of sentence was treated in the same way.

I looked at the scores given by each participant to each set of experimental items. If 85% or more of the experimental items in a set received scores that leaned towards the same direction, that is, were either accepted (received a score of 4 or 5) or rejected (received a score of 1 or 2), the consistency level for this set was set to be 3, the maximum. If the rate of acceptance/rejection was between 75% and 85%, the consistency level was set to 2; for a rate between 60% and 75% the consistency level was 1; and if only 60% or less of the experimental items in a set were either accepted or rejected, the consistency level was set to be 0, the minimum.

	Rate of acceptance / rejection			
	< 60%	60% - 75%	75% - 85%	> 85%
Consistency Level	0	1	2	3

Table 8 – Consistency levels

For example, if one participant gave scores of 4 or 5 to 5 out of the 8 items in the target TSCs set, this means 62.5% of the sentences in the set were accepted, so the consistency level would

be 1. If he/she gave scores of 1 or 2 to 10 out of the 11 items in the *in situ* interrogatives set, then a 90.9% of the sentences in the set were rejected, so the consistency level for this set would be 3.

Following this method, each participant received 5 consistency ratings, one for each type of sentence. This way, it is possible to study the consistency by type of sentence and compare them between groups to see if a specific group was more or less consistent than other regarding a given set of sentences.

Below, we can see a boxplot graph of the consistency ratings for each set of sentences divided by group. The higher the value is, the more consistent the participants in each group are (although there may still be participants consistently behaving in different way from one another!). The lower the value, the less consistent the participants are. But it is also important to consider the size of the box. The bigger the box, the more diversity between participants (some may be consistent while others may be inconsistent).

(14)   Boxplot graph for Consistency levels sorted by group and type of sentence
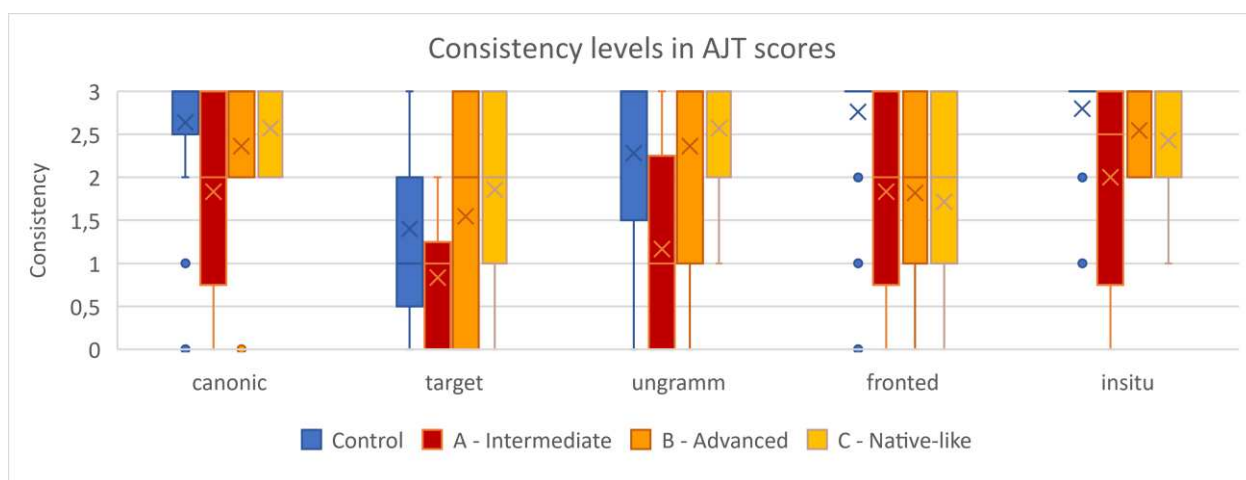
This graph shows that the participants in the control group are clearly more consistent than those in the experimental groups for all 5 sets of sentences (with the exception of ungrammatical TSCs, for which the native-like participants were slightly more consistent). It also shows that control participants are extremely consistent for interrogatives and for canonical versions of TSCs, but a bit less consistent for ungrammatical TSCs and even less so for target TSCs.

Within the experimental groups, native-like participants are quite consistent regarding canonical and ungrammatical TSCs, and also *in situ* interrogatives, but their consistency level decreases for target TSCs and fronted interrogatives. The advanced participants' consistency is

similar to that of native-like participants, but there is more diversity for ungrammatical and especially target TSCs. This box is the bigger of them all, which means the participants in this group were the most diverse, some were quite consistent but other were quite inconsistent. Finally, participants in the intermediate group are the least consistent in general, especially for target TSCs. There is also a lot of diversity within this group.

Looking at the type of sentence, what we see is that target TSCs were those rated most inconsistently and also that there is more diversity within each group regarding this type of sentence. This means, on one hand, that there were more individuals that both accepted and rejected sentences of this type, and on the other hand, that there were more differences in consistency between individuals of a given group regarding these sentences.

5.3.2. *Control group*

Let us now look at the individual results for the control group. Obviously, even if this group was more consistent than the experimental groups, as we saw in the previous section, there were differences between participants. Some were more assertive in their responses, others more voluble; some were more permissive and gave higher scores, some were very strict and gave lower scores to all types of sentences. Nonetheless, broadly speaking, with more or less data dispersion, higher or lower scores in general, the individual graphs show a relative behavior between the different types of structures very similar to what we have seen for the whole group. This can be seen in the following graph, which corresponds to what can be considered the prototypical control participant.

(15)   Boxplot graph for individual AJT scores – ER (prototypical control behavior)
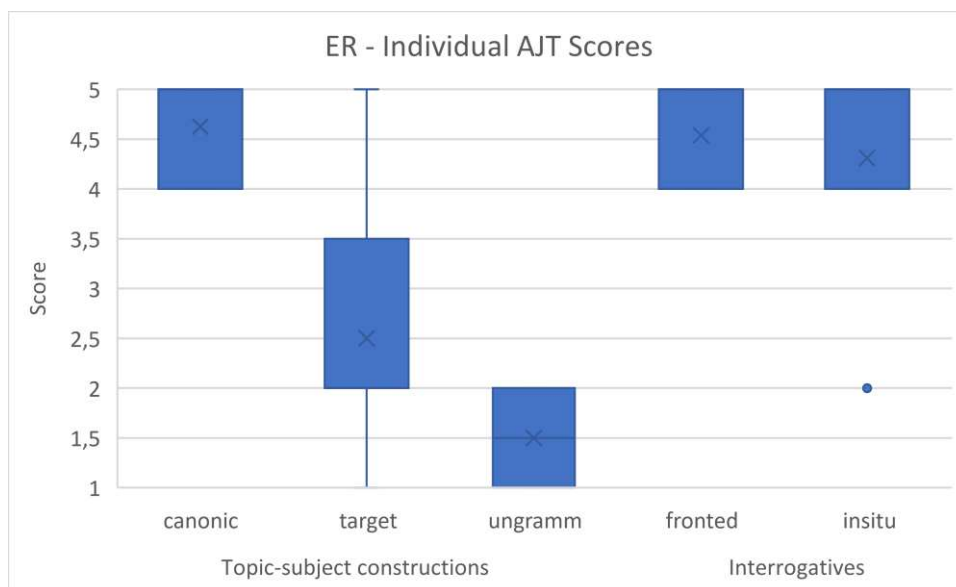
In this case, it was not reasonable to run two-way inferential tests to study the interaction between the type of sentence and the participant, since post-hoc test would give us a 4-digit number of combinations. So, the analysis was focused only on the participant factor.

A Kruskal-Wallis test (a non-parametric version of a one-way ANOVA) confirmed that the participant was a significative factor for the AJT scores of this group (H_{24}=126.61, p<10^{-15}). A Wilcoxon post-hoc test revealed which pairs of participants behaved in a significatively different manner. Participant ER, who I just mentioned as the prototypical control participant, had a behavior that did not differ from any of the other participants in this group. There were 300 different pairs, of which 32 (around 10%) behaved significatively differently from one another. I will not list all of them, but I will show some examples of extreme behaviors below.

The most "discordant" participants had behaviors that differed with the same 8 or 9 colleagues from the control group, whose behavior was quite similar between them. Below, you can see AJT scores by sentence for two of the most discordant participants (IB and MC, in darker blue) and compare them with the AJT scores of one of those 8 or 9 participants whose behavior was representative of this subgroup (JL, in lighter blue).

(16)   Boxplot graph for individual AJT scores – IB and MC *vs*. JL
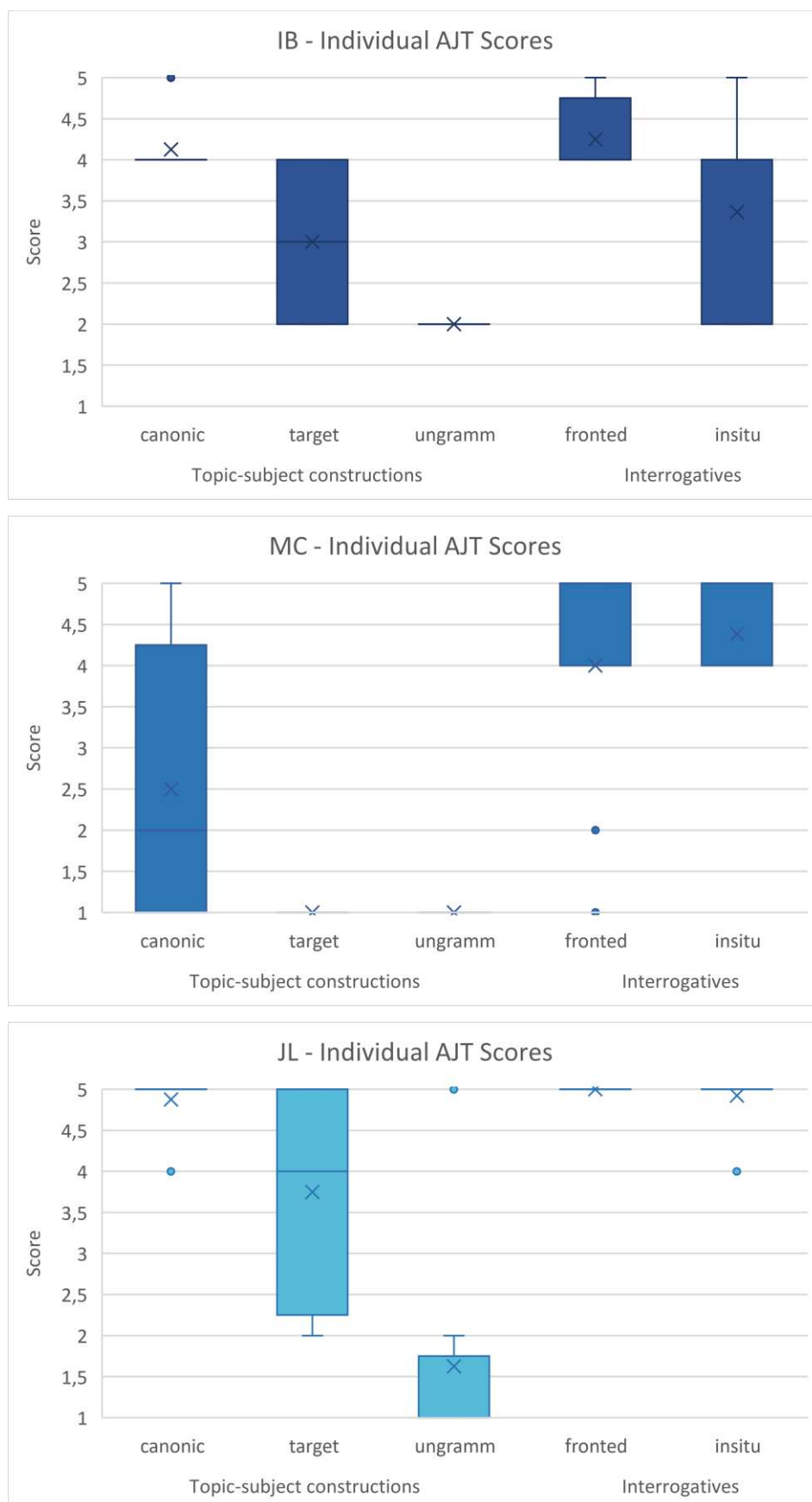
Note that MC clearly rejected both target and ungrammatical versions of TSCs and gave low ratings to some of the canonical test sentences as well. Although there were other participants who did not like target TSCs very much either, none was as categorical as MC.

On the other side of the spectrum, we have GV, who gave higher scores to target TSCs than to their canonical versions. No other control participant behaved exactly like this, but there were only 3 other participants in this group who mostly accepted these constructions. Most of them did not show a clear preference, they accepted some of them and rejected others.

(17) Boxplot graph for individual AJT scores – GV
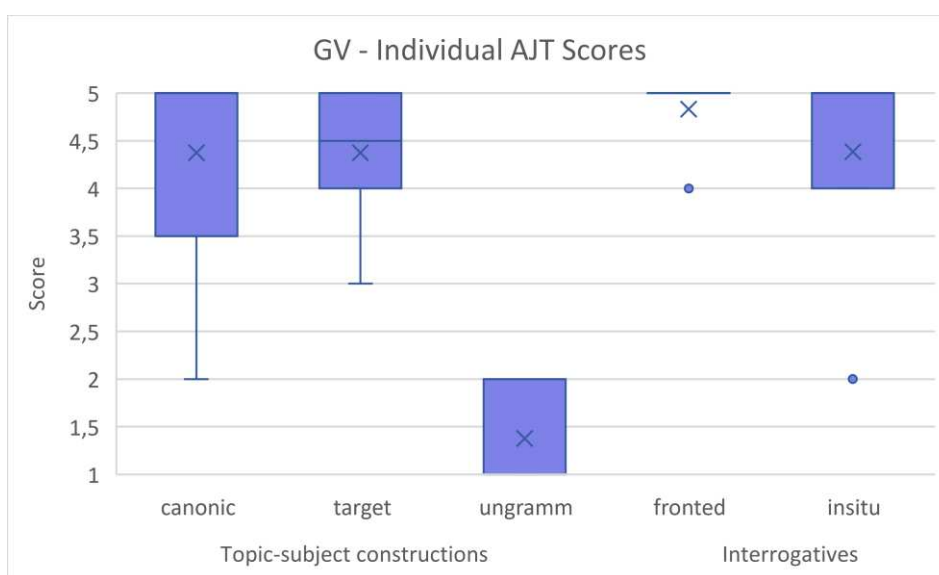
As we saw in the previous section, the participants in this group were quite consistent about the other types of sentences, especially the interrogatives. So, I will not go into more detail here.

5.3.3. *Experimental group*

Within the experimental group, the participants behaved in a more heterogeneous manner than what we just saw for the control group. Nonetheless it is not hard to find within this group participants that rated the five type of sentences in the same way our prototypical control participant did, see for example how MT behaved, a participant of the advanced experimental group.

(18)   Boxplot graph for individual AJT scores – MT (advanced group)
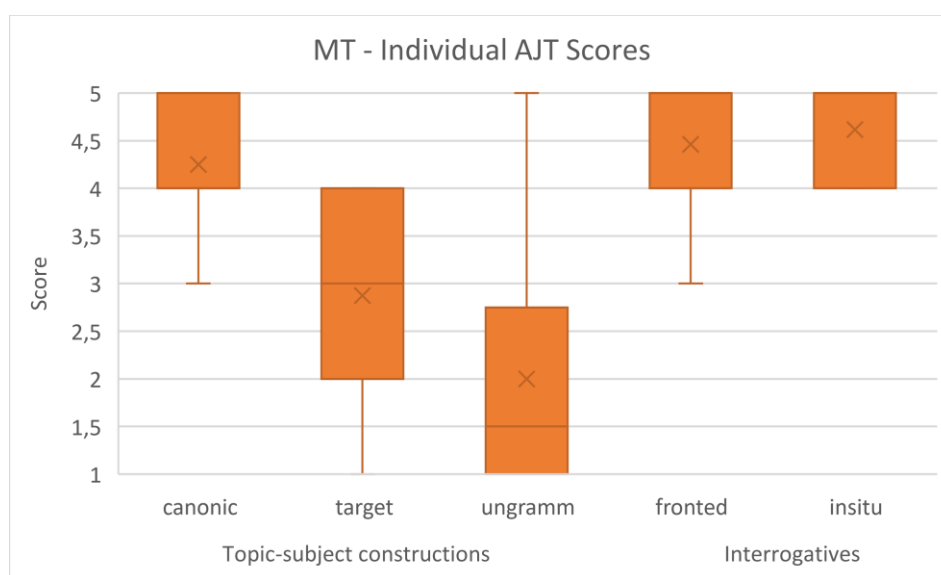
As discussed before, during the analysis of the consistency levels, the experimental group was less consistent, and the intermediate subgroup especially so. This is evident if we look at the individual behavior of some of these participants. MN below, who belongs to that subgroup, is the perfect example of this.

(19)   Boxplot graph for individual AJT scores – MN (intermediate group)
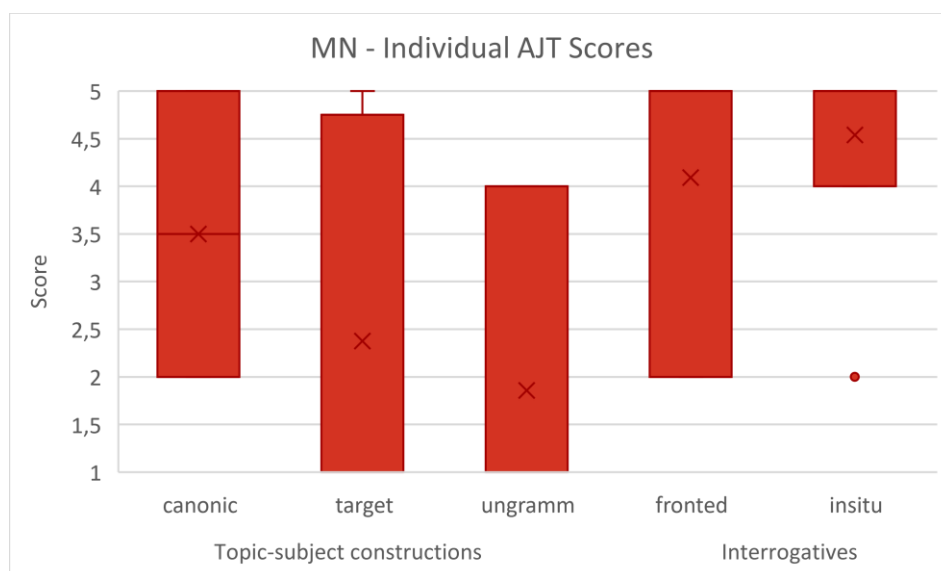
However, there were participants much more consistent in their responses, like EL, from the native-like group. Leaving the consistency of this participant aside (which is not representative), these were the kind of ratings I expected to find in the experimental group.

(20) Boxplot graph for individual AJT scores – EL (native-like group)
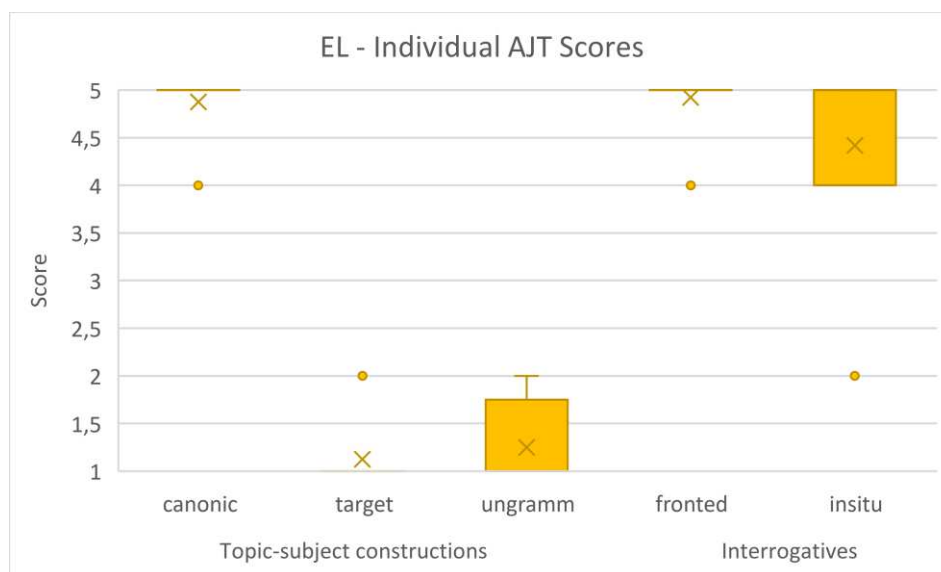
Like for the control group, applying a two-way test (type of sentence *vs*. participant) was unreasonable, since it would be impossible to check the number of possible combinations in the post-hoc test. So, again, a Kruskal-Wallis test was used instead, and the participant factor proved to be significative (H_{23}=78.746, $p<10^{-7}$). A Wilcoxon post-hoc test revealed that out of the 276 possible pairs of participants, there were 16 that differed significatively between them (only 6% of all the pairs). I will go back to this in a moment.

Since the experimental group was divided in subgroups by proficiency level, I ran the tests in each subgroup. For the intermediate group, a Kruskal-Wallis test revealed that the participant was not a significative factor (H_5=7.22, p=0.2), which means all the participants in this group behaved in a similar way. The main characteristic of this group is their data dispersion, as could be seen in MN's scores, see graph (19) above, which is a representative example. Only one of the participants in the intermediate group was more consistent in his answers, and he had the highest proficiency level in his group.

The advanced experimental participants were the most heterogeneous subgroup. A non-parametric version of a one-way ANOVA test confirmed that the participant was a significative factor (H_{10}=64.7, $p<10^{-9}$). In this case, there were 55 possible pairs of participants, of which 15 (27% of the total) behaved in a significatively different manner, according to a Wilcoxon post-hoc test. The test revealed there were three type of participants in this group: one that was more strict and gave lower scores in general, one that was more permissive and mostly accepted all types of sentences (but was confused about ungrammatical TSCs), and a third type of

participant who fell in the middle of those behaviors. This difference had no correlation with the participant's proficiency level. The graphs in (21) are examples of the first and second type of behavior (participant JN and MG, respectively). The graph seen in (18) corresponds to a participant of the third type, MT. All but one of the 15 pairs of participants that differed significatively in their responses were formed by a participant of the first and one of the second type.

(21) Boxplot graph for individual AJT scores – JN and MG (advanced group – strict *vs*. permissive)
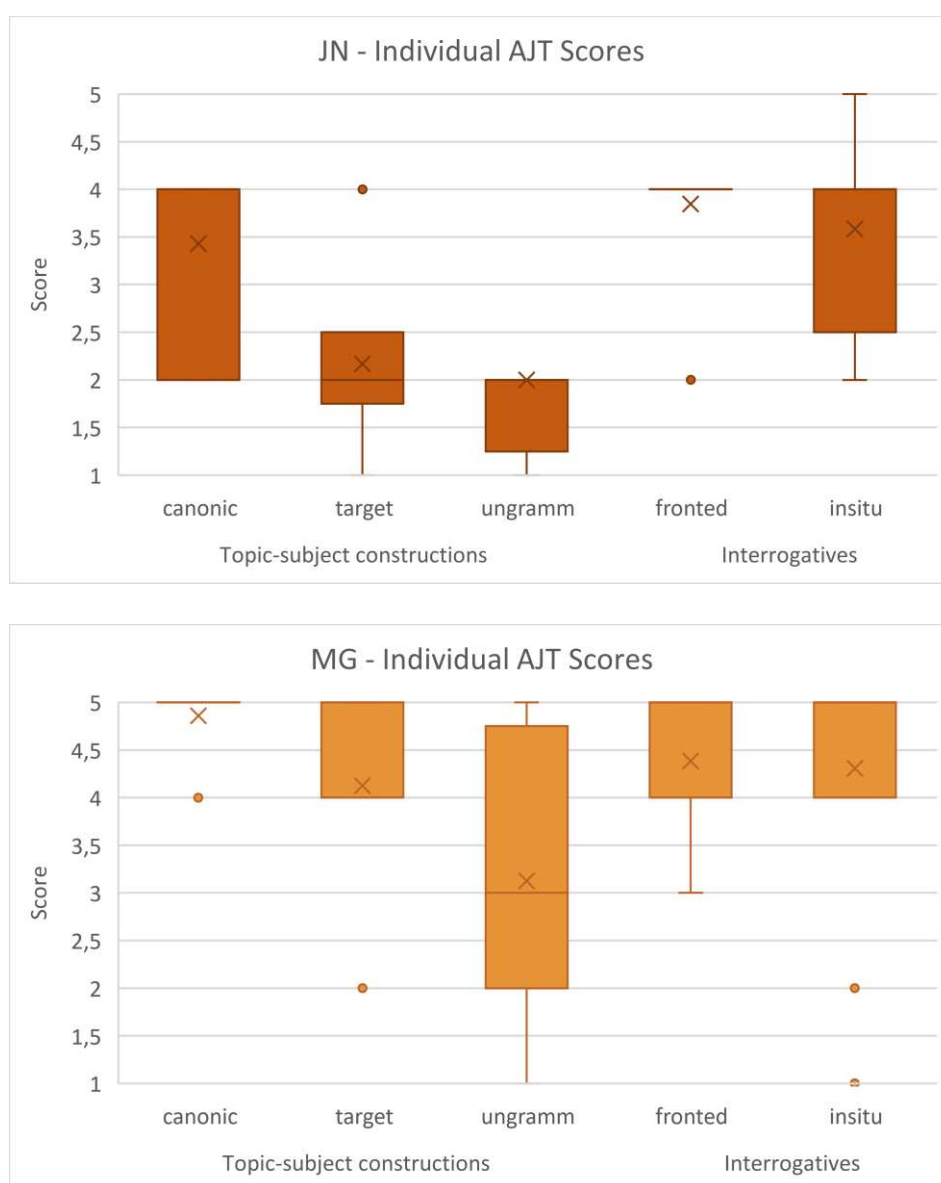
For the native-like experimental subgroup, the participant was not a significative factor (H_6=8.84, p=0.18). The graph below is an example of a participant from this subgroup. As we saw in the analysis of the group results, what is most characteristic of these participants is that

they mostly rejected target TSCs. This was evident in the example (20) above, and it is visible in the following example, too.

(22)   Boxplot graph for individual AJT scores – LL (native-like subgroup)
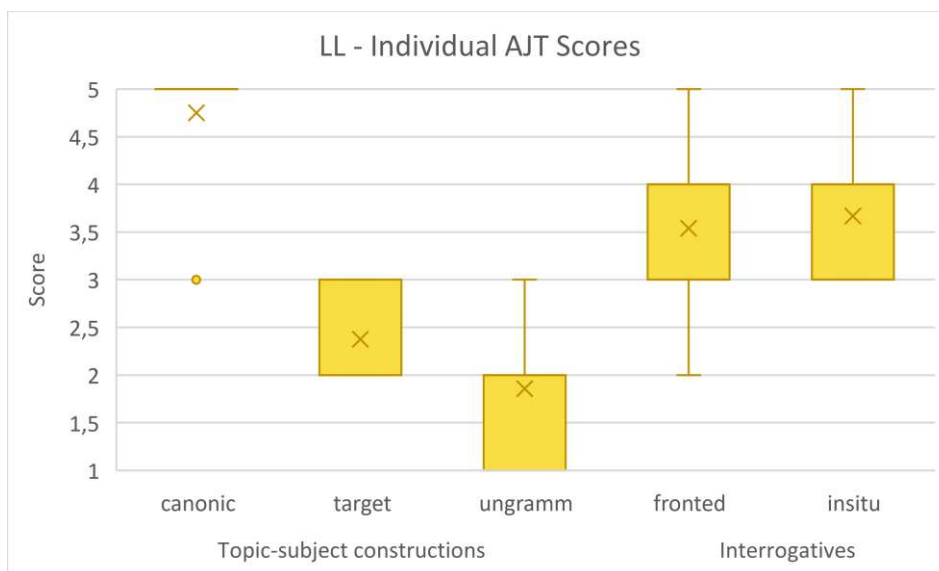
Going back to the results of the test applied to the whole experimental group, we discover that of the 16 pairs of participants with significatively different scores, all of them contained at least one participant of the advanced subgroup. Twelve of them were formed only by participants in this subgroup; three of them contained also a participant from the intermediate subgroup; and the last one contained a participant from the native-like subgroup, LL, whose scores could be seen in (22).

Regarding interrogative sentences, as we saw in the group results analysis, the individual results show that the participants in the experimental group generally accepted both fronted and *in situ* interrogatives. Whenever the individual graphs show lower scores for these sentences, they were due to double complementizers (in fronted interrogatives), or interrogatives with postverbal subject (which, unlike what was done for the group results, were not excluded from the individual analysis).

Regarding TSCs, what we see in this group is summarized as follows. All native-like participants and half of the participants in the advanced group mostly reject both target and ungrammatical TSCs. Only 3 participants (all from the advanced group) mostly accepted target TSCs and was confused about their ungrammatical versions. And the rest of the participants in the experimental group were not quite sure about these structures, either giving them middle ratings or accepting some while rejecting others.

5.3.4. Bilingual group

Recall, from section 4.3, that besides the participants in the experimental and control groups, there were 4 additional participants who happened to be bilingual in Spanish and Brazilian Portuguese. This fact excluded them from any of the other two groups, and their reduced number (plus the heterogeneity of their backgrounds) precluded them to form a third group. I will, however, briefly discuss their individual results here.

Within this group, there were two different types of behavior. Two of the bilingual participants scored 47 out of 50 in the proficiency test, they were stricter with TSCs and more consistent in their responses in general. The other two scored 43 in the proficiency test, they showed less consistency and were a bit more unsure about TSCs, although they rejected more than accepted these structures. None of them showed any problems accepting *in situ* interrogatives. The following graphs correspond to a participant with the first kind of behavior, CB, and a participant with the second kind of behavior, AF.

(23) Boxplot graph for individual AJT scores – CB and AF (bilinguals)

AF - Individual AJT Scores
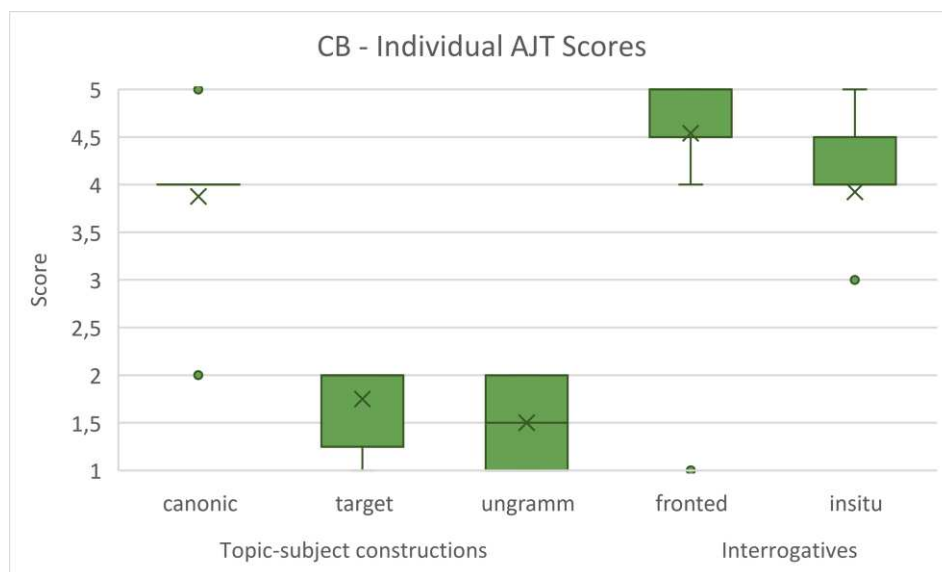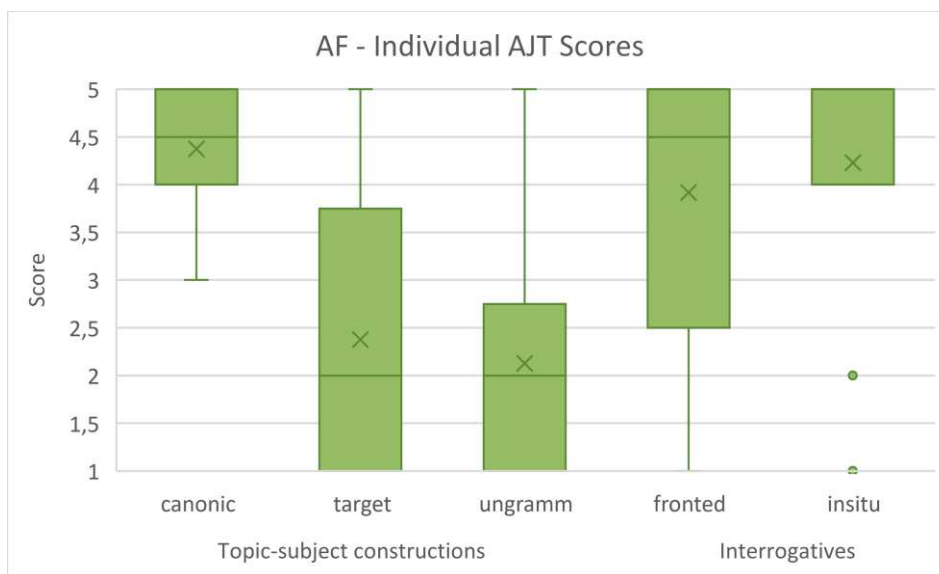
5.4. Follow-up: qualitative analysis

So far, the experimental results have been analyzed quantitatively. But what happened with the rejected sentences? Why were they rejected? What was the element in them that bothered the participants? That is exactly what I want to discuss in this section, firstly, because this, *per se*, provides interesting qualitative information, but also because it affects the quantitative analysis and may influence the results.

5.4.1. *Problematic cases*

As mentioned in chapter 4 (section 4.4), the main purpose of the follow-up task used in this experiment was to confirm that the rejected sentences were deemed unacceptable by the participants for the reasons that interest us, and not because of some other aspect of their syntactic structure, or because of a specific lexical choice. Whenever the follow-up task shows a sentence was rejected for a reason other than the syntactic structures under study, or whenever the reason for rejection is unclear, I will refer to those cases as "problematic cases".

Methodologically, it is crucial to take these cases into consideration, whether their impact on the quantitative results proves to be relevant or not. It is unreasonable not to, since they would be treated in the analysis just as ordinary cases (in which the cause of a sentence's rejection is exactly what is being studied). But now, thanks to the follow-up, we know for a fact that they

are not ordinary cases, and to treat them as such would be deliberately adding noise to our measurements.

There are three different scenarios for the problematic cases. Let me examine them, present some examples, and discuss how they should be treated.

In the first possible scenario, the rewritten sentence maintains the syntactic structure of interest (the topic-subject construction or the *in situ* question) and it is some other element of the sentence to be altered by the participant. This shows that the reason of rejection was not what it "should" be, and it suggests that, had we altered that other element bothering the participant, he/she may have had a positive response to the sentence.

(24) Examples of problematic cases from the experimental group

Experimental item	Follow-up modification
i. *Do que **se** trata o livro?*	*Do que _ trata o livro?*
of.the what **SE** revolves.around the book	of.the what _ revolves.around the book
'What is the book about?'	'What is the book about?'
ii. *Os alunos vão se reunir **onde**?*	*Os alunos vão se reunir **aonde**?*
the students go SE meet **where**	the students go SE meet **to.where**
'The students are going to meet where?'	'The students are going to meet where?'
iii. *Nos bairros da periferia **já** não passa taxi.*	*Nos bairros da periferia não passa **mais** taxi.*
in.the neighbourhoods of.the periphery **anymore** not passes taxi	in.the neighbourhoods of.the periphery not passes **more** taxi
'Taxis don't go to the outskirts of the city anymore.'	'Taxis don't go anymore to the outskirts of the city.'
iv. *Os livros **às vezes** enganam as capas.*	***Às vezes** os livros enganam as capas.*
the books **at.the times** lie the covers	**at.the times** the books lie the covers
'Book covers sometimes lie.'	'Sometimes book covers lie.'

Note, however, that I said it *may* have had a positive response. There is no certainty of this, which is why, although it may be tempting, we cannot treat these cases as we would those

deemed acceptable. Not to mention this would be equivalent to altering the participant responses, something unethical and completely out of the question.

Even though these are the most interesting cases, their incidence is negligible, since it only represents 1.5% of all our data.

In the second scenario, the sentence was rewritten exactly as it was originally presented or it was substituted by an "OK". This means the participant "changed his/her mind" and what seemed unacceptable at first, now seems alright on second thought. This can be interpreted as the first scenario, meaning that the response could have been positive. However, the fact that the first impulse was to reject the sentence cannot be ignored. First of all, it is probable that the first, more intuitive response to an experimental item is precisely the one that best reflects the participant's internal grammar. But also, it may have been the exposure to other sentences with similar syntactic structures during the experiment what made the participant reevaluate and accept now what before seemed wrong for some reason. In other words, the following sentences in the acceptability judgement test may have had a priming effect on the follow-up task. We can never know for sure why these sentences were rejected, but the fact remains that they were, and this must be taken into account.

Finally, in the third scenario, the follow-up response for a specific item was left blank. This is harder to interpret, since it may mean the same as the previous cases, but it may also mean the participant did not know how to fix the sentence to make it acceptable. Whenever participants directly asked me what to do if they changed their mind, I would tell them to rewrite the sentence as it was, so that I could know that is what happened. However, not everybody asked me, and I did not want to offer them this possibility from the start, so as not to influence their answers. Therefore, they may have opted for leaving the follow-up item blank as a way of saying they found nothing wrong with the sentence. This is a possibility, though I believe it is more likely that they were not able to find a successful way to rewrite the sentence (especially within the members of the experimental group). Again, there is no way of knowing for sure.

In any of these three possible scenarios, the only reasonable way to treat a problematic case is to simply disregard it. The sentence's score in the AJT either reflects the participant's grammatical intuition on a completely different matter (first scenario), or it provides unreliable information (second and third scenarios), and, therefore, it should be removed from the statistical analysis.

Although, as I have been arguing, the removal of these cases is methodologically very important, their incidence is less than 3% of all our data. This means their impact on the quantitative analysis, both descriptive and inferential, but especially the latter, is quite small, if present at all. In fact, I carried out all the analysis in section 5.2 both with and without the problematic cases and there were no significative differences in means, standard deviations and boxplot graphs and all the results of the statistical tests remained the same.

This is true, from a global perspective. If we look more closely, though, the problematic cases start to gain more relevance. As I just said, they represent 3% of all our data; however, if we consider only the number of rejected sentences, their incidence rises to 7% for the control group, and 12% for the experimental group, which does not feel as negligible as it did before.

An even closer look, focusing now on the individual results, reveals the chosen treatment for problematic cases can considerably change the boxplot graph of the scores for the acceptability judgement task. The following example shows the same graph before and after removing the scores in three problematic cases for one specific participant, CR. One of the problematic cases corresponds to the third sentence in the previous example, the other two were sentences rewritten exactly as the experimental item (second scenario).

(25) Boxplot graphs for CR individual AJT scores <u>before</u> and <u>after</u> removal of problematic cases

CR - Individual AJT Scores - **After**
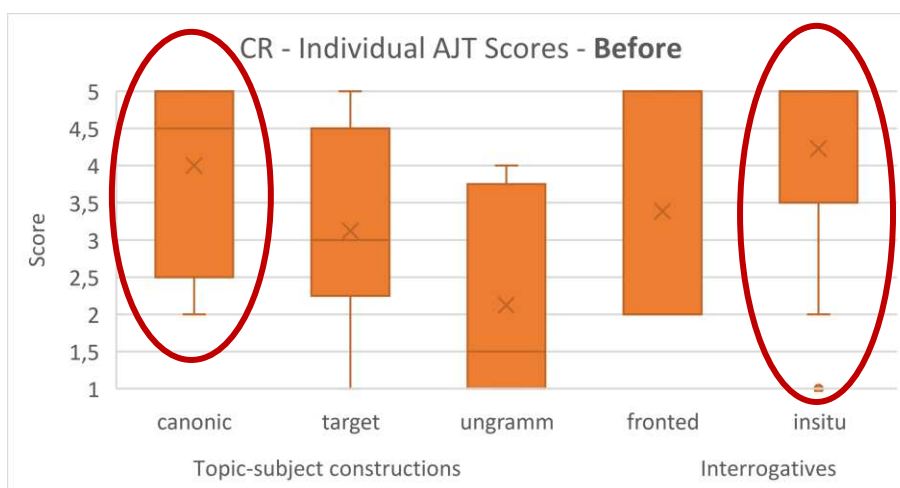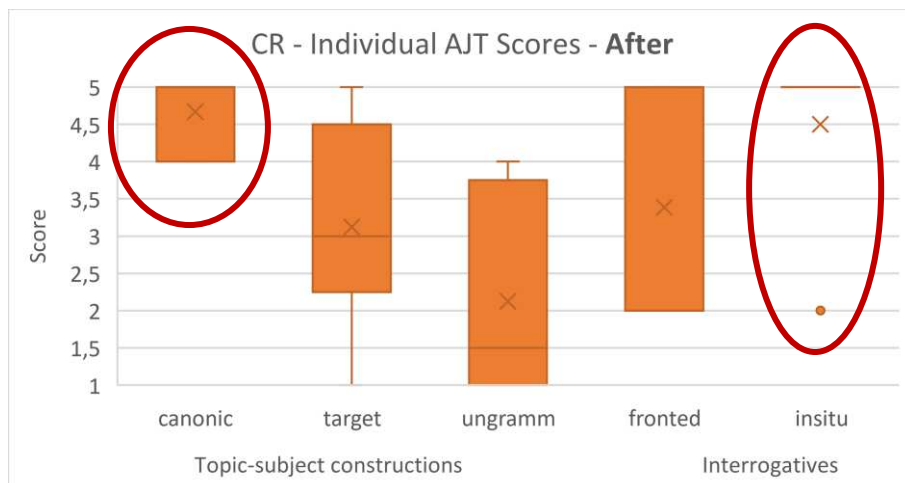
Such a difference should not be a surprise, since, in these specific individual results, the removed scores represented 2 out of 8 in the canonic structure set, and 1 out of 13 in the *in situ* set.

Two important conclusions can be drawn from this. First, it confirms that the treatment of problematic cases does matter, and not only from a methodological point of view. But second, it calls our attention to the reliability of the individual boxplot graphs. These graphs are quite useful, they provide a quick visual summary of each participant's scores, which would otherwise be hard to visualize. But they should be interpreted carefully, because unlike the same kind of graphs used for the group results, the individual ones are made with only 8 items for each of the first three boxes, and 13 for each of the other two. Since their outer aspect is the same, it is easy to consider the information portrayed in them as being as solid and robust as the information portrayed in the group graphs. This is not the case.

Now that the problematic cases have been discussed and dealt with, let me go on to the ordinary cases and show the specific strategies used by the participants.

5.4.2. *Interrogative sentences*

From the quantitative analysis, we know that interrogative sentences, whether fronted or *in situ*, generally received high scores from both groups. Here we have the AJT mean scores table again, just as a reminder.

Group		Fronted		In situ
Experimental	Mean	**4.12**	Mean	**4.31**
	SD	(1.11)	SD	(0.95)
Control	Mean	**4.62**	Mean	**4.67**
	SD	(0.77)	SD	(0.64)

Table 3 – Interrogatives Mean Scores (AJT)

Fronted interrogatives

The experimental group rejected 17.3% of the fronted interrogatives. As suspected, the main reason for rejection was the presence of a complementizer beside the interrogative pronoun (a doubly filled Comp), to which the participants in this group responded either by eliminating the complementizer or by adding a copula between this and the pronoun to turn the sentence into a cleft, as seen in the following example (26.i). In addition to this removal, the participants sometimes opted for postposing the subject or substituting it for a null subject, two additional strategies that reflect their Spanish grammars (26.ii).

However, interrogatives without a doubly filled Comp were also rejected, although not as much. In this case, some participants simply added the complementizer or a copula as well to create a cleft (26.iii). There was a particularly interesting case here, in which a fronted interrogative was modified to become an *in situ* interrogative (26.iv), this modification occurred twice within the bilingual group.

(26) Examples of amending strategies for fronted interrogative sentences in the experimental group

    <u>Experimental item</u>   <u>Follow-up modification</u>

i.   *O **que que** o Pedro vai comprar?*   a) *O **que** _ o Pedro vai comprar?*
     the **what that** the Pedro goes to.buy   the **what** the Pedro goes to.buy
     'What is Pedro going to buy?'   'What is Pedro going to buy?'

     b) *O **que é que** o Pedro vai comprar?*
     the **what is that** the Pedro goes to.buy
     'What is it that Pedro is going to buy?'

ii.   *Do **que que** <u>eles</u> precisam?*   a) *Do **que** _ precisam <u>eles</u>?*
     of.the **what that** they need   of.the **what** need they
     'What do they need?'   'What do they need?'

b) *Do **que** _ precisam _?*

   of.the **what** need

   'What do (they) need?'

iii. *Do **que** <u>se</u> trata o livro?*

   of.the **what** <u>SE</u> revolves.around the book

   'What is the book about?'

a) *Do **que que** trata<u>se</u> o livro?*

   of.the **what that** revolves.around.<u>SE</u> the book

   'What is the book about?

b) *Do **que é que** _ trata o livro?*

   of.the **what is that** revolves.around the book

   'What is it that the book is about?'

iv. ***Para quem** o Pedro enviou aquele pacote?*

   to whom the Pedro sent that package

   'To whom did Pedro send that package?'

*O Pedro enviou aquele pacote **para quem**?*

the Pedro sent that package **to whom**

'Pedro sent that package to whom?'

Apart from these, there were also some problematic cases: a few sentences were left unchanged, and others suffered minor changes that did not alter the main syntactic structure (addition of a preposition, removal of an article before a proper noun, position of an adverb or removal of a non-passivazing *se* pronoun).

There is one extra case which is, I believe, worth mentioning. One of the participants rejected the sentence *Quem você mais admira?* ('Who do you most admire?') and then altered the order between the interrogative particle and the pronoun in the follow-up task, generating the sentence **Você quem mais admira?* ('(And) you, who do you most admire?'). In Spanish[39] interrogatives, there is subject-verb inversion, unless the subject is focalized, which then must appear in sentence initial position. This topic structure is preferred when the subject is a tonic pronoun. And my guess is that the participant was reproducing this Spanish structure here, since the equivalent one in Brazilian Portuguese is ungrammatical. The only way for the subject to be in sentence initial position in a BP sentence like this is to use an *in situ* interrogative, *Você admira quem?* ('You admire who?'). Otherwise, it is necessary to introduce a topic-comment

---

[39] This is not true for all Spanish dialects. One of the distinguishing characteristics of Cuban Spanish, for example, is the absence of this inversion (*¿Qué tú quieres?* instead of *¿Qué quieres tú?* or *¿Tú qué quieres?* ('What do you want?'))

structure, just like in English, *E você, quem você mais admira?* ('And you, who do you most admire?').

In contrast, the control group rejected only 5.9% of the fronted interrogatives (*vs*. the 17.3% rejected by the experimental group), but the main reason of rejection remained the same, the 'extra' complementizer. In this case, though, the only strategy the control group used to rewrite these sentences was removing it (without the addition of a copula), see the following example (27.i). Only two sentences without this complementizer were rejected. One was left unchanged in the follow-up and the other one was transformed into an *in situ* question, as you can see in (27.ii) (although this particular sentence also contained a word that generated some problematic cases).

It is quite interesting that the lack of a complementizer beside the interrogative pronoun did not bother the participants in the control group but caused some rejections within the experimental group. Interrogative sentences with this doubly filled Comp are extremely frequent in spoken BP[40], which may have caused some of the non-native speakers (I include myself in this group) to adopt this structure as compulsory, resulting in the rejection of interrogatives without it.

The control group showed some problematic cases, but almost all of them were related with a mistake on my part, a Spanish verb (*compartir*, 'to share') that managed to slip into the experimental items. All the follow-up responses substituted this verb for a Portuguese equivalent, as can be seen in (27.iii).

(27) Examples of amending strategies for fronted interrogative sentences in the control group

	Experimental item	Follow-up modification
i.	*Por **que que** a Maria brigou com o Pedro?*	*Por **que** _ a Maria brigou com o Pedro?*
	why that the Maria fought with the Pedro	**why** the Maria fought with the Pedro
	'Why did Maria and Pedro fight?'	'Why did Maria and Pedro fight?'

---

[40] I believe this varies diatopically, not all Brazilian dialects accept interrogatives with the extra complementizer. However, this has not yet been studied (at least not to my knowledge).

  ii. ***Com quem*** *a Maria* <u>*comparte*</u> *o*   A Maria <u>compartilha</u> o quarto **com quem**?

    quarto?   the Maria shares the room **with whom**

    with whom a Maria shares the room   'Maria shares the room with whom?'

    'With whom does Maria share the

    room?'

  iii. *Com quem a Maria* ***comparte*** *o*   *Com quem a Maria* ***divide*** *o quarto?*

    quarto?

In situ interrogatives

The experimental group rejected 10.4% of all the *in situ* interrogative sentences. Apart from a few problematic cases I will shortly discuss, all the sentences were rewritten in the follow-up as fronted interrogatives, see example (28.i). Some of them, though, had additional properties such as a doubly filled Comp (28.ii), or a cleft structure (28.iii), but especially subject-verb inversion, see (28.iii and iv). As I mentioned before, subject-verb inversion is a characteristic of Spanish interrogative sentences that BP does not share, (unless the verb is unaccusative, in which case, subject-verb inversion[41] is acceptable).

(28) Examples of amending strategies for *in situ* interrogative sentences in the experimental

   group

    <u>Experimental item</u>   <u>Follow-up modification</u>

 i. *O acidente aconteceu **como**?*   ***Como** aconteceu o acidente?*

   the accident happened **how**   how happened the accident

   'The accident happened how?'   'How did the accident happen?'

---

[41] Note that, technically, there is no 'inversion' here, since the subject in this case is the verb's internal argument and it did not undergo any movement.

On another note, preference for pre-verbal or post-verbal subjects in BP fronted interrogatives with unaccusative verbs seems to be influenced by other factors. Take, for example, the pairs below:

 (1) i. *Como chegou o pacote?*   (3) i. *Como aconteceu o acidente?*

     how arrived the package   how happened the accident

   ii. *Como o pacote chegou?*   ii. *Como o acidente aconteceu?*

     how the package arrived   how the accident happened

 (2) i. *Como chegou você?*   (4) i. *Quando aconteceu o acidente?*

     how arrived you   when happened the accident

   ii. *Como você chegou?*   ii. *Quando o acidente aconteceu?*

     how you arrived   when the accident happened

Native speakers show a preference for post-verbal subjects in (1) but not in (2), which seems to be related to referentiality or animacy. Or even more subtly, they show a preference for the pre-verbal subject in (3) but show no preference in (4), where the only difference is the wh pronoun.

ii. *O Pedro ama **quem**?*　　　　　*__Quem que__ o Pedro ama?*

　　 the Pedro loves who　　　　　　**who that** the Pedro loves

　　 'Pedro loves who?'　　　　　　'Who does Pedro love?'

iii. *__Eles__ gostam **do que**?*　　　*__Do que é que__ gostam __eles__?*

　　 <u>they</u> like **of.the what**　　　**of.the what is that** like <u>they</u>

　　 'They like what?'　　　　　　'What is it that they like?'

iv. *__Elas__ usaram a pintura **para que**?*　*__Para que__ usaram __elas__ a pintura?*

　　 <u>they</u> used the paint **for what**　　**for what** used <u>they</u> the paint

　　 'They used the paint for what?'　　'For what did they use the paint?'

As for the problematic cases, again some sentences were left blank or unchanged, and others were subject to changes that did not alter their *in situ* structure (addition of a preposition – *onde* > *aonde*, '(to)where', or removal of an article – *do* > *de*, 'of(.the)').

There was a sentence to which the participants provided several different strategies of modification. One of them simply made it a fronted interrogative (29.a), another added subject-verb inversion (29.b), unknowingly turning the sentence's subject into its object, and yet another participant added a preposition on top of that (probably to avoid altering the subject's syntactic function), which made the final sentence ungrammatical in BP (29.c). But the two most interesting cases altered the sentence in a different way, it went from being an open question to a yes/no question in indirect speech. This was possible using a different meaning of the verb *indicar* 'recommend/state'. It was probably caused by this polysemy of the verb, but also because the interrogative constituent was followed by an adjunct (there were no other experimental items in this situation).

(29) Example of amending strategies for a specific *in situ* item in the experimental group

<u>Experimental item</u>　　　　　　<u>Follow-up modification</u>

*Ele indicou **quem** para a vaga?*　a)　*__Quem__ __ele__ indicou para a vaga?*

he indicated **who** for the vacancy　　**who** <u>he</u> indicated for the vacancy

'He recommended who for the job?'　　'Who did he recommend for the job?'

　　　　　　　　　　　　　　b)　*__Quem__ indicou __ele__ para a vaga?*

　　　　　　　　　　　　　　　　who indicated <u>him</u> for the vacancy

　　　　　　　　　　　　　　　　'Who recommended him for the job?'

c) ***A quem** indicou <u>ele</u> para a vaga?*

   to who indicated him for the vacancy

   '*Who did him recommend for the job?'

d) *Ele indicou **quem** <u>vai pegar a vaga</u>?*

   he indicated **who** <u>goes to.get the vacancy</u>

   'Did he indicate who is going to get the job?'

e) *Ele indicou **para quem** <u>é a vaga</u>?*

   he indicated **for who** <u>is the vacancy</u>

   'Did he indicate for who is the job?'

The control group only rejected 2.9% of the *in situ* interrogative sentences. Their amending strategy was fronting the interrogative particle (as in (28.i) or (29.a)) but, unlike the experimental group, there were no added modifications (no extra complementizers, no clefts). The only problematic cases here were due to lack of graphic accent on the interrogative word (*que > qué*, 'what').

Finally, regarding the interrogatives with post-verbal subjects, the experimental group rejected 37.5% of them all, with no distinction between inergative or unaccusative verbs, whereas the rejection rate for the control group was higher 52% and showed a preference for the structure with unaccusatives (out of the rejected sentences, 65% had inergative verbs and 35% had unaccusative verbs). Both groups "fronted" the interrogative particle in these sentences (meaning the subject was placed before the verb), but within the responses of the experimental group we also found some clefts, like the ones in the following example.

(30) Examples of clefts as amending strategies for interrogatives with post-verbal subject.

	Experimental item	Follow-up modification
i.	*Nadou **quem** na piscina?*	***Quem é que** nadou na piscina?*
	swam **who** in.the pool	**who is that** swam in.the pool
	'Who swam in the pool?'	'Who was it that swam in the pool?'
ii.	*Chegou **quem** na aula?*	***Quem foi quem** chegou na aula?*
	arrived **who** in.the classroom	**who was who** arrived in.the classroom
	'Who arrived at the classroom?'	'Who was who arrived at the classroom?'

5.4.3. *Topic-subject constructions*

Now, let us turn to topic-subject constructions. As mentioned in the quantitative analysis, the scores varied depending mainly on the type of sentence structure: canonical sentences were those better rated, target topic-subject constructions received intermediate ratings and the ungrammatical versions were the ones with lower scores. However, both groups behaved in a similar way. The table below shows again the AJT mean scores table for topic-subject constructions.

Group		Canonical		Target		Ungrammatical
Experimental	Mean	**4.02**	Mean	**2.53**	Mean	**2.06**
	SD	(1.17)	SD	(1.30)	SD	(1.23)
Control	Mean	**4.41**	Mean	**3.20**	Mean	**1.86**
	SD	(0.96)	SD	(1.37)	SD	(1.02)

Table 4 – Topic-subject Constructions Mean Scores (AJT)

Canonical

The rejection rates for canonical structures were higher than expected, 14.1% within the experimental group and 9% within the control group. Since these were grammatical BP sentences with no salient syntactic characteristic, the amending strategies in this case were quite varied. But there were two tendencies that can be discussed.

For the subset of sentences regarding locatives, the position of this constituent seemed to be the most recurrent cause of discomfort between the participants of both groups, who preferred to have the subject in pre-verbal position. The fact that this is so for the experimental group is interesting, since it reflects BP grammar and not Spanish grammar, which is more flexible with post-verbal subjects.

For the subset regarding genitives, a strategy that was used several times by the members of the experimental group was the addition of the pronoun *se*. This is a common way of avoiding the passive voice in Spanish, which is mostly restricted to a more formal register, usually written, and sounds unnatural in conversation. In contrast, this kind of *se*[42] is falling into disuse in vernacular BP (Galves, 1986) and the structures taking its place are precisely topic-subject constructions.

---

[42] Actually, the nature of this pronoun is not so clear. Here, I am assuming it corresponds to what is known in the literature as *se medio*.

The following example shows how the experimental group applied the aforementioned strategies (31.i and ii) and also two other modifications that I found interesting to touch upon, a topic-subject construction (31.iii) and a case of locative agreement (31.iv).

(31) Example of amending strategies for canonic structures – Experimental group

	Experimental item	Follow-up modification
i.	***Nessas salas*** *cabem muitos alunos.*	*Muitos alunos cabem* ***nessas salas***.
	in.those classrooms fit many students	many students fit **in.those classrooms**
	'Many students fit in those classrooms.'	'Many students fit in those classrooms.'
ii.	*Os picolés dessa geladeira derretem muito rápido.*	*Os picolés dessa geladeira derretem-**se** muito rápido.*
	the popsicles of.that fridge melt very quickly	the popsicles of.that fridge melt **SE** very quickly
	'That fridge's popsicles melt very quickly.'	'That fridge's popsicles melt very quickly.'
iii.	***Nas universidades está*** *sumindo o dinheiro para pesquisa.*	***As universidades estão*** *sumindo o dinheiro para pesquisa.*
	in.the universities is disappearing the money for research	**the universities are** disappearing the money for research
	'Universities have less and less money for research.'	'Universities have less and less money for research.'
iv.	*Nos supermercados grandes **caiu** o preço da carne.*	*Nos grandes supermercados **cairam** o preço da carne.*
	in.the big supermarkets dropped.3**SG** the price of.the meat	in.the big supermarkets dropped.3**PL** the price of the meat
	'Meat prices have dropped in big supermarkets.'	'Meat prices have dropped in big supermarkets.'

I will not discuss the problematic cases for canonic structures, but they were mainly due to lexical choices (*acabar > terminar* 'to finish').

Target topic-subject constructions

The experimental group rejected 56.8% of the target topic-subject constructions, and when considered by subset, the rejection rate was 65.6% for locative TSCs and 47.9% for genitive TSCs.

For locative TSCs, when the semantic subject and the locative constituent had the same grammatical number, the participants simply added the preposition *em* 'in' and, in most cases, left the rest of the original syntactic structure as it was. This can be seen in the following example (32.i). In a few cases, on top of adding the preposition, the locative was moved to a post-verbal position, the subject was moved to a pre-verbal position, or both things at the same time, as shown in (32.iii). When the locative constituent and the semantic subject did not have the same grammatical number, the amending strategies varied a bit more. However, the most recurrent strategies were the same as before, but included a modification of the verb's morphology too to achieve subject-verb agreement, see examples (32.ii and iv). Example (32.v) shows an interesting case with a different amending strategy (without addition of *em* 'in') that resulted in a topic-comment structure, there were two such cases in our data.

(32) Examples of amending strategies for locative target TSCs – Experimental group

	Experimental item	Follow-up modification
i.	***Aqueles quartos** só cabem duas pessoas.* **those rooms** only fit.3PL two people 'Only two people fit in those rooms.'	***Naqueles quartos** só cabem duas pessoas* **in.those rooms** only fit.3PL two people 'Only two people fit in those rooms.'
ii.	***Aquele bairro** mora só os ricos.* **that neighborhood** live.3**SG** only the.PL rich.PL 'Only rich people live in that neighborhood.'	***Naquele bairro** moram só os ricos.* **in.that neighborhood** live.3**PL** only the.PL rich.PL 'Only rich people live in that neighborhood.'
iii.	***Os jardins** floresceram <u>umas margaridas lindas</u>.* **the gardens** blossomed.3PL <u>some daisies beautiful</u> 'Beautiful daisies blossomed in the gardens.'	<u>*Umas margaridas lindas*</u> *floresceram* **nos jardins**. <u>some daisies beautiful</u> blossomed.3PL **in.the gardens** 'Beautiful daisies blossomed in the gardens.'

iv. ***As lojas estão** caido <u>o preço do celular</u>.*   <u>*O preço do celular* **está** caido **nas lojas**</u>.

     the stores are falling <u>the price of cellphone</u>   <u>the price of cellphone</u> **is** falling **in.the stores**

     'Cellphone prices are falling in the stores.'   'Cellphone prices are falling in the stores.'

v. *O consultório cheg**a** muitos pacientes de manhã.*   *O consultório, cheg**am** muitos pacientes de manhã.*

     the doctor's.office come.3**SG** many patients of.the morning   the doctor's.office come.3**PL** many patients of.the morning

     'A lot of patients come to the doctor's office in the morning.'   'As for the doctor's office, a lot of patients come there in the morning.'

For genitive TSCs, the amending strategy was to add the preposition *de* 'of' and, in addition, group together the genitive with the noun it modifies, placing the regrouped constituent at the beginning of the sentence, in subject position, see (33.i). In a few cases, a pronoun *se* was also added, as can be seen in (33.iii). That is what happened when the grammatical number of the semantic subject and the genitive coincided. As before, whenever those numbers were not equal, the strategies were mainly the same, but added a modification to the verb's morphology so that there was subject-verb agreement, see examples (33.ii and iv). Example (33.v) shows a different strategy in which the semantic subject is turned into a locative, and the genitive takes its place.

(33) Examples of amending strategies for genitive target TSCs – Experimental group

<u>Experimental item</u>   <u>Follow-up modification</u>

i. ***O portão** enferrujou <u>a fechadura</u>.*   <u>*A fechadura* **do portão** *enferrujou*</u>.

     the gate rusted <u>the lock</u>   <u>the lock</u> **of.the gate** rusted

     'The gate's lock rusted.'   'The gate's lock rusted.'

ii. ***Os carros** acab**aram** <u>a gasolina</u> no meio da viagem.*   <u>*A gasolina* **dos carros** *acab**ou** no meio da viagem*</u>.

     the cars finished.3**PL** <u>the gas</u> in.the middle of.the trip   <u>the gas</u> **of.the cars** finished.3**SG** in.the middle of.the trip

     'The cars run out of gas in the middle of the trip.'   'The cars run out of gas in the middle of the trip.'

iii. ***O violão*** *arrebentou* <u>*a corda*</u>.   <u>*A corda*</u> ***do violão se*** *arrebentou.*

    the guitar broke <u>the string</u>   <u>the string</u> **of.the guitar SE** broke

    'The guitar string broke.'   'The guitar string broke.'

iv. ***Aquele relógio*** *estrag**ou*** <u>*os dois*</u>   <u>*Os dois ponteiros*</u> ***daquele relógio se*** <u>*ponteiros*</u>.   *estrag**aram**.*

    that clock ruined.**3PL** <u>the two hands</u>   <u>the two hands</u> **of.that clock SE** ruined.**3PL**

    'The clock hands were ruined.'   'The clock hands were ruined.'

v. *O portão enferrujou **a fechadura**.*   *O portão enferrujou **na fechadura**.*

    the gate rusted **the lock**   the gate rusted **in.the lock**

    'The gate's lock rusted.'   'The gate got rust in its lock.'

The rejection rate for the control group regarding target topic-subject constructions was lower than that of the experimental group. These participants rejected 41.2% of the target constructions, the rate was 37% for locative TSCs and 45.5% for genitive TSCs. Note that, apart from the rate in the experimental group being higher (56.8%), their behavior differs further. The controls rejected more genitive than locative TSCs, whereas the experimental group did the opposite, as the following summary table shows.

Group	Locative TSCs	Genitive TSCs	Total
Experimental	65.6%	47.9%	**56.8%**
Control	37.0%	45.5%	**41.2%**

*Table 9 – Rejection rates (% of score 1 or 2 in the AJT) for **target** topic-subject constructions*

Nonetheless, the main amending strategies used by both groups were the same; with one exception, the addition of the pronoun *se*. This was a strategy only used by the experimental group for genitive TSCs (as mentioned before, this construction is disappearing from BP, but it is very common in Spanish).

There was a slight difference that is worth mentioning. Although it was a strategy also found within the non-native speakers, locative agreement structures appeared more frequently within the control group follow-up modifications. There are two examples below.

(34) Examples of locative agreement in amending strategies for locative target TSCs – Control group

Experimental item	Follow-up modification
i. ***As ruas do centro*** *não estão passando carro.*	***Nas ruas do centro*** *não estão passando carro.*
the streets of.the downtown not are passing cars	**in.the streets of.the downtown** not are passing car
'Cars do not go through downtown streets.'	'Cars do not go through downtown streets.'
ii. ***O consultório*** *chega muitos pacientes de manhã.*	***No consultório*** *chega muitos pacientes de manhã.*
the doctor's.office come.3SG many patients of morning	**in.the doctor's.office** come.3SG many patients of morning
'A lot of patients come to the doctor's office in the morning.'	'A lot of patients come to the doctor's office in the morning.'

But the major difference between groups was that the controls rarely diverged from the main amending strategies previously discussed, whereas the participants from the experimental group were more 'creative' in their follow-up modifications, to the point of altering the meaning of the sentence in some cases, as can be seen in the following example (35.i), or directly copying Spanish syntactic structures, like in (35.ii).

(35) Examples of diverging amending strategies for target TSCs – Experimental group

Experimental item	Follow-up modification
i. <u>*Os armários*</u> *arranharam* <u>*as portas*</u>.	<u>*As portas*</u> *arranharam* <u>*os armários*</u>.
<u>the closets</u> scratched <u>the doors</u>	<u>the doors</u> scratched <u>the closets</u>
'The closets' doors got scratched.'	'The doors scratched the closets.'
ii. *Meu filho já* ***está*** *nascendo os dentes.*	*A meu filho já* ***lhe estam*** *nascendo os dentes.*
my son already is birthing the teeth	**to** my son already **him are** birthing the teeth
'My son's teeth are already coming in.'	'My son's teeth are already coming in.'

Ungrammatical topic-subject constructions

The last subset of sentences is what I called 'ungrammatical' topic-subject constructions. Recall that these sentences share the same topic-subject structure as the previous kind but, unlike those, their main verb is intransitive instead of unaccusative. From what is known about topic-subject constructions, this should make them ungrammatical, hence the name of the subset.

The total rejection rate for these sentences within the experimental group was 76.1%, for locative TSCs it was 84.9% and for genitive TSCs it was 67.4%. The main amending strategies were the same that were used with target TSCs, with one exception I will shortly discuss.

For locative TSCs, when the semantic subject and the locative constituent shared the same grammatical number, the participants added the preposition *em* 'in', or, in some cases, a different preposition with locative meaning but that formed a better collocation with the verb (*caminhar pelo parque* 'walk through the park'). This can be seen in the example below (36.i). In contrast with the strategies for target TSCs, on top of adding the preposition, here it was more common to also rearrange the sentence constituents placing the verb between the subject and the locative (in that order), see (36.iii). In addition to those two strategies, there were also some instances of the passivizing pronoun *se* or an equivalent construction, all in response to the same experimental item, see (36.v)

Again, when the locative and the semantic subject did not share the same grammatical number, there was a modification in the morphology of the verb or of the subject in order to obtain subject-verb agreement. The strategies remained the same, see examples (36.ii) and (36.iv).

(36) Examples of amending strategies for locative ungrammatical TSCs - Experimental group

Experimental item	Follow-up modification
i. ***Aquelas praias*** *mergulham só os turistas.*	***Naquelas praias*** *mergulham só os turistas.*
those beaches dive only the tourists	**in.those beaches** dive only the tourists
'Only tourists dive in those beaches.'	'Only tourists dive in those beaches.'

ii. *Esses bares dançam só bébado*[43].   *Nesses bares dançam só bébados.*

   those bars dance.3PL only drunk.**SG**   **in.those bars** dance.3PL only drunk.**PL**

   'Only drunk people dance in those bars.'   'Only drunk people dance in those bars.'

iii. *A esteira correu o atleta.*   *O atleta correu na esteira.*

   the treadmill ran the athlete   the athlete ran **in.the treadmill**

   'The athlete ran on the treadmill.'   'The athlete ran on the treadmill.'

iv. *As ruas protestaram o sindicato.*   *O sindicato protestou nas ruas.*

   the streets protested.3PL the union   the union protested.3**SG in.the streets**

   'The union took to the streets to protest.' 'The union took to the streets to protest.'

v. *Aquela lagoa pesca muito pintado.*   *Naquela lagoa dá para pescar muito pintado.*

   that lake fishes many spotted (catfish) **in.that lake gives to to.fish** many spotted

   'In that lake you can catch a lot of   (catfish)

   catfish.'   'In that lake it is possible to catch a lot of

                                               catfish.'

As with their target counterparts, for ungrammatical genitive TSCs, the amending strategy was to add the preposition *de* 'of' and group together, in subject position, the genitive with the noun it modifies, see (37.i). When the grammatical number of the semantic subject and the genitive were not the same, the verb's morphology was also altered to achieve subject-verb agreement, see (37.ii). In this case, though, there were no instances of the pronoun *se*. Instead, we found the following modification: the genitive becomes the semantic subject, which in turn becomes a complement or an adjunct of the verb, see (37.iii and iv). This changed completely the intended meaning of the original sentence, which was not always easy to grasp, given its ungrammaticality. Although it happened only with the two experimental items shown in the example, it was the preferred strategy in those cases, becoming one of the main amending strategies. This was due to the specific semantic characteristics of these experimental items, which allowed that interpretation. And this is the exception I mentioned before.

---

[43] Note that what we have here in this example is a bare singular noun, which, in BP, admits a generic reading, equivalent to that of a bare plural (Ionin et. al, 2015). This may be motivating the participants to change the subject's morphology instead of the verb's (like they did in most cases).

(37) Examples of amending strategies for genitive ungrammatical TSCs - Experimental group

	Experimental item	Follow-up modification
i.	*A Maria* sempre chora *o bebê à noite.*	*O bebê **da Maria** sempre chora à noite.*
	the Maria always cries the baby at.the night	the baby **of.the Maria** always cries at.the night
	''Maria's baby always cries at night.'	'Maria's baby always cries at night.'
ii.	*Meus amigos cochilam o filho sempre à tarde.*	*O filho **dos meus amigos** cochila sempre à tarde.*
	my friends nap.3**PL** the son always at.the afternoon	the son **of.the my friends** nap.3**SG** always at.the afternoon
	'My friends' son always takes a nap in the afternoon.'	'My friends' son always takes a nap in the afternoon.'
iii.	*O Pedro brigou **as irmãs** o dia todo ontem.*	*O Pedro brigou **com as irmãs** o dia todo ontem.*
	the Pedro fought **the sisters** the day all yesterday	the Pedro fought **with the sisters** the day all yesterday
	'Yesterday, Pedro's sisters fought all day.'	'Yesterday, Pedro fought with his sisters all day.'
iv.	*O Pedro está viajando **os pais** pelo mundo inteiro.*	*O Pedro está viajando **com os pais** pelo mundo inteiro.*
	the Pedro is travelling **the parents** by.the world whole	the Pedro is travelling **with the parents** by.the world whole
	'Pedro's parents are travelling all over the world.'	'Pedro is travelling with his parents all over the world.'

Before going on with the discussion, I would like to underline the fact that the intended meaning for the experimental items in this subset of TSCs is the meaning they would have if they were indeed interpreted as TSCs. However, considering that they are ungrammatical sentences artificially built for this experiment, that TSC interpretation was quite hard to attain in some cases. This happened especially with items for which there was a much more natural option, (37.iii and iv) are a perfect example of this.

The experimental group diverged from the main amending strategies discussed above on several occasions. For example, the verb was substituted with a transitive one (38.i), or a tonic pronoun was added (38.ii). There were also two interesting problematic cases, two instances of alterations that preserved the TSC structure and merely modified the position of a temporal adverb, see (38.iii).

(38) Examples of diverging amending strategies for ungrammatical TSCs – Experimental group

Experimental item	Follow-up modification
i. *Os meus tios **pularam** o gato do telhado.* the my aunt.and.uncle **jumped** the cat off.the roof 'My aunt and uncle's cat jumped off the roof.'	*Os meus tios **pegaram** o gato do telhado.* the my aunt.and.uncle **got** the cat off.the roof 'My aunt and uncle got their cat off the roof.'
ii. ***Esses bares** dançam só **bébado**.* those bars dance only drunk.SG 'Only drunk people dance in those bars.'	***Eles** dançam só **bébados nesses bares**.* **they** dance only drunk.**PL in.those bars** 'In those bars, they only dance when they're drunk.'
iii. *Os livros <u>às vezes</u> enganam as capas.* the books at.the times lie the covers 'Book covers sometimes lie.'	*<u>Às vezes</u> os livros enganam as capas* at.the times the books lie the covers 'Sometimes book covers lie.'

Unlike what happened with target TSCs, the rejection rate for ungrammatical topic-subject constructions within the control group was higher than the rate within the experimental group. The control participants rejected 83% of the ungrammatical constructions, this rejection rate was 75% for locative TSCs and 91% for genitive TSCs. Again, the groups' behaviors differ further than just in the total rate of rejection. The controls rejected more genitive than locative TSCs, whereas the experimental group did the opposite, as can be seen in the table below. This is the same difference we saw for target TSCs.

Group	Locative TSCs	Genitive TSCs	Total
Experimental	84.9%	67.4%	**76.1%**
Control	75.0%	91.0%	**83.0%**

*Table 10 – Rejection rates (% of score 1 or 2 in the AJT) for **ungrammatical** topic-subject constructions*

The control group used the same main amending strategies than the experimental group, but they also diverged from them in some cases. For locative ungrammatical TSCs, there were some instances of locative agreement within the control group, and they also used the pronoun *se* (or the equivalent construction *dá para* 'one can' + infinitive) for a specific experimental item, see (36.v). But, in general, the experimental group diverged more from the main strategies than the controls. The situation was reversed for genitive constructions, though. In this case, it was the control group the one that showed more diversity in their amending strategies.

The following example shows some of the diverging amending strategies used by this group, many of which alter the sentence's intended meaning. In (39.i), the locative TSC is substituted with an existential sentence and the verb is nominalized. The use of a different verb, a transitive one, was somewhat common, see (39.ii). In (39.iii), the chosen strategy was to use a causative structure instead of the genitive TSC. In (39.iv), the preposition added was not *de* 'of', but one that transformed the genitive into a locative or a different kind of constituent. And in (39.v), the genitive remained as such, but it was linked to a locative constituent instead of being linked to the semantic subject.

(39) Examples of diverging amending strategies for ungrammatical TSCs – Control group

	Experimental item	Follow-up modification
i.	*As ruas **protestaram** o sindicato.*	***Houve protesto** do sindicato nas ruas.*
	the streets **protested** the union	**there.was protest** of.the union in.the streets
	'The union took to the streets to protest.'	'There was a union demonstration in the streets.'
ii.	*A Maria sempre **chora** o bebê à noite.*	*A Maria sempre **conforta** o bebê à noite.*
	the Maria always **cries** the baby at.the night	the Maria always **comforts** the baby at.the night
	'Maria's baby always cries at night.'	'Maria always comforts her baby at night.'

iii. *Meus amigos **cochilam** o filho sempre à tarde.*

    my friends **nap** the son always at.the afternoon

    'My friends' son always takes a nap in the afternoon.'

*Meus amigos **fazem cochilar** o filho sempre à tarde.*

my friends **make nap** the son always at.the afternoon

'My friends always make their son take a nap in the afternoon.'

iv. *Os livros às vezes enganam as capas.*

    the books at.the times lie the covers

    'Book covers sometimes lie.'

*Às vezes os livros enganam **pelas** capas*

at.the times the books lie **by.the** covers

'Sometimes books lie by their covers.'

v. *Meus tios **pularam** o gato do telhado.*

    the my aunt.and.uncle jumped.3**PL** the cat off.the roof

    'My aunt and uncle's cat jumped off the roof.'

*O gato **pulou** do telhado dos meus tios.*

the cat jumped.3**SG** off.the roof of my aunt.and.uncle

'The cat jumped off my aunt and uncle's roof.'

In addition to these, there were also two problematic cases worth mentioning. The first one is another case of a modification that did not alter the TSC structure of the original sentence, it only changed the verb's aspect from perfective to imperfective, see (40.i). In the second one, it seems as if the TSC structure was also preserved, but it is not clear. The participant altered the grammatical number of the semantic subject, leaving the rest of the sentence as it was, see (40.ii).

(40) Examples of problematic cases for ungrammatical TSCs – Control group

    <u>Experimental item</u>

    <u>Follow-up modification</u>

i. *Os jóqueis **relincharam** os cavalos muito alto.*

    the jockeys neighed.**PERF** the horses very loudly

    'The jockeys' horses neighed very loudly.'

*Os joóqueis **relinchavam** os cavalos muito alto.*

the jockeys neighed.**IMPERF** the horses very loudly

'The jockeys' horses neighed very loudly.'

ii. *Meus amigos cochilam **o filho** sempre à tarde.*
my friends nap.3PL **the son** always at.the afternoon
'My friends' son always takes a nap in the afternoon.'

*Meus amigos cochilam **os filhos** sempre à tarde.*
my friends nap.3PL **the sons** always at.the afternoon
'My friends' sons always take a nap in the afternoon.'

5.4.4. *Summing up*

In situ interrogatives were highly accepted by both groups, only 10% of these sentences were rejected by the experimental group, and the rate decreased to 3% in the control group. When rejected, *in situ* interrogatives were then fronted in the follow-up task, both by the experimental and the control group. However, the experimental group sometimes added an extra complementizer, or a complementizer and a copula, or placed the subject in post-verbal position.

The rejection rate for topic-subject constructions, leaving the ungrammatical cases aside, was higher in both groups. The experimental group rejected 57% of these sentences and showed a slight preference for genitive TSCs, whereas the control group rejected 41% and seemed a bit more comfortable with locative TSCs.

For locative TSCs, the main amending strategy was the same in both groups and consisted in adding the preposition *em* 'in' and modifying the verb's morphology when necessary to ensure subject-verb agreement. In addition to this, participants sometimes preferred to place the locative in postverbal position, the subject in preverbal position, or both. There were also some marginal instances of locative agreement structures, especially within the control group.

For genitive TSCs, the main strategy was again the same in both groups. This time it consisted in adding the preposition *de* 'of' and grouping together, in subject position, the genitive with the noun it modifies, the verb's morphology was also modified when necessary. Only the participants in the experimental group sometimes added a *se* pronoun as well.

The main difference between the two groups, regarding target TSCs, was that the control group rarely diverged from the main strategies aforementioned whereas the experimental group was more creative in their modifications.

The rejection rate for ungrammatical versions of TSCs was very high, 76% for the experimental group and 83% for the control group. Main amending strategies were the same used for target TSCs, but both groups also used other strategies, which depended quite strongly on the specific experimental items.

Rejection of fronted interrogatives and of canonical versions of TSCs was low but unexpected, since these were grammatical BP sentences. In both groups, it was due to the same reasons (an extra complementizer for interrogatives, and position of the locative constituent or lack of a *se* pronoun for canonical structures). However, the information provided by this is not relevant for this experimental study.

In the following chapters, we will turn to the discussion of the theoretical and empirical implications of our study.

PART III

Theoretical and Empirical Implications

Chapter 6
Empirical Considerations

6.1. Introduction

This chapter constitutes a discussion of some of the methodological difficulties encountered while conducting the experimental study, I comment on the problems I found in the experimental design and the other tools I used, considering possible solutions and alternatives.

Section 6.2 focuses on the participants, the recruitment difficulties and the classifying tools, namely, the language background questionnaire and the proficiency test. Section 6.3, on the other hand, is dedicated to the experimental design. There, I provide some insights and suggest improvements based on this experience and on the results analysis aiming for future studies.

6.2. Participants: recruitment and classifying tools

The biggest obstacle I encountered during the experimental study was the recruitment of participants for the experimental group. First of all, it was simply a matter of scarcity of Spanish native speakers willing to participate in the experiment. Campinas (and even São Paulo) is a less common destination for Spaniards than I had anticipated.

Secondly, the ideal experimental participant was a Spanish native speaker living in Brazil and acquiring Portuguese in a naturalistic manner. Clearly, such a participant was almost impossible to find, since taking some kind of formal training in the country's language is extremely common for foreigners, especially within the academic environment. This would not be as relevant if, after, or in addition to, the formal training, the participant had an immersive linguistic experience in the country. However, for many of the experimental participants in this study, previous or ongoing formal training constituted their main form of contact with Brazilian

Portuguese. Although these are not the ideal conditions for SLA research in general[44], they may be especially relevant for this study because one of the syntactic structures being tested, topic-subject constructions, violates prescriptive grammar's rules, although grammar books and L2 textbooks do not call attention to these structures (more on this in the next chapter, section 7.3).

Some participants in this group, however, had mainly learned BP by living in the country and spent their everyday lives immersed in the language and the culture. Interestingly enough, the kind of learning experience the experimental participants had was not a determinant factor for their scores in the proficiency test.

In addition, given the difficulty of finding experimental participants, there was much more variety in age and backgrounds within this group than within the control group. All of this resulted in a quite heterogeneous experimental group.

As I mentioned in the previous chapter, there were even a few bilingual participants who had moved from Spain to Brazil (some when they were quite young, one during middle age). It was a shame there were not more participants in this category, it would have been interesting to compare this group with the other two.

Control participants were easier to recruit. They were obviously more abundant, but they were also more willing to take part in a research experiment. Nonetheless, these participants were surprisingly self-conscious about their linguistic abilities and got quite nervous at the mention of a proficiency test.

6.2.1. *Linguistic Background Questionnaire*

The linguistic background questionnaire proved to be a useful tool. It was helpful to understand what the experience with BP was like for each experimental participant and to determine what kind of second language acquisition they were having.

However, it was unnecessary long. Whenever it needed to be done in the same session as the experiment (which was quite often), the questionnaire was always the last task, and I could tell it became tiresome for the participants. Considering the way in which we ended up using this questionnaire, it could have been reduced to a smaller number of questions, those that provided

---

[44] This is a personal observation; most SLA studies are actually done with participants who acquired the language through formal training and often times in their country of origin. I believe, though, that if we are interested in natural language acquisition, those are not the best testing conditions.

the essential information: what languages the participant spoke (indicating order of acquisition and order of dominance), what their daily exposure time to BP was, and what their learning method had been. To see the whole questionnaire, as applied to participants, see Appendix II.

The questionnaire was created to extract information about the participants in the experimental group, its questions did not apply to control participants, so they did not perform this task. Nonetheless, it was a mistake not to create a specific questionnaire for the control group. Given the suspected diatopic variation of topic-subject constructions and of the interrogative structures containing an extra complementizer, it would have been useful to know what BP variant each control participant spoke. This could have been determined by such a questionnaire, consisting of a couple of simple questions.

6.2.2. *Proficiency Test*

Proficiency tests are extremely useful (if not indispensable) tools in SLA research. For this experimental study, the results of the proficiency test provided a way of classifying the experimental participants and dividing them into more homogeneous subgroups. As the analysis in Chapter 5 showed, this classification revealed interesting differences in these subgroups' behaviors, which would otherwise have been concealed by the heterogeneity of the experimental group as a whole.

As mentioned in Chapter 4 (section 4.4.3), the proficiency test was the same used in Ionin et al. (2015) and it was kindly provided by the authors, who had previously confirmed its reliability in Montrul et al. (2011). This constituted one of the main reasons for choosing this particular test, the other one being the fact that it was based on the Spanish proficiency test created by the Spanish Ministry of Education, known as DELE (*'Diploma de Español como Lengua Extranjera'*). The test can be found in its entirety in Appendix III.

A posteriori, though, this test may not have been the one best suited for the specific purposes of our experimental study. One of the two syntactic structures under investigation here, topic-subject constructions, despite being common in vernacular BP, is not sanctioned by prescriptive grammar, and even less so in formal written contexts. But tests officially used by countries to certify the applicants' proficiency level in the country's language, like the one this test is based on, tend precisely to follow the rules of prescriptive grammar. As an example, one of the items in the test concerned a common orthographic mistake in BP, the accent necessary to distinguish

the feminine article or the preposition 'to' (which happen to share the same written form *a*) from their contraction (*à* = *a* + *a*, 'to + the').

This could have led to two different problems. First, non-native speakers who did not acquire the language through formal training and whose level of proficiency in vernacular BP is quite high might not get a score representative of their abilities. I suspect this is the case for some of the participants in the advanced subgroup, who did not get the highest scores in the proficiency test but showed great consistency in their AJT responses.

And second, this kind of proficiency test might trigger a prescriptive frame of mind, which can be detrimental to the acceptability judgements of vernacular sentences. Casual conversations with a couple of participants confirmed this, mainly with those in the control group, who were worried about taking a proficiency test in their own mother tongue. To avoid the problem as much as possible, I only gave the test to the participants after they had completed the acceptability judgement task and the follow-up task. This way, the prescriptive effect of the proficiency test could not interfere with the experiment.

In addition to this, Spanish grammar is not that distant from BP grammar and many of their lexical items share common and recognizable roots, making it relatively easy for native Spanish speakers to get high scores in a test which is intended for foreign students from any nationality.

Considering all this, I believe a proficiency test which included other syntactic structures typical of vernacular BP but not sanctioned by prescriptive grammar (like example 1.i below) and that focused on specific differences between Spanish and BP (example 1.ii) may have provided a better discrimination tool for this study.

(1)   Examples of possible sentences for a proficiency test.

   i.   *Paulo encontrou ela no cinema ontem.*

          Paulo found she in.the cinema yesterday

          'Paulo run into her at the cinema yesterday.'

   ii.   Ele acordou logo.[45]

          he woke.up soon

          'He woke up soon.'

---

[45] *Acordar* and *logo* are false friends. Spanish has words very similar to those, but with a different meaning ('remember' and 'afterwards', respectively). The sentence could then be mistakenly interpreted in a completely different way (something like 'He remembered later.').

6.3. Experimental design

6.3.1. *Sentence selection*

The way experimental items were selected was thoroughly explained in Chapter 4. But in this section, I want to make some comments thereon in light of the experiment results.

The overall philosophy was to diversify as much as possible. So, for TSCs, in addition to varying the type of constructions (locative *vs.* genitive) and the grammatical number of the relevant constituents, verb tense and aspect were also diversified, having some TSCs with a verb in the present or present continuous and others with a verb in past tense.

(2)   Diversification in TSCs experimental items

i.   *Essa casa* (SG) *bate bastante sol* (SG).   Locative TSC – present

that house beat.3SG plenty sun

'That house gets plenty of sunlight.'

ii.   *O quarto* (SG) *está mofando as paredes* (PL).   Genitive TSC – pres. Cont.

the room is molding the walls

'The room walls are getting moldy.'

iii.   *Os armários* (PL) *arranharam as portas* (PL).   Genitive TSC – past

the closets scratched.3PL the doors

'The doors of the closets were scratched.'

For interrogative sentences, variety was accomplished by questioning either argumental or adjunct constituents with different syntactic functions within the sentence (such as direct object, indirect object, or adverbial adjuncts of time, manner, place, etc.), by the occasional addition of an extra complementizer to fronted interrogatives, and also by varying the verb's tense and/or aspect.

(3)   Diversification in interrogative experimental items

i.   *O que que o Pedro vai comprar?*   Argument – DO

'What (is) that Pedro is going to buy?'

ii.   *A Maria emprestou a caneta para quem?*   Argument – IO

'Maria lent the pen to whom?'

iii.   *Os alunos vão se reunir onde?*   Adjunct – Place

'The students are going to meet where?'

iv. *Quando seus pais vão embora?*   Adjunct – Time

'When are your parents leaving?'

This resulted in a high number of experimental conditions, which, in turn, implied having a smaller number of items per condition (for each person). Considering that this was an exploratory study and that the results revealed none of the studied conditions were truly relevant, I wonder whether it would have been better to have a simpler experimental design. This could have been done by reducing the conditions and building a more homogeneous set of sentences. For example, having only one type of TSC or fixing the grammatical number of their constituents, and having interrogatives questioning only argumental constituents (and avoiding extra complementizers).

If time had allowed it, it would have been interesting to conduct a previous study with BP native speakers to select which TSCs and *in situ* interrogatives were more acceptable and determine what characteristics made these sentences more natural to native speakers. The results in this pre-study could then have helped in the sentence selection for the main experiment.

Finally, extra constituents were added to some of the sentences when it felt necessary to aid in its comprehension and to make them sound like real sentences (as opposed to experimental items in a study), see the example below.

(4)   Sentences with extra constituents

i. *Os muros da minha casa mofam <u>sempre que chove</u>.*   Canonical

'The walls in my house go mouldy whenever it rains.'

ii. *As TVs <u>da sala de espera</u> sumiram a imagem <u>de repente</u>.*   Target TSC

'The image in the waiting room TVs suddenly disappeared.'

iii. *Essa avenida desfila as escolas <u>de samba</u> <u>no carnaval</u>.*   Ungramm. TSC

'Samba schools march down that avenue during Carnival.'

With target and ungrammatical versions of TSCs, this was mainly done in order to facilitate the expected interpretation, although in some cases, extra constituents were crucial to ensure comprehension (like in 4.iii). With their canonical versions though, it does not seem that necessary. All this may have overcomplicated the sentences and may have even allowed unexpected interpretations on some occasions.

6.3.2. *Experimental paradigm*

The chosen experimental paradigm was an acceptability judgement task complemented by a follow-up task. As I argued in Chapter 4, this paradigm constitutes a direct way to obtain negative evidence, so it is still a fundamental tool for syntactic experimental studies despite the many criticisms it has received over the years.

I stand by this choice, especially considering how fruitful it was to pair the acceptability judgements with a correction follow-up task that provided information about the reasons behind negative judgements. However, I cannot help but agree to some extent with this paradigm's detractors. Acceptability judgements require the participants to tap into their metalinguistic knowledge and this leads to a prescriptive frame of mind that can negatively affect the results. This type of task can also turn out to be a bit artificial.

In retrospect, there might have been a way of making the test less artificial. Since some authors argue that common ground plays an important role in the licensing of *in situ* interrogatives (see Pires and Taylor 2007, Oushiro 2009, Zocca 2010, specifically for BP), and topic-subject constructions needed extra information to ensure the desired interpretation (especially for target and ungrammatical versions), I believe the experimental design would have benefited from the addition of a small paragraph placing the experimental items in context before asking for a judgement. This way, the actual sentences being judged could have been much simpler in terms of syntactic structure, as the need to provide more information to ensure their comprehension would have already been fulfilled by the preceding paragraph, eliminating the need for extra constituents.

(5)   Example of experimental items for AJT with context

   i.   Context:

   Duas amigas conversando. Uma diz: 'Two friends talking. One says:

   - Ontem me ligaram a Maria e   - Yesterday, Maria called me and
   depois o Pedro. O Pedro queria que   then Pedro did too. He wanted me to
   eu fosse com ele ao cinema!   go to the cinema with him!

   E a outra responde: And the other one replies:'

   Item to be judged:

   - E a Maria queria o que?   '- And Maria wanted what?'

ii.   Context:

Hoje eu estava assistindo TV enquanto esperava no consultório do meu dentista. Daí, começou a chover forte e depois de um trovão, ...	'Today I was watching TV while waiting at my dentist office. Then it started to rain heavily and after a thunder, …'

Item to be judged:

...as TVs da sala sumiram a imagem.	'… the waiting room TVs lost the signal.'

When using acceptability judgements, though, it is very hard to circumvent the problem of prescriptive grammar influence. What can be done is mitigate the effect by combining this type of tasks with others involving production or spontaneous use of language.

In this case, for example, the AJT could be complemented with a production task involving elicitation. For the topic-subject constructions, the same kind of contexts mentioned before could be now used to request from the participant the production of a sentence. In some cases, the beginning of the sentence would be given to try to 'force' the syntactic structures, in others, there would be no hint whatsoever, some visual stimuli could be shown as well. For interrogative questions we could also use the same kind of contexts and then request the participant to ask about a specific element.

(6) Example of experimental items for AJT with context

i.   Context:

Um amigo fala para você:	'A friend tells you:
- A Maria, o João e o Pedro foram no shopping pra comprar um presente de aniversário pra Aline. A Maria comprou um livro e o João comprou um brinquedo.	- Mary, John and Peter went to the mall to buy a birthday present for Alice. Mary bought a book and John bought a toy.'

Cue to elicit sentence:

Pergunte acerca do Pedro.	'Ask about Peter.'

ii.   Context:

Os moços da mudança foram muito pouco cuidadosos com as minhas coisas. Eles até deixaram cair o caixote das xícaras!	'The guys from the moving agency were very uncareful with my stuff. They even dropped the box with the teacups.'

Cue to elicit sentence:

O que aconteceu?	'What happened?'

Possible hint:

Minha xícara…	'My teacup …

Visual stimulus:

A different approach could be to combine the AJT with an interpretation task. This would have been interesting for the topic-subject constructions, since I believe Spanish speakers not only do not produce this kind of structures but may even have trouble understanding some of them.

6.3.3. *Reaction Times*

The results analysis showed that measuring reaction times was not as useful for this study as I would have liked. There were some differences by group and by type of sentence independently, but, contrary to my expectations, there was no interaction between the two factors.

So, reaction times did not help distinguish between type of sentence and how each group treated them. I expected to find significatively longer reaction times for target and ungrammatical TSCs in the experimental group compared to the controls, but not for *in situ* interrogatives. This was

not the case, though, none of these differences reached significance. Both groups showed a very similar relative behavior towards the five kind of experimental items.

There are many factors that may influence reaction times when rating sentences, syntactic structure is clearly one of them, but it is definitely not the only one. Another important factor is the length of the sentence. This was not ignored while designing the experiment, but it may have not received the careful attention it needed. It is unreasonable to compare the reaction times for two sentences that differ drastically in length. And, although most of the experimental items in this study were not radically different in this respect, there were some extreme pairs in the lists of sentences, as can be seen below.

(7)   Examples of experimental items pairs with considerable length difference
   i.   *A festa é quando?*   vs.   *A água dos reservatórios acabou durante o inverno.*
        'When's the party?'   'Water reservoirs got empty during winter.'
   ii. *Quem que morreu?* vs.   *Os parques caminham muitas pessoas no final de semana.*
        'Who died?'   'Many people walk in the park in the weekends.'

Since the differences were always between the items testing TSCs and the interrogative items, this, might have been one of the factors contributing to the longer reaction times seen for the former when compared to the latter. Again, sentence length (in number of syllables, for example) would have been easier to control with simpler and shorter sentences.

On another note, there was also no correlation between the sentence's score and the time it took the participants to rate the sentence. This can be seen in the following graphs, which show for each possible score (from 1 to 5) all the reaction times of sentences that received that score from participants in the experimental group on the left and from the control group on the right.

(8)   Reaction times by sentence score for both groups.
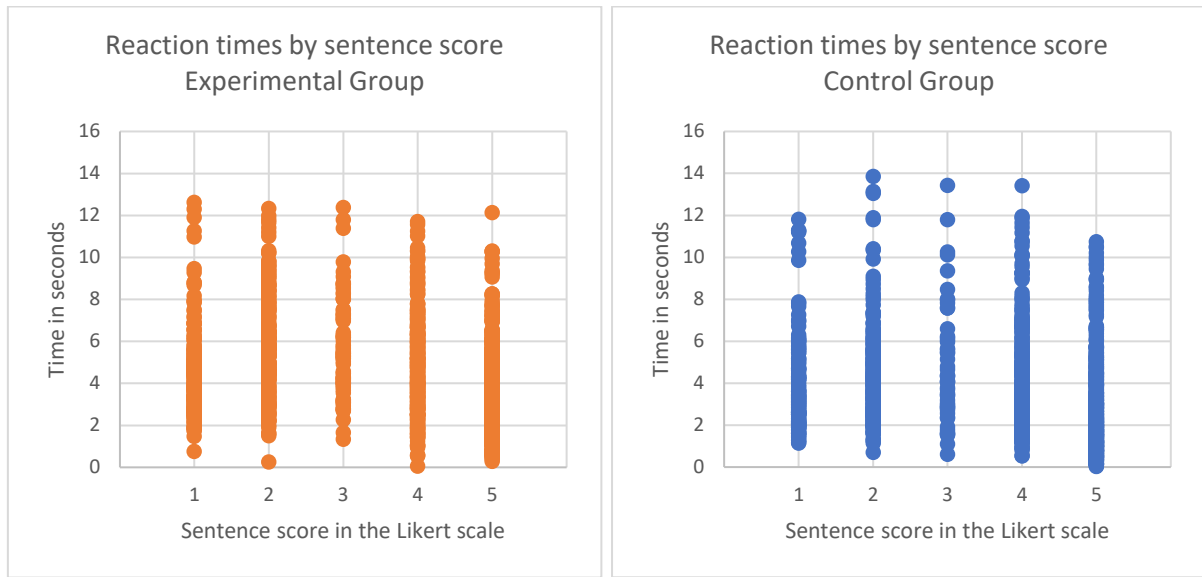
Given this situation, it would have been interesting to find a way to incorporate the information provided by reaction times into the analysis of AJT scores. For example, creating a score modulated by the reaction time needed by the participant to produce the AJT score. The modulation I have in mind would be done in such a way that a bigger reaction time would soften the original score. A sentence that received an AJT score of 5 but took a long time to process would see its new score somewhat reduced, whereas one receiving a lower AJT score also with a long reaction time would receive a slight boost in its new score.

Unfortunately, sorting out all the details to come up with a robust new measurement in these lines would far exceed the primary goals of the dissertation, but it should be kept in mind for the future.

Chapter 7

Theoretical Implications and Conclusions

7.1. Introduction

In this chapter, I discuss the theoretical implications of our experimental study results.

I start by arguing in favor of the hypothesis previously put forth, showing the different ways in which the study results support it, section 7.2 is devoted to this. Section 7.3 constitutes a debate on the impact of prescriptive grammar, which, though hard to deny, I will argue it cannot be made responsible for the results alone.

In section 7.4, I explore in what ways this study can contribute to the theoretical discussions about current theories on second language acquisition. Whereas section 7.5 is devoted to the proposals found in L1 literature for the two syntactic structures under study and how they may be affected by our study.

Section 7.6 presents some ideas for future work on this same topic and a proposal for studying the impact of feature strength in SLA. Finally, in section 7.7, I review the work carried out throughout the study, summarize our results and offer some concluding remarks.

7.2. So, what about our hypothesis?

First things first, I would like to remind the reader that this experimental study was designed to test a hypothesis about the differences in how, or when, Spanish native speakers acquire BP wh *in situ* and topic-subject constructions. This hypothesis was put forward in Chapter 4 (section 4.2) and is now restated here for ease of reading.

(1)   *Hypothesis: There is a significant difference in the acquisition of BP wh* in situ *and topic-subject constructions by Spanish native speakers. More specifically,*

> *the former is acquired earlier whereas the latter is acquired much later or may*
> *never be incorporated into the L2 grammar in its steady state.*

It is only reasonable, then, to start exploring the theoretical implications of this study by discussing this hypothesis and whether the results verify or falsify it. In what follows, I will argue that, although a global overview of the study results does not seem to support the hypothesis, a deeper and more complete analysis shows that, in fact, they sustain it.

As already noted, the study results were not exactly as expected. To sum up briefly, the experimental and control groups behaved in a more similar way than what I anticipated, as can be seen in the following graph. (Nonetheless, recall that the differences seen in this graph for canonical and ungrammatical TSCs, though small, were statistically significative.)

(2)   Boxplot graphs for AJT scores from both experimental and control group
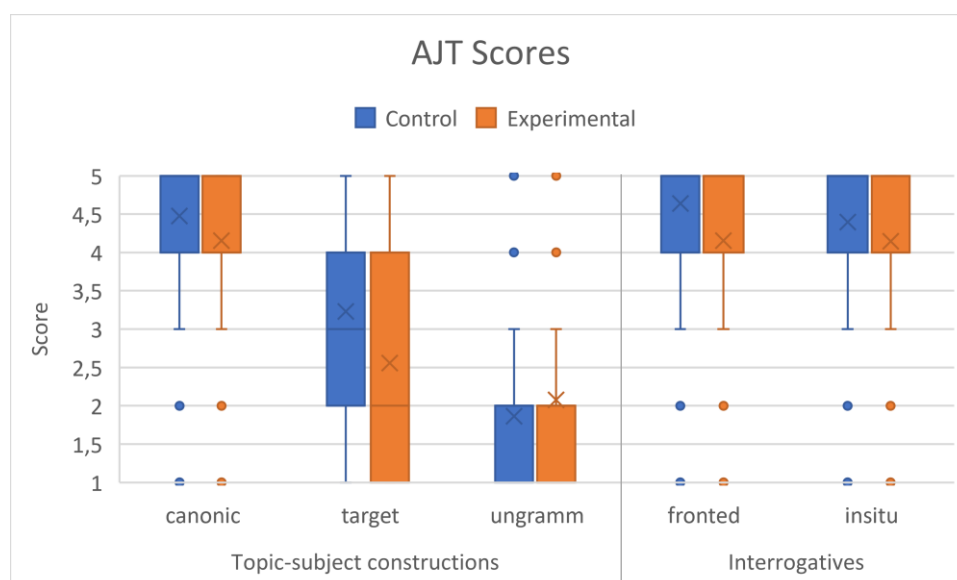
What I had anticipated was to find similar behaviors regarding interrogative sentences but to see steeper differences regarding topic-subject constructions. More precisely, I thought control participants would rate target versions of TSCs mainly as acceptable (although maybe not as much as their canonical versions) and rate ungrammatical versions mainly as unacceptable, whereas experimental participants would only rate canonical versions as acceptable and both target and ungrammatical TSCs as unacceptable, having trouble to distinguish between the two.

Those were my initial expectations. Regarding the experimental participants, I see now that these expectations were probably too narrow and based specifically on my own personal experience and on my own judgement of the two BP syntactic structures under study. On second thought, it is not only reasonable but maybe even what should be expected, that non-native

speakers react to unknown structures with uncertainty rather than with clear rejection, especially those with lower proficiency levels.

Regarding BP native speakers, the scores attributed to TSCs are not that surprising if we take into consideration that these structures may be subject to diatopic and diastratic variation, or to diachronic variation. This variation manifests itself in the AJT scores among the group (some participants being more accepting of these structures than others) or within individuals (the same person accepting some of these sentences but rejecting others). Prescriptive grammar effects may also be a factor contributing to TSCs scores. As argued in Kato (2005), in the I-language of a Brazilian literate individual, parametric values present in his/her peripheral grammar may be competing with the values present in his/her core grammar.

Although this first global analysis of the results does not support the hypothesis, it does not necessarily falsify it either. It is true that, as a whole, the experimental group showed higher acceptance for *in situ* interrogatives than for topic-subject constructions. However, this is also true for the control group. Apparently, they behaved practically in the same way. If the members of the experimental group rate topic-subject constructions just like controls do, we cannot affirm they have not acquired these structures yet, as was suggested in section 4.2, since they are treating them like native speakers do. But are they really treating them in the same way?

A first hint of the answer came actually from a slight simplification of the results. When we look at the rejection rates (the percentage of items in a determined category that received scores 1 or 2 in the AJT), we see a bigger difference between groups regarding TSCs than *in situ* interrogatives. The experimental group rejected 10% of *in situ* interrogatives while the control group rejected 3% of them. But the rejection rate for target TSCs within the experimental participants was 57%, quite higher than the rate within the controls, which was 41% (the difference here is more than double the difference for *in situ* interrogatives). And if we look at the rejection rates for ungrammatical TSCs, we find that the rate is lower for the experimental group, 76%, than for the control group, 83%, which shows non-native speakers have a bit more difficulty than natives distinguishing grammatical from ungrammatical TSCs.

But to be able to answer the previous question more thoroughly, it was necessary to dig a little deeper. On one hand, the experimental group was divided into smaller subgroups by proficiency level (see the graph in (3), which summarizes these subgroups' behavior); on the other hand, the individual consistency of the participants was also studied (see the graph in (4), which

already appeared in Chapter 5, go back for a detailed explanation of how consistency was measured). Once this was done, the picture became clearer.

(3)   AJT scores – Experimental subgroups by proficiency level
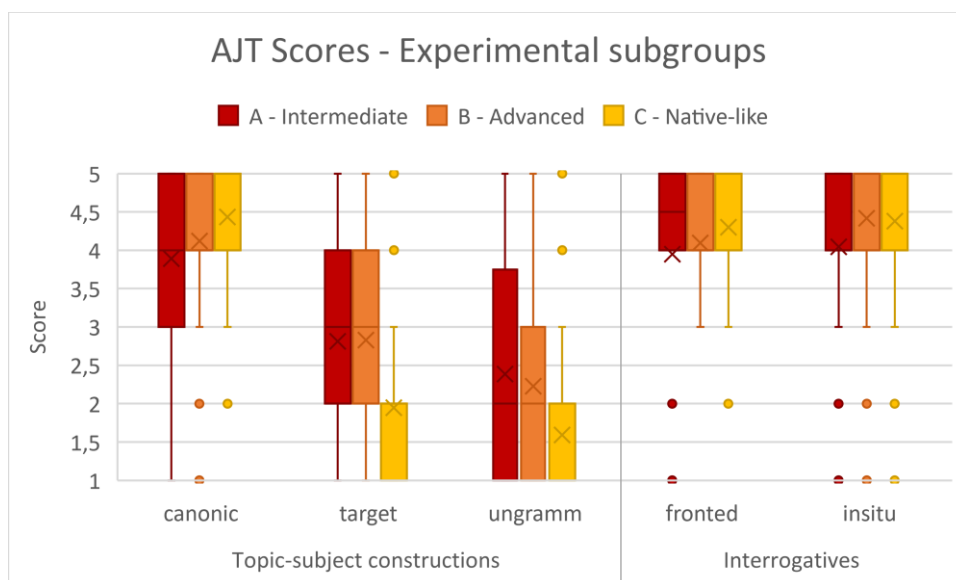
(4)   Consistency levels (on a scale from 0 to 3) by group and type of sentence
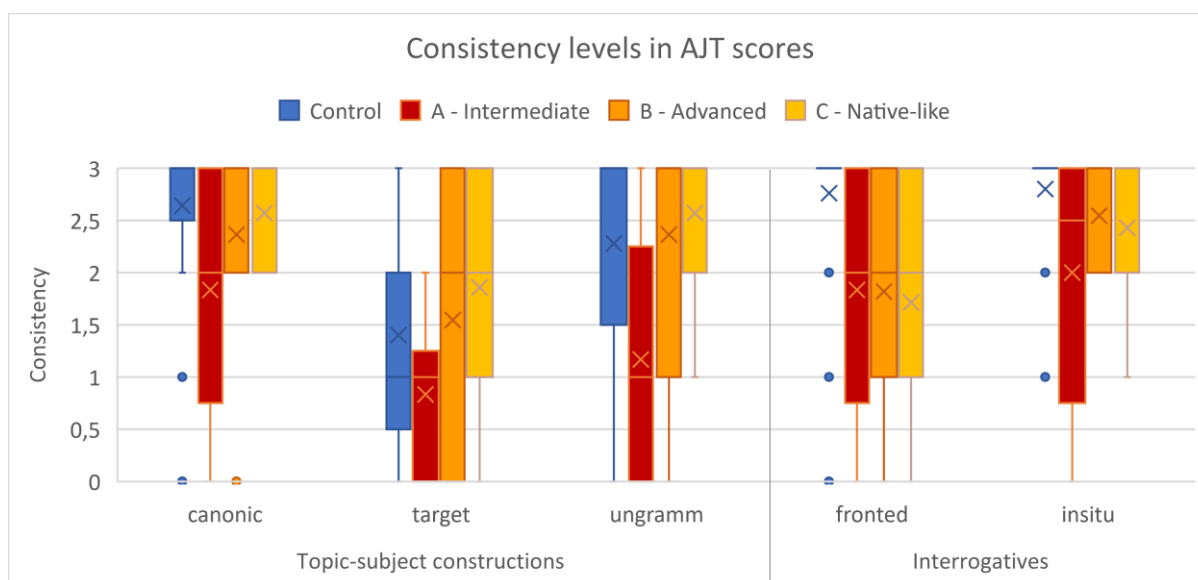
Note that subgroup A, the least proficient one, behaved towards TSCs exactly as I was saying before, that is, showing uncertainty about these structures rather than rejection. This can be inferred both from their scores and their consistency levels. Look at how data dispersion is quite high for this subgroup's scores of TSCs and also how their consistency levels are the lowest of all groups, especially so for target TSCs. This suggests their ratings are almost random, which in turn, suggests they have not acquired this syntactic structure.

On the other hand, note that, despite having low consistency levels for interrogative sentences, subgroup A gave quite high ratings to these sentences, 4 or 5, except for a few isolated cases. This suggests wh *in situ* is acquired quite quickly[46], since even the participants that had been in contact with BP for only a couple of months showed remarkably high acceptance of these sentences.

Now let us focus on subgroup C, which was formed by experimental participants with native-like proficiency in BP. With respect to TSCs, this subgroup rated target and ungrammatical versions as clearly unacceptable and did so with conviction, since their consistency levels for these structures were even higher than those in the control group. Hence, it can be inferred from this, that Spanish native speakers do not acquire TSCs, not even after having spent a considerable amount of time living in Brazil and immersed in the language.

As for interrogatives, subgroup C rated both fronted and *in situ* questions almost like the controls did. This subgroup and subgroup B gave higher ratings to these sentences than those attributed by subgroup A. This suggests that the acquisition of wh *in situ* is solidified with the increase in proficiency level, which is confirmed by higher consistency levels for *in situ* interrogatives within these two subgroups, B and C. Consistency for fronted interrogatives did not improve with proficiency level, though. However, recall that qualitative analysis of the follow up task confirmed inconsistencies and low ratings for fronted interrogatives were due to the presence of a doubly filled Comp in some of the sentences, as I previously discussed.

These conclusions are further supported by the results from the bilingual participants. These were people that had all moved from Spain to Brazil over 30 years ago, one of them was middle aged but the other three were young children when they moved. All of them consistently gave high scores to *in situ* interrogatives. As for TSCs, two of them clearly rejected them and were as consistent about it as they were with interrogatives. The other two were not as consistent, but still generally gave low scores to these sentences, except for one or two (those corresponding to more crystallized expressions).

Our qualitative analysis of the follow-up task also seems to support the conclusions drawn so far. Rephrasing of the rejected *in situ* interrogative sentences was more consistent and similar

---

[46] Recall from Chapter 1 that I am assuming different syntactic analyses for wh *in situ* in Spanish (apparent *in situ* from Etxepare and Uribe-Etxebarria, 2005) and in Brazilian Portuguese (true *in situ*, from Grolla 2005). If we were to assume the same kind of syntactic analysis for these interrogatives, the study results could then be interpreted as the non-native speakers applying their L1 grammar to L2 data, which could also explain high acceptance of the structure even with very low proficiency levels. This will be further discussed in section 7.4.

amongst the groups than that of target and ungrammatical versions of topic-subject constructions. It was in the rephrasing of these sentences where we saw the experimental participants diverge more easily from the main amending strategies used by both groups.

Having all of this into consideration, I strongly believe the study results sustain the proposed hypothesis. There is indeed a significative difference in the acquisition of BP *in situ* interrogatives and topic-subject constructions by Spanish native speakers: wh *in situ* is quickly acquired whereas topic-subject constructions not only are not acquired, but are increasingly rejected as proficiency level rises.

7.3. Prescriptive grammar effects *vs*. natural rejection

Despite firmly believing the proposed hypothesis does indeed hold, I cannot help but wonder how much impact prescriptive grammar had on this experimental study.

The results clearly show the rejection rate for topic-subject constructions increased with proficiency level within the experimental group. I attributed this to natural rejection of these syntactic structures by Spanish native speakers, and concluded they failed to acquire them, even after years of cultural and linguistic immersion in the country (as the bilinguals' results showed). I stand by this claim. However, it would be naïve to ignore completely the effects prescriptive grammar could have had on the results.

In chapter 6, I discussed how the specific proficiency test used in this research might be a bit more tied to prescriptive grammar than would have been ideal (see section 6.2.2). I also talked about the shortcomings of the chosen experimental paradigm, which can lead the participants into a prescriptive frame of mind, due to its metalinguistic nature (see section 6.3.2), as some of the follow-up responses confirmed. And more importantly, in the participants' description in chapter 4, I noted that most of them had acquired BP through formal training rather than in a more naturalistic manner, and that some of them were daily exposed to the language only a 30% of their time or less (see section 4.3).

The combination of all these factors may give rise to a different interpretation of the study results, at least regarding topic-subject constructions. The increase in rejection rates for TSCs within the experimental group as the proficiency level rises could be interpreted as a consequence of their formal training and the prescriptive effects this may entail. In other words, one could argue that what the results are showing is simply that the more the amount of formal

training (which corresponds to higher scores in the proficiency test), the higher the rejection rates for TSCs.

However, it is crucial to bear in mind that this syntactic structure is rarely discussed in traditional grammar books or Portuguese textbooks, either for native or non-native speakers. In fact, topic-subject constructions are not stigmatized like other BP vernacular constructions (such as those involving nominal agreement, clitic position, tonic pronoun in object position, use of relative pronouns, etc.). Although TSCs clearly violate syntactic rules of what prescriptivism considers "good" Portuguese, language purists have not payed much attention to these structures. Therefore, they constitute a somewhat unconscious violation of prescriptive grammar. Would Portuguese teachers (as L1 or L2) actually correct TSCs when they come up in student's homework? If this construction is part of the teacher's own grammar, chances are they will probably not even notice them.

Moreover, some of our experimental participants acquired Brazilian Portuguese naturalistically instead of through formal training and still rated TSCs mainly as unacceptable. And the four bilingual participants, who had learned BP naturalistically, did not like TSCs either, even after many years in the country. These facts point towards a deeper reason behind TSCs rejection rates in these groups.

In addition to this, none of the L2 Portuguese textbooks I checked made any specific reference to *in situ* interrogatives either, and the teachers I consulted confirmed this is not a topic taught in class. Yet our Spanish native speakers had no trouble dealing with this structure, as opposed to the difficulties they showed regarding TSCs.

Let us focus now on the control group. Since proficiency level is not a tool that can be used with native speakers in the same way it is used with non-native speakers, the argument linking proficiency with rejection towards TSCs and attributing this only to the influence of formal training cannot be applied to the control participants. This does not mean, though, that prescriptive grammar had no effect at all on the control group. The effect is hard to dismiss completely, considering the sociocultural context of the country and seeing firsthand the slight anxiety these participants showed when presented with a proficiency test in their own mother tongue.

But, in this case, I believe there are far more important factors at play. To begin with, there is the above-mentioned fact that TSCs are not yet on the prescriptivism radar, and native speakers who use them do not perceive there is any grammar violation in them. But also, if it were only

a matter of linguistic prescriptivism, I would have expected more consistency from control participants in their ratings of these syntactic constructions (and lower ratings as well). Given how variable the ratings were for each individual, I think what the results show are the effects of a syntactic phenomenon under variation, as I will argue in section 7.5.

Unfortunately, with the information available to us, it is impossible to ascertain the real impact of prescriptive grammar effects and the extent to which it affects our results. However, although it is hard to deny any kind of prescriptive influence whatsoever, I find it is much harder to support the opposite, i.e., that the obtained results are only (or even mainly) due to the influence of prescriptive grammar, as I have argued throughout this section.

7.4. Implications for SLA theories

Just to recap briefly, recall that the results of this experimental study support the hypothesis of a significative contrast in the acquisition of BP *in situ* interrogatives and topic-subject constructions by Spanish native speakers advanced as the departure point of this research. So, the next logical step to take, as scientists, is to ask ourselves why. In other words, what is this difference due to and how can it be explained? This will organically lead us to the study's implications for SLA theories.

One possible answer to the questions above is to account for the contrast in terms of frequency. Frequency-based research on SLA has gained strength during the last decades, due in part to the introduction and recent popularity of neural network approaches. Some authors argue that frequency has a major role in SLA and that second language learning is mainly driven by input (Ellis 2002).

Applying this line of thinking, it could be argued that wh i*n situ* is acquired earlier than TSCs simply because it appears more often in the input that L2 learners receive. First of all, to support such a claim, it would be necessary to conduct a corpus-based analysis that showed *in situ* interrogatives occur in fact more frequently than TSCs. But even if this turned out to be true (which we do not know as of yet), the problem with this kind of argument is deeper. Comparing two different syntactic structures in terms of frequency can be problematic. One structure may need a higher rate of occurrence than the other to trigger the necessary changes in the interlanguage grammar to be considered acquired by the speaker. And this is without even considering that other structures may interfere, which, assuming the analysis in Avelar and

Galves (2011, 2016), is probably the case of TSCs. Since T's EPP phi-independence is also responsible for hyper-raising and ambiguity in tough constructions (among other BP peculiarities), the occurrence of these kinds of sentences in the input may contribute to the implementation of that change in the grammar, and hence, to the acquisition of TSCs. And on top of that, frequency does not exclude the influence of other factors, which could have a stronger role in the acquisition process, as the poverty of the stimulus argument suggests. There might be frequency effects, but it is extremely hard to prove that they are the true engine behind the acquisition process.

An explanation for the difference in acquisition of our two type of sentences that seems much more plausible comes from their syntactic structures.

Remember that BP *in situ* interrogative sentences were extensively accepted within the whole experimental group, and their scores slightly increased with higher proficiency level. This can be considered as evidence of the early acquisition of this BP structure by Spanish native-speakers.

According to Etxepare and Uribe-Etxebarria's analysis (2005) for Spanish, wh *in situ* in this language is highly restricted and only apparent. It involves overt movement of the wh constituent to a focus position and subsequent remnant movement to obtain the final word order. This contrasts with BP *in situ* interrogatives if we adopt a no movement analysis like Grolla's (2005), for example (see Chapter 1 for a more detailed explanation). Assuming Chomsky (1995), this means that, in Spanish, the uninterpretable wh-feature in C is strong, whereas in BP, the strength of this feature can be optionally relaxed. So, for a Spanish native speaker, acquiring BP's true *in situ* interrogatives would imply modifying the strength setting of this uninterpretable formal feature in his/her L2 grammar.

It can be argued, then, that our results support Full Access approaches like the one defended by Hettiarachchi and Pires (2016), and it is not incompatible either with the Feature Reassembly hypothesis advanced by Lardiere (2008, 2009). And, since the L2 learner seemed to be able to access UG and modify the strength of a formal feature, this can be considered as evidence against Representational Deficit approaches (such as Hawkins and Chan's Failed Functional Features Hypothesis, 1997, or the Interpretability Hypothesis adopted in Hawkins and Hattori, 2006). But hold this thought for a moment, I will come back to this later.

Now recall how the experimental group reacted to topic-subject constructions. The least proficient participants showed very inconsistent and variable behavior towards target and

ungrammatical TSCs (and even towards their canonical versions). The participants in the middle also showed variation in their ratings for TSCs, their behavior towards target TSCs was very similar to that of the previous group, but their ratings for ungrammatical TSCs were lower and those for canonical versions were higher. They were less inconsistent, or rather, there were some individuals with more consistent behavior in this group. Finally, the near native participants consistently rejected both target and ungrammatical TSCs but accepted their canonical versions. Overall, this suggests that TSCs have not been acquired and that a reanalysis of these structures using the participants' L1 grammars did not result in an acceptable syntactic structure either.

Since Spanish is clearly a subject-prominent language, these results combined with the qualitative analysis of the follow-up task seem to side with the idea that interlanguage grammars show typological transfer from the L1, as defended by Jung's study (2004) about SLA of topic prominence, rather than with the opposing side of the debate, represented by Fuller and Gundel (1987), who argue in favor of a universal early topic-comment stage.

Avelar and Galves (2011, 2016) claimed the position of Spec-TP in BP is inserted before Comp and without Case features, and they proposed the phi-independence of T's EPP to account for topic-subject constructions and several other BP syntactic phenomenon regarding the subject position, subject-verb agreement and pronouns (go back to Chapter 2 for a thorough discussion on this topic).

Spanish (and most Romance languages too) is more restrictive on what can occupy the subject position and on the formal features that a constituent must have to trigger verbal agreement. As discussed by Fernández-Soriano (1999), there are some impersonal verbs in Spanish that allow locatives to occupy the subject position (Spec-TP, to be more precise), but they never trigger verbal agreement and they must have a preposition (*En Barcelona ha ocurrido un accidente*, 'In Barcelona, there has been an accident'). The need for a preposition is explained by the impossibility to insert DPs without a Case feature in this language (unlike in BP). To justify that these locatives do not agree with the verb despite occupying the subject position, the author proposes that a non-nominative constituent is able to satisfy T's EPP, while nominative Case is assigned to a postverbal DP that triggers verb agreement. In other words, that, in Spanish, T's EPP feature and its phi-features agreement/nominative Case assignment can be fulfilled by different elements.

Note that this is not the same as claiming the phi-independence of T's EPP and the possibility of inserting DPs without a Case feature, that is a much stronger claim. So, although in some very specific cases, this language allows T's phi-features to be satisfied by an element that is not the same satisfying its EPP feature, T's EPP feature is phi-dependent in Spanish (unlike in BP).

Therefore, our results suggesting that non-native speakers acquired *in situ* interrogatives but did not acquire TSCs can be explained by accounting for this contrast drawing upon the differences behind their syntactic structures summed up above. That is, claiming that the mere modification of a shared formal feature's strength value (from strong to optional) constitutes a very mild change in comparison to a change that implies going from a phi-dependent T's EPP to a phi-independent one, and which would set off a number of other changes in the L2 grammar.

I personally find this explanation to be the most satisfactory, but there is yet another way of looking at the contrast between wh *in situ* and TSCs: blaming variation.

From the results obtained in this study, topic-subject constructions do not seem to be uniformly implemented in BP native speakers' grammars, which means TSCs are a syntactic phenomenon under variation. According to DeKeyser (2012, p. 58) when native speakers show variation regarding the structures under study, "*the native and non-native ranges of variation are almost bound to overlap substantially, whether for the same reasons or not*". SLA research on variation is usually focused on the variation found in L2 grammars rather than exploring the effects of input variation. But research on L1 acquisition shows that input variation can delay the child's acquisition (see Miller and Schmitt 2012), which could be extrapolated here to SLA.

So, the fact that BP native speakers show variation regarding TSCs may be the reason behind the different behavior shown by our participants towards these structures compared to wh *in situ*. As I see it, there are two problems with this argument. First, variation in the input on TSCs did not exactly result in variation in the L2 grammars because non-native speakers with high proficiency clearly rejected these structures. It was only those with lower proficiency level who showed variable behavior revolving TSCs, and given the extreme inconsistency demonstrated by these participants (even for interrogative sentences) this variability seems to be better accounted for as a consequence of other factors, like lack of self-confidence in the participants' own linguistic knowledge of the L2 (they seem to be answering randomly).

And second, the *in situ* strategy seems to be licensed under specific discursive and/or pragmatic circumstances (Pires and Taylor, 2007; Figueiredo Silva and Grolla, 2016), but it is never compulsory, fronting always remains a possibility. Hence, there is some variation here as well, it is not the same as with TSCs, granted. But it is a variation that had no impact on our experimental participants whatsoever.

Therefore, I strongly believe the contrast in acquisition between wh *in situ* and TSCs is best accounted for by attributing it to the former being associated with the implementation of a far less drastic change in L2 grammar than the latter.

Let me resume now the discussion I left half-baked before. Since wh *in situ* does exist in Spanish, although it is much more restricted than in BP, we cannot be completely sure that the experimental participants were not applying their L1 grammar to account for these structures. I do not believe this was the case, because if it were, experimental and control behavior towards *in situ* interrogatives could not have been as similar to one another. But the hint of a doubt remains.

If it were so, then our results would constitute support for Representational Deficit approaches (such as Hawkins and Chan's Failed Functional Features Hypothesis, 1997; or the Interpretability Hypothesis adopted in Hawkins and Hattori, 2006). From this perspective, the wh-feature is subject to a critical period and its strength cannot be modified. This constrains us to assume BP learners' mental representations of *in situ* interrogatives do not involve a true *in situ* structure, but a reanalysis that accommodates the syntactic structure to the L1 grammar. That is, for example, an analysis of BP wh *in situ* in the same lines as the one proposed by Etxepare and Uribe-Etxebarria for Spanish (2005): the wh constituent is overtly moved to a focus position and the rest of the sentence undergoes remnant movement to the left periphery.

With the data at our disposal right now, it is not really possible to discern between the two possible interpretations of the *in situ* results (as sustaining Representational Deficit or Full Access approaches). One way to check what is in fact happening would be to see how Spanish native speakers react to violations of the Sentence Final Requirement (Uribe-Etxebarria, 2002), which claims the wh constituent in a Spanish *in situ* interrogative must appear last in the sentence. If they apply this restriction to BP wh *in situ*, it would be an indication that non-native speakers are analyzing BP data under L1 grammar properties, hence backing Representational Deficit approaches. If they do not, this would show they had acquired the structure and

implemented its subtle syntactic properties in their L2 grammars, tipping the balance in favor of Full Access approaches instead.

To conclude the discussion on our study's implications for SLA research, there are a couple of perhaps more tangential conclusions that can be inferred from the results.

Recall again that highly proficient non-native speakers rejected TSCs, this could also be interpreted as supporting the idea that the L2 grammars of non-native speakers in their steady (or end) state might diverge from the grammar of a native speaker, and that they might do so precisely due to the influence of the L1. Remember that extremely high proficient participants, some of which had lived in the country for several years (more than 30 years in the case of the bilingual participants) still rejected TSCs almost categorically. But let me underline something, the fact that our most proficient participants L2 grammars diverged from those of native speakers is only evidence that end state grammars *can* diverge from native grammars, not that they *must* do so, i.e., that convergence is not possible. That is a much stronger claim, one with which I do not agree and which I believe is extremely hard to prove. The possibility of divergence, however, is compatible with the Full Access Full Transfer hypothesis (introduced by Schwartz and Sprouse, 1994) and also with the Feature Reassembly hypothesis (proposed by Lardiere, 2008, 2009).

Still keeping our focus on the study results regarding TSCs, there is another interesting detail I would like to comment on. As noted in chapter 5, during the discussion of the study results, there was a small asymmetry between locative and genitive TSCs. Experimental participants showed a slight preference for genitive TSCs, which went against my expectations. I believed Spanish native speakers would have disliked more this type of construction than the other, since an analysis of BP genitive TSCs in terms of L1 grammar hinders their semantic interpretation, something that does not happen with locative TSCs. Applying L1 grammar, the topic-subject constituent can be interpreted as having agency over the verb, which alters the sentence meaning (sometimes to a degree that can make it incomprehensible).

However, in light of the results (in particular the follow-up's qualitative analysis), it seems that apparent similarities were what "activated" the participants' L1 grammar more easily. Locative TSCs superficially differ from their Spanish semantic equivalents "merely" in that the topic-subject constituent lacks a preposition and it agrees with the verb, which is apparent only when the semantic subject does not share the same grammatical number. In comparison, the

difference for genitive TSCs is sharper, it involves reorganizing constituents and adding a pronoun *se* (which alters the verb's argumental structure). Compare (5) and (6) below.

(5)   i. *Aquele quarto só cab**e** duas pessoas.*   (BP locative TSC)

            that room only fit.3SG two people

       ii. ***En** ese cuarto sólo cab**en** dos personas.*   (Spanish semantic equivalent)

            in that room only fit.3PL two people

      'That room only fits two people / Only two people fit in that room.'

(6)   i. *<u>A mesa</u> quebrou <u>o pé</u>.*   (BP genitive TSC)

            the table broke the feet

       ii. *<u>La pata</u> **<u>de</u>** <u>la mesa</u> **se** ha roto*.   (Spanish semantic equivalent)

            the leg of the table SE has broken

     'The table leg broke.'

Therefore, our data can be interpreted as evidence in favor of apparent similarities being a factor contributing to transfer from the L1. This goes in line with the idea of psychotypology playing an influential role on the source of transfer in SLA. Psychotypology is a concept introduced by Kellerman (1983) and it refers to typological proximity as perceived by the learner, not as objectively determined. Although the relationship between psychotypology and cross-linguistic influence is usually applied in the context of L3 acquisition (Ionin et. al, 2015; Santos, 2013; Rothman, 2011; among many others), it is also relevant for a discussion in terms of L2 acquisition.

7.5. Implications for pre-existing L1 analyses

This research was conceived within the framework of SLA, but that does not mean its conclusions have to be necessarily restricted to that field. In this dissertation's preliminaries, I theoretically discussed the proposals put forth in the literature to account for the two BP syntactic structures under study. Now, I want to explore the ways in which our results could contribute to these analyses, supporting one in favor of another when there are contradicting proposals, providing hints for new possibilities, etc.

7.5.1. Interrogatives

Recall from chapter 2 that the proposals regarding BP wh *in situ* took two different approaches: one assuming a movement analysis and one assuming a true *in situ* analysis. The movement approach was adopted by Kato (2013), who proposed that the wh constituent is only apparently *in situ*, arguing that it moves to an internal focus position and the final order is due to remnant movement of the rest of the sentence. Grolla (2005), on the other hand took the *in situ* approach, using child acquisition data as argument in favor of her proposal.

I would like to remind the reader that the results of our on-line measurement of reaction times showed a significative difference between *in situ* and fronted interrogatives, fronting taking longer than leaving the wh constituent *in situ*. From a language processing point of view, these results suggest that in BP *in situ* interrogatives, the wh constituent does in fact remain *in situ*. Hence, our study seems to offer support to the no movement approach. Nevertheless, we should keep in mind that the connection between the reaction time to assign an acceptability score to a given sentence and the sentence syntactic structure is not straightforward, so the theoretical inference above should be treated with caution.

The results regarding interrogative sentences questioning the subject in postverbal position were not as determinant as I expected, but they also seem to point to a true wh *in situ* in BP.

These interrogatives were included in the experimental study to find evidence in favor of a no movement approach. The idea was to compare this kind of sentences with unergative and unaccusative verbs, with the intuition that participants (at least native speakers) would reject the former but accept the latter. Since the subject is the internal argument for unaccusative verbs, in a sentence like *Chegou quem?* ('Who arrived?'), under a no movement approach, the wh element *quem* has remained *in situ*. Whereas with an unergative verb, the subject is the external argument, so a sentence like *Telefonou quem?* ('Who called?') should be ungrammatical. Under a movement analysis like Kato's (2013), both sentences would be grammatical since the wh constituent moves to an internal focus position and then remnant movement of the verb creates the apparent *in situ* structure.

Our results showed a significative difference between the unergative and unaccusative conditions for experimental and control participants together, but there was no interaction between the type of verb and the group. However, looking at the control results, there was a visible difference, as can be seen again in the following graph, and a post-hoc test revealed that significance was almost reached. However, considering the number of these sentences was quite

small, these results are not as robust as is desirable, and they should be interpreted having this in mind.

(7)   AJT scores for interrogatives questioning the subject in postverbal position – Control group

Control Group
Postverbal Subject Qs
(*Nadou quem?* vs. *Chegou quem?*)
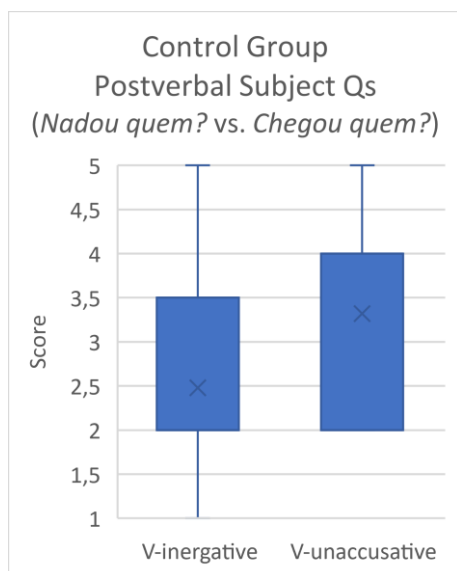
So, from our study results, we were able to determine two pieces of evidence pointing in the direction of a no movement analysis. Yet, the theoretical implications drawn from them do not stand on completely solid ground. Therefore, further research is needed to settle the ongoing debate between the *in situ* and movement approaches to account for the BP wh *in situ* syntactic phenomenon.

7.5.2. *Topic-subject constructions*

Perhaps the most interesting byproduct of this experimental research comes from the behavior towards topic-subject constructions displayed by BP native speakers.

Recall that our results revealed there was a great deal of variation on native speakers' ratings of TSCs. There was variation between individuals, ones being more accepting than others; but also within the same individual, almost each one accepting some TSCs while rejecting others.

A summary of these results is offered below. The graph on the left shows the scores attributed to TSCs by the control group in the AJT (on the left). The one in the right represents the individual consistency level for each type of sentence in a scale from 0 – totally inconsistent to

3 – totally consistent. Consistency levels for interrogative sentences were included for comparison. (See section 5.3 for a detailed explanation of how consistency was computed).

(8)   Control group – Results summary
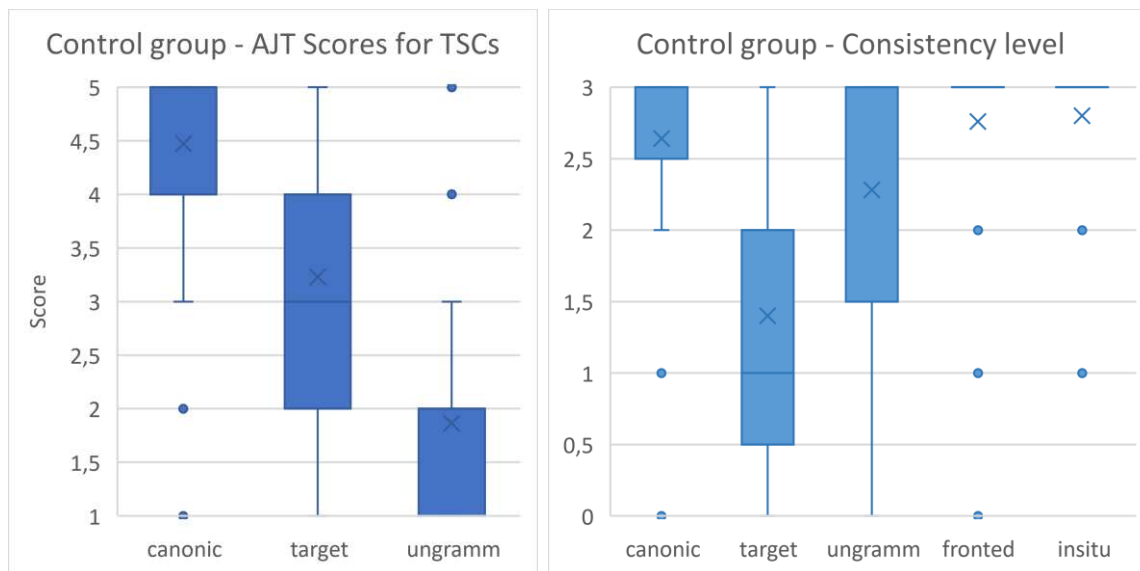
Two different implications arise from those results. The fact that some participants tended to reject these structures implies TSCs do not belong to all native speakers' grammars (yet?). The fact that some speakers mainly accepted TSCs and others were very inconsistent in their judgments towards these sentences imply these structures are not being acquired in a uniform manner (i.e., with the same set of properties), which, in turn, indicates TSCs are a syntactic phenomenon under variation in BP.

This leads us to a remarkably interesting question. What are the restrictions responsible for this unsteady behavior? In other words, what are the differences between a grammar that fully accepts TSCs, a grammar that fully rejects them, and one that falls in between, accepting some and rejecting others? Diving deeper into this discussion, other questions come up. What characteristics distinguish an acceptable TSC from an unacceptable one in one of those 'fluctuating' grammars? And, are these characteristics always the same for all these grammars or do they vary from one to another? Or could it be possible for our results to be explained appealing to Yang's idea of competing grammars (Yang, 2000)? These are all fascinating questions to which, unfortunately, this study cannot provide answers. More research is needed to understand the place that TSCs occupy within native grammars.

What we can do now, though, is speculate and put forward some hypotheses. I would like to start with an attempt to guide the discussion concerning the first question. According to Avelar

and Galves' proposal for TSCs (2011, 2016), a grammar that implemented the change necessary to accommodate all these structures is one in which the EPP in T is phi-independent and in which DPs can be inserted in the derivation without Case features. As the authors argue, this change would not only account for TSCs but for other BP syntactic singularities as well, like verb agreement with non-argumental DPs, hyper-raising, ambiguity in tough-constructions, variation in subject-verb agreement, and morphological uniformity between pronouns in nominative and non-nominative positions (check chapter 2, section 2.4.3, for more details). Assuming this proposal, a grammar that only sanctions some TSCs and not others could not be like one in which all these constructions are acceptable. In such a grammar, T's EPP could not be phi-independent, but then, how could the accepted TSCs be accounted for? Maybe in those cases, the acceptability of some TSCs derives from some change in the grammar at the v-VP level. This would not constitute such a substantial change and it would explain why a grammar like this is less permissive regarding the type of verb that licenses TSCs than one which has implemented the phi-independence of T's EPP. Finally, a grammar that consistently rejects TSCs could be similar in this respect to Spanish native grammars, allowing T's EPP and its phi-features to be fulfilled by two different elements, as Fernandez-Soriano (1999) suggests. This would imply that locatives and genitives in the corresponding constructions would need to be inserted with a preposition (to ensure no DP remains without Case features), and that the verb would have to agree with a postverbal DP with the relevant features. Only one of our control participants showed categorical rejection towards these constructions, but the rephrasing in his follow-up responses followed to the letter the above-mentioned restrictions.

To offer a more concrete hypothesis regarding the kind of grammars that accept some TSCs and reject others, it would be necessary to understand what properties underlie this distinction and to ascertain whether the distinction is the same for all of those grammars or not, that is, we would need to answer the other set of questions I mentioned above. But tackling those questions requires careful examination of the accepted/rejected constructions in search for patterns that could enlighten the answer, and it demands further investigations to test those patterns (if found). A possible pattern to check would be the contrast between locative and genitive TSCs, since our results revealed a slight preference for the locative constructions within the control participants. Research focused on this matter should also include the syntactic phenomena related to agreement and Case that Avelar and Galves used as evidence in favor of their proposal. If native speakers who fluctuated in their judgement of TSCs do the same with these

other structures (or directly reject them), this would support the hypothesis entertained in the previous paragraph regarding this kind of grammars.

Analysis of TSCs reaction times also opens the door to interesting discussions. Recall that reaction times for these structures were significatively longer than for interrogatives. This was so even for fronted interrogatives, which include a long distance dependency (a filler-gap dependency between the wh element and the place where it is semantically interpreted), a characteristic that is known to increase the sentence's processing cost. So, what does the difference in reaction times come from? Is there a dependency of this kind in TSCs as well? It is true (as discussed in chapter 6) that in this experimental study, items designed to test TSCs were a tad longer in number of syllables than those designed to test interrogatives. So, the difference could simply come from that, but it could also come from TSCs syntactic structure, we cannot be sure. In any case, expanding the study of these sentences to the field of psycholinguistics would give us more information on the way native speakers process them, which, in turn, could offer some insight into their syntactic structures.

For example, an eye tracking experimental study could provide evidence in favor of Andrade and Galves' proposal (2014) and further information to account for the licensing of TSCs. These authors argue that, for a TSC to be grammatical, there must be a semantic relation between the genitive or locative topic-subject constituent and the theme (part/whole or inalienable possession for locative TSCs and content/container for genitive TSCs) and that they are inserted in the derivation as a small clause selected by the verb (again, check chapter 2, section 2.4.2 for more details). An eye tracker experiment could confirm this relation if after looking at the theme, the participants go back to look at the topic-subject constituent. Or it could tell us something about the verb's role in this. Or perhaps it could explain the slight contrast between locative and genitive TSCs present in this study's results. The experiment could also compare target TSCs with their canonical versions and with ungrammatical versions built by using locatives/genitives without that semantic relation with the theme.

7.6. Next steps for future research

Given the complexity of studying two syntactic structures as different from one another as the ones on which this research is focused, the first thing that should be done in future work on this topic is to divide the problem into smaller pieces, taking one step at a time.

What I mean by this is to start, for instance, by conducting a thorough experimental study only about the L2 acquisition of BP interrogatives. A study that, in addition to the fronted vs. *in situ* contrast tackled here, would also take into consideration any other subtleties that have escaped me in this study, for example, the insertion of an extra complementizer (doubly filled Comp) or the lack of subject verb inversion.

Or a study that placed its focus only on wh *in situ* and their L2 acquisition, but that took context into consideration. Some authors argue that licensing of wh *in situ* is not only a matter of syntax but also of discursive/pragmatic context (see Biezma 2016 for Spanish, Pires and Taylor 2007 or Figueiredo Silva and Grolla 2016 for BP). Interlinguistic differences regarding wh *in situ* can then be attributed to different conditions on the discursive/pragmatic context too. Since BP seems to be less restrictive than Spanish, it would be interesting then to compare how Spanish native speakers deal with wh *in situ* in different kinds of contexts in their L1 and in BP as their L2, and compare them as well with BP native speakers' behavior towards these constructions. I believe such a study could bring more information to the theoretical discussions on this topic for both languages as L1.

Going on the other direction, we could concentrate only on topic-subject constructions and study TSCs L2 acquisition. Such a study could then include other structures that usually fall under the umbrella term topic-subject construction but that were not considered for this research, for example, sentences like the ones below, which were mentioned back in chapter 4.

(9)   a. *A revista está xerocando.*

          the magazine is photocopying

          'The magazine is being photocopied.'

       b. *A cerca pintou.*

          the fence painted

          'The fence was painted.'

It should also include those other syntactic structures that Avelar and Galves (2011, 2016) considered to be a consequence of the same change in the grammar that TSCs are a consequence of, namely, T's EPP phi-independence.

My suggestion for both of these research proposals, either the study focusing on interrogatives or the one focusing on TSCs, would be to implement some of the empirical modifications that I proposed in chapter 6. That is, adding a small context to each of the experimental items so that the sentences themselves could be less complex, controlling sentence length in number of

syllables if reaction times are going to be part of the experimental design, combining different types of tasks to complement acceptability judgements such as comprehension and/or production tasks, taking into consideration prescriptive grammar effects (by analyzing textbooks and/or interviewing L2 teachers) and figuring out ways to try to avoid them (like devising a new kind of proficiency test), etc.

In addition to that, given how relevant the data collected from bilingual participants was, including a higher number of this kind of participants in subsequent studies could also prove to be quite relevant (at least for TSCs). Assuming Meisel (2008), there is a substantial difference between simultaneous acquisition of two first languages and successive acquisition of two or more languages. The term 'bilingual', however, is sometimes used to refer to people in both of those categories (especially when the successive acquisition of a second language was extremely successful). Meisel claims there is extensive research proving that simultaneous acquisition of two or more first languages is an instance of multiple L1 acquisition. Nevertheless, our three bilingual participants who had learned both languages as L1 showed a behavior towards TSCs much more similar to L2 learners than to native speakers. So, it would be quite interesting to see if this is a consistent pattern within a bigger sample of these speakers. And if it were, we could ask ourselves why: could this be due to attrition between the two languages, is T's EPP phi-independence especially susceptible to attrition?

Including those other 'bilingual' speakers, those who are actually very successful L2 learners, in the SLA research I was proposing above would also be useful, I believe. It could provide information about steady state grammars and about the quality of a change in the grammar that is so hard to implement for foreign speakers.

Yet another possibility would be to turn the focus towards native speakers and explore how they deal with topic-subject constructions. This is still a relatively new topic of study, at least from a comparative and/or experimental approach, and research could be done in many different directions and take many different forms.

TSCs could be studied from an experimental perspective following the lines of what I just suggested for SLA research, for example. An experimental study on native speaker behavior towards TSCs should also take into consideration the fine-grained differences that distinguish a grammar in which these constructions are fully acceptable and a grammar in which only some of them are accepted while others are rejected. And it should explore as well what properties

are involved and if they are always the same or they vary from speaker to speaker, as discussed in section 7.5.2.

They could also be studied taking a more psycholinguistic approach, delving into the processing of these syntactic structures. Following, for example the experiment I proposed at the end of the previous section. But perhaps what is most necessary right now is formal studies on the variation regarding TSCs, that is, research taking a sociolinguistic approach, since we suspect a strong influence of regional and socioeconomic factors.

On an even more ambitious note, though, there is an idea that (citing Machado de Assis, a brilliant Brazilian writer) one day hung herself up on the trapeze of my head, started jumping around and will not leave me until I decipher it or it devours me.

Recall that in the first version of the Minimalist Program (Chomsky, 1995), movements in the derivation were attributed to the strength of the formal features of the functional head receiving the moved element. If the feature is strong, the movement is 'visible' (it is morpho-phonologically realized); if the feature is weak there is no open movement and the feature is only checked in the covert component. Since natural languages differ in their use of syntactic movement, the strength of the relevant feature can be considered a parameter that reflects interlinguistic variation.

Assuming this approach to syntactic movement, the following hypothesis could be proposed:

(10)   *Hypothesis: In L2 acquisition, resetting a formal feature's strength from strong to weak is faster than the opposite.*

Support for this hypothesis comes from Ian Roberts' proposal (1999) about the 'weak' value being unmarked or being the default value. He says "*we see that weak features give rise to simpler representations than strong features, and so robust positive evidence is required for strong features, while weak features represent the default (or unmarked) value*" (Roberts, 2018, p. 129).

To test this far more general hypothesis, it would be necessary to conduct a much more ambitious research looking for L2 acquisition asymmetries between pairs of languages with a strong/weak contrast for some formal feature.

Fronted vs. *in situ* interrogatives in pairs of languages like English and Chinese and other typically *in situ* languages are the result of such an asymmetry and have already been extensively studied in SLA literature (Hawkins & Chan, 1997; Hawkins & Hattori, 2006;

Hettiarachchi & Pires, 2016; Choi, 2016; were those reviewed in this dissertation, but there are many others). Apart from this, subject-verb inversion or lack thereof in interrogatives in Spanish vs. BP, as exemplified in (11), can also be an example of that kind of asymmetry, and it has not been studied as much.

(11) a. *O que o Pedro comprou?*

   the what the Pedro bought

   b. *¿Qué ha comprado Pedro?*

   what has bought Pedro

  'What did Pedro buy?'

Another famous example comes from the pair English vs. French, also related to verb movement, but this time in declarative sentences like those in the examples below (from Pollock, 1989).

(12) a. *John likes not Mary.

   b. *Jean (n') aime pas Marie.*

(13) a. *John kisses often Mary.

   b. *Jean embrasse souvent Marie.*

   c. John often kisses Mary.

   d. *Jean souvent embrasse Marie.*

(14) a. *My friends love all Mary.

   b. *Mes amis aiment tous Marie.*

   c. My friends all love Mary.

   d. *Mes amies tous aiment Marie.*

These differences were the focus of discussion in Pollock's seminal paper, which studied the asymmetries from a theoretical perspective. But this pair of languages has also been the subject of many SLA studies (White, 1989, 1990, 1992; Hawkins et. al, 1993; Ayoun, 1999; Yuan, 2001; among others).

An in-depth study of the existing SLA experimental literature related to asymmetries of this kind, comparing and reinterpreting results if necessary, could be then complemented with an original experimental study. The central idea would be to explore thoroughly the impact of feature strength in second language acquisition. This is not a new idea, of course; however, to my knowledge, studies on this topic have been mainly conducted only in one direction or the other (not so much on both at the same time), but have not been compared between each other,

or at least, not exhaustively. I believe that a comparison in such terms could bring new light into the matter.

7.7. Concluding remarks

Taking an experimental approach, this research was set out to compare BP *in situ* interrogatives and topic-subject constructions in their second language acquisition by Spanish native speakers, from a generativist perspective. The main goals were to study the acquisition of the two structures, to confirm that it differs substantially, which was verified by our study results, and to consider the potential reasons behind this contrast.

The first step in this work was to understand better the syntactic structures underlying both wh *in situ* and topic-subject constructions, which transported us to purely theoretical fields in which we wandered through the different proposals regarding these constructions.

A review of the literature on *in situ* interrogatives led us to discuss Rizzi's *Q-Criterion* (1996), pointing out the conceptual problems with his idea of *dynamic agreement*, and to contrast it with Cheng's *Clausal Typing Hypothesis* (1991). This offered a deeper insight into the main proposals to account for the special properties of this syntactic phenomenon in BP, those defended in Mioto and Kato (2005), Grolla (2005), and Kato (2013). Since what underlies Mioto and Kato's proposal (2005) is Rizzi's *Q-criterion*, it inherits the same issues we questioned about Rizzi (1996), and their analysis was dismissed in favor of the others. The remaining proposals are representative of the ongoing debate between the *in situ* approach, adopted in Grolla (2005), and the movement approach advocated by Kato (2013), in which wh-phrases are overtly moved to a focus position, the final word order being a result of subsequent remnant movement. Based on a comparison between wh *in situ* in BP and in Spanish, assuming Etxepare and Uribe-Etxebarria's (2005) analysis for the latter, we argued in favor of Grolla's proposal. This author brings language acquisition and cue-based theories into the picture, suggesting that the language is first set to be [+wh movement] and later reconsidered as being [±wh movement] when the relevant cues appear in the input, making her analysis the one better suited to the SLA approach adopted in this research, and the only one embracing the optionality of BP wh *in situ*.

We then delved into the studies on topic-subject constructions, starting with the pioneer work by Pontes (1987), of course, but also with the, perhaps lesser known but brilliantly insightful,

Galves (1986). These authors brought attention to topic-subject structures and the contrast they marked between BP and other Romance languages, but the first concise proposals did not come until later, such as Galves' minimalist analysis (1998). The discussion in this dissertation, however, was mainly focused on three of the latest proposals advanced to account for these constructions' syntactic structure. First, we examined Munhoz and Naves' work (2012), where the effort is mainly aimed at the argumental structure of unaccusative verbs that license TSCs suggesting they are monoargumental for genitive TSCs and biargumental for locative TSCs (both arguments being internal). We challenged this analysis with arguments of our own, namely that it does not explain how the argumental DP in genitive TSCs is formed nor why it splits in two; and with arguments from Andrade and Galves (2014), who point out the opacity of the relation between TSCs and the verb's argumental structure under this proposal, and claim it fails to offer a unified analysis for both types of TSCs. We showed that this is precisely what Andrade and Galves (2014) do by suggesting the licensing of TSCs depends, not only on the verb being unaccusative, but also on it selecting a secondary predicate or small clause (*alla* Den Dikken, 2006), that syntactically encodes the necessary semantic relationship between topic and theme (part-whole for genitives, content-container for locatives). We noted that their proposal explains where and how the topic-subject originates in the derivation, but that it does not account for the place of arrival and the triggering of verb agreement. Here is where Avelar and Galves' work (2011, 2016) comes in. They locate the topic-subject constituent in Spec-TP and claim TSCs and other BP syntactic phenomena is accounted for by assuming this position is created without the action of phi-features and postulating this language allows DPs to be inserted without Case feature. Finally, we argued that these last two proposals are compatible and that, together, they provide a full explanation of TSCs' syntactic structure and its derivation.

Once the syntactic structures under study had been thus thoroughly examined, we went on to offer a general outlook of the main hypotheses put forth within the literature on SLA acquisition. Special emphasis was placed upon Representational Deficit approaches, under which the main claim is the impossibility to acquire functional features absent from the L1, and their contrast with Full Access approaches, which advocate that UG is fully accessible in SLA and hence functional categories or feature values that differ from those of the L1 can indeed be incorporated into the interlanguage grammar. This was done because the main hypotheses used in SLA research on the availability of the wh-feature fell under one of those approaches. Hawkins and Chan (1997) and Hawkins and Hattori (2006) respectively argue in favor of the Failed Functional Features Hypothesis and its refined version, the Interpretability Hypothesis,

claiming that L2 learners are unable to acquire the uninterpretable wh-feature, due to its being subject to a critical period, just like any other uninterpretable feature determining parametric cross-linguistic differences must be. Hettiarachchi and Pires (2016), on the other hand, replicate Hawkins and Hattori's study with a different L1-L2 pair and conclude the relevant uninterpretable feature was acquired by the L2 learners, providing evidence for Full Access to UG in SLA. Choi (2016) offers yet another possibility under this approach, arguing in favor of the Feature Reassembly Hypothesis proposed by Lardiere (2008, 2009). This proposal takes a new direction in SLA and tackles the problem in terms of selection and reorganization of the features that each language chooses from the universal inventory and associates with lexical items and functional categories in different ways. Topic-subject constructions had not been studied from an SLA perspective, so we discussed two contrasting proposals related to topic prominence in SLA, one arguing for an initial topic-comment stage (Fuller and Gundel, 1987), and another bringing evidence of typological transfer from the native language (Jung, 2004).

Based on this preliminary review of the relevant literature and on my own personal experience as a Spanish native speaker learning BP, a hypothesis regarding the difference in the acquisition of BP wh *in situ* and topic-subject constructions was formulated, postulating that the former is acquired much earlier than the latter by Spanish native speakers, who may even never incorporate these constructions into their L2 grammars. To test this hypothesis, we conducted an experimental study on a group of 24 Spanish native speakers, 26 BP native speakers and 4 bilingual participants, all of which were subject to a BP test that served to divide the experimental group intro three subgroups according to their proficiency level. Off-line measurements in the form of acceptability judgements on a 5-point Likert scale were combined with on-line information by measuring reaction times. Roughly half the experimental items were sentences meant to assess the participants' behavior towards *in situ* interrogatives compared to fronted ones, different argumental and non-argumental elements were questioned. The other half were designed to do the same for topic-subject constructions, comparing the target structure, that is, a locative or genitive TSC with an unaccusative verb, with ungrammatical versions of these structures built using unergative verbs, and also with canonical versions, in which the locative/genitive constituent appeared in a PP and did not trigger verb agreement. The acceptability judgement task was followed by a correction task where the participants had to rephrase the sentences that they had rejected, making them acceptable according to their own judgement. This was intended to understand the reasons driving the sentences' rejection and it proved to be an extremely useful tool.

The study results were analyzed in depth, by groups and individually, both quantitatively (using descriptive and inferential statistical methods) and qualitatively. From a global overview of the results, the experimental and control groups seemed to react to the test in a similar way, accepting both type of interrogatives and canonical TSCs, rejecting ungrammatical TSCs and being quite inconsistent about target TSCs. However, a comprehensive analysis revealed an altogether different picture. Non-native speakers' behavior was extremely dependent on their proficiency level. Least proficient participants showed great inconsistency towards TSCs (especially in their target versions) but not as much towards interrogatives, they varied widely in their ratings of the former but mainly accepted the latter. Native-like speakers, on the other hand, quite consistently rejected target and ungrammatical versions of topic-subject constructions. Ratings and consistency for TSCs within the middle subgroup was similar to that of least proficient speakers although their ratings and consistency for *in situ* interrogatives improved. These results suggest an early acquisition of *in situ* interrogatives whereas topic-subject constructions fail to be acquired and seem to be increasingly rejected as the proficiency level improves. Results from our bilingual participants and a qualitative evaluation of the follow-up task further supported our findings.

One could argue that these results may be due to prescriptive grammar effects instead. Notwithstanding the impact this may have had, it is not reasonable to consider prescriptive grammar or formal instruction (in the case of non-native speakers) as the main motive behind our findings, for several reasons. First, TSCs have not yet fallen under prescriptivism's radar, so they have not been stigmatized, constituting a rather unconscious violation. Second, Spanish native speakers acquiring BP in a naturalistic manner also rejected these constructions. And third, our experimental group as a whole had no trouble accepting *in situ* interrogatives although they are not a topic taught in class either. Therefore, we can conclude that our study sustains the proposed hypothesis, i.e., there is indeed a significative difference in the acquisition of BP wh *in situ* and topic-subject constructions by Spanish native speakers, in that the former are quickly acquired but the latter do not seem to be incorporated in L2 grammars in their steady state.

Once this contrast was verified, we turned to the discussion of its plausible explanations. Frequency was ruled out as a main factor, on the basis of the poverty of the stimulus argument, and because it is not valid on its own, it would require an additional corpus-based analysis. Input variation regarding TSCs was also discarded, since it only resulted in variation at the early stages of SLA, the construction being clearly rejected later on. Assuming the analyses discussed

in the preliminary theoretical study, we argued the distinction is best accounted for by attributing it to the differences between the two syntactic structures underlying the constructions under study. We suggested that the acquisition of one requires the implementation of a substantial change in the L2 grammar, namely T's EPP phi-independence and the possibility to insert DPs without Case features, whereas the necessary change for the acquisition of the other is less drastic, it only involves the loosening of a shared formal feature's strength value (from strong to optional).

We suggested, as well, that the study results regarding wh *in situ* can be considered evidence in favor of Full Access approaches and against Representational Deficit accounts of wh-feature SLA, because L2 learners were able to access UG and modify the strength value of this uninterpretable formal feature. However, it is not possible to determine from our data whether Spanish native speakers were analyzing *in situ* interrogatives in terms of their L1 grammar or according to BP native grammars' rules, so this SLA theoretical implication does not seem to stand on completely solid ground. As for the implications of these results for the debate between movement vs. no movement approaches in BP wh *in situ* anlyses, we argued reaction times of fronted compared to *in situ* interrogatives provided support for the *in situ* approaches and a slight preference for unacussative vs. unergative verbs in interrogative sentences questioning the subject in postverbal position also points towards the same direction.

Turning then to TSCs results, we claimed that interlanguage grammars show typological transfer from the L1. We also interpreted the slight preference for genitive vs. locative TSCs within non-native speakers as evidence of transfer from the L1 due to perceived similarities between the L1 and the L2, which is in line with the idea of psychotypology being an influential factor for L1 transfer in SLA. Regarding the implications for L1 theories on TSCs, our results indicate that these constructions are not ingrained in all native grammars, and when they are, they do not seem to have been acquired uniformly, which means we are dealing with a syntactic phenomenon under variation. This constitutes the most interesting byproduct of our research because it opens the door to fascinating discussions on the characteristics of the grammars that accept some topic-subject constructions while rejecting others, and on the restrictions that distinguish a grammatical construction from an ungrammatical one, which may vary from one speaker to another. So, I would like to finish this work by calling for further research that will help establish the place topic-subject constructions occupy in native grammars.

REFERENCES

ANDRADE, A.; GALVES, C. (2014). A Unified Analysis dor Subject Topics in Brazilian Portuguese. *Journal of Portuguese Linguistics*, v. 13.1, p; 117–147

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*, v. 3, p. 55–75.

AVELAR, J.; GALVES, C. (2016). From European to Brazilian Portuguese: A parameter tree approach. *Cadernos de Estudos Linguísticos*, v. 58 n. 2, p. 237–256.

_____; _____ (2011). Tópico e concordância em português brasileiro e português europeu. In: Costa, A.; Falé, I.; Barbosa, P. (Eds.). *Textos selecionados – XXVI Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa: APL. p. 49–65.

AYOUN, D. (1999). Verb movement in French L2 acquisition. *Bilingualism: Language and Cognition*, v. 2.2, p. 103–125.

BARDEL, C.; FALK, Y. (2007). The role of the second language in third language acquisition: The case of Germanic Syntax. *Second Language Research*, v. 23.4, p. 459–484.

BECK, M.-L. (1998). L2 acquisition and obligatory head movement: English-speaking learners of German and the local impairment hypothesis. *Studies in Second Language Acquisition*, v. 20, p. 311–48.

BELLETI, A. (2004). Aspects of the low IP area. In: Rizzi, L. (ed.). *The Structure of IP and CP. The Cartograph of Syntactic Structures*. Oxford: OUP, p. 16–51.

BIBERAUER, T.; RICHARDS, M. (2006) True Optionality: When the grammar doesn't mind. In: Boeckx, C. (ed.). *Minimalist Essays*. Amsterdam: John Benjamins, 35–67.

BIEZMA, M. (2016). *Givenness* and the difference between *wh*-fronted and *wh*-in-situ questions in Spanish. In: Berns, J.; Jacobs, H.; Nouveau, D. (Eds.). *Romance Languages and Linguistic Theory: Selected papers for Going Romance 29*. Nijmegen: John Benjamins, 2018. p. 21–39.

BLEY-VROMAN, R. (1990). The Logical Problem of Foreign Language Learning. *Linguistic Analysis*, v. 20.1-2, p. 3–49.

BRESNAN, J. (2007). Is Syntactic Knowledge Probabilistic? Experiments with the English Dative

Alternation. In: Featherston, S.; Sternefeld , W. (Eds.), *Roots: Linguistics in Search of Its Evidential Base*, Berlin, New York: Mouton de Gruyter. P. 75–96.

CASIELLES-SUÁREZ, (2003). Left-dislocated structures in Spanish. *Hispania*, v. 86.2, p. 326–338.

CHENG, L. (1991). *On the Typology of Wh-questions*. Ph.D. Dissertation, MIT.

_____; ROORYCK, J. (2000). Licensing wh-in-situ. *Syntax*, v. 3.1, p. 1–19.

CHOI, M. (2016). *The acquisition of wh-in-situ constructions in second language acquisition.* Ph.D. Dissertation, Georgetown University.

CHOMSKY, N. (2008). On Phases. In: Freidin, R.; Otero, C. P.; Zubizarreta, M. L. (Eds.). MIT Press. p. 133–166.

_____ (2001). Derivation by Phase. In: Kenstowicz, M. (Ed.) *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press. p. 1–52.

_____ (2000). Minimalist inquiries: The framework. In: Martin, R.; Michaels, D.; Uriagereka, J. (Eds) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: The MIT Press, p. 89–155.

_____ (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

_____ (1977). On wh-movement. In: Culicover, P.; Wasow, T.; Akmajian, A. (Eds.). *Formal Syntax*. New York: Academic Press. p. 71–132.

_____; LASNIK, H. (1977). Filters and Control. *Linguistic Inquiry*, v. 8, p. 425–504.

CLAHSEN, H.; HONG, U. (1995). Agreement and null subjects in German L2 development: new evidence from reaction-time experiments. *Second Language Research*, v. 11, p. 57–87.

CLYNE, M. (2003). *Dynamics of Language Contact: English and Immigrant Languages.* Cambridge: Cambridge University Press (Cambridge Approaches to Language Contact).

CORDER, S. P. (1967). The Significance of Learners' Errors. *International Review of Applied Linguistics in Language Teaching*, v. 5, p. 161–170.

COOK, V. (1986). The basis for an experimental approach to second language learning. In: _____ (Ed.). *Experimental approaches to second language learning*. Oxford: Pergamon Press. p. 3-21.

DE BOT, K.; WELTENS, B. (1995). Foreign Language Attrition. *Annual Review of Applied Linguistics*, v. 15, p. 151–164.

DeKEYSER, R. M. (2012). Interactions Between Individual Differences, Treatments, and Structures in SLA. *Language learning*, v. 62, Suppl. 2, p. 189–200.

_____ (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, v. 22, p. 499–533.

DEN DIKKEN, M. (2006). *Relators and Linkers: the syntax of Predication, Predicate Inversion and Copulas.* Cambridge, MA: MIT Press.

ELLIS, N. C. (2002). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, v. 24, p. 143–188.

ELLIS, R. (1991). Grammaticality Judgements and Second Language Acquisition. *Studies in Second Language Acquisition*, v. 13.2, p. 161–186.

EPSTEIN, S.; FLYN, S.; MARTOHARDJONO, G. (1998). The strong continuity hypothesis: some evidence concerning functional categories in adult L2 acquisition. In: Flynn, S.; Martohardjono, G.; O'Neil, W. (Eds.). *The generative study of second language acquisition.* Mahwah, NJ: Lawrence Erlbaum. p. 61–77.

ETXEPARE, R.; URIBE-ETXEBARRIA, M. (2005). In-situ wh-phrases in Spanish: locality and quantification. *Recherches linguistiques de Vincennes*, v. 33, p. 9–34.

EUBANK, L. (1993/1994). On the transfer of parametric values in L2 development. *Language Acquisition*, v. 3, p. 183–208.

FERNÁNDEZ-SORIANO, O. (1999). Two types of impersonal sentences in Spanish: Locative and dative subjects. Syntax, v. 2.2, p. 101–140.

FIGUEIREDO SILVA, M. C.; GROLLA, E. (2016). Some syntactic and pragmatic aspects of WH-in-situ in Brazilian Portuguese. In: Kato, M.; Ordóñez, F. (Eds). *The Morphosyntax of Portuguese and Spanish in Latin America*. Oxford: Oxford University Press. p. 259–285.

FLYNN, S.; VINNITSKAYA, I.; FOLEY, C. (2004). The cumulative enhancement model for language acquisition: Comparing adults and children's patterns of development in first, second and third language acquisition of relative clauses. *International Journal of Multilingualism*, v. 1, p. 3–16.

FOLEY, C.; FLYNN, S. (2013). The role of the native language. In: Herschensohn, J; Young-Scholten, M. (Eds.). *The Cambridge Handbook of Second Language Acquisition*. Cambridge: Cambridge University Press. p. 97–113.

FULLER, J. W.; GUNDEL, J. K. (1987). Topic-Prominence in Interlanguage. *Language Learning*, v. 37.1, p. 1–18.

GALVES, C. (1986). A sintaxe do Português Brasileiro. Work presented in *Simpósio sobre a Diversidade Linguística no Brasil*. UFB, Salvador.

_____ (1998). Tópicos e sujeitos, pronomes e concordância no Português do Brasil. Cadernos de Estudos Linguísticos, v. 34, p. 19–32.

GETHIN, A. (1990). Antilinguistics: A critical assessment of modern linguistic theory and practice. Oxford: Intellect books.

GOAD, H.; WHITE, L. (2006). Ultimate attainment in interlanguage grammars: a prosodic approach. *Second Language Research*, v. 22, p. 243–268.

_____; _____; STEELE, J. (2003). Missing inflection in L2 acquisition: defective syntax or L1-constrained prosodic representations? *Canadian Journal of Linguistics*, v. 48, p. 243–263.

GROLLA, E. (2005). Sobre a aquisição tardia de QU in situ em português brasileiro. *D.E.L.T.A.*, v. 21, p. 57–73.

GUIMARÃES, M. (2017). *Os fundamentos da teoria linguística de Chomsky*. Petrópolis: Vozes.

HAWKINS, R. (2005). Revisiting Wh-movement: The availability of an Uninterpretable [wh] Feature in Interlanguage Grammars. In: Dekydtspotter, L. et al. (Eds.). *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004)*. Somerville, MA: Cascadilla Proceedings Project. p. 124–137.

_____ (2003). *Representational deficit theories of (adult) SLA: Evidence, counterevidence and implications*. Plenary paper presented at EUROSLA, Edinburgh, Sept.

_____; TOWELL, R.; BAZERGUI, N. (1993). Universal Grammar and the acquisition of French verb movement by native speakers of English. *Second Language Research*, v. 9.3, p. 189–233.

_____; CHAN, C. (1997). The partial availability of Universal Grammar in second language acquisition: the failed functional features hypothesis. *Second Language Research*, v. 13, p. 187–226.

_____; HATTORI, H. (2006). Interpretation of English multiple wh-questions by Japanese speakers: a missing uninterpretable features account. *Second Language Research*, v. 22.3, p. 269–301.

HAZNEDAR, B.; SCHWARTZ B. D. (1997). Are there optional infinitives in child L2 acquisition? In: Hughes, E.; Hughes, M.; Greenhill, A. (Eds.). *Proceedings of the 21st Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press. p. 257–68

HERSCHENSOHN, J. (2013). Age-related effects. In: Herschensohn, J; Young-Scholten, M. (Eds.). *The Cambridge Handbook of Second Language Acquisition*. Cambridge: Cambridge University Press. p. 317–337.

_____ (2009). Fundamental and gradient differences in language development. *Studies in Second Language Acquisition*, v. 31, p. 259–289.

HERSCHENSOHN, J.; YOUNG-SCHOLTEN, M. (2013) *The Cambridge Handbook of Second Language Acquisition*. Cambridge: Cambridge University Press.

HETTIARACHCHI, S.; PIRES, A. (2016). L2 Acquisition of Wh-Features and Syntactic Constraints: Evidence for Full-Access Approaches. In: Perkins, L. et al. (Eds.), *Proceedings of the 6th Conference on Generative Approaches to Language Acquisition North America (GALANA 2015)*. Somerville, MA: Cascadilla Proceedings Project. p. 48–59.

HOLMBERG, A. (2010). Null subject parameters. In: Biberauer, T.; Holmberg, A.; Roberts, I; Sheehan, M. (Eds.). *Parametric variation: null subjects in minimalist theory*. Cambridge: Cambridge University Press. p. 88–124.

HORNSTEIN, N.; NUNES, J.; GROHMANN, J. (2005). *Understanding Minimalism*. Cambridge: Cambridge University Press.

HORNSTEIN, N.; ROSEN, S.; URIGARIEKA, J. (2002). Integrals. In: Urigarieka, J. (Ed.). *Derivations: Exploring the dynamics of syntax*. London and New York: Routledge. p. 179–191.

IONIN, T.; GOLLA, E.; SANTOS, H.; MONTRUL, S. (2015). Interpretations of NPs in generic and existential contexts in L3 Brazilian Portuguese. *Linguistic Approaches to Bilingualism*. Amsterdam: John Benjamins, v. 5:2, p. 215–251.

IONIN, T. (2013). Morphosyntax. In: Herschensohn, J; Young-Scholten, M. (Eds.). *The Cambridge Handbook of Second Language Acquisition*. Cambridge: Cambridge University Press. p. 505–528.

JIMÉNEZ, M. L. (1997). *Semantic and Pragmatic Conditions on Word Order in Spanish*. Ph.D. Dissertation. Georgetown University, Washington D.C.

JUNG, E. H. (S.) (2004). Topic and Subject Prominence in Interlanguage Development. *Language Learning*, v. 54:4, p. 713–738.

KATO, M. (2013). Deriving "wh-in-situ" through movement in Brazilian Portuguese. In: Camacho, V.; Jiménez Fernández, A. L.; Martín-González, J.; Reyes-Tejedor, M. (Eds.). *Information Structure and Agreement*. Amsterdam: John Benjamins. p. 175–192.

_____ (2005). A gramática do letrado: questões para a teoria gramatical. In: Marques, M. A.; Koller, E.; Teixeira, J.; Lemos, A. S. (Eds.). *Ciências da linguagem: 30 anos de investigação e ensino*. Braga: Centro de Estudos Humanísticos (Universidade do Minho). p. 131 145.

_____; RAPOSO, E. (1994). European and Brazilian Portuguese Word Order: Questions, Focus and Topic Constructions. Work presented in *Linguistic Symposium on Romance Languages XXIV*, Washington, D.C.

KELLERMAN, E. (1983). Now you see it, now you don't. In: Gass S.; Selinker, L. (Eds) *Language transfer in language learning.* Rowley, MA: Newbury House. p. 112–134.

LAKOFF, G. (1977). Linguistic Gestalts. *Papers from the Regional Meeting of the Chicago Linguistic Society*, v. 13, p. 236–287.

LARDIERE, D. (2009). Some thoughts on the contrastive analysis of feature in second language acquisition. *Second Language Research*, v. 25, p. 171–225.

_____ (2008). Feature assembly in second language acquisition. In: Liceras, J.; Zobl, H.; Goodluck, H. (Eds.), *The role of features in second language acquisition*. Mahwah, NJ: Erlbaum. p. 106–140.

_____ (2000). Mapping features to forms in second language acquisition. In: Archibald, J. (Ed.). *Second language acquisition and linguistic theory*. Oxford: Blackwell. p. 102–29.

LENNEBERG, E. H. (1967). *Biological Foundations of Language*. New York: John Wiley & Sons Inc.

LI, C. N.; THOMPSON, S. A. (1976). Subject and topic: A new typology of language. In: Li, C. N. (Ed.). *Subject and topic*. New York: Academic Press. p. 457–489.

LICERAS, J. (1997). The then and now of L2 growing pains. In: Díaz, L.; Pérez, C. (Eds.), *Views on the acquisition and use of a second language*. Barcelona: Universitat Pompeu Fabra. p. 65–85.

LIGHTFOOT, D. (1999). *The Development of Language: Acquisition, Change and Evolution.* Massachusetts: Blackwell Publishers Inc.

LIPSKI, J. M. (2012). Geographical and Social Varieties of Spanish: An Overview. In: Hualde, J. I.; Olarrea, A.; O'Rourke, E. (Eds.), *The Handbook of Hispanic Linguistics.* Chichester: John Wiley & Sons. p. 1–26.

LOPES-ROSSI, M. A. (1996). *A Sintaxe Diacrônica das Interrogativas-Q do Português*. Ph.D. Dissertation. IEL, Unicamp.

MANDELL, P. B. (1999). On the Reliability of Grammaticality Judgement Tests in Second Language Acquisition Research. *Second Language Research*, v. 15.1, p. 73–99.

McCARTHY, C. (2007). *Morphological variability in second language Spanish*. Ph.D. Dissertation, McGill University.

_____ (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research*, v. 24.4, p. 459–486.

MEISEL, J. (2010). Age of onset in successive acquisition of bilingualism: Effects on grammatical development. In: Kail, M. (Ed.) *Language Acquisition across Linguistic and Cognitive Systems*. Amsterdam: John Benjamins. p. 225–247.

_____ (2008). Child second language acquisition or successive first language acquisition? In: Haznedar, B.; Gavruseva, E. (Eds.) *Current trends in child second language acquisition: A generative perspective*. Amsterdam: John Benjamins. pp. 55–80.

_____ (1997). The acquisition of the syntax of negation in French and German: contrasting first and second language acquisition. *Second Language Research*, v. 13, p. 227–63.

MILLER, K.; SCHMITT, C. (2012). Effects of variable input in the acquisition of plural in two dialects of Spanish. *Lingua* v. 120. p. 1178–93.

MIOTO, C. (1994). As interrogações no português brasileiro e o critério-WH. *Letras de Hoje*, v. 96, p. 19-33.

_____; KATO, M. (2005). As interrogativas Q do Português Europeu e do Português Brasileiro atuais. *ABRALIN*, v. 4.1-2, p. 171–196.

MIYAGAWA, S. (2001) The EPP, Scrambling, and wh-in situ. In: Kenstowicz, M. (Ed.) *Ken Hale. A Life in Language*. Cambridge, MA: The MIT Press. p. 293–338.

_____ (2010). *Why agree? Why move? Unifying agreement-based and discourse-configurational languages*. Cambridge: MIT Press, 2010. (Linguistic Inquiry Monographs, v. 54).

MONTRUL, S.; DIAS, R.; SANTOS, H. (2011). Clitics and object expression in the L3 acquisition of Brazilian Portuguese: Structural similarity matters for transfer. *Second Language Research*, v. 27, p. 21–58.

MONTRUL, S. (2009). Re-examining the fundamental difference hypothesis: what can early bilinguals tell us? *Studies in Second Language Acquisition*, v. 31.2, p. 225–57.

_____ (2008). *Incomplete Acquisition in Bilingualism. Re-examining the Age Factor*. Amsterdam: John Benjamins.

MUNHOZ NAVES (2012). Construções de tópico-sujeito: uma proposta em termos de transferência argumental e de transferência de traços de C. *Signum: Estudos Linguísticos*, v. 15(1), p. 245-265.

MYLES, F. (2013). Theoretical approaches. In: Herschensohn, J; Young-Scholten, M. (Eds.). *The Cambridge Handbook of Second Language Acquisition*. Cambridge: Cambridge University Press. p. 46–70.

NEELEMAN, A.; WEERMAN, F. (1997). L1 and L2 word order acquisition. *Language Acquisition*, v. 6, p. 125–70.

NUNES, J. (2017). Circumventing phi-minimality: on some unorthodox cases of A-movement in Brazilian Portuguese. In: Lopes, R.; Avelar, J.; Cyrino, S. (Eds.). *Romance languages and linguistic theory*, v. 12. Amsterdam: John Benjamins. p. 159–184.

OUSHIRO, L. (2009). As Interrogativas Q-in situ têm restrições pragmáticas? In: *Anais do VI Congresso Internacional da ABRALIN*. João Pessoa: Ideia, v. 1, p. 2445–2452.

PEIRCE, J. W. et al. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, v. 51.1, p. 195–203.

PILATI, E. N. S.; NAVES, R. R. (2012). Cisão da categoria pronominal, transferência de traços de C para T e a expressão do sujeito no português brasileiro. Work presented in: *II Congresso Internacional de Linguística Histórica*. São Paulo.

PIRES, A.; TAYLOR, H. (2007). The syntax of wh-in-situ and common ground. In: *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, v. 43.2. Chicago: Chicago Linguistic Society. p. 201–215.

POLLOCK, J.-Y. (1989). Verb Movement, Universal Grammar and the Structure of IP. *Linguistic Inquiry*, v. 20, p. 365–424.

PONTES (1987). *O tópico no Português do Brasil*. Campinas: Pontes.

PRÉVOST, P.; WHITE, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, v. 16, p.103–33.

RIZZI, L. (1996). Residual verb second and the Wh criterion. In: Belletti, A.; Rizzi, L. (Eds) *Parameters and Functional Heads: Essays in Comparative Syntax*. Oxford: OUP. p. 64–90.

_____ (2006). On the from of chains: Criterial positions and ECP effects. In: Cheng, L; Corver, N. (Eds.). *WH-movement: Moving on*. Cambridge, MA: The MIT Press. p. 97–133.

ROBERTS, I. (2018). *Diachronic and Comparative Syntax*. Abingdon: Routledge.

_____ (2012). Macroparameters and minimalism: a program for comparative research. In: Galves, C.; Cyrino, S.; Lopes, R.; Sândalo, F.; Avelar, J. (Eds.). *Parameter Theory and Linguistic Change*. Oxford: Oxford University Press. p. 319–334.

_____ (1999). Verb movement and markedness. In: Graff, M. (Ed.) *Language Change and Language Creation: Creolization, Diachrony and Development*. Cambridge, MA.: The MIT Press. p. 287–327.

_____; HOLMBERG, (2010). Introduction: parameters in Minimalist Theory In: Biberauer, T.; Holmberg, A.; Roberts, I; Sheehan, M. (Eds.). *Parametric variation: null subjects in minimalist theory*. Cambridge: Cambridge University Press. p. 1–57

ROSS, J. R. (1967). *Constraints on variables in syntax*. Ph.D. Dissertation. MIT.

ROTHMAN, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The Typological Primacy Model. *Second Language Research*, v. 27, p. 107–127.

_____ (2010). On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics in Language Teaching*, v. 48, p. 245–273.

_____; CABRELLI AMARO, J.; DE BOT, K. (2013). Third language acquisition. In: Herschensohn, J; Young-Scholten, M. (Eds.). *The Cambridge Handbook of Second Language Acquisition*. Cambridge: Cambridge University Press. p. 372–393.

SANTOS, H. (2013). Cross-linguistic influence in the acquisition of Brazilian Portuguese as a third language. PhD. Dissertation, University of Illinois at Urbana-Champaign.

SCHACHTER, J. (1999). On the issue of completeness in second language acquisition. *Second Language Research*, v. 6, p. 93–124.

SCHMID, M. S. (2010). Languages at play: The relevance of L1 attrition to the study of bilingualism. *Bilingualism: Language and Cognition*, v. 13.1, p. 1–7.

SCHWARTZ, B. D.; SPROUSE, R. (1994). Word order and nominative case in nonnative language acquisition: a longitudinal study of (L1 Turkish) German interlanguage. In: Hoekstra, T.; Schwartz, B. D. (Eds.), *Language acquisition studies in generative grammar*. Amsterdam: John Benjamins. p. 317–68

SELINKER, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, v. 10, p. 209–241.

SINGLETON, D. N.; RYAN, L. (2004). *Language Acquisition: The Age Factor*. Clevedon: Multilingual Matters.

THOMAS, M. (2013). History of the study of second language acquisition. In: Herschensohn, J; Young-Scholten, M. (Eds.). *The Cambridge Handbook of Second Language Acquisition*. Cambridge: Cambridge University Press. p. 26–45.

TSIMPLI, I. M. (2003). Interrogatives in the Greek/English interlanguage: a minimalist account. In: Mela-Athanasopoulou, E. (Ed.). *Selected papers on theoretical and applied linguistics.* Thesaloniki: Aristotle University. p.214–225.

_____; DIMITRAKOPOULOU, M. (2007). The Interpretability Hypothesis: evidence from wh-interrogatives in second language acquisition. *Second Language Research*, v.23.2, p.215–242.

_____; MASTROPAVLOU, M. (2008). Feature interpretability in L2 acquisition and SLI: Greek clitics and determiners. In: Liceras, J. M.; Zobl, H. Goodluck, H. (Eds.). *The Role of Formal Features in Second Language Acquisition*. London: Routledge. p. 142–83.

TSIMPLI, I. M.; ROUSSOU, A. (1991). Parameter resetting in L2? *UCL Working Papers in Linguistics*, v. 3, p. 149–69.

URIBE-ETXEBARRIA, M. (2002). In-situ questions and masked movement. *Linguistic Variation Yearbook*, v.2, p. 215–255.

VAINIKKA, A.; YOUNG-SCHOLTEN, M. (2013). Stagelike development and Organic Grammar. In: Herschensohn, J; Young-Scholten, M. (Eds.). *The Cambridge Handbook of Second Language Acquisition*. Cambridge: Cambridge University Press. p. 581–604.

_____; _____ (1994). Direct access to X'-theory: evidence from Korean and Turkish adults learning German. In: Hoekstra, T.; Schwartz, B. D. (eds.). *Language acquisition studies in generative grammar*. Amsterdam: John Benjamins. pp. 265–316.

XU, L. J. (1986). Free empty category. *Linguistic Inquiry*, v. 17, p. 75–93.

_____; LANGENDOEN, D. T. (1985). Topic structures in Chinese. *Language*, v. 61, p. 1–27.

WHITE, L. (2003). *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

_____ (1992). Long and Short Verb Movement in Second Language Acquisition. *Canadian Journal of Linguistics/Revue Canadienne De Linguistique*, v. 37.2, p. 273–286.

_____ (1990). The Verb-Movement Parameter in Second Language Acquisition. *Language Acquisition*, v.1.4, p. 337–360.

_____ (1989). *Universal grammar and second language acquisition*. Amsterdam: John Benjamins.

WHONG-BARR, M. K. (2006). What transfers? In: Unsworth, S.; Parodi, T.; Sorace, A.; Young-Scholten, M. (Eds.). *Paths of development in L1 and L2 acquisition: in honor of Bonnie D. Schwartz*. Amsterdam: John Benjamins. P. 187–199.

YUAN, B. (2001). The status of thematic verbs in the second language acquisition of Chinese: against inevitability of thematic-verb raising in second language acquisition. *Second Language Research*, v.17.3, p. 248–272.

ZOCCA, C. L. (2010). *Divide et Impera – Separating Operators from their variables*. PhD. Dissertation. University of Connecticut.

APPENDIX I – Lists of Sentences

LISTA 1	
Minha xícara quebrou a asa.	Tossiu quem na reunião?
Meu filho já está nascendo os dentes.	Quem que morreu?
Os carros acabaram a gasolina no meio da viagem.	Quem vai dormir naquele quarto?
As árvores estão apodrecendo as raízes.	A Maria queria o que?
O meu vizinho está sempre latindo o cachorro.	Eles gostam do que?
O Pedro brigou as irmãs o dia todo ontem.	O Pedro ama quem?
Meus amigos cochilam o filho sempre à tarde.	O que que o Pedro vai comprar?
Os jóqueis relincharam os cavalos muito alto.	Do que que eles precisam?
A chave da casa está enferrujando.	Quem você mais admira?
Os muros da minha casa mofam sempre que chove.	A Maria emprestou a caneta para quem?
O pé das bailarinas sangrou durante o espetáculo.	Eles têm que devolver o dinheiro para quem?
Os vidros dos celulares arranharam com o uso.	Para quem que você deu o livro?
Essa casa bate bastante sol.	Para quem o Pedro enviou aquele pacote?
O meu apartamento apareceu uns vazamentos bem feios.	A festa é quando?
As ruas do centro não estão passando carro.	Os alunos vão se reunir onde?
Aqueles quartos só cabem duas pessoas.	A Maria almoçou com quem?
Aquela lagoa pesca muito pintado.	Vocês estão bravos por quê?
Essa piscina nada só os mais rápidos.	O Pedro voltou como?
Esses bares dançam só bêbado.	Elas usaram a pintura para que?
Os parques caminham muitas pessoas no final de semana.	Quando que você agendou a visita?
No meu jardim dá manga.	Onde que fica a loja?
Na praça desceram várias pessoas do bondinho.	Com quem que o Pedro viajou?
Nos documentos consta o meu consentimento.	Por que as crianças mentiram?
Nesses quartos estão entrando as folhas com o vento.	Como eles resolveram o problema?
Chegou quem na aula?	Para que vocês estão fazendo isso?

LISTA 2	
O violão arrebentou a corda.	Nadou quem na piscina?
Aquele relógio estragou os dois ponteiros.	Quem nasceu nesse hospital?
Os turistas sempre irritam a pele na praia.	Quem que vai cantar no show?
Os manifestantes incharam os pés com as pancadas.	Vocês estão comendo o que?
A Maria sempre chora o bebê à noite.	O Pedro se lembrou do que?
O Pedro está viajando os pais pelo mundo inteiro.	Ele indicou quem para vaga?
Os meus tios pularam o gato do telhado.	O que a Maria escreveu no caderno?
Os livros às vezes enganam as capas.	Do que se trata o livro?
A tosse da minha avó piorou bastante.	Quem que sua mãe vai convidar para festa?
Os picolés dessa geladeira derretem muito rápido.	O aluno pediu ajuda para quem?
O zíper das minhas calças está rasgando aos poucos.	A menina mandou a carta para quem?
As folhas daquelas árvores já brotaram de novo.	Para quem o autor dedica o livro?
Minha cidade aconteceu uma catástrofe.	Para quem que você levou o presente?
Aquele bairro mora só os ricos.	A greve vai começar quando?
As TVs da sala de espera sumiram a imagem de repente.	O Pedro estacionou o carro onde?
As caixas de chocolates vinham muitos bombons amargos.	Sua amiga está dançando com quem?
A esteira correu o atleta.	A professora cancelou a aula por quê?
Essa avenida desfila as escolas de samba no carnaval.	O acidente aconteceu como?
As ruas protestaram o sindicato.	Esse medicamento serve para que?
Aquelas praias mergulham só os turistas.	Quando seus pais vão embora?
Na minha mão está saindo uma mancha.	Onde nos encontramos amanhã?

Na universidade chegaram muitos estrangeiros.	Com quem a Maria comparte o quarto?
Nos supermercados grandes caiu o preço da carne.	Por que que a Maria brigou com o Pedro?
Nos parques de Amsterdam florescem tulipas.	Como que o menino explicou a situação?
Desapareceu quem na floresta?	Para que que o Pedro utilizou o programa?

LISTA 3

Minha pipoca não derreteu bem a manteiga.	Tossiu quem na reunião?
Essa planta brota as flores só no verão.	Quem que morreu?
Os casacos rasgaram a manga na máquina.	Quem vai dormir naquele quarto?
Os pacientes estão piorando os sintomas com o frio.	A Maria queria o que?
O meu vizinho está sempre latindo o cachorro.	Eles gostam do que?
O Pedro brigou as irmãs o dia todo ontem.	O Pedro ama quem?
Meus amigos cochilam o filho sempre à tarde.	O que que o Pedro vai comprar?
Os jóqueis relincharam os cavalos muito alto.	Do que que eles precisam?
O teclado do computador quebrou.	Quem você mais admira?
Os chifres da vaquinha já estão nascendo.	A Maria emprestou a caneta para quem?
A água dos reservatórios acabou durante o inverno.	Eles têm que devolver o dinheiro para quem?
Os frutos das amoreiras apodreceram.	Para quem que você deu o livro?
Meu nariz está saindo sangue.	Para quem o Pedro enviou aquele pacote?
O consultório chega muitos pacientes de manhã.	A festa é quando?
As lojas estão caindo o preço do celular.	Os alunos vão se reunir onde?
Os jardins floresceram umas margaridas lindas.	A Maria almoçou com quem?
Aquela lagoa pesca muito pintado.	Vocês estão bravos por quê?
Essa piscina nada só os mais rápidos.	O Pedro voltou como?
Esses bares dançam só bêbado.	Elas usaram a pintura para que?
Os parques caminham muitas pessoas no final de semana.	Quando que você agendou a visita?
Nessa praça não bate vento.	Onde que fica a loja?
Na parede apareceram umas manchinhas pretas.	Com quem que o Pedro viajou?
Nos bairros da periferia já não passa taxi.	Por que as crianças mentiram?
Nessas salas cabem muitos alunos.	Como eles resolveram o problema?
Chegou quem na aula?	Para que vocês estão fazendo isso?

LISTA 4

O portão enferrujou a fechadura.	Nadou quem na piscina?
O quarto está mofando as paredes.	Quem nasceu nesse hospital?
Os bateristas sangraram a mão.	Quem que vai cantar no show?
Os armários arranharam as portas.	Vocês estão comendo o que?
A Maria sempre chora o bebê à noite.	O Pedro se lembrou do que?
O Pedro está viajando os pais pelo mundo inteiro.	Ele indicou quem para vaga?
Os meus tios pularam o gato do telhado.	O que a Maria escreveu no caderno?
Os livros às vezes enganam as capas.	Do que se trata o livro?
O fio do meu telefone arrebentou.	Quem que sua mãe vai convidar para festa?
Os dentes do meu pai já estão estragando.	O aluno pediu ajuda para quem?
O nariz dos velhinhos irrita fácil no inverno.	A menina mandou a carta para quem?
Os pés das grávidas incham nas últimas semanas.	Para quem o autor dedica o livro?
Meu sítio já deu muita pitanga.	Para quem que você levou o presente?
Esse andar desce duas pessoas do elevador.	A greve vai começar quando?
Todos os documentos constam o salário do trabalhador.	O Pedro estacionou o carro onde?
Esses potes de açúcar entraram formigas.	Sua amiga está dançando com quem?
A esteira correu o atleta.	A professora cancelou a aula por quê?

Essa avenida desfila as escolas de samba no carnaval.	O acidente aconteceu como?
As ruas protestaram o sindicato.	Esse medicamento serve para que?
Aquelas praias mergulham só os turistas.	Quando seus pais vão embora?
Na minha escola aconteceu um problema com a diretora.	Onde nos encontramos amanhã?
Nessa rua moram vários políticos.	Com quem a Maria comparte o quarto?
Nas universidades está sumindo o dinheiro para pesquisa.	Por que que a Maria brigou com o Pedro?
Naqueles pacotes vinham muitos plásticos inúteis.	Como que o menino explicou a situação?
Desapareceu quem na floresta?	Para que que o Pedro utilizou o programa?

APPENDIX II – Language Background Questionnaire

Nome: _____ Data de nascimento: ___/___/_____

Este é um breve questionário acerca de sua experiência linguística. Está formado por 8 perguntas e demora uns 5 ou 10 minutos no máximo para responder. Se você tiver qualquer dúvida ou dificuldade de compreensão, por favor, não tenha medo de perguntar. Estou aqui para ajudar!

1. Qual é sua língua materna e que outras línguas você fala? Por favor, liste todas na ordem em que você as adquiriu (começando pela sua língua materna):

1	2	3	4	5

2. Agora liste de novo todas as línguas que você sabe em ordem de dominância (ou seja, começando por aquela que faz você se sentir mais confortável, a que você considera que domina melhor):

1	2	3	4	5

3. Atualmente, que porcentagem de tempo você está exposto (em média) a cada língua?
(Todas as porcentagens devem somar 100%)

Língua	1	2	3	4	5
% Tempo					

4. Com que idade você começou a aprender estas línguas?

Língua	1	2	3	4	5
Idade					

5. Para cada uma das línguas, quais fatores você acha que contribuíram para o seu aprendizado?
(Por exemplo: interação com família, leitura, instrução formal, auto-instrução, assistir TV, ...)
Por favor, liste os fatores em ordem de relevância.

Língua	Fatores para o aprendizado
1	
2	
3	
4	
5	

6. Com que idade você se tornou fluente nestas línguas?
(Se você não se considera fluente marque com uma x)

Língua	1	2	3	4	5
Idade					

7. Para cada uma das línguas, quanto tempo (em anos e meses) você passou nos seguintes ambientes?

Língua	Anos e meses que você passou em...		
	...um país onde se fala a língua	...uma família que fala a língua	...uma escola ou trabalho onde se fala a língua
1			
2			
3			
4			
5			

8. Em uma escala de zero a dez, qual você considera que é seu nível de proficiência para cada uma das línguas?

Língua	Nível de proficiência		
	Fala	Leitura	Escrita
1			
2			
3			
4			
5			

Sinta-se à vontade para fazer qualquer comentário ou observação a respeito de sua experiência linguística que você considere relevante e não tinha já sido perguntado.

_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____

APPENDIX III – Brazilian Portuguese Proficiency Test

Nome _____

Teste de Múltipla Escolha

Cada uma das seguintes sentenças contém um espaço em branco indicando que uma palavra ou frase foi apagada. Selecione aquela que melhor complete a sentença.

1. O chefe quer que você dê uma boa _____ para ter chegado tão atrasado à reunião.

(a) solução   (b) satisfação   (c) explicação   (d) saída

2. O Banco do Brasil decidiu _____ dinheiro à população pobre a juros baixos.

(a) dar   (b) entregar   (c) exigir   (d) emprestar

3. O professor pediu aos alunos que _____ na aula.

(a) prestassem   (b) se distraíssem   (c) fixassem   (d) ouvissem

     atenção   o olhar   sua voz

4. Quando o marido desaparecido há anos entrou pela porta, a mulher não podia _____ nos seus olhos.

(a) acreditar   (b) confiar   (c) esquecer   (d) entregar

5. Depois de anos de trabalho sem folga, o empregado decidiu tirar _____.

(a) homenagem   (b) férias   (c) vagabundagem   (d) licenças

6. Aquela briga _____ más lembranças à minha memória.

(a) levou   (b) veio   (c) transportou   (d) trouxe

7. Quando fez quinze anos, ao invés de uma festa de _____ a menina pediu uma viagem de presente.

(a) formatura   (b) bodas   (c) nascimento   (d) aniversário

8. Despediu-se de todos os convidados antes de _____.

(a) ir embora   (b) desaparecer   (c) ficar em casa   (d) cumprimentar

9. A empresa está passando por uma crise econômica e despediu um terço do seu quadro de _____.

(a) pessoas   (b) funcionários   (c) gente   (d) trabalho

10. Gostei muito do apartamento que visitamos, mas antes de assinar o contrato temos que saber _____.

(a) a cobrança   (b) o pagamento   (c) o preço   (d) o valor

    da dívida   do imposto   do aluguel   da taxa

11. O amor que sentiam um pelo outro quando se conheceram tinha se transformado em _____.

(a) depressão   (b) elogio   (c) comiseração   (d) ódio

12. O rapaz olhava para os olhos grandes e o bonito sorriso da moça. Não conseguia tirar os olhos do _____ dela.

(a) rosto   (b) cabelo   (c) ombro   (d) colo

13. Não quero convidar a Maria para a minha festa porque ela é muito fofoqueira e vive _____ de todo mundo.

(a) se desentendendo   (b) se aproveitando   (c) falando mal   (d) fugindo

14. Perdi os meus óculos e não sei o que fazer. Já _____ por toda a casa e não acho.

(a) pesquisei   (b) procurei   (c) rastreei   (d) investiguei

15. Este coquetel não é aberto ao público. Você só entra se tiver _____.

(a) aceitação   (b) amigos   (c) convocação   (d) convite

16. Estou contente porque minha irmã teve bebê e agora tenho um _____.

(a) sobrinho   (b) neto   (c) primo   (d) cunhado

17. O promotor de justiça tinha todos os dados sobre o crime e podia apresentar a
_____.

(a) fiança   (b) testemunha   (c) denúncia   (d) versão

18. O presidente da empresa confiava muito no seu diretor e deu _____ para que
ele tomasse as decisões.

(a) mão aberta   (b) carta branca   (c) vista grossa   (d) olho gordo

19. A peça trata de uma menina que poderia ser feliz, não fossem os problemas
_____ de sua aparente loucura.

(a) concorrentes   (b) decorrentes   (c) conseqüentes   (d) resultados

20. Com o aumento do desemprego e da inflação, o governo já pensava em alternativas para
_____ econômico.

(a) reverter o   (b) adiantar o   (c) fomentar o   (d) endossar o

    quadro   histórico   aumento   quadro

Teste de Cloze

No seguinte texto, algumas palavras foram substituídas por espaços em branco numerados
do 1 ao 30. Para começar, leia o texto completo para compreendê-lo. Depois, leia de novo e
escolha, da folha de respostas, a palavra correta para preencher cada um dos espaços.
Marque suas respostas com um X na letra escolhida.

Índios do Brasil - Identidade e diversidade[47]

[...]

O Brasil possui uma imensa diversidade étnica e lingüística, estando ____1____ as maiores do
mundo. ___2___ 215 sociedades indígenas, mais cerca de 55 grupos de índios isolados, sobre
os quais ____3____ não há informações objetivas. 180 ___4___, pelo menos, são faladas
___5___ membros destas sociedades, as ___6___ pertencem a mais de 30 famílias lingüísticas
diferentes.

No ____7____, é importante frisar que as ____8____ culturas das sociedades indígenas
modificam-___9___ constantemente e reelaboram-se com o ___10___ do tempo, como a
cultura de ____11____ outra sociedade humana. E é preciso considerar ___12___ isto

---

[47] Texto retirado de http://www.brasilturismo.com/meioambiente/indiosdobrasil/identidadeediversidade.php

aconteceria mesmo que não ____13____ ocorrido o contato com as sociedades de origem européia e africana.

No que diz ____14____ à identidade étnica, as mudanças ocorridas em ___15___ sociedades indígenas, como o fato de ____16____ português, vestirem roupas iguais ___17___ dos outros membros da sociedade nacional com que estão em ___18___, utilizarem modernas tecnologias (como câmeras de vídeo, máquinas fotográficas e aparelhos de fax), não fazem ___19___ que percam sua identidade étnica e ____20____ de ser indígenas.

A diversidade cultural pode ser enfocada tanto ___21___ o ponto de vista das diferenças existentes entre as sociedades indígenas e as não-indígenas, ____22____ sob o ponto de vista das diferenças entre as muitas sociedades indígenas que ____23____ no Brasil. Mas está sempre relacionada ao contato entre realidades socioculturais diferentes e à necessidade de ____24____ entre elas, especialmente num país pluriétnico, como é o ____25____ do Brasil.

É necessário reconhecer e ____26____ a identidade étnica específica de cada ___27___ das sociedades indígenas em particular, compreender suas línguas e suas formas tradicionais de organização social, de ocupação da terra e de uso dos ___28___ naturais. Isto significa o respeito pelos direitos coletivos especiais de cada uma delas e a ____29___ do convívio pacífico, por ___30___ de um intercâmbio cultural, com as diferentes etnias.

1. (a) com   (b) entre   (c) sobre

2. (a) estão   (b) são   (c) hão

3. (a) ainda   (b) sempre   (c) já

4) (a) dialetos   (b) língua   (c) línguas

5) (a) pelos   (b) naqueles   (c) destes

6) (a) quem   (b) quais   (c) cujas

7) (a) mesmo   (b) tanto   (c) entanto

8) (a) variantes   (b) todas   (c) variadas

9) (a) o   (b) se   (c) las

10) (a) passar   (b) passo   (c) passagem

11) (a) qualquer   (b) toda   (c) mesma

12) (a) quanto   (b) qual   (c) que

13) (a) haveria   (b) houvesse   (c) havia

14) (a) cuidado   (b) consideração   (c) respeito

15) (a) várias   (b) aquelas   (c) mesmas

16) (a) falando   (b) falarem   (c) falam

17) (a) às   (b) a   (c) aos

18) (a) relação   (b) conversa   (c) contato

19) (a) de   (b) com   (c) a

20) (a) deixem   (b) deixam   (c) deixaram

21) (a) sob   (b) frente   (c) sobre

22) (a) quando   (b) cujo   (c) quanto

23) (a) vivem   (b) vive   (c) vivam

24) (a) confronto   (b) convívio   (c) comoção

25) (a) fato   (b) trato   (c) caso

26) (a) verificar   (b) valorizar   (c) avaliar

27) (a) vez   (b) um   (c) uma

28) (a) fontes   (b) meios   (c) recursos

29) (a) busca   (b) perseguição   (c) procura

30) (a) entre   (b) meio   (c) através