



Università
Ca'Foscari
Venezia

Corso di Dottorato di ricerca
in Lingue, culture e società
moderne e Scienze del linguaggio
ciclo 32°

Tesi di Ricerca

To Be an or Not To Be An Island

The status of adjuncts
SSD: L-LIN/01

Coordinatore del Dottorato
ch. prof. Enric Bou Maqueda

Supervisore
ch. prof.ssa Alessandra Giorgi

Dottorando
Chiara Dal Farra
Matricola 835035

Contents

| | |
|---|----|
| Acknowledgments | 9 |
| 1. INTRODUCTION | 12 |
| 2. THEORETICAL BACKGROUND | 18 |
| 2.1 On islands | 18 |
| 2.1.1 <i>What is an island?</i> | 18 |
| 2.1.2 <i>Strong and weak islands</i> | 21 |
| 2.1.3 <i>Why islands?</i> | 22 |
| 2.1.4 <i>Island circumvention</i> | 25 |
| 2.2 On adjuncts | 27 |
| 2.2.1 <i>What are adjuncts?</i> | 27 |
| 2.2.2 <i>Why extraction matters</i> | 30 |
| 2.3 Previous approaches to adjunction | 30 |
| 2.3.1 <i>Pre-minimalist approaches</i> | 31 |
| 2.3.2 <i>Takahashi (1994)</i> | 33 |
| 2.3.3 <i>Multiple Spell-Out</i> | 34 |
| 2.3.4 <i>Set-Merge and Pair-Merge</i> | 37 |
| 2.3.5 <i>Late Merge</i> | 39 |
| 2.3.6 <i>Phases and the Phase Impenetrability Condition</i> | 42 |
| 2.3.7 <i>Labeling algorithm</i> | 43 |
| 2.3.8 <i>The role of Agree</i> | 45 |
| 2.3.9 <i>Summing up</i> | 45 |
| 3. ON EXTRACTION: WHERE, WHEN AND HOW | 48 |
| 3.1 Counter-examples | 48 |
| 3.1.1 <i>Extraction from different types of adjuncts</i> | 48 |
| 3.1.2 <i>Extraction, cross-linguistically</i> | 52 |
| 3.2 Factors affecting extraction in Italian | 54 |
| 3.2.1 <i>Presence of finite tense in the adjunct</i> | 56 |
| 3.2.2 <i>Argument/adjunct distinction</i> | 60 |
| 3.2.3 <i>D-linking effect</i> | 61 |
| 3.2.4 <i>Resumptive pronouns</i> | 64 |
| 3.2.5 <i>Level of attachment of the adjunct</i> | 65 |

| | | |
|------------|--|-----|
| 3.2.6 | <i>Type of verbs used</i> | 66 |
| 3.2.7 | <i>Syntactic integration of the adjunct</i> | 68 |
| 3.2.8 | <i>Position of the adjunct in the sentence</i> | 71 |
| 3.3.9 | <i>Summing up</i> | 78 |
| 4. | EXPERIMENTS | 82 |
| 4.1 | Experimental syntax | 82 |
| 4.2 | Previous experimental settings | 85 |
| 4.3 | Experiment A | 89 |
| 4.3.1 | <i>A closer look to resumption and its role in extraction</i> | 89 |
| 4.3.2 | <i>Materials and procedure</i> | 93 |
| 4.3.3 | <i>Results</i> | 96 |
| 4.3.4 | <i>Discussion</i> | 98 |
| 4.3.4 | <i>Summary</i> | 104 |
| 4.4 | Experiment B | 105 |
| 4.4.1 | <i>Materials and procedure</i> | 105 |
| 4.4.2 | <i>Results</i> | 108 |
| 4.4.3 | <i>Discussion</i> | 112 |
| 4.4.4 | <i>Summing up</i> | 115 |
| 4.5 | Experiment C | 115 |
| 4.5.1 | <i>Materials and procedure</i> | 116 |
| 4.5.2 | <i>Results</i> | 119 |
| 4.5.3 | <i>Discussion</i> | 120 |
| 4.5.4 | <i>Summing up</i> | 127 |
| 4.6 | Conclusions | 128 |
| 4.6.1 | <i>Issues of the tests and of statistical analysis</i> | 129 |
| 4.6.2 | <i>What about reductionist theories?</i> | 131 |
| 5. | AN ANALYSIS OF EXTRACTION | 134 |
| 5.1 | Differences among adverbial clauses | 135 |
| 5.1.1 | <i>Experimental evidence reviewed</i> | 136 |
| 5.1.2 | <i>Cross-linguistic and experimental data. Asymmetries part II</i> | 142 |
| 5.1.3 | <i>So far</i> | 148 |
| 5.2 | Attachment of adjuncts | 150 |
| 5.2.1 | <i>Where do they merge?</i> | 152 |

| | |
|--|-----|
| 5.2.2 <i>Is a hierarchy of adverbial clauses possible?</i> | 156 |
| 5.2.3 <i>Potential extension of attachment height</i> | 167 |
| 5.3 Macroevent | 168 |
| 5.3.1 <i>The Single-Event Condition</i> | 169 |
| 5.3.2 <i>The Single Event Condition is not enough</i> | 173 |
| 5.4 Combining the two: ingredients for extraction | 176 |
| 5.4.1 <i>Previous explanations</i> | 176 |
| 5.4.2 <i>The proposal</i> | 184 |
| 5.4.3 <i>Summing up</i> | 203 |
| 6. INTER-SPEAKER VARIATION | 206 |
| 6.1 Inter-speaker variability and what it can do for generative grammar | 206 |
| 6.2 Experimental data of individual variation | 207 |
| 6.2.1 <i>Experiment A</i> | 208 |
| 6.2.2 <i>Experiment B</i> | 211 |
| 6.2.3 <i>Experiment C</i> | 216 |
| 6.2.4 <i>Summing up</i> | 220 |
| 6.3 Inter-speaker variation: where else is it found? | 220 |
| 6.4 Accounting for diversity | 225 |
| 7. CONCLUSION | 231 |
| References | 237 |
| Appendix A | 252 |
| Appendix B | 255 |
| Appendix C | 257 |

List of Figures

| | |
|--|-----|
| Figure 4.1. Raw acceptability judgments on a 1-7 scale in the adjunct introduced by after for each condition presented. | 97 |
| Figure 4.2. Raw acceptability judgments on a 1-7 scale in the adjunct introduced by before for each condition presented. | 97 |
| Figure 4.3. Raw acceptability judgments on a 1-7 scale in the adjunct introduced by without for each condition presented. | 97 |
| Figure 4.4. Raw acceptability judgments on a 1-7 scale representing the average ratings for the island condition, with short filler-gap dependency and no resumption. | 102 |
| Figure 4.5. Raw acceptability judgments on a 1-7 scale representing the average ratings for the island condition, with short filler-gap dependency and no resumption. | 103 |
| Figure 4.6. Z-score acceptability judgments comparing extraction from each adjunct type for the island with non-finite tense condition. | 113 |
| Figure 4.7. Z-score acceptability judgments comparing extraction from each adjunct type for the parasitic gap with non-finite tense condition. | 114 |
| Figure 4.8. Z-score acceptability judgments comparing extraction from each adjunct type for the island and the gap condition. | 122 |
| Figure 4.9. Z-score acceptability judgments comparing extraction from each adjunct type for the parasitic gap without resumptive pronoun condition. | 122 |
| Figure 6.1 Individual raw scores of each participant for extraction from after in the island with no embedding and no resumption condition in Experiment A. | 209 |
| Figure 6.2. Individual raw scores of each participant for extraction from before in the island with no embedding and no resumption condition in Experiment A | 210 |
| Figure 6.3. Individual raw scores of each participant for extraction from without in the island with no embedding and no resumption condition in Experiment A | 210 |
| Figure 6.4. Individual z-score acceptability judgments for each participant for the island condition with non-finite tense of item in (2a) in Experiment B | 213 |
| Figure 6.5. Individual z-score acceptability judgments for each participant for the island condition with non-finite tense of item in (2b) in Experiment B. | 213 |

| | |
|--|-----|
| Figure 6.6. Individual z-score acceptability judgments for each participant for the island condition with non-finite tense of item in (3) in Experiment B. | 214 |
| Figure 6.7. Individual z-score acceptability judgments for each participant for the island condition with non-finite tense of item in (4) of Experiment B..... | 215 |
| Figure 6.8. Individual z-score acceptability judgments for each participant for the parasitic gap condition with non-finite tense of item in (5) of Experiment B. | 216 |
| Figure 6.9. Z-score acceptability judgments comparing extraction from each adjunct type for the island and the gap condition of Experiment C. | 217 |
| Figure 6.10. Individual z-score acceptability judgments for each participant for the island condition with gap. | 217 |
| Figure 6.11. Individual z-score acceptability judgments for each participant for the island condition with gap. | 218 |
| Figure 6.12. Individual z-score acceptability judgments for each participant for the island condition with gap. | 218 |
| Figure 6.13. Individual z-score acceptability judgments for each participant for the island condition with gap. | 218 |
| Figure 6.14. Individual z-score acceptability judgments for each participant for the island condition with gap. | 219 |
| Figure 6.15. Individual z-score acceptability judgments for each participant for the island condition with gap. | 219 |

List of Tables

| | |
|---|-----|
| Table 4.1. Means of z-score ratings for each condition and each adjunct of Experiment A. .. | 96 |
| Table 4.2. Means of z-score ratings for each condition and item for Experiment B | 111 |
| Table 4.3. Mixed-effect model for each adjunct type for Experiment B. | 112 |
| Table 4.4. Means of z-score ratings for each condition and adjunct of Experiment C..... | 119 |
| Table 4.5. Mixed-effect model results for Experiment C. | 120 |

Acknowledgments

After three years, here I am. Writing acknowledgments as the last thing to be added to this thesis, a few hours before the deadline. The feeling is rather strange. Three years have passed by so fast (and yet, sometimes, so slow), and beside enriching me as a person and as a researcher and letting me grow in ways that I did not expect, they had also lead me to some amazing people, to which I am very grateful.

First and foremost, I want to thank my supervisor Alessandra Giorgi. I was still a BA student when I realised that she was one of the best teacher that I ever met. These PhD years supported this intuition. When she is teaching, you can see a spark and a passion towards her job that are not commonly found. For this reason, I was keen on talking with her and the hours spent discussing about linguistics are definitely a memory that always helped me. As a supervisor, she has constantly been extremely supportive, encouraging me and enriching me not only as a linguist, but also as a person. I have to thank her for not letting me to give up on doubt, for believing in me much more than I believed in myself, for challenging me in every discussion, for pushing me when I needed it, and for dealing with my stubbornness. Grazie per l'assertività.

I would like to sincerely thank Guglielmo Cinque, for discussing with me several aspects of my thesis, for his helpful suggestions and for the interest he demonstrated. Mostly, however, I want to thank him for his kindness and his passion and curiosity. During these years he has been a great source of inspiration.

I am extremely grateful to Roland Hinterhölzl for sharing his ideas and his help in several occasions and to Nicola Munaro, for his advices and his trust in me as an assistant to his linguistic classes.

In general, I have an intellectual debt with the whole Linguistic Department in Venice: it is thanks to many Professors here that I fell in love with linguistics. To the ones already mentioned, I would like to add Giuliana Giusti, Anna Cardinaletti, Francesca Volpato, Chiara Branchini and Jared Bernstein.

I thank Michelle Sheehan, for reading my thesis and for her helpful comments that helped me to improve it. Many thanks also to Ángel Gallego. I wish to thank many other people with whom I have discussed my work in the past years: Klaus Abels, Valentina Bianchi, Dustin Chacón, Peter Cole, Gabriella Hermon, Kyle Johnson, Dave Kush, Jim McCloskey, Jairo Nunes and Hedde Zeijlstra. Moreover, parts of this work have been presented at some conferences

(ConSOLE2018, III GETEGRA workshop, WCCFL36, LSA2019). I want to thank the audiences for the feedback I received in these occasions.

I would like to add a few more acknowledgments, even though not directly linked to the work of this thesis. In one way or another, I benefited from these occasions and the people I met. Different works have been presented at some other conferences (WOLP2017, DGfS2018, GLOW41, ISGS8 and SLE2018). I am indebted to the audiences of these occasions for the challenges. In particular, thanks to Elena Callegari for being so supportive during my first conference and to Vieri Samek-Lodovici (also for his help when this PhD was not a thing, yet). Thanks to the organizers and instructors of the CreteLing Summer School 2018 for their incredible work. I also would like to thank Chiara Gianollo and Marco Coniglio for inviting me at the workshop in Villa Vigoni *Dimensions of wh-words*, and for letting me escape from the islands for a bit. Thanks to Cecilia Poletto, for her help and for the passionate discussions in different topics.

I would like to share a few more thanks: to all my colleagues in Ca' Bembo, as well as to some colleagues in the rest of the world. Special mentions go to: Luana and Silvia for sharing that yogurt in Crete, Marloes for so many adventures (hopefully for some more yet to come), Chiara for all the things that we did not say, Andrea for the healthy dose of cynicism, Chiara and Nicholas for the special projects, Łukasz for the great memories in Olinda, Alba for taking the wrong road to the Griffith Observatory. But most and foremost: thanks to Elena, with whom I shared all my carrier as an undergraduate and a graduate student. The whole experience would have been missing something without you – the importance of your constant support, your critiques, the laughter and the joy of being able to share the good and the bad feelings, I really cannot stress enough.

Finally, and crucially, to my family and my friends. For the endless love and support during these years, despite my silences, my stubbornness and my ability to disappear. For the patience in answering my odd language-related questions. I wish I could thank each one of you in the way you deserve, and not with a few random lines in an incomprehensible thesis (focus, anyone?).

1. INTRODUCTION

A crucial aspect of natural languages is that they can give rise to potentially unbounded dependencies: the distance between a gap and its antecedent does not have a fixed upper limit, other than those imposed by processing requirements. However, in some cases these dependencies are not as unrestricted as it seems: they are subject to constraints, as was first noted by Ross (1967). The term used to refer to these structures is *islands*, i.e. structural domains that impose constraints on certain grammatical operations such as *wh*-questions, topicalization and relativization. Notice the asymmetry between sentences in (1):

- (1) a. Who do you believe that Simon saw _ ?
b. *Who do you believe the claim that Simon saw _ ?
c. * Which girl did John arrive after Simon kissed _ ?

In (1a), extraction is grammatical, but in (1b) something goes wrong and the dependency between the gap and its antecedent cannot be formed: an island is being crossed, in this case a Complex Noun Phrase. (1c) is another case of illicit dependency formation: an Adjunct Island. This thesis is devoted to the latter type of islands, which are traditionally taken to be strong ones, because they do not allow any kind of extraction. More precisely, my work is dedicated to the *exceptions* to the island condition, namely to those cases that allow the creation of the filler-gap dependency between the matrix sentence and the adjunct. My primary focus is on cases of extraction from prepositional adjuncts in Italian, but I will show that these exceptions are at work also cross-linguistically, sometimes with different factors involved. The adverbial clauses I will be dealing with are of the central kind, not the peripheral ones (see Haegeman 2012), namely those that are more integrated in the structure.

Islands are to be thought – quite literally – as domains that are separated from the rest of the sentence, and are hence domains from which it is difficult to get away from. Adjuncts are often considered in the same way: domains that are not entirely – or not at all – part of the matrix sentence, but are detached from it. Such a detachment is captured by means of different

representational formats, as in the case of the Set-Merge Pair-Merge asymmetry (Chomsky 2004, 2008), or different timing of merge that can either be too late (Stepanov 2007) or too early (Uriagereka 1999), or the absence of a label (Nunes & Hornstein 2008). Extraction is helpful in this sense: if it is always cross-linguistically impossible, it means that these domains are invisible to the main sentence. On the other hand if extraction from adjuncts is sometimes allowed, we are forced to adopt a different theory that allows the adjunct and the matrix clause to be connected. Therefore, extraction can help us in the analysis of an object that is still quite mysterious in the syntactic domain, despite being so frequent, i.e. adjuncts. A completely satisfying explanation of these elements is still lacking, as admitted many times in recent literature: “Adjuncts are a pesky species” (Boeckx 2008: 34 fn 8); “It is fair to say that what adjuncts are and how they function grammatically is not well understood.” (Hornstein & Nunes 2008: 57); “There has never, to my knowledge, been a really satisfactory theory of adjunction, and to construct one is no slight task.” (Chomsky 2004: 17).

For such a reason, I focus on empirical data of Italian, also by means of formal experimental studies. These are particularly important in order to establish the degree of acceptability – or unacceptability – of some items, as well as to establish the degree of influence some factors have on the possibility of extraction. The three experiments that will be discussed allowed me to answer a crucial question: is extraction from prepositional adjuncts allowed in Italian, as long as some factors are controlled for? Spoiler alert: the answer is ‘it depends’.

Extraction is not blocked to the same degree among all adjuncts. This result allows me to get to a first crucial conclusion: adjuncts are not a uniform class, and accounting for them under a uniform condition is not appropriate. Moreover, the tests allowed me to check also the importance of two of the factors which I claim to be relevant with respect to the possibility of extraction: (i) the role of finite tense, and (ii) the role of resumptive pronouns. The former turns out to be really able to affect the acceptability of the sentence, whereas the latter has different effects depending on the adjunct, in that it can both help rescue the sentence or make the sentence worse. A third crucial result of my experiments is the presence of inter-speaker variability, which is often mentioned in the literature but rarely addressed. In this study I underline the importance of taking into account individual variability and being able to account for its presence. Therefore, the approach that is finally proposed explains why it is present only with some of the items of the test and the crucial correlation between the amount of variability and the type of complementizer used. The aims of my thesis can be summarized as follows:

- To check whether extraction from prepositional adjuncts is allowed in Italian, also by means of formal experimental studies;
- To identify the factors that are able to influence the acceptability of extraction, from islands in general and from adjuncts specifically. These factors are at work not only in Italian, but also in other languages;
- To examine with formal experiments the role of two of these factors (finite tense and presence of resumptive pronouns) in Italian;
- To propose an account explaining the main results found in the test: (i) the difference among adjuncts, and (ii) the presence of individual variation.

I propose that the difference among adjuncts is due to a difference in their point of merger: some are high and some are low. The higher they are, the more difficult it is to extract from them. This is true not only for peripheral adverbial clauses (Haegeman 2004, 2012), but also for central ones. Such a difficulty is linked to the creation of a macroevent reading, which is open only for low adjuncts, but not for higher ones. I also propose that individual variability is due to the ability of the speaker to create a macroevent.

The outline of this thesis is as follows: chapter 2 offers a theoretical background, dedicated first to islands in general, explaining the distinction between weak and strong, their causes, and the mechanisms are able to circumvent the island, i.e. resumptive pronouns, sluicing and pied-piping. This chapter also introduces adjuncts, explaining their differences with arguments and their main properties, and it provides an overview of some of the main approaches accounting for the impossibility of extraction, starting with the CED (Huang 1982) and other pre-minimalist analyses and getting to the PIC (Chomsky 2000, 2001) and more recent developments.

In chapter 3 I show that the theories mentioned above are too strict because none of them is able to account for the fact that extraction from adjuncts is sometimes licensed. In this chapter I first discuss examples of extraction from different types of adjuncts and focus later only on extraction from prepositional adjuncts, both in Italian and cross-linguistically. In the second part of the chapter I show that extraction is affected by many factors which are able to strengthen or weaken the island effect – an aspect that does not fit well with the analyses that consider these domains as strong islands. I discuss several of these factors, which are sometimes common to different types of islands and sometimes at work only with adjuncts. I show that the factors mentioned are valid also for other languages, offering the support of experimental

evidence, when available. Some of the factors are well-known in the literature – as in the case of finite tense, D-linking effect and the asymmetry between argument and adjunct extraction – but some are not well-studied. In particular, this is the case of the placement of the adjunct in the sentence, for which I offer an analysis involving a comparison with parenthetical structures, and I show that they share several crucial properties.

Chapter 4 reports and discusses the results of three different experiments I conducted. The main aim is to understand whether extraction from adjuncts is always blocked, and to which degree. Moreover, I investigate the role of some factors: (i) the distance of the filler-gap dependency (Experiment A), (ii) the presence of a resumptive pronoun (Experiment A and C), and (iii) the presence of finite tense (Experiment B). Experiment A and B focus on three adjuncts, introduced by *after*, *before* and *without*. Experiment C increases their number and considers also *in order to*, *until* and *because*. For each experiment, I present the relevant results focusing in particular on the fact that adjuncts behave differently, both when it comes to islands and to (supposedly) acceptable sentences such as parasitic gaps. At the end of the chapter I present a first analysis where I claim that the island effect cannot be blamed entirely on processing costs, as is the case with reductionist theories.

Chapter 5 presents my analysis of extraction. In the first part I discuss the results of my experiments, as well as other experiments in different languages to show that differences among adjuncts of the central kind are consistently found. In particular, I focus on results of tests involving adjuncts in Norwegian (Bondevik 2018) and Swedish (Müller 2019) that are very similar to my own. I claim that the difference between adjuncts is explained with a syntactic approach that involves different points of merger in the structure, which can be higher or lower. Here, different tests involving scope of adverbs and Condition C effects are presented in order to establish where they merge in the structure. I also discuss a potential hierarchy between adverbial clauses, and compare the result to the acceptability of extraction from them. In the second part of the chapter, I discuss the crucial role that a single event reading has on the possibility of extraction, presenting the approach of Truswell (2007, 2011) but showing also that a semantic approach on its own is not sufficient to explain the data presented. Therefore, I show how this can be related to the attachment site of adjunct, i.e. that a macroevent reading is available only for lower adjuncts, whereas it is blocked for higher ones.

Chapter 6 presents and discusses one more crucial aspect of my experiments, namely the presence of individual variation – the fact that the same item is rated (very) differently depending on the speaker. I also present data from other experimental studies, showing that

inter-speaker and intra-speaker variability are commonly found – even though they are rarely addressed formally. This leads me to discuss the potential issues of focusing on statistical analysis only, and in particular of focusing on the mean value. I underline the importance that individual judgements are kept into account, and that the theory must be able to accommodate these data as well. Finally, I show how the approach developed in chapter 5 is able to account for it. Individual variation is not based here on a difference in the grammar of the speakers, but rather on their capacity of forming a macroevent.

In chapter 7 I briefly discuss the main results and I illustrate some of the (many) open questions and future investigations that this work left me.

2. THEORETICAL BACKGROUND

This introductory chapter sets the theoretical background for the main topics I will discuss in the thesis. Therefore, a first section is dedicated to a general description of different types of island phenomena. The distinction between weak and strong islands will be taken into account, as well as the mechanisms that are able to repair an island effect, i.e. resumptive pronouns, pied-piping, and sluicing. I will also briefly present and discuss the *why* question, namely why do we have these kinds of constraints, following the traditional analyses. Section 2 considers one type of island only, namely adjuncts. This section is dedicated to an introduction to adjuncts and focuses also on the importance of taking into account the possibility of extraction from these domains in order to establish whether they are connected to the main clause or not. In section 3 I will present and analyse some of the main theories that explain why adjuncts are opaque domains, in the sense that extraction from them is impossible – or, rather, that it should be.

2.1 On islands

2.1.1 *What is an island?*

One of the basic properties of a natural language is its ability to create dependencies between a gap and its antecedent. Look at sentence (1) for an example: the antecedent, or filler, is the *wh*-constituent, whereas the gap is the empty object following the verb and signalled by an underscore.

(1) Who did Mary see _ ?

In (1) a dependency is formed between the verb and the interrogative pronoun, in that the latter gets its semantic interpretation and receives its theta-role from the verb. These kinds of

dependencies are obtained with different phenomena such as *wh*- questions, relative sentences, (some types of) topicalizations, and some cases of cleft sentences.¹ These dependencies are potentially unbounded, in that there is not a determined upper limit applied to the distance that can occur between the gap and its antecedent:

- (2) a. Who did Mary see _ ?
 b. Who do you believe [that Mary saw _] ?
 c. Who does Simon think [that you believe [that Mary saw _]]?
 d. Who does Liz know [that Simon thinks [that you believe [that Mary saw _]]] ?
- (3) a. Chi hai incontrato _ ?
 ‘Who did you meet?’
 b. Chi Andrea crede [che tu abbia incontrato _]?
 ‘Who does Andrea believe that you met?’
 c. Chi Sara ha detto [che Andrea crede [che tu abbia incontrato _]]?
 ‘Who did Sara say that Andrea believes that you met?’
 d. Chi Elisa immagina [che Sara abbia detto [che Andrea crede [che tu abbia incontrato _]]]?
 ‘Who does Elisa think that Sara said that Andrea believes that you met?’

Even though they may appear a bit odd, sentences in (2) and (3) are grammatical in both English and Italian. In all of the above cases, a dependency is established between the *wh*-constituent at the beginning of the sentence and the gap at the end of the sentence, namely the thematic position of the *wh*-constituent. In most of the cases we are dealing with long-distance filler-gap dependencies. As can be seen, the depth of the embedding increases and the distance between the gap and its filler gets bigger and bigger. This is what is meant by *potentially unbounded movement*, i.e. the two elements can appear at an indefinite distance apart and are still able to operate over indefinitely large domains. Basically, as shown by sentences in (2) and (3), long-distance dependencies are not affected by the number of clauses or by the number of words standing in between the gap and its filler. Of course, such a dependency is only *potentially*

¹ Since this thesis deals mainly with dependencies in questions formations both for the experimental part and the theoretical one, these are the only type of dependencies that will be discussed and exemplified in this chapter.

unbounded, in the sense that there has to be an upper limit – because of the restricted capacity resources of our memory, if not because of anything else.

However, since the seminal work by Ross (1967), we know that in some cases it is either difficult or impossible to establish a dependency between a gap and its antecedent. Even though movement is unbounded, this does not imply that it is also unconstrained. These cases usually includes sentences in which the gap is inside a certain type of structure:

- (4) a. *Chi Maria dubita [l'affermazione che Andrea abbia incontrato _] ?
*‘Who does Mary doubt the claim that Andrea met?’
- b. *Chi hai incontrato [il ragazzo che ama _] ?
*‘Who did you meet the boy who loves?’
- c. *Di quale macchina [l'autista _] ha causato un incidente ?
*‘Of which car did the driver cause an accident ?’
- d. *Chi Elisa è arrivata [dopo che Andrea ha incontrato _] ?
*‘Who did Elisa arrived after Andrea met?’

These cases are known as *islands*, i.e. domains from which it is difficult to get away from. Sentences in (4) are all ungrammatical examples from Italian, with their ungrammatical counterparts in English. The issue with these sentences is not the fact that the length of the dependency is too big, but the fact that for some reason the dependency cannot be formed. To put it briefly, islands are structural domains that impose constraints on some grammatical operations. It is as if these domains are detached from the rest of the sentence. For such a reason, several elements cannot escape the island, when they are not ‘helped’ by something else.² Sentences in (4) are examples of different types of islands; usually islands are defined on the basis of the structure that creates them. Therefore, (4a) is a Complex Noun Phrase island, (4b) a Relative Clause island, (4c) a Subject island, and (4d) an Adjunct island.³

Since Ross (1967), much work has been done in the field of islands. This does not include just syntax, but also semantics, pragmatics, phonology and their interaction at the interfaces. Different languages have been taken into account, and cross-linguistic differences have been discussed. What appears to be true is that islands are present in every language, but

² Island repair mechanisms will be briefly discussed in §2.1.4.

³ See Szabolcsi (2006), Szabolcsi & den Dikken (2003), Szabolcsi & Lohndal (2017) for a complete overview of islands.

different languages may have different islands, and different island effects. That is to say, island constraints are language-specific, and they undergo parametrization.

2.1.2 Strong and weak islands

The basic distinction we can draw is the one between strong and weak islands, which does not rely on the strength of the effect, but rather it is connected to the (im)possibility of extracting certain elements.⁴ Following Szabolcsi & Lohndal (2017), we may think of strong islands as absolute, meaning that no extraction is allowed. On the other hand, weak islands are selective: some phrases can extract whereas some others cannot. Cinque (1990) offers a diagnostic for the distinction between strong and weak islands, where the difference is based on PP and DP gap. If a domain can contain neither a PP gap nor a DP one, this will be a strong island. On the other hand, it is possible to find a DP gap in a weak island, but not a PP one – even though this might have a weaker effect than in the case of strong islands. Therefore, in the case of weak islands, the distinction between a DP gap and a PP one is stronger than in strong islands.

- (5) a. ?Which girl do you wonder [whether Bill kissed _]?
b. *How do you wonder [whether Bill kissed Sue _]?
- (6) a. *Which girl did John arrive [after Bill kissed _]?
b. *How did John arrive [after Bill kissed Mary _]? (Boeckx 2012: 16)

Sentences in (5) represent a weak island: extraction of an adjunct leads to a total unacceptability, as in the case of (5b), whereas extraction of an argument is not as degraded, as shown by (5a). Such a difference is not felt in example (6), where both an argument (6a) and an adjunct (6b) are extracted from an adjunct. In fact, we are dealing with an example of a strong island, i.e. no extraction is possible, independently from the nature of the extracted element. Notice however that the distinction between weak and strong islands is not always completely straightforward,

⁴ As noted by Cinque (1990: 161, footnote 2), the terms are misleading: “it originated from the rather ‘weak’ ill-formedness produced by complement extraction from *wh*-islands. It is inappropriate on at least two counts: (i) because weak islands [...] are not islands at all for the extraction of complements; and (ii) because, when they act as islands they give rise to *strong* ill-formedness”.

and sometimes it is not easy to distinguish between the two (see Szabolcsi & den Dikken 2003; Szabolcsi & Lohndal 2017 for an overview).

2.1.3 *Why islands?*

The why question is particularly complex in the case of islands: why do languages have these kinds of constraints? Do we have to blame movement itself, or is there a different problem? First, I will briefly review some of the main hypotheses accounting for strong and weak islands, to turn later to approaches explaining islands in more general terms. Some of these approaches will be analysed and further explained later on. For the purpose of this chapter, it is sufficient to stress the fact that for both weak and strong islands the traditional approach has been to find a common explanation able to account for the islandhood of these two (macro)classes of islands. Of course, this ‘dream of unity’ is not always pursuable, as can be seen from Stepanov (2001, 2007) for the empirical issues of treating subjects and adjuncts as a natural and unique class, and from Boeckx (2012) for the theoretical side of the same problem and for a more general overview.

An important distinction has to be drawn from the beginning: one type of approach offers an explanation for weak islands, whereas the other focuses on strong ones.^{5,6} Weak islands are usually accounted for with intervention effects created by the presence of certain elements along the filler-gap dependency. This was the influential proposal by Rizzi (1990), Relativized Minimality described as in (7):

(7) *Relativized Minimality*

In ...X...Z...Y a local relation between X and Y is disrupted when there is Z such that

- (i) Z is of the same type of X; and
- (ii) Z intervenes hierarchically between X and Y.

⁵ That is to say, the distinction is split *at least* into two. The split is not at all so straightforward, but for the sake of the argument here I will take into account mainly approaches arguing for unity.

⁶ Moreover, notice that in the course of this chapter, I will focus mostly on syntactic approaches. But these are not the only ones developed to account for islands: semantic approaches (Szabolcsi & Zwarts 1993; Truswell 2007, 2011) and pragmatic ones (Erteschik-Shir 1973; Goldberg 2006) are proposed as well.

Such a proposal nicely captures a whole bunch of islands and is able to explain them under the same condition: there is an element that intervenes between the source and the target of movement. Movement is affected because of the intervention effect created by the element that stands in the way. In particular, intervention arises when an element that has to be moved tries to cross another element with sharing features in order to reach a higher position that has the same features. Basically, if the two elements are of a similar nature, there will be an intervention effect, and movement will be blocked or degraded.⁷

Strong islands, on the other side, are usually explained with principles holding in narrow syntax; it is not a matter of something standing in the way, but rather the impossibility of crossing certain domains. This is the case of the Condition on Extraction Domain (Huang, 1982), which states that strong islands are domains that are not properly governed, but also of the subsequent Adjunct Condition, which extends the ban on extraction to different types of islands, such as adjuncts, subjects and relative clauses. But the same ban of movement is adopted in the Subadjacency Condition (Chomsky 1973), which imposes locality restrictions. In fact, Subadjacency states that movement cannot cross more than one bounding node and that island effects arise when too many of these are crossed. The deviance of a sentence depends on the number of nodes crossed: crossing one bounding node at a time is okay, but crossing more than one leads to degradation. Notice that bounding nodes are parametrized depending on the language taken into account.⁸ Later on, the role of bounding nodes has been taken over by barriers (Chomsky 1986), and by phases after the advent of minimalism (Chomsky 2000, 2001). In this case the island effect is captured through the Phase Impenetrability Condition (PIC), which will be analysed and explained later.

As can be seen in the hypotheses described here, the general claim is that the island effect should be blamed on movement. In fact, all those approaches predict that movement out of islands is entirely blocked. However, as first stated by Ross (1967), and subsequently underlined by Boeckx (2012) we should not think of islands as domains out of which *any* type of movement is prohibited or blocked. In fact, there are several counterexamples coming from different languages showing that it is not the case that movement is always blocked. This thesis is dedicated to one of these cases, i.e. extraction from adjuncts. Notice that this can be restated

⁷ See Starke (2001) and Rizzi (2001, 2004, 2013) for the refined version of Relativized Minimality, i.e. featural Relativized Minimality. Here, the degree of unacceptability is affected by the degree of featural distinctness of the intervening element.

⁸ See Rizzi (1982) for discussion about Italian and Torrego (1984) for Spanish.

in different terms: there are no absolute islands, and there are no domains that totally ban all dependencies of all kinds. Rather, we should talk about domains that are more or less selective in their islandhood.

We could take the why question to an additional level, and ask why islands – any kind of islands – exist at all, i.e. their ultimate cause. In such a question three main approaches can be distinguished: a derivational approach, a representational one and the processing account (see Boeckx 2008, 2012 for a more detailed overview). Derivational approaches consider islands as constraints on syntactic processes. Under this kind of approach, islands cannot be derived and – if formed – they always lead to an unacceptable output. This is what happens with Subjacency and subsequent analyses of Chomsky (1973, 1986, 2000, 2001). On the other hand, the representational approach sees islands as conditions on the outputs of the computational system, i.e. they are syntactically possible, but unacceptable at the interface between syntax and PF/LF. The common point of these approaches is that islands are conditions of the faculty of language in the narrow sense, either of narrow syntax (derivational approaches) or of the interfaces (representational approaches). Processing accounts – also called *reductionist* accounts – take a complete different turn. The island effect depends on the processing load created by islands, which are not ungrammatical *per se* but are too complex, and for this reason lead the parser to a breakdown. The cause of the island effect is connected to different factors, which go from attention span, to limited processing resource capacity, to working memory (Kluender & Kutas 1993; Hofmeister & Sag 2010). The crucial point is that reductionist accounts blame the island effect entirely on extra-grammatical explanations, i.e. grammar does not have a role when it comes to islands. This means that the grammar gets a much simpler version under a reductionist account because we do not need to acquire a whole new class of constraints. These approaches will be taken into account again later on, and their explanatory power will be discussed considering the experimental results of islands.⁹

⁹ There is actually an additional group of theories that should be mentioned: grounded theories (the terminology is coming from Phillips (2011)). These are a mixture of grammatical and reductionist approaches, in the sense that they consider both grammar and extra-grammatical factors as responsible for the island effect. According to grounded theories, the immediate cause of island effects depends on formal grammatical constraints, but they are ultimately caused by extra-grammatical factors, i.e. by properties of the human sentence parser. Moreover, grounded theories argue that the origin of island effects could be connected to parsing efficiency considerations, which were later grammaticized into a set of constraints (see Berwick & Weinberg 1984).

2.1.4 Island circumvention

Interestingly, there are some strategies able to repair an island, or to better put it, to circumvent the island effect.¹⁰ As well as island effects, these mechanisms are language-specific: what works in a language might not be useful in another. The fact that islands improve under some circumstances was already noted by Ross (1967). I will briefly present these phenomena, but I will not focus on any of them, since they are not primarily relevant to the main topic of the thesis.

The first class has to do with resumptive pronouns. Notice that usually we find a gap inside an island. In some languages, if this gap is ‘filled’ with a resumptive pronoun, the island effect disappears or improves. There is a distinction between languages with grammatical resumptive pronouns, as Modern Irish and Hebrew, and languages with so-called *intrusive* resumptive pronouns, as Italian and English. In the former, the presence of a resumptive pronoun circumvents the island completely, whereas in the latter it only makes the sentence better. Compare for this purpose sentences in (8) and (9), taken from Ross’s original examples (1967: 433)¹¹:

- (8) a. *King Kong is a movie which you’ll laugh yourself sick if you see _ .
b. King Kong is a movie which you’ll laugh yourself sick if you see **it**.
- (9) a. *Palmer is a guy who for _ to stay in school would be stupid.
b. Palmer is a guy who for **him** to stay in school would be stupid.

(8a) and (9a) show that if the island contains a gap the sentence is out, whereas if instead of the gap we find a resumptive pronoun the sentence is better, as in (8b) and (9b). Ross (1967) discusses the differences between chopping and copying rules, where the former are subject to the usual island constraints but the latter are not, thanks to the presence of the resumptive

¹⁰ I am following Boeckx’s (2012) terminology, in that talking about a circumvention effect is more neutral with respect to the theory that one is following: ‘circumvention’ does not imply that islands actually require rescue, and it does not convey that the ultimate cause of islands is movement.

¹¹ However, notice that Ross (1967: 432) himself specifies that sentences similar to (8b) and (9b) are perfectly grammatical in *a dialect* of English, thus he is not referring to standard English.

pronoun. The (presupposed) ameliorating effect of resumption of the intrusive kind will be discussed in more details in the next chapters, as will be also considered experimentally.

As a second strategy for island circumvention, pied-piping is usually discussed. That is, if the moved element brings along the island domain in its movement, the sentence is fine, as shown in (10b) as opposed to (10a).

- (10) a. *Whose did you buy [_ book]?
b. [Whose book] did you buy _ ?

Last, if movement occurs out of an elided domain, i.e. if the island is elided, the sentence improves, as noted by Ross (1969: 276). This corresponds to *sluicing*: a *wh*-constituent is realized, but the rest of the island is not. The island effect is said to disappear if the island itself gets deleted through ellipsis.¹²

- (11) a. *She kissed a man who bit one of my friends, but Tom doesn't realize which one of my friends she kissed a man who bit _ .
b. ?She kissed a man who bit one of my friends, but Tom doesn't realize which one.

Interestingly, as noted by Merchant (2001), if we take into account the phenomenon of sluicing it seems that it is really not the case that movement out of certain domains is inherently blocked. The same condition applies to the other rescuing strategies: if the island effect can be reduced or bypassed by some mechanisms we might conclude that it is not movement *per se* that is completely prohibited, and hence that we should not think of islands as absolute.¹³ As stated by Boeckx (2008: 162, *italic mine*) “no island is a true island – where a true island would be a *truly impenetrable domain*. As Ross suspected, all islands must be relativized to the kind of processes available in the grammar”.

¹² See Merchant (2001) for a detailed analysis of the rescuing effect of sluicing.

¹³ Of course, this depends on the approach that one is following. Depending on the analysis that is given on resumption and ellipsis, different conclusions may hold as well.

2.2 On adjuncts

I will now focus on one of the island types, i.e. adjuncts. As introduced in §2.12, adjuncts belong to the class of strong islands, which means that both arguments and adjuncts extractions are prohibited. Following Szabolcsi & Lohndal's (2017) terminology, we might define them absolute islands. As will be clearer in the next chapter, this is not entirely correct. But let's first take a step back in order to introduce adjuncts and the issues that are connected with these structures.

The difference between arguments and adjuncts is one of the most fundamental distinction in syntax: where the former are obligatory, fixed and selected, the latter are optional, flexible and unselected. Yet, no theory of adjunction reaches a satisfactory explanation within the current approaches. Such a distinction was explained in Government and Binding with two different structure-building operations, namely substitution and adjunction, which were replaced by Set-Merge and Pair-Merge in more recent minimalist terms (see Chomsky 2004, 2008). In such a case, the basic difference is that Set-Merge places syntactic objects on the 'same plane', whereas with Pair-Merge syntactic objects are placed in separate planes. I will focus on the Pair-Merge/Set-Merge distinction, as well as on some others hypotheses regarding adjunction, in §2.3. I will explain the importance of taking into account the possibility of extraction in order to have a better grasp on the phenomenon and its place and role in the main structure.

2.2.1 *What are adjuncts?*

First, I will introduce adjuncts focusing in particular on prepositional adjuncts, also referred to as adverbial clauses (see Haegeman 2012, among others). These are a subtype of subordinate clauses, i.e. their occurrences depend on a matrix sentence. The difference between adjuncts and arguments concerns several properties that will be illustrated here. Arguments are obligatory and fixed elements that are required and selected by the verb and the theta-criterion. If they are dropped the sentence will be ungrammatical, as seen in (12a) and (12b). On the other hand, adjuncts are not required by the verb or by any other element in the matrix sentence. This means that adverbial clauses are not selected and are not assigned a theta role, and that they are optional. Therefore, they can always be omitted, and there is not a fixed number; they can be iterated. In a sentence similar to (12a) three adjuncts can be found, but none of them is needed

in order to form a correct sentence. They can appear in a relatively free order, either with respect to other adverbial clauses in the sentence, or with respect to their position in the sentence itself, as can be seen in sentences (13).¹⁴

- (12) a. Margherita ha visto *(Sara) al parco con il suo cane questa mattina.
'Margherita saw Sara at the park with her dog this morning.'
- b. Alberto ha conosciuto *(sua moglie) dopo aver finito l'università.
'Alberto met *(his wife) after he finished university.'
- (13) a. Giulia ha risparmiato per un anno per viaggiare in Sudamerica.
'Giulia saved money for a year to travel to South America.'
- b. Giulia, per viaggiare in Sudamerica, ha risparmiato per un anno.
'Giulia, to travel to South America, saved money for a year.'
- c. Per viaggiare in Sudamerica, Giulia ha risparmiato per un anno.
'To travel to South America, Giulia saved money for a year.'

Importantly, adjuncts do not affect the category of the phrase they are attached to (see Hornstein & Nunes 2008). The role of adverbial clauses is to provide additional information regarding the circumstances under which the events described in the matrix clause took place, i.e. they have an adverbial function. The type of information that can be encoded is different depending on the adjunct, which can have different roles, such as temporal (14a), causal (14b), conditional (14c), purpose (14d). The classification depends on the semantic relation adverbial clauses establish towards their host.

- (14) a. Matilde ha mangiato un panino prima di andare al ristorante.
'Matilde ate a sandwich before going to the restaurant.'
- b. Marco si è arrabbiato perché il cane gli ha mangiato le scarpe.
'Marco got mad because the dog ate his shoes.'
- c. Samuele non andrà a Madrid se non passerà l'esame.

¹⁴ Notice that sentences in (13) are not entirely equal: changing the position of adjuncts in the sentence requires also a change in their pronunciation. In particular, (13b) has to be realized with comma intonation, the typical realization of parentheticals (see Selkirk 2005). This issue will be discussed in §3.2.8.

‘Samuele will not go to Madrid if he won’t pass the exam.’

d. Abbiamo preso il treno per arrivare prima.

‘We took the train to get there earlier.’

The interpretation of adjuncts does not simply rely on the lexical choice of the preposition introducing the adjunct itself. In fact, some conjuncts can be connected to different interpretations. It is the case of *while*, which can have a temporal reading, as in (15a), or a contrastive one, (15b). The difference depends on its placement in the structure, and in particular on a peripheral reading opposed to a central reading (Haegeman 2012). This factor will be taken into account in §3.2.7.

(15) a. Mentre Lucia suona la batteria, Pietro prepara la cena.

‘While Lucy plays the drums, Pietro is making dinner.’

b. Elena ha passato tutti gli esami mentre Giulia nessuno.

‘Elena passed all her exams, while Giulia none of them.’

Adverbial clauses are found in every language, but they are introduced via different marking systems depending on the language: (i) separate words or adverbial subordinators, as in the case of *senza* (without), *prima* (before), *per* (in order to) in Italian and English; (ii) special verb forms, as in the case of infinitives and participials in English and Italian; (iii) special word orders, as in the case of German.¹⁵

This thesis focuses on adjuncts of the first type, those introduced by a preposition. Prepositional adjuncts can either precede or follow the matrix sentence. Such a variation depends both on the language and on the semantic relation of the adjunct clause with the main clause. As noted by Diessel (2001) there is a correlation between the position of the adverbial clause and the matrix one, and the position of the subordinator. In fact, languages using a final subordinator tend to place the adverbial clause before the main one.¹⁶ On the other hand, in languages where the preposition precedes the adverbial clause, adverbial sentences can either precede or follow the main sentence. This is the case of Italian. In such a case, the position of the sentence can also be connected to a different meaning, as mentioned above.

¹⁵ For an overview of the realization of different types of adverbial clauses see Lin (2015).

¹⁶ As first noted by Greenberg (1963: 83) “In rigid OV languages adverbial clauses always precede the main clause.”

2.2.2 *Why extraction matters*

The role of extraction is crucial as far as adjunction is considered. In fact, establishing whether extraction is never allowed, i.e. if adjuncts are invariably strong islands, can help us define their place and role in the structure. Empirical cross-linguistic evidence regarding the (im)possibility of extraction from adjuncts can help us understand and define a theory for adjunction, which is still lacking nowadays (see Chomsky 2004). Notice that if extraction from adjuncts is always cross-linguistically impossible, it means that these domains are invisible to the main sentence. There are different hypotheses that can explain such a behaviour. First of all, we could follow the Pair-Merge hypothesis (Chomsky 2004, 2008) and propose that adjuncts are attached to a different plane. A second hypothesis has to do with linearization: the Multiple Spell-Out approach (Uriagereka 1999) predicts that adjuncts go under Spell-Out too soon with respect to the rest of the sentence, and for this reason they become opaque domains. Third, we might also consider that adjuncts are inserted too late to participate in the derivation, following the Late Merge hypothesis (Lebeaux 1988; Stepanov 2001, 2007). Last, the role of labeling could be blamed, either because it is absent (Nunes & Hornstein 2008; Hornstein 2009), or too complex (Oseki 2015). Each of these approaches will be explained in the next section.

On the other hand, however, if extraction from adjuncts is sometimes allowed we are forced to draw different conclusions. In fact this would mean that it is not correct to consider adjuncts placed on a different plane than the rest of the sentence, and therefore that two different representational formats are not actually needed, as well as different timing of merge.

2.3 Previous approaches to adjunction

In this section I will present some of the main syntactic approaches to adjunction, from the Condition on Extraction Domain (Huang 1982) to more minimalist accounts. I will briefly discuss some of the issues of such theories, but I will let aside their empirical problems, which will be considered in §3.1.

2.3.1 Pre-minimalist approaches

One of the first approaches proposed to explain the islandhood of adjuncts was the *Condition on Extraction Domain* (Huang 1982), a principle affecting narrow syntax.¹⁷ According to the CED, strong islands are domains that are not properly governed.

(16) *Condition on Extraction Domain* (Huang 1982: 505)

A phrase A may be extracted out of a domain B only if B is properly governed.

(16) states that extraction is only possible out of a domain that is properly governed, i.e. governed by a lexical head. Thus, the CED is able to account for the difference between complements and non-complements. In fact, the former are governed by a lexical head V and thus extraction is allowed, whereas the same is not true for the latter, including subjects and adjuncts.¹⁸ In fact, subjects are governed by a non-lexical head T – or INFL in previous accounts – whereas adjuncts are not governed by any head. The lack of proper government predicts that extraction can never take place from these domains.¹⁹ Under this account, the (strong) islandhood of subjects and adjuncts is unified and accounted for under the same condition; therefore, they are considered as a natural class.

The *Barriers* system (Chomsky 1986) captures the same complement/non-complement distinction of the CED. In this case the crucial role is assigned to L-marking, i.e. θ -marking by a lexical head. If a domain is L-marked, as in the case of complements, it allows extraction. Subjects and adjuncts, on the other side, are not θ -governed by a lexical head. The fact that they are not L-marked makes them barriers for extraction, which is therefore blocked. Movement is constrained by a locality condition which imposes that not more than one barrier can be crossed.

¹⁷ Cattell (1976) should also be taken into account as one of the first approaches accounting for subject and adjunct islands with a unified approach. Cattell's idea is that syntactic configurations are defined on the basis of predicates and functions of the verb. This definition excludes both extraction from subjects, since these are outside the predicate, and from adjuncts, which are optional, i.e. not a function of the verb.

¹⁸ Later on the CED has been extended to other domains, such as relative clauses, complex NP phrases, sentential subjects and sentential complements.

¹⁹ As noted by Lasnik & Saito (1992) the CED can also account for languages in which extraction from subjects is allowed, as in the case of Japanese. In these languages, it can be argued that the subject does not move from the VP, i.e. it is properly governed and extraction is expected. On the other hand, in languages where the subject moves to TP, there is no proper government and hence no extraction is allowed.

Maximal projections are barriers to movement. In some cases these can be circumvented by adjoining the moved phrase to that projection as an intermediate landing site. In fact, movement proceeds successive cyclically and is therefore composed of a series of local steps. However, under other circumstances, a maximal projection cannot be used as a landing site and the moved element is forced to cross the barrier in one fell swoop. When this happens, the sentence is degraded. The level of degradation depends on the number of barriers crossed: if no barrier is crossed the sentence is fine, crossing one barrier will lead to a slight degradation, whereas if more than one barrier is crossed the sentence is completely ungrammatical. Chomsky (1986) considers an adverbial PP as an inherent barrier that also transmits barrierhood to IP: two barriers are crossed at the same time. Crucially, the account predicts a different degradation of the sentence depending on the attachment site of the adverbial clause: if the adjunct is attached to IP the sentence will be more degraded, but if it attaches to VP it will be a minor degradation because the barrierhood of IP is voided.

The problem with these approaches is that they rely on notions such as lexical government, and it is not clear *why* this should play a role in blocking extraction; its role is actually stipulative. Moreover, with the advent of minimalism and the rejection of some notions such as government, both the CED and the Barriers system had to be discarded, at least in the terms adopted before.

A different approach was proposed in Kayne's (1983) *Connectedness* theory, where the unified approach for adjuncts and subjects is abandoned, as predicted by the canonical government configuration given in (17).

(17) W and Z (Z a maximal projection, and W and Z immediately dominated by some Y) are in a *canonical government configuration* iff:

(i) V governs NP to its right in the grammar of the language in question and W precedes Z; or

(ii) V governs NP to its left in the grammar of the language and Z precedes W.

(Kayne 1983: 225)

The prediction of such theory is that there is a left-right asymmetry resulting in different behaviours with respect to extraction. If extraction takes place from a left branch it will be ungrammatical. This applies to subjects in a language like English, which are placed on a left

branch and hence are not in the canonical government configuration with their sisters.²⁰ It is different in the case of right branches, which in English are governed by V to the right. The latter is the case of adjuncts. Following the Connectedness theory, extraction from adjuncts is therefore predicted to be licensed, whereas extraction from subjects is not.²¹

2.3.2 Takahashi (1994)

The first minimalist account for strong islands has been developed by Takahashi (1994), whose major claim is that no extraction is possible out of a domain which has been previously moved. His hypothesis builds on two minimalist assumptions (Chomsky 1995), namely Shortest Move and Chain Uniformity. The former requires that movement should be as short as possible, and therefore that potential landing sites are not skipped. However, if a domain is the result of movement, the intermediate step to its edge is ruled out because of Chain Uniformity, which requires copies created by movement to be identical. If a constituent has moved, this condition is not respected because movement creates a difference between the original copy and the displaced one. The Uniformity Corollary of Adjunctions (UCA) in (18) states that when this condition is violated, sentences are blocked post-syntactically.²²

(18) *Uniformity Corollary on Adjunction* (UCA) (Takahashi, 1994: 25)

Adjunction is impossible to a proper subpart of a uniform group, where a uniform group is a non-trivial chain or a coordination.

Takahashi unifies subjects and adjuncts under the same condition, but this comes at a price with respect to the latter. On the one hand, the UCA does really well in the explanation of the islandhood of subjects, accounted for with the ban on extraction out of a previously moved

²⁰ In general, the proposal of Kayne (1983) does well in the explanation of the islandhood of subjects in different languages. In fact, if in SVO languages it is not common, extraction in SOV and VSO languages is sometimes allowed (see discussion in Truswell 2011: 27 and Stepanov 2007: 89).

²¹ For this reason, Longobardi (1985) proposed a refined version of Connectedness, adding the requirement that “W governs Z” to (19). This has the effect of entirely blocking extraction from adjuncts as well.

²² Takahashi does not specify which portion of the grammar would intervene.

domain.²³ On the other hand, the UCA is not really able to account for adjuncts, which are treated as a variety of coordinate structures. However, such an assumption is stipulative and not motivated (see Stepanov 2007).²⁴

2.3.3 Multiple Spell-Out

The analysis proposed by Uriagereka (1999) and Nunes & Uriagereka (2000) is a structure-building account that captures the different behaviour of complements and non-complements. It is based on the fact that subjects and adjuncts are nonprojecting phrasal sister of a phrase. Uriagereka's (1999) proposal is to derive CED effects from the Linear Correspondence Axiom (Kayne 1994). In this case, the timing of merging of adjuncts is taken into account: the adjunct is claimed to be merged too early, in the sense that a PF convergence requirement forces an early transfer of the adjunct in order to establish the correct linear order following the LCA.

Let me take a step back and explain some basic minimalist concepts needed here. According to the traditional analysis, there is a basic structure-building operation – Merge – that takes syntactic objects from the Numeration and combines them two at a time, building the structure in a bottom-up fashion. A structure is first formed by the operations of the computational system, and is followed by Spell-Out, the point in which the structure formed in the syntax splits into two interface levels: PF and LF. The former is dedicated to the phonological part, whereas the latter to the semantic one. The traditional analysis assumes also that each element is introduced one step at a time in the derivation, i.e. Merge deals with monotonic units. This is the simplest procedure for the linearization operation, which maps constituents into a linear order of terminals. In Uriagereka (1999) a key role is assigned to the Linear Correspondence Axiom (Kayne 1994), which states that structures are linearized on the basis of dominance, i.e. that there is a perfect match between hierarchical structure and linear order.²⁵ Linearization works well when dealing with right branches, i.e. complements, because

²³ In fact, as noted in footnote 19, this would also explain why in some languages the ban on extraction out of subjects is not at work, as in the case of Japanese. If subjects are not moved from their VP position, we do not expect extraction to be bad.

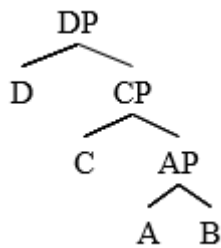
²⁴ Moreover, notice that it is not the case that extraction from coordinate structures is impossible, as can be seen in sentence in (i):

(i) Here's the whiskey which I [went to the store] and [bought _]. (Ross 1967: 168)

²⁵ We can think of Linearization as a requirement for the PF interface: it is an operation imposed on the phonological component as a legibility requirement of the articulatory-perceptual interface (see Higginbotham 1983: 151).

in these cases hierarchical structures can be mapped on the basis of linear order. Following a bottom-up approach, every element that is added to the structure will simply precede that part of the structure at linearization. This would be the case of (19), where the linear order can easily be established applying the LCA and the dominance relation, and results in (20).

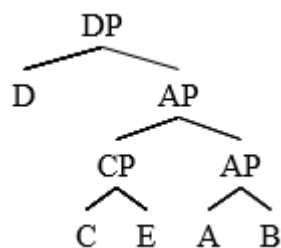
(19)



(20) $D > C > A > B$

However, (20) is an ideal situation but not all of the structures found in syntax are that simple. Uriagereka (1999) focuses on complex left branches and complex adjuncts, claiming that these structures do not allow the syntactic derivation to be strictly monotonic. This means that merge cannot apply in the usual way, i.e. with each element introduced one step at a time. It is the case of (21):

(21)



In example (21) linear order cannot be established: E does not dominate AP, but nonetheless it precedes AP. (22) and (23) – as restated by Stepanov (2007: 86) – describe the problem of (21), and of islands in general:

- (22) If a phrase marker X was assembled in parallel with a phrase marker Y, and then X and Y were Merged, whereupon Y projects, no extraction is ever possible from X.
- (23) X is assembled in parallel with Y iff there exists a derivational point at which X and Y coexist in the derivational space and are unconnected.

Basically, Uriagereka assumes that the two units CP and AP in (21) are assembled in parallel in separate derivational workspaces. As a requirement for the PF interface, i.e. the need to be linearized, adjuncts and complex subjects have to be linearized early in order to become a monotonic unit. Complex units are thus first built in a separate derivational workspace and then are sent to the external system responsible for linearization. This is what happens to CP in (21). The unit gets atomized, and it gets back to syntax as if it was a single, atomic, lexical element, that can be introduced in a monotonic fashion into the rest of the structure and therefore allows the linearization process by means of the LCA.²⁶

The result of early spell-out plus linearization is that the unit gets atomized and becomes opaque, i.e. an island. In fact, these elements become linearized objects; they are frozen in place and the only thing that is recoverable is their label, but not what is inside them. The label is needed in that it encodes the required information that allows a category to undergo a syntactic operation. After early spell-out, thanks to their label, the linearized object is still accessible to the computational system even though its constituents are gone, i.e. it is still visible to linearization and it can move. However, it is treated as a lexical item, which also means that the computational system does not have access to its constituent, but only to its label: they are interpretable but not accessible. In this way adjuncts and complex subjects become invisible to syntax – opaque domains in Uriagereka’s term – and for this reason extraction is never allowed.

I will now provide an example of how the Multiple Spell-Out approach works, taking into account a prepositional adjunct, as in (24):

²⁶ The problem of units being assembled in separate derivational workspace gets a different solution in Johnson (2003). Instead of early spell-out, he applies what he calls *Renumeration* which results in a change of the tree building algorithm. Only non-complements are forced to go through an intermediate step, i.e. Renumerate, before being introduced to the phrase marker.

- (24) a. *Chi Alice ha visto Elisa [mentre Samuele salutava _] ?
 ‘Who did Alice see Elisa while Samuele greeted?’
- b. K= [_{vP} Alice ha visto Elisa]
 L= [_{PP} mentre Samuele salutava chi]]]
- c. [_{CP} +Q [_{TP} Alice [_{vP} [_{vP} ha visto Elisa] [_{PP} while <mentre, Samuele, salutava, chi>]]]]]

Sentence in (24a) is analysed as (24b); it is made of two sentences, which are built independently and assembled in parallel. On the one hand, the terminal elements of the matrix sentence can be assembled without abandoning the derivational space, which means that merge can proceed with pairs and linear order can be established. The same is not true for the adjunct, a complex, nonmonotonic structure that cannot be linearized because it is not possible to determine precedence relations following the LCA. In fact, the elements of the *vP* do not c-command the ones inside the adjunct, thus the lexical elements of K cannot be linearized with respect to the lexical elements of L. Therefore, before K and L merge, spell-out must apply to L: the adjunct goes to spell-out separately and its terminal elements are linearized internal to it. After this operation the adjunct is atomized, i.e. it becomes a monotonic unit and is assigned a label, so that it has the address for its appropriate plugging in, and it can merge with *vP*, namely with K. At this point we have a structure similar to (24c), where a strong feature in the CP of the matrix clause that cannot be checked because there are no *wh*-elements available. In fact, the only *wh*-constituent is realized in the adjunct, which is no longer accessible to the computational system after it was spelt-out. Thus, extraction cannot take place and the derivation crashes: (24a) is ruled out.

As noted by Boeckx (2012), beside some technical issues, the Multiple Spell-Out approach has to deal with a major empirical drawback, i.e. the presence of some mechanisms that are able to attenuate or eliminate the island effect, as well as the presence of some counter-examples. I will illustrate this problem in details in the next chapter.

2.3.4 Set-Merge and Pair-Merge

A different minimalist explanation is offered in Chomsky (2004) on the basis of two primitive mechanisms for structure-building: Set-Merge and Pair-Merge. These mechanisms derive from substitution and adjunction in the Government and Binding theory. In the case of Set-Merge,

two Syntactic Objects (SOs) are combined and they generate unordered sets. On the other hand, Pair-Merge generates ordered sets, as in (25).

- (25) a. Set-Merge $(\alpha, \beta) = \{\alpha, \beta\}$
b. Pair-Merge $(\alpha, \beta) = \langle \alpha, \beta \rangle$

(26) *Set-Merge*:

“For structure building, we have so far assumed only the free symmetrical operation Merge, yielding syntactic objects that are sets, all binary: call them simple. The relations that come “free” (contain, c-command, etc.) are defined on simple structures.”

(Chomsky 2004: 117)

(27) *Pair-Merge*:

“But it is an empirical fact that there is also an asymmetric operation of adjunction, which takes two objects β and α and forms the ordered pair $\langle \alpha, \beta \rangle$, α adjoined to β .”

(Chomsky 2004: 117-118)

These mechanisms are used to explain the difference between arguments and adjuncts. The basic distinction is that Set-Merge generates objects that are on the primary plane, while in the case of Pair-Merge objects are on a different plane. A structure in the ‘primary plane’ retains all the usual properties, such as c-command. The same is not true for adjunction: since this is placed on a ‘separate plane’, properties as c-command are not valid anymore. This explains the difference between arguments and adjuncts, and accounts for the fact that adjuncts are invisible to the syntactic component. Basically, such a distinction would explain why it is impossible to extract from adjuncts. As stated in Chomsky (2008: 146) “the adjunct-island subcase follows if an adjunct is not in the search domain of the probe. That in turn follows from the approach to adjuncts in Chomsky 2004, taking them to be entered into the derivation by pair-Merge instead of set-Merge to capture the fundamental asymmetry of adjunction”. A mechanism like Pair-Merge, which can be applied to adjuncts only, is the reason why these structures are not visible to the main spine of the tree, and therefore explains why extraction cannot take place, and no c-command relations can be defined between the matrix sentence and the adjunct. Chomsky (2004) also postulates the existence of an additional operation called Simplification (SIMPL),

whose role is to apply at Spell-Out and make in-situ adjuncts visible, converting ordered pairs $\langle \alpha, \beta \rangle$ to unordered sets $\{\alpha, \beta\}$ at Spell-Out.

However, the need of a Pair-Merge operation has been questioned by many scholars (Collins 2002; Epstein, Kitahara & Seely 2012; Chomsky 2013; Oseki 2015). This is so because of both theoretical and empirical reasons. The latter will be taken into account in the next chapter; the former are connected to the fact that Pair-Merge exists solely for adjunction: it is construction-specific. Second, stating that adjuncts are on a ‘separate plane’ with respect to the rest of the sentence does not really explain why, it is simply a descriptive restatement of the facts. Chomsky himself (2004: 117-118, quoted in (27)) claims that Pair-Merge is motivated by “an empirical fact”.

2.3.5 Late Merge

I will now turn to a different explanation having to do with the timing of merger: Late Merge. Differently from early Spell-Out, the timing of merge is here blamed for being too late. The assumption of this approach is that adjuncts are inserted after everything else in the structure, i.e. too late to participate in the derivation. This is the position taken by Stepanov (2001, 2007) who builds on Lebeaux (1988) and derives a stronger version of the timing of adjunction. In fact, Lebeaux’s (1988) idea is that adjuncts can be merged late, i.e. post-cyclically – or even acyclically – whereas Stepanov claims that there is no choice: adjuncts must be inserted late, as shown by some empirical facts.^{27, 28}

Stepanov’s (2007) hypothesis, differently from the others seen so far, does not account for the islandhood of subjects and adjuncts under the same condition. In fact, he claims that this kind of analysis makes the wrong predictions: both the CED and the Multiple Spell-Out approach predict that extraction from adjuncts and subjects is equally unacceptable – cross-linguistically. However, Stepanov (2007: 89) offers many cross-linguistic data regarding

²⁷ See for example the case of anti-reconstruction effects in adjuncts (Lebeaux 1988, 1991) and pseudo-opacity effects in adjunct *wh*-extraction (Bošković & Lasnik 1999).

²⁸ A similar Late Merge hypothesis can be found in Ishii (1998), who builds on the difference between arguments, which are selected, and adjuncts, which are not. He postulates the presence of a Derivational Selectional Restriction, which requires selectional restriction to be satisfied as early as possible. This rule therefore requires that all of the arguments are selected first, whereas adjuncts have to be merged after the rest of the structure is built. However, notice again that this rule is stipulative and construction-specific: why would it be that there are two different timing for selected and non-selected domain?

languages where subjects are transparent and adjuncts are opaque, as in the case of Japanese, Navajo, Turkish and Paluan. He claims, though, that there are no languages that behave the other way around, i.e. languages with opaque subjects and transparent adjuncts. He concludes that subjects and adjuncts ask for two different types of explanation.

According to Stepanov (2007), the best way to account for the islandhood of adjuncts is to assume that it is a consequence of narrow syntactic properties of the grammar rather than a constraint at the interfaces. According to his idea, this is connected to the special status of adjuncts in the syntactic derivation: they are merged late, after the whole cycle of the derivation is complete. In particular, he assumes that the Least Tampering Condition (Chomsky 2000) given in (28) regulates the timing of adjunction and forces all instances of set-Merge to apply first. In the Late Adjunction Hypothesis, only when these are completed will it be possible to proceed with adjunction.

(28) *Least Tampering Condition*

Given a choice of operations applying to a syntactic object α , the computation must select one that does not change the set of c-command relations in that object.²⁹

(Stepanov 2001: 102)

The Late Adjunction Hypothesis applies only to adjuncts with no uninterpretable features. If an adjunct contains uninterpretable feature in its label, it will enter the derivation by adjunction.³⁰ It is the case of the prepositional adjunct in (29), which is thematic and does not contain uninterpretable features in its label. The derivation proposed by Stepanov (2001) is described below.

- (29) a. * Chi Alice ha visto Elisa mentre Samuele salutava?
 ‘Who did Alice see Elisa while Samuele greeted?’
 b. [CP Alice ha visto Elisa] [AdjP mentre Samuele salutava chi]

²⁹ The Least Tampering Condition therefore distinguishes set-Merge, which creates new structure on top of the syntactic object, and Pair-Merge which instead add structure inside the syntactic object, and therefore changes the set of c-command relations.

³⁰ In Stepanov (2000) the fundamental distinction is between *wh*-adjuncts and non *wh*-adjuncts. The former are inserted cyclically because they are selected, as in the case of *why* and *how*. Only the latter are true adjuncts and are inserted post-cyclically.

By the Least Tampering Condition and the Late Adjunction Hypothesis, the adjunct has to be merged after the rest of the sentence already merged and all the other processes are complete. This means that the interrogative pronoun inside the adjunct cannot be merged by the time the interrogative feature Q of the matrix complementizer is merged with the TP. Basically, the *wh*-constituent is inserted in a point of the derivation in which the two phrase markers are unconnected to each other. Following (30) we have a problem:

- (30) A (strong) feature must be checked and eliminated (almost) immediately upon insertion into the structure. (Stepanov, 2000: 13)

When *wh*-movement should take place, the adjunct containing the moved element is not yet connected to the main sentence, and therefore to the target of the movement. This is so because of the timing of merger, given that the adjunct is the last thing that is attached to the rest of the sentence in overt syntax. Therefore, the relevant feature of the matrix C cannot be satisfied by *wh*-movement at the derivational stage depicted, because the only available candidate for such movement, namely, *chi* (who), is not part of the same phrase marker. As a result, this feature remains unchecked, and the derivation of this sentence crashes.

This idea has several counter-arguments, both empirical and theoretical. Empirical issues regarding extraction out of adjuncts will be considered later, but it is important to stress a point first made by Boeckx (2008, 2012). The general claim of Stepanov (2000, 2001, 2007) is that adjuncts and subjects cannot be considered as a natural class, and therefore that they cannot be explained under the same condition because of their different behaviour, especially when it comes to extraction. However, the fact that the adjunct condition seems more robust cross-linguistically could easily be connected to the fact that the *notion* of adjunct is in itself more robust cross-linguistically than the one of subject: what counts as a subject may vary, but not what counts as an adjunct.³¹ Therefore, what is being classified as extraction out of a subject might not really be extraction from the same type of domains.³² Notice also that there are many issues created by the notion of late merge. In fact, late merge would non trivially increase the generative capacity of grammars. Moreover, it would also end up to over-generate, predicting

³¹ Boeckx (2012) traces back this idea to Chomsky (1965, 1977), who already claimed that if adjuncts can be considered a primitive of the theory, there is no primitive notion of subject.

³² Moreover, notice that there are some experimental data proposed in Jurka (2010), Jurka, Nakao & Omaki (2011) showing that extraction from subject in German and Japanese is not possible, and that it actually leads to a degraded sentence.

that adjunct movement is potentially able to void any island, because of its unboundedly counter-cyclic nature.³³

2.3.6 Phases and the Phase Impenetrability Condition

As noted above, phases are the natural ‘transformation’ of bounding nodes (Chomsky 1973) and barriers (Chomsky 1986) in minimalist terms. They all capture the same restriction, namely that an efficient computation applies within limited domains. As was the case of barriers and cycles, the basic idea is that movement can never be long-distance, but it is rather composed of successive smaller steps. This is the only way movement can proceed; and when it is not respected the deviation leads to an island effect, because movement is longer than it should and it crossed a limit, i.e. the bounding node (Chomsky 1973), the barrier (Chomsky 1986), the cycle, or the phase (Chomsky 2000, 2001).³⁴ To focus only on the latter, and on the minimalist turn of the hypothesis, I will briefly discuss the idea of the Phase Impenetrability Condition. The general idea first developed in Chomsky (2000, 2001) is that syntactic derivations proceed in smaller units, the phases. Once a phase is completed, its content is transferred to the external systems: it gets spelled-out and it goes to the interfaces. This happens in order to alleviate the burden imposed on the computational system. Once this happens, the content of that phase is no longer visible, it gets opaque. This is precisely what the PIC expresses: the *complement* of a phase is no longer accessible to syntactic computations, it is like a closed window which is going to the external systems in order to be interpreted, phonologically and semantically.

(31) *Phase Impenetrability Condition* (PIC)

Spell-Out the Complement of Ph(ase) as soon as Ph is completed.

(Chomsky 2000: 108)

Notice that it is not the whole phase that is no longer accessible, but only the complement of that phase. This is crucial in that otherwise the PIC would have a way too strong effect, that of blocking any kind of movements and freezing all the elements inside of it. Therefore, an

³³ See discussions contained in Sportiche (2018) for both these points.

³⁴ See Bošković (1994) Abels (2003) and Grohmann (2003) on the opposite idea, namely that there is a ban also on movements that are too short, i.e. the so-called anti-locality. The idea behind such theories is that some syntactic operations are prohibited because they connect elements that are structurally too close to each other, and not because they are too far away.

addition has been made by Chomsky (2000, 2001), namely the role of phase edge composed by the phase head and its specifier. The phase edge is an escape hatch: before the content of the phase gets transferred, movement of any element within that phase to the phase edge is allowed, depending on the presence of features that have to be interpreted. A phase head is able to project additional specifier positions, as many as needed, and to trigger movement of any constituent thanks to the presence of EPP features, i.e. generic features that are capable of attracting any constituent. In this way movement can happen also from phase to phase. Basically, any phase-based theory is forced to rely on extra machinery postulation in order to account for extraction, as in the case of the edge feature.³⁵

In the case of adjuncts, the PIC works well or not depending on how we consider adjuncts, namely if they can be considered phases or not and if they are provided with an escape hatch in their edge. Nowadays, there is still no agreement on the issue. In some cases adjuncts – or to be more precise, Ps – are considered phases (see Abels 2003; Gallego 2010; Svenonius 2010), but there is no general consensus. It suffices to say that the issue requires many stipulations for the time being, since there is no independent diagnostic for phasehood (yet).³⁶

Moreover, as noted by some scholars (see in particular Boeckx 2012: 60-61; Abels 2012), this version of the PIC is able to account for the possibility of movement, but is not able to *restrict* which kind of movement cannot take place: “once an escape hatch is provided, it becomes necessary to provide a restrictive theory of when this escape hatch can be exploited, for, otherwise, all opacifying effects of the PIC can be voided”. The issue in the current theory is that it can be too permissive.

2.3.7 Labeling algorithm

I will now turn to the label adjuncts can be assigned, a non-trivial issue that may have a role in the islandhood of certain domains. The current theory predicts that, once two elements α , β are merged, they get a label, as shown in (32):

³⁵ See also Müller (2010) for an attempt to derive the CED from phases, and in particular from the PIC. Müller, however, takes a complete different turn from Chomsky (2000, 2001), in that he considers all phrases to be phases. See Boeckx (2012) for the issues this approach leads to, and for more general problems with Müller’s proposal.

³⁶ However, notice that there are some language-specific diagnostics that are used for such scope. As argued in Harwood (2015) in English this is the case of ellipsis, VP-fronting and reconstruction. Thanks to Michelle Sheehan for pointing this out.

(32) $\{\gamma \{\alpha, \beta\}\}$

Usually, the label γ is predictable: it is either the projection of α or of β , depending on the projecting head. This is not the case for adjuncts, where the situation is once again not that simple. The issue is that two maximal projections are being merged, I will call them XP and YP. This is not common: usually merge is between a maximal projection and a head: there is an asymmetric condition at the base, and this is what helps with the labeling. Adjuncts are different, and this makes it difficult to establish a way to assign a label to the new constituent. I will briefly review here some of the major proposals. Notice that the most common one is to return a phrase of the same kind the adjunct was attached to:

- (33) a. $[_{XP} [_{XP} \dots X^\circ \dots] YP]$
b. $[_{VP} [_{VP} [_{VP} [\text{greeted that place}] \text{ this week}] \text{ with sadness}]]$

The notation in (33) captures some fundamental properties of adjunction,³⁷ but is not without problems. In particular, under (33) adjunction leaves the maximality of the target XP intact. Normally, a projection is maximal only when it no longer projects. (33) does not follow this: the head of the input is also the head of the output.³⁸

Chomsky (2004) captures the properties of adjunction with a more complex label than regular merge. The label of adjuncts is more complex than the one of complements, in that it is made by both the maximal projections that are being merged. In some accounts, the presence of such complex label forces Spell-Out to take place immediately in order to simplify it. Such an operation requires that the content of the adjoined phrase is transferred to the external systems, after which adjuncts become islands, similar to a sort of Multiple Spell-Out approach (see Narita 2011; Johnson 2009, who claims that adjunction can lead to doubly rooted structures).

There is a third possibility, as far as labeling is concerned: that there is no labeling at all. This is the proposal made by Hornstein & Nunes (2008) and Hornstein (2009). Notice that the role of labeling is to allow merge to apply to previous merged objects. This is always obligatory for arguments, which must be integrated into the structure because their interpretive lives depend on it: they can modify events only via designated relations, i.e. by means of θ -

³⁷ See Hornstein & Nunes (2008) for a detailed discussion of such properties.

³⁸ See Hornstein & Nunes (2008) and Hornstein (2009) for a discussion.

roles. Without labeling these relations would not be possible. On the other side, adjuncts are much more well-behaved: they don't need a label in order to modify the event. The absence of label captures the less-integrated nature of adjuncts to the structure they are inserted in, hence their islandhood. A similar proposal is adopted in Oseki (2015), with a different explanation of the island status of adjuncts: given the absence of labeling, the adjunct has to undergo early Transfer, and its content becomes opaque.

2.3.8 The role of Agree

The final approaches I am taking into account investigate the role of Agree in the islandhood of adjuncts. The basic claim is that a mechanism such as Agree is restricted to selected domains, i.e. arguments, but cannot target adjuncts. This is so because arguments have ϕ features, whereas in adjuncts these are inherent (see Boeckx 2003), and so is their Case (Gallego 2010). When a domain is in an Agree relationship with a head that asymmetrically c-commands it, that domain is transparent and the filler-gap dependency can be created. On the other hand, if there is no Agree the domain will be opaque. The proposal is adopted by several scholars to explain adjuncts (see Rackowski & Richards 2005; den Dikken 2018 for an analysis of all strong islands involving Agree).³⁹ In fact, the typical situation is that verbs do not show an agreement with adjuncts to their projections, meaning that no filler-gap dependency can be created across them.

2.3.9 Summing up

This section took into account some of the major hypotheses concerning the view of adjuncts as strong islands. With the exception of Kayne (1983) and Stepanov (2001, 2007), all of the mentioned approaches propose a unified condition for the strong islandhood of adjuncts and subjects, assuming that extraction is unacceptable in both cases. The hypotheses taken into account blame the island effect on different part of the grammar. Pre-minimalist approaches such as the CED (Huang 1982) and the Barriers system (Chomsky 1986) consider the islandhood as a derivational issue for syntax, meaning that extraction from adjuncts is banned by grammar in the narrow sense. On the other hand, the problem might also be connected to

³⁹ In particular, Boeckx (2003, 2012) analyses the island effect with a decomposition of movement into Move, Agree and Match. His suggestion is that the island effect is not due to constraints on Move, but rather on Agree.

one of the interfaces with syntax – a representational approach in Boeckx’s (2012) terms. This is the case of the Multiple Spell-Out (Uriagereka 1999; Nunes & Uriagereka 2000), which accounts for these islands with a PF requirement and in particular with the fact that adjuncts and subjects are domains that get locked up too soon than the rest of the sentence. A third approach, finally, connects the island status of adjuncts to the fact that they have a special status in the syntactic derivation, different from the main clause. This is what happens with the Late Merge approach (Stepanov 2001, 2007), where adjuncts are merged post-cyclically, or in Lebeaux’s (1988) approach, who claims that adjuncts are merged acyclically. In Chomsky (2004, 2008) adjuncts have a special status in the derivation, that is taken to be realized with the structure-building operation Pair-Merge. The main difference with respect to the other theories is that in this case adjuncts are assumed to be on a ‘separate’ plane, which makes them invisible for the main spine of the tree. Hornstein & Nunes (2008) and Hornstein (2009) derive the impenetrability of these domains on a labeling algorithm, and the absence of an Agree relation is also taken to be influential (Boeckx 2003; Rackowski & Richards 2005; den Dikken 2018).

What all these approaches have in common is the fact that they make a very strong prediction for adjuncts, namely that extraction from these domains is completely banned, either because of syntax, of PF interface, or of a problem in their syntactic derivation. In this chapter I briefly presented some theoretical issues for every approach, letting aside the crucial point of empirical issues. As discussed in §2.2.2, if there are some cases cross-linguistically presenting acceptable extraction from adjuncts, these theories would suddenly become too strong. Of course, things are never that easy for theories, and many examples of extraction are found. The next chapter will be devoted to present some of the cases, first from a cross-linguistic point of view and then with a focus on Italian.

3. ON EXTRACTION: WHERE, WHEN AND HOW

As anticipated in §2.2.2, extraction has a crucial role when it comes to adjuncts. In fact, it can help us establish the connection of these domains to the main structure (if any) and thus it allows us to get a better grasp on the phenomenon. As seen in §2.3, most of the approaches to adjuncts state that in one way or another these are not really related to the main sentence. Whether this is due to narrow syntax itself (Huang 1982; Chomsky 1986), to the timing of merger (Lebeaux 1988; Stepanov 2001, 2007), to the relation with the interfaces (Uriagereka 1999, Nunes & Uriagereka 2000), to a special structure-building operation (Chomsky 2004), to the absence of labeling (Nunes & Hornstein 2008) or to the absence of an Agree relation (Rackowski & Richards 2005), the prediction they make is always very clear – beside being a very strong one. Adjuncts are detached from the matrix sentence – on a ‘separate plane’ in Chomsky’s (2004) terms – which means that extraction can never take place. This chapter claims and shows that this is not the case. It does so first presenting different types of adjuncts from which extraction has been claimed to be licensed, and later focusing only on extraction from prepositional adjuncts from a cross-linguistic point of view. Finally, it turns to extraction from these domains in Italian. Using empirical data, I will present and discuss the role of some factors that are able to affect the island effect, some of them being well-known in the literature, whereas others are more recent and not well-studied yet.

3.1 Counter-examples

3.1.1 Extraction from different types of adjuncts

As explained in §2.2, adjuncts belong to the class of strong islands, because, relying on Cinque’s (1990) diagnostic, there is no distinction between arguments and adjuncts when extraction is taken into account. This can be seen in (1), where both sentences are ungrammatical.

- (1) a. *Which girl did John arrive [after Bill kissed _]?
 b. *How did John arrive [after Bill kissed Mary _]? (Boeckx 2012: 16)

Several approaches have been developed to account for the strong islandhood of adjuncts (as discussed in §2.3), and the generalization emerging from these analyses is that extraction from adjuncts – and from subjects as well, in the majority of the accounts – is never a good option.¹ However, it has been shown by many scholars that extraction is not always unacceptable, from a cross-linguistic point of view.² Apparently, the first counter-examples were presented by Belletti & Rizzi (1981) and by Chomsky (1982). These are given in (2) and (3):³

- (2) What did John go to New York [to buy _]? (Belletti & Rizzi 1981: 132)

- (3) a. Here is the influential professor that John went to college [in order to impress _].
 b. The man that I went to England [without speaking to _]
 (Chomsky 1982: 72)

Sentences in (2) and (3) exemplify some acceptable cases of extraction out of a prepositional adjunct, the one that will be investigated most in this chapter and in the whole thesis. However, notice that extraction can take place from different types of adjuncts. A much studied case regards extraction from bare present participial adjuncts (see Borgonovo & Neeleman 2000;

¹ With the exception of the Connectedness Theory (Kayne, 1983) which states that extraction from left branches (subjects) is not possible, whereas extraction from right ones (adjuncts) is allowed. The issue in this case is that the theory is not able to restrict movement.

² Notice that the same holds for subjects, as already discussed by Stepanov (2001, 2007) among others. Even though these domains are supposed to be strong islands, they do not always behave as such and extraction is sometimes licensed.

³ Belletti & Rizzi (1982: 132) also provided the translation of sentence in (2) for Italian which is given in (i), and they judged it as unacceptable. However, notice that (ii) gets better.

(i) *Che cosa è andato a New York [per comprare _] Gianni?
 what is gone to New York to buy John

(ii) Cosa è andato a New York [a comprare _]?

The amelioration is obtained by changing the introducing conjunct: instead of using *per* I introduced the adjunct with *a*. There is no difference in the English translation, both would be translated with *in order to*.

Truswell 2007, 2011 for English; Demonte 1988; Fábregas & Jiménez-Fernández 2016a for Spanish).^{4,5}

- (4) a. Which play did you fall asleep [watching _]?
b. What did John drive Mary crazy [complaining about _]?

- (5) a. ¿Qué llegó [silbando _] María?
what arrived whistling María?
'What did María arrive whistling?'
b. ¿Qué entró [diciendo _] Juan?
what entered saying Juan
'What did John came in saying?'

(Fábregas & Jiménez-Fernández 2016a:1308)

Extraction is also reported out of secondary adjectival predicates (see Borgonovo & Neeleman 2000; Fábregas & Jiménez-Fernández 2016b)

- (6) a. ¿ Con quién llegó [enfadada _] María?
with whom arrived angry María
b. Whom did Mary arrive [angry with _] ?

(Fábregas & Jiménez-Fernández 2016b: 42)

- (7) a. What did John come back [addicted to _]?
b. What did John leave [satisfied with _]?

(Borgonovo & Neeleman 2000: 203)

Third, as introduced before, extraction might be allowed also from prepositional adjuncts – adverbial phrases in Haegeman's (2012) terms. Sentences in (8) are the first examples used to show that extraction from adjuncts cannot simply be considered as a binary phenomenon, i.e.

⁴ However, see Kohrt, Sorensen & Chacón (forthcoming) for experimental data claiming against the possibility of extraction from gerunds.

⁵ The Spanish data presented in Fábregas & Jiménez-Fernández (2016a) are supported by experimental data in which Spanish speakers were asked to judge the acceptability of these sentences.

either grammatical or ungrammatical. On the contrary, the acceptability of these sentences is something more gradient and requires a different approach in order to be fully captured. Chomsky (1982: 72) states that these sentences “range in acceptability from fairly high to virtual gibberish”.

- (8) a. Here is the influential professor that John went to college [in order to impress _].
 b. The article that I went to England [without reading _]
 c. The book that I went to college [because I liked _]
 d. The man that I went to England [without speaking to _]

(Chomsky 1982: 72)

Last, a final example of the possibility of extraction regards parasitic gaps. This case is different from any other case presented so far. In fact, extraction in these sentences is always valid and parasitic gaps form grammatical sentences.⁶ They have been studied as exceptions to the strong islandhood of adjuncts since Chomsky (1982), Kayne (1983), Engdahl (1983) and Cinque (1990). See sentences in (9):

- (9) a. Quale libro Anna ha recensito _ [prima di leggere _]?
 which book Anna reviewed before of reading
 b. ‘Which book did Anna review _ [before reading _]?’

Parasitic gaps are constructions in which a single *wh*-phrase is associated with multiple gaps in a sentence: one is inside an adjunct and the other is outside. Basically these constructions are made of a ‘good’ and a ‘bad’ gap, which together make an acceptable sentence, as can be seen in (9a), (9b) and (10a). However, if the gap inside the island is the only one, the sentence is not grammatical, as in (10b). On the other hand, if the only present gap is the good one, the sentence is fine (10c).⁷

⁶ Although, as we will see in chapter 4, this is not always the case once experimental data are taken into account.

⁷ The examples presented here take into account only parasitic gaps realised with adjuncts. Notice however that they are possible also inside subjects, as in (i):

(i) Which boy did Mary’s talking to _ bother _ most ? (Engdahl 1983: 5)

- (10) a. What did the student read _ [without understanding _]?
 b. *What did the student read the textbook [without understanding _]?
 c. What did the student read _ [without taking too long]?

The general claim is that the gap inside the island is licensed by the presence of a higher c-commanding gap, created by movement. Since the acceptability of the gap inside the island depends on the presence of another gap, it is called *parasitic*. Several experimental studies showed that their acceptability ratings are usually comparable to declarative sentences (see Phillips 2006, Wagers & Phillips 2009). I will come back to these constructions in chapter 4, where they are used as experimental data.

3.1.2 Extraction, cross-linguistically

I will focus now on extraction from prepositional adjuncts. I will first offer some cross-linguistic examples, which show that the acceptability of these sentences is common in different languages, as recognized by different scholars. Notice that extraction is usually related to infinitival adjuncts only, i.e. adjuncts with a verb realized in its non-finite tense, one of the most influential factors that will be discussed in §3.2.1. Some examples from English are listed in (11). These structures are generally judged as acceptable by native speakers (see Hegarty 1992; Truswell 2007, 2011; Chaves 2012; Sheehan 2013; Brown 2017).

- (11) a. That's the symphony that Schubert [died without finishing _].
 b. Which report did Kim [go to lunch without reading _]?
 c. Who did you [go to Girona in order to meet _]? (Chaves 2012: 468)
 d. Which paper did John travel halfway round the world [to submit _]?
 e. What did you come round [to work on _]? (Truswell 2011: 29)

Extraction from prepositional adjuncts is commonly found cross-linguistically. These cases have been discussed by Cinque (1990) for Italian, Uriagereka (2011) for Spanish, by Anward (1982), Faarlund (1992), Hansen & Heltoft (2011), Müller (2017, 2019) for Mainland

Scandinavian languages and by Biskup & Šimík (2018) for Czech (and other Slavic languages discussed therein), among others.⁸ I will give some examples below.

- (12) Anna, che me ne sono andato via [senza neanche salutare _], ...
Anna, who I went away without even greeting
'Anna, who I left without even greeting, ...' (Italian; Cinque 1990: 103)
- (13) ?Ese es el partido que lloró (amargamente) [por haber perdido _].
this is the game that he cried bitterly for having lost
'That's the game that he cried bitterly for having lost.'
(Spanish; Uriagereka 2011: 117)
- (14) a. [Den saka] ventar vi her [mens de ordnar _].
this things wait we here while they fix
'We are waiting here while they fix this thing.'
(Norwegian; Faarlund 1992: 117, cited in Müller 2019: 60)
- b. [Krig og fred] husker jeg ikke [om _ har blitt oversatt til esperanto].
'War and Peace' remember I not if has been translated to Esperanto
'War and Peace, I don't remember if has been translated to Esperanto'
(Norwegian; Engdahl 1982: 167)
- (15) [Den vase] får du ballade [hvis du taber _].
this vase get you trouble if you drop
'You are in trouble if you drop this vase.'
(Danish; Hansen & Heltoft 2011: 1814, cited in Müller 2019: 58)

⁸ Notice that Mainland Scandinavian languages show island effects in other structures, such as subject islands and coordinate structure islands (Müller 2019). Moreover, the same possibility does not apply to Insular Scandinavian languages like Faroese and Icelandic which do not allow extraction from adjuncts.

- (16) Det här berget måste man träna mycket [för att man ska kunna bestiga _].
 this here mountain must one exercise a-lot for that one shall be-able to-climb
 ‘One has to exercise a lot in order to be able to climb that mountain.’
 (Swedish; Müller 2017: 75)
- (17) To je ten člověk, který1 [když _ promluví], (v ten moment) všichni ztichnou.
 it is the man.nom which.nom when speak.he in the moment all.nom fall.silent
 ‘This is the man such that when he begins to speak, all fall silent.’
 (Czech; Biskup & Šimík 2018: 2)

The sentences above show that there are several circumstances in different languages allowing extraction from adjuncts. Therefore, the data presented here show that the proposal by Stepanov (2001, 2007) is built on empirically unsound bases. In fact, his claim is that there are four possibilities as far as subjects and adjuncts are considered: (i) languages with transparent adjuncts and transparent subjects; (ii) languages with opaque adjuncts and opaque subjects; (iii) languages with opaque adjuncts and transparent subjects, and (iv) languages with transparent adjuncts and opaque subjects. Most of the theories mentioned in §2.3 assumes that only languages of the type (ii) exist, because no extraction is predicted to be allowed from both domains.⁹ The theory proposed by Stepanov (2001, 2007) predicts that two types of languages are present: type (ii), as well as type (iii). In this chapter I showed that we must be able to account for a third type of languages, i.e. type (iv). This seems to be the case of English and Italian, where extraction from subjects is not an acceptable option, whereas extraction from adjuncts can be – under the right circumstances.¹⁰

3.2 Factors affecting extraction in Italian

In the second part of the chapter I will focus on extraction from prepositional adjuncts in Italian, investigating the factors that are able to influence the island effect. First of all, notice that the

⁹ Connectedness theory (Kayne 1983) is an exception. In this case, the type of language predicted by the theory is (iv): transparent adjuncts and opaque subjects.

¹⁰ This leaves out, I believe, the forth possibility, i.e. type (i)languages with transparent subjects and transparent adjuncts.

usual restriction holds for Italian as well: in many cases, extraction out of an adjunct is not acceptable.

- (18) *Quale ragazza Lidia è arrivata [dopo che Elia aveva baciato _]?
which girl Lidia is arrived after that Elia had kissed
'Which girl did Lidia arrived [after Elia kissed _]?'

However, the island effect can be manipulated by means of several factors: islands tend to be selective in the degree of their unacceptability, rather than being absolute (see also Boeckx 2012 on this point). This means that the strength of the island effect can be weaker (or non-existent) or stronger depending on such factors. Consider as an example sentence in (19), where the island effect is not present, or is not very strong.¹¹

- (19) Quale ragazza Gianni è partito [senza salutare _]?
which girl Gianni is left without greeting
'Which girl did John leave without greeting?'

Let's analyse the case in (19) in details. This sentence represents extraction of an argument from a prepositional adjunct introduced by *without*. The *wh*-constituent is a complex one formed by *which* + NP. The matrix verb is intransitive and describes an achievement, whereas the verb in the adjunct is a transitive one describing an activity, and is realized in its non-finite form. If we change some of the factors involved in (19), though, we obtain a much stronger island effect. In the following sections I will present many of these factors, but notice that some of them are not restricted to adjuncts only: they can be applied to other island types. The role of these factors is important, as noted by Chomsky (1986: 1): "With regard to the theory of movement, it appears that a number of different factors enter into informant judgments, including lexical choices to which they are sensitive; it is therefore necessary to proceed on the basis of some speculations concerning the proper idealization of complex phenomena: how they should be sorted out into a variety of interacting systems (some of which remain quite obscure),

¹¹ The judgement of this sentence and many Italian examples in this chapter are my own. However, they are shared by other Italian native speakers who were informally asked for some judgments. Sentence in (19) and similar ones will be taken into account under a formal experimental approach in chapter 4.

which ones may be tentatively be put aside to be explained by independent (sometimes unknown) factors, and which others are relevant to the subsystems under investigation.”

The aim of this chapter is mainly to introduce and explain such factors with empirical data, without going into a detailed analysis, as I will come back to most of them in chapters 4 and 5. Some of the factors will be investigated experimentally and therefore discussed in the next chapter, whereas others will be used for the final theory of extraction. Whenever some of the factors I mention are not formally investigated in one of my tests, I tried to complete the analysis with additional cross-linguistic experimental studies. Notice that when this is the case, the experiments I mention are not usually focused on adjuncts only. In fact, these domains are rarely the main focus of researchers, who tend to investigate the class of islands as a whole. Therefore, not a lot of cases will be adjunct-specific.

3.2.1 Presence of finite tense in the adjunct

One of the major factors able to affect the acceptability of a sentence with extraction from an adjunct is the presence of finite tense, whose effect is well-known and discussed in the literature (see Ross 1967; Cinque 1990; Manzini 1992; Truswell 2011; Szabolcsi & Lohndal 2017 among others). This effect can be applied to different types of island domains, which are better when the gap is present in a non-finite clause. Tense is able to interact with different kinds of islands, both weak and strong, and it makes extraction worse.

- (20) a. What do you wonder [how to repair _]?
b. *What do you wonder [how I repaired _]? (Manzini 1992: 117)

Sentences in (20) show a *wh*-island. Notice that in (20a) the island domain does not have finite tense, whereas (20b) does. Extraction is licensed in the former case, but not in the latter. Thus (20b) is an ill-formed sentence exactly because of the presence of tense.¹²

The finiteness effect can be applied to adjuncts as well: if the adjunct contains a finite verb, the island effect is much stronger. This can be noticed from the ungrammaticality of the Italian example in (21a), the tensed counterpart of sentence in (19) repeated here as (21b), and also from the asymmetries in the English examples in (22).

¹² For experimental evidence confirming the role of finiteness in islands in English see Michel & Goodall (2012) and Müller (2019).

- (21) a. *Quale ragazza Gianni è partito [senza che salutasse _]?
 which girl Gianni is left without that he greeted
 ‘Which girl did John leave without that he greeted?’
- b. Quale ragazza Gianni è partito [senza salutare _]?
 which girl Gianni is left without greeting
 ‘Which girl did John leave without greeting?’
- (22) a. *Which topic did you leave [because James talked about _]?
 b. Which topic did you leave [without talking about _]?
 c. *A person who Robin got real drunk [after she thought about _]
 d. A person who Robin got real drunk [after thinking about _].
- (Culicover 2008, cited in Boeckx 2012: 30)

The role of tense seems so crucial that Truswell (2011: 11 footnote 4) noted that probably “the widespread belief that adjunct parasitic gaps are more acceptable than extraction out of adjuncts is partially due to lack of control for factors such as tense”. And in fact, in examples involving parasitic gaps the verb contained in the adjunct is always in its non-finite form, whereas the same does not hold for island: in several cases examples involving extraction from adjuncts with finite tense have been provided, since Huang (1982), as can be noted in sentences in (23).¹³

- (23) a. *What does Jeff do the housework [because Cindy injured _]?
 (Sprouse 2007: 130)
- b. *Who did you leave the party [because Mary kissed _]?
 (Sprouse 2007: 125)
- c. *Which competition did Lucy read the book [so that she would win _]?
- d. *Which magazine did Sophie play the violin [before she read _]?
 (Brown 2017: 120)
- e. *In campo c’è un giocatore con il quale ti dovrebbero dare istruzioni chiare [se gli altri giocatori saranno scorretti _].

¹³ Examples (23c) and (23d) were taken from Brown (2017), who used them as ungrammatical controls in a test assessing the acceptability of extraction from participial adjuncts and pseudocoordinated structures in English, German and Norwegian. Their role was therefore to provide a baseline for the unacceptability of such sentences.

- f. *‘There’s a player playing in the field towards whom they should give you clear instructions if the other player are unfair’. (Sprouse et al. 2016: 312)

This is a crucial issue that needs to be further addressed: if the presence of tense in the embedded verb makes extraction impossible, then there are at least two hypotheses that could be elaborated, at least from a syntactic point of view.¹⁴ First, we might assume that there is something special about the head T or its maximal projection TP, which somehow is able to interact with adjunction and thus it induces the island effect. This is the proposal of Cinque (1990), who assumes that a tensed inflection has the effect of blocking the upward percolation of the *wh*-feature.¹⁵ The main issue with such an approach is to explain why tense interferes with adjuncts but not with complements, since extraction from a tensed complement is completely acceptable (see also Truswell 2007b). Secondly, we might follow Chomsky (2004, 2008) and hypothesize that a finite sentence should be considered a CP, whereas if T is defective it should be analysed as a *v*P. I will leave the issue open for now, and come back to it in the next chapter, where the effect of tense is investigated from an experimental point of view, and in chapter 5 where it will be discussed from a theoretical point of view.

Notice that the effect of tense could also be related to an additional factor, i.e. the presence of a different subject than the one of the matrix sentence, as in (22a) and (20b) for English and in (24) for Italian. The presence of a subject – or an overt argument – is known in processing accounts as an additional factor increasing the processing burden (Gibson 1998; Warren & Gibson 2002). However, this is unlikely to be the ultimate cause of unacceptability given that sentence (24) is the counterpart of (21a), which does not have an explicit subject but is equally bad. Moreover, extraction from an untensed adjunct with a subject has been reported, as in case of (25).¹⁶

- (24) *Quale ragazza Gianni è partito [senza che Luca salutasse _]?
which girl Gianni is left without that Luca greeted
‘Which girl did John leave without that Luca greeted?’

¹⁴ See Truswell (2007a, 2011) for a semantic-based explanation.

¹⁵ See also Manzini (1992) who unifies the role of Tense and of Definiteness, which have both the effect of blocking the formation of dependencies on the basis of Case-addresses, usually used by DPs to escape from islands.

¹⁶ See also Michel & Goodall (2012) for experimental evidence against the role of subjects/overt arguments as increasing the processing difficulties.

(25) ?This is the problem that we went on holiday [without anyone mentioning _].

(Truswell 2007b: 138)

However, notice that such an effect is not at work cross-linguistically. There are some languages which reportedly are able to extract also from adjuncts containing verbs in their finite tense. This is the case of Czech and other Slavic languages (Biskup & Šimík 2018) and of Mainland Scandinavian languages (Müller 2017, 2019).¹⁷ It has also been claimed that in some cases the effect of tense is not working for English and Spanish as well (Etxepare 1998; Hornstein 2001; Chaves 2012).¹⁸

(26) a. Which email account would you be in trouble [if someone broke into _]?

b. Which problem would you be devastated [if someone had already solved _]?

c. This is the formula that I would be devastated [if someone had already discovered _].

(Chaves 2012: 468)

(27) a. Que libro dijiste [que si Ricardo leía alguna vez] abandonaria la Linguística?

b. Which book did you say that if Ricardo ever read he would abandon linguistics?'

(Hornstein 2001: 91)

(28) Potkal jsem člověka, kterému1 [když to dáš _], tak (to) tě tam pro1/2 pustí]].

met aux.I man.acc which.dat when it give.you so it.nom you.acc there let.in.3

'I met a man such that if you give it to him, he/she/they will let you in.'

(Biskup & Šimík 2018: 2)

¹⁷ See Müller (2017, 2019) for an analysis discussing when extraction from tensed adjuncts is actually licensed and for experimental evidence showing that finiteness is not a factor at work in Swedish.

¹⁸ Judgments of sentences in (26) are from Chaves himself. Notice that the only examples with finite tense in the adjunct are from prepositional adjuncts introduced by *if*, which might be connected with different properties than other adjuncts. In fact, this is not the only occasion in which *if* behaves differently, as in the case the adjunct is in its preverbal position (see Hornstein 2001), which will be discussed later in this chapter.

- (29) Sportspegeln somnar jag [om/när jag ser _].
 sports program-the fall-asleep I if/when I watch
 ‘I fall asleep if/when I watch the sports program.’

(Swedish; Anward 1982: 74, cited in Müller 2019: 53)

- (30) [Den saka] ventar vi her [mens de ordnar _].
 this things wait we here while they fix
 ‘We are waiting here while they fix this thing.’

(Norwegian; Faarlund 1992: 117, cited in Müller 2019: 54)

3.2.2 *Argument/adjunct distinction*

A second factor that should be kept in mind is the different effect of extraction of an argument or of an adjunct. This distinction is well-known and it was used in Cinque (1990) and subsequent literature as a diagnostic for distinguishing between weak and strong islands. As stated in §2.1.2, in the case of weak islands the acceptability of extraction depends on what kind of element has to be extracted: extraction of an argument is much better than extraction of an adjunct, which is rarely acceptable. Since adjuncts are traditionally considered strong islands (see Szabolcsi & Lohndal 2017), extraction should be equally banned for both arguments and adjuncts. However, there is a clear asymmetry in extraction of a DP with respect to extraction of a PP. The former is much more acceptable than the latter, as can be seen in (31) for Italian and in (32) for English. This effect was discussed also by Tanaka (2015), who shows that extraction from adjuncts in English is only possible with arguments extraction, but not with adjuncts.^{19,20}

¹⁹ See also experimental evidence discussed in Müller (2019), showing that extraction of adjuncts from adjunct clauses in Swedish is degraded.

²⁰ Notice that the same restriction applies to parasitic gaps: only DP can be extracted, but extraction of any other category is unacceptable (see Cinque 1990: 102):

- (i) *[_{AP} Quanto importanti] si può diventare _ [senza sentirsi _]?
 How important can one become without feeling
 (ii) *[_{PP} A chi] hai lasciato la lettera _ [dopo esserti rivolto _]?
 To whom did you leave the letter after talking to

- (31) a. Quale ragazza Gianni è partito [senza salutare _]?
 which girl Gianni is left without greeting
 ‘Which girl did John leave without greeting?’
- b. *Con quale ragazzo Maria e Sara sono partite [senza aver parlato _]?
 with which boy Maria and Sara have left without having spoken
 ‘With which boy did Mary and Sara leave without having spoken?’
- (32) a. Which book did Mary improved her thesis [after reading _]?
 b. *How much water did you make the pasta [after boiling _]?

Therefore, the behaviour of adjuncts is more similar to weak islands than to strong ones, in that extraction of an adjunct leads to a stronger degradation than extraction of an argument. Not only because it is possible to extract from them – at least in cases where the verb of the adjunct has a non-finite form – but also because for weak islands there is a common partition between what can and what cannot be extracted, and such a distinction holds in many cases for adjuncts as well. In fact, as discussed by Szabolcsi & den Dikken (2003) and Tanaka (2015), the division between weak and strong islands is not always so straightforward and some cases are rather blurred. This might be the case of adjuncts, at least when certain factors are controlled for. This fact was recognised several times in Government and Binding era, where some scholars reported that extraction from adjuncts was not always banned (see the first cases with Belletti & Rizzi 1981; Chomsky 1982 among others). However, the same is not true for minimalist theories, as seen in §2.3. As I showed, the majority of the accounts does not predict the possibility of extraction, which is taken to be completely ruled out. I will come back to this point at the end of this chapter and in §5.1.3.

3.2.3 D-linking effect

The role of D-linking with respect to extraction from different types of islands has often been recognised since Pesetsky (1987), given that it has the effect of making *wh*-phrases immune to weak islands. If extraction involves the use of a complex *wh*-constituent – a *which* phrase – the sentence is much better than extraction of a simple *wh*-word. Notice that D-linking is not the

only term used to refer to this effect, which has also been defined *referentiality* in Cinque (1990) and Rizzi (1990) or *specificity* in Starke (2001).^{21, 22}

- (33) a. *What do you wonder [whether Mary read _]?
b. Which book do you wonder [whether Mary read _]?
- (34) a. Quale ragazza Gianni è partito [senza salutare _]?
which girl Gianni is left without greeting
'Which girl did John leave without greeting?'
b. *Chi Gianni è partito [senza salutare _]?
who Gianni is left without greeting
'Who did John leave without greeting?'

The effect of D-linking predicts that it is easier to extract elements with more specific/informative fillers, i.e. that extraction is facilitated by an additional property.²³ In this respect, there is a discrepancy between two types of *wh*-phrases, the first being the class of bare *wh*-constituents and the second formed by complex *wh*-phrases made by *which* + NP. Only the latter have the special property of referring to members of a set shared by hearer and speaker (Pesetsky 1987). The main point of D-linking is that the range of variables is already presupposed by the speaker, meaning that it can either be given in the immediate discourse

²¹ See also Aoun, Choueiri & Hornstein (2001: 130) on the role of complex *wh*-constituents in Modern Standard Arabic. They claim that extraction is allowed only with complex *wh*-constituents – provided that a resumptive pronoun is involved as the tail of a dependency – but not with simple *wh*-constituents. See also Tucker, Idrissi, Sprouse & Almeida (2019) for an experimental analysis confirming such a fact.

²² Anagnostopoulou (1994) proposed a hierarchy of *wh*- phrases in order to account for such an effect, as stated below:

- (i) Overt partitive *wh*-phrases (which of your books);
- (ii) *Which*-phrases (which book);
- (iii) *What*-phrases (what book);
- (iv) Bare *wh*-phrases (who, what).

²³ Interestingly, it is the reverse situation in the case of extraction from an island: the less specified, the better, as can be seen in (i) and (ii) (Chaves 2012: 478).

- (i) Which rebel leader did you hear [rumors that the CIA assassinated _]? >
- (ii) Which rebel leader did you hear [the rumor that the CIA assassinated _]?

context or that is shared knowledge. Therefore, crucially, the range of felicitous answers is limited to a set that is shared by both the speaker and the hearer.^{24,25}

Under processing accounts, the role of D-linking has been connected to the fact that a D-linked phrase is expressing a pre-established set of possible answers, and therefore is able to narrow down the set of alternatives that the parser has to consider, basically having the property of reducing the computational effort and helping the working memory in maintaining the filler (Hofmeister & Sag 2010). However, Rottman & Yoshida (2013) showed with a formal experiment that relative clauses are immune to D-linking effects, confirming a previous result in Phillips (2006). This has crucial consequences for processing explanations of D-linking effect. In fact, if D-linking effect is not found with all types of islands, it cannot be simply connected to a processing explanation, which makes the opposite claim: D-linking should be able to uniformly alleviate the island effect in different types of islands. Notice that a division among strong versus weak islands and a difference in their D-linking effect is not helpful: relative clauses are strong islands, which apparently are immune to D-linking effect (Phillips 2006; Rottman & Yoshida 2013), but so are adjuncts, which can indeed improve thanks to the D-linking effect.

There is also a semantic approach to D-linking (Szabolcsi & Zwarts 1993), which states that D-linked *wh*-constituents help with the interpretation when the set questioned consists of individuals, and for such reason it allows for a coherent semantic interpretation of the sentence. On the other hand, many syntactic approaches explain this effect with the addition of an extra feature able to overcome the intervention effect created otherwise – see Haegeman 2010 with the δ feature. The main problem of these accounts is that often these additional featural specifications are ad hoc, and therefore not too good an option.

²⁴ The effect is supported by experimental data where sentences with D-linked phrases were read faster and judged more acceptable than sentences with bare *wh*-phrase (see Hofmeister & Sag 2010; Alexopoulou & Keller 2013). See also Goodall (2015), who showed that D-linking indeed have an ameliorating effect on islands, but it also improves non-island, grammatical sentences. It is also shown that such an amelioration is uniform among the two conditions. See Goodall (2015) for the consequences this findings have on favouring working memory explanations rather than grammatical ones.

²⁵ Notice that the effect of D-linking is at work also with subject islands (see Jiménez Fernández 2009; Bianchi & Chesi 2012), as can be seen in the asymmetry presented in (i) and (ii):

- (i) ¿[De qué cantante] te parece que [algunas fotos _] han escandalizado a la audiencia?
of which singer to-you seem that some pictures have shocked to the audience
- (ii) ??¿De qué cantante te parece que [las fotos _] han escandalizado a la audiencia?
of which singer to-you seem that the pictures have shocked to the audience

(Jiménez Fernández 2009: 127)

Notice however that even if D-linking has a (strong) ameliorating effect on the island, there are some cases with simple *wh*-words that are well-formed sentences. Sentences (35) and (36) are completely acceptable cases for Italian and English as well.

- (35) a. Cosa sei andato/tornato [a fare _]?
 what are (you) gone/come.back to do
 b. What did go/come back to do?’
- (36) a. What did John die [whistling _]?
 b. What did John drive Mary crazy [whistling _]? (Truswell 2011: 30)

3.2.4 Resumptive pronouns

An additional factor able to influence the island effect is the presence of resumptive pronouns in the island domain. As already stated in §2.1.4, there is a difference between languages with grammatical resumptive pronouns and languages with intrusive resumptive pronouns. The latter are not licensed by the grammar, but nonetheless they tend to be used in certain occasions, especially in spontaneous speech, whereas the former are licensed by the grammar. In this section and in the rest of the thesis I will focus only on resumptives of the intrusive kind, given that these are the available option in Italian and English. In these languages the presence of a resumptive pronoun is often reported to improve the status of sentences containing an island violation (Ross 1967; Kroch 1981; Asudeh 2011 for English, Bianchi 2004; Belletti 2006 for Italian). Thus, compare (37a) to (37b) for English, and (38a) to (38b) for Italian.

- (37) a. *The guy who I hate almost everything _ does.
 b. The guy who I hate almost everything he does. (Kroch 1981: 125)
- (38) a. *L’uomo a cui sono certo del fatto che _ parleranno...
 the man to whom I.am.sure of.the fact that they.will.talk
 b. (?)? L’uomo che sono certo del fatto che gli parleranno...
 the man whom I-am-sure-of the fact that to him they will talk
 (Belletti 2006: 130)

However, such an ameliorating effect is often not detected in several experimental tests, assessing either production or acceptability (Ferreira & Swets 2005; Heestand, Xiang & Polinski 2012; Polinsky, Clemens, Morgan, Xiang & Heestand 2013; Beltrama & Xiang 2016). I will therefore discuss the effect of resumption in chapter 4, where experimental data are presented, and where this issue will be taken into account in more detail.

3.2.5 Level of attachment of the adjunct

From this section on, I will focus on effects that are specific for adjuncts and cannot be extended to other types of islands. A first condition regards the height of the adjunct we are dealing with (see Uriagereka 2011; Boeckx 2012: 24), so crucial that “the adjunct condition is robust only insofar as we try to extract elements from adjuncts that are very high in the syntactic hierarchy”. Notice that there are asymmetries between adjuncts that presumably are very high in the structure, and those that are lower. This is the case of *because* clauses, presumably very high (see Uriagereka 2011). In fact, extraction from an adjunct introduced by *because* is always unacceptable, as can be seen in (39) for English and Italian, and in the asymmetries presented in (40) for Spanish, where the sentences are equal but have a different conjunct introducing the adjunct.²⁶ I will not discuss more about such an effect here, as I will come back at this issue in chapter 5.

- (39) a. *Who did John leave the room [because Mary kissed _]?
 b. *Chi Marco è scappato [perché aveva baciato _]?
 who Marco is left because (he) had kissed

- (40) a. ?*Ese es el partido que lloró (amargamente) [porque habíamos perdido _].
 this is the game that he cried bitterly because we had lost
 ‘That’s the game that he cried (bitterly) because we had lost.’
 b. ?Ese es el partido que lloró (amargamente) [por haber perdido _].
 this is the game that he cried bitterly for having lost
 ‘That’s the game that he cried (bitterly) for having lost.’ (Uriagereka 2011: 117)

²⁶ Notice that there is a second aspect which should be taken into account: in both sentences containing extraction from *because*, finite tense is realised in the adjunct – it has to be realised: *because* asks for finiteness. Untensed verbs cannot be combined with this adjunct.

3.2.6 Type of verbs used

Verbs are important not only with respect to finite or non-finite tense, but also to the kind of verb that is used. Borgonovo & Neeleman (2000) noticed such an effect for English, which was later analysed in details in Truswell (2007, 2011) for English and Fábregas & Jiménez-Fernández (2016a) for Spanish. They focused mainly on bare present participial adjuncts. The interesting fact is that not all types of verbs allow extraction, as shown by the contrast in (41) for English.²⁷

- (41) a. *What did John dance [imagining _]?
b. What did John arrive [whistling _]?
c. What did John die [singing _]? (Borgonovo & Neeleman 2000: 199-200)

Borgonovo & Neeleman (2000) noticed that extraction from gerunds adjuncts – predicative adjuncts in their terms – is possible only if the verb is unaccusative or ergative, but not if it is transitive or unergative. Truswell (2007, 2011) developed their idea, claiming that the possibility of extraction depends on the aspectual properties of the matrix predicate. Of the classes individuated by Vendler (1957), i.e. states, activities, achievements and accomplishments, Truswell states that extraction is allowed only with achievements and accomplishments. Moreover, extraction is allowed thanks to a coherent relation between the matrix clause and the adverbial one.

- (42) a. What does John arrive [whistling _]?
b. What did John drive Mary crazy [complaining about _]?
c. *What does John work [whistling _]?
d. *What did John write the cheque [complaining about _]?
(Truswell, 2007: 1359)

In all of the sentences in (42), the verb of the adjunct describes an activity, whereas matrix verbs are of different classes; (42a) is an achievement, (42b) an accomplishment, on the other hand

²⁷ Notice that also the verb of the adjunct can have an effect:

- (i) What did John come home [trying to understand _]?
(ii) *What did John come home [beginning to understand _]? (Truswell 2007a: 1361)

(42c) an activity and (42d) an accomplishment. Notice that sentences in (42) differ only in the choice of the matrix verbs: there is nothing else we could blame for the unacceptability of sentences (42c) and (42d), other than the type of verbs used. The same effect has been noted for Spanish, where extraction is possible only with achievement main verbs (Fábregas & Jiménez-Fernández, 2016a).^{28, 29}

- (43) a. ¿Qué llegó [silbando _] María?
 what arrived whistling María?
 ‘What did María arrive whistling?’
 b. *¿Qué corría [escuchando _] María?
 what ran listening María?
 ‘What did María run listening?’

(Fábregas & Jiménez-Fernández 2006a: 1313)

- (44) a. Quale ragazza Gianni è partito [senza salutare _]?
 which girl Gianni is left without to greet
 ‘Which girl did John leave without greeting?’
 b. *Quale ragazza Gianni ha raggiunto la vetta [senza salutare _]?
 which girl Gianni has reached the top without greeting
 ‘Which girl did John reach the top without greeting?’

As can be noted in the asymmetry in (44), the same effect can be found in Italian: once we change the matrix verb of (44a), the possibility of extraction is reduced, as in (44b). I will come back to this effect in chapter 5, where its semantic and syntactic properties will be taken into account.

²⁸ See Truswell (2007, 2011) for a semantic approach of this phenomenon, and Fábregas & Jiménez-Fernández (2016 a, b) for a different analysis involving syntax.

²⁹ See Müller (2019) for experimental evidence showing that the possibility of extraction in Swedish is significantly ameliorated thanks to a coherent relation between matrix and adjunct clause. Moreover, see Brown (2017) for experimental evidence showing that the type of matrix predicate makes a difference. Specifically, she shows that in English intransitivity is the key factor in the licensing of extraction, contrary to telicity alone. In fact, sentences with unaccusative matrix predicates are significantly more acceptable than unergative or transitive verbs.

3.2.7 Syntactic integration of the adjunct

The syntactic integration of the adjunct in the sense of Haegeman (2004, 2012) has also a crucial role. She showed that it is crucial to distinguish between peripheral and central adverbial clauses because such a difference is important in order to fully account for their properties, also when it comes to extraction. In fact, the two classes do not behave in a homogeneous way; extraction is possible out of central adverbial clauses, but not from peripheral ones. The two types are very different from one another, both in their external and in their internal syntax. First, the role of peripheral adverbial clauses is to modify and structure the discourse, whereas central adverbial clauses modify the event structure of the main clause. An example is given in (45a) for a central adverbial clause and in (45b) for a peripheral one.

- (45) a. According to Smith, a group of Arkansas state troopers who worked for Clinton while he was governor wanted to go public with tales of Clinton's womanising.
- b. While his support for women priests and gay partnerships might label him as liberal, this would be a misleading way of depicting his uncompromisingly orthodox espousal of Christian belief. (Haegeman 2012: 160)

The sentences in (45) show the difference between the two readings; the adjunct is introduced by the same preposition – *while* – but its meaning is different depending on its position in the sentence: in (45a) it expresses a temporal specification of the main event (central reading), whereas in (45b) it is connected to a contrast reading that applies to a higher level, the discourse one (peripheral reading), which basically provides background proposition. The same applies to Italian, as shown in (46), where the adjunct is realised in the reverse order with respect to English.

- (46) a. Mentre Lucia suona la batteria, Pietro prepara la cena.
'While Lucy plays the drums, Pietro is making dinner.'
- b. Elena ha passato tutti gli esami mentre Giulia nessuno.
'Elena passed all her exams, while Giulia none of them.'

The two readings do not apply only to a conjunct like *while*, but can be extended to different adjuncts as well. This is the case of *so that*, which may express a purpose clause in the central reading or a result clause in the peripheral one, and of *because*, which expresses cause or reason in the central reading and a rationale for the speaker in the peripheral one. Not all of the conjuncts allow for a double reading; some allow only for a central one, as in the case of the temporal expressed by *before* and *after*, whereas others allow only for a peripheral one, as with *although*, *whereas* and *given that* (see Haegeman 2004, 2012 for a detailed discussion).

The distinction between central and peripheral adverbial clauses led Haegeman to assume that their external syntax is different in the level and degree of integration to the main clause. In fact, adverbial clauses are not merged at the same point in the structure: on the one hand central adverbial clauses modify either the TP or the domain internal to the TP and are meant to modify the event described in the matrix clause, on the other peripheral adverbial clauses modify a higher domain, and are in fact external to the TP. The former are therefore merged lower and earlier in the derivation, showing that they are closely related to the main clause (they are part of the main clause), whereas the latter are adjoined to the CP, hence their connection to the main clause is a much looser one. They are also merged after the main clause is fully projected. Peripheral adverbial clauses have a peculiar prosody, signalling that they are detached from the main clause, and an appropriate punctuation, in that in written texts peripheral adverbial clauses are always separated by a comma from the rest of the sentence; the same does not apply to central peripheral clauses.³⁰

Moreover, they also differ in their internal syntax: central adverbial clauses refer to a part of the speech act expressed in the superordinate clause, i.e. adjunct and matrix are a unit, while peripheral ones have properties similar to main clauses, and give rise to main clause phenomena. For this reason peripheral adverbial clauses display a full clausal structure, whereas central ones lack a piece, given that they are “part of and they modify the proposition with which they are associated” (Haegeman 2004: 76). The difference proposed in Haegeman

³⁰ The fact that adverbial clauses attach to different merging points is also shown by effect of scope of different phenomena, such as focus, interrogative and negative operators or of matrix tense. See the discussion in Haegeman (2004, 2012) showing that only central adverbial clauses are within their scope, unlike peripheral adverbial clauses. There are also differences in the syntactic processes that affect them: central adverbial clauses are affected by operations that affect the TP and VP domain, such as VP substitution and VP ellipsis, but the same is not true for peripheral adverbial clauses. See also discussions in Frey (2011), who shows that central adverbial clauses can carry the stress of the entire utterance, whereas peripheral adverbial clauses requires two different stresses on the matrix and the adjunct clause.

regards the presence (for peripheral) and the absence (for central) of illocutionary force in the adjunct. Moreover, internal and external syntax of adverbial clauses are somehow correlated, in the sense that the level of attachment determines their internal syntax (Haegeman 2012; Endo & Haegeman 2019).

The two types of adverbial clauses are important for the possibilities of extraction. In fact, this is open only in case the adverbial clause is a central one, i.e. supposedly integrated in the structure. Extraction is (sometimes) allowed in these cases, whereas it is never found in a peripheral adverbial clauses which, following Haegeman's (2012) hypothesis, are not sufficiently integrated in the main clause. In any case, there is a clear degradation if extraction takes place from peripheral instead of central adverbial clauses, as shown by the asymmetries between (47) and (48).³¹

- (47) a. the details and the whole, which an artist cannot be great unless he reconciles
(Jespersen 1931, cited in Haegeman 2004: 70)
- b. a stranger, from that remote and barbarian Isle which the Imperial Roman
shivered when he named, paused. (Poutsma 1926, cited in Haegeman 2004: 70)
- (48) a. ??This is the paper which I enjoyed the conference very much, whereas I
disliked _ .
- b. ??This is the paper which, whereas I disliked _ , I enjoyed the conference very
much. (Postal 2001, cited in Haegeman 2004: 71)

Notice that these differences are not related only to extraction, but also to parasitic gap licensing, which is allowed only with central adverbial clauses and not with peripheral ones.

- (49) a. This is the paper which I memorised _ while I was copying _ .
- b. ?This is the paper which I myself enjoyed _ very much, while/whereas you
will probably dislike _ . (Postal 2001, cited in Haegeman 2004: 71)

³¹ This is supported by experimental data from Swedish discussed in Müller (2017, 2019), who shows that the level of syntactic integration of the adjunct is capable of influencing extraction possibilities, confirming that it is (sometimes) possible with central adverbial clauses such as those introduced by *in order to* or *after*, but it is never possible with peripheral ones such as those introduced by *although*. As discussed in Müller, however, this is not a sufficient factor. See §5.2.2 for an overview of the study.

An explanation of such asymmetry may be connected to the fact that peripheral clauses are not sufficiently integrated with their host clause and therefore they are too independent to allow extraction, violating some sort of local relation. A second possibility has to do with the different external syntax of adjuncts, i.e. their level of attachment. As already proposed by Haegeman (2012), central adverbial clauses are attached lower in the structure whereas peripheral adverbial clauses are higher and external to the main sentence. I will take into account this possibility and discuss it further in chapter 5. The third option is that their internal syntax has to be blamed, in the sense that in peripheral adverbial clauses it is too rich given the presence of illocutionary force, which is not present in adverbial clauses of the central type.

3.2.8 Position of the adjunct in the sentence

The last factor I am going to take into account is the position of the adjunct in the sentence. Once again, this is crucial with respect to extraction phenomena. As I will show, only if extraction takes place from an adjunct in a postverbal position it is allowed; if it takes place from a preverbal position the island effect is much stronger. This can be seen in the asymmetries between (50b) and (51b).

- (50) a. Gianni è partito senza salutare Maria.
 ‘John left without greeting Mary.’
 b. Quale ragazza Gianni è partito [senza salutare _]?
 ‘Which girl did John left without greeting?’
- (51) a. Senza salutare Maria, Gianni è partito.
 ‘Without greeting Mary, John left.’
 b. *Quale ragazza, [senza salutare _], Gianni è partito?
 ‘Which girl, without greeting, did John left?’

In a sentence like (50a) the adjunct is in a postverbal position – as seen in §2.2.1, this is the most common placement in the structure in Italian, where adjuncts are usually placed at the beginning of a sentence or – mostly – at the end (see Diessel 2001; Valmala 2009). In such a case, the adjunct does not have to be intonationally separated from the main clause and extraction is licensed, as in (50b). Notice that the same cannot be said for (51) where the adjunct

is in a preverbal position: it requires an appropriate intonation, and extraction leads to an ungrammatical result (51b). If we change the order of the adjuncts in Italian, we clearly obtain a stronger violation when the adjunct is in the middle or in the beginning rather than, most naturally, at the end of the sentence. Notice that in the cases above it is not a matter of syntactic integration of the adjunct. In fact, if we change the position of a conjunct like *without* we do not obtain a different reading: the adjunct is always related to the event of the host clause, i.e. it is an adverbial clause of the central type.

Interestingly, the same distinction holds for Spanish, as noted by Uriagereka (2011)³²:

- (52) a. ?Ese es el partido que lloró (amargamente) [por haber perdido _].
 this is the game that he cried bitterly for having lost
 ‘That’s the game that he cried (bitterly) for having lost.’
- b. ??Ese es el partido que [por haber perdido _] lloró (amargamente).
 this is the game that for having lost he cried bitterly
 ‘That’s the game that for having lost he cried (bitterly).’

(Uriagereka 2011: 11)

Moreover, the same asymmetry is found in English. Examples in (53) were used in an experimental study conducted in English, whose aim was to test the island effect of different domains (Heestand et al. 2012).

- (53) a. *This is the dish that, [although the chef overcooked _], the guests were not upset.
- b. This is the dish that, although the chef overcooked the sauce, the guests enjoyed. (Heestand et al. 2012: 147)

The ratings of adjuncts were surprising: both the supposedly grammatical control sentence like (53b) and the island violation in (53a) were rated very poorly and sentences without island violations had only a slight advantage over sentences containing an island. However, the placement of the adjunct is not the usual one, given that it is realised in the middle of the

³² Notice that not even in the Spanish case in (52) it is possible to connect the unacceptability to a peripheral reading: purpose clauses are classified as central adverbial clauses (Haegeman 2012).

sentence. Most likely, such a preverbal position might explain the lack of amelioration in the control items. Moreover, if we follow Haegeman (2004, 2012) an additional explanation may hold, namely the fact that the syntactic integration of the adjunct is not sufficient and for this reason extraction is not licensed. In fact, *although* is one of the conjuncts which only has a peripheral reading (Haegeman 2004). However, notice that the lack of syntactic integration is not able to account for the bad results of the control item in (53b).

The data presented here therefore show that the placement of the adjunct in the clause is one of the factors able to affect the island effect. I will account for such an asymmetry in this section and offer a first account, since this factor will not be considered in the following chapters. The explanation I am considering here regards the interpretation that adjuncts are connected with. I will explore the possibility that they can obtain a status different than their normal adverbial clause one, and ‘acquire’ a parenthetical status. These two interpretations predict different behaviours with respect to extraction phenomena, which is always impossible if adjuncts are interpreted as parenthetical.

Let me first take a step back in order to introduce parenthetical structures, which are fragment of discourse that can be interpolated in a sentence. They can have different forms and be simple words, as in the case of (54a) or more complex structures as (54c).

- (54)
- a. John, probably, likes to go to parties.
 - b. John, you know, likes to go to parties.
 - c. John, as everybody knows, likes to go to parties.

Crucially, parenthetical structures share many properties with adjuncts. I will briefly review here the main points. First of all, neither parentheticals nor adjuncts can occur in isolation, as can be seen in sentences in (55c) and (56c) respectively. Rather, they must be inserted in a matrix clause in order for the sentence to be grammatical.

- (55)
- a. A Elia, come tutti sanno, piace cucinare.
‘Elia, as everybody knows, likes to cook.’
 - b. A Elia piace cucinare.
‘Elia likes to cook.’
 - c. *Come tutti sanno.
*‘As everybody knows.’

- (56) a. Alice è andata a Londra perché Noemi l'aveva invitata.
 'Alice went to London because Noemi invited her'.
 b. Alice è andata a Londra.
 'Alice went to London.'
 c. *Perché Noemi l'aveva invitata.
 *'Because Noemi invited her.'

On the other hand, the matrix clause does not need a supplement or an adjunct, as can be seen in (55b) and (56b), but it can be accompanied by an extra level of the sentence which has the effect of adding or changing meaning and information of that sentence. Notice that an adjunct is never required in a sentence, as in the case of a parenthetical, and this means that it is not selected by anything, contrary to arguments.

Parentheticals and adjuncts also share distributional properties: the former can be realised in different positions in the sentence, with the same intonation and the same meaning. Adjuncts behave in a similar way: their most common position in a language like Italian or English is lower in the sentence (Diessel 2001; Valmala 2009), but they can be moved:

- (57) a. Giulia ha risparmiato per un anno per viaggiare in Sudamerica.
 'Giulia saved money for a year to travel to South America.'
 b. Giulia, per viaggiare in Sudamerica, ha risparmiato per un anno.
 'Giulia, to travel to South America, saved money for a year.'
 c. Per viaggiare in Sudamerica, Giulia ha risparmiato per un anno.
 'To travel to South America, Giulia saved money for a year.'

Crucially, when the adjunct is realised in a preverbal position, the intonation connected to it is different: it is separated from the main clause with an intonational break. Usually, if the adjunct appears in a sentence initial position, it is separated from the matrix clause with a comma in written text, or with a prosodic break in speech (see also Valmala 2009). This is similar not only to peripheral adverbial clauses, but also to parenthetical structures. These are all realized with a pause between what precedes and what follows, the so-called comma intonation (Selkirk, 2005). This is the intonation found in (57b) and (57c), and not obligatorily in (57a).

Moreover, a peculiar fact regarding parentheticals is that their truth value is independent from the truth value of the host (see Giorgi 2014, 2016), as can be seen in (58a): even though

the supplement is false, the sentence remains true. If we apply the same test to an adverbial clause, we can notice that it works in the same way: the truth value of the adjunct can be detached from that of the clause. Notice at this point sentence (58b): the adjunct is false, but the main sentence is still true. This argument supports the claim that adjuncts are – or can be – largely independent.

- (58) a. The Romans, who arrived before one hundred AD, found a land of wooded hills. (Selkirk 2005, cited in Giorgi 2016: 473)
- b. The Romans found a land of wooded hills after they arrived before one hundred AD.

However, parentheticals have properties that cannot be extended to adjuncts. In particular, parentheticals can be defined as illocutionary independent from their hosts (Giorgi 2014, 2016), as can be seen in (59). The same effect cannot possibly be adopted for prepositional adjuncts, especially in case they are central ones. As Haegeman (2004, 2012) shows, only peripheral adverbial clauses have the complete CP realised and show main clause phenomena.

- (59) a. She may have her parents with her, in which case where am I going to sleep? (Huddleston & Pullum 2002, cited in Giorgi 2016: 473)
- b. My friend, who God forbid you should ever meet, ... (Andrews 1975, cited in Giorgi 2016: 473)

Therefore, there is not a complete overlapping of properties and behaviours: adjuncts cannot actually be parenthetical structures, but they can ‘adopt’ such an interpretation when they are moved in a position different than their base one. The parenthetical interpretation can easily explain the impossibility of extraction from preverbal adjuncts, independently from the approach that one adopts. For example, if we follow Giorgi’s (2014, 2016) analysis, the issue is connected to the presence in the syntax of a prosody oriented head corresponding to the [+comma] feature (Selkirk, 2005). In this case, comma is a head K projecting its own constituent, which contains the parenthetical and which is read off at the interface with prosody as the characteristic comma intonation of parentheticals. Crucially, the content of the K head is not lexical, but a prosodic oriented one: KPs are not permeable to movement. This means that

extraction can never take place out of parentheticals. The parenthetical interpretation holds also with different theories, such as those proposing that parentheticals are non-integrated (Espinal 1991; Burton-Roberts 2006), which claim that these structures are totally external to the syntactic structure of the main clause, resembling Chomsky's (2004) proposal of separate planes. In this case, it is clear that extraction cannot be licensed because the two clauses would not be placed on the same plane.³³

There is also an additional, crucial distinction between adjuncts occurring in a sentence final or sentence initial position that can be explained with a parenthetical interpretation. Only the latter present Condition C violation, as can be seen from sentences in (60) (see also Valmala 2009: 952):

- (60) a. Se Gianni_i arriva tardi, pro_i viene in macchina.
 'If John_i arrives late, he_i'll go by car.'
 b. *pro_i viene in macchina, se Gianni_i arriva tardi.
 *'He_i'll go by car, if John_i arrives late.'

The asymmetries between (60a) and (60b) show that there is a difference in the position of the adjunct because only in (60b) the pronominal subject of the matrix binds the R-expression in the adjunct. This means that the adjunct is in a position low-enough to permit binding between the two elements. The same is not true for (60a), where no Condition C effect arise. This can be explained with a parenthetical interpretation, since the adjunct would be outside of the clause.³⁴ Notice also that in some cases the interpretation of an adjunct changes depending on its position in the sentence. It is the case of (61), as noted by Cinque (1990: 66).

- (61) a. Per questa ragione, hanno arrestato Gianni.
 'Because of this reason, they arrested John.'

³³ An analysis of parenthetical structures goes beyond the aims of this work, so I will not dive in any further.

³⁴ See Borgonovo & Valmala (2009) and Valmala (2009) for a different approach for sentence initial adjuncts. In their account, whenever adverbial structures appear at the beginning of a sentence, they are either interpreted as topics or foci and therefore they occupy the respective positions claimed since Rizzi (1997). If we are dealing with topics, they can either be base generated or moved there, whereas foci are moved from the post-verbal position. This means that adjuncts in a sentence initial position occupy a specifier position, rather than an adjoined one.

- b. Hanno arrestato Gianni per questa ragione.
 ‘They have arrested John for this reason.’

The fact that (61a) and (61b) do not share the same meaning suggests that the former cannot be reconstructed into the same position it occupies in (61b), i.e. a TP internal one. A parenthetical interpretation could easily explain this behaviour.³⁵ Notice that in (61a) the adjunct is separated from the rest of the sentence with a comma.

Notice that there might also be a different explanation able to account for the asymmetry of the placement of the adjunct. Assuming that the low position is the base one for adjuncts, one might follow Takahashi (1994) and subsequent literature, and state that extraction is never possible out of a previously moved domain. This could also explain the higher degree of unacceptability of the case in (51b), repeated here as (62b). But notice what happens in case the adjunct in its base position gets realised with comma intonation, in (63b):

- (62) a. Senza salutare Maria, Gianni è partito.
 ‘Without greeting Mary, John left.’
 b. *Quale ragazza, [senza salutare _], Gianni è partito?
 ‘Which girl, without greeting, did John left?’
- (63) a. Quale ragazza Gianni è partito [senza salutare _]?
 which girl Gianni is left without greeting
 b. *Quale ragazza Gianni è partito, [senza salutare _]?

Contrary to (63a), extraction is not accepted. This means that it is not simply a matter of moved versus unmoved domains, otherwise we would expect that there is no difference in the acceptability of sentences in (63), where both adjuncts are in a low and (supposedly) unmoved position and the only difference has to do with the prosody connected to them.³⁶

³⁵ In Cinque (1990), (61a) is analysed as an instance of clitic left dislocation, which does not involve *wh*-movement and cannot reconstruct since the empty category contained in TP is not antecedent-governed because its antecedent is too far away.

³⁶ Of course, such an argument is more or less valid depending on the syntactic account of parenthetical structure, i.e. if they are taken to be moved domain or not. In the former case, the account of Takahashi (1994) would still be correct.

To sum up, I argue that adjuncts are usually realized in a low, postverbal position, which I propose here is their base position (following Diessel 2001; Valmala 2009). Whenever they are in a low position they can simply be considered adjuncts, in fact they are not connected to a special prosody, and if extraction takes place from such a position the sentence is either fine or less degraded. On the other hand, if they are realised in a preverbal position their interpretation is not that of a normal adjunct but more similar to a parenthetical one. In fact, in such cases extraction is not licensed, as shown in cross-linguistic examples.³⁷

3.3.9 *Summing up*

This chapter brought some empirical data showing that extraction from adjuncts can be licensed, under certain circumstances and in different languages. In particular, I claim that extraction from the class of prepositional adjuncts is acceptable, but that it is affected by several factors, which are able to influence the island effect making it weaker/non-existent or stronger. I explored and discussed some of these factors, showing that they are at work in different languages and with different types of islands – whereas others are specific of adjunct domains.³⁸

³⁷ However, notice that there are some counterexamples to the impossibility of extraction from a preverbal position (see Etxepare 1998; Hornstein 2001):

- (i) Que libro dijiste [que si Ricardo leía alguna vez] abandonaría la Lingüística?
‘Which book did you say that if Ricardo ever read he would abandon linguistics?’
 - (ii) *Que libro quieres [que [si algún lee _]] abandone la Lingüística ?
‘Which book do you desire that if anyone reads, he would abandon linguistics?’
 - (iii) Which book did you say that [if Ricardo ever read _] he would abandon linguistics?
 - (iv) *Which book did you say that Quinn would abandon linguistics [if he ever read _]?
- (Hornstein 2001: 91)

As can be seen in (ii) and (iv) extraction is not acceptable if the adjunct occupies a final position. Notice that these sentences deal with conditional clauses only: it may be possible that these adverbial structures are different than others. I leave this issue open for now.

³⁸ A factor I did not mention is adjacency, which seems to facilitate extraction. This is at work in Spanish (Fábregas & Jiménez-Fernández 2016a), as seen in the contrast in (i) and (ii).

- (i) ¿Qué llegó [silbando _] María?
- (ii) *¿Qué llegó María [silbando_]?
what arrived María whistling
‘What did María arrive whistling?’ (Fábregas & Jiménez-Fernández 2016a: 1315)

The only difference between (i) and (ii) regards the intervention of the subject, which in (ii) is between the verb of the adjunct and the matrix one. Fábregas & Jiménez-Fernández (2016a) attributes this contrast to the need of adjacency, but it may well be that the contrast relies rather on the position of the subject. I thank Anna Cardinaletti (p.c.) for this comment.

A similar contrast is discussed in Borgonovo & Neeleman (2000), but in their examples intervention of adverbs is taken into account. They noticed that if an adverb intervenes between

Therefore, given the possibilities of extraction explored and presented in this chapter, we can draw an important conclusion with respect to the theories discussed in §2.3: they are too strong, in that none of them is able to predict or account neither the possibility of extraction, nor the factors that are able to manipulate the island effect. Moreover, there are also other data that I did not consider here, which show that adjuncts are inserted in a way that is able to feed movement operations.³⁹ These facts are a serious drawback to the theoretical approaches in §2.3, which all predicted that extraction is never allowed – except for the Connectedness theory (Kayne 1983), which has the opposite issue: it allows too much. We know that there has to be a limit, giving that extraction from adjuncts is not usually allowed, and is definitely degraded with respect to extraction from complements. We could follow two paths: either we establish that the theories mentioned have to be completely discarded, or we save something from them. The second option would mean creating escape hatches that allow for some elements to exit the island domain (see the edge feature in Chomsky 2000, 2001), and hence rescue the island. For instance, taking Chomsky’s (2004) Pair-Merge, we may assume that there is something able to connect them to the main clause, creating a bridge to the island and letting something cross it. The danger of such method is making ad hoc mechanisms that cannot be generalised to different constructions. I will leave aside these questions for now, and come back to them in chapter 5.

Second, notice that the possibility of extraction described in this chapter confirms that we cannot really think of syntactic domains as *absolute* islands, i.e. that they block extraction at any cost and in any circumstances. Islands are selective, and adjuncts do what we expect under such a view: they allow movement in certain circumstances, when the factors described in this chapter are controlled for. This also leads to an additional questions, namely the possibility of considering adjuncts as weak islands instead of strong ones (see also Tanaka 2015;

the matrix verb and the adjunct one, extraction is degraded. They attribute such an effect to the position of the adjunct in the structure, claiming that for an adjunct to be transparent, it has to be lower than adverbs.

(iii) What did John die *(peacefully) [whistling _]?

(iv) What did John come back *(on Tuesday) [bringing _]?

However, this effect is discussed for bare present participial adverbials but not for prepositional adjuncts, where the verb of the adjunct and the matrix one can never be adjacent because of the presence of the conjunct introducing the adjunct. Moreover, it does not seem that there is an adjacency requirement between the matrix verb and the preposition introducing the adjunct, given the Italian example in (v).

(v) Quale ragazza Gianni è scappato velocemente [senza salutare _]?

‘Which girl did John quickly run away without greeting?’

³⁹ This is the case of reconstruction effects (see Johnson 2003; Oseki 2015).

Müller 2019 on this point). In fact, if they were strong islands as traditionally stated, it would be hard to explain why the factors above have an effect on extraction: we do not expect strong islands to be ameliorated thanks to the presence – or absence – of some elements. The ameliorating effect obtained with the factors described above points towards a different analysis: adjuncts are weak islands rather than strong ones – I will come back at this issue in chapter 5. Therefore, the theories mentioned in §2.3 need to be modified in order to accommodate extraction under certain circumstances. As stated by Phillips (2006: 797 footnote 1) “the acceptable island-violation examples may simply indicate that standard theories of islands need to be modified in order to capture fine-grained differences among islands, for example, by incorporating semantic or discourse properties. If that is the appropriate analysis, then such cases would not show that *wh*-dependencies can enter islands, but rather that we do not yet fully understand island constraints.” If these factors are able to ameliorate the island effect, it also means that the filler-gap dependency is created, otherwise if the parser never constructs such dependency and their acceptability should not vary thanks to these factors.

Last, I also want to stress the possibility that the islandhood of these domains is not connected to a so-called Adjunct Condition, but may actually be related to a cumulative effect due to different factors taken altogether. That is to say that the degradation of extraction should actually be attributed to the interplay of several constraints, namely the factors that were discussed in this chapter. This approach could be extended to other islands as well, as in the case of subjects (see Jiménez Fernández 2009; Haegeman, Jiménez-Fernández & Radford 2014). In fact, it has been discussed that for there are several factors able to affect the acceptability of extraction from subjects,⁴⁰ even though most of them are different from the ones presented in the course of this. I won't discuss any further this point, since an analysis of subject islands goes beyond the scope of this dissertation.

⁴⁰ See Bianchi & Chesi (2012) on this point.

4. EXPERIMENTS

In this chapter I adopt a different approach, focusing on empirical data from an experimental point of view, which is crucial given the types of sentences I am dealing with. In fact, as will be explained in §4.1, for these kinds of sentences a binary distinction is not sufficient and gradience has to be involved. Therefore, I will follow a more formal approach in order to come back to the theory in chapter 5. I will first introduce the role of experimental syntax, which has been developed in the last ten years for islands (see especially Sprouse 2007) and discuss the issues connected with such an approach. In §4.2 I will briefly present some experiments for Italian and other languages focusing on the investigation of extraction from adjuncts. I will then present my experiments, which will be discussed from §4.3 on. The three experiments I conducted investigate the (un)acceptability of extraction from prepositional adjuncts in Italian. This is the red line that characterizes each test, and to which some more factors are added: the role of intrusive resumptive pronoun (Experiment A and C), the role of tense (Experiment B) and the role of embedding (Experiment A). The results of two of these test were published (Experiment A and B, see Dal Farra 2019a, b) although here their analyses are expanded.

4.1 Experimental syntax

Acceptability judgments play a crucial role in generative syntax, since they offer us the empirical bases for the development of theories. The majority of these judgments has been based mainly on informal methods which allowed to easily collect the relevant data. This has been the main method for at least fifty years. With no doubt this is the easiest way to gather data, but not always the most adequate one, as will be explained in the chapter. In fact, in many cases the acceptability of a sentence is a clear binary phenomena: they are either black or white, hence the options of acceptability and unacceptability are sufficient in order to discuss the

relevant data.¹ However, things are not so simple. In some cases there is no clear cut division between fully acceptable and fully unacceptable: there is a gradience in between that tells us that binary judgments sometimes are an idealization. As stated already in Chomsky (1965, 1975: 131, italics mine), acceptability is more a matter of degree, rather than a binary phenomenon, and thus “an adequate linguistic theory will have to recognize *degrees* of grammaticalness”. This is especially true when certain complex structures are involved: they seem to require a much more gradient approach, in that speakers’ judgments are not always neat. This is the case of (some) islands, among other constructions. As discussed in chapter 3, the relevant empirical domain is made of subtle judgments between structures whose (un)acceptability varies a lot depending on many factors. Often these structures seem to vary depending on the speaker: what I consider a completely good sentence, to someone else is perceived as unacceptable – or vice versa – and many cases fall in between acceptability and unacceptability. Some judgments are more sound than others, and this is especially true for the unacceptable cases presented in §3.2. For such reason I used formal experimental data.

Before going on, let me explain something more about experimental syntax. The fact that some sentences are not simply acceptable or unacceptable, i.e. that a binary option is not the appropriate choice, shows that grammar is not (always) categorical. We should actually use a spectrum (see Schütze 2016), whose extremities are represented by completely unacceptable and perfectly acceptable, but with a rich partial (un)acceptability area in between. That middle was not often recognised in previous stages of the research.² It is sufficient to think about the fact that different violations have different strength effects: how can we account for such a difference if all we have is a clear-cut division between acceptable and unacceptable? Clearly, these could not be captured under such an approach.³

For these reasons, in the last twenty years many researches have focused on those middle points of the spectrum, especially thanks to experimental syntax. Linguists started using the

¹ Throughout this chapter I will mostly talk about acceptability judgments, which are different from comprehensibility judgments. The latter refer to the interpretability of a sentence, whereas acceptability refers to the well-formedness and naturalness of a sentence, i.e. factors deeply influenced by the syntactic form.

² On the contrary, it was explained by means of properties of extra-grammatical factors such as the limits of working memory or the attention span. Basically, middle points were in the middle not because of the grammar, but because of factors that had nothing to do with grammar itself, and which could change depending on the individual.

³ As shown by all the factors involved only in adjuncts and their role in the acceptability of extraction phenomena explained in §3.2.

formal methods and statistical analyses of experimental psychology. At the beginning of this new approach, the main questions linguists tried to answer were (i) whether the data informally collected were reliable and (ii) whether the use of more formal methodologies would lead to uncover certain differences or patterns that were still unnoticed with a traditional methodology. As far as (i) is considered, it has been showed that the answer is positive (see Sprouse & Almeida 2012). The answer to (ii) is also a positive one, crucially. Thus, formal methods have become more and more reliable over the past few years, and they are being used always more.

There are many reasons to prefer formally collected data over informal ones (see Chacón, forthcoming, for an overview). First, informally collected judgments could lead to results that are not completely reliable, in the sense that some variables could be left out or not adequately controlled for. Second, informal judgments could be susceptible to subconscious biases from either the linguist or the speaker. This is not the case of formal experimental methods, which are able to correct the majority of these factors by means of several mechanisms such as controlled stimuli, filler items, randomized presentation order, and multiple observations per condition across items. In this way it is possible to overcome the idiosyncratic properties of both stimuli and participants, and hit the goal of the study. Third, formally collected data help us diagnose whether different factors are really able to affect the acceptability of a sentence, as well as the degree. This is a crucial point: thanks to the design of the tests, we are potentially able to isolate and quantify the contribution of factors we are interested in. This is done by isolating certain factors that are to blame for the (un)acceptability of some sentences, and checking out how the acceptability is affected. This means that they could help us distinguish the effects of grammatical factors and those of extra-grammatical ones, thus allowing us to overcome this division. Forth, formal acceptability judgments can help us reveal a potential gradability, an effect which cannot be detected on the basis of informal judgments. Last, they can also help uncover variability between participants.

Of course, informally collected data are convenient and much more practical than formally collected ones. We cannot think of relying only to the latter, since this would go beyond our needs in the majority of cases and since it has been shown that informally judgments are reliable (see Sprouse & Almeida 2012). This is clear especially for sentences that are not particularly complex. On the other hand, more complex cases require a different technique, and that's where formal experiments come in. Their purpose is to move beyond informal judgments and collect more fine-grained data to shed light on several phenomena. We may consider them "the end point for linguists characterizing phenomena" (Chacón, forthcoming: 21).

I will use formally collected data for two reasons mainly. First, the possibility of isolating some factors and quantifying their contribution to the acceptability of a sentence. This is particularly helpful in the case of islands, where so many factors seem to be involved and able to affect the acceptability of these constructions, as discussed in §3.2. Therefore, the use of formal methodologies in this sense could help us enrich the empirical bases for a theory of islands. Second, the empirical domain of this test consists of subtle comparative judgments which are not clearly acceptable or unacceptable, they rather fall in between in a grey area of partial acceptability. Moreover, items differ not only according to the sentence, but sometimes also according to the participant.

This way of collecting data however has an unwelcomed result: that of crashing the differences among subjects. The amount of variation among speakers remains usually unexplored (but see Kush, Lohndal & Sprouse 2018, 2019; Bondevik 2018). This is so because the result of these judgments is an average. That average, based on mean values, helps us establish a standard grammar but this might be just an idealization, given the presence of individual differences. I will analyse this issue in details in chapter 6. In the meantime, notice that acceptability judgments studies are sometimes criticized because averaging the data of different speakers might obscure the fact that people have different I-languages (see Dikken, Bernstein, Tortora & Zanuttini 2007).

4.2 Previous experimental settings

In this section, I will briefly discuss some tests investigating islands and present their results, focusing only on adjuncts. Only one of these experiments is dedicated to Italian, whereas the other take into account different languages. This will be helpful also to show the different behaviours of adjuncts from a cross-linguistic point of view and with a formal methodology. Most of the tests I present take into account different island types, since it is not very common to test one type of islands only (yet). For this reason, the tests I will discuss here are rarely restricted to adjuncts, with some exceptions (see Müller 2017, 2019 and Bondevik 2018 for adjunct clauses in Swedish and Norwegian, respectively; Kohrt, Sorensen & Chacón 2018 for gerunds in English).

Adjunct islands have been formally investigated in several tests for English (see Heestand, Xiang & Polinski 2012; Polinsky, Clemens, Morgan, Xiang & Heestand 2013; Sprouse, Wagers & Phillips 2012; Sprouse, Caponigro, Greco & Cecchetto 2016). In Heestand

et al. (2012) and Polinsky et al. (2013) the effect of structure and of resumption was investigated in both declarative sentences and *wh*-questions. The adjuncts considered were introduced by *while* and *although*. They found that the difference between the control group and that with island violation was small, especially in the case of declaratives. An example of the items they presented is given in (1).

- (1) a. This is the dish that, although the chef overcooked __, the guests were not upset.
b. Which dish did Gina think that, although the chef overcooked __, the guests were not upset? (Heestand et al. 2012: 147)

Notice that the adjunct has a verb in its finite form (see §3.2.1), and moreover the position of the adjunct is preverbal. As shown in §3.2.8, this is not a common placement for adjuncts in English and in Italian (see Diessel 2001; Valmala 2009). The fact that this is not the base position is also signalled through punctuation: the adjunct is preceded and followed by commas – as is the case of parentheticals. Nonetheless, they noticed that the adjunct islands they investigated were very weak, pointing to the conclusion that adjuncts are a heterogeneous group, for which different effects can be detected. Notice also that *although* is a peripheral adverbial clause, according to Haegeman’s (2012) classification, whereas *while* can have a double reading. Since no examples with this adjunct are given, we cannot establish which reading was used.

Sprouse et al. (2012) discuss a similar result, in the sense that they did not find a significant effect of structure in adjuncts similar to (2a).⁴ Sprouse et al. (2016) investigated the acceptability of extraction from adjuncts with both relative-clause formations (2b), and *wh*-dependencies (2a). In the former case, they did not find an effect.⁵

- (2) a. What do you worries [if the lawyer forgets __ at the office]?
b. I called the client who the secretary worries [if the lawyer insults __].
(Sprouse et al. 2016: 318-319)

⁴ Even though they found a significant interaction effect of structure and length, as expected under Sprouse’s (2007) factorial design.

⁵ Notice that this result is particularly puzzling: the island effect is present with one type of A’ dependency, i.e. *wh*-question, but not with relativization. However, we expect that all A’ dependencies should be equally degraded when crossing an island domain.

In both sentences the position of the adjunct is the expected one, i.e. postverbal, but notice that the adjunct contains a finite verb, and there is an intervening subject. Moreover, the *wh*-constituent is a simple one (§3.2.3).

Müller (2017, 2019) focused on extraction from different types of adjuncts in Swedish: *in order to*, *after*, *although* and *because*. Some of the factors mentioned in §3.2 were controlled for, in particular their level of syntactic integration, and these were shown to have a significant effect. She found out that extraction is unacceptable only in some cases, namely when (i) the adverbial clause is a peripheral one, as in the case of *although*, (ii) the relation expressed between the matrix clause and the adjunct is not contingent, i.e. not expressing a causality or enablement (see Truswell 2007, 2011), as in *although* and some cases of *after*, and (iii) the internal syntax of an adjunct is presumably too rich, as in the case of *because*. I will discuss in more details her test and the results in chapter 5.

Kush et al. (2019) investigate various islands in Norwegian, and the adjunct they focus on is introduced by *if*. When testing the role of focalization, they do not find an effect, differently from the case *wh*-questions (see Kush et al. 2018). Based on their results, Bondevik (2018) analyses extraction from different adjunct clauses in Norwegian, introduced by *if*, *when* and *because*. She found that there is a clear difference between them, for both the general acceptability and the variability within and between participants. The data briefly presented here will be picked up again and discussed in more details in §5.1.2, when dealing with the differences among adjuncts.

The island effect of adjuncts in Italian has been formally investigated only in Sprouse et al. (2016). They tested different types of islands – *whether*, subject, adjunct and complex NP – with both *wh*-questions and relative clauses. In this experiment, sentences containing extraction from adjuncts were judged as completely ungrammatical, for both *wh*- questions and relative clauses. However, such a result is deeply affected by the type of sentences that were presented to participants. Examples are given in (3a) for the *wh*-question and (3b) for the relative clause formation.

- (3) a. Cosa ti infuri [se uso _ in classe]?
what yourself infuriate if I use in class
'What are you going to get infuriated if I use in class?'

- b. Hanno intervistato il manager su cui ci abbasseranno lo stipendio
 (they) have interviewed the manager on whom to-us (they) will-lower the stipend
 [se Andrea farà nuove rivelazioni _].
 if Andrea will-make new revelations
 ‘They interviewed the manager who they will lower our salary if Andrea
 makes new revelations on.’ (Sprouse et al. 2016: 321, 323)

Let us analyse (3a). The unacceptability of this kind of question is expected for many reasons connected to the factors explained in §3.2. As I discussed before, these can deeply influence the islandhood of a domain, making the sentence much more unacceptable. This is exactly the case of (3a); first, notice the verb of the adjunct: it has a finite tense instead of non-finite one. This difference alone would already be enough to explain – and expect – the ungrammaticality of this example and, subsequently, of similar sentences adopted for the test.⁶ Second, the *wh*-constituent proposed is a simple one, which I have shown is not the best option. A complex one of the form *which* + NP helps the sentence to get better.

Thus, it seems that these factors are not usually taken into account, and therefore we already have a strong predictions regarding the low scores that participants assigned to the sentences. These factors are not controlled for because the traditional view ‘teaches’ us that adjuncts are strong islands, namely that extraction out of them is not acceptable. In several experiments whose aim is to test the role of extraction from adjuncts it is as if the worst options are taken into account, and therefore the island effect is more than expected. What is being checked in these cases is not the possibility of extraction from adjuncts, but its *impossibility*.

However, what does it change if we fix the factors that were applied to these sentences? Are these still unacceptable, or is there a difference? In the next sections I will answer these questions, presenting three formal experiments I conducted for Italian. Notice that the items I presented to participants were carefully prepared in order to obtain the weakest possible effect when dealing with the possibility of extraction from adjuncts. This means that the majority of factors discussed in §3.2 were manipulated and controlled for. Moreover they were applied in all sentences so that the only difference among them was the type of adjunct used, which were

⁶ A much better option is a sentence like (i), where the verb of the adjunct is non-finite:

- (i) Cosa non sei sicuro [se fare _]?
 what not (you) are sure if doing
 ‘What are you not sure if doing?’

all adverbial clauses of the central type, since, as discussed in §3.2.7, peripheral adverbial clauses are not integrated well enough to the main sentence (Haegeman 2004, 2012). In some cases I isolated some of the effects in order to check whether these would really be able to affect the (un)acceptability of the sentences, i.e. the presence of resumption and the effect of finite tense in the adjunct.

4.3 Experiment A

The aim of Experiment A was to investigate (i) the possibility of extraction from adjuncts in Italian, (ii) the effect of embedding, and (iii) the effect of resumptive pronouns. The first is the common question for all the experiments that will be presented in this chapter, and have been formally investigated for Italian only once (Sprouse et al. 2016). As I explained in §2.2, the possibility of extraction has a crucial role with respect to adjuncts, which are traditionally considered strong islands and hence accounted for with very strong analyses. The claim I made throughout chapter 3 is that extraction from prepositional adjuncts in Italian – as well as for other languages – is sometimes licensed and that acceptability depends on several factors able to influence the strength of the island effect. One of these is the presence of resumptive pronouns (see §3.2.4), which is also one of the mechanisms that can circumvent the island effect (see §2.1.4). Their role in Italian has been formally investigated only once before (Beltrama & Xiang 2016). The second effect I am going to consider regards the length of the filler-gap dependency, i.e. the role of embedding, which is supposed to negatively influence the acceptability of the sentence, especially in processing accounts. In fact, given that the length of the filler-gap dependency increases, working memory has to maintain active the information for a bigger amount of time and therefore the effort is stronger.

4.3.1 A closer look to resumption and its role in extraction

In order to better explain the role of resumptive pronouns, I will first briefly describe them. We can define resumptive pronouns (RPs) as obligatory bound pronouns appearing in the tail position of a filler-gap dependency, and obligatorily interpreted as co-referent with the filler of the dependency (McCloskey 2006). They are often said to be a rescuing mechanisms for islands – in Boeckx’s (2012) terms, they have a circumvention effect. However the best way to capture their effect is to distinguish between a *rescuing* effect, and an *ameliorating* one. Such a

distinction relies on the type of language we are considering. The presence of resumptives can be considered as a mechanism for island circumvention in languages with grammatical RPs, where they can freely alternate with gaps in most long-distance dependencies, and they are obligatory in island contexts. It is the case of languages like Irish, Hebrew and Arabic (Aoun, Choueiri & Hornstein 2001; Shlonsky 1992; McCloskey 2006). In these cases, it is an example of island circumvention because the presence of RPs inside an island always lead to a grammatical result (4a), whereas the gap to an unacceptable sentence (4b).

- (4) a. teach nach n-aithneochthá [cá rabh sé].
house not recognize where was it
‘A house_i that you wouldn’t recognize where it_i was.’
- b. *teach nach n-aithneochthá [cá rabh _].
house not recognize where was (McCloskey 2006: 99)

On the other hand, resumption can be considered an ameliorating factor in languages where RPs are not licensed by the grammar – they are *intrusive* resumptive pronouns – but nonetheless they appear in some contexts, especially in spontaneous speech. This is the case of English and Italian. In these languages, the presence of RPs is often reported to improve the status of sentences containing an island (Ross 1967; Kroch 1981; Asudeh 2011 among others for English, Bianchi 2004; Belletti 2006 for Italian).

- (5) a. The guy who I hate almost everything **he** does.
b. *The guy who I hate almost everything _ does. (Kroch 1981: 125)

- (6) a. ?L’uomo che sono certa del fatto che **gli** parleranno ...
the man to whom I am sure of the fact that to him they will talk
b. *L’uomo che sono certa del fatto che _ parleranno ...
the man to whom I am sure of the fact that they will talk
(Belletti 2006: 130)

The effect of resumption can be seen comparing (5a) with (5b) and (6a) with (6b) for English and Italian respectively. Notice that standard Italian does not form relative clauses with the resumption strategy (Belletti 2006), and hence these kinds of sentences are accepted at a

substandard level. It is therefore striking that its presence is able to ameliorate the sentence. However, the effect of resumption is only reportedly ameliorating: the judgments given have been informally-collected, but are not investigated by means of formal experiments, for both grammatical and intrusive resumptives. Once formal studies have been introduced it turned out that their effect was rarely significant, i.e. it was not detected in several tests (Ferreira & Swets 2005; Heestand et al. 2012; Polinsky et al. 2013; Beltrama & Xiang 2016).

The effect of resumptive pronouns of the intrusive kinds has been largely studied for English, with several methods: study of corpora (Prince, 1990), forced choice tasks (Ackerman et al. 2014), acceptability judgments (Alexopoulou & Keller 2007; Heestand et al. 2012; Polinsky et al. 2013; Beltrama & Xiang 2016) and comprehensibility judgments (Beltrama & Xiang, 2016). I will briefly explain the results from these tests, in order to show that there is a contrast between speakers' production and acceptability which makes it difficult to detect the ameliorating effect of RPs in island contexts.

Intrusive resumptive pronouns are consistently found in spontaneous speech and laboratory-based speech production studies. From a study on corpora, Prince (1990) found several examples of RPs in islands. Her major claim is that these elements “are ways of salvaging a sentence that the speaker has started without realizing that it is impossible or at least difficult to finish it grammatically” (Prince 1990: 483). Basically, their role and ability is “making the best out of a bad job” (see also Kroch 1981). In fact, in this study the 67% of spontaneously produced RPs occurred in island conditions. Example (7) is taken from one of these corpora production and displays a case of extraction from an adverbial clause in English.

- (7) Apparently, there are such things as bees in the area which if you are stung by **them**, you die. (Prince 1990: 483)

In a forced choice task, Ackermann et al. (2014) tested the effect of RPs using two forced choices: in the first participants were asked to complete a sentence by selecting a segment containing either a gap or a pronoun, whereas in the second they also had to choose the most acceptable option between a sentence with a RP and another one with a gap. They found out that the presence of resumptive pronouns is always the preferred option, in both tasks.

However, once the task assigned to participants is different than the ones mentioned above, in several experimental investigations the ameliorating effect was not found: islands with RPs are not more acceptable than their gapped counterparts in acceptability judgments.

The study of Ferreira & Swets (2005) is particularly interesting: the goal of the test was to cue participants to the production of sentences containing a resumptive pronoun in a *wh*-island, similar to sentence in (8):

(8) This is the donkey that I don't know where **it** lives. (Ferreira & Swets, 2005: 109)

The same participants who produced these sentences were later asked to judge the acceptability of these sentences, and those containing RPs were significantly less acceptable than the control sentences: RPs in islands were not accepted, even by the same speakers who produced them.⁷

Moreover, in several acceptability tests investigating their role in both online and offline tasks, RPs were never more acceptable than gaps. This is the case of Heestand et al. (2012) who tested the role of resumptives in Complex NP islands and adjuncts islands with an offline task, and in relative clauses in an online task. In all cases they did not find an effect, even though reaction times were a little bit faster in the RP condition; they made it easier to detect the unacceptability of these constructions. Polinsky et al. (2013) investigated their role in an online reading task and a reading judgment task on both declaratives and questions. They focused on relative clauses and adjuncts, and they confirmed that resumptives do not have a rescuing effect on islands, especially in the case of questions.

Evidence that the instructions of the test have a crucial role is found in Beltrama & Xiang (2016). In this test two different tasks were assigned to participants, namely to judge the *acceptability* or the *comprehensibility* of sentences containing island violations formed by NP relative clause and resumptive pronouns, as well as counterpart sentences containing gaps. Interestingly, the effect of RPs was significant for both Italian and English only in the comprehensibility task, but not in the acceptability one – which is usually the one employed for these tests. Moreover, this effect was present only when the target sentence was embedded in a conversation, therefore it seems that the role of context is crucial. When the task was to judge the acceptability of similar sentences, there was no difference between RPs and gaps.

⁷ Although a main issue with this test is that they didn't compare the acceptability of sentences containing a RP with their gapped counterpart; the comparison was only between sentences containing an island and a resumptive pronoun and grammatical sentences. Therefore, it is not particularly surprising that they found a significant effect.

4.3.2 Materials and procedure

For Experiment A I used a 2x2x2 factorial design, with 8 conditions resulting from fully crossing the three factors: (i) STRUCTURE; (ii) LENGTH and (iii) RESUMPTION. For the STRUCTURE factor, extraction from adjuncts was compared to parasitic gaps. Hence, the comparison was between unacceptable cases of extraction from an island and (supposedly) acceptable cases of parasitic gaps. This is different from many other experimental studies investigating the island effect, where the island condition is usually compared to extraction from a declarative sentence (see Sprouse 2007; Sprouse et al. 2016 among others). As seen in §3.1.1, parasitic gaps are constructions in which a single *wh*-phrase is associated with multiple gaps in a sentence, one inside an island, and the other outside. The former is parasitic on the latter, and the sentence is acceptable thanks to its presence. Since several experimental studies showed that these structures are fully acceptable and comparable to declarative sentences (see Phillips 2006; Wagers & Phillips 2009), I decided to use parasitic gaps in order to have minimal contrast with the sentences containing an island. As for the LENGTH factor, sentences with a short distance filler-gap dependency were compared to longer dependencies. Last, for the RESUMPTION factor, sentences with a gap were compared to sentences with a resumptive pronoun. See (9) for the entire paradigm of an item.

(9) a. Parasitic gap, short, gap

Quale ragazzo Silvia ha guardato _ senza salutare _?

which boy Silvia has looked without greeting

‘Which boy did Silvia look without greeting?’

b. Parasitic gap, short, resumption

Quale ragazzo Silvia ha guardato senza salutar**lo**?

which boy Silvia has looked without greeting him

‘Which boy did Silvia look without greeting him?’

c. Parasitic gap, long, gap

Quale ragazzo tutti dicono che Silvia ha guardato _ senza salutare _?

which boy everybody say that Silvia has looked without greeting

‘Which boy did everybody says that Silvia looked without greeting?’

d. Parasitic gap, long, resumption

Quale ragazzo tutti dicono che Silvia ha guardato senza salutar**lo**?

- which boy everybody say that Silvia has looked without greeting him
 ‘Which boy did everybody says that Silvia looked without greeting him?’
- e. Island, short, gap
 Quale ragazzo Silvia è partita senza salutare _?
 which boy Silvia is left without greeting
 ‘Which boy did Silvia leave without greeting?’
- f. Island, short, resumption
 Quale ragazzo Silvia è partita senza salutarlo?
 which boy Silvia is left without greeting him
 ‘Which boy did Silvia leave without greeting him?’
- g. Island, long, gap
 Quale ragazzo tutti dicono che Silvia è partita senza salutare _?
 which boy everybody say that Silvia is left without greeting
 ‘Which boy did everybody says that Silvia left without greeting?’
- h. Island, long, resumption
 Quale ragazzo tutti dicono che Silvia è partita senza salutarlo?
 which boy everybody say that Silvia is left without greeting him
 ‘Which boy did everybody says that Silvia left without greeting?’

The 8 conditions were tested in three types of prepositional adjuncts introduced by *dopo* ‘after’, *prima* ‘before’, and *senza* ‘without’, for a total of 24 items.

Fillers were also included in a ratio of 1:1, for a total of 24 fillers. Fillers were similar to items in the sense that they were of a comparable length. Moreover, I used fillers with varying acceptability: even though most of the sentences were either completely grammatical or ungrammatical, the acceptability of some was expected to fall somewhere in between. To do so, I presented varying degrees of acceptability, such as morpho-syntactic mismatches or word order mistakes. In this way participants are encouraged to use a large portion of the scale rather than focusing only on some parts. Notice also that fillers consisted of both declaratives and questions, which were included so that the target items would not be the only questions in the experiment.

Every subject was therefore tested on 48 items total. Items were presented in a pseudo-randomized order, so that the same condition never appeared twice in a row. The items were divided and distributed into two different blocks, Block A and Block B. These were presented

in different order among participants: some had to judge the sentences in Block A first, while others sentences of Block B. This allowed to obtain a more randomized order. All participants saw all stimuli.

Experiment A was a pilot study, meant to answer one question mainly, i.e. the possibility of extraction from (some) adjuncts. A positive answer would allow me to investigate further (and with more participants) this issue, including also different factors. I expected parasitic gaps to be a better option than islands, for all the adjuncts considered, and I also expected this difference to be significant. As for the other factors, my expectations were that longer embedding would be worse than shorter ones, and that the presence of resumptives would not help the islands, based on previous results.

Twelve people participated in the experiment. The subject pool was uniform: the age range was between 22 and 28 years old and all of the participants came from the North-East of Italy. All of them had at least a bachelor degree. They participated voluntarily in the experiment.

Items were presented in a written form. Participants were asked to judge the acceptability of the sentences presented on a 7-point Likert scale, where 7 indicates perfect acceptability, and 1 total unacceptability. Before beginning the experiment, participants were provided with instructions on how to use the scale: they were asked to rate 6 or 7 sentences they found perfectly acceptable, to give 1 or 2 to sentences they found completely unacceptable, and to assign 3-5 to sentences that were somewhere in between. They were also instructed to judge the sentences on the basis of their native-speaker intuition, rather than any prescriptive rules, and to go with their first instinct instead of spending time thinking about their answers. Moreover, the first five experimental items were used as a pre-test phase and were then excluded from the statistical analysis. The practice items were not marked as such: participants did not know these were practice items.

As for statistical analysis, raw ratings of each individual subject, including both target and filler items, were first transformed into *z*-scores in order to avoid potential scale biases between participants (for example, using only one end of the scale, or choosing different ranges of values among participants). Linear mixed-effects models were then run on the transformed data with the R statistical package *lme4* (Bates et al. 2015). The fixed effect predictors included STRUCTURE, LENGHT and RESUMPTION, as well as their interaction. The random effects included subjects and items. All predictors were sum coded before the data analysis, with [+ island], [+embedding] and [-resumption] coded as 1, and [- island], [- embedding], and [+ resumption] coded as -1.

4.3.3 Results

Transformed results from Experiment A are summarized in Table 4.1. Statistical analysis revealed that there is no main effect of RESUMPTION, i.e. there is no significant difference among sentences with resumptive pronouns and their counterparts with a gap, in any of the adjuncts considered: *after* ($\beta = -0.15$, $se = 0.08$, $p > 0.06$), *before* ($\beta = 0.01$, $se = 0.11$, $p > 0.92$) and *without* ($\beta = 0.15$, $se = 0.13$, $p > 0.91$). I will present the results for the gap and the resumptive model separately.

| | <i>after</i> | <i>before</i> | <i>without</i> |
|---------------------------|--------------|---------------|----------------|
| parasitic gap, short, gap | 0.68 | 0.43 | 0.98 |
| parasitic gap, short, RP | 1.08 | 1.28 | 1.13 |
| parasitic gap, long, gap | 0.42 | 0.85 | 0.42 |
| parasitic gap, long, RP | 0.72 | 0.45 | 0.52 |
| island, short, gap | -0.53 | -0.28 | 0.45 |
| island, short, RP | -0.57 | -0.43 | -0.54 |
| island, long, gap | -1.04 | -0.28 | 0.3 |
| island, long, RP | -0.86 | -0.43 | -0.31 |

Table 4.1. Means of *z*-score ratings for each condition and each adjunct of Experiment A.

Raw ratings results summarizing each adjunct type and condition are presented in Figures 4.1, 4.2 and 4.3, where the y-axis indicates average ratings and error bars indicate standard errors.

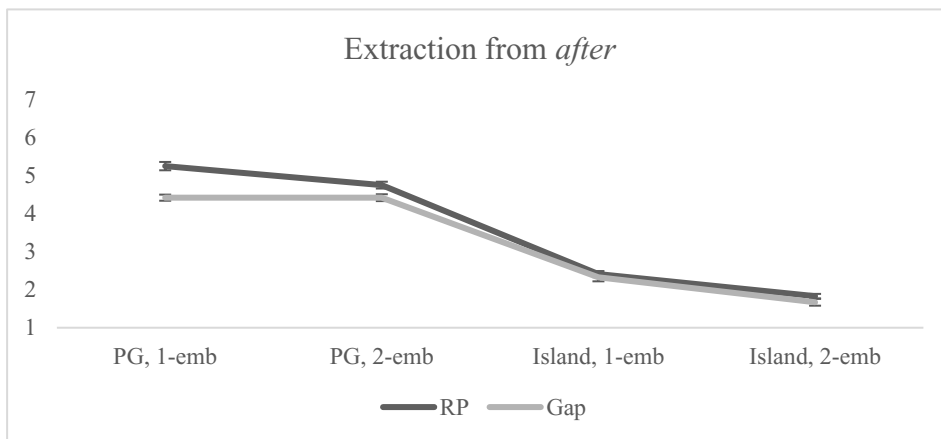


Figure 4.1. Raw acceptability judgments on a 1-7 scale in the adjunct introduced by *after* for each condition presented.

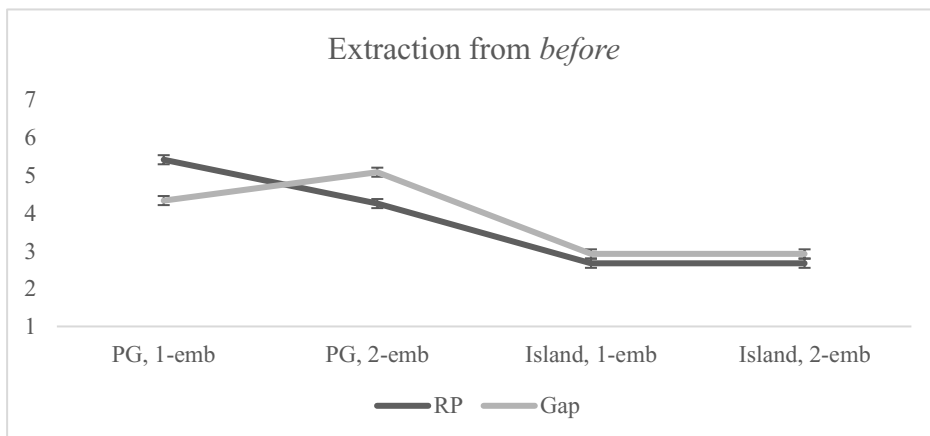


Figure 4.2. Raw acceptability judgments on a 1-7 scale in the adjunct introduced by *before* for each condition presented.

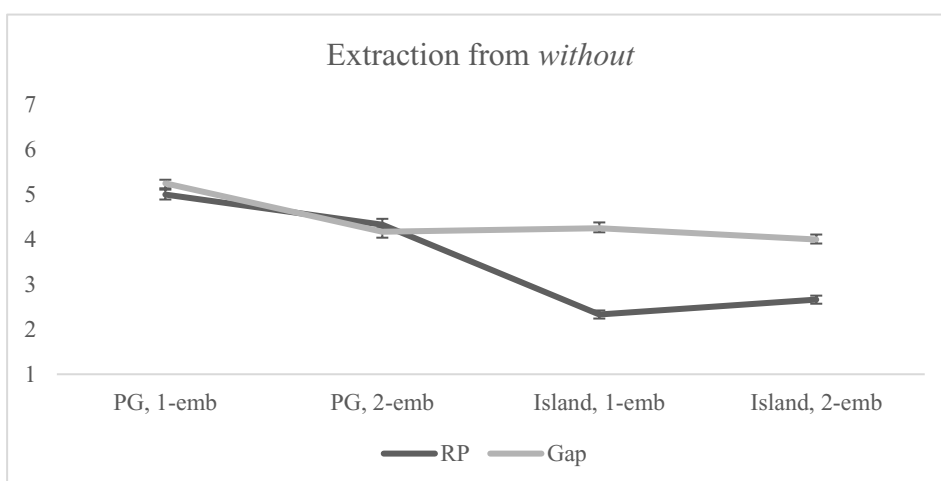


Figure 4.3. Raw acceptability judgments on a 1-7 scale in the adjunct introduced by *without* for each condition presented.

4.3.3.1 Structures with a gap

When the sentence is realized with a gap and not a resumptive pronoun, the effect of STRUCTURE is significant for both the adjunct introduced by *after* (linear mixed effect estimate: $\beta = -0.68$, $se = 0.11$, $p < 0.001$) and by *before* ($\beta = -0.46$, $se = 0.12$, $p < 0.004$): sentences containing an island were always rated significantly lower than their parasitic gap counterparts. Interestingly, the case of *without* is different: the effect of STRUCTURE is *not* significant ($\beta = -0.16$, $se = 0.08$, $p > 0.2$).

There is no main effect of LENGTH for any of the adjuncts analyzed: *after* ($\beta = -0.17$, $se = 0.11$, $p > 0.14$), *before* ($\beta = 0.10$, $se = 0.12$, $p > 0.40$) and *without* ($\beta = -0.18$, $se = 0.08$, $p > 0.06$). This means that there is no significant difference between long and short dependencies.

Interaction of STRUCTURE and LENGTH is not significant for any adjunct: *after* ($\beta = 0.02$, $se = 0.11$, $p > 0.83$), *before* ($\beta = -0.11$, $se = 0.12$, $p > 0.39$) and *without* ($\beta = 0.10$, $se = 0.08$, $p > 0.22$).

4.3.3.2 Structures with resumption

When the sentences are realized with a resumptive pronoun, there is a main effect of STRUCTURE for every adjunct: *after* ($\beta = -0.81$, $se = 0.09$, $p < 0.0003$), *before* ($\beta = -0.60$, $se = 0.12$, $p < 0.007$), and *without* ($\beta = -0.59$, $se = 0.11$, $p < 0.003$). In the case of *without* the presence of a resumptive pronoun actually makes the sentence significantly worse than its parasitic gap counterpart, differently from what happened in the gap structure.

LENGTH, on the other hand, does not have an effect for any adjunct: *after* ($\beta = -0.16$, $se = 0.09$, $p > 0.86$), *before* ($\beta = -0.19$, $se = 0.12$, $p > 0.15$) and *without* ($\beta = 0.12$, $se = 0.11$, $p > 0.28$), as well as the interaction of STRUCTURE and LENGTH (*after*: $\beta = 0.02$, $se = 0.09$, $p > 0.86$, *before*: $\beta = 0.19$, $se = 0.12$, $p > 0.14$ and *without*: $\beta = 0.17$, $se = 0.11$, $p > 0.13$).

4.3.4 Discussion

In this section I discuss the main results of the pilot study, focusing first on each factor and then on some interesting and unexpected results. I take into account the statistical analysis I presented in the previous sections and additional observations as well.

4.3.4.1 *The role of structure*

As seen in §4.3.3.2, in all of the adjuncts considered, the effect of STRUCTURE is significant in the resumptive model. This means that sentences containing the island effect are rated significantly lower than their parasitic gap counterpart. The first conclusion we can draw is that parasitic gap sentences can be used as acceptable counterparts of islands. Second, the island status of adjuncts does not change thanks to resumptive pronouns: they do not seem to be able to rescue the island violation, at least not significantly. I will discuss this result in a dedicated section.

The crucial result as far as the STRUCTURE factor is considered is from the gap model, and this will lead to an additional observation regarding the (not really) uniform status of adjuncts. Extraction from *after* and *before* is considered bad, significantly worse than its parasitic gap counterpart, i.e. there is a main effect of STRUCTURE. For both *after* and *before* this is true independently from the presence of a resumptive pronoun or of a gap. It is different in the case of *without*: when the adjunct is realised with a gap the sentence is not rated significantly worse than the parasitic gap. This means that there is no island effect, i.e. this sentence is not interpreted as an island for extraction. Such a striking result could also be connected to the small number of participants, and for this reason it will be investigated again in Experiment B and C. Crucially, the results obtained in the gap model tell us that the three adjuncts I took into account are not treated in the same way: they are connected with different results. I will discuss such an effect in §4.3.4.4.

4.3.4.2 *What did John think that Mary said about embedding?*

The prediction for the LENGTH factor was that if extraction took place from a more embedded adjunct, it would lead to a degraded result. However, this effect was not significant, for both the gap and the resumption model. The lack of an effect of LENGTH is unexpected under different approaches. For instance, this is the case of Sprouse (2007 and much subsequent literature), where the factorial design he developed predicts the island effect to be a result of two factors, i.e. structure and length. Taken together, the two determines a super-additive effect that ‘create’ the island. The effect of the length of the filler-gap dependency is also particularly important in reductionist approaches, which claims that the longer the distance from the gap,

the more difficult it is to connect gap and extracted element. Therefore, the more working memory has to be used and kept active, making the parsing yet more difficult.

However, in this test the length of the dependency is not significant: sentences with a longer filler-gap dependency are not rated worse than their counterparts with shorter dependencies. This could be connected to the fact that the depth of embedding was never too much: participants were presented with sentences that had at most two levels of embedding, as shown in (10b). Therefore, it may be the case that this factor was not stressed enough in order to show its effect on the island.

- (10) a. Quale ragazzo Silvia è partita [senza salutare _]?
which boy Silvia is left without greeting
'Which boy did Silvia leave without greeting?'
- b. Quale ragazzo [tutti dicono che [Silvia è partita senza salutare _]]?
which boy everybody say that Silvia is left without greeting
'Which boy did everybody says that Silvia left without greeting?'

Notice however that the same effect was tested in Beltrama & Xiang (2016) for 2 and 3 levels of embedding. Interestingly, they obtained the same result: the depth of embedding did not have any significant effect in the island model, for both Italian and English.⁸

4.3.4.3 *The role of resumption*

The prediction for the RESUMPTION effect was that the presence of RPs would not have rescued the island, at least not significantly. This prediction was based mainly on the results obtained in other tests assessing the acceptability of these constructions in languages with intrusive RPs (Heestand et al 2012; Polinsky et al. 2013; Beltrama & Xiang 2016). In fact, there is no significant effect: sentences containing a gap are not rated worse than their counterparts with RPs. This is true for both the case of *after* and *before*. The case of *without* has actually a different situation, in that when the RP is employed the sentence is worse than the counterpart with a gap. In this case, extraction from *without* is judged as unacceptable and presents a main

⁸ Although the depth of embedding came out as significant in the non-island model, i.e. in grammatical sentences.

effect of STRUCTURE, which was not found in the gap model.⁹ In a way, this is similar to the results found in Heestand et al. (2012), where the effect of resumption in islands was tested with an online task where it did not have a rescuing effect, but made it easier and faster for participants to judge the unacceptability of these sentences.

Therefore, the results obtained here are in line with results of many other tests investigating the acceptability of similar sentences (see Heestand et al 2012; Polinsky et al. 2013), in that the ameliorating effect was not detected. On the other hand, there was actually a degrading effect for one of the adjuncts. I will discuss in more details the lack of an effect of resumption in the third test, Experiment C. In fact, further investigations will be made to assess this point, extending both the sample of participants and that of adjuncts, given that those investigated in this test showed such different results from one participant to another.

4.3.4.4. *Differences among adjuncts*

A crucial result of this test regards the difference among the adjuncts taken into account. They behave differently with respect to extraction phenomena, which can be seen in the case of extraction from the island condition with no embedding and no resumption. In fact, extraction in the case of *without* is much more acceptable than in the other adjuncts, to the point that the effect of structure was not significant in the gap model, i.e. there was not an island effect. On the contrary, extraction from *after* and *before* lead to much degraded sentences, and thus it seems to be completely ruled out. Their distinction can be seen in Figure 4.4.

⁹ Notice that the same effect is not found in parasitic gap sentence: the presence of a resumptive pronoun here does not change the (un)acceptability of the sentence.

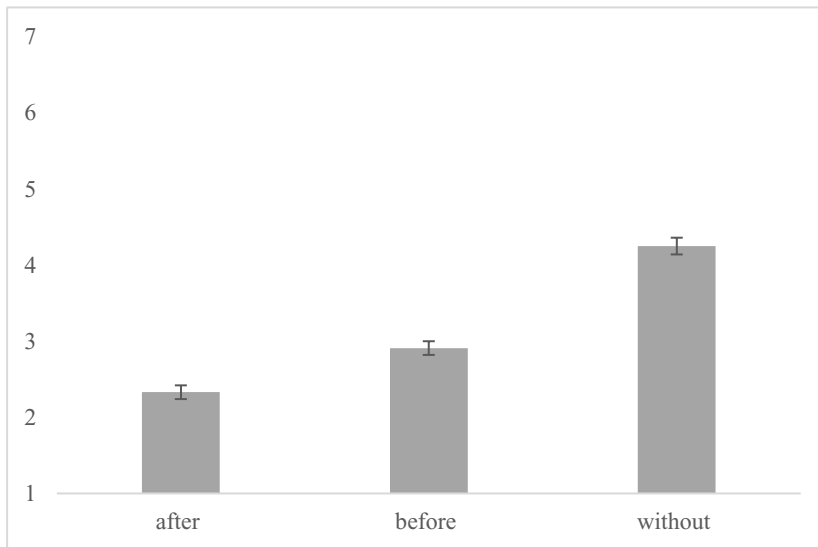


Figure 4.4. Raw acceptability judgments on a 1-7 scale representing the average ratings for the island condition, with short filler-gap dependency and no resumption.

Notice that their distinction depends entirely on the type of adverbial clauses. Samples of sentences used for the test are in (11).

- (11) a. Quali dolci Luca è scappato [dopo aver rubato _]?
 which sweets Luca is run away after having stolen
 ‘Which sweets did Luca run away after having stolen?’
- b. Quale libro Anna si è preoccupata [prima di trovare _]?
 which book Anna herself is worried before finding
 ‘Which book did Anna worry before finding?’
- c. Quale ragazzo Silvia è partita [senza salutare _]?
 which boy Silvia is left without greeting
 ‘Which boy did Silvia leave without greeting?’

In all the sentences in (11) the island condition with the gap and the shortest filler-gap dependency is presented. The sentences are equal to each other with respect to different factors: extraction takes place from a prepositional adjunct with the verb in its non-finite tense, extraction is of a DP argument and is realised with a complex *wh*-constituent. Moreover, the adjunct is in its final position, i.e. there should not be a parenthetical interpretation able to block extraction entirely. Therefore, all of the factors described in §3.2 are carefully manipulated in

order to obtain the weakest island effect possible.¹⁰ Moreover, the conjuncts used in these cases do not give rise to a central versus peripheral reading contrast: *after*, *before* and *without* modify only the events described in the matrix verb, and therefore we expect that they are sufficiently integrated in the host clause to license extraction.

Even more crucial, I believe, is the case of parasitic gap sentences. Remember that these sentences should always be grammatical, because the gap contained in the island is licensed thanks to the presence of the higher c-commanding gap outside the adjunct. That they are well-formed sentences has long been discussed in the literature on the topic, starting from Engdhal (1983). Hence my prediction for the sentences in question was that they would be considered equally acceptable by everybody. However, if we look at the results for the condition with no embedding and no resumption, a slight distinction can be seen among adjuncts.

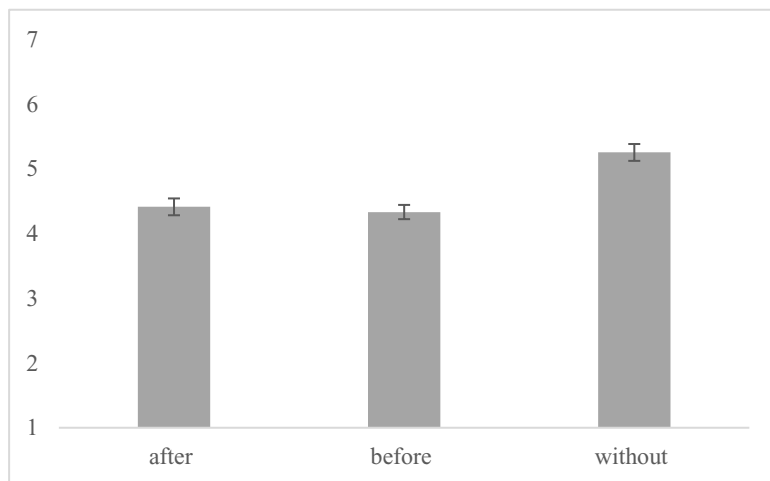


Figure 4.5. Raw acceptability judgments on a 1-7 scale representing the average ratings for the island condition, with short filler-gap dependency and no resumption.

Apparently, not all parasitic gaps are accepted in the same way. Crucially, the case of *without* is always the better option, whereas the cases of *after* and *before* are not so good, replicating the results of the island condition seen above. I will now present the examples used in the test, in (12).

¹⁰ Notice that the only difference relies in the type of matrix verb of (11b), which is different than the achievements described in (11a) and (11c), here we are dealing with a verb describing an activity.

- (12) a. Quali dolci Luca ha mangiato _ dopo aver rubato _ ?
 which sweets Luca has eaten after having stolen
 ‘Which sweets did Luca eat after having stolen?’
- b. Quale libro Anna ha cercato _ per mesi prima di trovare _ ?
 which book Anna has looked for months before finding
 ‘Which book did Anna look for months before finding?’
- c. Quale ragazzo Silvia ha guardato _ senza salutare _ ?
 which boy Silvia has looked without greeting
 ‘Which boy did Silvia look without greeting?’

As can be seen, sentences containing the parasitic gap are very similar, in that all of them have transitive verbs in the matrix and in the adverbial clause, the DP argument is extracted from prepositional adjuncts and it is the object of the matrix verb. Moreover, in all of the sentences the *wh*-constituent is a complex one, formed by *which* + NP. Finally, the verb in the adjunct is in its non-finite tense and in every sentence there is an overt subject shared by the matrix verb and the one in the adjunct. In (12b) there is an additional adjunct attached to the matrix sentence; this was the only difference among the items, and which could have had an effect.

Considering both the case of the island and the parasitic gap condition, and given that the structures are equal in most aspects, we should attribute the differences to the adjunct itself. As a crucial result of this pilot study, I claim that prepositional adjuncts are not a uniform group as is traditionally stated, but a heterogeneous one that could be sorted according to some criteria. Notice that the distinction I am talking about does not refer to the one individuated by Haegeman (2004, 2012), among peripheral and central adverbial clauses. In fact, the distinction regards only adjuncts of the central type, meaning that this group needs a more fine-grained analysis.

4.3.4 Summary

The main results from this study can thus be summarized as follows:

- (i) Resumptive pronouns do not help to rescue the island effect. On the contrary, they can strengthen the effect, at least in one case, i.e. *without*. Note that this may be connected to independent reasons, such as the fact that RPs are perceived as substandard in Italian;

- (ii) Different adjuncts lead to different islands effects;
- (iii) Different adjuncts lead also to different acceptability for parasitic gaps;
- (iv) The length of the dependency has no main effect.

Experiment A was a pilot test that allowed me to individuate the results summarized here, and particularly the difference among adjuncts with respect to extraction. Being a pilot test, the sample of participants was reduced. For this reason, the statistical effect might be different once a bigger sample is taken into account. Therefore, in the two tests that follow I extended the sample in order to obtain sounder results.

4.4 Experiment B

The aim of Experiment B was to check whether the conclusions discussed above hold by means of a larger sample. The test is slightly different, in that I diminished the number of dependent variables and tested only two conditions. The main questions this second experiment addresses are (i) is there an effect of structure? (ii) is the presence of finite tense in the adjunct so crucial for the acceptability of the sentence? And (iii) are there differences among adjuncts as in Experiment A? Materials and procedure are similar to those of the first experiment, and I investigated the same adjuncts I took into account for the first test. This choice allowed me to compare the results of the two tests, especially when it comes to the differences among adjuncts.

4.4.1 *Materials and procedure*

Experiment B has a 2x2 factorial design where two factors are crossed: STRUCTURE and TENSE. In the STRUCTURE factor, extraction from an adjunct is compared to extraction from a parasitic gap. For the TENSE factor, the embedded sentence contains either a finite tensed verb or a non-finite one. As discussed in §3.2.1, finite tense in the island domain is cross-linguistically considered one of the major factors affecting the possibility of extraction.¹¹ However, notice that the effect of finiteness is rarely tested with formally-collected judgments (see Michel & Goodall 2012).

¹¹ There are some counter-examples, especially as far as Slavic and Mainland Scandinavian languages are considered (see §3.2.1).

The four conditions are tested in three types of adjuncts, for a total of 12 sentences. As in Experiment A, adjuncts are introduced by *dopo* (after), *senza* (without), and *prima* (before). The examples in (13) provide the full paradigm for an item. Notice that once again factors were manipulated in order to obtain the weakest effect: extraction is always of an argument and not of an adjunct, and the *wh*-constituent is always a complex one.

- (13) a. Parasitic gap, Non-finite tense
 Quale ragazzo Silvia ha guardato _ senza salutare _ ?
 which boy Silvia has looked without greeting
 ‘Which boy did Silvia look without greeting?’
- b. Parasitic gap, finite tense
 Quale ragazzo Silvia ha guardato _ senza che salutasse _ ?
 which boy Silvia has looked without that she greeted
 ‘Which boy did Silvia look without that she greeted?’
- c. Island, non-finite tense
 Quale ragazzo Silvia è partita senza salutare _ ?
 which boy Silvia is left without greeting
 ‘Which boy did Silvia leave without greeting?’
- d. Island, finite tense
 Quale ragazzo Silvia è partita senza che salutasse _ ?
 which boy Silvia is left without that she greeted
 ‘Which boy did Silvia leave without that she greeted?’

There were two sentences for each condition in all adjuncts, for a total of 24 items. I included fillers in a ratio of 1:1. The total is 24 fillers, which are of comparable length and varying acceptability: most of them are either completely grammatical or ungrammatical, but the acceptability of some sentences falls in between. In doing so, participants are encouraged to use a large portion of the scale, instead of focusing only on its extremities. Notice that I included both declaratives and questions as fillers, so that the target items were not the only questions in the experiment.

Every subject was therefore tested on a total of 48 items, which were presented in a pseudo-randomized order, so that the same condition never appeared consecutively. The items were divided and distributed into two different blocks, Block A and Block B. These were

presented in different order among participants: some had to judge the sentences in Block A first, while others sentences of Block B. This allowed to obtain a more randomized order. All participants saw all stimuli.

The predictions for Experiment B were that the same difference among the adjuncts could be found, i.e. that extraction from *without* would turn out to be more acceptable than extraction from *after* and *before*. As for the role of TENSE, I expected it to be able to influence the acceptability of the sentences, uniformly across adjuncts.

Eighty-three native speakers of Italian participated in the experiment. They were undergraduate students at the University of Venice (age-range 18-25) with no training in formal syntax. Students were all enrolled in an introductory linguistics course, which did not introduce them to island effects. Furthermore, the experiment was administered prior to the discussion of syntactic structures. Before the experiment, participants were asked to provide some information regarding their age, gender, as well as their native language and dialect. Among these eighty-three, three were excluded from the statistical analysis because they had mistaken the scale.

Items were presented in a written form. Participants were asked to judge the acceptability of the sentences on a 7-point Likert scale, where 7 indicates perfect acceptability, and 1 total unacceptability. Before beginning the experiment, participants were provided with instructions on how to use the scale: they were asked to rate 6 or 7 sentences they found perfectly acceptable, to rate 1 or 2 sentences they found completely unacceptable and incomprehensible, and to assign 3-5 to sentences that were somewhere in between. They were also asked to judge the sentences on the basis of their native-speaker intuitions, rather than any prescriptive rules, and to go with their first instinct instead of spending time thinking about their answers. Moreover, the first five experimental items were used as a pre-test phase and were then excluded from the statistical analysis. The practice items were not marked as such: participants did not know these were practice items. All participants saw all stimuli.

For statistical analysis, raw ratings of each individual subject, including both target and filler items, were first transformed into *z*-scores. In this way, potential scale biases between participants, such as the use of extremes of the scale only, are avoided. In other words, the raw scores are transformed to denote how many standard deviations from a participant's mean score a judgement is. Linear mixed-effects models were then run on the transformed data with the R statistical package *lme4* (Bates et al. 2015). The fixed effect predictors included STRUCTURE and TENSE, as well as their interaction, and the random effects included subjects and items. All

predictors were sum coded before the data analysis, with [+ island] and [+ tense] coded as 1, and [- island] and [- tense] coded as -1.

4.4.2 Results

4.4.2.1 Preliminary results

Preliminary results of Experiment B revealed some peculiarities for the adjunct introduced by *after*, which will be illustrated here. As for the STRUCTURE factor, the statistical analysis showed that sentences with islands were rated lower than the parasitic gap condition. The odd result, though, is that the effect was significant for both adjuncts introduced by *before* (linear mixed effect estimate: $\beta = -0.52$, $s.e. = 0.08$, $t = -5.98$, $p < .0003$) and by *without*, although in the latter case the effect is much weaker ($\beta = -0.44$, $s.e. = 0.09$, $t = -4.619$, $p < 0.02$). On the other hand, for the adjunct introduced by *after*, the statistical analysis did not reveal an effect of STRUCTURE: islands were rated lower than parasitic gaps (linear mixed effect estimate: $\beta = -0.22$, $s.e. = 0.13$, $t = -1.61$), but this did not reach statistical significance ($p > 0.07$). Given the results of Experiment A, this is unexpected and it may be due to two different reasons: first of all, the comparison here is not the typical one where extraction takes place out of an island or out of a declarative, which is always grammatical. In this case, the structures compared to islands are parasitic gaps, which are usually considered grammatical, but they are also slightly degraded with respect to declaratives. Secondly, it may also be connected to the kind of sentences proposed in the test. For this reason, I decided to run linear mixed-effects model separately on the two set of sentences proposed for the adjuncts introduced by *after* in order to shed some light on the issue.

As for the TENSE factor, statistical analysis revealed that there is an effect such that sentence with finite tense in the adjunct were rated lower than sentences with non-finite tense. In the case of *before* the effect was significant ($\beta = -0.18$, $s.e. = 0.09$, $t = -2.03$, $p < 0.03$), as well as in the case of *without* ($\beta = -0.35$, $s.e. = 0.09$, $t = -3.62$, $p < 0.01$). Once again, though, the case of *after* is different (and quite unexpected): even though sentences with finite tense were rated lower than sentences with non-finite tense, the effect was not significant ($\beta = -0.22$, $s.e. = 0.13$, $t = -1.61$, $p > 0.37$).

Finally, the interaction between structure and tense revealed once again the peculiarity of the items presented in the *after* adjunct, and therefore it stressed the necessity of running a

separate analysis for the two sets of sentences. Basically, the interaction between the two factors STRUCTURE and TENSE was significant in the case of *before* ($\beta = 0.16$, $s.e. = 0.06$, $t = 2.79$, $p < 0.003$) and *without* ($\beta = 0.21$, $s.e. = 0.03$, $t = 7.664$, $p < 0.003$), but not in the case of *after* ($\beta = 0.09$, $s.e. = 0.14$, $t = 0.61$, $p > 0.39$), where results tell us that the score is not affected by the effect of structure over tense. Individual analyses are hence required to understand whether this is due to one set of sentences only.

4.4.2.2 A study of 'after'

As expected, the picture is clearer with a separate statistical analysis for the two kinds of sentences investigated in extraction out of the adjunct introduced by *after*. Two separate linear mixed-effect models were run for the two sets, and the results were quite different, proving that the preliminary results have to be integrated. The sets presented can be found in (14) and (15).

- (14) a. Quali dolci ha mangiato Luca dopo aver rubato?
 which sweet has eaten Luca after having stolen
 'Which sweets did Luca eat after having stolen?'
 b. Quali dolci ha mangiato Luca dopo che ha rubato?
 which sweet has eaten Luca after that he has stolen
 'Which sweets did Luca eat after that he stole?'
 c. Quali dolci Luca è scappato dopo aver rubato?
 which sweet Luca is run away after having stolen
 'Which sweets did Luca run away after having stolen?'
 d. Quali dolci Luca è scappato dopo che ha rubato?
 Which sweet Luca is run away after that he has stolen
 'Which sweets did Luca run away after that he stole?'
- (15) a. Quale esame Elisa ha registrato dopo aver superato?
 which exam Elisa has registered after having passed
 b. Quale esame Elisa ha registrato dopo che aveva superato?
 which exam Elisa has registered after that she had passed
 c. Quale esame Elisa ha festeggiato dopo aver superato?
 which exam Elisa has celebrated after having passed

- d. Quale esame Elisa ha festeggiato dopo che aveva superato?
which exam Elisa has celebrated after that she had passed

As far as sentences in (14) are considered, results were as expected: there are significant effects for both STRUCTURE and TENSE. For the former, the island condition was rated significantly lower than the parasitic gap one (linear mixed effect estimate: $\beta = -0.5$, $s.e. = 0.09$, $t = -5.43$, $p < 0.002$). As for the TENSE factor, when a finite verb was used in the adjuncts, sentences were rated significantly lower (linear mixed effect estimate: $\beta = -0.13$, $s.e. = 0.09$, $t = -1.5$, $p < 0.03$). The interaction of STRUCTURE and TENSE was significant ($\beta = 0.09$, $s.e. = 0.15$, $t = 0.58$, $p < 0.009$).

Sentences in (15) are different: none of the factors reached significance, and the island condition was rated higher than the parasitic gap one (linear mixed effect estimate: $\beta = 0.06$, $s.e. = 0.08$, $t = 0.788$, $p > 0.25$). Sentences with finite tense were rated lower than those with non-finite tense, but not significantly ($\beta = -0.06$, $s.e. = 0.08$, $t = 0.715$, $p > 0.21$). The interaction of STRUCTURE and TENSE was significant ($\beta = 0.08$, $s.e. = 0.03$, $t = 2.96$, $p < 0.008$).

The first set of sentences does not behave in an unexpected way: both factors of STRUCTURE and TENSE were significant. However, the second set had some surprising results, in that not only did the factor STRUCTURE not reach significance, but the island condition was actually rated higher than the parasitic gap one. We can see that the parasitic gap and the island condition are equal to each other in sentences with non-finite tense in the adjunct. This is due to a combined effect: (i) island sentences are rated higher, and (ii) parasitic gaps are rated very poorly. Therefore the results are quite unusual, and possibly connected to the kind of sentence involved here: either we are dealing with a floor effect, or the parasitic gap goes under an island interpretation. The difference could also be attributed to the lexical choices, rather than some other factors connected to the syntax. Therefore, it seems that the speakers did not analyse the parasitic gap sentences as such: there is no higher gap able to c-command and license the lower one, so that the sentence actually becomes an island. At the same time, it may be that in the island condition the reverse situation holds: probably because of the lexical choices, the islands is actually analysed as a parasitic gap, and in (15c) *which exam* is interpreted as the object of both the verb in the adjunct, and the one of the matrix clause. Most probably, this is related to

the type of verb used in the matrix sentence, given that *to party* in Italian can be transitive. In all the other island cases, the matrix verb was never transitive.¹²

4.4.2.3 Final results

Given the different results between the two set of sentences introduced by *after*, I excluded from the final analysis those of (15), which were completely unexpected and are presumably connected to a different interpretation. In this way, the overall analysis would be more reliable. Overall results of z-scores are summarized in Table 2 and 3. These show an expected pattern since both STRUCTURE and TENSE are significant factors, as well as their interaction.

| | <i>after</i> | <i>before</i> | <i>without</i> |
|-----------------------|--------------|---------------|----------------|
| parasitic gap, -tense | 0.24 | 0.44 | 0.89 |
| parasitic gap, +tense | -0.21 | -0.24 | -0.23 |
| island, -tense | -0.94 | -0.93 | -0.41 |
| island, + tense | -1.03 | -0.97 | -0.71 |

Table 4.2. Means of z-score ratings for each condition and item for Experiment B

For the STRUCTURE factor, statistical analysis revealed that sentences with islands are significantly less acceptable than parasitic gaps, in all adjunct types (for the case of *after*: $\beta = -0.5$, $s.e. = 0.09$, $p < 0.002$, *before*: $\beta = -0.52$, $s.e. = 0.08$, $p < .0003$, and *without*: $\beta = -0.44$, $s.e. = 0.03$, $p < 0.01$).

As for the TENSE factor, sentences with a finite verb in the adjunct were rated significantly lower than sentences with non-finite tense (*after*: $\beta = -0.13$, $s.e. = 0.09$, $p < 0.03$, *before*: $\beta = -0.18$, $s.e. = 0.09$, $p < 0.03$, and *without*: $\beta = -0.35$, $s.e. = 0.03$, $p < 0.01$).

Finally, also the interaction between structure and tense was significant in all the cases considered (*after*: $\beta = 0.09$, $s.e. = 0.15$, $p < 0.009$, *before*: $\beta = 0.16$, $s.e. = 0.06$, $p < 0.003$, and *without*: $\beta = 0.21$, $s.e. = 0.03$, $p < 0.003$).

¹² These results seems to fit nicely with Truswell's (2007, 2011) semantic approach of cases of acceptable and unacceptable extraction from adjuncts, which will be discussed in the next chapter.

| | <i>after</i> | <i>before</i> | <i>without</i> |
|--------------------------|--------------|---------------|----------------|
| Main effect of STRUCTURE | *** | *** | * |
| Main effect of TENSE | * | ** | ** |
| STRUCTURE x TENSE | ** | ** | *** |

Table 4.3. Mixed-effect model for each adjunct type for Experiment B.

4.4.3 Discussion

These results show that there are several ways which allow us to talk about variation among the adjuncts proposed: (i) the effect of STRUCTURE, (ii) the case of parasitic gaps. I will briefly discuss each of them in the following sections.

4.4.3.1 Structure and differences among adjuncts

As in Experiment A, the factor of STRUCTURE was significant, i.e. the sentences with extraction from the island condition are rated significantly lower than their counterparts with the parasitic gap. This is true for all of the adjuncts taken into account – except for one set introduced by *after* which has been excluded from the final statistical analysis and which will be discussed in a separate section. Notice that even though the factor of STRUCTURE was significant in all the adjuncts considered, nonetheless it can be weaker or stronger depending on the adjunct. In fact in the case of *without* it is much weaker ($p < 0.01$) with respect to the both *after* ($p < 0.002$) and *before* ($p < 0.0003$), where extraction is completely ruled out.

Notice that the same conditions discussed for Experiment A hold here: the only thing that differentiate the sentences is the type of conjunct introducing the adverbial clause. In fact, the sentences investigated were equal in several aspects: all of the them represent a case of extraction from adjuncts, the extracted element is a DP argument made by a complex *wh*-constituent rather than a simple *wh*-word, in every example there is an intervening overt subject, and the verbs used for both the matrix and the embedded sentence are similar. All of the factors discussed in §3.2 have been manipulated in order to make the island effect as weak as possible. Therefore, the immediate question we may ask is why the STRUCTURE have different effects according to the adjuncts. A natural answer is once again to assume that adjuncts are not to be

considered a uniform group, but a heterogeneous one. Apparently, for some reasons which will be discussed in chapter 5, the island effect can be weaker (or not-existent) or stronger, depending on the adjunct used. In the cases investigated here, we can state that *without* is much weaker than *after* and *before*. The fact that (some) adjunct islands have weaker effect was already detected by some authors, from a theoretical point of view (see Cinque 1990; Haegeman 2012) as well as from an experimental one (Heestand et al. 2012; Polinsky et al. 2013). Figure 4.6 shows such difference: means of *z*-score ratings are compared for the three adjuncts in the island condition with non-finite tensed adjunct.

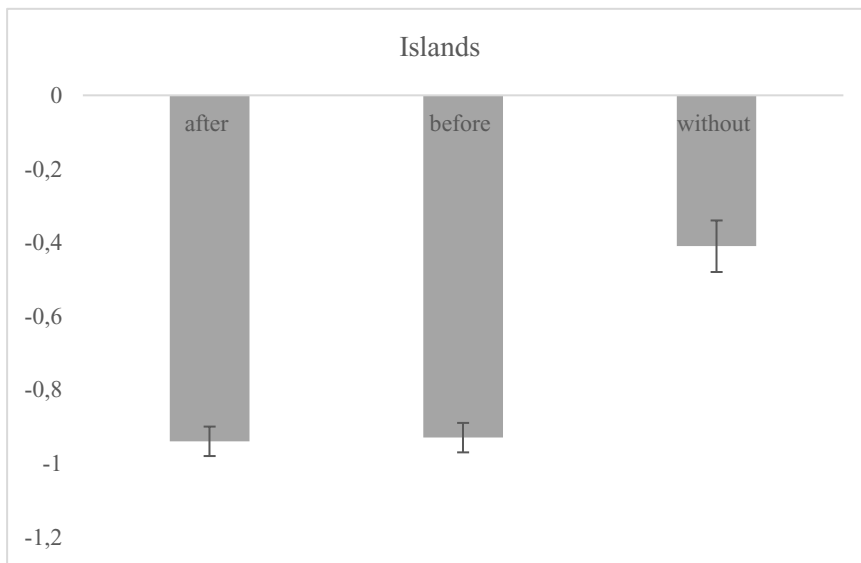


Figure 4.6. Z-score acceptability judgments comparing extraction from each adjunct type for the island with non-finite tense condition.

As in Experiment A, the same distinction crucially holds for parasitic gaps, i.e. structures which should be equally acceptable and licensed. However, acceptability judgments show that there is a difference among the adjuncts considered: extraction from *without* is completely grammatical, but the same cannot be said of the other adjuncts, which are much more degraded.

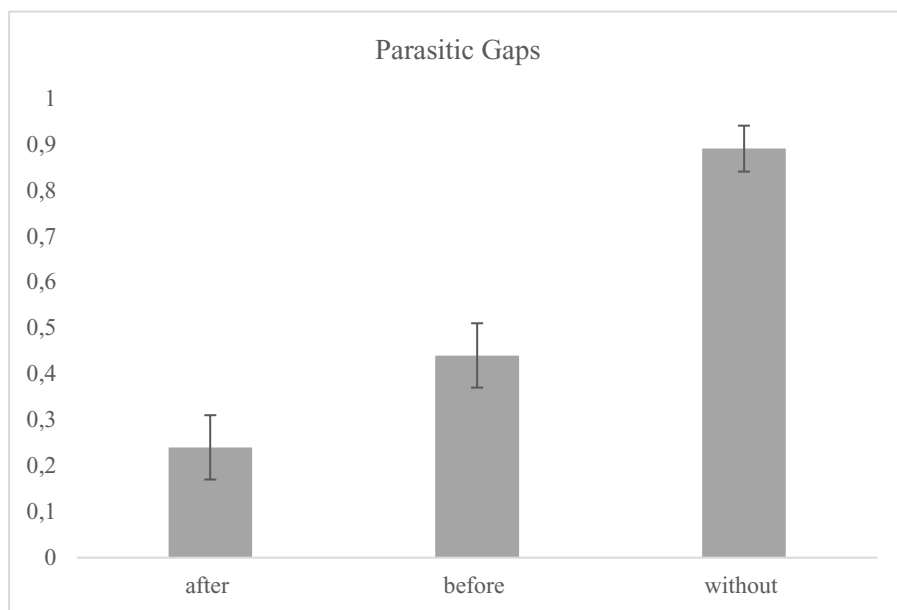


Figure 4.7. Z-score acceptability judgments comparing extraction from each adjunct type for the parasitic gap with non-finite tense condition.

The fact that such a distinction affects not only degraded or unacceptable cases of extraction, but also cases that should be licensed is an important piece of evidence proving that there is a difference among adjuncts we should keep into account.

4.4.3.2 *The role of tense*

As introduced in §3.2.1, the presence of finite tense is one of the factors able to affect islands: whenever the embedded clause contains a tensed verb, the island effect is strengthened. The result of this test confirms such an effect: whenever extraction took place from an adjunct with finite tense, the sentence was much more degraded than its non-finite counterpart and it was rated significantly lower. This is true for all the adjuncts considered. Its role was crucial not only in island constructions, but also in parasitic gap sentences – probably even more in this case. I will further discuss why tense has such an influential role in chapter 5. Notice that the effect of tense has been rarely investigated with formally-collected judgments. Michel & Goodall (2012) tested its effect in English, and they obtained similar results: in the case of adjuncts, there was a highly significant finiteness effect in *wh*-questions (island constructions), but not in yes-no questions (grammatical control sentences). This is important in that it shows

that the presence of finite tense does not affect all kinds of sentences, i.e. that is not intrinsically difficult on its own, contrary to what is predicted by processing accounts (Hofmeister 2007).¹³

4.4.4 Summing up

Results from Experiment B can be summarized as down below:

- (i) adjuncts are not to be considered a uniform group, but a heterogeneous one. The island effect can be weaker or stronger, depending on the adjunct used. In the cases investigated here, we can state that *without* is much weaker than *after* and *before*, confirming results from Experiment A;
- (ii) the same difference among adjuncts is valid for parasitic gaps, namely for grammatical structures;
- (iii) the effect of finite tense really is crucial and able to affect significantly the acceptability of a sentence containing extraction;

Notice that the sentences investigated here were equal in several aspects: the difference between them should be attributed to the adjunct itself, and in particular to its syntax. Thus, it seems that we cannot account for adjuncts by means of a uniform condition, but we have to take a step further and divide this class according to some criteria. In order to check whether this is true, more adjuncts types should be investigated, which is the aim of the last experiment.

4.5 Experiment C

The main purpose of Experiment C is to extend the sample of prepositional adjuncts in order to check whether the variation detected in the previous tests can be related also to new adverbial clauses, and therefore investigating further whether it is possible to consider adjuncts as a heterogeneous group rather than a uniform one. This issue is particularly interesting in order to give a more exhaustive explanation of prepositional adjuncts. Moreover, if extraction really affects adjuncts differently, such a test would be particularly revealing for a more fine-grained theory of adjuncts. Therefore, I decided to check once again the effect of structure comparing islands and parasitic gaps, but increasing the number of adjuncts involved. Moreover, given

¹³ This is different from the results obtained for the D-linking effect discussed in §3.2.3, where it was shown that D-linking had an ameliorating effect not only for islands, but also for grammatical questions (Goodall 2015), hence favouring a processing explanation.

some suggestions,¹⁴ I decided to consider again the effect of the presence of a resumptive pronoun with respect to a gap, in order to establish whether the presence of resumptive pronouns helps the island condition, or if it rather makes the sentence with the island worse, as was the case of *without* in Experiment A. Therefore, this test allows me to (i) compare (supposedly) acceptable cases of parasitic gaps and unacceptable sentences containing island violations, (ii) compare different types of prepositional adjuncts and determine whether they form a uniform category in extraction phenomena, (iii) check whether resumptive pronouns help the island effect.

4.5.1 *Materials and procedure*

A 2x2 factorial design is employed, with 4 conditions resulting from crossing two factors: (i) STRUCTURE and (ii) RESUMPTION. For the STRUCTURE factor, the unacceptable cases of extraction from an island are compared to the (supposedly) acceptable cases of parasitic gaps. For the RESUMPTION factor, sentences with a gap are compared to sentences with a resumptive pronoun. See (18) for an example of an entire paradigm of an item.

(18) a. Parasitic gap, gap:

Quale libro Maria ha comprato _ [per leggere _] ?

which book Maria has bought in order to read

‘Which book did Mary buy in order to read?’

b. Parasitic gap, RP:

Quale libro Maria ha comprato _ [per leggerlo]?

which book Maria has bought in order to read it

‘Which book did Mary buy in order to read it?’

c. Island, gap:

Quale libro Maria è andata a casa [per leggere _]?

which book Maria is gone to home in order to read

‘Which book did Mary go home in order to read?’

¹⁴ Thanks to Anna Cardinaletti and Giuliana Giusti (p.c.).

d. Island, RP:

Quale libro Maria è andata a casa [**per leggerlo**]?

which book Maria is gone to home in order to read it

‘Which book did Mary go home in order to read it?’

The 4 conditions will be tested in six different types of prepositional adjuncts, three are the same of Experiment A and B – *after*, *before*, and *without* – and three are new – *in order to*, *because* and *until*. The total is 24 items. Sentences in (19) presents some of the items in the island condition with the gap.

- (19) a. Quale libro Sergio è tornato a casa [**per** leggere _]?
‘Which book did Sergio come back home (in order) to read?’
- b. Quale ragazzo Elisa è partita [**senza** salutare _]?
‘Which boy did Elisa leave without greeting?’
- c. Quale esame Paolo si è preoccupato [**prima** di superare _]?
‘Which exam did Paolo worried before passing?’
- d. Quale ragazza Piero è corso via [**dopo** aver baciato _]?
‘Which girl did Piero run away after having kissed?’
- e. Quale film hai insistito [**finché** non ho guardato _]?
which movie have (you) insisted until not (I) have watched
‘Which movie did you insist until I watched?’
- f. Quale libro ti sei arrabbiato [**perché** ho rovinato _]?
‘Which book did you get angry because I ruined?’

I created 4 lexicalizations for each conditions, yielding 16 stimuli for each adjunct and a total of 96 stimuli. These were distributed among 4 lists using a Latin Square design. This meant that each list consisted of 4 tokens per sentence type, for a total of 96 items and 4 different surveys. Items were presented in a pseudo-randomized order, so that the same condition never appeared consecutively. The items were divided and distributed into two different blocks, Block A and Block B. These were presented in different order among participants: some had to judge the sentences in Block A first, while others sentences of Block B. This allowed to obtain a more randomized order. The experiment was balanced with respect to the types of adjuncts tested. Each of them were tested an equal amount of times.

Fillers are included in a ratio of 1:1, for a total of 24 fillers, which are of comparable length and varying acceptability: most of them are either completely grammatical or ungrammatical, but the acceptability of some sentences falls in between. In doing so, participants are encouraged to use a large portion of the scale, instead of focusing only on a small portion at its extremities. Notice that I included both declaratives and questions as fillers, so that the target items were not the only questions in the experiment. Every subject was therefore tested on a total of 48 items.

Fifty-nine native speakers of Italian participated in the experiment. They were undergraduate students at the University of Venice (age-range 18-25) with no training in formal syntax. Students were all enrolled in an introductory linguistics course which did not introduce them to island effects. Furthermore, the experiment was administered prior to the discussion of syntactic structures. Before the experiment, participants were asked to provide some information regarding their age, gender, as well as their native language and dialect.

Items were presented in a written form. Participants were asked to judge the acceptability of the sentences on a 7-point Likert scale, where 7 indicates perfect acceptability, and 1 total unacceptability. Before beginning the experiment, participants were provided with instructions on how to use the scale: they were asked to rate 6 or 7 sentences they found perfectly acceptable, to give 1 or 2 to sentences they found completely unacceptable and incomprehensible, and to assign 3-5 to sentences that were somewhere in between. They were also asked to judge the sentences on the basis of their native-speaker intuition, rather than any prescriptive rules, and to go with their first instinct instead of spending time thinking about their answers. Moreover, the first five experimental items were used as a pre-test phase and were then excluded from the statistical analysis. The practice items were not marked as such: participants did not know these were practice items.

For statistical analysis, raw ratings of each individual subject, including both target and filler items, were first transformed into *z*-scores in order to avoid potential scale biases between participants. Linear mixed-effects models were then run on the transformed data with the R statistical package *lme4* (Bates et al. 2015). The fixed effect predictors included *STRUCTURE* and *RESUMPTION*, as well as their interaction, and the random effects included subjects and items. All predictors were sum coded before the data analysis, with [+ island] and [- resumption] coded as 1, and [- island] and [+ resumption] coded as -1.

4.5.2 Results

I will now present the results from this experiment, first taking into account the general behavior of all the adjuncts considered and later focusing on each adjunct presented.

| | <i>after</i> | <i>before</i> | <i>without</i> | <i>in order to</i> | <i>until</i> | <i>because</i> |
|--------------------|--------------|---------------|----------------|--------------------|--------------|----------------|
| parasitic gap, -RP | 0.08 | 0.35 | 0.52 | -0.51 | -0.37 | 0.2 |
| parasitic gap, +RP | 0.22 | 0.31 | 0.34 | 0.10 | 0.29 | 0.15 |
| island, -RP | -0.94 | -0.8 | -0.45 | -0.64 | -1.07 | -0.6 |
| island, +RP | -0.95 | -0.74 | -0.7 | -0.45 | -0.43 | -0.8 |

Table 4.4. Means of z-score ratings for each condition and adjunct of Experiment C.

First, I will take into account the STRUCTURE factor. Statistical analysis revealed that sentences with islands are always less acceptable than parasitic gaps, for all adjunct types. What is interesting, though, is the fact that this result was significant only in some of the cases, but not all. Islands were rated significantly worse than parasitic gaps in the case of *after* (linear model estimates: $\beta = -1.19$, $s.e. = 0.10$, $p < 0.0001$), *before* ($\beta = -1.04$, $s.e. = 0.12$, $p < .0003$), *without* ($\beta = -1.02$, $s.e. = 0.11$, $p < 0.0001$), and *because* ($\beta = -0.91$, $s.e. = 0.10$, $p < 0.0001$), but the difference was not significant for both *in order to* ($\beta = -0.59$, $s.e. = 0.11$, $p > 0.018$) and *until* ($\beta = -0.74$, $s.e. = 0.11$, $p > 0.072$).

As for the RESUMPTION factor, results are rather mixed. In some cases, the presence of a resumptive pronoun makes the sentence significantly better in the case of *in order to* ($\beta = 0.64$, $s.e. = 0.11$, $p < 0.018$) and *until* ($\beta = -0.7$, $s.e. = 0.11$, $p < 0.0001$), and it makes the sentence better but without reaching significance in the case of *after* ($\beta = 0.14$, $s.e. = 0.10$, $p > 0.46$). On the other hand, the presence of a resumptive pronoun makes the sentence significantly worse in the case of *without* ($\beta = -0.17$, $s.e. = 0.11$, $p < 0.01$), and it makes the sentence worse without reaching significance in the case of *before* ($\beta = 0.04$, $s.e. = 0.12$, $p > 0.8$) and *because* ($\beta = -0.04$, $s.e. = 0.10$, $p > 0.11$).

Finally, the interaction between structure and resumption was not significant in almost all the adjuncts considered (*after*: $\beta = 0.17$, $s.e. = 0.15$, $p > 0.2$, *before*: $\beta = -0.05$, $s.e. = 0.17$, $p > 0.7$, *without*: $\beta = 0.05$, $s.e. = 0.16$, $p > 0.74$, *until* $\beta = 0.11$, $s.e. = 0.15$, $p < 0.48$ and *because*

$\beta = 0.13$, $s.e. = 0.14$, $p > 0.3$), with the exception of *in order to*, where the interaction resulted significant ($\beta = 0.47$, $s.e. = 0.16$, $p < 0.008$).

| | | <i>after</i> | <i>before</i> | <i>without</i> | <i>in order to</i> | <i>until</i> | <i>because</i> |
|----------------|------------------------|--------------|---------------|----------------|--------------------|--------------|----------------|
| Main effect of | STRUCTURE | *** | *** | *** | - | - | *** |
| Main effect of | RESUMPTION | - | - | * | * | *** | - |
| | STRUCTURE x RESUMPTION | - | - | - | ** | - | - |

Table 4.5. Mixed-effect model results for Experiment C.

We can already observe from Table 4.5 and 4.6 that adjuncts do not behave in a uniform way, confirming the results from the previous tests. I will discuss the main effects and the consequence of these results, taking into account both what was expected and what was unexpected.

4.5.3 Discussion

4.5.3.1 Different adjuncts leads to different effects

First I am going to analyse the results from the adjuncts which were considered for Experiment A and B as well, in order to check whether the same results hold in this test or whether there is something different. First of all, this test confirmed the results from both Experiment A and B in the role of STRUCTURE for adjuncts introduced by *after* and *before*: sentences with parasitic gaps are always better than their island counterpart, and the difference between the two is highly significant. Therefore, participants consistently judge extraction from these adjuncts to be unacceptable. Differently from the previous experiments, the effect of STRUCTURE resulted highly significant also in the case of *without*. Remember that in Experiment A it was not significant in the gap model, whereas in Experiment B it was significant but the effect was much weaker than in the other adjuncts. In this test, there is no difference as long as structure is considered: extraction from islands always lead to unacceptable sentences.

What happens with the new adjuncts, then? The case of *because* aligns well with these results: the parasitic gap is rated significantly better than the island, hence the island effect is present. However, both *in order to* and *until* behave in a different way, as can be seen summarized in Table 4.5. In neither of them is the effect of STRUCTURE significant. This means that although the sentences containing an island were rated lower than their counterparts containing a parasitic gap, the difference is not significant. We might conclude that these adjuncts are not sensitive to the island effect (in this test, at least). However, this is not entirely true. In fact, we should rather consider the sentences islands were compared to: parasitic gaps. The first question that should be asked is whether these constructions are really acceptable and comparable to extraction from declaratives, as often argued in the literature both on the theoretical side and the experimental one. Both Experiment A and Experiment B showed that there is a difference among adjuncts – which can be seen not only in the case of islands, but also in that of parasitic gaps. This means that different adjuncts licence (or do not license) parasitic gaps differently. It may well be the case that for adverbial clauses introduced by *in order to* and *until* it is not the case that the effect of island is not present, but rather that parasitic gaps are judged as unacceptable, resulting in the lack of a significant difference among islands and parasitic gaps. In fact, if we look at the mean values of parasitic gaps (in the gap model), we can notice that they are not good sentences: *in order to* has a mean value of 3 (raw value) and -0.51 (z-value), whereas the mean value of *until* is 3.27 (raw value) and -0.37 (z-value). If we compare these values with the most acceptable case, i.e. parasitic gap formed with *without*, we can notice a big difference: the mean the raw value is 5.17 and the z-score equals to 0.52. Therefore, there is something going on with parasitic gaps. I will leave the question aside for the time being.

At first sight it looks like that the results of this test lead to an equal condition, suggesting that no significant difference among adjuncts can be discussed. However, look at Figure 4.8 representing the comparison of the mean z-scores in the island condition with no resumption, for all adjunct types used.

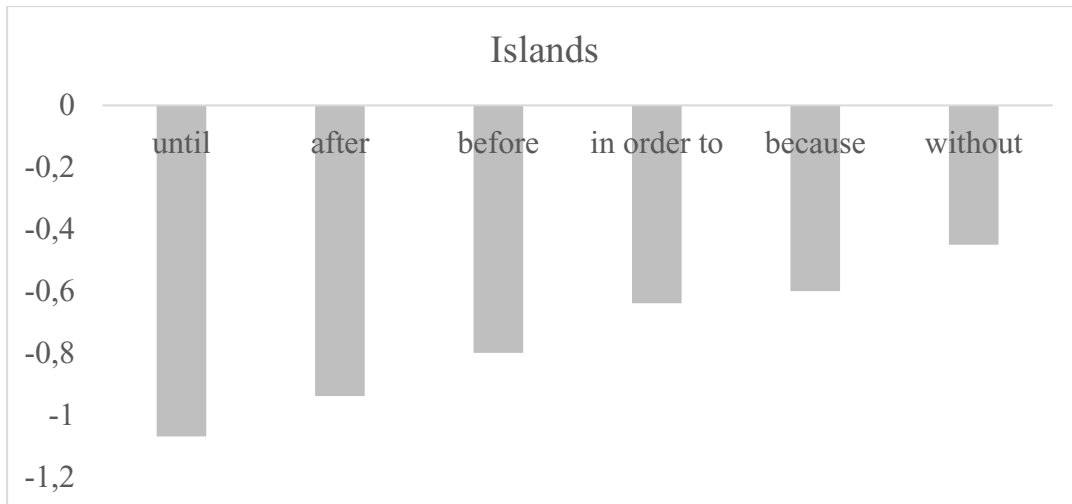


Figure 4.8. Z-score acceptability judgments comparing extraction from each adjunct type for the island and the gap condition.

As can be seen, we still have to take into account a (sometimes major) difference between the adjuncts considered: even though they also showed a significant effect of STRUCTURE (with the exceptions of *in order to* and *until*, which are due to the low acceptability of the parasitic gap condition), Figure 4.8 clearly shows that they have different behaviours.

Now look at Figure 4.9, where the comparison is between the mean z-values of the parasitic gap condition, with no resumption.

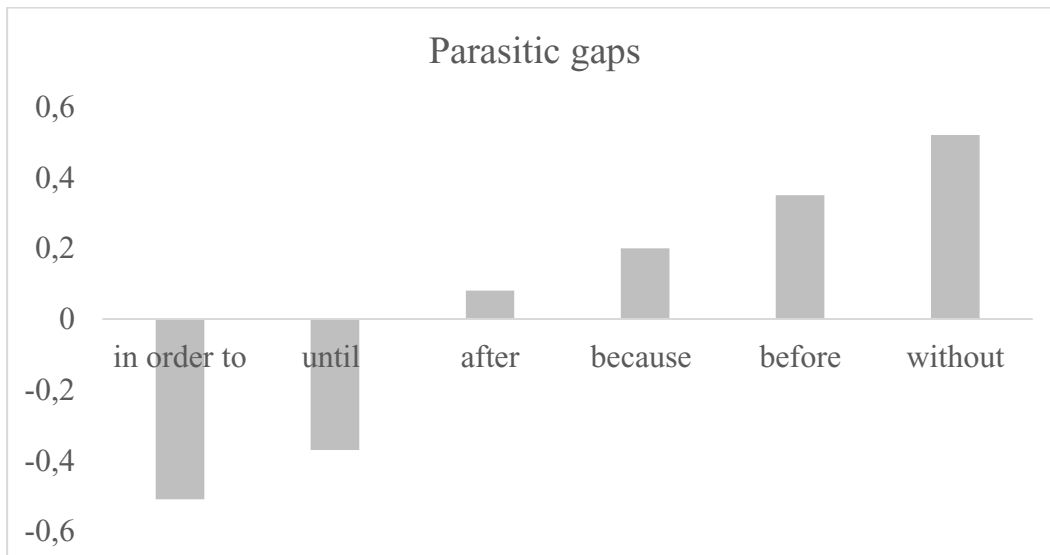


Figure 4.9. Z-score acceptability judgments comparing extraction from each adjunct type for the parasitic gap without resumptive pronoun condition.

The difference among adjuncts is even clearer in the case of parasitic gaps. It is also more unexpected, given what is usually said about these structures, i.e. that they are grammatical. From this test, as well as from the previous ones, we may draw different conclusions with respect to parasitic gaps: (i) they are not equally acceptable; (ii) they are not always grammatical sentences; (iii) they might show a deeper difference among adjuncts, which is not always clear in the case of islands. Probably (iii) is related to the fact that these are usually much more degraded sentences and it is more difficult to perceive a difference among them, as in a sort of floor effect. Moreover, notice that we may have a problem with the use of parasitic gaps as the comparison items for island structures. In fact, even though usually acceptable, these turned out not to be always good, and more degraded than declaratives. This means that different results could be obtained with a different structure compared to the island.

4.5.3.2 *Effect of Resumption*

When it comes to the effect of resumption, little has changed in Experiment C with respect to the first pilot test (Experiment A). I expected the presence of resumptive pronouns in the island to be a non-conditioning factor, i.e. that sentences containing a resumptive pronoun would not turn out to be better than those where a gap is used. The results obtained in this case are not at all as straightforward as expected. The presence of resumptive pronouns goes in different directions, depending on the adjunct considered. It is not as in the case of STRUCTURE, where there is only one direction to follow, which tells us that islands are always rated worse than parasitic gaps. Here things are not so simple: some adjuncts prefer the gapped option, others the filled one (following the copying rules of Ross 1967). In some cases the effect is significant, but in others it is not.

It turns out that in two of the adjuncts considered, RPs make the sentence significantly better than its counterpart with a gap. This is the case of *in order to* and *until*. Notice that these are the same adjuncts which did not show a significant STRUCTURE effect. Since this was most probably due to the low acceptability of parasitic gaps, it is difficult to establish whether there is a correlation here. In both the case of *after* and *because*, sentences with resumptives were slightly better than those with a gap, but the effect is not significant. On the other hand, in *before* the opposite situation holds: sentences with a gap are better than their counterpart with a resumptive pronoun, but without reaching significance. Finally, the case of *without* stands on its own: sentences with a gap were significantly better than those with a resumptive pronoun,

confirming results from Experiment A. It seems that *without* cannot tolerate the use of resumptives.

Given these different behaviours, it is difficult to account for the lack or presence of the rescuing effect of resumption in the island. There is an interesting pattern to be observed: significance is obtained in those cases where the acceptability (or rather, unacceptability) is at the lowest, both in the case of islands and of parasitic gaps. On the other hand, the significance at the reverse is obtained in the case of *without*, which is also the adjunct that mostly allows for extraction in both cases, as showed in all experiments. In this case, the presence of the resumptive makes the sentence significantly worse than the gapped counterpart. There are also in between cases: namely adjuncts where the effect of resumption is not significant – on one side or the other – which are the adjuncts in the middle of our continuum of the acceptability scale. Resumption in the case of *after* leads to an improvement, but in both *before* and *because* the sentence gets worse.

The lack of an ameliorating effect connected to the resumptive might depend on several factors of the test. First, it may be due to an effect of structure. Notice that usually resumptives are mostly found in declaratives, whereas I used only *wh*-questions, and this might explain the lack of an effect. As shown in Heestand et al. (2012) and Polinsky et al. (2013), RPs are much more acceptable in declaratives than they are in questions. Moreover, the effect of resumptives should be considered in an online task, rather than an offline one (see Heestand et al. 2012). In fact, since resumptives are mostly a spoken phenomenon, when it comes to Italian and English, i.e. languages with intrusive RPs, an auditory presentation may lead to a different result.¹⁵ It is also possible that the lack of an effect of resumptives in these tests is connected to the type of task asked to participants, and the environment in which the tests took place. Instructions of the test do seem to have a crucial role, as shown by Beltrama & Xiang (2016). In this test two different instructions were assigned to participants, which were asked to judge either the *acceptability* or the *comprehensibility* of sentences containing different types of islands and resumptive pronouns, as well as counterpart sentences containing gaps. Interestingly, the effect of RPs was significant for both Italian and English in the comprehensibility task, but not in the acceptability one – which is usually the one employed for these tests. Moreover, this effect was present only when a context sentence preceded the item.

¹⁵ Although it was shown in Polinsky et al. (2013) that the modality of presentation does not actually affect the acceptability of resumption.

Crucially, there is a discrepancy between the production of intrusive resumptive pronouns and their acceptability. Such a discrepancy makes it difficult to assess these kinds of structures in an experimental context; even though RPs are systematically realized in spontaneous speech – as noticed in studies of corpora (see Prince, 1990) – nonetheless they tend not to be accepted. This tendency is particularly clear in the example of Ferreira & Swets (2005), where the production of RPs was elicited in island contexts and the same speakers who produced these sentences did not accept them when asked to rate their acceptability.

Given that resumptive pronouns in Italian, as well as in English, are considered an informal and mostly spoken phenomenon, it is hard to state that the lack of an effect cannot be linked to their status. We should remember that these kinds of elements are generally perceived as part of a ‘bad’ version of a language, hence as substandard varieties, as first noted by Ross (1967: 434), who wrote that “resumptives are common in everyone’s speech, but are regarded as substandard by normative grammarians”. In fact, intrusive RPs tend to be connected to informal and spoken registers, which means that speakers might be inclined to classify them as not entirely acceptable, even though they use them. Participants may be influenced by normative considerations, and therefore reject these forms, following a prescriptive behavior. If this is really the case, it is obvious that to obtain such an effect the environment in which these sentences are tested is crucial: we do not expect to see an amelioration among those who were prescriptively instructed that these elements should not be used.¹⁶ It is much easier to find them in spontaneous speech, which is usually less controlled than an acceptability task where participants are required to actually think about that construction. It is also important to note that people who participate in this kind of studies are usually enrolled at university. This is important because it may be more difficult to find an effect of resumption in higher levels of education, precisely because people have been instructed not to like them. Notice also that there is a difference in this respect among trained linguists and linguistically-naïve people, namely the ones usually employed for these tests. In fact, linguists’ introspective judgments tell us that sentences with island violations at least sounds better than their gapped counterpart.

It is still not clear how to consider these data and how to account for them in an experimental study, and further investigations are required to assess this point, extending the sample of participants and checking some of the factors mentioned here. Given the consistent

¹⁶ In fact, after participating in the pilot test, some speakers gave me feedback about the sentences they had to judge. When talking about the resumptives, many told me that it was a repetition of something that was already present in the sentence, and hence wrong.

use of RPs in spontaneous production, it is possible that intrusive resumption does not fix a problem in the derivation, and only affects the comprehension of a sentence. In fact, in many cases the explanation of the ameliorating effect of RPs is based on reductionist approaches, rather than on syntax. The use of resumptive pronouns tends to be connected to unfavorable processing conditions and to poor planning in production and the help they provide cognitively. This may be due to the fact that RPs help (i) to construct a coherent parse, and (ii) to retrieve part of a non-local dependency, given that finding the tail of a filler-gap dependency may be particularly difficult in island contexts. Basically, we may analyze RPs as anaphoric elements that help us in the identification of the tail of a dependency and offer explicit morphological cues. Gap could be considered as bound variables, and RPs as anaphoric elements linked to their antecedents. Following this reasoning, gaps are more sensitive to syntactic islands, whereas RPs find their contextually salient discourse antecedent despite the intervening island boundary. Such an explanation would explain the impact of the presence of the context sentence in Beltrama & Xiang (2016).

Quite interestingly, notice that even in languages with grammatical resumption, the presence of an RP does not always ameliorate the island, contrary to what is commonly stated (see McCloskey 2006 for Irish; Shlonsky 1992 for Hebrew and Arabic; Aoun, Choueiri & Hornstein 2001 for Arabic). As it turns out, once these structures are formally investigated, results are quite different, and it seems that resumption does not always lead to an increased grammaticality. Results from several tests show that there is an ameliorating factor of resumption, but this is smaller than expected and it is also subject to variation depending on the construction taken into account. In general, RPs are more used inside islands, but cannot freely vary in other constructions (see McCloskey 2017). This is the case shown in Tucker, Idrissi, Sprouse & Almeida (2019) for Modern Standard Arabic.¹⁷ They show that resumption can indeed ameliorate an island, but that it depends on the island; their rescuing value is not extended to all of the islands considered. Adjuncts showed a clear amelioration, which was not so clear in the case of *whether* island. On the other hand, in the case of Complex Noun Phrases there is no difference between the structure with resumption and the gapped counterpart. Moreover, they showed that the role of resumptive pronouns depends also on the *wh*-constituent they are linked to: in case it is a simple *wh*-constituent, resumption is not acceptable, but it is when it is related to complex *wh*-constituents of the form *which* + NP. It is also the case of

¹⁷ See references therein for additional evidence regarding Hebrew.

Keshev (2016) for Hebrew, who tested extraction from three types of adjuncts: temporal adverbial clauses (*after*, *before* and *while*) with finite verbs, *in order to* with non-finite verbs, and *instead of*. Sentences with resumptive pronouns were compared to sentences with gaps. Resumption was rated higher in all of the adjuncts proposed, but it reached significance only in two of them, i.e. temporal adjuncts and *instead of*. The same holds for Irish, as McCloskey (2017) shows with an informally collected corpus in which resumption is favoured only in contexts that are claimed to be difficult to process. This means not only that different types of islands lead to different resumption effects (Tucker et al. 2019), but also that different types of adjuncts show different levels of acceptability (Keshev 2016; Keshev & Meltzer 2018).

A last explanation of a lack of an effect due to RP is connected to a theoretical explanation. Remember that there is a difference among grammatical resumptive pronouns, and intrusive resumptive pronouns. We may think that the difference between intrusive and grammatical relies on syntax for grammatical RPs and the PF interface. In the latter case they are not licensed by the grammar, but they are still a way to mark the base position of the dependency since they are basically bundles of features. On the one hand, grammatical RPs can be explained with a base generation + movement approach where the *wh*-constituent and the resumptive pronoun are both present in syntax, but only the former moves leaving the latter stranded in its base position (see also Boeckx 2003, 2008, 2012; Belletti 2006). If this was the case, we could distinguish intrusive RPs, which only affect PF but have no association with the *wh*-element in syntax, contrary to grammatical RPs. Intrusive RPs are simply the overt versions of gaps and therefore sensitive to islands: they do not fix a problem in the derivation.

4.5.4 Summing up

Results from Experiment C can be summarized as below:

- (i) the island effect was more present than in the previous test, since in this case not even extraction from *without* led to a non significant effect. The two cases that obtained a different result were those of *in order to* and of *until*, but this is likely to be an effect due to the poor acceptability of the parasitic gap;
- (ii) the results showed once again that adjuncts behave differently, especially in the supposedly acceptable examples of parasitic gaps;
- (iii) The case of resumptive pronouns complicates the analysis: in some cases we obtain a significant effect (in those where both the island and the parasitic gaps were not

accepted at all, i.e. *in order to* and *until*), but in one other, *without*, the effect is significant with the reverse result: the presence of RP makes the sentence significantly worse than the gap.

4.6 Conclusions

I am going to sum up the main results from the three experiments explained in this chapter. First of all, results were pretty similar for all the tests – even the pilot one discussed in §4.3. The first result that should be discussed regard the possibility of extraction from the adverbial clauses I proposed, namely the investigation of the effect of STRUCTURE. Two adjuncts, *after* and *before*, lead to sound results: extraction in these cases is consistently unacceptable, in all three experiments. Extraction from *without* behaves differently: in Experiment A there was no effect, whereas in Experiment B there was, but it was much weaker than with the other adjuncts. On the other hand Experiment C showed a highly significant effect. As expected, extraction from *because* is not licensed. In two cases, i.e. *in order to* and *until*, the result was the same but not significant. However, this has to be connected to the lower acceptability of parasitic gaps formed with these adjuncts, rather than to an increased acceptability of island sentences. This means that sentences with an island were always rated lower than those with a parasitic gap, and that such difference was significant in most cases. Even though we have to talk about an island effect, we still have to consider that there is a difference between adjuncts, which can be seen not only from island structures, but also in the case of parasitic gaps, in all the experiments. This is more unexpected, given that parasitic gaps are traditionally considered acceptable sentences. However, there is a big difference in the acceptability of these sentences. I showed and argued that such difference is connected to the adverbial clause types only, given that the sentences presented were equal. Therefore, one of the main conclusions is that we cannot think of prepositional adjuncts as forming a uniform category: there are differences among them that leads to different effects with respect to extraction phenomena, both in islands and in parasitic gaps.

As explained in §4.1, experimental syntax is particularly useful in that it permits us to isolate a potential factor able to influence the island effect, and test whether this is really true with formally-collected judgments. After discussing the several factors that are apparently able to strengthen the island effect in §3.2, I decided to formally investigate two of them: the presence of finite tense, and the presence of a resumptive pronoun. The former has rarely been

tested (Michel & Goodall 2012), and results of Experiment B confirmed that when the verb of the adjunct is finite, extraction is affected in the bad way, as was often claimed in the literature (Cinque 1990; Manzini 1992; Szabolcsi & Lohndal 2017, among others). Resumption is usually considered a mechanism able to rescue or ameliorate the island, depending on the language taken into account. As many other tests assessing the acceptability of sentences containing islands and resumptives (Heestand et al 2012; Polinsky et al. 2013; Beltrama & Xiang 2016), this result was not found in most of the adjuncts considered. However, notice that it does not lead to a uniform conclusion – as expected if adjuncts do not behave uniformly. In fact, in Experiment C in two adjuncts the presence of RPs made the sentence significantly better, whereas in case of *without*, it made the sentence significantly worse than its gapped counterpart. These findings make it difficult to explain this difference in syntactic terms.

Therefore, experimental syntax did help uncover previously unobserved patterns: the differences among adjuncts, that holds not only in islands but also in parasitic gaps, as well as the differences among speakers, which I did not mention in the course of the chapter but which will be analyzed in details in the next sections. We are left with two main consequences that we have to deal with, and that will be analyzed in chapter 5 and chapter 6 respectively. Before proceeding with the analysis, I will first consider some issues connected with these tests, and also the consequences my findings have on reductionist theories.

4.6.1 Issues of the tests and of statistical analysis

A first issue with the experiments regards the empirical premises of the factorial design itself. Notice that I decided to compare island structures to sentences involving parasitic gaps, given that this would have let me obtain a minimal contrast. The premise therefore was that the (supposedly) unacceptable case – the island – was to be compared to a (supposedly) acceptable case – the parasitic gap construction. Several scholars have often considered these sentences as completely acceptable, with different theoretical explanations that will be set aside for the moment. The basic description, however, is that the presence of an illicit gap inside an island is licensed thanks to a higher ‘good’ gap outside the island. The acceptability of parasitic gaps was also shown in different experimental works (see Wagers & Phillips 2009). Results from both Experiment A and B showed that this indeed was the case: sentences with islands were rated significantly lower than those with a parasitic gap (even though the degree of significance is different depending on the adjunct – see the special status of *without*). What was surprising

was that there was a difference in the degree of acceptability among these constructions. This led to a first assumption, i.e. that we have to talk about a difference among adjuncts and be ready to account for it. Such a conclusion was confirmed in Experiment C, where more adjuncts have been taken into account. Notice that in this case the statistical analysis gave confusing results with respect to the role of structure, which turned out not to be significant in two cases. However, as discussed before, this is not due to a higher acceptability of island structures, but rather to a lower acceptability of parasitic gaps. In these cases they seem not to be allowed at all, and for such reason no significance is obtained. Therefore, different results could have been found if a safe and sound structure would have been used for the comparison with islands. On the other hand, it would have not let us uncover a very interesting pattern, previously not considered.

The main aim of these tests was to analyse the effect of several independent variables on the score of the sentences, i.e. the dependent variable. In this sense, score was considered as a function of different fixed effects, such as structure, the presence of a resumptive pronoun, or the presence of finite tense. In the statistical analysis, the score was predicted by these effects. However, things are not so simple: the formula used is extremely reduced (and it has to be, in a sense), in that there are many variables that are necessarily left out and that nonetheless are (supposedly) able to influence the thing we were looking for. It is highly unrealistic realising a complete model. Therefore, the second major issue of these tests has to do with the amount of variables that has to be left out, for the sake of simplicity and the limited number of items. These tests were meant to check four points, divided into different experiment, the main one being the effect of islands in different types of adjuncts. I also wanted to check whether some effect could strengthen or weaken the island effect, for such a reason some variables were introduced, such as (i) embedding (Experiment A), (ii) resumption (Experiment A and C) and (iii) the role of finite tense in the adjunct (Experiment B). But of course many variables were left out, because there are simply too many things to look at that might have a role in these structures, and it is impossible to check them out with some tests. This has to do not only with the factors discussed in §3.2, but also with different aspects considering the way items were presented. For instance, in all these tests, the sentences were never introduced by a context, which could make them more comprehensible (see Tanaka 2015; Beltrama & Xiang 2016; Lindahl 2017; Kush et al. 2019; Bondevik 2018; Müller 2019). This means that one more (possible) ameliorating effect had to be left out. The same goes with the task assigned, i.e. acceptability versus comprehensibility of the structures, and to the type of A' dependency

investigated, i.e. *wh*-questions rather than topicalizations or relative clauses (see Sprouse et al. 2016; Kush et al. 2018 compared to Kush et al. 2019).

4.6.2 *What about reductionist theories?*

The last issue I am going to briefly discuss in this chapter regards the effect these results have on reductionist theories, confirming that these cannot be the sole source of explanation for the existence of islands. First, let me step back and describe briefly reductionist theories, introduced in §2.1.3. Different from grounded theories, reductionist accounts blame the island effect solely on extra-grammatical factors: grammar does not prohibit these kinds of movements, which are thought to be difficult from the point of view of processing, of extra-grammatical factors. There are actually many things connected to processing which could go wrong in islands, such as working memory or lexical semantic processing (Kluender & Kutas 1993), limits on attention spans (Deane 1991), or problems with the comprehension of the sentence. These are only a few of the explanation offered by reductionist accounts, but there is no unified approach.

Many scholars have already discussed some of the main issues with reductionist approaches,¹⁸ but there are a few things to point out in these experiments. First, it is worth mentioning the lack of an effect of embedding, which is not what a processing based explanation expects: the longer the filler-gap dependency, the more difficult the structure. This was not the case of Experiment A – confirmed also by results of Beltrama & Xiang (2016). Second, more crucially, why different adjuncts should lead to different acceptability results from a processing point of view? Notice that variability is found in the same structures of the same language, with equal *wh*-dependencies based on the same memory mechanisms, therefore a reductionist account should predict no difference in island effects if we are dealing with the same structure, as in the case of adjunct islands. This is not what happens here: the fact that some islands have a weaker effect remains unaccounted for and thus constitutes another major challenge for reductionist approaches. Third, parasitic gaps are a puzzle for processing accounts: as showed in all the experiments, they are mostly accepted even though they should be more difficult to process. If adjunct islands (and other islands as well) are unacceptable because it is too difficult to create a gap inside an adjunct, then it should be equally difficult (or more difficult) if the gap is licensed by another well-formed gap. Nonetheless they tend to be

¹⁸ See Boeckx (2012: 37) for a good overview, claiming at the end that “they forget to point out that this assortments of processing factors finds a direct reflex in grammar”.

licensed, although they are subject to variation – which is once again not accounted for in reductionist accounts. Thus, parasitic gaps undermine the prediction of reductionist theories stating that island effects are there because they are too difficult (see also Engdahl 1983; Phillips 2006). Moreover, it has been shown that the island effect cannot be attributed to the working memory capacity (e.g., Sprouse, Wagers & Phillips 2012), i.e. one of the main candidate to explain islands with processing accounts.¹⁹ Therefore, the conclusion is that it is unlikely that the ultimate source of islandhood is to be connected solely to processing factors. Grammatical factors have (also) a role, and this will be considered in the next chapter.

¹⁹ Not to mention the fact that island effects are present also in languages with *wh*-in-situ (Huang 1982; Lasnik & Saito 1992).

5. AN ANALYSIS OF EXTRACTION

In this chapter I will deal with a theoretical explanation accounting for the (im)possibility of extraction from adverbial clauses. As discussed in chapter 2, such a possibility can help us establish the type of relationship an adjunct have with the main spine of the tree, i.e. if the two are connected somehow or if they are on two separate planes. If extraction is licensed, we know that there has to be a connection between matrix clause and adjunct. Most of the analyses reviewed in §2.3 predict that this is never possible. However, extraction from adjuncts is sometimes allowed, as shown in chapter 3 with both Italian and cross-linguistic examples. Such a possibility is restricted and manipulated by means of many factors, most of them discussed in §3.2. These can either be applied to different island types, or to adjuncts only. The possibility of extraction was addressed in chapter 4 with some experimental studies that showed that extraction phenomena are influenced by the type of adjunct taken into account, in that some of them allow it, whereas in others it is completely blocked. An interesting result regards the differences among adverbial clauses. Therefore, in this chapter I will mainly address two questions: (i) why is there a difference among adjuncts when extraction is taken into account – notice that this is not restricted to supposedly unacceptable cases, i.e. islands, but affects also constructions that should be grammatical, namely parasitic gaps;¹ and (ii) why is extraction allowed in some cases but so degraded in others?

In this introduction I sketch the analysis I will propose in the following sections. The ingredients for extraction I will talk about are mainly two:

- Point of merger in the structure: attachment to a lower portion of the tree favours extraction, whereas attachment to higher ones blocks it. This also accounts for the different behaviours of adverbial clauses, noted both experimentally and theoretically.

¹ We could take a step further and ask whether the difference between adjuncts is connected not only to extraction, but to some other phenomena as well. I will try to tackle such a question in the course of the chapter, however the issue will not be discussed in details.

- Possibility of forming a macroevent: a semantic condition allowed by syntax, since it is dependent on the preceding condition. This accounts for the impossibility of extraction from higher adjuncts, as well as the differences among speakers (see chapter 6).

Both solutions have already been proposed in the literature. The macroevent approach was first proposed by Borgonovo & Neeleman (2000), and later expanded by Truswell (2007, 2011). Different attachment heights have been hinted in Narita (2011), Boeckx (2012, 2014), Sheehan (2013) and Brown (2016, 2017). I will show that simply being attached to a certain height is a necessary, but not sufficient condition. The additional requirement is the possibility of forming a macroevent, which is allowed only to adjuncts merged at a certain height and that do not have a too complex internal structure (see also Haegeman 2003, 2004, 2012; Endo & Haegeman 2019). I will explore in details the proposal in the following sections. Moreover, I will discuss how extraction is allowed in these cases, taking into account a reanalysis of the adjunct into a complement.

5.1 Differences among adverbial clauses

In the literature, prepositional adjuncts have mostly been considered as an undifferentiated and homogenous group. In the tests presented here, the main focus has been on the possibilities of extraction and on the licensing of parasitic gaps. As I showed, different adjuncts have different results with respect to these properties. It thus seems that we cannot talk about adjuncts as a uniform group, in that they do not behave in the same way and some sub-types can be individuated. This means that we need a criterion able to sort out these differences. As I will show in the first part of this chapter, my claim is that a possible criterion regards their level of attachment to the structure. First, I will briefly review the results from my own experiments, and then I will move on to different tests that have been made recently and that show similar results (Bondevik 2018; Müller 2019). This will allow me to claim that they are truly different from one another and are not part of a uniform group.

Notice that I will be interested only in prepositional adjuncts. Moreover, I will not talk about the difference already studied by Haegeman (2012) between central and peripheral adverbial clauses. Thus, I will refer only to adverbial clauses of the central kind, which were the ones investigated in the experiments. I will argue that there are differences among this group as well, and that their different behaviours reflect differences of their internal and external syntax.

5.1.1 Experimental evidence reviewed

In this section I will briefly report the relevant results of the three experiments I conducted and that were discussed in chapter 4. The main goal was to check whether extraction from prepositional adjuncts could be licensed. To do so, I compared island structures – traditionally considered as strong islands – to parasitic gaps, usually thought of as grammatical sentences. Such a comparison allowed me to obtain a minimal contrast between the sentences, although, as explained in §4.6.1, it might not have been the best way to check the island effect. Remember that the sentences I proposed were equal to each other, as can be seen in (1), where some of the items presented for Experiment C for the island gapped condition are presented.

- (1) a. Quale libro Sergio è tornato a casa [per leggere _]?
‘Which book did Sergio come back home (in order) to read?’
- b. Quale ragazzo Elisa è partita [senza salutare _]?
‘Which boy did Elisa leave without greeting?’
- c. Quale libro Gianni si è preoccupato [prima di trovare _]?
‘Which book did John worry before finding?’
- d. Quale ragazza Piero è corso via [dopo aver baciato _]?
‘Which girl did Piero run away after having kissed?’
- e. Quale film hai insistito [finché non ho guardato _]?
‘Which movie did you insist until I watched?’
- f. Quale libro ti sei arrabiato [perché ho rovinato _]?
‘Which book did you get angry because I have ruined?’

All sentences in (1) represent cases of extraction from an adverbial clause, with different conjuncts introducing them. The verb in the adjunct is always a transitive one in its non-finite form (given §3.2.1 and experimental evidence discussed in §4.4.3.2), which shares the subject with the matrix sentence, except in (1e) and (1f). Because of the conjunct used, in these cases it is not possible to use the verb in its non-finite form: *until* and *because* ask for finiteness – thus we can already notice that they behave differently. In all cases in (1) extraction of an argument is considered (see §3.2.2), which is realized as a *wh*-constituent in a D-linked form of the type *which* + NP (as discussed §3.2.3). Moreover, the adjunct is always in its postverbal position and should thus get its normal interpretation (§3.2.8). Notice that at least in one case,

namely in (1f), I used an adjunct that is presumably very high in the structure (§3.2.6; Uriagereka 2011; Boeckx 2012). The verbs of the matrix clauses are similar to one another, given that they are mostly unaccusatives. Finally, when it comes to the syntactic integration of the adverbial clause, I used conjuncts that either have only a central adverbial reading following Haegeman (2004, 2012), as in the case of *after*, *before* and *until*, or that have both a central and a peripheral reading, but only the central one was tested. This is the case of *in order to*, which is used in its purpose reading, and of *because* which introduces an event reason. Lastly, the case of *without* is never taken into account in Haegeman (2003, 2004, 2009, 2010, 2012), but it is related to the event expressed in the matrix, as can be seen from (2), showing that a *without* clause can be clefted.

- (2) È senza salutare Sara che Marco è partito.
'It's without greeting Sara that Marco left.'

Basically, all the adverbial clauses in (1) should be sufficiently integrated in the matrix clause in order to modify the event described there, and therefore potentially able to allow extraction.

As I showed, the sentences proposed were all equal to each other, except for the cases of *until* and *because* for which the only option in Italian is to have the verb in the adjunct in the finite form. I voluntarily left out all the possible distinctions that were not connected to the type of adverbial clause and I manipulated all the factors investigated in §3.2 in order to obtain a weaker island effect. In so doing, if a difference were found, it would have let me claim that it was purely connected to the kind of adjunct taken into account. This was the expectation after the first pilot experiment discussed in §4.3, which showed a difference among the three adverbial clauses investigated: on the one hand, extraction from temporal adjuncts introduced by *after* and *before* led to a main effect of structure, i.e. extraction from them is unacceptable. On the other, this effect was not significant in the case of *without*, showing that extraction is licensed in this case.² Experiment B in §4.4 confirmed such a difference among the adjuncts just mentioned. In this case, the effect of structure reached significance also in the case of

² For the time being, I am focusing only on the effect of structure, leaving aside other differences among these adjuncts, such as the role of resumptive pronouns discussed in both Experiment A and C. As I showed, however, different adjuncts leads to different effects also with respect to the use of resumptives (see also Tucker et al. 2019). I am also letting aside a major difference between adjuncts, i.e. the presence of inter-speaker variability, which will be taken into account in chapter 6.

without, but it was much weaker than the other adjuncts. In Experiment C I incremented the sample of adverbial clauses, doubling them. Results regarding *after* and *before* were confirmed: extraction always leads to unacceptability. A strong effect was also found in *without*, differently from the previous tests. Among the new ones, *because* does not seem to allow for extraction, as expected given (i) the role of finite verbs in the adjunct, which cannot be controlled for in this case, and (ii) the fact that it is presumably attached higher in the structure. The peculiar cases of *in order to* and *until* were addressed as issues of the test, and in particular because of the fact that parasitic gaps formed with these conjuncts are not always accepted. In both cases there was no effect of structure, meaning that islands were not rated significantly lower than parasitic gaps. This could have been expected in the case of *in order to*, which, if integrated with the main clause, creates a contingent relation with the matrix verb, following Truswell's (2007, 2011) semantic analysis. On the other hand, it was not expected with *until*, given the mandatory finiteness of the verb, as well as the additional difficulty created by the presence of the negative word. A second look at the results of parasitic gaps allowed me to establish that the lack of an effect was not due to a higher acceptability of the islands, but rather to a lower acceptability of the parasitic gaps. In any case, the crucial result holds in this case as well: there is a difference among adverbial clauses when it comes to extraction. It is clear if we look at Figure 5.1, presented in §4.5.3.1 and repeated here. The figure shows the mean of z-scores judgments from Experiment C for all the adjuncts taken into account. They are ordered from the most unacceptable to the most acceptable.



Figure 5.1. Z-score acceptability judgments comparing extraction from each adjunct type for the island and the gap condition of Experiment C

The difference among adverbial clauses is very much clear; extraction from *until* is twice as bad as extraction from *without*, which is – as expected – the weakest case of island effect. Let's have a more in depth look at Figure 5.1 and at the asymmetries it shows. The islandhood of *until* is expected, for both the fact that the verb has to be realised in its finite form and the presence of a negative word in the items, i.e. *not*. In fact, the presence of some elements such as negative words could be considered an additional factor able to influence the acceptability of extraction. Notice however that this conjunct did not have an alternative, neither for the finiteness of the verb, nor for the use of the negative word. Also the cases of *after* and *before* are expected: given the results from previous tests, we already knew that extraction from temporal adverbial clauses was not accepted. Notice that there seems to be a difference in the acceptability between the two: extraction from *before* (the mean of z-scores is -0,8) is more accepted than extraction from *after* (mean of -0,94). Such a difference is not found in Experiment B (where extraction from *after* has a mean of z-scores of -0,94, and extraction from *before* of -0,92) and is very little in Experiment A as well. *In order to* quite surprisingly is almost equal to *because*, which I expected to show a bigger island effect, due to (i) the obligatory use of a finite verb, (ii) the different subject between matrix and adverbial clause and (iii) the fact that *because* is taken to be attached higher (see Uriagereka 2011; Boeckx 2012). I will leave aside this issue for now. The last one, namely *without*, is the one that best allows for extraction, which was also expected: given previous results we knew that this case shows a minor effect, even though in Experiment C it resulted highly significant. Such a fact can be connected to the existence of different island effects. As stated in Kush, Lohndal & Sprouse (2019), we might distinguish between a classic island effect and a subliminal one, where we are not simply considering the significance of the effect, but also its (un)acceptability on the absolute scale. Thus, the classic island effect refers to an effect that is not only significant, but also very unacceptable on the absolute scale. In the subliminal effect, on the other hand, the difference between the island and its non-island counterpart is significant, but sentences are nonetheless judged within the range of acceptability on the absolute scale (see discussion in Almeida 2014; Keshev & Metzger-Ascher 2018).

As already stated, the surprising fact is that differences among adjuncts do not involve only illicit cases of extraction, i.e. islands. In fact, the same asymmetries are found in parasitic gaps as well, namely cases of extraction that should be licit, or to better put it, that should be made licit thanks to the presence of a higher gap. The results from all experiments showed that this is not really the case: acceptability is – very – different depending on the adjuncts

considered. Notice once again that the only difference among the sentences I presented can be attributed to the kind of adverbial clause investigated. This can be seen from sentences in (3), which are the direct counterparts of those presented in (1).

- (3) a. Quale libro Sergio ha acquistato _ per leggere _?
 ‘Which book did Sergio buy _ in order to read _?’
- b. Quale ragazzo Elisa ha guardato _ senza salutare _?
 ‘Which boy did Elisa look _ without greeting _?’
- c. Quale libro Gianni ha cercato _ a lungo prima di trovare _?
 ‘Which book did John look for _ before finding _?’
- d. Quale ragazza Piero ha salutato _ dopo aver baciato _?
 ‘Which girl did Piero greet _ after kissing _?’
- e. Quale film ti interessava _ finché non hai guardato _?
 ‘Which movie were you interest in _ before looking _?’
- f. Quale libro hai sistemato _ perché avevi rovinato _?
 ‘Which book did you fix _ because you had ruined _?’

Following many scholars starting from Engdahl (1983), parasitic gaps are considered grammatical structures. This means that sentences in (3) should be all considered equally good, given that they are very much alike. Let me go through their properties. All cases above represent the extraction of an argument, realized as a complex *wh*-phrase made by *which* + NP. Both the verbs in the matrix clause and the adjunct are transitive, both are realized with a finite tense and the subject is shared by the two clauses. Notice however that there are some differences in the sentences presented in (3). The matrix verb of (3c) is connected to an adverbial clause but it is also preceded by an adjunct *for long*, which might influence the acceptability of the sentence, making it less acceptable. In the adjunct of (3e) we also find the negative word *not* present in the island counterpart (1e). As stated before, this might be an additional factor affecting the acceptability of the sentence. A part from that, things should be neat. This is – of course – not what happens: parasitic gaps are accepted (or not accepted) differently depending on the adverbial clauses. This can be seen in Figure 5.2, presented first in §4.5.3.1 and reproduced here as well. It shows the mean value of the z-scores judgments assigned to the parasitic gap condition with no resumption of Experiment C.

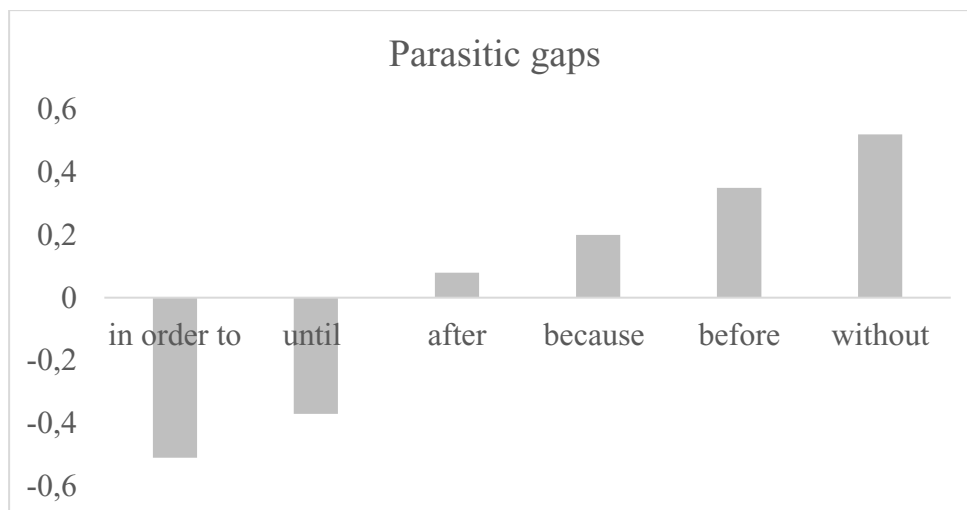


Figure 5.2. Z-score acceptability judgments comparing extraction from each adjunct type for the parasitic gap without resumptive pronoun condition.

The first thing that can be noted is that for two adjuncts the parasitic gap condition seems to be really bad. It is the case of *until*, which was expected if we consider the presence of the negative element, and especially the low acceptability of its island counterpart. Following what we know about parasitic gaps, the latter should not be an influencing factor. More surprisingly in this respect is the case of *in order to*, which is the worst, so much so that it is comparable to its island counterpart. The case of *after* is also pretty bad, a result which is also confirmed by the previous tests, where it was the worst option for the three adjuncts. Notice that in the case of parasitic gaps there is a bigger difference between *after* and *before* than in the case of islands given above. In fact, the parasitic gap formed with *before* is mostly accepted – even though in some of the items such as (3c) there is an additional adjunct connected to the matrix verb. The case of *because* is halfway between the two temporal adverbial clauses, whereas *without* is always the best option – even though in Experiment C it is not as good as in the previous ones.

Such an asymmetric pattern has never been formally observed before, to my knowledge.³ As anticipated above, the main conclusion we can draw is that prepositional adjuncts are not a uniform group: there is something going on either with their internal or external syntax, or both – as claimed in Haegeman (2012) with respect to central and peripheral

³ Notice that a variation between adverbial clauses was already observed by Engdahl (1983: 4) where she individuates a hierarchy, partially reproduced here.

(i) manner > temporal > purpose > when/because/conditional if

Notice that the first three adjuncts are non-finite, whereas the last three are finite. Hence, the variation can be partially attributed to the role of tense. However, this is not helpful for the first part of the scale.

adverbial clauses. In the next section, I will dwell some more on experimental data from prepositional adjuncts showing that they are different. In doing so, cross-linguistic data will be taken into account in order to show that these differences are alive and kicking in more languages.

5.1.2 Cross-linguistic and experimental data. Asymmetries part II

The aim of this section is to take into account different data on extraction from adjuncts from an experimental point of view, namely with formally-collected judgements. I will briefly review the results obtained in some tests. Notice that there are not many experiments working with different types of adverbial clauses and investigating these constructions only.

First, I will consider the cases of Heestand et al. (2012) and of Polinsky et al (2013), which were already briefly discussed in §4.2. The role of these tests was to investigate the effect of different types of islands, also considering adjuncts. In both tests the adverbial clauses considered were introduced by *while* and *although*. Taking into account Haegeman (2004), the former can have a double reading, i.e. a central one when it modifies the event, or a peripheral one with the role of contrasting utterances. Unfortunately no examples are given in Heestand et al. (2012), so we don't know which one of the two readings was proposed to the participants. On the other hand, *although* has only a peripheral reading, hence the prediction is that the adverbial clause is not sufficiently integrated to allow extraction. Examples of the items presented are reported in (4):

- (4) a. This is the dish that, although the chef overcooked __, the guests were not upset.
b. Which dish did Gina think that, although the chef overcooked __, the guests were not upset? (Heestand et al. 2012: 147)

As already discussed in §4.2, we do not expect sentences of this kind to be accepted, because of different reasons: (i) the verb in the adjunct is finite; (ii) the position of the adjunct is preverbal and is realized with a comma intonation; (iii) *although* is an adverbial clause of the peripheral type. Adjuncts were investigated both with *wh*-questions (4b) and declaratives (4a) in an online test. Considering the factors mentioned above, results are surprising: there is only a slight amelioration between sentences with islands and grammatical controls. Even though this effect is most probably due to the position of the adjunct (see §3.3.8), both questions and

(especially) declaratives have a mean value of more than 3.5 on the basis of a 7-point Likert scale.⁴ Reaction times were also very similar. Their conclusion is that the adjuncts they investigated are very weak and hence showed a great transparency, and that adjuncts are a heterogeneous group.

The next findings I am going to present come from Mainland Scandinavian languages. It has already been noted that these languages allow extraction from adjunct clauses, even when finite tense is present (Anward 1982 for Swedish; Faarlund 1992 for Norwegian; Hansen & Heltoft 2011 for Swedish, all cited in Müller 2019). Swedish and Norwegian have thus been investigated in this respect (Müller 2017, 2019; Bondevik 2018; Kush et al. 2019) with several formal experimental tests in order to check whether extraction was really allowed. Müller (2017, 2019) tested extraction from different adjuncts in Norwegian with an acceptability judgment test. Crucially, she tested topicalized constituents and all of the items were preceded by a context sentence, as example in (5). The adjuncts could trigger either a peripheral or a central reading and were compatible with a contingent interpretation or not. The adverbial clauses considered were temporal clauses (5), purpose clauses (6), concessive clauses (7), causal clauses and conditional clauses.

- (5) Two persons are having a preparty and talk about which wine they should drink.
A: Vilket vin ska vi dricka ikväll? Det vita eller det röda vinet som du har kvar?
'Which wine should we drink tonight? The white one or that red wine that you have left?'
B: Hellre det vita.
'Rather the white one.'
Det där röda vinet mådde jag lite illa efter att jag hade druckit sist.
this there red wine-the felt I a-little sick after that I had drunk last time
'I felt a little sick after I had drunk that red wine last time.' (Müller 2017: 74)
- (6) Et här berget måste man träna mycket för att man ska kunna bestiga.
this here mountain must one exercise a-lot for that one shall be-able to-climb
'One has to exercise a lot in order to be able to climb that mountain.' (Müller 2017: 84)

⁴ Though notice that the two types of adjuncts are unified in the analysis, meaning that we cannot know whether it is an effect of one of them only or if their results were really similar and comparable.

- (7) Men den boken kunde jag inte lösa uppgifterna fastän jag hade läst.
but this book could I not solve assignments-the although I had read
'But I could not solve the assignments even though I had read this book.'

(Müller 2017: 91)

Results point towards a distinction between these adjuncts: they show a different behaviour when extraction is taken into account. Extraction from purpose clauses introduced by *in order to* is well-formed (with a mode and median of 4 on a 5-point Likert scale). This was expected both following the approach of Truswell (2007, 2011) and the classification of Haegeman (2012). Temporal clauses were introduced by *after* and *when*, and the results are mixed: in the case of *after*, when they expressed purely a temporal relation, as in (8), extraction was not allowed. On the other hand, when the temporal relation was enriched and expressed a contingent relation triggering a causal interpretation of the events, as in (9), extraction was allowed. Both readings were introduced thanks to manipulation of contexts. The same is not true when the adverbial clause is introduced by *when*, where independently from the interpretation extraction did not obtained good results.

- (8) Den filmen fick jag gå hem efter att vi hade sett.
that movie must.PAST I go home after that we had seen
'I had to go home after we had seen that movie.'

- (9) Den filmen börjar man alltid gråta efter att man har sett.
that movie start one always to-cry after that one has seen
'One always starts crying after having seen that movie.'

(Müller 2017: 76)

A similar result is obtained with conditional clauses introduced by *if*. Two readings are available: both the central and the peripheral one. Only with the former is extraction allowed, even though the difference between the two is not as straightforward as expected, given the results of the temporal clause.⁵ On the other hand, results clauses introduced by *so that*, which present a coherent relation with the matrix clause but are peripheral adverbial clauses, do not allow extraction. The same applies to concessive clauses introduced by *although*, as well as

⁵ See Müller (2019) for a detailed discussion.

causal clauses introduced by *because*, which does not show a difference in extraction licensing depending on the central versus peripheral reading.⁶ The overall results from Swedish do not sustain previous analyses claiming that extraction is always acceptable in Mainland Scandinavian languages, but they show that is rather conditioned by some factors – as I showed is the case for Italian in §3.2. The factors individuated by Müller (2019) are (i) the syntactic integration of the adverbial clauses, i.e. the division between central and peripheral adverbial clauses (see §3.2.7), (ii) their consequent internal syntax, that should not be too rich (see Haegeman 2012), and (iii) the adjunct and the matrix clause events are related by a contingent relation (based on Truswell 2011). The results from this experiment show that extraction is only possible from central adverbial clauses, given that they are syntactically integrated in the matrix clause. The same is not possible for peripheral adverbial clauses. Extraction is allowed in cases of semantic cohesion, namely when the matrix and the adjunct are in a contingent relation. This can be seen from the examples of temporal clauses: when the relation is strictly temporal, extraction is bad, but when a causal reading is forced by the context, extraction is allowed.

Kush et al. (2019) investigate different island effects in Norwegian with a follow-up study, in order to check whether the same results of Kush et al. (2018) are obtained. In the latter, extraction from islands is tested through *wh*-questions, whereas in the former with topicalization.⁷ Sentences were presented either with no context or with an introductory sentence. The only type of adjuncts investigated here are conditionals introduced by *if*. See sentence in (10).

⁶ This is a puzzling result, since *because* is an adverbial clauses of the central type and it expresses a contingent relation of causation, therefore predicting that extraction is allowed following an analysis à la Truswell (2007, 2011). Müller (2017) explains such low acceptability claiming that the internal syntax of central adverbial clauses is more elaborated than is usually the case with adverbial clauses of the central kind, as explained in §3.2.7 (Haegeman 2012). However notice it might also be connected to its point of merger, given that an adjunct like *because* is thought to be attached higher in the structure (see Uriagereka 2011; Boeckx 2012).

⁷ Notice that we usually treat different A' dependencies such as *wh*-questions, relative clause formations and topicalizations as a uniform phenomenon, i.e. subject to the same syntactic rules. On the other hand, it is true that these are connected to different semantic and pragmatic constraints. Therefore, according to Kush et al. (2019) testing different A' dependencies can help us establish whether the effect is connected to syntax only, as predicted if the same results are obtained with different A' dependencies, or if extra-syntactic factors can be blamed as well. Their results push towards the latter conclusion.

- (10) Bakdøren blir han nervøs om de lar stå ulåst.
Back.door.DEF gets he nervous if they leave stand unlocked
'The backdoor he gets nervous if they leave unlocked.' (Kush et al. 2019: 400)

The results showed that there is a significant effect, but it is smaller than in the other islands tested – especially subject islands. Moreover, scores are bimodally distributed: it is judged as acceptable in many trials, but unacceptable in others. They conclude that conditional *if* might not be a structural island in Norwegian, thus hinting that adjuncts are not uniform syntactic islands. The presence of context had a crucial role: it helped the realization of topicalization, introducing a discourse adding a contrastive value to the fronted argument. In this case, they did not find any island effect for the conditional, which was judged almost as acceptable as topicalization from embedded declaratives. They concluded that conditionals in Norwegian are not islands for topicalization.⁸

Bondevik (2018) studied different types of adverbial clauses, introduced by *if*, *when* and *because*, together with subject islands and *whether*-islands.^{9,10} Following Kush et al. (2019) Bondevik tested topicalized structures, rather than *wh*-questions. Moreover, all the items were introduced by a context sentence aimed to provide contextual motivation for the items. An example is given in (11).

⁸ It seems that extraction is really affected by the type of A' dependency given previous results obtained in Kush et al. (2018), where there was a significant effect for extraction from adjuncts in the form of *wh*-questions. In the case of topicalization, especially when introduced by a context sentence, the effect is not there. See also results from Sprouse, Caponigro, Greco & Cecchetto (2016) that apparently confirm that we might have to treat A' dependencies differently. In this test, extraction from adjuncts (among other island types) was investigated with both *wh*-dependencies and relative clauses formation. Only in the former a significant effect was found, but crucially not in the latter. Since my experiments only focused on *wh*-questions, I won't deal with the issue here, but clearly further studies are needed in order to assess this point.

⁹ In fact, based on previous tests on Norwegian, it is known that subject islands always show a strong effect, whereas *whether*-islands do not show an effect at all in the majority of cases. This means that the two types could offer a good comparison for adjuncts as well. See Kush et al. (2018, to appear) for a more in-depth analysis.

¹⁰ The author specifies that all the adjuncts taken into account can be classified as conditional (following Bhatt & Pancheva 2006), and that this was the available reading for the items proposed in the test.

- (11) a. *Context sentence:*
 Mette er ikke fornøyd med sommertemperaturene I Nord-Norge,...
 Mette is not satisfied with summer-temperatures-DEF in North Norway,...
 ‘Mette is not happy with the summer temperatures in Northern-Norway,...’
- b. men [vintertemperaturene] blir hun boende fordi hun liker _ .
 but winter-temperatures-DEF becomes she living because she likes
 ‘but the winter temperatures she stays there because she likes’.
- (Bondevik 2018: 48)

These sentences were tested with three experiments, where the second and the third were meant to investigate the presence of variation, given that the number of items for each conjunct was incremented, allowing to extend the observation per participants and per adjunct types. This means that variation was considered not only between participants – and within participants – but also within adjunct items.¹¹ Results confirm that islands are significantly worse, in every case. There is however a crucial difference between the items tested. Extraction from *because* seems not to be accepted, and it is almost as bad as subject islands. Extraction from *when* and from *if* is sometimes licensed, especially in the latter case, which is similar to the results of *whether*-island. Bondevik found that acceptability is deeply affected by the conjunct introducing the adverbial clauses, to the point that the three classes above were shown to be statistically different from each other. The conclusion is that it doesn’t make sense to talk about ‘adjunct’ as a natural and coherent class in Norwegian, because extraction is not blocked uniformly, as would be expected if it were a uniform island type.

The results I presented here force us to ask whether it makes sense to talk about an adjunct islands class, since it has been shown that significantly different results are obtained depending on the type of adverbial clause used.¹² Notice that in several cases I presented here (Bondevik 2018; Kush et al. 2019; Müller 2017, 2019), extraction is investigated by means of

¹¹ As for the difference and the results of the participants, I will report and discuss them in chapter 6.

¹² Interestingly enough, it has also been shown that different types of adjuncts are extractable to a different degree. Kiziak (2010: 77) individuates a continuum focusing on German, where a causal adjunct is more difficult to extract than a temporal one and of a manner one (*warum* > *wann* > *wie*).

topicalization, rather than *wh*-dependencies, as in my experiments. This might tell us that the type of dependency investigated is able to affect the possibility of extraction.¹³

5.1.3 *So far*

There are three main consequences I want to discuss before going on with the analysis. These consequences are the result of the experiments I conducted and the cross-linguistic data reviewed so far in this chapter. The first one regards the class of adjuncts: these are usually considered part of a uniform class and therefore explained by means of the same analysis. However, the data just discussed show that this approach is not correct. There are (major) differences among adverbial clauses, which has been shown both experimentally (Bondevik 2018; Müller 2017, 2019; Dal Farra 2019a, b) and theoretically (Sheehan 2013; Brown 2017; Endo & Haegeman 2019). My claim is that the only way to account for their different behaviours is first to admit that we cannot treat them as one and the same mechanism; there are differences among them that should be captured syntactically.¹⁴ This is the main topic of next section, where I propose an analysis of such difference based on the attachment level of adjuncts.

The second important issue regards the status of adjuncts with respect to movement, i.e. their degree of islandhood. Recall that they are usually considered strong islands (see Szabolcsi & Lohndal 2017) on the basis of the diagnostic offered in Cinque (1990), in that extraction of both arguments and adjuncts is not allowed. However, I have shown that this is too strong: extraction in some cases is accepted – even though it is maybe best to describe it as *marginally* acceptable in the light of the results of Experiment C (see also Tanaka 2015). Data from Swedish (Müller 2017, 2019) and especially Norwegian (Bondevik 2018) support this view. The argument/adjunct asymmetry presented in §3.2.2 and repeated here in (12) shows that adjuncts do behave more similarly to weak islands than to strong ones.

¹³ See also Engdahl (1997) and Lindahl (2017) for the extraction from relative clauses. The latter shows that topicalization is the most frequent type of extraction in spontaneously uttered extractions in Swedish.

¹⁴ Notice that in the literature there are several proposals in order to capture the differences between adjuncts and arguments, i.e. Pair-Merge and Set-Merge (Chomsky 2004, 2008), doubly-rooted structures (Johnson 2009; Oseki 2015), and labeled versus non-labeled – or too complex labeled – structures (Nunes & Hornstein 2008). However, there are no syntactic mechanisms able to capture the dual nature of the ‘same’ phenomenon, i.e. adjunction.

- (12) a. Quale ragazza Gianni è partito [senza salutare _]?
 which girl Gianni is left without greeting
 ‘Which girl did John leave without greeting?’
- b. *Con quale ragazzo Maria e Sara sono partite [senza aver parlato _]?
 with which boy Maria and Sara have left without having spoken
 ‘With which boy did Mary and Sara leave without having spoken?’

As already argued, a distinction based solely on the adjunct/argument asymmetry in the distinction of islands is not always straightforward and often is too coarse (Szabolcsi & den Dikken 2003; Szabolcsi & Lohndal 2017). For this reason, Tanaka (2015) offers a three-way distinction based on Cinque (1990) and the A’ dependencies he analyses. There are strong islands, as well as two types of weak ones. As usual, strong islands are absolute: they prohibit any kind of dependency formation, independently from the argument/adjunct distinction. Weak islands can either be selective or non-selective. The former allows for *argument* extraction of both NP and PP type, as long as they are referential (or D-linked following Pesetsky 1987, see §3.2.3), but extraction of adjuncts is not permitted. Non-selective weak islands, on the other hand, allow only for referential NPs argument extraction. According to Tanaka (2015), finite adjuncts belong to the class of strong islands, whereas non-finite adjuncts to the class of selective islands.¹⁵ Following Tanaka (2015), I may argue the same, as long as non-finite adjuncts are considered: rather than strong islands, I will consider them weak islands. Notice though that the results of the experiment discussed in chapter 4 suggest rather that some adjuncts really are strong islands, in that extraction was never licensed, as in the case of those introduced by *after*. On the other hand, other adjuncts are weak islands – or they are not islands at all – as seems to be the case of those introduced by *without*. However, as discussed in chapter 3, it is important to remember an additional issue, namely the fact that there are several factors able to manipulate the island effect, making it more or less acceptable depending on their manipulation. If adjuncts really were strong islands as traditionally stated, it would be hard to explain why the factors above have an effect on extraction: we do not expect strong islands to be ameliorated thanks to the presence – or absence – of some elements. In fact, remember that the factors able to ameliorate the effect are valid throughout different adverbial clauses, and not

¹⁵ Notice that also Boeckx (2008) considers only clausal adjuncts to be strong islands. However such a view is problematic with data from Swedish (Müller 2019), Norwegian (Bondevik 2018; Kush et al. to appear) and Czech (Biskup & Šimík 2018).

only in the case of *without*. From experimental evidence, it appears that a binary distinction is not sufficient, and we should rather think of a scale of island strength (see Lindahl 2017; Müller 2019).

Third, there is empirical evidence showing that extraction from adjuncts is not always blocked. As already explained, this means that the majority of the theories discussed in §2.3 is not suitable, and therefore that we must come up with an explanation that is able to account for such data. Notice that this theory needs to compromise between being too strict, and therefore not allowing any kind of movement, and being too permissive. We have to be able to explain extraction, but this has also to be restricted to some cases, contra what the Connectedness theory (Kayne 1983) predicted. Moreover, such theory should also be able to accommodate the ameliorating factors discussed in §3.2. In the next sections I will first discuss what might be a good criterion able to account for the different behaviour of adjuncts, which is also connected to the (im)possibility of extraction and to a crucial factor which will be addressed in §5.3. §5.4 will be dedicated to the final proposal, after a brief review of some of the theories that deal with extraction cases from adjuncts.

5.2 Attachment of adjuncts

In the previous sections I showed that adjuncts do behave in different ways when it comes to extraction. This is true not only for Italian, but also for Swedish and Norwegian, that is to mention only languages that have been studied with formal experiments. Such difference between adjuncts can also be seen in the case of parasitic gaps, which are supposedly grammatical sentences. I have shown that this is not really the case: the type of adverbial clause used matters a great deal in the acceptability of these structures. At this point, it is difficult to state that it makes sense to talk about a class of ‘adjunct islands’, because it looks like these elements are not a natural class. The results obtained in the tests discussed in chapter 4 and §5.1 point towards a different direction. Notice that such a variation is completely unexpected in every theory of adjunction discussed in §2.3, in that adjuncts are always adjoined – or merged too late, or to a separate plane – to the matrix clause in the same way. None of the approaches taken into account is able to explain the variation between adjuncts: the theory is not fine-grained enough to accommodate the data concerning variation. We need a further subcategorization between adverbial clauses. The criterion I am proposing here is their point of merger in the structure. I will argue that adverbial clauses can be connected to different heights

in the structure, as already proposed by some scholars, and that their place of attachment has an impact on different phenomena, including the possibility of extraction.

A first division has already been discussed in Haegeman (2004, 2012), as reviewed in §3.2.7. This regards central and peripheral adverbial clauses, which are shown to behave differently in many respects, both in their external and their internal syntax. In fact, central adverbial clauses modify the TP or the domain internal to the TP, and are thus within the scope of operators inside TP or CP such as focus operators and interrogative ones. Peripheral adverbial clauses, on the other hand, are not affected by these operators, meaning that they are outside the domain of TP. Haegeman (2012) points out that such a difference is reflected in their external syntax: their points of merger are different as well as their timing of attachment. Central adverbial clauses are merged inside the TP, peripheral ones are merged in parallel with the CP. Basically, the latter are attached higher and later in the structure. There is also a correspondence between external and internal syntax, in that the former influences the latter. Peripheral adverbial clauses have a richer internal syntax (in Haegeman this means that they have illocutionary force), whereas central ones are reduced. Extraction from peripheral adverbial clauses is always ungrammatical, whereas extraction from central ones is sometimes allowed (see Haegeman 2004; Müller 2019 for experimental evidence).

My proposal comes in similar terms. We need to establish a more fine-grained distinction between central adverbial clauses. As we have seen, there has to be a criterion that allows us to sort them out. Notice that I am focusing only on adjuncts that are inside the TP and are thus able to modify it, or the domain internal to it. The level of attachment has often been (part of) the solution adopted by different scholars in order to account for different phenomena. Notice that analyses of adjuncts merged to different levels have already been proposed in the literature (Narita 2011; Sheehan 2013; Fabregas & Jimenez-Fernandez 2016a; Brown 2016, 2017; Müller 2019). In a sense, this was already part of the proposal of the Condition on Extraction Domain (Huang 1982), where extraction is ruled by means of two attachment sites. Phrases are either merged low, as sister to heads, or merged higher, to the maximal projection of a head. In the first case, we are dealing with complements which allow extraction thanks to the government relation; in the latter we are dealing with subjects and adjuncts, domains that are not in a proper government relation with lexical heads and hence do not allow extraction.¹⁶

¹⁶ This was the solution adopted also by Borgonovo & Neeleman (2000) to account for the transparency of predicative adjuncts. Their proposal is still based on the role of L-marking in licensing extraction. The fact that these are predicative is crucial, in that they are part of the

That there is a difference with respect to merging points was also stated in Chomsky (1986) and the *Barriers* system. He proposes that moving an element from a VP adjunct implies the crossing of one barrier only, whereas if movement takes place from a TP adjunct two barriers are crossed, leading to a further degradation. The prediction is that extraction from an adjunct is never fully acceptable, but it is only mildly degraded in case adjuncts are attached lower. Therefore, a crucial role is assigned to the attachment site of the adjunct, in that it is responsible for the availability of extraction. Chomsky based the difference (also) on the ambiguity of (13a).

- (13) a. They were too angry [to hold the meeting].
 b. Which meeting were they too angry [to hold]? (Chomsky 1986: 33)

(13a) is ambiguous between two readings: (i) states that they could be so angry that it is not possible to hold the meeting, whereas (ii) that they are so angry that it is not safe for us to hold the meeting. Interestingly, when extraction takes place, only the former reading is available. The difference in meaning is connected to a difference in the height of attachment, where (i) emerges when the adjunct is internal to the VP, whereas (ii) when it is external to the VP. When extraction takes place, as in (13b), only the reading in (ii) is blocked because crossing two barriers, as would be the case for a VP-external adjunct, leads to a severe degradation.

5.2.1 *Where do they merge?*

Central adjuncts are usually thought to be merged in different positions: TP, *v*P, *v* or VP, depending on the proposal. From tests regarding scope, we know that the highest position they can occupy is TP: not only are they able to modify the TP or the domain internal to it, but they are also within the scope of the matrix tense and of operators inside TP or CP. In fact, as claimed by Haegeman (2012), given that a central adverbial clause can be in the scope of a focusing operator, it means that it can be clefted, (14a) and (15a), whereas this is not allowed with a peripheral clause, as can be seen in (14b) and (15b). They are also in the scope of interrogative operators of the matrix clause, as in (16).

thematic structure of that sentence. In order for them to be transparent they need to be in a complement position under VP, i.e. lower in the structure, or as they state in a lower than usual position.

- (14) a. È mentre sei al mare che puoi rilassarti.
 (it) is while you are at the sea that you can relax
 ‘It is while you are at the beach that you can relax.’
- b. *È mentre Silvia ha finito il dottorato che Elisa si deve ancora laureare.
 (it) is while Silvia has completed the PhD that Elisa is have still to graduate
 *’It’s while Silvia completed her PhD that Elisa still has to graduate.’
- (15) a. È perché parla con me che Gianni suda.
 (it) is because (he) talks with me that Gianni sweats
 ‘It is because he talks to me that Gianni is sweating.’
- b. #È perché la sua macchina è in cortile che Gianni è a casa.
 (it) is because the his car is in backyard that Gianni is at home
 ‘It is because his car is in the backyard that Gianni is at home.’

(Cecchetto & Donati 2015: 119-120)

- (16) Hai visitato il MoMA mentre eri a New York?
 have (you) visited the MoMA while (you) were in New York
 ‘Did you visit the MoMA while you were in New York?’

Moreover, central adverbial clauses can be affected by operations that affect the TP or VP domain. It is the case of VP substitution and VP ellipsis. The latter is shown in (17), where the temporal clause introduced by *while* can either have a strict (17i) or a sloppy identity (17ii). The same is not possible with peripheral adverbial clauses, given that they are too high to permit it (see Haegeman 2012: 171).

- (17) a. Mentre la sua tesi veniva discussa, Gianni ha lasciato la stanza ma Luca no.
 while the his thesis was being discussed, John has left the room, but Luca not
- (i) ‘Luca didn’t leave the room while John’s thesis was being discussed.’
- (ii) ‘Luca didn’t leave the room while his own thesis was being discussed’.

Condition C effect obtained in (18) and (19) tells us that the position of the adjunct has to be below that of the subject, meaning that they are at least below TP. If this is the case, we can

explain the ungrammaticality of (18), because adjunction is to a place that is c-commanded by the subject (see also Narita 2011), and therefore the R-expression triggers a Principle C effect.¹⁷

(18) pro^*_i si è sentita male perché/dopo che/senza che Lisa_i aveva mangiato quel piatto.
She is felt sick because Lisa had eaten that dish.

(19) pro^*_i si stanca, se Gianni_i lavora troppo.
(he) gets tired if Gianni works too.much (Cecchetto & Donati 2015: 121)

Moreover, bound variable reading for pronouns inside different types of adverbial clauses is available, and it can be bound by the quantifier in both the direct and indirect object position, meaning that these adverbials are c-commanded by both theme and goal DPs, as is shown in (20).

(20) a. They will accept every paper_i for publication *if/after* it_i is positively evaluated by two reviewers.
b. They will give every female teacher_i a substantial salary increase *after/if/so* (*that*) she_i accepts more teaching load. (Valmala 2009: 961)

Notice that the same is not true for *because* clauses, which appear to be outside the c-command domain of the internal argument. As (21) shows, the pronoun cannot be bound by a quantifier in both direct and indirect object position.

(21) a. *They accepted every paper_i for publication *because* it_i had been positively evaluated by two reviewers.

¹⁷ This is true only when the R-expression appears inside an adverbial clause of the central type, contrary to what happens if it appears inside a peripheral one, as seen in (i). This supports one again the structural asymmetry between adverbial clauses shown by Haegeman (2003, 2004, 2012).

(i) ? pro_i deve aver lavorato molto se Gianni_i ha le occhiaie.
(he) has worked a.lot if Gianni has the shadows
'He_i must have worked hard, if Gianni_i has dark circles under his eyes.'

(ii) ? pro_i è a casa, perché la macchina di Gianni_i è in cortile.
(he) is at home because the car of Gianni is in backyard
'He_i is at home, because Gianni_i's car is in the backyard.'

(adapted from Cecchetto & Donati 2015: 121)

- b. *They gave every female teacher_i a substantial salary increase *because* she_i had accepted more teaching load.
- c. *They gave a substantial salary increase to every female teacher_i *because* she_i had accepted more teaching load. (Valmala 2009: 961)

As (21) shows, causal clauses are in a higher position than the one occupied by both the direct and the indirect object. (22) shows that, nonetheless, it is still lower than the position of the subject, given that Condition C effects are obtained.

- (22) a. Every female teacher_i has been hired because she_i is intelligent.
- b. *She_i has been hired because Mary_i is intelligent. (Valmala 2009: 961)

Evidence from adverbial scope tests also supports the difference in height, as noted by Brown (2016, 2017). Consider as an example a subject-oriented adverb like *reluctantly* and a VP adverb like *loudly*. The latter can scope either over the first predicate only, or over both matrix and adjunct as a whole, as in (23a) and (23b). This tells us that the adjuncts are not merged higher than *vP*.

- (23) a. Mary reluctantly ate an ice cream before she whistled the national anthem.
- b. Monica reluctantly arrived whistling the national anthem. (Brown 2017: 62)

On the other hand, the VP adverb can modify the matrix predicate or it can scope over both matrix and adjunct as a whole. This is possible only in the case of (24a), but not in (24b). This means that the prepositional adjunct in (24a) is attached to a lower position than *loudly*, such as VP, whereas attachment is higher in (24b), where the adverb cannot scope over the entire complex: it has to merge higher than VP – but lower than the position of *reluctantly*, probably to *vP*.

- (24) a. Monica loudly arrived whistling the national anthem.
- b. #Mary loudly ate an ice cream before she whistled the national anthem. (Brown 2017: 63)

5.2.2 *Is a hierarchy of adverbial clauses possible?*

In order to fully capture the properties of different adverbial clauses and check whether there are some restrictions on their relative order, I discuss here a hierarchy of adverbial clauses.¹⁸ First, it can be noticed that whenever one of the adjuncts is connected to the other rather than to the matrix clause, there are severe restrictions regarding their ordering. The adjunct modifying the other adjunct has to follow it, as in (25a) and (26a), otherwise the sentence is ungrammatical, as in (25b) and (26b). This is the expected relation.

- (25) a. Maria è andata a casa dopo aver parlato con Sara senza raccontarle la grande novità.
'Mary went home after talking to Sara without telling her the big news.'
- b. *Maria è andata a casa senza raccontarle la grande novità dopo aver parlato con Sara.
'Mary went home without telling her the big news after talking to Sara.'
- c. Dopo aver parlato con Sara Maria è andata a casa senza raccontarle la grande novità.
'After talking to Sara Mary went home without telling her the big news.'
- d. *Senza raccontarle la grande novità Maria è andata a casa dopo aver parlato con Sara.
'Without telling her the big news Mary went home after talking to Sara.'
- (26) a. La polizia ha fermato Gianni per interrogarlo per poi lasciarlo andare.
'The police stopped John to interrogate him and subsequently let him go.'
- b. *La polizia ha fermato Gianni per lasciarlo andare per interrogarlo.
'The police stopped John to subsequently let him go to interrogate him.'
- c. ??Per interrogarlo la polizia ha fermato Gianni per poi lasciarlo andare.
'To interrogate him the police stopped John and subsequently let him go.'

On the other hand, if we try to combine adjuncts which are independent from one another and both related to the matrix sentence, there are some interesting restrictions. In the

¹⁸ Notice that the kind of sentences I will present are particularly difficult to judge. Therefore, judgments come primarily from linguists.

sentences that follow I combined the different adverbial clauses that were used in Experiment C, i.e. those introduced by *after*, *before*, *without*, *in order to*, *because* and *until*. I discuss the results of the relative ordering of all the adverbial combined with each other after the main clause, as well as their position on the left of the matrix, namely what was tentatively analysed as an adverbial clause with a parenthetical interpretation. I will then present a hierarchy that sums up the results.¹⁹

Remember that usually adverbial clauses appear in a sentence-final position (Diessel 2001; Valmala 2009). This is often considered the basic order for adjuncts with respect to the matrix sentence in VO languages. In fact, notice that in such a position matrix and adjunct are part of the same intonation contour, i.e. they are not separated by the so-called comma intonation, or by a real comma in written text. This is also a crucial difference with peripheral adverbial clauses, which “are prosodically set off from the associated clause. Peripheral adverbial clauses are also set off by a comma, whether they be sentence-initial or sentence-final. Initial central adverbial clauses may be followed by a comma, but this is not always the case, and sentence-final central adverbial clauses are usually not preceded by a comma” (Haegeman 2012: 168). In the test mentioned here, therefore, the sentences are all intended to be part of the same intonation contour, especially when both of them appear in a sentence final position. When this was not possible, I signalled that a comma must be present, in order for the sentence to be acceptable.

Senza > dopo:

- (27) a. Mario è andato a casa dopo aver parlato con Sara, senza fare la spesa.
‘Mario went home after talking to Sara, without going grocery shopping.’
- b. Mario è andato a casa senza fare la spesa dopo aver parlato con Sara.
‘Mario went home without going grocery shopping after talking to Sara.’
- c. Dopo aver parlato con Sara, Mario è andato a casa senza fare la spesa
‘After talking to Sara, Mario went home without going grocery shopping.’
- d. *Senza fare la spesa Mario è andato a casa dopo aver parlato con Sara.
‘Without going grocery shopping Mario went home after talking to Sara.’

¹⁹ See also Schweikert (2004) on the ordering of prepositional phrases in the structure of the clause, and on the difficulties of individuating a hierarchy. Notice that his focus is on prepositional adjuncts not of the clausal type.

Senza > prima

- (28) a. Gianni è tornato a casa prima che scoppiasse il temporale, senza fare la spesa.
'John came back home before the storm without going grocery shopping.'
- b. Gianni è tornato a casa senza fare la spesa prima che scoppiasse il temporale.
'John came back home without going grocery shopping before the storm.'
- c. Prima che scoppiasse il temporale, Gianni è tornato a casa senza fare la spesa.
'Before the storm, John came back home without going grocery shopping.'
- d. *Senza fare la spesa Gianni è tornato a casa prima che scoppiasse il temporale.
'Without going grocery shopping John came back home before the storm.'

*Dopo, prima²⁰

- (29) a. *Gianni è partito dopo aver visto Silvia prima che scoppiasse il temporale.
'John left after seeing Silvia before the storm.'
- b. *Gianni è partito prima che scoppiasse il temporale dopo aver visto Silvia.
'John left before the storm after seeing Silvia.'
- c. *Dopo aver visto Silvia Gianni è partito prima che scoppiasse il temporale.
'After seeing Silvia John left before the storm.'
- d. *Prima che scoppiasse il temporale Gianni è partito dopo aver visto Silvia.
'Before the storm John left after seeing Silvia.'

Senza ? per

- (30) a. Gianni è andato a casa per prendere le chiavi senza pensarci due volte.
'John went home to take the keys without thinking twice.'
- b. Gianni è andato a casa senza pensarci due volte per prendere le chiavi.
'John went home without thinking twice to take the keys.'
- c. Per prendere le chiavi, Gianni è andato a casa senza pensarci due volte.
'To take the keys John went home without thinking twice.'
- d. Senza pensarci due volte, Gianni è andato a casa per prendere le chiavi.
'Without thinking twice John went home to take the keys.'

²⁰ I have not been able to come up with an example with both *before* and *after* referring to the matrix clause and not to each other. Notice that (29a) can be rescued with a coordination between the adjuncts.

Senza > perché:

- (31) a. Serena è partita senza salutare Gianni perché aveva prenotato l'albergo.
'Serena left without greeting John because she booked the hotel.'
- b. Serena è partita perché aveva prenotato l'albergo, senza salutare Gianni.
'Serena left because she booked the hotel, without greeting John.'
- c. Senza salutare Gianni, Serena è partita perché aveva prenotato l'albergo.
'Without greeting John, Serena left because she booked the hotel.'
- d. *Perché aveva prenotato l'albergo Serena è partita senza salutare Gianni.
'Because she booked the hotel Serena left without greeting John.'

Senza > finché:

- (32) a. *Lisa ha continuato la sua carriera per anni finché non si è stufata senza avere rimpianti.
'Lisa pursued her carrier for years until she got sick of it without having regrets.'
- b. Lisa ha continuato la sua carriera per anni senza avere rimpianti, finché non si è stufata.
'Lisa pursued her carrier for years without having regrets until she got sick of it.'
- c. Finché non si è stufata, Lisa ha continuato la sua carriera per anni senza avere rimpianti.
'Until she got sick of it, Lisa pursued her carrier for years without having regrets.'
- d. Senza avere rimpianti, Lisa ha continuato la sua carriera finché non si è stufata.
'Without having regrets, Lisa pursued her carrier for years until she got sick of it.'

Per > dopo:

- (33) a. *Sono andata a casa dopo aver visto Luca per prendere le chiavi.
'I went home after seeing Luca to take the keys.'
- b. Sono andata a casa per prendere le chiavi dopo aver visto Luca.
'I went home to take the keys after seeing Luca.'
- c. Dopo aver visto Luca sono andata a casa per prendere le chiavi.
'After seeing Luca I went home to take the keys.'
- d. *Per prendere le chiavi sono andata a casa dopo aver visto Luca.
'To take the keys I went home after seeing Luca.'

Dopo > finché:

- (34) a. *Sono rimasta in giro finché non ho visto Luca dopo aver fatto la spesa.
'I stayed out until I saw Luca after going grocery shopping.'
- b. Sono rimasta in giro dopo aver fatto la spesa finché non ho visto Luca.
'I stayed out after going grocery shopping until I saw Luca.'
- c. *Finché non ho visto Luca sono rimasta in giro dopo aver fatto la spesa.
'Until I saw Luca I stayed out after going grocery shopping.'
- d. Dopo aver fatto la spesa sono rimasta in giro finché non ho visto Luca.
'After going grocery shopping I stayed out until I saw Luca.'

Dopo > perché

- (35) a. Silvia ha mangiato una mela dopo aver fatto la spesa perché aveva fame.
'Silvia ate an apple after going grocery shopping because she was hungry.'
- b. Silvia ha mangiato una mela perché aveva fame, dopo aver fatto la spesa.
'Silvia ate an apple because she was hungry after going grocery shopping.'
- c. Dopo aver fatto la spesa Silvia ha mangiato una mela perché aveva fame.
'After going grocery shopping Silvia ate an apple because she was hungry.'
- d. *Perché aveva fame Silvia ha mangiato una mela dopo aver fatto la spesa.
'Because she was hungry Silvia ate an apple after going grocery shopping.'

Per prima??

- (36) a. Sono andata a casa prima di salutare Luca per prendere le chiavi.
'I went home before greeting Luca to take the keys.'
- b. Sono andata a casa per prendere le chiavi prima di salutare Luca.
'I went home to take the keys before greeting Luca.'
- c. Prima di salutare Luca sono andata a casa per prendere le chiavi.
'Before greeting Luca I went home to take the keys.'
- d. *Per prendere le chiavi sono andata a casa prima di salutare Luca.
'To take the keys I went home before greeting Luca.'

Prima > finché??

- (37) a. *Sono rimasta in giro finché non ho visto Luca prima di fare la spesa.
'I stayed out until I saw Luca before going grocery shopping.'

- b. Sono rimasta in giro prima di fare la spesa finché non ho visto Luca.
'I stayed out before going grocery shopping until I saw Luca.'
- c. Finché non ho visto Luca, sono rimasta in giro prima di fare la spesa.
'Until I saw Luca, I stayed out before going grocery shopping.'
- d. *Prima di fare la spesa sono rimasta in giro finché non ho visto Luca.
'Before going grocery shopping I stayed out until I saw Luca.'

Prima > perché:

- (38)
- a. Silvia è scappata prima di fare la spesa perché aveva visto Gianni.
'Silvia run away before going grocery shopping because she saw John.'
 - b. *Silvia è scappata perché aveva visto Gianni prima di fare la spesa.
'Silvia run away because she saw John before going grocery shopping.'
 - c. Prima di fare la spesa, Silvia è scappata perché aveva visto Gianni.
'Before going grocery shopping, Silvia run away because she saw John.'
 - d. *Perché aveva visto Gianni Silvia è scappata prima di fare la spesa.
'Because she saw John Silvia run away before going grocery shopping.'

Per > finché:

- (39)
- a. Silvia ha aspettato per andare a fare la spesa finché non ha visto Luca.
'Silvia waited to go grocery shopping until she saw Luca.'
 - b. *Silvia ha aspettato finché non ha visto Luca per andare a fare la spesa.
'Silvia waited until she saw Luca to go grocery shopping.'
 - c. *Per andare a fare la spesa Silvia ha aspettato finché non ha visto Luca.
'To go grocery shopping Silvia waited until she saw Luca.'
 - d. Finché non ha visto Luca, Silvia ha aspettato per andare a fare la spesa.
'Until she saw Luca, Silvia waited to go grocery shopping.'

Per > perché:

- (40)
- a. Luca è tornato a casa per prendere le chiavi perché stava aspettando Gianni.
'Luca came back home to take the keys because he was waiting for John.'
 - b. *Luca è tornato a casa perché stava aspettando Gianni per prendere le chiavi.
'Luca came back home because he was waiting for John to take the keys.'

- c. Per prendere le chiavi Luca, è tornato a casa perché stava aspettando Gianni.
'To take the keys Luca came back home because he was waiting for John.'
- d. *Perché stava aspettando Gianni Luca è tornato a casa per prendere le chiavi.
'Because he was waiting for John Luca came back home to take the keys.'

Finché > perché:

- (41) a. Luca non può uscire finché non ha finito i compiti perché gli è stato proibito.
'Luca cannot go outside until he finishes his homework because they forbid it.'
- b. *Luca non può uscire perché gli è stato proibito finché non ha finito i compiti.
'Luca cannot go outside because they forbid it until he finishes his homework.'
- c. Finché non ha finito i compiti Luca non può uscire perché gli è stato proibito.
'Until he finishes his homework Luca cannot go outside because they forbid it.'
- d. *Perché gli è stato proibito Luca non può uscire finché non ha finito i compiti.
'Because they forbid it Luca cannot go outside until he finishes his homework.'

As can be noticed, not all orders are accepted. This is true not only when both the adjuncts follow the verb, but also when one of them is in a preverbal position. In the latter condition, only in some cases it is possible to have an adverbial clause at the beginning of a sentence without a special comma intonation.²¹ In sentences (27)-(41), in the majority of the cases, when an adjunct can be realized in a sentence initial position, it also needs to be intonationally separated from the main clause. Such fact agrees with the parenthetical interpretation that was discussed in §3.2.7. This was noted also by Lobo (2002), who distinguishes two types of adverbial clauses: non-peripheral adverbial clauses, for which the unmarked order is the final one without any intonational break, and peripheral ones, where the unmarked position is the initial one and the final can be used but with an intonational break. For the former, the initial position is also allowed, but in some cases it is easier than in others: for temporal and conditional it is perfect, whereas for reason and purpose it is more marked, as is the case for Italian.

On the other hand, when the adjuncts are both following the verb, in their most natural position, not all orders are possible. On the contrary, in the majority of the cases there is a clear

²¹ This was already noted by McCloskey (2006), who showed that temporal clauses introduced by *after*, *before* and *since* cannot be realized as left-peripheral adjuncts.

ordering between the two – except in the case of the temporal clauses and in *without* and *in order to*, where the order is not so straightforward. We find the hierarchy in (42):

- (42) Senza/per > prima/dopo > finché > perché
Without/in order to > before/after > until > because

The canonical order of central adverbial clauses has been investigated also by Valmala (2009), who considers the order of adjuncts expressing time, cause, condition, and purpose in a sentence final position, and with neutral intonation. She obtains a paradigm very similar to (42), represented in (43):

- (43) Time > Condition/Purpose > Cause

The adjunct expressing time (in Valmala only *before* is taken into account) always precedes the other adjuncts (it can follow only with a comma intonation – in which cases they are taken to be parentheticals). Event condition adverbial introduced by *if* always precedes cause, but it can either precede or follow purpose introduced by *so that*. Purpose clauses always precede cause clauses. See (44) for the entire paradigm:

- (44) a. I will hide the Wii before the children arrive if it is too late so that they don't get distracted because they have a lot of homework to do.
b. I will hide the Wii before the children arrive so that they don't get distracted if it is too late because they have a lot of homework to do. (Valmala 2009: 959)

Crucially, (43) parallels the order I found for Italian, even if there are some differences in the choice of adverbial clauses. First, Valmala did not check *without* clauses, which are the highest in my order. Second, she checked time adverbials introduced only by *before*, whereas I had both *before* and *after*. Since the two of them cannot appear in the same sentence, it seems that they are in complementary distribution.²² Third, she checked both purpose and conditional adverbials, whereas in my test *in order to* was introducing a final adverbial clause. Despite the

²² Although notice that not all adverbials expressing time are on the same level, see *while*.

differences in some of the adverbial clauses selected, notice that the final hierarchies are really similar to each other.

Quite interestingly, notice that if we combine the hierarchy found in (42) – here repeated as (45) – and we compare it to the possibility of extraction obtained in Experiment C, we have an interesting result. See Figure 5.3, where the results of extraction in the island condition of Experiment C for each adjuncts are represented. We notice that (42) is respected in the reverse order: it is easier to extract from the higher part of the hierarchy, and more difficult from the adjuncts at the lower parts. The exception is represented by *because*, from which extraction is not as bad as expected.²³

(45) Senza > per > prima/dopo > finché > perché

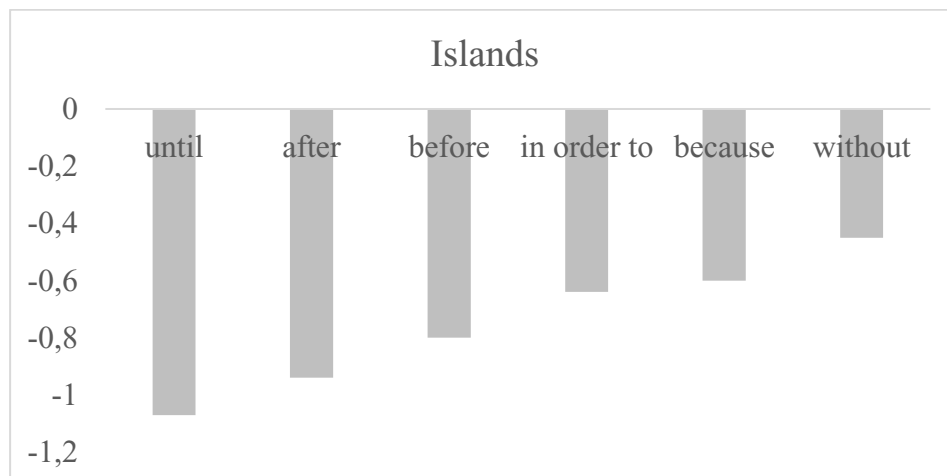


Figure 5.3. Z-score acceptability judgments comparing extraction from each adjunct type for the island and the gap condition.

It can be noticed that, with the exception of *because*, the order is exactly the mirror image of the one hypothesized on the basis of extraction phenomena, where the lower the adjunct, the easier the extraction. (45) is the mirror order with respect to the one expected, since it is easier to extract from *without* and more difficult to extract from *until* (leaving *because* aside for the time being). Therefore, the hierarchical relation should be the reverse one, with *without* being

²³ Notice that *because* cannot be used as a parenthetical, it can never appear at the beginning of a sentence (to do so, we have to change the conjunct that introduces the sentence: *visto che*, *siccome* are okay but not *because*; however notice that these conjuncts cannot be used as answers to a why questions: only *because* is acceptable in that context).

the furthest from the verb and *because/until* the closest.²⁴ An available option in order to obtain the correct order is to hypothesize a movement à la Cinque (2006): the VP moves to the specifier position and in doing so, it pied-pipes the projection that immediately dominates the specifier position from which it is moving. Basically, the pied-piping triggered by the movement of the verb manages to reverse their order.

In order to test the order in (45), we can check whether both orders are possible with parasitic gaps. If not, it means that there really are some restrictions.²⁵ I will use such test for some of the adjuncts proposed, focusing in particular on *without* and its relation with *after* and *in order to*.

- (46) a. L'unica persona che ho guardato _ [senza salutare _] [dopo aver visto _ trattare male i suoi genitori] è Gianni.
 'The only person that I looked [without greeting _] [after seeing _ treating his parents badly] is John.'
- b. *L'unica persona che ho guardato _ [dopo aver visto _ trattare male i suoi genitori] [senza salutare _] è Gianni.
 'The only person that I looked [after seeing _ treating his parents badly] [without greeting _] is John'

Sentences in (46) shows that *without* can only precede *after*, but the other way around is not acceptable. The unacceptability of (46b) is probably due to the fact that the adjuncts headed by *after* is not in its base position, but it has been moved. Therefore, extraction is blocked because of additional movement.

- (47) a. L'unica persona che Gianni ha guardato _ [senza salutare _] [per far arrabbiare _ ancora di più] è Sara.
 'The only person that John looked [without greeting _] [in order to make _ angrier] is Sara.'
- b. *L'unica persona che Gianni ha guardato _ [per far arrabbiare _ ancora di più] [senza salutare _] è Sara.

²⁴ See similar findings in Valmala (2009).

²⁵ Many thanks to Guglielmo Cinque (p.c.) for suggesting this test.

‘The only person that John looked [in order to make _ angrier] [without greeting _] is Sara.’

As shown in (47), there is a restriction in the order of *in order to* and *without*: the latter can only follow the former (47a), otherwise the sentence is not grammatical (47b). Thus, it can help us establish their relative order, which was not individuated in the hierarchy in (42). Therefore, from the test developed here, it is clear that *in order to* is lower than *without*. The hierarchy can be modified as in (48):

(48) per > senza > prima/dopo > finché > perché
in order to > without > before/after > until > because

Interestingly, the same results emerge when extraction is tested from two adverbial clauses and no gap in the matrix sentence. We would expect these sentences to be completely ungrammatical, since there is not only one but two cases of extraction from adjuncts. Sentences in (49a), (50a) and (51a) show that this is not the case: as long as the order in (48) is respected, they are acceptable.²⁶

- (49) a. L’unica persona che sono uscita [senza salutare _] [dopo aver visto _ trattare male i suoi genitori] è Gianni.
‘The only person that I left [without greeting _] [after seeing _ treating his parents badly] is John.’
- b. *L’unica persona che sono uscita dopo aver visto trattare male i suoi genitori senza salutare è Gianni.
‘The only person that I left [after seeing _ treating his parents badly] [without greeting _] is John

²⁶ Notice that extraction in sentences (46)-(51) was not tested with *wh*-dependencies, which were experimentally shown to be more difficult than topicalization (Kush et al. 2019) or relativization (Sprouse et al. 2016).

- (50) a. L'unica persona che Gianni se ne è andato [senza salutare _] [per far arrabbiare _ ancora di più] è Sara.
 'The only person that John left [without greeting _] [in order to make _ angrier] is Sara.'
- b. *L'unica persona che Gianni se ne è andato [per far arrabbiare _ ancora di più] [senza salutare _] è Sara.
 'The only person that John left [in order to make _ angrier] [without greeting _] is Sara.'
- (51) a. L'unica persona che sono uscito [per aiutare _] [dopo aver visto _ in difficoltà] è Gianni.
 'The only person that I went out [to help _] [after seeing _ in trouble] is John.'
- b. *L'unica persona che sono uscito [dopo aver visto _ in difficoltà] [per aiutare _] è Gianni.
 'The only person that I went out [after seeing _ in trouble] [to help _] is John.'

5.2.3 Potential extension of attachment height

An analysis of adjuncts with different points of merger might have the potential to account for factors different than extraction. It can be the case, for instance, of object control in PRO-headed adjuncts. A few words about PRO in adverbial clauses are needed: PRO is usually the null subject of a non-finite adjunct clause, as in (52b). We refer to this relationship as obligatory subject control. Object control in the same kind of sentences is not allowed, whereas it is obligatory with complements, as can be seen from (52a) and (52b) and their Italian counterparts in (53a) and (53b).

- (52) a. John_i persuaded Bill_k [PRO*_{i/k} to leave].
 b. John_i saw Mary_k [before/without/after PRO_{i/*k} leaving the room].
 (Hornstein 2009: 26)
- (53) a. Gianni_i ha convinto Sara_k [PRO*_{i/k} a partire].
 b. Gianni_i ha visto Sara_k [senza PRO_{i/*k} lasciare la stanza].

This means that the subject c-commands the adjunct, which is therefore assumed to be merged below TP (see Hornstein 2009). However, there are some cases in which object control is possible even inside an adjunct, as shown in (54a) and (54b) with rationale clauses.

- (54) a. Sara_i ha punito Gianni_k [per PRO*_{i/k} aver mentito].
b. Sara_i punished John_k [for PRO*_{i/k} lying].
c. They_i punished John_k [for PRO*_{i/k} driving drunk]. (Hornstein 1999: 76 fn 10)

The fact that some adjuncts permit object control might be revealing of their attachment site, as stipulated by Hornstein (1999). Possibly, some adjuncts are attached lower than others and for this reason they allow control into them. Cases similar to (51b) and (52b) can be explained because of the assumption that objects cannot c-command adjunct, given that they are too high in the structure. This means that in cases similar to (54), the object might be able to c-command the adjunct. I am not aware of other examples of adjuncts allowing object control other than the case of *in order to*. Considering also the hierarchy presented in (48), it seems therefore that the adjunct headed by *in order to* is to be found very low in the structure.²⁷

5.3 Macroevent

In the preceding sections I showed that adjuncts are related to different parts of the structure, and I have argued that such a difference is what helps in extraction phenomena: the lower they merge in the structure, the easier it is to extract from them, whereas when they attach higher extraction is more difficult. The attachment site is therefore a requirement for the possibility of extraction. Notice that this is not enough. In my experiments I showed that the easiest case for extraction is when the adverbial clause is introduced by *without*, which means that this adjunct is lower in the structure. However, is it not always the case that extraction is allowed from it, as can be seen from the example in (55). Moreover, as shown in Experiment C, it is sometimes only marginally acceptable.

²⁷ Taking into account the results from Experiment C, it may be the case that *in order to* is attached too low in the structure to allow for extraction, both in the case of the island and in the parasitic gap.

- (55) a. *Quale ragazza Gianni ha raggiunto la vetta [senza salutare _]?
 which girl Gianni has reached the top without greeting
 ‘Which girl did John reach the top without greeting?’
- b. Quale ragazza Gianni è partito [senza salutare _]?
 which girl Gianni is left without greeting
 ‘Which girl did John leave without greeting?’

A further requirement is therefore needed to allow extraction. Notice that the issue of (55a) is the type of verb used in the matrix. Unlike (55b), in (55a) an achievement is used. I will come back at this example at the end of the chapter. I follow the analysis of Truswell (2007, 2011), and I claim that such condition depends on the possibility of creating a macroevent: when this is possible, extraction is licensed, otherwise it is blocked. My claim is different than the one of Truswell in that here it is not purely a semantic requirement, but it is rather a consequence of the syntactic placement of the adjunct in the structure (see also Fabregas & Jimenez-Fernandez 2016a, b; Brown 2016, 2017). This means that the macroevent can be constructed only if the adjunct is in a low position in the structure, but if it is in a high one it will not be formed (or will be formed, but extraction is not allowed anyway. This is probably the case of *because*). I will first explain Truswell’s (2007, 2011) account, and show its effect on experimental works. I will also briefly show independent evidence for the construction of macroevents. I will then focus on some of the main issues of such an approach, arguing that is not enough by itself in order to account for extraction.

5.3.1 *The Single-Event Condition*

An influential analysis first presented in Borgonovo & Neeleman (2000) is that the semantic relation between the matrix clause and the adverbial one is deeply connected to the possibility of extraction. Truswell (2007, 2011) claims that extraction from some adjuncts is not blocked and presents some examples such as those in (56). His main focus is *wh*-extraction from gerunds in English (bare present participial adverbs in his terms). Examples in (56) are to be compared to (57), where extraction is not allowed.²⁸

²⁸ All the examples in (56), (57) and (61) are taken from Truswell (2007a, b, 2011).

- (56) a. Which book did John design his garden [after reading _]?
 b. What did John arrive [whistling _]?
 c. What did John drive Mary crazy [whistling _]?
 d. What are you working so hard [in order to achieve _]?
 e. What did you come in [to talk to us about _] today?
 f. This is the book that John designed the garden [after reading _].
 g. I bet I know what John drove Mary crazy [trying to fix _].
 h. It is this goal that students are working so hard [in order to achieve _].
- (57) a. *What does John work [whistling _]?
 b. *What did John write the cheque [complaining about _]?
 c. *What did John see the spectacle [looking through _]?

What distinguishes (56) from (57) is the Single-Even Condition, formulated as in (58). This condition postulates that movement is constrained by event-based locality domains.

(58) *Single Event Condition*

An instance of *wh*-movement is legitimate only if the minimal constituent containing the head and the foot of the chain can be construed as describing a single event grouping. (Truswell 2011: 157)

(58) explains why the extractions in (56) are allowed: it is possible to combine the event of the matrix verb with the one of the adjunct and form a single event, a macroevent. The same is not true for (57) and for this reason extraction is ruled out. The Single Event Condition needs additional requirements, given in (59) and (60) (Truswell 2011: 157).

(59) *Event grouping*

\mathcal{E} is a set of core events and/or extended events such as that

- (i) every two events $e_1, e_2 \in \mathcal{E}$ overlap spatiotemporally;
- (ii) a maximum of one event $e \in \mathcal{E}$ is agentive.

(60) An event is *agentive* iff

e is an atomic event, and one of the participants in e is an agent; e consists of subevents e_1, \dots, e_n , and one of the participants in the initial subevents e_1 is an agent.

Extraction is licensed provided that the two events can be combined into a single one in a single event grouping. This is facilitated by (i) their spatio-temporal overlapping; (ii) the fact that no more than one event is agentive, and/or (iii) a causal relation connecting matrix and adjunct predicates. The latter option can be obtained also by a spatio-temporal overlap, which is the usual interpretation of gerunds, but not of all kinds of adverbial clauses. In fact, it is influenced by the introducing conjunct. Therefore, a crucial condition according to Truswell (2007, 2011) is that the possibility of forming a macroevent depends on the semantic relation between the matrix and the adjunct. In fact, it has to be a contingent relation, i.e. it has to express causation or enablement. This is usually the case of in *in order to* clauses, where such a relation is linguistically encoded by the conjunct itself. On the other hand, it is not the case of temporal ones, such as (61b). However, notice that even when a temporal adjunct is used, a different interpretation can be obtained, allowing a contingent reading and therefore permitting extraction.

(61) a. What are you working so hard in order to achieve?

b. What did you go home after thinking?

c. Who did John go home after talking to?

d. What did John drive Mary crazy whistling?

e. Which gardening book did John redesign his garden [after reading with great care]?

The temporal interpretation of (61c) can be pragmatically enriched with a causal interpretation, i.e. that John went home as a consequence of talking to someone, i.e. it has a contingent relation, and not simply as a temporal relation between matrix and adjunct. Similarly, (61d) can express the same relation, and the gerund is hence interpreted as the cause of the event of the matrix clause. In order to obtain a single event grouping the two (sub)events have to be connected by a causal relation: the subevent of the adjunct, either realized as a gerund or as an adverbial clause, causes the subevent of the matrix verb. Therefore, the kind of relation expected in order

to permit extraction is e_{adjunct} cause $e_{\text{finite verb}}$, namely the verb of the adjunct is what causes the event expressed in the matrix clause.²⁹

In Truswell (2007, 2011) extraction from tensed adjuncts is totally ruled out. This is so because tensed adjuncts contain their own event operator, Op, that binds the event variable of the adjuncts and hence blocks the construction of a macroevent. Such an operator is also present in non-finite adjuncts, but it is merged after the adjunct is merged to the main structure, thus allowing the construction of a macroevent.

Notice that Truswell (2007, 2011) is a semantic approach, and not a syntactic one. This means that unacceptable cases of extraction are ruled out post-syntactically at the interface, by means of a semantic filter at LF. Or that the filter at LF can license movement out of the adjuncts if the conditions for the single event formation are met, i.e. if semantic conditions expressed in (58)-(60) are met. (58) acts as a semantic filter, as a kind of last resort mechanism able to rescue gaps that violate locality conditions. What leads to the island effect in extraction from non-coherent adjuncts is that the formation of a macroevent is impossible, and we are thus left with two independent events that cannot be combined. Such result is filtered out at the semantic interface. A second option is that the violation of the event grouping leads to an increased difficulty due to processing costs (Truswell 2011: 123). When a multiple-event reading is triggered, an additional event has to be processed, but at the same time the parser is resolving the filler-gap dependency.

The fact that a macroevent can be formed with matrix and adjunct verbs is confirmed by adverbial test scope. High adverbs like *reluctantly*, i.e. subject-oriented ones, can refer to the matrix verb alone, or to the complex made by matrix and adjunct. This is the case of (62a) and (62b), and shows that the adverb is in a higher position than both verbs, i.e. that the adjuncts are lower than the adverb itself.

- (62) a. Mary reluctantly arrived whistling the national anthem.
b. John reluctantly cried after Mary hit a friend. (Brown 2017: 61)

²⁹ The role of Single Event is also supported by experimental evidence, as shown by Tanaka (2015) and Storoshenko (2016) – especially with respect to the role of telicity – for English, Müller (2017, 2019) for Swedish. It is also the case that extraction from parasitic gaps is more acceptable when the matrix and the adverbial clauses express a close relation in English (Wagers & Phillip 2009). See also Kohrt, Sorensen & Chacón (forthcoming) for experimental evidence on English failing to find an effect, most probably due to the kind of sentences they presented.

The same is not possible with a lower adverb, typically a VP one like *quickly*. In this case, the adverb can refer to both predicates only in (63a), showing both that (i) the gerund is in a lower position than the adverb itself, and (ii) that the entire complex can be unified as a whole, i.e. as a macroevent. The same is not true for (63b), where the adverb can only refer to the matrix predicate. This shows that the attachment site is different and also that when this is the case, it is not possible to form a macroevent.

- (63) a. Mary quickly arrived whistling the national anthem.
b. John quickly cried after Mary hit a friend. (Brown 2017: 61)

In this section I described the major proposal of Truswell, namely that the possibility of extraction from adjuncts depends on the possibility of forming a macroevent made by the subevent described in the matrix and the subevent of the adverbial. Notice that the formation of macroevents finds supporting independent evidence in other cases, such as telic pairs (as described by Giorgi & Pianesi 2001; Higginbotham 2009) and analyses of pseudo-coordination (De Vos 2005; Wiklund 2007).

5.3.2 *The Single Event Condition is not enough*

In this section I will discuss the main issues of an analysis à la Truswell (2007, 2011), from both an empirical point of view and a conceptual one. I will start with the former, and for this reason I will go back to some of the results of my experiment, showing that a single event analysis is not able to account for the differences among adjuncts.

First, remember that my experiments revealed a crucial difference among adjuncts. This was evident in the case of adjuncts introduced by *after*, *before* and *without*. In this case, notice that we expect a difference under Truswell approach: temporal connectives introduce different values, and *after* and *before* do not allow for an important condition, i.e. spatio-temporal overlapping. However, notice that the kind of sentences used with *after* could be pragmatically enriched with a causal reading, as shown in (64).

- (64) Quali dolci Luca è scappato dopo aver rubato?
'Which sweets did Luca run away after stealing?'

As explained above, it is not the case that *after* always introduces a temporal connection between matrix and adjunct verbs. In fact, it can also express a causality link between the two, as in the case of (61), where the event of running away is caused by the event of stealing: e_{adjunct} cause e_{matrix} . The two are therefore related by a causal link, expressing a contingent relation, and Truswell (2011) predicts that extraction is favoured in such case. The results show that this is not true: extraction from *after* leads to unacceptability, which cannot be rescued by the causal link between the two: the semantic relation is not enough. This indicates that there is a role of syntax.³⁰

That there is an effect of syntax is shown also by results of Müller (2019), where extraction from both central and peripheral clauses was tested in Swedish. Extraction in the former case was allowed, provided that the single event grouping was respected. On the other hand, extraction from peripheral adverbial clauses was always unacceptable, even when a causal link was constructed, as in the case of result clauses. This shows that a semantic filter by itself is not sufficient, otherwise it would not be possible to explain the difference between peripheral and central adverbial clauses expressed in Haegeman (2012). Moreover, I also want to stress the fact that the majority of the factors described in §3.2 cannot be explained by means of a semantic approach. This is true for example for the position of the adverbial clause in the structure (§3.2.8), or its syntactic integration (§3.2.7), or the asymmetries between adjuncts and arguments (§3.2.2).

Furthermore, languages like Mainland Scandinavian (Müller 2019, among others, also shown experimentally) and Czech, as well as other Slavic languages (Biskup & Šimík 2018) also allow extraction from finite adjuncts, as shown in examples in (65), contrary to what predicted by Truswell.³¹

³⁰ Tanaka (2015) and Müller (2017, 2019) confirm the role of causality with temporal connectives, showing that whenever a causal interpretation can be obtained, extraction is either allowed or significantly better than its temporal counterpart. However, notice that in their tests sentences were introduced by a context meant to express the intended meaning of the sentences, and this might have facilitated the causality reading.

³¹ Notice that some English examples of extraction from a finite adjunct have been considered grammatical (Chaves 2012: 468):

- (i) Which email account would you be in trouble [if someone broke into _]?
- (ii) Which problem would you be devastated [if someone had already solved _]?

- (65) a. Potkal jsem člověka, kterému1 [když to dáš _], tak (to) tě tam pro1/2 pustí]].
 met aux.I man.acc which.dat when it give.you so it.nom you.acc there let.in.3
 ‘I met a man such that if you give it to him, he/she/they will let you in.’
 (Biskup & Šimík 2018: 2)
- b. Sportspegeln somnar jag [om/när jag ser _].
 sports program-the fall-asleep I if/when I watch
 ‘I fall asleep if/when I watch the sports program.’
 (Swedish; Anward 1982: 74)

However, notice that it may be the case that these languages either lack or have a reduced non-finite embedding. If this is the case of Mainland Scandinavian or Slavic languages, it might be that adjuncts that seem to be finite, actually share the semantics of non-finite clauses.³² Notice also that this is not necessarily against the macroevent view, but it may simply go against the presence of an event operator whenever tense is realised, which was the explanation given by Truswell.

Sentences in (66) present another puzzle for Truswell. In fact, (66a) and (66b) are semantically identical but extraction is fine only in (66b). This asymmetry is not expected under Truswell approach, and it shows that the contrast is syntactic rather than (purely) semantic.

- (66) a. ?*Which trial did the DA prove [the suspect to have been at the scene of the crime] [in order to conclude _]?
 b. Which trial did the DA prove [that the suspect was at the scene of the crime] [in order to conclude _]?
 (Narita 2011: 112)

I will now turn to the conceptual issues of a semantic-based approach. Remember that the Single Event condition is presented as a semantic filter, able to rescue sentences with extraction when it is respected. This is a post-syntactic mechanism: the meaning of a sentence determines whether movement is acceptable or not. Following the current theories this is rather problematic: syntax is totally independent from both phonology and semantics. This also means that what is supposed to be a syntactic constraint regarding movement is replicated in a different

³² Thanks to Michelle Sheehan for pointing this out.

module, the semantic one. As noted by Brown (2017), adding a semantic constraint also means giving rise to redundancies in the grammar.

5.4 Combining the two: ingredients for extraction

Section §5.4 showed that the role of the macroevent is crucial, but not sufficient to explain extraction and the differences among adjuncts. Therefore, I propose that the possibility of forming a macroevent is open only for certain types of adjuncts, i.e. those attached low enough. Only when they are low in the structure, it will be possible to form a macroevent, and hence extraction will be allowed. On the other hand, if an adjunct is higher in the structure, it will not be possible to unify the two events in a single one, and hence extraction will not be licensed. I am therefore claiming that the semantic constraint of Truswell (2007, 2011) is the result of a syntactic condition, i.e. the point of merger of the adjunct. Notice that such proposal fit nicely with Haegeman's (2012) distinction between central and peripheral adverbial clauses. Crucially, notice that the data regarding peripheral adverbial clauses cannot be explained under Truswell: the difference proposed in Haegeman (2012) is a structural one having to do with both the internal and external syntax of the adjunct. I have been focusing on the latter for central adverbial clauses. The lack of integration of peripheral adverbial clauses is a result of their attachment site; at the same time, the impossibility of forming a macroevent – even when a causal relation is created between matrix and adjunct – is most probably related to the same reason: the attachment site of that adjunct is just too high to permit its creation.

5.4.1 Previous explanations

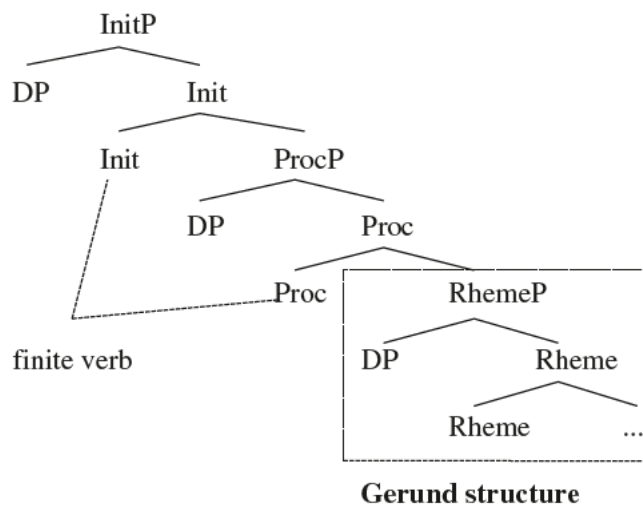
Some scholars already proposed that the semantic approach of Truswell is the result of syntax. I will briefly review here some of the main theories. First, in Fábregas & Jiménez-Fernández (2016a) the adjunct is taken to be a complement, in the sense that it occupies a structural position that allows for a coherent interpretation of matrix and adjunct, i.e. the formation of macroevent. The proposal is syntactic in nature, in that it is not the semantic relation between matrix and adjunct that allows extraction, but their position in the structure. Their work focuses on extraction from gerunds in Spanish. The macroevent results from a specific structure, which they apply to Ramchand's (2008) decomposition of the Aktionsart, represented in (67). Events are structures with internal complexity, and in Ramchand each subevent corresponds to a head.

The maximal amount is made by an *initiation* phrase, codifying the cause component, a *process* phase, codifying the event part, a *rheme* phrase, which introduces descriptors or modifiers of the main event, and a *result* phrase, present with telic verbs.

(67) [InitP [ProcP [RhemeP [ResP]]]]

The transparency of adjuncts is directly linked to (67), because they are taken to be one of the subevents in the verbal domain, whose internal parts are to be identified with those of the matrix predicate. Even though the matrix verb and the gerund look like two distinct verbs from a phonologic and a morphologic point of view, syntactically they are part of one and only verbal structure. Basically, the gerund identifies with one of the subevent of the matrix verb – RhemeP in their proposal – and for this reason it can be merged as a complement of ProcP, as shown in (68) (adapted from Fábregas & Jiménez-Fernández 2016: 1309).

(68)



Such a position is subject to restrictions: the heads of the subevents introduced by the matrix must be different from those that the adjunct lexicalizes. If this is not the case, two subevent heads are present, i.e. two different Aktionsart structures. It thus naturally follows that with some verbs extraction is just not licensed, capturing such restriction not only semantically, but also syntactically. If the matrix verb already fills the RhemeP position by itself, as in the case of activities and accomplishments, the gerund will not be able to integrate in the structure in (64) and extraction is not permitted. On the other hand, extraction is allowed with achievements

as main verbs: in this case the RhemeP position is not filled by the matrix verb and is hence free for the gerund, which can be integrated in the same verbal structure.³³ Extraction is therefore allowed because gerunds are in a complement structure, and not in an adjunct one.³⁴ They are the internal constituents of a predicate. An analysis as the one in (68) also captures an additional restriction for Spanish, i.e. that the gerund has to be adjacent to the matrix verb.^{35,36}

The problem with such an approach is that it cannot be applied to prepositional adjuncts, i.e. the kind investigated in this thesis. These are very different from bare present participial adjuncts. In fact, they are syntactically reduced structures and they are usually in a semantic relation with the matrix verb – as can be noticed from the fact that they usually overlap spatio-temporally (Truswell 2011). On the other hand, adverbial clauses have an internal structure that is more complex than gerunds, and the proposal of Fábregas & Jiménez-Fernández (2016) cannot be easily applied. There is a potential (big) issue in considering gerunds as complements, namely the fact that they are *adjuncts* in nature: they are not selected and are not obligatory. Moreover, notice that assuming (68) with the only restriction being the type of verb of the matrix, the prediction it is making is that in all the other cases extraction is not only licensed, but also obligatory. The problem is that it is not the case that all achievement verbs permit extraction. We are therefore left with a problem regarding how to restrict it.

A different approach is adopted in Brown (2016, 2017), where a crucial role of islandhood is assigned to the status of the head the adjunct is attached to, i.e. phasal or non-

³³ Such a restriction is apparently language-specific. It is the case of Spanish (Fábregas & Jiménez-Fernández 2016), whereas English seems to be more permissive. As shown by Truswell (2007: 1360) extraction is allowed also when the matrix verb is not an achievement, as in (i), where we are dealing with a mixture of states and activities (a Davidsonian state).

(i) What was John lying in bed [reading _]?

³⁴ A complement analysis of transparent adjuncts is offered also by other scholars (see Borgonovo & Neeleman 2000; Wiklund 2007; De Vos 2005 for the same analysis for pseudocoordination).

³⁵ Note the asymmetry in (i) and (ii):

(i) ¿ Qué entró [diciendo _] Juan?
what entered saying Juan

(ii) * ¿ Qué llegó Maria [silbando _]?
what arrived Maria whispering

(Fábregas & Jiménez-Fernández 2016a: 1315)

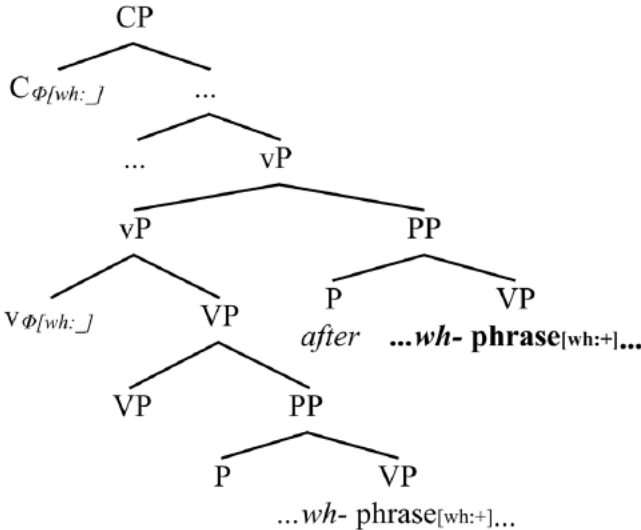
Such an asymmetry also supports the fact that there has to be a syntactic explanation; it would not be possible to analyse it as a semantic condition, showing once again that Truswell (2011) is not sufficient. However, notice also that the unacceptability of (ii) might be connected to the position of the subject, rather than to the non-adjacency of the two verbs. Thanks to Anna Cardinaletti (p.c.) for pointing this out.

³⁶ Notice that this requirement is once again language specific, and does not apply to English:

(i) What did John enrage his neighbours [whistling _]?

phasal. Brown derives islandhood from phase theory (Chomsky 2000, 2001), individuating what she calls ‘blackholes’ in the structure, i.e. nodes dominated by specifiers of a phase head (basically SpecCP and SpecvP). These are areas in the clause that block movement. Two attachment sites are proposed: higher adjuncts merge to vP, whereas lower ones to VP. The latter allows for single-event constructions and is therefore transparent, whereas in the former only non-single event constructions can be formed, meaning that extraction is not allowed. In her proposal, transparency applies to participial adjuncts, whereas opaqueness to tensed adjuncts, as can be seen in (69) (adapted from Brown 2017: 58).

(69)



Phase heads are assumed to contain uninterpretable copies of all the features present in the derivation, which must be checked under an Agree relation in order not to violate Full Interpretation. When an uninterpretable copy is found, movement to the specifier position of the phase head is triggered, and from that position the feature can be checked. In the structure we are looking at, the situation is as in (69): the phase head vP contains the uninterpretable copy of the *wh*-feature. Therefore, *v* searches under its c-command nodes for the interpretable *wh*-feature, and when it finds the feature, it triggers its movement to its specifier, so that the checking relation can be satisfied. This happens when the adjunct is attached to VP, given that the *wh*-feature is in the c-command domain of the phase head vP, which also means that movement applies and that the *wh*-phrase is still available for further cyclic movement even after Spell-Out. The same is not true for a higher adjunct attached to vP: in this case the *wh*-

phrase is outside the c-command domain, too high for v to trigger movement. The derivation does not converge because Full Interpretation is violated.

Such an account relies heavily on the presence of a *wh*-feature in the phase head v , which is potentially problematic since such feature is supposed to be present in any case, in any kind of sentence. When uninterpretable features are not needed, however, they should not be active on the phase head. Moreover, the distinction in Brown (2017) is between high adjuncts introduced by prepositions such as *before* and *after* and low VP adjuncts introduced by a null preposition. We already know that such a distinction is neither efficient nor sufficient. Moreover, notice that basing the difference of phase head versus non phase head is problematic because extraction from an edge is not restricted by anything, given that the PIC itself does not place any restriction (see Gallego 2010; Boeckx 2012; Abels 2012; Tanaka 2015 on the same point). In fact, the edge is not selective for different moving categories. In other words, whenever there is an edge extraction is allowed, and we want something more restrictive than that.

Sheehan (2013) and Narita (2011) have similar ideas accounting for the transparency of adjuncts, both linked to the role of Multiple Spell-Out (Uriagereka 1999) and the opacification of certain domains. Notice that in Uriagereka's (1999) account the only domain that can be opacified is the adjunct, in order to create a monotonic object that can be inserted in the structure strictly following the LCA (Kayne 1994). Sheehan (2013) and Narita (2011) both explore the possibility of atomizing the main spine, i.e. the matrix clause, and not only the adjunct. Sheehan (2013) considers the contrast in (70), which is problematic under Uriagereka's approach: the acceptability of extraction in (70a) shows that the adjunct is not opaque, and (70b) that the main spine is not opaque, given that extraction has taken place from there.

- (70) a. Which exam did you give John some tuition [(in order) to help him _ pass]?
b. What did you give John _ [(in order) to help him pass the exam]?

(Sheehan 2013: 146)

The contrast in (70) is explained in terms of different attachment levels, as can be seen from sentences in (71)

- (71) a. How much tuition did she give him_i _ to help John_i pass the exam?
b. *How much tuition did she give every student_i _ to help him_i pass the exam?

- c. *Which exam did she give him_i some tuition [to help John_i pass _]?
- d. Which exam did she give every student_i some tuition [to help him_i pass _]?

(Sheehan 2013: 146)

The contrast between (71a) and (71c) shows that when extraction has taken place from the main clause, the adjunct is attached higher than the indirect object of the matrix clause, as can be seen from the absence of Condition C effect, present in (71c). In the latter, extraction has taken place from the adjunct, showing that this is attached lower than the indirect object in the structure. In Sheehan (2013) the former is attached to ResP, and the latter to P_{have}P – following Ramchand’s (2008) analysis of the verbal domain. Their attachment sites are also connected to linearization and atomization issues: when a transparent, lower adjunct is attached to P_{have}P, it is the main spine that gets atomized, not the adjunct, and for this reason extraction from the adjunct is permitted, whereas extraction from the main spine should not be.³⁷ On the other hand, when the adjunct does not contain a gap it is attached higher, to ResP. When this is the case, adjuncts are merged too late to participate in the derivation, in order to avoid multiple applications of Spell-Out. They are combined with the rest of the sentence only at PF, meaning that they can only adjoin to a phase head. This is due to the fact that only ResP, but not P_{have}P, is a phase head. Early merger is forced to target a different position than the one of late merge (not clear why). Such proposal requires a bit of stipulation: first it is not clear why precisely that phase head is chosen: Sheehan (2013: 146) does not offer argument in favour of ResP as a phase, and as she herself admits “in the absence of an independent diagnostic for phasehood, the explanation risks circularity”. Second, a particular type of attachment is required in order to explain islandhood, namely Late Merge, which involves some issues.³⁸ Third, it is not clear why early adjunction is prohibited to a phase head.

Narita (2011) has a similar proposal as far as linearization is concerned: XP-XP merger is blocked because of linearization. One of the two phrases has to undergo Spell-Out prior to merge so it can be inserted as a lexical item in the derivation. Such phrase becomes opaque. Either the adjoined material or the main spine gets transferred, depending on the presence of unchecked features.³⁹ If the domain contained in the adjunct does not have any unchecked

³⁷ See Sheehan (2013) for more details on this point.

³⁸ See §2.3.5 and Sportiche (2018) for more details regarding the introduction of Late Merge in grammars.

³⁹ Similar proposals based on the need to transfer immediately the adjunct in order to reduce the complexity of the structure can be found in Boeckx (2012, 2014) where transfer depends on

features, it will be a convergent domain able to constitute a phase. In this case, the adjunct can undergo reduction, thus becoming a LI, i.e. an opaque domain that does not allow extraction. This is what happens to the *because* clause in (72a), as confirmed by the impossibility of extraction in (72b), and to other finite adjuncts.

- (72) a. The man criticized Mary [because she failed the exam.]
b. *This is the girls that John failed the test [because he was thinking about _]
(Narita 2011: 106-107)

Certain adjuncts, however, are allowed not to undergo reduction. Crucially, the distinction is based on attachment height in the structure: the division is above and below v/v^* . Higher adjuncts like *because* are attached higher than v/v^* and are opaque in that they undergo reduction. Lower adjuncts are attached lower than v/v^* – specifically, they can be as low as in the sister of the object of the matrix verb – and are transparent, as can be seen from (70).

- (73) Which book did John design his garden [after reading _]? (Truswell 2011: 31)

In a case like (73), the adjunct headed by *before* – a lower adjunct in Narita’s (2011) account – cannot undergo reduction. In fact, the adjunct contains a *wh*-phrase, which Narita considers a phase headed by a covert Q feature, following Cable (2010). But a phase whose edge contains Q cannot undergo reduction. Therefore, it is the matrix clause that undergoes Spell-Out, while the adjunct remains accessible in the derivation.

One of Narita’s major issue is to explain why in a sentence as (72b) the adjunct can undergo reduction even though a Q feature is present, which means that it should be in the phase edge of the adjunct, therefore not allowing Spell-Out. This is left open. A crucial issue, in common with Sheehan (2013) is to predict which one of the two phrases in a XP-XP configuration is chosen for Spell-Out and reduction (see Boeckx 2012).⁴⁰

The last approach I am considering here is the one by Oseki (2015), who proposed that adjunction gives rise to two-peaked structures.⁴¹ In his proposal, a sister of adjuncts is

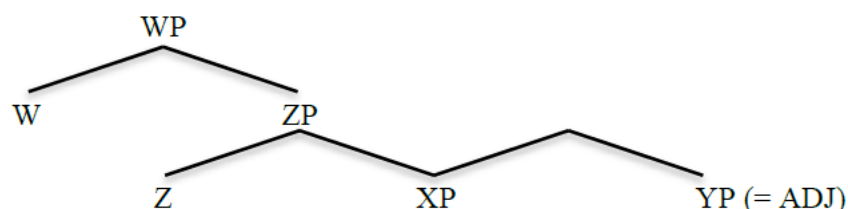
the absence of labelling of adjuncts, which creates problems with the interpretation of these structures. The same can be said for proposals where adjuncts are treated as doubly-rooted structures (see Oseki 2015).

⁴⁰ For some more issues of Narita’s (2011) approach see Müller (2019: 136).

⁴¹ Similar to the proposal of Johnson (2009).

dominated by two different mothers and because of this special configurations adjuncts seem to be on a different plane (Chomsky 2004). These kinds of two-peaked structures are naturally derived from structure-building of adjunction, namely {XP, YP}. See structure in (74) for an example (adapted from Oseki 2015: 307).

(74)

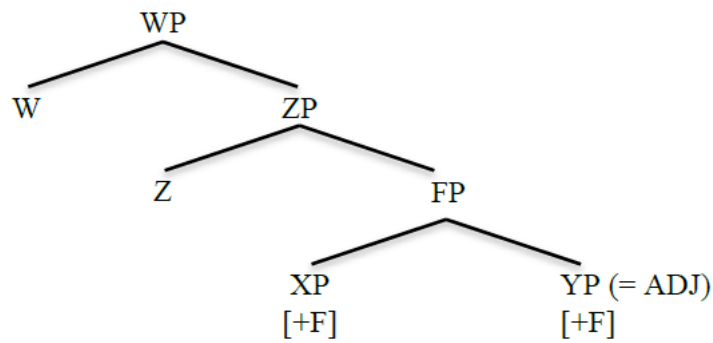


The main issue with a structure like (74) is that there is an unlabeled syntactic object, which, following the traditional analysis (Chomsky 2000; Hornstein 2009) is not accessible to Merge.⁴² In fact, following the Labeling Algorithm (Chomsky 2013),⁴³ merging two phrases as {XP, YP} creates an ambiguous relation and no label can be assigned. Oseki follows Epstein, Kitahara & Seely (2012) and claims that because of the absence of labeling the structure needs to undergo Spell-Out, meaning that the adjunct becomes an opaque domain, and cannot participate further to syntax. However, there are adjuncts that are visible to syntax: those that enter a feature-sharing relation. According to Feature-Sharing, it is possible to label a {XP, YP} structure if X and Y share the same feature [+F]. In that case, if label is assured, the resulting structure will be a normal one, not a two-peaked as in (75) (adapted from Oseki 2015: 309).

⁴² The issue can also be seen under a different light: rather than because of the absence of labeling, Spell-Out could also be induced because of a label that is too complex, when more than one element project. If this is the case, the need is that of reducing such label to a unique element. See the analysis of small clauses of Moro (2000), or Chomsky (2004) for adjunction as Pair-Merge.

⁴³ When merging one head and one phrase {X, YP} LA unambiguously selects the closest head as the label for the entire syntactic object, since minimal search locates the unique head X. However, when dealing with merging of two heads {X, Y} or two phrases {XP, YP} things are not so simple: heads are equidistant and minimal search fails to find a unique head which can assign the label to the entire syntactic object.

(75)



Thus, in Oseki (2015) the fundamental issue is that adjuncts enter an Agree relation, which means that they will be visible to syntax and syntactic processes. Agree is what establishes Feature Sharing between the two phrases, and the agreeing feature is what assigns the label (Chomsky 2013). This is equal to say that transparent adjuncts have an Agree or Feature Sharing relation with something like v or Aspect of the main clause (though this is probably not visible in the morphology). However, it is not clear which kind of feature would be shared by adjunct and matrix clause.

5.4.2 *The proposal*

In the proposal I am going to develop two crucial points have to be addressed: (i) extraction from adjuncts is sometimes licensed, but its effect is different depending on the adjunct taken into account; (ii) adjuncts behave differently and cannot be considered as a uniform class. Let me now take a step back, and explain which analyses are able to explain the two points above. As we know, the potential problems with islands can either be connected to syntax in its narrow sense, or to the external systems interfacing with syntax, i.e. the phonological component or the semantic one. In PF, the problem regards the linearization of the copies, whereas in LF we could have issues with the interpretation of copies (with reconstruction). There are many theories blaming islandhood on such problems with the interfaces; in these cases the island effect is not due to movement itself, but rather to a filter of one of the external systems interacting with syntax and judging such movement as illicit. In case of PF, the problem is mostly related to the presence of a gap in an island domain (see Merchant 2001; Lasnik 2001; Hornstein, Lasnik & Uriagereka 2007 among others). Such a view is connected to some mechanisms of island

circumvention: resumption and ellipsis, where either the entire island is not realised (ellipsis), or there is an extra element that is able to rescue the island (resumption). However, we saw that this is not really the case: the presence of a resumptive pronoun does not rescue the island in Italian⁴⁴ – or it has different degrees of rescuing effect depending on the adjunct. As shown in Experiment A, in some cases the effect obtained is the other way round: the presence of a resumptive pronoun is significantly worse than the gap in the case of *without*. Overall, a PF treatment in the case of adjuncts does not seem to be able to account for either of the two points above. There is no reason to assume that gaps are present only with a subset of adjuncts, whereas with others there is a resumptive element. Notice that this argument also goes against a theory involving the presence of an empty resumptive *pro* à la Cinque (1990), which is not able to account for the differences among adjuncts, either. The problem is not the presence/absence of a gap itself, but rather the way this is connected to the rest of the structure: it is a matter of syntax.⁴⁵

However, given the two points above, the theory accounting for the (im)possibility of extraction cannot be too rigid: it cannot be applied *only* to the core property of language, otherwise we would not be able to explain why extraction is sometimes licensed – unless we postulate the existence of an escape hatch, and we individuate the reasons why such an escape hatch is valid only insofar some adjuncts are considered. Moreover, the issue with escape hatches is that most times they are created ad hoc in order to justify movement from one phase to the other, and, once they are introduced, the problem is to *restrict* movement (see Boeckx 2012; Tanaka 2015). The ideal theory has to be flexible enough in order to account for both variation among adjuncts and variation among speakers – which will be investigated in the next chapter. Clearly, it can't be based only on syntax in its narrow sense.

I claim that the best way to account for the dual behaviour of adjuncts is to propose that they are syntactically different. The best account explains why adjuncts seem to be simultaneously inside and outside the workspace of the main sentence. If they are outside no movement is possible, whereas if they are inside extraction can be allowed. There are several hypotheses able to explain why they are outside: Late Merge (Lebeaux 1988; Stepanov 2001, 2007) states that the timing of merger of adjuncts is different than the one of the main spine,

⁴⁴ Nor in other languages with resumptive pronouns of the intrusive kind (as in the case of English: Ferreira & Swets 2005; Heestand et al. 2012; Polinsky et al. 2013; Beltrama & Xiang 2016) and it seems to be the case also of languages with grammatical RPs, once this effect is formally investigated (Keshev 2016; Keshev & Meltzer 2018; Tucker et al. 2019).

⁴⁵ See also Boeckx (2012) for some theoretical issues of a theory involving the PF side.

because they are inserted late, being the last part that gets attached to the structure. In a Pair-Merge analysis (Chomsky 2004, 2008), they occupy a different plane than the main sentence, whereas in Multiple Spell-Out (Uriagereka 1999; Nunes & Uriagereka 2000) adjuncts are forced to an early linearization and are consequently opacified. Finally, in some analyses (Nunes & Hornstein 2008; Hornstein 2009) the issue depends on the fact that adjuncts are merged to the main structure without a label, which means that they cannot be probed, whereas in others the issue depends on the absence of an Agree relationship between the matrix and the adjunct (Rackowski & Richard 2005). It is difficult to adopt one of these approaches since, as explained in the previous sections, all of them are too strong. In any case, in order to adopt one of these analyses it would be necessary that they explain the data regarding the dual behaviour of adjuncts *and* the differences among them, which they do not. Moreover, notice also that most of the theories above are not only inadequate on empirical bases, but also on theoretical ones. For instance, assuming the possibility of Late Merge is potentially very problematic: once this is allowed, it ends up increasing non-trivially the generative capacity of grammars (see Sportiche 2018 for an analysis). On the other hand, Pair-Merge is an ad hoc mechanism: it is construction-specific because it exists only for adjunction.

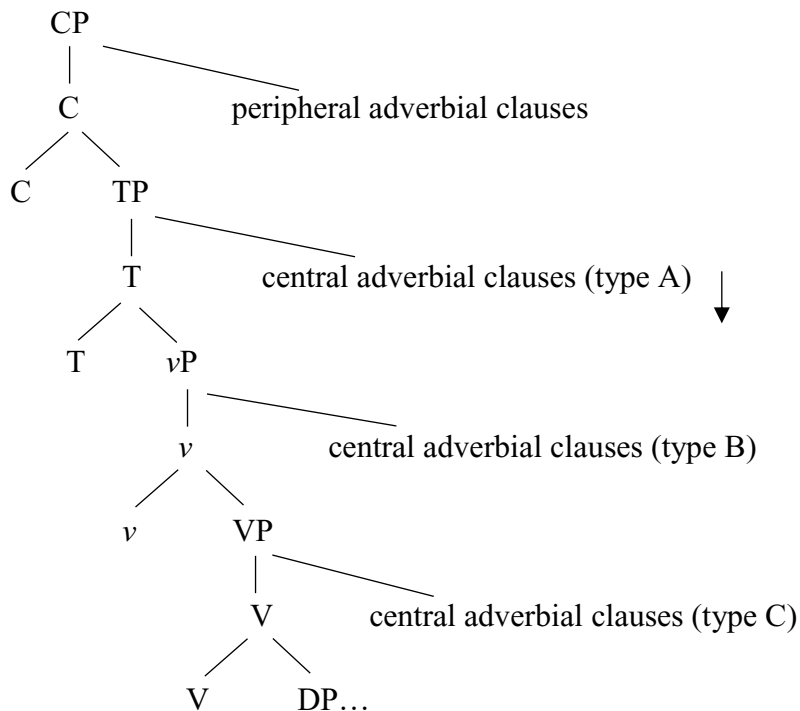
From the results of the test discussed in §4 and the analysis explained in the previous sections of §5, I claim that at least two factors play a crucial role in the licensing of extraction: (i) the type of adjunct and (ii) the possibility of forming a macroevent made by the matrix verb and the one in the adjunct. The latter is the semantic condition proposed by Truswell (2007, 2011). However, I have shown that a semantic condition by itself is not able to explain all the data and I have claimed that syntax has a role as well. In particular, I propose that it is a syntactic condition that allows or disallows for a macroevent reading; factor (ii) depends on factor (i). The formation of a single macroevent is allowed only with adjuncts low in the structure, whereas it is not an option if the adjunct is attached high. In the latter case, extraction is not licensed, whereas in the first it can be. Let me take a step back and describe the two options with more details.

As discussed in §5.2, adjuncts are not a uniform class and I showed that the best way to account for their difference is to assume that there is a difference in their representational format, specifically in their attachment height in the syntactic spine. Some are attached to the *vP* level, and hence are high in the structure, and some are attached lower, to *VP*.⁴⁶ Notice that

⁴⁶ A difference in their levels of attachment successfully explains also the different behaviours of central and peripheral adverbial clauses, as shown in Haegeman (2004, 2012).

there is no special format required here in order to account for the properties of adjuncts and no ad hoc stipulations are required: adjuncts are always merged in the same way, but to different points in the structure.⁴⁷ See the structure in (76):

(76)



As can be seen, there are different levels available for adverbial clauses. First, the distinction is based on peripheral and central adverbial clauses. The former are external to the TP, as shown by Haegeman (2012) and are too external to the sentence to allow extraction. This is the case of all adverbial clauses with a peripheral reading only – as in the case of *although* – but also of adverbial clauses with a double reading when the peripheral one is used, which would be the case of some instances of *if* and *because*. From the TP below, the syntactic spine is dedicated to different types of adverbial clauses of the central type, indicated here as Type A, B and C for the time being. This analysis will be expanded later on.

The possibility of forming a macroevent made by the (sub)event of the matrix and the (sub)event of the adjunct is deeply linked to the height of the adjunct: it is not only a semantic

⁴⁷ A difference in the attachment site was already hinted in Chomsky (1986) within the Barriers system. In particular, the system he developed predicted a difference in the acceptability of extraction depending on the merging point: adjuncts attached to IP cross two barriers, whereas adjuncts to VP cross only one barrier. The former situation leads to unacceptability, whereas the latter to a mild degradation.

condition as expressed in Truswell (2007, 2011), but it first depends on a syntactic one. When the adjunct is high, no single event reading can be created, and hence extraction is not allowed. On the other hand, with low adjuncts the possibility of forming a macroevent is open. Notice that it is only a *possibility*, not an obligation: single eventhood is only one of the possible readings available, but by no means it is the only one. This depends (i) on the types of verbs, as Truswell showed on the basis of Vendler (1957) and Dowty (1979), (ii) on the lexical choices of the sentence, and (iii) on the type of relation that is created between the matrix and the adjunct (contingent versus non-contingent).⁴⁸ Therefore, when the adjunct is low a macroevent can be formed, in which case extraction is licensed. But when the adjunct is low there is another option: the macroevent can not be formed, for (at least) one of the reasons above. If this happens, extraction is not licensed. To sum up, extraction is never licensed from higher adjuncts, it can be licensed in lower ones – but it does not have to be: it depends on the macroevent reading. Therefore, the single event reading creates a sort of bridge between the adjunct and the main clause, allowing extraction to take place. But such a bridge is an option only insofar the adjunct is in an expected structural position, that is to say when it is low enough.⁴⁹

The single event reading is connected to the attachment site of the adjunct – among other things. But there are a few more requirements that have to be respected in order to allow for a macroevent reading. For instance, the adjunct does not have to create its own event substructure in order for the macroevent to take place. The kind of complementizer matters also insofar the formation of the macroevent is considered, not only because of its attachment site, but also because of its semantic properties, i.e. its interpretation. It is not the case that an element heading an adjunct – whatever complementizer we are dealing with – can be categorized as a bounding node inherently, as was the case in Chomsky (1986), where all PPs were considered as inherent barriers. At the same time, different points of merger are not established only by means of the lexical choice that is made for the introducing conjunct. My claim is therefore that the attachment sites are affected by the interpretation. In this sense, the semantic interpretation is the result of a syntactic position in the clause (see also Brown 2017). Remember at this point that some conjuncts can have different interpretations, depending on (i) their peripheral or

⁴⁸ There is a forth influencing factor in the construction of a macroevent, and this depends on the speaker. I will discuss about it in chapter 6.

⁴⁹ Ernst (2002) also talks about a different attachment site for adjuncts. He specifies that event internal modifications are restricted to lower parts of the clause, which goes hand in hand with saying that adjuncts that convey a coherent single-event relation to the matrix are attached in a lower position.

central reading (Haegeman 2012), (ii) their placement in the structure, as shown by (77), and (iii) potentially their point of merger.

- (77) a. Per questa ragione, hanno arrestato Gianni.
'Because of this reason, they arrested John.'
- b. Hanno arrestato Gianni per questa ragione.
'They have arrested John for this reason.'
- (Cinque 1990: 66)

Notice that such a correlation between the interpretation of a conjunct and the position of the adverbial clause may be extended to other types of adjuncts as well. It is the case of some adjuncts *wh*-phrases such as *how* and *why*. These are linked to different interpretations. *How* can either be connected to a manner/instrumental interpretation or to a causal one. The former is related to the event, whereas the latter is fact related, i.e. it is linked to the whole proposition rather than only to the predicate. Interestingly, the difference in the interpretation is reflected in the possibility of moving the *wh*-constituent, as can be seen by the asymmetry between (78a) and (78b):

- (78) a. How can this cost so much now? It was still affordable yesterday!
- b. *How do you think that this can cost so much now? It was still affordable yesterday!
- (Starke 2001: 29)

There is a structural difference among the two, and this can be seen also in the case of *because*, which can introduce either cause or motivation. In the latter case it is event related, i.e. it is linked to the verb. Consider sentence presented in (79a), which can either be interpreted as (79b) or as (79c).

- (79) a. Brutus didn't go to Rome [because his wheel broke].
- b. Brutus didn't go to Rome and the reason is that his wheel broke.
- c. Brutus did go to Rome and the reason is unrelated to the fact that his wheel broke.
- (Starke 2001: 30)

Notice that it is not only a matter of scopal ambiguity between the negation and the adjunct. In fact, whereas in (79c) *because* can only be interpreted as motivation, under reading in (79b)

there is still an ambiguity in its interpretation, in that it can either be motivation or cause. We may argue that it is not possible that a conjunct able to introduce different readings – not only with respect to the central or peripheral ones – is connected to the same merging points, since when it does not share the same function. Pushing the claim of Huddleston & Pullum (2006), I argue that adverbial clauses introduced by the same conjunct cannot be merged at the same level in the structure if they do not have the same function. Again, notice that the interpretation that is being assigned to the adjunct influences the acceptability of extraction, as was demonstrated by experimental evidence in Tanaka (2015) and Müller (2019), who shows that extraction from *after* is not possible when it is connected to a purely temporal interpretation, but it is licensed when it is enriched with a causal reading. In the analysis proposed here, therefore, the difference in the interpretation of an adjunct and the following difference in the acceptability of extraction depends on the attachment site of the adjunct, and the consequent possibility of forming a macroevent.

I will now expand the analysis proposed showing how all the patterns obtained in the experiments explained in chapter 4 can be derived following the analysis proposed here. The attested patterns I will discuss are not only those coming from my experiments, but also from the other formal experiment which focused on different types of adjuncts – and which were reviewed in §5.1.2. Notice that I will consider only those cases in which different adjuncts were investigated and which were analysed separately from one another (see Müller 2017, 2019; Bondevik 2018). However, many more formally collected judgements would be required in order to expand the analysis to different languages and involving more adverbial clauses than those presented here.

First, I focus on the case of *without*. As we saw, in Italian extraction from *without* is the most accepted, not only in islands but also in parasitic gaps. Following the analysis proposed here, it means that *without* is attached low in the structure, namely to VP. This leaves open the possibility of forming a macroevent, and when this happens extraction is licensed, as shown in (80a).

- (80) a. Quale ragazzo Maria se ne è andata senza salutare _ ?
 which boy Maria went away without greeting
 ‘Which boy did Mary go away without greeting _?’

- b. *Quale ragazza Gianni ha raggiunto la vetta [senza salutare _]?
 which girl Gianni has reached the top without greeting
 ‘Which girl did John reach the top without greeting?’

Notice once again that there are some cases in which the macroevent is not formed, that is to say when (i) lexical choices are not appropriate, as in (80b); and (ii) the relation between matrix and adjunct is not contingent. In these cases, since the macroevent cannot be formed, no extraction can take place, despite the fact that the adjunct is merged low. Sentence in (80b) is clearly unacceptable. Following my analysis, there is something syntactically wrong about it – it is not only a matter of semantics and macroevent formation. I will come back at this example and its unacceptability at the end of this section. Notice moreover that there is a third option, namely the fact that in some cases the macroevent can be formed (at least following Truswell’s (2007, 2011) analysis and the points considered above), but nonetheless extraction is not acceptable. I will dwell some more on such a possibility in the next chapter, where individual variation will be considered.

Let’s have a closer look at *without* and its properties. Notice that adjuncts usually provide additional information regarding the circumstances under which the event described in the matrix sentence took place. In some cases, though, it seems that they provide the same information of the matrix: it is not additional information, it is a specification of the event itself. This is what happens with *without*: it intervenes on the eventive structure of the matrix sentence, modifying the subevents of the matrix predicate. Notice that it actually has a role in the partition of the event, in the sense that it describes a particular type of event. Its role is to subtract the subevents of the predicate it introduces from the ones described in the matrix clause. In (81), for instance, it is stating that among all the events of leaving of the world, the ones relevant for here are the leaving-without-greeting.

- (81) a. Quale ragazzo sei partita senza salutare?
 ‘Which boy did you leave without greeting?’
 b. partire senza salutare
 c. <e,e,e ... e,e,e>

The property of *without* in a sentence such as (81) is that of telling us that there is an intersection between the events of leaving and those of not-greeting, and that such an intersection is not

empty. In (81) even though the syntax wants to close the unit, this is not possible: before closing the domain all the information is needed and the event here is not complete. (81a) is analysed as (81b) and (81c), in a single event.

I will turn to temporal adverbial clauses. The ones I investigated were headed by *after* and *before* and the experiments showed that they do not allow for extraction. Ideally, hence, there are different options: (i) either they are higher than *without* and extraction is not licensed because a structural requirement is not respected, (ii) it is more difficult to form a macroevent, and therefore extraction is not licensed even though they are low, or (iii) their internal structure is too complex. Let me go through these possibilities. First, it is difficult to see a difference between *without* and *after/before* in terms of structural position. This is so because tests involving c-command – such as violations of Condition C effects, as seen in (82) – do not give us different results among the two, and it is therefore difficult to establish that they really occupy two different positions.

(82) pro^*_i si è sentita male dopo che/senza che Lisa_i aveva/avesse mangiato quel piatto.

She is felt sick after/without Lisa had eaten/having eaten that dish.

On the other hand, considering the hierarchy presented in (48) and repeated here as (83), it is clear that there is a restriction in their ordering, in that *without* can only precede temporal adverbial clauses.

(83) per > senza > prima/dopo > finché > perché
in order to > without > before/after > until > because

Second, it may be that the macroevent is more difficult to form. First, *after* and *before* naturally exclude one of the conditions presented by Truswell (2011), namely that the two events overlap spatiotemporally. However, in some cases the temporal interpretation can be enriched with a causal reading, i.e. the relation between the matrix and the adjunct can be a contingent one, as in the case of sentences in (84).

(84) a. Quali dolci Luca è scappato dopo aver rubato?
'Which sweets did Luca run away after stealing?'

- b. Quale ragazza Luca è scappato dopo aver investito?
 ‘Which girls did Luca run away after running over?’

In (84) both adjuncts can express a causal link between the matrix and adverbial. In (84a) the event of running away is caused by the event of stealing: $e_{\text{adjunct}} \text{ cause } e_{\text{matrix}}$. The contingent relation expressed here – which is predicted to favour extraction in Truswell (2011) – seems not to be at work in Italian: extraction from *after* leads to unacceptability, and this cannot be rescued by the causal link between the two: the semantic relation is not enough. This indicates that there is a role of syntax. On the other hand, remember that in the case of English (Tanaka 2015) and Swedish (Müller 2019), the additional causal reading also allowed for extraction. Interestingly, in the Swedish case this was true only for the adjunct introduced by *after*, but not in the case of *when* which always lead to degradation.⁵⁰

Third, it could be that a richer internal structure does not allow for extraction and/or the macroevent formation. Notice that different analyses involving a complex internal structure for temporal adverbial clauses have been proposed. In fact, their role is to provide additional information regarding the circumstances under which the event described in the matrix sentence took place. Specifically, the role of temporal adjuncts is that of describing a relation between events, offering a comparison between the event time and the reference time, which must be different. For such reason an analysis like the one proposed by Brugè & Suñer (2008) hypothesizes the presence of an additional head inside these adjunct, a comparative head. In other approaches, the presence of operators in temporal adverbial clauses have been proposed (see Haegeman 2012 for an overview).⁵¹ All of these factors could explain the unacceptability of extraction showed in the experiments presented.

I will now talk about the case of *in order to*. Extraction from this adjunct in Italian always lead to a severe degradation, in both islands and parasitic gaps. This was unexpected, not only because it expresses a contingent relation between matrix and adjunct, but also because

⁵⁰ In Norwegian, on the other hand, Bondevik (2018) showed that extraction from *when* is sometimes allowed.

⁵¹ An additional explanation can be connected to the theory that CPs cannot be the complements of a preposition, for a restriction of the type in (i) (see Stowell 1981; Alrenga 2005):

(i) * $[P \text{ CP}]$

According to some scholars, it is necessary that the preposition actually select a nominal complement. This led to the hypothesis that they are actually selecting a silent head of the types discussed in Kayne (2005). As an example, temporal connectives such as *after*, *before* and *until* select a silent TIME (see Haegeman 2012: 204).

several informally collected judgments from English present this adverbial clause and extraction is always acceptable. Swedish also behaves differently, and extraction is mostly accepted (see Müller 2019). In my analysis, this would mean that *in order to* is attached too high – at least in Italian. However, the hierarchy tells us something different: *in order to* is the first adverbial clause in the hierarchy, meaning that it is actually very low in the structure. Data from object control confirm such an account: as shown in §5.2.3, usually object control is unacceptable inside adjuncts. However, *in order to* behaves differently and it is the only one that allows for such reading, as shown in (85) for both Italian and English.

- (85) a. Sara_i ha punito Gianni_k [per PRO*_{i/k} aver mentito].
 b. Sara_i punished John_k [for PRO*_{i/k} lying].
 c. They_i punished John_k [for PRO*_{i/k} driving drunk]. (Hornstein 1999: 76 fn 10)

This shows that *in order to* is low enough to be c-commanded by the object, therefore possibly below VP. My claim is that it is attached so low, that the macroevent cannot be formed and no extraction is licensed in Italian. Possibly, the same explanation could be extended to parasitic gaps: these structures are not licensed with *in order to* exactly because of its too-low attachment.

The last two adjuncts I am going to talk about from my experiments are *because* and *until*. In the latter, extraction is severely degraded from both islands and parasitic gaps. This means that its point of merger is too high to allow for extraction: if they are connected to vP, following my analysis it is not possible to form a macroevent and hence no extraction is licensed. Moreover, notice that this adjunct – which has been formally investigated only in Italian – has additional difficulties such as the mandatory presence of a negative word, as well as the finite verb in the adjunct. Both these aspects can be said to be problematic because they add internal structure to the adjunct, hence making extraction more difficult, so much so that it is not allowed even in parasitic gaps.

Finally, the case of *because* was a bit peculiar: extraction was bad, but not as bad as expected. Notice also that *because*-adjuncts were investigated also for Swedish and Norwegian, and in both languages extraction was never an option. This is a puzzling result, since *because* is an adverbial clauses of the central type and it expresses a contingent relation of causation, therefore predicting that extraction is allowed following an analysis à la Truswell (2007, 2011). This shows that the attachment height really has a role in the formation of a macroevent. In fact, we know that it is attached high in the structure – higher than other adjuncts (see Uriagereka

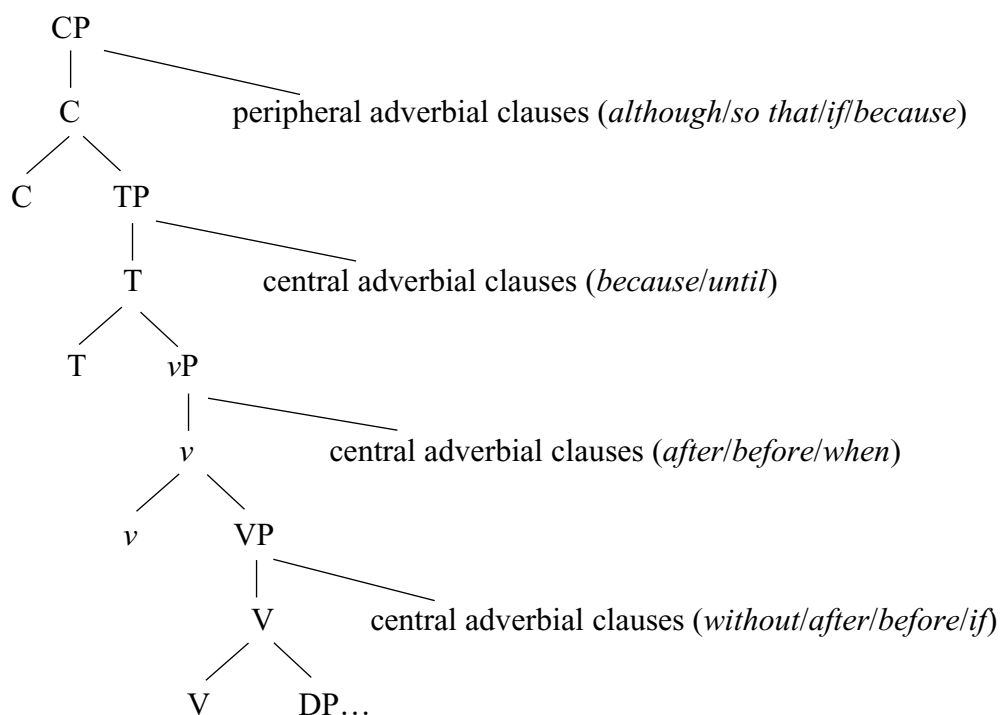
2011; Boeckx 2012). In particular, we know that it has to be higher than direct and indirect object (see 86a) – differently from other adjuncts in (86b) –, but still lower than the subject as shown by violations of Condition C effect in (86c).

- (86) a. *They accepted every paper_i for publication *because* it_i had been positively evaluated by two reviewers.
 b. They will accept every paper_i for publication *if/after* it_i is positively evaluated by two reviewers. (Valmala 2009: 961)
 c. pro*_i si è sentita male perché Lisa_i aveva mangiato quel piatto.
 She is felt sick because Lisa had eaten that dish.

This shows that *because* is attached to *vP* – or even higher. Moreover, in Italian, *because* requires the tense of the adjunct to be realised in its finite form, therefore adding an additional difficulty for extraction. Notice also that it has been claimed that a conjunct like *because* applies to a larger semantic domain and combines with a larger syntactic domain (see Sawada & Larson 2004), which means that it has extra layers of structure.

The discussion of the patterns leads us to a structure of the kind presented in (87) – including adverbial sentences used for other experiments described in §5.1.2.

(87)

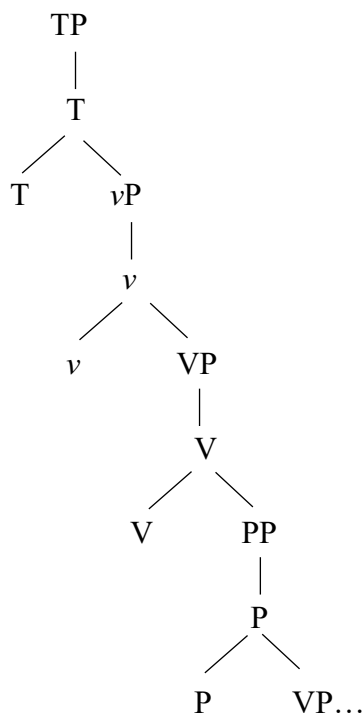


I will now go back to the issue of the possibility of extraction, and specifically to sentence in (88), where extraction from *without* is a possible option in Italian. Following the analysis just proposed, this is possible because this adjunct is attached lower in the structure, namely in VP. This means that the macroevent can be constructed and therefore that extraction is licit. How is it so, however?

- (88) Quale ragazzo sei partita senza salutare?
 ‘Which boy did you leave without greeting?’

Syntactically speaking, (88) could be analysed with a structure similar to (89) – as well as all the other transparent adjuncts, i.e. those that are low and allow for the creation of a macroevent. In this case, it is as if they are selected by the verb, and their position is actually in the complement of V, following a Larsonian analysis.⁵²

- (89)



⁵² See Chesi (2007) for an analysis of this kind based on a top-down approach.

My claim is that the construction of a macroevent leads to a reanalysis of the adjunct as if it was a complement. This means that extraction in these cases is naturally allowed, given what we know about complements. Not always is a reanalysis possible: it is actually subject to several restrictions, some of which were already spelled-out, namely (i) the height of the adjunct, and (ii) the construction of a macroevent. But there are more aspects that should be kept into account, such as (i) the internal structure of the adjunct – which should not be too rich; (ii) the matrix verb used; and (iii) individual differences. I will consider each of the points above in the following sections, except for (iii), which will be investigated in the next chapter. Notice that all of them can also be considered as conditions for a macroevent analysis.

As for the internal structure of adjuncts, I will focus first on the role of finite tense, to show how this can interact with extraction. Its effect was presented in §3.2.1 and experimentally examined in Experiment B. I showed that this factor has a significant impact on the acceptability of a sentence: when the adjunct has finite tense, extraction is degraded. There are several explanations that could work here. First, we might assume that there is something special about the head T or its maximal projection TP, which somehow is able to interact with adjunction and thus it induces the island effect.⁵³ This is the proposal of Cinque (1990), who assumes that a tensed inflection has the effect of blocking the upward percolation of the *wh*-feature. However, if we adopt such an approach we cannot explain why tense interferes with adjuncts but not with complements, since extraction from a tensed complement is completely acceptable. Second, we might follow Truswell (2011) and state that tensed adjuncts contain their own event operator that binds the event variable of the adjuncts and hence it blocks the construction of a macroevent. Once again, this view would require an explanation of why such an operator is at work only with adjuncts, or why it is not present in complements. Third, we could adopt Chomsky's (2004, 2008) analysis and hypothesize that a finite sentence should be considered a CP, whereas if T is defective it should be analysed as a *v*P. I will consider the last approaches and claim that it is actually connected to the internal syntax of adjuncts. In order for the macroevent to be created, their internal structure should not be too complex: the less structure they contain, the better the extraction. Notice that both the presence of an operator and the analysis of the tensed adjunct as a CP would explain this. A richer internal structure can block extraction, in the sense that it blocks the creation of a single event reading. This might be

⁵³ Notice that TP has a long story of special syntactic treatment; for instance, in Chomsky (1981) the Extended Projection Principle is applied only to SpecTP, whereas in Chomsky (1986) TP was considered as an exception in the establishment of barriers to movement.

connected to different factors such as the presence of finite tense, or the presence of additional operators.^{54,55} Notice that the external syntax of adverbial clauses is shown to influence the internal syntax as well (Haegeman 2004, 2012; Endo & Haegeman 2019): the more structure it contains, the higher it is merged.

With respect to the role of verbs, I present once again the asymmetry in (90):

- (90) a. Quale ragazzo Maria se ne è andata senza salutare _ ?
 which boy Maria went away without greeting
 ‘Which boy did Mary go away without greeting _ ?’
- b. *Quale ragazza Gianni ha raggiunto la vetta [senza salutare _]?
 which girl Gianni has reached the top without greeting
 ‘Which girl did John reach the top without greeting?’

The difference between the two sentences relies on the type of verb used in the matrix. Such a difference is already known since Truswell (2011) and Fábregas & Jiménez-Fernández (2016a). The effect could be related to the impossibility of respecting the structure in (89). In the case of an achievement such as the one in (90a), we could state that the macroevent is already formed, and is made by the matrix verb and the internal object (which does not really have such role anymore).⁵⁶ Crucially, notice that if the matrix verb already contains an object, the reanalysis of the adjunct into a complement is not possible anymore and hence extraction is blocked. Therefore, if an internal argument is already present in the structure, there will be no space for the adjunct to be reanalysed into a complement, ruling out extraction in these cases and explaining the asymmetry between (90a) and (90b).⁵⁷ In fact, the adverbial clause will be forced to remain in its position in the specifier of VP, as in (91).

⁵⁴ However, notice that there are some languages that seems to allow extraction from finite adjuncts (see Biskup & Šimík 2018; Müller 2019)

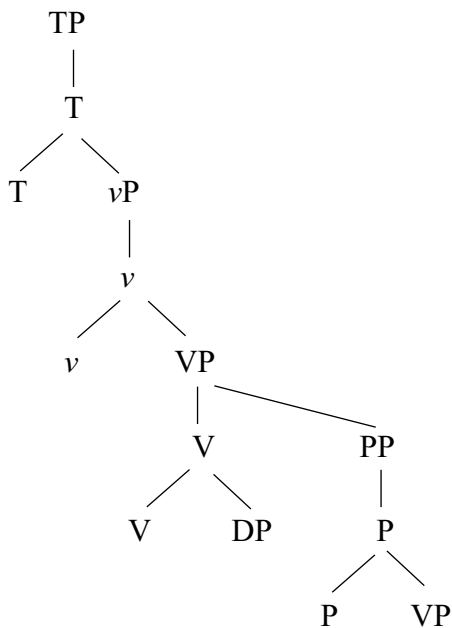
⁵⁵ To such point, notice also that several analysis of temporal adverbial clauses propose that some kind of operators are present (see Brugè & Suñer 2008; Haegeman 2012 for an overview).

⁵⁶ As for accomplishments, notice that extraction is sometimes licensed even in this case:

(i) Who did John go home after talking to? (Truswell 2007: 5)

⁵⁷ See Fábregas & Jiménez-Fernández (2016a) for a similar analysis of gerunds. Notice however that in the analysis presented here, (89) is not the only option: several restrictions have to be respected for extraction to be allowed, and, most importantly, the reanalysis of adjuncts into complements is not obligatory. This means that there can be cases in which extraction is not licensed.

(91)



In Italian, a reanalysis of the kind in (91) is obtained only with *without* – at least following the results of the experiments (even though it is subject to individual differences). But we have seen that in other languages other types of adjuncts can be transparent. This is the case of *if* in Norwegian, and of *after* with a contingent relation or *in order to* in Swedish. The latter is transparent also in English – at least from what we know from informal judgments. My claim is therefore that these kinds of adjuncts are all low enough in the structure to create a macroevent and to undergo a reanalysis of their role in the structure: they become complements. Basically, following structure in (87), we could state that the lower the adjunct, the easier it is for it to be analysed as a complement. Notice that such a possibility is open only for those attached to VP. In fact, higher adjuncts merged to vP or above do not allow for such a reanalysis: they cannot be a part of the VP-shell, but are simply attached to it. Such a position naturally explains that a macroevent can not be formed under these circumstances.

Interestingly, there are some sentences such as the ones in (92) which are technically extraction from adjuncts, but which are completely acceptable in both Italian and English.

- (92) a Chi sei andato a trovare?
 who are you gone to visit
- b. ‘Who did you get there to visit?’

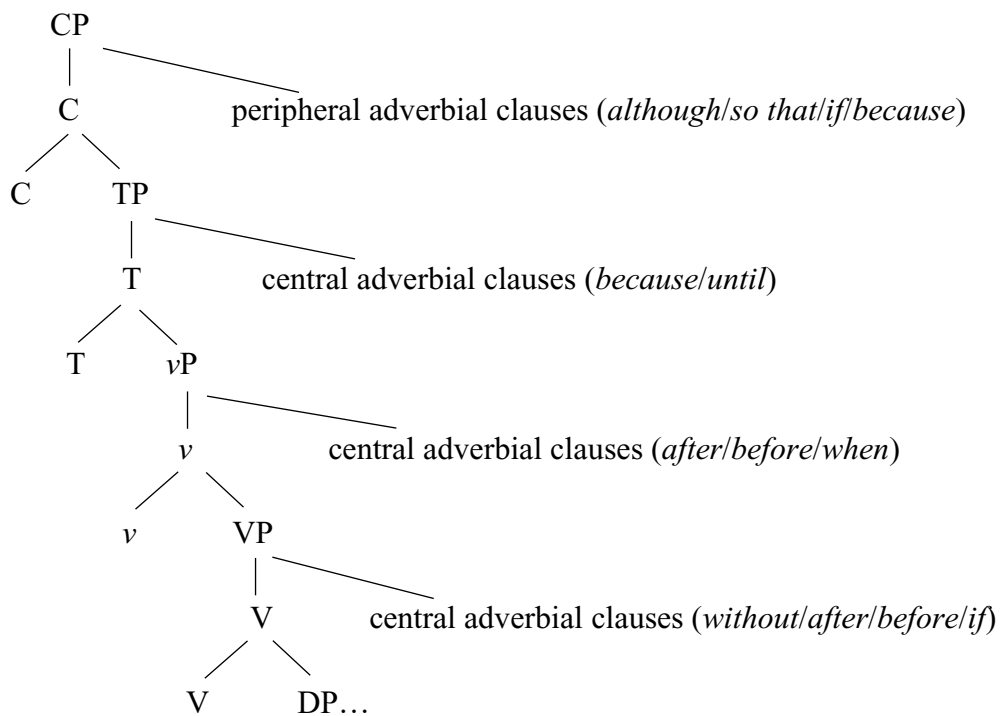
- c. Cosa è andato a New York a comprare?
‘What did John go to New York to buy?’
- d. *Che cosa è andato a New York [per comprare _] Gianni?

Sentences like (92) represent a perfect case of extraction, even though the *wh*-constituent used is a simple one for both English and Italian. A semantically light verb like *to go* favours a complement reanalysis of the kind in (89), and hence also the formation of the macroevent. I claim that some adjuncts – as *a* in Italian and (*in order*) *to* in English – are born as adjuncts but tend to become complement – and they are always analysed in such way. This is the case of (92a), where extraction is always acceptable. Notice that if we change the introducing conjunct, as in (92b) extraction is not accepted; probably because the macroevent is more difficult to construct in such a case.⁵⁸

At this point, how do we treat parasitic gaps, though? Remember that these structures are usually considered grammatical, but my experiments showed that things are not so simple: there is a difference among adjuncts also in this case, in that there is a different degree of acceptability which tends to respect the one identified for islands. Why should there be a difference? In the usual treatment of parasitic gaps, the gap inside the island is licensed thanks to the presence of a higher c-commanding gap – actually thanks to the presence of movement. I am going to pursue the analysis of adjuncts attached to different merging points: their level of attachment is what explains the difference in the acceptability in structures that are usually considered grammatical. In fact, if we go back and consider the results from the experiments, it will be clear that the ordering of adjuncts proposed in (87) and repeated here in (93) is respected.

⁵⁸ We could also blame the unacceptability entirely to *in order to*: remember that I proposed that in Italian this is attached too low to allow for extraction.

(93)



Let me briefly explain why. First, parasitic gaps are not acceptable if the adverbial clause is of the peripheral type (see Haegeman 2012), i.e. if the adjunct is attached as high as CP, extraction is severely degraded. Second, parasitic gap licensing with *until* is very degraded. This follows if this adjunct is attached high.⁵⁹ The same goes for adjuncts attached to vP – I am referring in particular to the case of *after*. On the other hand, parasitic gaps were considered better with adjuncts placed to VP – in the experiments presented, this was the case of *before* and especially *without*. Notice that *in order to* was left aside, and remember that parasitic gaps formed with such adjuncts were really degraded. I will follow the analysis presented above, and claim that it is so because this adjunct is actually attached too low in the structure.

I want to stress a second condition that is probably at work with parasitic gaps as well. It may very well be the case that it is not (only) movement that licenses the parasitic gap, but that is rather the role of the macroevent that can be created in these structures as well – provided that structural conditions just described are respected. In fact, notice that extraction from parasitic gaps has been shown to be more acceptable when the matrix and the adverbial clauses express a close relation in English (Wagers & Phillip 2009). Moreover, notice also that parasitic gaps are not acceptable with peripheral adverbial clauses, which could be explained because

⁵⁹ I will leave aside the case of *because*: the parasitic gap sentences were not unacceptable, but not perfect either.

the macroevent reading is not licensed in these cases. Therefore, with central adjuncts the acceptability of parasitic gaps has different degrees depending on the type of adjunct, i.e. if it is low or high, and again on the formation of a macroevent.

At this point, notice that a structure of the type in (89) is not able to accommodate the parasitic gap case: if the adjunct is being reanalysed as a complement, a crucial issue relies on the fact that the verb does not have an internal argument, either overt or covert. In the case of parasitic gaps, though, there is – there has to be – an internal argument shared by matrix and adjunct. This means that (89) is not the right analysis for these cases. Supposedly, in parasitic gaps the adjunct remains in its ‘usual’ position, that is to it is attached to VP and is not part of the VP-shell. The fact that parasitic gaps are nonetheless more acceptable than islands may be connected to θ -roles: since matrix and adjunct share not only the external argument but also the internal one, the macroevent reading could be facilitated (see the analysis of sideward movement proposed in Nunes (1995, 2001)).

The importance of the macroevent could be extended also to exceptions to the Coordinate Structure Constraint, as in (94):

- (94) a. What did you go to the store and buy _ ?
b. How much can Josh drink _ and still stay sober?
(Wagers & Phillips 2009: 400 fn. 6)
c. Here’s the whiskey which I went to the store and bought. (Ross 1967: 168)

These sentences are often explained with narrative relationships holding between the coordinates under specific circumstances. For instance, Wagers & Phillips (2009) assume that these might arise in specific circumstances expressing purpose, outcome and temporal contiguity, i.e. cases where the formation of a macroevent is facilitated.

Last, let me consider why a macroevent can be assigned such a crucial role. I think that the formation of a macroevent could have crucial consequences for two theories mentioned before, that is to say the assignment of a label and the role of Agree. I will briefly explain why. When the adjunct is high, and a single event reading is therefore not available, the label could be either too complex (see also Chomsky 2004) or absent (Nunes & Hornstein 2008). If the macroevent reading is not available, the label cannot be assigned and the adjunct would not be integrated into the main structure, or it would give rise to two-peaked structures (Johnson 2009;

Oseki 2015). On the other side, lower adjuncts forming a macroevent could be assigned a label by means of a mechanism similar to Feature-Sharing (Oseki 2015).

A second potential consequence regards the role of Agree. In some proposals, the island effect is related to the absence of an Agree relation between the island and a c-commanding domain. Basically, when a domain is in an Agree relation with an asymmetrically c-commanding head, such domain is transparent, which means that a filler-gap dependency can be created – and therefore that extraction is licensed. On the contrary, the issue arises when the domain is not in an Agree relation with a c-commanding head. The absence of such relation is what turns that domain into an island, making extraction impossible and the domain opaque (see Rackowski & Richards 2005; den Dikken 2018). Possibly, the macroevent can have a crucial role in the establishment of an Agree relation between the matrix sentence and the adverbial clause: when a macroevent cannot be created, an Agree relation between the matrix verb and the adjunct is formed. However, when the formation of a macroevent cannot be formed, there is no Agree relation between matrix and adjunct, i.e. the adjunct is opaque and no extraction is possible out of it. Notice that, once again, the attachment height is important: as already stated by Rackowski & Richards (2005), adjuncts do not enter into an Agree relation with *v*. But it may be the case that Agree is actually established once all the conditions above are respected, namely: (i) when the adjunct is low; (ii) when a macroevent can be created, and (iii) when the adjunct can undergo a complement reanalysis.^{60,61}

5.4.3 *Summing up*

In this chapter I first considered the difference among adverbial clauses, focusing on experimental results obtained in my tests, as well as on other tests in which adjuncts were investigated. I showed that, once formal judgments are considered, extraction from adjuncts obtains very different results in both islands and, more surprisingly, parasitic gaps. I claimed that such a difference can be captured with a syntactic criterion that allows us to sort out this class, i.e. their position in the syntactic spine. Adjuncts can be lower or higher. I discuss how such difference in their height can influence the possibility of extraction, in that this is allowed from lower adjuncts but not from higher ones. Therefore, the condition I propose is a syntactic

⁶⁰ Notice at this point that the Agree relation is usually restricted to selected domains, i.e. arguments.

⁶¹ See den Dikken (2018) for a similar analysis involving the first two points presented here.

one. However, I showed that such a condition is not sufficient in itself in order to account for the possibility of extraction. The second crucial aspect that has to be taken into account is the possibility of forming a macroevent, i.e. of grouping together the event of the matrix verb and the verb of the adjunct, following the proposal of Truswell (2011). I also showed that such a possibility is directly linked to the height of the attachment of the adjuncts: higher adjuncts cannot form a macroevent with the matrix, whereas lower ones do. This is what allows for extraction from these domains. The macroevent can be formed provided that some conditions are respected: the adjunct has to be low, and it does not have an internal structure that is too rich. In the discussion of extraction, I proposed a tentative analysis whereby when the adjunct is low and can create a macroevent, it is reanalysed as if it was a complement, therefore being a part of the VP-shell following a Larsonian analysis. On the other side, when the adjunct is high – or is low but a macroevent cannot be formed – it is simply attached to the VP or vP or TP depending on the adjunct, but cannot be part of the VP-shell. In these cases, extraction is not allowed.

6. INTER-SPEAKER VARIATION

In this chapter I am going to focus on an additional aspect that resulted from my experiments. I am talking about inter-speaker variation, i.e. the difference in the judgment of the same sentence depending on the speaker. Individual variability has long been noted in the literature – not only regarding extraction from adjuncts. This issue was often addressed anecdotally by several scholars, but despite being noted in many different phenomena, it has rarely been investigated.

In what follows, I will present the relevant data from the three experiments explained in chapter 4, focusing only on the differences between participants and not on the aspects that were already discussed. I will then present some more data collected with other formal experiments investigating extraction from island domains (Kush, Lohndal & Sprouse 2018, 2019; Bondevik 2018). Finally, I will present an account that can explain the presence of inter-speaker variation in extraction from adjuncts.

6.1 Inter-speaker variability and what it can do for generative grammar

In many cases linguists have to work with, and rely on, acceptability judgments of native speakers. These judgments give us the basis for theories and allow us to build models that are able to explain them. With formal experiments, these judgments are examined with a series of statistical analyses often starting with the mean of the sentences investigated. Quite often the mean is considered as a sort of ‘golden number’ and *the* source of genuine data: anything that does not agree with it is pushed aside and considered as noise. See the case of Featherston (2007: 284, original italics), who declares that “[t]he group produces a clear, statistically significant pattern revealing a syntactic generalization, but the judgements of each individual informant are noisy and much less visibly systematic” and “the real appropriate measure is between any single informant’s judgements and the *mean value of a group of informants*”. However, although statistical analysis is necessary in order to establish whether the effects investigated are significant or not, it does not give us all the information we need. Consider for

example the case of the differences among adjuncts, especially as far as Experiment C is concerned: despite the fact that the effect of structure was significant in most of the adverbial clauses investigated, there is a clear difference among them as shown and discussed in §5.1.1. Moreover, the amount of variation among speakers tend to remain – or *has* to remain in cases similar to Featherston (2007) – unexplored. The problem of averaging the data of different speakers implies that we disregard the fact that people have different grammars (different I-languages, in fact), and that the potential difference with respect to the mean is not noise at all, it's just a difference that is present and that should be evaluated and accounted for. Notice that one of the basics of chomskian linguistics is to focus on the *I-language of the individual*, not on the E-language of the speech community. As put in den Dikken, Bernstein, Tortora & Zanuttini (2007: 342,343) “when we generalize from a group of speakers to all speakers of that language, we are making a gross overgeneralization, since we know that grammars can vary from one individual to the next”. Such an overgeneralization can cause a problem: “all potentially interesting points of variation are then cast aside as ‘noise’, and the net result is lots of grey averages. Universal Grammar thus becomes Universal Grey, and that would hardly be reflective of the real patterns that micro-analysis would allow one to identify”. Therefore, in what follows, I will discuss the presence of inter-speaker variation in this study and show that it is fundamental that we take this issue into account in order to uncover certain patterns. Individual variation can also help us choosing the correct theory.

6.2 Experimental data of individual variation

In this section I will present the relevant data coming from the experiments that were presented in chapter 4. When I informally started collecting judgments for sentences presenting extraction from adjuncts, I already noticed that there was something peculiar about these judgments: in some of the cases my informants did not agree. This aspect added an extra degree of difficulty to the topic: how to account for such a diversity? Could it be possible to detected such variation also by means of formal experiments? Notice that considering a standard grammar – or a standard language – we do not expect that speakers do not agree on the level of the (un)acceptability of a sentence. I am not only talking about differences in the degrees of unacceptability or acceptability, but also on the basic acceptable/unacceptable distinction itself. Crucially, the difference is not a small one: when the task assigned to the speakers is to rate the sentences on the basis of a 7-point Likert scale, where 7 indicates a perfect sentence and 1 a

completely unacceptable one, it can happen that the entire scale from 1 to 7 is used for the same item by different speakers. Clearly this is not the ideal situation for a *standard* grammar: if speakers do not agree, it can become difficult to talk about a thing such as a *standard Italian*, or whatever other language. Moreover, taking into account the mean of a certain sentence may be entirely misleading, when inter-speaker variation is present. I will come back to this issue after having explained the relevant data.

6.2.1 Experiment A

Experiment A was the pilot study I conducted in order to check whether extraction from three types of adjuncts was (un)acceptable to the same extent. In this experiment, I checked three main effects: (i) structure (comparing islands and parasitic gaps), (ii) level of embedding, and (iii) role of resumption. Before taking into account experimental data, let me introduce the sample of participants. Since this was a pilot study, the sample was quite restricted: 12 participants took part in the test. The subject pool was uniform; the age range varied between 22 and 28 years old, with all of the participants coming from the same region of Italy in the North-East. All of them had (at least) a bachelor degree. The uniformity of the subject pool is expected to reduce the possibility of variation among participants (see Schütze 2016). However, this is not quite what happens. Their task was to judge the acceptability of these sentences on a 7-point Likert scale, 1 being the worse option and 7 being the best one.¹ As far as inter-speaker variation is concerned, I will focus only on the first point, i.e. the role of structure, and in particular on the items containing extraction from the island condition, not the parasitic gap ones. Therefore, I will present the data of the items containing an island violation with the minimum level of embedding and no resumption. The items for each adjunct are presented in (1).

- (1) a. Quali dolci Luca è scappato [dopo aver rubato _]?
which sweets Luca is run away after having stolen
'Which sweets did Luca run away after having stolen?'

¹ For further details on Experiment A see §4.3.

- b. Quale libro Anna si è preoccupata [prima di trovare _]?
which book Anna herself is worried before finding
'Which book did Anna worried before finding?'
- c. Quale ragazzo Silvia è partita [senza salutare _]?
which boy Silvia is left without greeting
'Which boy did Silvia leave without greeting?'

As far as results are considered, let's just keep in mind that the effect of structure was significant for extraction both from *after* and from *before*, but not in the case of *without*. This means that the sentence containing the island was rated significantly worse than the one with the parasitic gap, namely that there is a statistically relevant island effect. Once we check individual scores, however, things are not so neat. Figures 6.1, 6.2 and 6.3 below show individual judgments of each participant for sentences in (1a), (1b) and (1c), respectively.

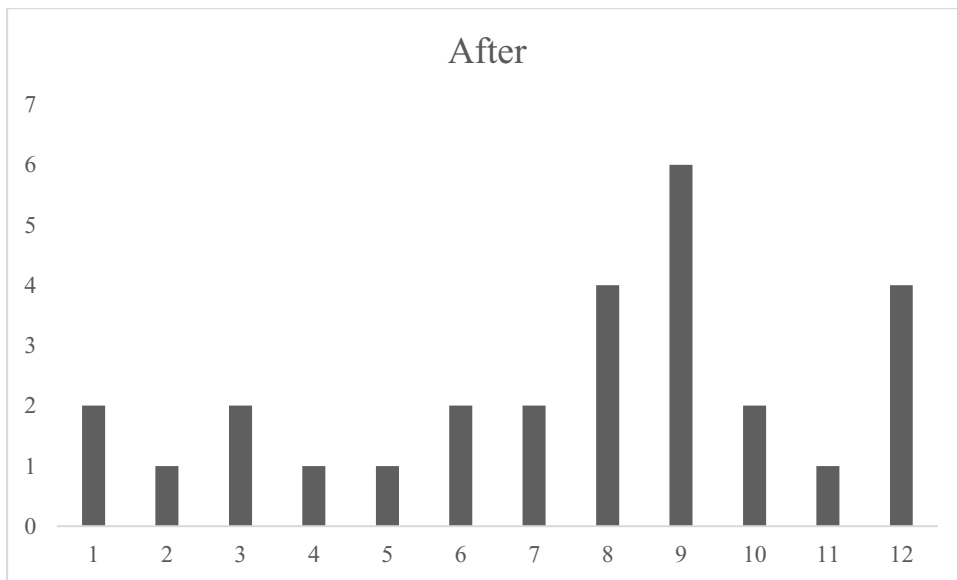


Figure 10 Individual raw scores of each participant for extraction from *after* in the island with no embedding and no resumption condition in Experiment A.

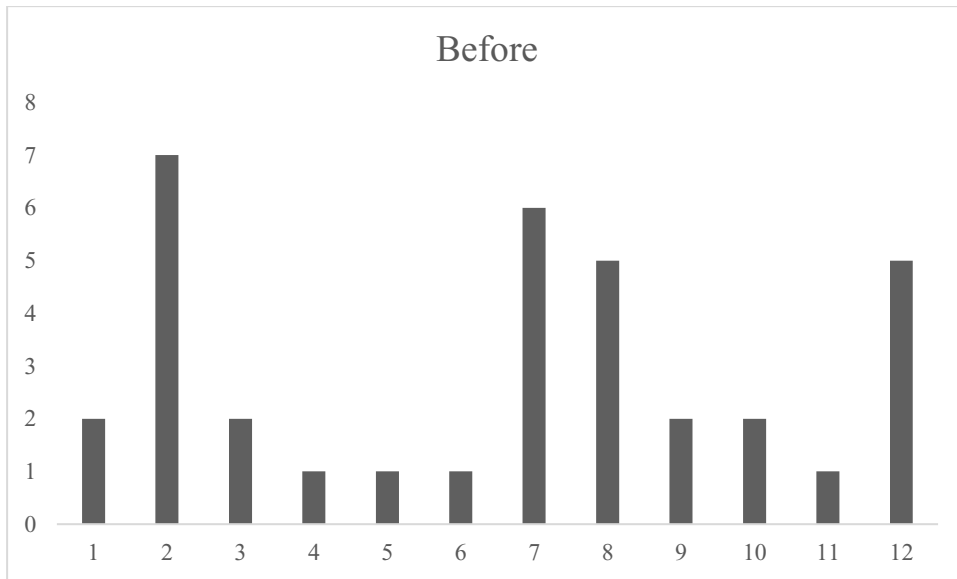


Figure 11. Individual raw scores of each participant for extraction from *before* in the island with no embedding and no resumption condition in Experiment A

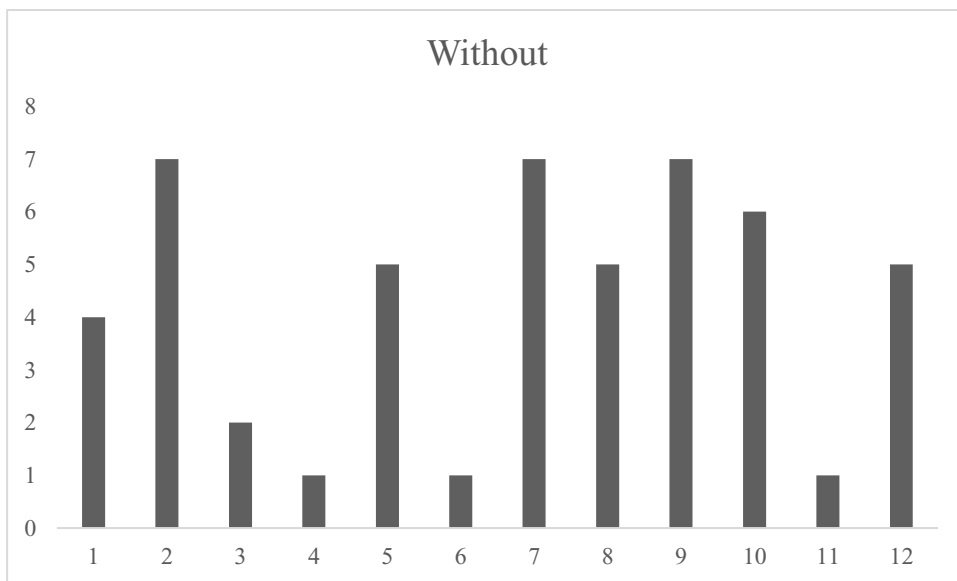


Figure 12. Individual raw scores of each participant for extraction from *without* in the island with no embedding and no resumption condition in Experiment A

As can be seen, extraction from adjuncts behaves differently in one more aspect other than those discussed in §4.3: variability among speakers. Figure 6.1 shows that in the case of extraction from *after*, participants mostly agree on the unacceptability of the sentence, with the majority of the ratings being on the (very) low end of the scale. Things start to change in the case of extraction from *before*, as shown in Figure 6.2. Even though the sentence is bad for most of the participants, there is still someone who accepts it. The distribution of ratings tends to be

bimodal, with one group accepting the sentence, and the other rejecting it. Figure 6.3 shows that *without* behaves in a different manner as well: extraction is much more accepted in this case, both because a bigger number of participants accept the sentence, and because higher ratings are used. Still, to some of the participants extraction is unacceptable. What is more puzzling is that in this item, the entire range on the 7-point Likert scale was used.

Crucially, variation is not found in all of the items but mostly in those that tend to be more accepted, as in the case of *without*. Those that are much more unacceptable do not present such variation, as in *after*, where speakers agree on the unacceptability of the sentence. Interestingly, some speakers rated similarly extraction from *without* and from *before*, which tend to have similar scores. In any case, extraction from *without* is clearly more accepted, also when it comes to the mean of z-scores (*before* = -0.28, *without* = 0.45).² Since this test had only a small sample, I will not dwell much more on the analysis here. I will focus more on such analysis for Experiment B and C, given that these provided a bigger sample.

6.2.2 Experiment B

The main aim of Experiment B was to check the possibility of extraction from prepositional adjuncts, considering also the role of finite tense. The adjuncts were the same as Experiment A: *after*, *before*, and *without*. This choice allowed me to compare the results of the two tests, especially when it comes to the differences among adjuncts. The sample of participants was much bigger than the one of the pilot study, with 80 participants in total. Their age range was between 18 and 25 years old, and they were all enrolled in an introductory linguistics course at the University of Venice. At the time the test was taken, they were not introduced to syntax. In this test there were two items dedicated to each condition, which means that variation can be checked once more. Participants had to rate the sentences according to a 7-point Likert scale. Results were similar to those of Experiment A: extraction is unacceptable with both *after* and *before*, but not so much in the case of *without*. Notice that in this experiment it was significant also for the latter condition, but with a much weaker effect than the others. I will now focus on

² Although notice that the same asymmetries between adjuncts in Experiment A are found in the condition with longer filler-gap dependencies: variability among speakers reflects the one found in the condition analysed here. Extractions from *before*, and especially from *without*, have a wide range of rates, whereas extraction from *after* is mainly in the lower part of the scale.

inter-speaker variability, checking mostly what happens in the island condition with non-finite tense in the adjunct.

It turns out that even in this case some of the items presented in the test have a great amount of variability. I will first take the case of *without* as an example: if we look at the statistical analysis discussed in §4.4, we know that there is an effect of structure, and that this effect is significant. However, such effect is much weaker than in the cases of *before* and *after* ($p < 0.02$ versus $p < 0.0001$). If we look at the mean values it is even clearer that the island effect is not as strong as expected. From these basic statistic data a crucial piece of information is missing from the puzzle, i.e. inter-speaker variation. I will repeat here the two items with extraction from *without*, in the island condition. See the sentences in (2):

- (2) a. Quale ragazzo Silvia è partita senza salutare?
which boy Silvia has left without greeting
'Which boy did Silvia leave without greeting?'
- b. Quale film Silvia è uscita senza aver visto?
which movie Silvia is gone out without having seen
'Which movie did Silvia leave without having seen?'

The mean value of z -scores of (2a) is -0.37, whereas the mean value of (2b) is -0.46. In the judgment task, in sentence (2a) the entire range 1-7 was used. Judgments are bimodally distributed: even though many participants consider (2a) completely ungrammatical and therefore assign 1 or 2, other speakers judge it as perfect, whereas for the majority of the speakers (2a) falls in the area of partial acceptability. This can be seen in Figure 6.4 and 6.5, where individual z -scores of all the participants are depicted for the island condition with non-finite tense.

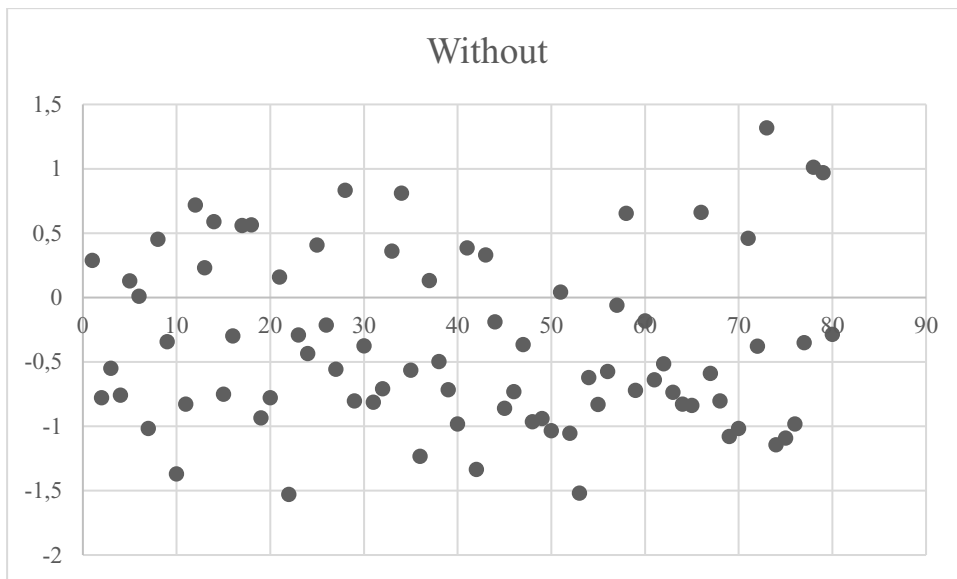


Figure 13. Individual z-score acceptability judgments for each participant for the island condition with non-finite tense of item in (2a) in Experiment B

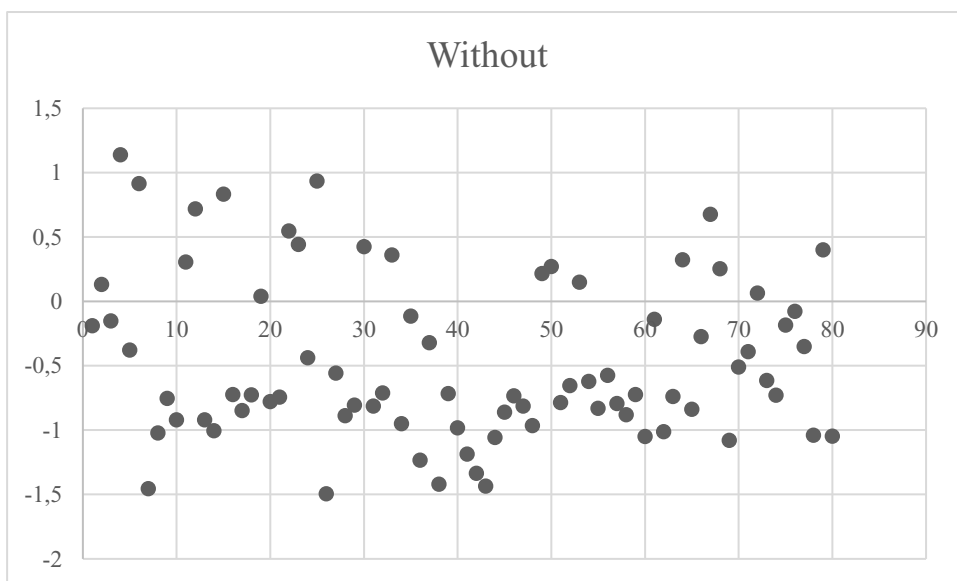


Figure 14. Individual z-score acceptability judgments for each participant for the island condition with non-finite tense of item in (2b) in Experiment B.

Notice that these are *z*-scores, namely transformed raw score where potential scale biases between participants are avoided, as in case only extremes of the scale is used. It can be noticed that judgments are spread all over the graph, or to better put it, they are both above and below the 0 point. We take the 0 to be the tipping point between acceptance and rejection (see also Kush et al. 2019; Bondevik 2018). We can see that in the case of *without* acceptability judgments are not uniform and are actually spread both above and below the 0 line, namely

they are in both the ‘good’ and the ‘bad’ areas of the graph: they are bimodally distributed and there is no agreement between speakers. In cases similar to these ones, it is clear that talking about a ‘mean value’ is not that useful, in that such value would not be a good representation, but only the numerical expression of an ‘ideal’ value based on diverse ratings. If only the mean score is taken into account, the extremities are being reduced – if not cancelled – and we are discarding the existence of differences among speakers.

Crucially, once again variability is not found in all of the items: mostly it is restricted to the adjunct introduced by *without*. Items presenting extraction from the other adjuncts introduced by *after* and *before* do not show the same amount of variability among speakers: in these cases, participants seem to agree more on the unacceptability of extraction, and in fact judgments are mostly uniform. I will use sentences in (3) and (4) to show this.

(3) Quali dolci Luca è scappato [dopo aver rubato _]?
 which sweets Luca is run away after having stolen
 ‘Which sweets did Luca run away after having stolen?’

(4) Quale quaderno Anna ha riflettuto prima di trovare?
 which notebook Anna has thought before finding
 ‘Which notebook did Anna think before finding?’

In the case of (3) the mean value of z-scores is -0.94, whereas for (4) the mean value equals to -0.87. Figures 6.6 and 6.7 show all of the individual z-scores for the items in (3) and (4) respectively.

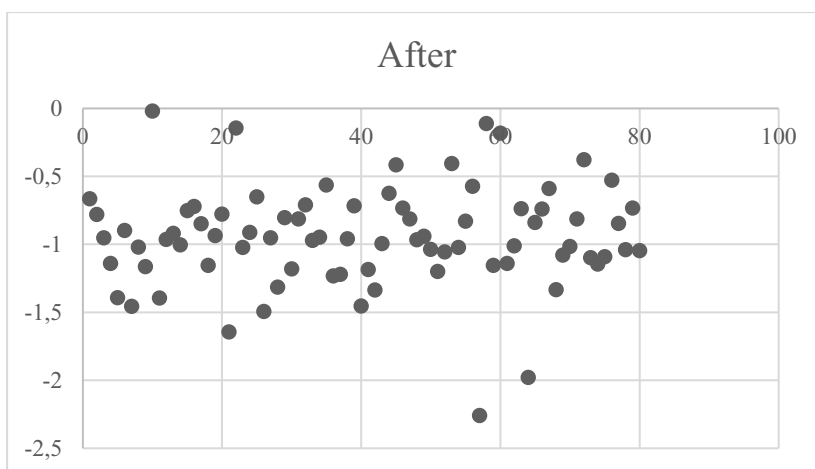


Figure 15. Individual z-score acceptability judgments for each participant for the island condition with non-finite tense of item in (3) in Experiment B.

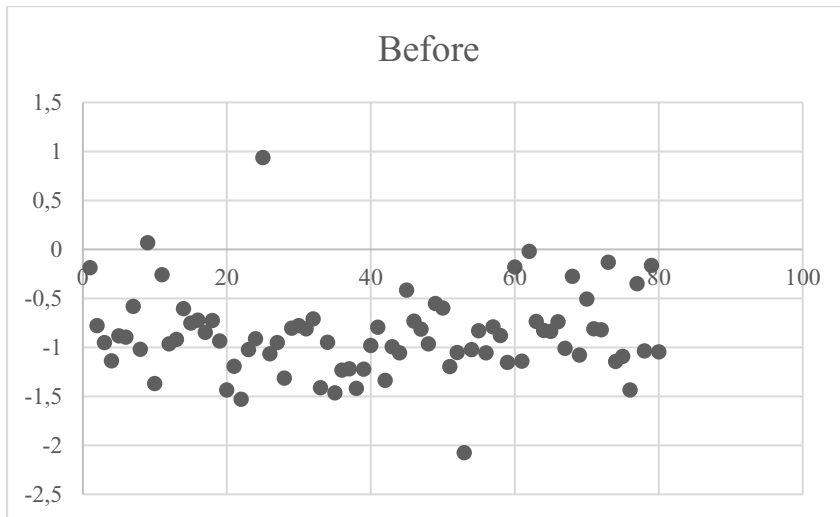


Figure 16. Individual z-score acceptability judgments for each participant for the island condition with non-finite tense of item in (4) of Experiment B

The difference between Figures 6.4 and 6.5 and Figure 6.6 and 6.7 is clear: in neither Figure 6.6 nor 6.7 can we find the same variability among speakers. All the judgments are below the zero point, indicating that there is agreement on the unacceptability of extraction in these cases. Truly enough, there is a difference regarding the degree of the unacceptability of these sentences, but at least there is agreement on the fact that they are unacceptable. Further studies would be required in order to further assess the difference in the degrees of acceptability.

In order to show how the judgments are uniformly distributed above the 0 point once a licit sentence is taken into account, let me just take into account one case of parasitic gap construction, namely the parasitic gap counterpart of (2a), presented in (5).³

- (5) Quale ragazzo Silvia ha guardato senza salutare?
 ‘Which boy did Silvia look without greeting?’

³ I will consider the parasitic gap construction introduced by *without*, given the sound results obtained. As discussed for the three experiments in §4, parasitic gaps are not accepted by everyone, and there seems to be an effect of the type of adjunct taken into account. Since extraction from *without* is the best case scenario, the analysis will show one of these items.

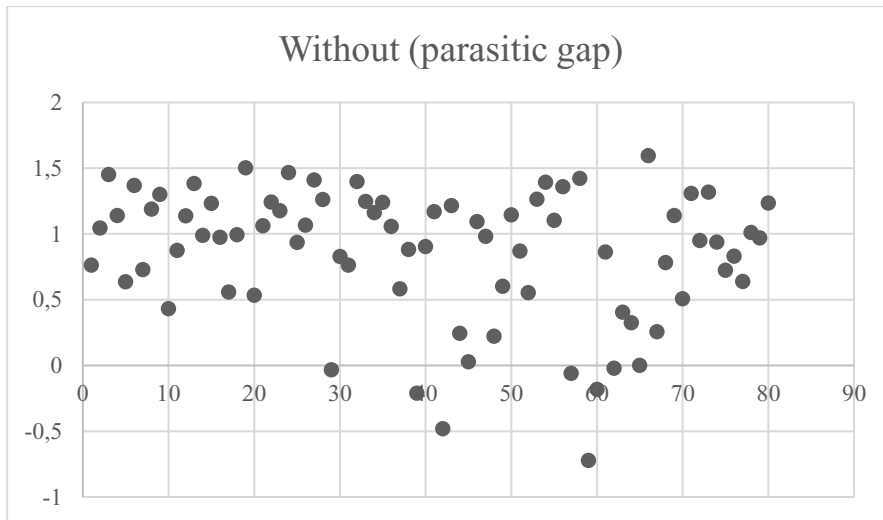


Figure 17. Individual z-score acceptability judgments for each participant for the parasitic gap condition with non-finite tense of item in (5) of Experiment B.

The reader will notice that in this case, basically all of the judgments are to be found above the 0 point, and that there is agreement in assigning a good score on such items. Once again inter-speaker variability is found only in a small portion of the items. As happened with Experiment A, variation is found in those items that are not too bad, as is the case with extraction from *without* in the island condition, but is not present in *after* or *before*, which are generally perceived as bad.

6.2.3 Experiment C

In Experiment C the main purpose was to extend the sample of prepositional adjuncts, and check whether extraction was allowed in a new set of adjuncts. For this reason, I considered not only *after*, *before* and *without*, but also *in order to*, *until* and *because*. This was needed in order to check whether adjuncts do really behave differently. The two effects I considered are (i) the role of structure, and (ii) the role of resumption. I will focus on variability only insofar the former is considered, i.e. in items with the island condition and no resumption realized. The subject pool consisted of 59 participants (age range 18-25), all enrolled in an introductory linguistics course at the University of Venice, at a time when syntax was not yet introduced. They had to rate the acceptability of some sentences on the basis of a 7-point Likert scale, as with the previous experiments. Results showed that in the majority of the adjuncts, the effect of structure is highly significant. It was not significant only in the cases of *in order to* and *until*, but this was connected to the low acceptability of parasitic gaps, rather than to a higher

acceptability of islands (see §4.5.3). However, a difference among adjuncts is still visible, when the mean z-score of the island condition for each adjunct is compared. See Figure 6.9.

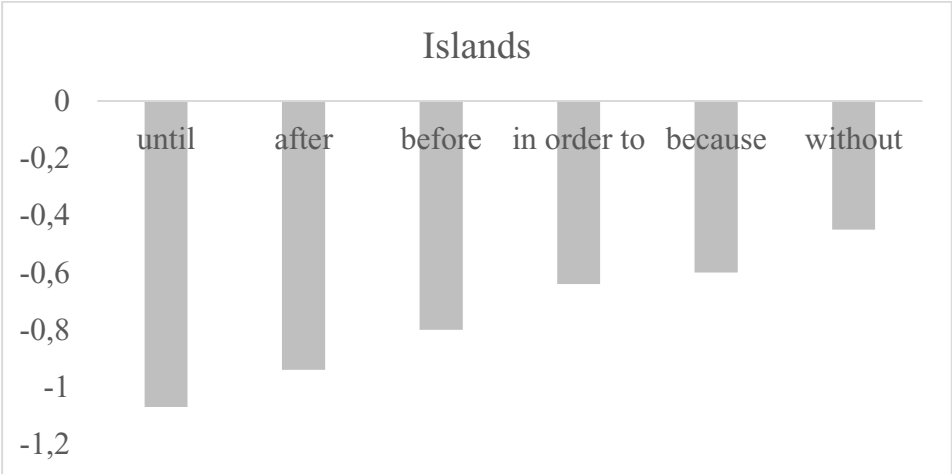


Figure 18. Z-score acceptability judgments comparing extraction from each adjunct type for the island and the gap condition of Experiment C.

Let's see what happens when inter-speaker variation is taken into account. I will present the data from each adjunct in the island condition with no resumptive pronoun.

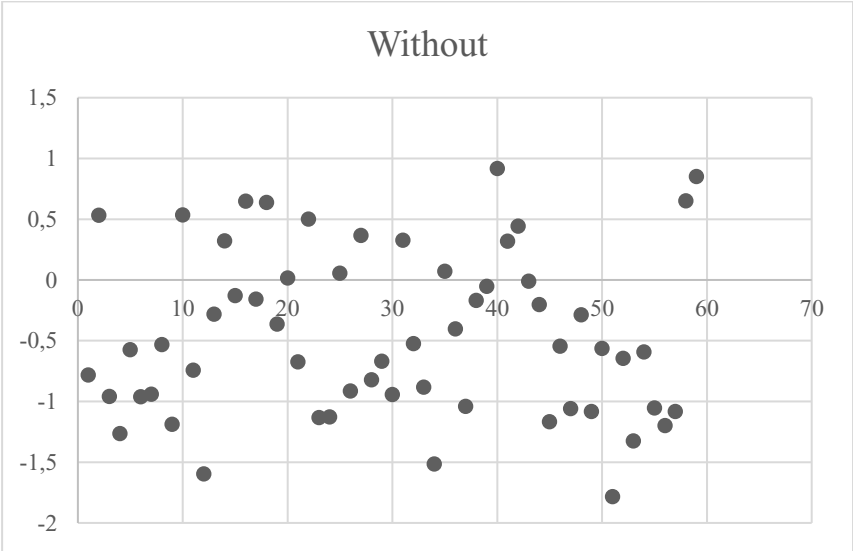


Figure 19. Individual z-score acceptability judgments for each participant for the island condition with gap.

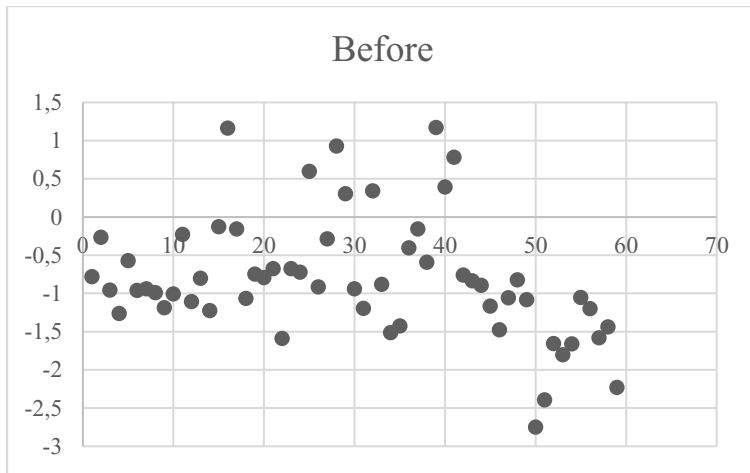


Figure 20. Individual z-score acceptability judgments for each participant for the island condition with gap.

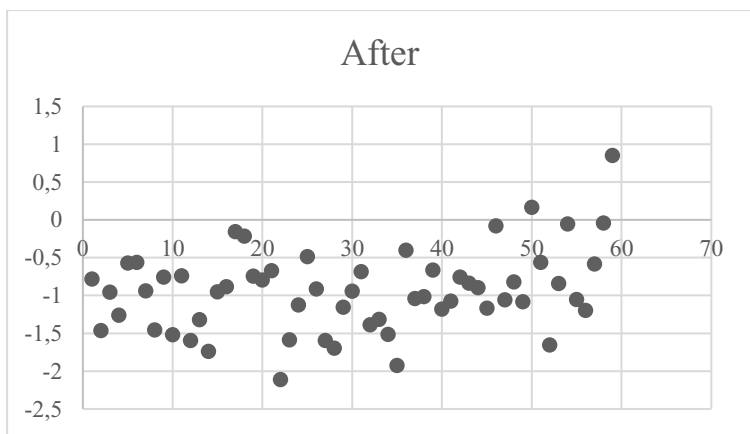


Figure 21. Individual z-score acceptability judgments for each participant for the island condition with gap.



Figure 22. Individual z-score acceptability judgments for each participant for the island condition with gap.

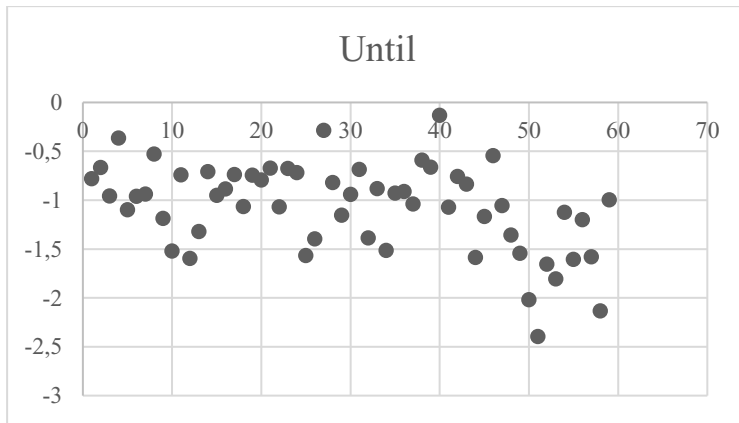


Figure 23. Individual z-score acceptability judgments for each participant for the island condition with gap.

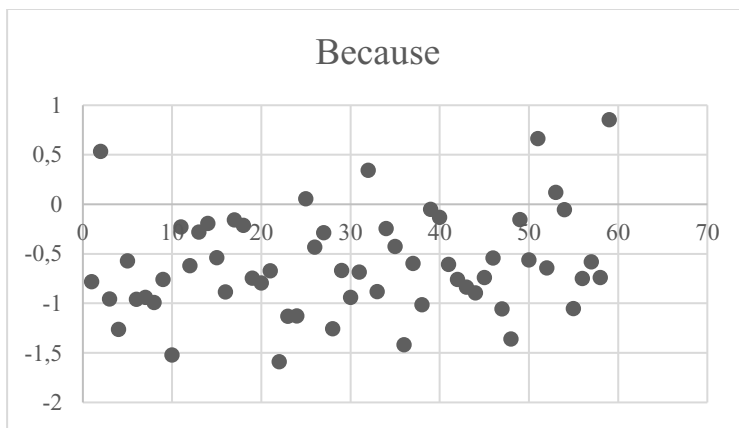


Figure 24. Individual z-score acceptability judgments for each participant for the island condition with gap.

All in all, Experiment C seems to have less inter-speaker variability than the other experiments: as shown in Figures 6.10-6.15, there is a certain agreement on the low acceptability of extraction from these adjuncts. The item that shows more variability is, once again, the case of *without*, even though it presents less variability than the previous experiments. In this case, most of the judgments are below the 0 point. In fact, notice that the statistical analysis showed a highly significant effect of structure also in the case of *without*, different from what happened in both Experiment A and B. Extraction from *without* is nonetheless also the item that most allows extraction, as can be seen from Figure 6.9 where the mean of z-scores of all adjunct types are compared. A really low amount of judgments are above the 0 point in the case of *before*, *in order to* and *because*, but not enough in order to talk about a bimodal distribution. Both *after* and *until* show a clear uniformity in rejecting the extraction.

6.2.4 *Summing up*

The generalizations we can draw from these data are that (i) adjuncts behave differently in one more aspect, i.e. the presence of inter-speaker variability, and (ii) variability is not found in all of the items, but only in those where extraction is more accepted. Therefore, when extraction is not licensed, speakers seem to uniformly agree on the unacceptability of the items. However, when extraction is borderline accepted, we find individual variation. That is to say, there is a correlation between the complementizer type and the amount of variability. As said before, such behaviour is unexpected: if we formulate a grammar made of constraints, we expect that those restrictions are ‘obeyed’ by everybody. This does not seem to be the case here, and in many other phenomena that will partially be reported in §6.3.

Before going on with the analysis I want to stress that inter-speaker variability is exactly what reductionist theories predict. Remember that reductionist theories blame the island effect on different modules, all having to do with the processing costs that are connected to a sentence. Since different people are rating those sentences, and people have different capabilities as far as working memory and attention span are considered, different results are expected depending on the rater. On the other hand, and as already stated, syntactic theories do not predict variation, given that the syntactic module is (should be) the same for everyone. Therefore, the results discussed here seem to point in the direction of reductionist theories. However, this is not the case because there is a crucial aspect that we have to keep in mind, namely the fact that variability is found only in *some* of the items of the test. This shows that the effect is not due to individual differences having to do with processing accounts, given that the sentences are equal in all aspects, except for the type of adverbial clauses from which extraction is taking place. Thus, there seems to be something going on independently from syntax or processing.

6.3 Inter-speaker variation: where else is it found?

Inter-speaker variation is quite spread both when it comes to formally collected judgments and to anecdotally reported ones. I will report here some cases of both, focusing especially on islands. Despite being so widespread, it has been rarely investigated in the literature. Until quite recently, it was not formally addressed even in experimental syntax. I will review in the second part of this section some of them (Kush et al. 2018, 2019; Bondevik 2018), but I will first briefly consider some other cases of individual variation to show that it is consistently found.

As an example, consider Szabolcsi & Lohndal (2017), who discuss about the differences between strong and weak islands. They say that extraction of a PP from a *wh*-island – therefore from a weak island – presents variation among speakers of English, as in (6a). As they write: “there is dialectal variation among speakers of English regarding the acceptability of PP-extraction out of *wh*-islands. Unfortunately, no systematic empirical study of this variation exists” (Szabolcsi & Lohndal 2017: 8). The same situation is found in the acceptability of tensed constituents with *wh*-complements, represented in (6b): some English speakers accept it, whereas to other it is entirely unacceptable.⁴

- (6) a. % About which topic did John ask [whether to talk _]?
 b. % Which man did John ask [who invited _]?

(Szabolcsi & Lohndal 2017: 7-8)

A second example is reported in Fox (1999, 2002), who discusses condition C effect in sentences similar to (7) and (8), where the PPs are analysed as adjuncts

- (7) a. *[How many stories about Diana_i's brother] is she_i likely to invent _ ?
 b. [How many stories about Diana_i's brother] is she₁ likely to reinvent _ ?
- (8) a. *[How many houses in John_i's city] does he_i think you should build _ ?
 b. [How many houses in John_i's city] does he_i think you should rebuild _ ?

(Fox 1999: 167)

Fox reports that there is variation as far as (7b) and (8b) are considered: not all of the speakers accept such sentences. In fact, for some of them these are slightly degraded, whereas (7a) and (8a) are always unacceptable. There is therefore variability among speakers, which he resolves in the following way: those speakers that tend not to accept (7b) and (8b) analyse the PP as a complement, rather than as an adjunct. Therefore Condition C effect is present, given that complements are inserted in the structure in a way that is able to feed the movement, but it is not present when it is analysed as an adjunct in the (b) sentences. This is due to the fact that

⁴ In the literature, the % symbol is commonly used to signal that there is variation among speakers in the acceptability of that sentence.

adjuncts are merged late, following Lebeaux's (1988) approach, i.e. adjuncts are merged after everything else in the derivation and Condition C effect does not arise.

Another case of variation is presented in Salzmänn (2013: 88), focusing on the use of dative resumptives in relative clauses among and within Alemannic varieties of German. The peculiarity of such variation is that for some speakers the use of a resumptive pronoun is mandatory in order to have a grammatical sentence, whereas for other it is not and they can freely alternate the gapped version and the one with resumption.

- (9) a. de Bueb, wo mer **em**/_ es Velo versproche hand.
 the boy C we he.DAT a bike promised have
 'the boy we promised a bike.'
- b. d Frau, won i von **ere**/_ es Buech überchoo han.
 the woman C I from she a book got have
 'the woman from whom I got a book.'

Salzmänn's analysis relies both on (i) inter-speaker variation; and (ii) intra-speaker variation, i.e. variation within the grammar of an individual. This is so because most speakers consider grammatical both the gapped and the resumptive version. In his idea it can be argued that different grammars exist: (i) grammar with obligatory dative resumptives; (ii) grammar without dative resumptives; and (iii) grammar where both options are available.

Finally, Fábregas & Jiménez-Fernández (2016a: 1311) reports that extraction from gerunds in Spanish of the type in (10) is not equally accepted by everyone. They report that these constructions are considered marked by native speakers. "There is, at first sight, some disagreement among individual speakers with respect to how acceptable some of the extractions are; some want to add *, some prefer??., some do not notice anything remarkable about them."

- (10) a. ¿Qué llegó [silbando _] María?
 what arrived whistling María?
 'What did María arrive whistling?'
- b. ¿Qué entró [diciendo _] Juan?
 what entered saying Juan
 'What did John came in saying?'

(Fábregas & Jiménez-Fernández, 2016a:1308)

The cases presented here were only a small subset, but many more examples of individual variation can be consistently found in the literature on the topic.⁵ I will now report some cases of variability in islands detected by means of formal experiments. To start with, Kush et al. (2018) investigated the effect of different types of islands in Norwegian with *wh*-dependencies: complex NP, adjunct, subject, relative clause and *whether* islands. The peculiarity of the results is connected to the last class, which presented a consistent individual variation. Basically, participants had different degrees of sensibility to this kind of island and the judgments displayed a big amount of variability both across and within participants. Participants were divided into ‘true accepters’, i.e. those who consistently assigned high ratings to the sentences containing this island, rejecters, and finally inconsistent raters, who tend to accept one token and reject the other. Their proposal is therefore to accommodate this inter-individual variation with the existence of grammars which differ in some aspects. Therefore, different theories could be used in order to explain such difference. For example, if we follow a Subjacency approach, which predicts that movement has to proceed successive-cyclically, we are forced to conclude that an intermediate landing site – in this case a CP – is occupied by another element, and therefore that long distance movement is blocked. The theory could be relaxed assuming the possibility of having multiple stacked CP phrases in the left periphery, which was proposed to account for island violations in Mainland Scandinavian languages (Christensen & Nyvad 2014; see also Sabel 2002). Such theory could account for individual variation in two ways; either hypothesising that one group has a grammar that allows stacked CPs, and consequently allows extraction, whereas the other does not have such a possibility, and therefore cannot accept extraction. Following this explanation, the difference relies in the grammar of individuals. The second option is to assume that all participants have the same grammar which allows for stacked CPs, but some participants fail to adopt this strategy in some cases; when this happens extraction is blocked. With the latter option it would be easier to accommodate the data regarding the group of inconsistent raters, but it would still be difficult to explain – and understand – why there is such a difference. The second theory Kush et al. explore regards scope intervention and relativized minimality effects (Rizzi 1990, 2004). In this case, the variability depends on the analysis of *whether*: if it is taken to be an operator, extraction will be blocked because of an intervention effect, whereas if it is not analysed as an operator extraction is licensed. In this case, therefore, the difference would be based on how individuals

⁵ See Schütze (2016) for an overview of some of these cases, and den Dikken (2018) for more examples, among many others.

analyse some elements in the sentence, but it is not clear to what such difference is connected to.

A similar situation was found in a follow-up study (Kush et al. 2019) investigating the island effect of different domains through topicalization with two different experiments, one without an introducing context and the second one with context. *Whether*-island mainly presented the same results of Kush et al. (2018): it was judged as acceptable by most of the speakers (true accepters), but there were also some inconsistent raters. They noticed that in this case participants rated the first example of a *whether*-island violation as degraded, but tended to accept later examples.⁶ Therefore, they think these results are compatible with an extra-syntactic explanation of the effect. Moreover, in the case of extraction from adjuncts – the only one investigated in the study is the conditional *if* – they found judgments bimodally distributed: for some extraction was acceptable, whereas for others it was not. They conclude that *conditional* adjuncts are not uniform syntactic islands.⁷ On the other side, judgments coming from relative clauses were mostly uniform in the first experiment: even though some participants accepted extraction, the pattern was not widespread enough. However, once context was included, they found that judgments were bimodally distributed, which lead them to the conclusion that this difference between participants is connected to a difference at the level of their *grammar*: “[s]ome participants could lack the alternate structure in their mental grammars. Such a state of affairs might arise due to dialect variation or through accidents of idiosyncratic exposure” (Kush et al. 2019: 413).

Finally, the results of Bondevik (2018) are also a very good example for inter-speaker variation. She tests extraction from three different types of adjuncts in Norwegian introduced by *if*, *when* and *because*. Extraction from *if* is acceptable (in line with Kush et al. 2019), extraction from *when* is in the middle and extraction from *because* is unacceptable. The first two cases present individual variation. She notices that there is an uncertainty in the acceptability of extraction especially in the case of *when*,⁸ where the distribution is clearly

⁶ This could be connected to a sort of satiation effect. See experimental studies in Snyder (2000) and critiques in Sprouse (2007).

⁷ Actually, their conclusion is the following: “If conditional adjuncts (and some others) are not syntactic islands in Norwegian – and no alternative structural analysis for this subset is tenable, then it would stand to reason that no adjuncts are syntactic islands in Norwegian.” (Kush et al. 2019: 413). Notice that if we take adjuncts not to belong to a uniform class, we are not forced to follow such a conclusion.

⁸ Notice that this finding is supported in the literature: the author reports that she found substantial disagreement across researchers when it comes to extraction from different types of

bimodal, with a group of participants accepting extraction and the other rejecting it. In this case, she found both between participants and within participants variation which could not be accounted for with inter-items variation.⁹ She follows the analysis of Kush et al. (2018) and states that such difference could be related to the intervention effect between moved element and trace because of the analysis of some complementizers as operator. Basically, *because* is analysed as an operator by every speaker, whereas *if* as a non-operator. *When*, on the other side, has a mixed status and can be interpreted as both. Interestingly, she notices that there is a correlation between the amount of individual variation and the acceptability of extraction from that complementizer “variation mirrors the differences between the complementizers that are seen in the interaction plots: the island receiving the highest z-scores (*om* ‘if’) also shows the most variation, the island receiving the mid-range z-scores (*når* ‘when’) shows slightly less variation and the island receiving the lowest z-scores (*fordi* ‘because’) shows even less variation” (Bondevik 2018: 78). Such results nicely fit with my own: the higher degree of acceptability, the higher amount of variability.

6.4 Accounting for diversity

The previous sections showed clearly that we cannot discard the existence of inter-speaker variability; on the contrary, we should profit from it and be able to explain why it is found so consistently.¹⁰ What should be considered is that we cannot reduce grammar to a standard, non-existent language investigated through golden means: there is no standard equal to everybody. It is, I believe, a frontier in the field of linguistics that should be addressed by both a theoretical and an experimental side. As noted by Kush et al. (2018) the latter has mainly focused on drawing inferences from differences in average acceptability calculated at the group level. But this can lead researchers to overlook information that is theoretically relevant or to draw conclusions about tendencies in the data that do not actually exist. It is not fair to state, as Featherston (2007: 309) did, that “[t]he big questions in generative linguistics are those which refer to all speakers, not just one speaker, and to the whole language, if possible to all languages,

adjuncts in Norwegian. The judgments of similar sentences containing extraction from adjunct introduced by *when* in particular was marked as unacceptable by some, but as grammatical by others (see Bondevik 2018: 32 for an overview).

⁹ See the details of the three different experiments she made to assess this point in Bondevik (2018).

¹⁰ See also Schütze (already 1996, and revised 2016) on this point.

not just to a single lexical string'. On the contrary, as den Dikken et al. (2007: 350, italics mine) put it, "we cannot emphasize enough the importance of taking individual judgments seriously. 'Outliers' should not be cast aside as 'noise' as a matter of course. Whenever an apparent 'outlier' presents itself, one should try to ascertain whether it might be correlated with some other 'outlier' from the same informant. What matters, after all, is what *patterns* there are in the data. One apparent 'outlier' may team up or correlate with another apparent 'outlier' elsewhere, and the correlation between these 'outliers' may be enormously revealing from a theoretical point of view". Differences among speakers may turn out to reveal patterns that remain uncovered until variation is taken into account, because what is being defined as 'standard' obscures them. Therefore "[i]n cases where people disagree, that fact cannot be ignored; the theory must be able to describe *every* speaker's competence, and thus must allow for variation wherever it occurs. This is why establishing the extent of interspeaker agreement is important. Theories are now being based on sentences whose status turns out not to be unanimous" (Schütze 2016: 37).

A little note before going on with the explanation of my data is required. One may think, as Featherston (2007) did, that taking into such a big account judgments of individual speakers means rejecting the possibility of the existence of a Universal Grammar. I want to stress that it is clearly not the case. We take UG to be an abstract grammatical system innately present in the mind of every human, which equals to say that the grammar of each individual is a reflection of such UG, i.e. that it meets its principles and parameters. It is also true that languages – human languages – keep changing over time, which also means that different members of a community of speakers of the same language will internalize grammars that *are* different. That is to say *slightly* different, not entirely: there are some small aspects that can change depending on some factors. But of course, "[s]uch grammars are obviously very close, ensuring perfect mutual intelligibility, but may differ in some fundamental ways in their I-language constitutions, thus leading different speakers to analyse a particular string in different ways, with each particular hypothesis leading to different analyses of other input strings and, importantly, to different output utterances regulated by the parameters in question" (den Dikken et al. 2007: 342). Of course we do speak the same language in its core, but sometimes there are differences that cannot be left aside. There is a sort of hard core that cannot be changed (let's say the principled part of UG, maybe syntax at its pureness), but this core is surrounded by a nebula where some aspects of the language change – and we still have to establish which ones – by means of a variation related to many different dialectal, historical and idiolectal aspects. Therefore, in some

cases we are forced to explain these differences in terms of slightly different grammars. Most probably, this also means that within the same individual two grammars of the same language can coexist, and therefore sometimes this makes it difficult to judge some sentences, as in a sort of attrition.¹¹

I will now focus on my own data, and explain why variation is found in the case of extraction from adverbial clauses. Given the theory developed in the previous chapter, the sources of variation that can lead to (un)acceptability are two: (i) the level of attachment of the adjunct, and (ii) the construction of a macroevent. As far as the first option is considered, variation may be attributed to the fact that different speakers have different attachment sites for adjunct. For some speakers the only available option is the higher one, *vP*, which means that extraction will not be allowed because the adjunct is too high to permit the construction of a macroevent. For others, both the *vP* and *VP* positions are available. This means that low adjuncts allow extraction because the single event reading is available, but higher ones do not. Notice that in this case, the difference is attributed to the syntactic component: speakers behave differently because their syntax is different. Here such an explanation is not very successful. In fact, it is not clear why the low position of adjuncts is not permitted for many speakers. Moreover, it would be difficult to accommodate the data of inconsistent raters, namely those who accept extraction in one token but not in the other. My tests were not suited in order to check intra-speaker variation, given the limited amount of items testing the same conditions in the same adjuncts, but we know that this is present from other formal studies (Kush et al. 2018, 2019; Bondevik 2018). In these cases we would be forced to stipulate that the same speaker can attach the same adjunct in two different position. Depending on the attachment site, then, the speaker allows or rejects extraction from the domain.

If we take the second option into account, namely the possibility of constructing a macroevent, we obtain a much nicer picture. The difference does not rely on the attachment site

¹¹ See also Schütze (2016: 194-195): “[u]nder a principles-and-parameters approach, we might expect the periphery, that part of the grammar *not* specified by UG but somehow learned, to vary with people’s learning abilities and experience, but we would presumably not expect variation on matters directly within the scope of innate universals. While it is hard to find data bearing directly on this point, my suspicion is that it is untrue. [...] Are we forced to conclude, then, that UG exhibits individual differences, that we are not all born with identical principles and parameters? This is certainly a possibility, and would not be a particularly surprising result – in general, people do exhibit individual differences on many, perhaps all, innately specified behaviours, while sharing the gross features. In fact, Chomsky (1991) takes it as a truism that genetically based UG will be subject to individual variation. It is merely a theoretically simplifying assumption that UG is invariant.”

of the adjuncts, which is the same for every speaker, but rather on the ability of analyzing the two events as belonging to the same one, of analyzing the two verbs as part of one and the same event. If a macroevent cannot be constructed, no extraction will be possible. My claim is that not all the speakers have the same ability to do so: those who do allow for extraction, whereas those who cannot will not license it. The fact that the (sub)event of the verb in the adjunct can be a part of the (sub)event in the matrix clause and altogether we obtain one single event is the real source of individual variability. This explanation accounts not only for inter-speaker variation, in that not all the individuals have the same ability of construing two single events as a single one, but for intra-speaker variability as well. The same speaker might be able to analyze the two events as a single one in one token but not in the other. This may very well be so, since a single event reading is linked to the lexical choice of items. It explains also the gradiency of these judgments, if we take the construction of a macroevent to be a continuum, rather than present/absent in the same speaker. Truswell himself (2011) noted that the macroevent can be influenced by many factors that have to do with semantics, pragmatics and cognition, and therefore that we expect to find individual variation, both inter-speaker and intra-speaker. In fact, all of these factors are supposed to influence the possibility to form a macroevent, and hence the possibility to extract from an adjunct.

The possibility of extraction is therefore dependent on (i) the type of adjunct, i.e. high or low, and (ii) the speaker. Even though a contingent reading of the events is available, and it is triggered thanks to a context sentence, participants vary in their ability to see such a link between the two events, and therefore to construct a single event. Basically, speakers have to adopt or coerce the relevant reading in order for extraction to be allowed. This explanation is in line with my own results, but also those of (Müller 2017, 2019; Bondevik 2018; Kush et al. 2018, 2019).

A second aspect that such a theory can account for is the correlation found between the amount of variability of an item and the complementizer used. The reader will remember that individual variability was very much present in the case of *without*, but absent when extraction from *after* and *before* is taken into account, especially in the case of Experiment B. The same situation was found in Bondevik (2018) with *if* and *when*, but not from *because*. We then expect there to be a reason why with some adjuncts variability is reduced – or non-existent. In the proposal I developed this is connected to the fact that the points of merger of some adjuncts are

just too high to allow for the creation of the macroevent.¹² This hypothesis hence predicts that with some of the adjuncts variability cannot be found.

To sum up, when we find individual variation the difference is not in the syntactic structure that is being assigned by the speaker, but rather it depends on how easily the speaker is able to accommodate the relevant reading, in this case the creation of a macroevent. The macroevent can be created only once the other condition is respected, namely that the adjunct has to be sufficiently low in the structure. That is the part that does not change depending on the speaker. What changes is the ability of the individual to create such a single event. This is linked not only to how easily the two events can be accommodated into a single one because of the capacity of creating links of the speaker, but also because of our knowledge of the world.

Kush et al. (2019) notice that the account of Truswell (2011) is presented as a universal condition, and therefore that it is not able to explain why extraction is freer in certain languages, as seems to be the case with Norwegian. If the semantic approach is taken to be the only source of explanation, then we would not be able to explain the difference among languages. If, on the other side, the attachment of the adjunct really do count, we could explain such an effect assuming that adjuncts are attached to different parts of the structure which differ according to the language. This requires much more experimental work dedicated to extraction from different types of adjuncts and in different languages.

Second, it seems that there is a difference in extraction depending on the type of A' movement considered: *wh*-dependencies were shown to lead to a much more degraded status than topicalization or relativization.¹³ Neither the single event reading, nor the attachment site of the adjunct are able to provide an explanation of such a different behaviour. In order to assess further this point, much more investigation is needed, to verify whether the same difference can be found in other languages as well. It is also possible that such a difference between A' dependencies is not revealing something about extraction and its condition, but rather about topicalization, relativization and *wh*-questions.

¹² See Truswell (2011) for a different approach which takes into account the difference in the semantics of adverbial clauses.

¹³ As shown in Sprouse et al. (2016) with extraction from conditional adjuncts introduced by *if* in English, which was investigated with both *wh*-dependencies and relativization. The former was degraded and showed an effect of structure, whereas the latter did not. It was shown also by Kush et al. (2018, 2019), where extraction from different types of islands was investigated through *wh*-dependencies first, and topicalization later. The latter was much more accepted.

7. CONCLUSION

The main question I addressed in this thesis regards the possibility of extraction from adjuncts. I started thinking about this issue when I came up with some Italian examples that seemed to allow for extraction, and I was led to check whether my intuition was shared by other speakers and to explain why. Extraction – or to better say the impossibility of extraction – from these domains has been largely discussed since the seminal work of Ross (1967) on *islands*, i.e. domains from which it is difficult or impossible to create A' dependencies. Traditionally, adjuncts belong to the class of strong islands (see Szabolcsi & Lonhaldal 2017), because they do not allow movement of any constituent. Therefore, the so-called Adjunct Condition – which accounts not only for the islandhood of adjuncts, but also of subjects – explains such restriction with rules based on narrow syntax or on the interfaces with the external systems. I showed that these theories do not contemplate the possibility of extraction from adjuncts, in that for all of them adjuncts – and strong islands in general – are absolute, that is to say that movement is never an option: what is inside an island is trapped there.

However, in chapter 3 I showed that this is not the case, focusing on Italian and comparing it to other languages. Once empirical data are taken into account, it is clear that things are not as the theories mentioned above expect: there are several cases of grammatical extraction from different types of adjuncts, and these are found cross-linguistically. In this thesis, I focused in particular on the case of prepositional adjuncts and I showed that there are several factors able to weaken or strengthen their island effect. If adjuncts really are strong islands, this is unexpected: strong islands are supposed not to allow any kind of movement. But this is not the real situation: it matters if extraction takes place from an adjunct with finite verb and not from a non-finite one, it matters if the constituent being extracted is an argument or an adjunct, it matters if it is D-linked. The verbs used are also important, and so is the syntactic integration the adjunct has with the main clause and its position in the structure. Since so many factors are at work – in Italian but also in other languages – I proposed that adjuncts are to be considered as a type of weak islands, rather than strong ones (see also Tanaka 2015; Müller

2019). Consider in particular the fact that there is actually an asymmetry between extraction of an argument and that of an adjunct, which was the diagnostic used to distinguish between the two classes of islands since Cinque (1990). The second option is to analyse the Adjunct Condition as a series of constraints that, taken together, ‘create’ the adjunct island effect, namely that there is not a uniform condition, but it is rather the sum of different factors that matters.¹ Independently of the analysis, the data I presented force us to state that adjuncts are not absolute islands, but they are selective – and I believe that the same analysis can be applied to different islands. So to say, islands “are to be understood dynamically” (Boeckx 2012: 111). This means that the theory accounting for them cannot be too strict and has to be able to adapt to different situations. This is not possible with the approaches mentioned above.

I addressed the possibility of extraction also by means of formal experiments, which allowed me to compare extraction from different types of adjuncts and investigate the role of some of the factors mentioned above. In particular, in the three experiments proposed here, I checked the role of resumptive pronouns (Experiment A and C), the role of embedding (Experiment A), and the role of finite tense in the adjunct (Experiment B). The crucial result obtained is that adjuncts behave differently – both when extraction takes place from an island, and from a parasitic gap. This was shown with statistical analysis in Experiment A, where the island effect of *without* was not significant, and in Experiment B, where the effect was much weaker than the one obtained with *after* and *before*. In Experiment C, all cases of extraction were highly significant, with the exception of *in order to* and *until*, but this was due to the low acceptability of the structures islands were compared to, namely parasitic gaps. The difference between adjuncts can be seen once their mean values are compared: extraction is allowed – or not allowed – at different degrees depending on the adjunct considered. This result aligns with the results of other formal experiments conducted for prepositional adjuncts in Norwegian (Bondevik 2018) and Swedish (Müller 2019).

Therefore, it is not only the case that adjuncts are not strong islands, but also that it is difficult to account for them as a uniform class, given that they give rise to different degrees of deviance even if the items presented are equal to each other – the only difference being the type of complementizer used to introduce the adverbial clause. Among the factors investigated, the length of the filler-gap dependency was not relevant, for any of the adjuncts considered – a result that aligns with the ones of Beltrama & Xiang (2016) and goes against a reductionist

¹ See Haegeman et al. 2014 for an analysis of this kind of subject islands, and Müller 2019 for adjuncts

approach to islands. On the other hand, the role of finite tense was significant for all the adjuncts considered, confirming an aspect often noted in the literature (Cinque 1990; Manzini 1992). Notice that in this respect, Italian behaves in the same way as English, but it is different from other languages such as Swedish (see Müller 2017, 2019) and Czech (Radek & Simik 2018): there are languages that seem to allow extraction even when the verb of the adjunct is finite. The last effect I took into account was the role of resumptive pronouns, which is often considered a rescuing strategy for islands. Their saving value, however, is not easily detected with formal experiments. This was the case of Experiment A, where not only the presence of RP did not help the island, but it made the effect significantly worse in the case of *without*. Experiment C replicated the same result for *without*, which apparently does not tolerate the presence of resumptive pronoun. Notice that this is also the conjunct that most allows for extraction.

There is another crucial result from my experiments that was addressed in chapter 6, the presence of inter-speaker variation. I showed that in some of the items, speakers do not agree on the rate that is to be assigned and hence judgments are bimodally distributed among those who accept extraction and those who don't. Crucially, such variability among speakers is not found with all the items but is restricted to some cases only. In particular, I show that there is a correlation between the amount of variability and the acceptability of extraction from a conjunct: the higher the acceptability rate is, the more variability can be found. In other words, variability is found only with some of the adjuncts – mostly with *without*.

The solution I propose is the following: adjuncts do not form a uniform class and have to be sorted out by means of a syntactic criterion. The criterion I propose regards their point of merger: we can distinguish adjuncts on the basis of their level of attachment to the structure. Some adverbial clauses are attached to a position external to the main sentence, above TP, and are therefore not integrated enough to allow for extraction. These are peripheral adverbial clauses (Haegeman 2004, 2012). Central adverbial clauses, on the other hand, can either be merged to *vP*, or to *VP*. Such a difference in their placement accounts for their different behaviours with respect to extraction phenomena – both with islands and parasitic gaps – and potentially also with respect to some other mechanisms, such as object control. The difference is that adjuncts merged higher do not permit extraction, whereas adjuncts merged lower *can*. They can, but they don't have to because extraction is actually connected to a second aspect, i.e. the formation of a macroevent (Truswell 2007, 2011). Only when this is the obtained reading, there can be extraction. The construction of a macroevent, where the (sub)event of the

matrix verb and the (sub)event of the adjunct are analysed together as a single event, is not only a semantic condition, as was proposed by Truswell, but actually depends on a syntactic aspect. In fact, I claim that a macroevent can be created only with adjuncts that are merged lower, but not with those merged higher, capturing the fact that the higher the adjunct, the more difficult the extraction. As explained in chapter 5 and 6, however, it is not the case that every time an adjunct is merged low a macroevent can be formed. This reading is available if some other conditions are respected: (i) the internal structure of the adjunct should not be too rich – an aspect that is probably related to the external syntax of the adjunct, namely to its point of merger in the structure (see Haegeman 2012; Endo & Haegeman 2019); (ii) the lexical choice of the items; (iii) the relation created by the two verbs, i.e. causation or enablement, and (iv) individual capacity of constructing a macroevent. The last aspect is crucial – as shown in chapter 6 – to account for individual variability. I showed that in my experiments the best way to account for its presence is not a difference in the syntax of speakers, but rather in their capacity of linking events and see them as connected. This accounts also for the fact that there is a correlation between the amount of variability and the type of complementizer: since only some of the adverbial clauses allows for the creation of a macroevent because they are merged low, we expect that those merged higher do not even present inter-speaker variation, and such prediction is borne out. This is true not only for my experiments of Italian, but also for those of Norwegian (Bondevik 2018). Moreover, such an analysis can account for cases of intra-speaker variation as well: it can be that in some cases the same speaker is able to create the macroevent, but in others she/he cannot, depending on the item.

I conclude this thesis with some open questions, and space for future researches both for the experimental side and the theoretical one. I will mention the former first. As discussed in chapter 4, it seems that extraction is influenced also by the type of A' dependency that is being tested. In my experiments only *wh*-questions were used, but some studies used topicalization (Müller 2019; Bondevik 2018) or relativizations compared to *wh*-questions (Sprouse et al 2016), or topicalizations compared to *wh*- questions (Kush et al. 2018, 2019). The unexpected results is that the island effect is different depending on the dependency, in that typically it is much more present in *wh*-questions than in relativization or topicalization. This is unusual for a syntactic theory: if all the dependencies involve movement, extraction should be blocked to the same amount for all of them. Second, the role of the context seems to be crucial. In some studies testing the possibility of extraction, the item is preceded by an introducing sentence that provides a context and in at least two tests the role of context was

significant.² However, the effect of context is different depending on the islands that are being tested: it is not the case that all types of islands are ‘saved’ by the presence of context. Therefore, it would be interesting to check whether there is a difference also among adjuncts. Third, and last, experimental studies should focus more on cases of inter-speaker and intra-speaker variation. This would allow us to observe new patterns and establish whether the theories we are using are able to accommodate the data. In order to do so, more items of the same type and focusing on the same structure should be presented especially as far as intra-speaker variation is considered.

As for the theoretical side, the consequences of my proposal first imply a revision of the basic theory of adjuncts of the central kind – given that the claim is that they are not a uniform class. This means that there should be some other phenomena able to show such a difference. The one I briefly mentioned regards object control on the adjunct – which in general is not allowed but with some exceptions. As far as I know, these exceptions are limited to the case of *in order to* (see Hornstein 1999, 2001): a potential explanation is that such adjunct is attached even lower than *without*, in a position where it can be c-commanded by the object. It may be the case, therefore, that the attachment sites for adjuncts are not only two – as proposed here.

Moreover, the role of the macroevent could be more important than previously assumed, in that it can influence extraction also from other structures such as parasitic gaps and coordinate structures. Finally, the consequences of the construction of a macroevent could involve additional explanations that are adopted by some scholars to explain islands. For instance, it could influence the kind of label that is assigned to the adjunct – a totally non-trivial aspect (see Nunes & Hornstein 2008 for a discussion), but also of the role of Agree (Rackowski & Richards 2005; den Dikken 2018). The macroevent could affect both these aspects – provided that the macroevent itself is affected by the attachment site of the adjunct – in that it could have a key role in assigning a label (or simplifying it) or in establishing an Agree relation. I will leave these – very – open questions for future investigations.

² Kush et al. (2019) from relative clauses; Beltrama & Xiang 2016 on the role of resumptive pronouns on the comprehensibility of different types of islands.

References

- Abels, Klaus. 2003. *Successive cyclicity, anti-locality and adposition stranding*. Storrs: University of Connecticut. (Doctoral dissertation.)
- Abels, Klaus. 2012. *Phases: An essay on cyclicity in syntax*. Berlin: Walter de Gruyter.
- Ackerman, Lauren, Frazier Michael and Masaya Yoshida. 2014. Resumptive pronouns salvage island violations in forced-choice tasks. (Poster presented at the CUNY Sentence Processing Conference.)
- Alexopoulou, Theodora & Frank Keller. 2007. Locality, cyclicity and resumption: At the interface between the grammar and the human sentence processor. *Language* 83. 110–160. DOI: <http://dx.doi.org/10.1353/lan.2007.0001>
- Alexopoulou, Theodora & Frank Keller. 2013. What vs. who and which: Kind-denoting fillers and the complexity of whether-islands. In Jon Sprouse & Norbert Hornstein (eds.) *Experimental Syntax and Island Effects*, 310-340. Baltimore, MD: Cambridge University Press.
- Almeida, Diogo. 2014. Subliminal wh-islands in Brazilian Portuguese and the consequences for syntactic theory. *Revista da ABRALIN* 13(2). 55-93.
- Alrenga, Peter. 2005. A sentential subject asymmetry in English and its implications for complement selection. *Syntax* 8(3). 175-207.
- Anagnostopoulou, Elena. 1994. *Clitic dependencies in Modern Greek*. Salzburg: University of Salzburg. (Doctoral dissertation).
- Andrews, Avery. D. III. (1975). *Studies in the syntax of relative and comparative clauses*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).
- Anward, Jan. 1982. Basic Swedish. In Elisabeth Engdahl & Eva Ejerhed (eds.), *Readings on unbounded dependencies in Scandinavian languages*, 47-75. Stockholm: Almqvist & Wiksell.
- Aoun, Joseph, Lina Choueiri & Norbert Hornstein. 2001. Resumption, Movement, and Derivational Economy. *Linguistic Inquiry* 32(3). 371–403.
- Asudeh, Ash. 2011. Local grammaticality in syntactic production. In Emily Bender and Jennifer Arnold (eds.), *Language from a cognitive perspective: Grammar, usage & processing. Studies in honor of Thomas Wasow*. Stanford: CSLI Publications.
- Bates, Douglas, Maechler, Martin, Bolker, Ben, Walker, Steve. 2015. Fitting Linear Mixed-

- Effects Models using lme4. *Journal of Statistical Software*, 67(1). 1-48.
- Belletti, Adriana. 2006. Extending Doubling to Non Local Domains: Complete vs. Partial Copying+ Deletion and Related Reconstruction Issues. In Patrick Brandt & Eric Fuss (eds.), *Form, Structure and Grammar*, 129-136. Berlin: Akademie-Verlag.
- Belletti, Adriana & Luigi Rizzi. 1981. The syntax of ‘ne’: Some theoretical implications. *The Linguistic Review* 1. 117–54.
- Beltrama, Andrea, and Ming Xiang. 2016. Unacceptable but comprehensible: The facilitation effect of resumptive pronouns. *Glossa* 1(1). 1–24.
- Berwick, Robert & Amy Weinberg. 1984. *The grammatical basis of linguistic performance*. Cambridge, MA: MIT Press.
- Bhatt, Rajesh & Roumyana Pancheva. 2006. Conditionals. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, 638-687. Oxford and Boston: Blackwell Publishing.
- Bianchi, Valentina. 2004. Resumptive Relatives and LF Chains. In Luigi Rizzi (ed.), *The Cartography of Syntactic Structures Vol. II*, 76-114. Oxford: Oxford University Press.
- Bianchi, Valentina & Cristiano Chesi. 2012. Subject islands and the Subject Criterion. In Valentina Bianchi & Cristiano Chesi (eds.), *Enjoy Linguistics! Papers offered to Luigi Rizzi on the occasion of his 60th birthday*, 25-53. Siena: Centro Interdipartimentale di Studi Cognitivi sul Linguaggio (CISCL Press).
- Biskup, Petr & Šimík, Radek. 2018. *Extraction from clausal adjuncts in Czech*. (Paper presented at Formal Approaches of Slavic Linguistics 27, Stanford University, 4-6 May 2018).
- Boeckx, Cedric. 2003. *Islands and chains*. Amsterdam: John Benjamins.
- Boeckx, Cedric. 2008. *Bare Syntax*. Oxford: Oxford University Press.
- Boeckx, Cedric. 2012. *Syntactic Islands*. Cambridge: Cambridge University Press.
- Boeckx, Cedric. 2014. *Elementary syntactic structures: Prospects of a feature-free syntax*. Cambridge: Cambridge University Press.
- Bondevik, Ingrid. 2018. *Investigating the universality of adjunct islands through formal acceptability experiments. A comparative study of English and Norwegian*. Trondheim: Norwegian University of Science and Technology. (MA thesis.)
- Borgonovo, Claudia & Ad Neeleman. 2000. Transparent adjuncts. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 45. 199–224.
- Borgonovo, Claudia & Vidal Valmala. 2009. *Constructions and the theory of grammar*. University Laval and University of the Basque Country. (Manuscript.)
- Bošković, Željko (1994). D-structure, theta-criterion, and movement into theta positions.

- Linguistic Analysis* 24. 247–86.
- Bošković, Željko & Howard Lasnik. 2001. How strict is the cycle? *Linguistic Inquiry* 30. 691-703.
- Brown, Jessica, M.M. 2016. Blackholes and subextraction from adjuncts in English and Norwegian. In *Proceedings of Chicago Linguistics Society 51*, 67-81. Chicago, IL: Chicago Linguistics Society.
- Brown, Jessica M.M. 2017. *Heads and adjuncts. An experimental study of subextraction from participials and coordination*. Cambridge: University of Cambridge. (Doctoral thesis.)
- Brugè, Laura & Avellina Suñer. 2008. Building up Complex Temporal Constructions. *Working Papers in Linguistics* 18. 7-63.
- Burton-Roberts, Noel. (2006). Parentheticals. In Keith Brown (ed.), *Encyclopaedia of Language and Linguistics*, vol. 9 (2nd ed), 179–182. Amsterdam: Elsevier Science.
- Cable, Seth. 2010. *The grammar of Q: Q-particles, wh-movement, and pied-piping*. Oxford: Oxford University Press.
- Cattell. 1976. Constraints on movement rules. *Language* 52. 18-50.
- Cecchetto, Carlo & Caterina Donati. 2015. *(Re)labeling*. Cambridge, MA: MIT Press.
- Chacón, Dustin. Forthcoming. Acceptability (and other) experiments for studying comparative syntax. In Grant Goodall (ed.), *The Cambridge Handbook of Experimental Syntax*.
- Chaves, Rui. 2012. On the grammar of extraction and coordination. *Natural Language and Linguistic Theory* 30. 465-512.
- Chesi, Cristiano. 2007. Five reasons for building phrase structures top-down from left to right. *Nanzan Linguistics* 3. 45-80.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 232-286. New York: Holt, Rinehart & Winston.
- Chomsky, Noam. 1977. On wh-movement. In Peter Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 71-132. New York: New York Academic Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1991. Universal grammar. Letter to the editor. *The New York Review of Books* 38(21).

- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by Step: essays on Minimalist syntax in honour of Howard Lasnik*, 89-115. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1-52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and beyond*, 104-131. New York: Oxford University Press.
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria-Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory*, 133-166. Cambridge, MA: MIT Press. DOI: <https://doi.org/10.7551/mitpress/9780262062787.003.0007>.
- Chomsky, Noam. 2013. Problems of Projections. *Lingua* 130. 33-49.
- Christensen, Ken R. & Anne Mette Nyvad. 2014. On the nature of escapable relative islands. *Nordic Journal of Linguistics* 37. 29-45.
- Cinque, Guglielmo. 1990. *Types of A'-dependencies*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 2006. Complement and adverbial PPs: Implications of clause structure. In Guglielmo Cinque (ed.), *The Cartography of Syntactic Structures*, 145-166. New York: Oxford University Press.
- Collins, Chris. 2002. Eliminating Labels. In Samuel David Epstein & Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program*, 42-64. Oxford: Blackwell.
- Dal Farra, Chiara. 2019a. The case of adjunct islands in Italian. In Astrid van Alem, Anastasiia Ionova & Cora Pots (eds.), *Proceedings of the 26th Conference of the Student Organization of Linguistics in Europe*, 264-278. Leiden: Leiden University Centre for Linguistics.
- Dal Farra, Chiara. 2019b. Variability in Adjunct Islands: An Experimental Test in Italian. In Richard Stockwell, Maura O'Leary, Zhongshi Xu, & Z.L. Zhou (eds.), *Proceedings of the 36th West Coast Conference of Formal Linguistics 36*, 115-120. Somerville, MA: Cascadilla Proceedings Project.
- Deane, Paul. 1991. Limits to attention: A cognitive theory of island phenomena. *Cognitive Linguistics* 2. 1-63.
- Demonte, Violeta. 1988. Remarks on secondary predicates: C-command, extraction and reanalysis. *The Linguistic Review* 6. 1-39.
- De Vos, Mark. 2005. *The syntax of verbal pseudo-coordination in English and Afrikaans*. Leiden: Universiteit Leiden. (Doctoral dissertation).
- Diessel, Holger. 2001. The ordering distribution of main and adverbial clauses: A typological study. *Language* 77. 345-365.

- Dikken, Marcel den. 2018. *Dependency and directionality*. New York, NY: Cambridge University Press. DOI: <https://doi.org/10.1017/9781316822821>.
- Dikken, Marcel den, Judy Bernstein, Cristina Tortora & Raffaella Zanuttini. 2007. Data and grammar: Means and individuals. *Theoretical Linguistics* 33(3). 335-352.
- Dowty, David. 1979. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and Montagues PTQ*. Dordrecht: Reidel.
- Endo, Yoshio & Liliane Haegeman. 2019. Adverbial clauses and adverbial concord. *Glossa: a journal of general linguistics* 4(1). 1–32. DOI: <https://doi.org/10.5334/gjgl.589>
- Engdahl, Elisabeth. 1983. Parasitic gaps. *Linguistics and Philosophy* 6. 5-34.
- Engdahl, Elisabeth. 1997. Relative clause extractions in context. *Working Papers in Scandinavian Syntax* 60. 59-86.
- Epstein, Samuel D., Hisatsugu Kitahara & Daniel Seely. 2012. Structure building that can't be. In Myriam Uribe-Etxebarria & Vidal Valmala (eds.), *Ways of Structure Building*, 253-270. Oxford: Oxford University Press.
- Ernst, Thomas. 2002. *The Syntax of Adjuncts*. Cambridge: Cambridge University Press.
- Erteschik-Shir, Nomi. 1973. *On the nature of island constraints*. Cambridge, MA: Massachusetts Institute of Technology (Doctoral dissertation.)
- Espinal, M. Teresa. (1991). The representation of disjunct constituents. *Language*, 67(4). 726–762.
- Etxepare, Ricardo. 1998. On the Structure of Declarative Clauses. In Armin Schwegler, Bernard Tranel & Myriam Uribe-Etxebarria (eds.), *Romance Linguistics: Theoretical Perspectives, Current Issues in Linguistics Theory*, 160. Amsterdam/Philadelphia: John Benjamins.
- Faarlund, Jan Terje. 1992. *Norsk syntaks i funksjonelt perspektiv* [Norwegian syntax in a functional perspective]. Oslo: Universitetsforlaget.
- Fábregas, Antonio & Ángel L. Jiménez-Fernández. 2016a. Extraction from gerunds and the internal syntax of verbs. *Linguistics* 54. 1307-54.
- Fábregas, Antonio & Ángel L. Jiménez-Fernández. 2016b. Extraction out of adjectival secondary predicates in English and Spanish: A nanosyntactic account. *Questions and Answers in Linguistics vol.3* (2). 41-56.
- Featherston, Sam. 2007. Data in generative grammar: The stick and the carrot. *Theoretical Linguistics* 33(3). 269–318.
- Ferreira, Fernanda & Benjamin Swets. 2005. The production and comprehension of resumptive pronouns in relative clause island contexts. In Anne Cutler (ed.), *Twenty-first century psycholinguistics: Four cornerstones*, 101–122. Malway, NJ: Lawrence Erlbaum

Associates.

- Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30. 157–196.
- Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, MA: MIT Press.
- Frey, Werner. 2011. Peripheral adverbial clauses, their licensing and the prefield in German. In Eva Breindl, Gisella Ferraresi & Anna Volodina (eds.), *Satzverknüpfung – zur Interaktion von Form, Bedeutung und Diskursfunktion*, 41–77. Berlin: de Gruyter.
- Gallego, Ángel J. 2010. *Phase theory*. Amsterdam, Philadelphia: John Benjamins.
- Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1). 1–76.
- Giorgi, Alessandra. 2014. Prosodic signals as syntactic formatives in the left periphery. In Anna Cardinaletti, Guglielmo Cinque & Yoshio Endo (eds.), *On peripheries: Exploring clause initial and clause final positions*, 161–188. Tokyo: Hituzi Syobo Press.
- Giorgi, Alessandra. 2016. Integrated parentheticals in quotations and free indirect discourse. In Alessandro Capone, Ferenc Kiefer & Franco Lopiparo (eds.), *Indirect reports and pragmatics*, 471–488. Springer International Publishing.
- Giorgi, Alessandra & Fabio Pianesi. 2001. Ways of terminating. In Carlo Cecchetto, Gennaro Chierchia & MariaTeresa Guasti (eds.), 211–277. *Semantic interfaces: Reference, Anaphora and Aspect*. Chicago: CSLI Publications.
- Goldberg, Adele. 2006. *Constructions at work. The nature of generalization in language*. New York: Oxford University Press.
- Goodall, Grant. 2015. The D-linking effect on extraction from islands and non-islands. *Frontiers in Psychology* 5:1493. DOI: <https://doi:10.3389/fpsyg.2014.01493>.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.
- Grohmann, Kleanthes (2003). *Prolific domains: On the Anti-locality of movement dependencies*. Amsterdam: John Benjamins.
- Haegeman, Liliane. 2003. Conditional clauses: External and internal syntax. *Mind & Language* 18(4). 317–339.
- Haegeman, Liliane. 2004. The syntax of adverbial clauses and its consequences for topicalisation. In Martine Coene, Gretel De Cuyper & Yves D'hulst (eds.), *Current Studies in Comparative Romance Linguistics*, 61–90. Antwerp: University of Antwerp.
- Haegeman, Liliane. 2010. The internal syntax of adverbial clauses. *Lingua* 120. 628–648.

- Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and the composition of the left periphery*. Oxford: Oxford University Press.
- Haegeman, Liliane, Ángel L. Jiménez-Fernández & Andrew Radford. 2014. Deconstructing the subject condition: Cumulative constraint violation and tolerance thresholds. *The Linguistic Review* 31. 73-150. DOI: <https://doi.org/10.1515/tlr-2013-0022>.
- Hansen, Erik & Lars Heltoft. 2011. *Grammatik over det danske sprog* [Danish reference grammar], vol. 3. København: Det danske Sprog- og Litteraturselskab: Syddansk universitetsforlag (kommissionær).
- Harwood, William. 2015. Being progressive is just a phase: celebrating the uniqueness of progressive aspect under a phase-based analysis. *Natural Language and Linguistic Theory* 33(2), 523-573.
- Heestand, David, Ming Xiang, and Maria Polinsky. 2012. Resumption still does not rescue islands. *Linguistic Inquiry* 42. 138–52.
- Hegarty, Michael (1992). *Adjunct extraction and chain configurations*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation.)
- Higginbotham, James. 1983. A note on phrase-markers. *Revue Québécoise de Linguistique* 13.147-166.
- Higginbotham, James. 2009. *Tense, Aspect and Indexicality*. Oxford: Oxford University Press.
- Hofmeister, Philip. 2007. *Representational complexity and memory retrieval in language comprehension*. Stanford, CA: Stanford University. (Doctoral dissertation.)
- Hofmeister, Philip & Ivan Sag. 2010. Cognitive constraints and island effects. *Language* 86. 366-415.
- Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30(1). 69-96.
- Hornstein, Norbert. 2001. *Move! A Minimalist theory of construal*. Oxford: Blackwell.
- Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and Universal Grammar*. Cambridge: Cambridge University Press.
- Hornstein, Norbert, Howard Lasnik & Juan Uriagereka. 2007. The dynamics of islands: Speculations on the locality of movement. *Linguistic Analysis* 33. 149-175.
- Hornstein, Norbert & Jairo Nunes. 2008. Adjunction, labeling, and Bare Phrase Structure. *Biolinguistics* 2(1). 57-86.
- Huang, James. 1982. *Logical relations in Chinese and the theory of grammar*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation.)
- Huddleston, Rodney & Geoffrey Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge, MA: Cambridge University Press.

- Huddleston, Rodney & Geoffrey Pullum. 2006. Coordination and subordination. In Bas Aarts & April McMahon (eds.) *A handbook of English linguistics*, 198-219. Oxford: Blackwell.
- Ishii, Toru. 1998. Derivational selectional restriction and reconstruction. *English Linguistics* 15. 28-49.
- Jiménez Fernández, Àngel. 2009. On the composite nature of subject islands: a phase-based approach. *SKY Journal of Linguistics* 22. 91-138.
- Johnson, Kyle. 2003. Towards an etiology of adjunct islands. *Nordlyd* 31. 187-215.
- Johnson, Kyle. 2009. Why movement? (Paper presented at Chicago Linguistic Society meeting, April 2009).
- Jurka, Johannes. 2010. *The importance of being a complement: CED effects revisited*. College Park: University of Maryland. (Doctoral dissertation.)
- Jurka, Johannes, Chizuru Nakao, & Akira Omaki. 2011. It's Not the End of the CED as We Know It: Revisiting German and Japanese Subject Islands. In Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer & Barbara Tomaszewicz (eds.), *Proceedings of the 28th West Coast Conference on Formal Linguistics*, 124-132. Somerville, MA: Cascadilla Proceedings Project.
- Kayne, Richard. 1983. Connectedness. *Linguistic Inquiry* 14. 223-49.
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Kayne, Richard. 2005. *Movement and silence*. New York: Oxford University Press.
- Keshev, Maayan. 2016. *Active dependency formation in syntactic islands: Evidence from Hebrew sentence processing*. Tel-Aviv: Tel-Aviv University. (Master's thesis).
- Keshev, Maayan & Aya Meltzer-Asscher. 2018. Active dependency formation in islands: How grammatical resumption affects sentence processing. *Language*, 93(3). 549-568.
- Kiziak, Tanja. 2010. *Extraction asymmetries: Experimental evidence from German*. Amsterdam/Philadelphia: John Benjamins.
- Kluender, Robert & Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8. 573-633.
- Kohrt, Annika, Trey Sorensen & Dustin A. Chacón. Forthcoming. The real-time status of semantic exceptions to the adjunct island constraint. Minneapolis: University of Minnesota. (Draft).
- Kroch, Anthony. 1981. On the role of resumptive pronouns in amnestying island constraint violations. In Eric Williams et al. (eds.), *Proceedings of the 16th Chicago Linguistic Society (CLS 16)*, 125-135. Somerville, MA: Cascadilla Press.

- Kush, Dave, Terje Lohndal & Jon Sprouse. 2018. Investigating variation in island effects: A case study of Norwegian Wh-extraction. *Natural Language & Linguistic Theory* 36(3). 743-779. DOI: <https://doi.org/10.1007/s11049-017-9390-z>.
- Kush, Dave, Terje Lohndal, & Jon Sprouse. 2019. On the Island Sensitivity of Topicalization in Norwegian: An Experimental Investigation. *Language* 95(3). 393-420.
- Lasnik, Howard. 2001. When you can save a structure by destroying it? In Minjoo Kim & Uri Strauss (eds.), *Proceedings of North East Linguistic Society 31*, 301-320. University of Massachusetts, Amherst: GLSA.
- Lasnik, Howard & Mamoru Saito. 1992. *Move a. Conditions on its applications and output*. Cambridge, MA: MIT Press.
- Lebeaux, David. 1988. *Language acquisition and the form of grammar*. Amherst: University of Massachusetts, Amherst. (Doctoral dissertation).
- Lebeaux, David. 1991. Relative clauses, licensing and the nature of the derivation. In Susan Rothstein (ed.), *Syntax and Semantics 25: Perspectives on phrase structure*, 209-239. New York: Academic Press.
- Lin, Jingxia. 2015. Adverbial Clauses. In James D. Wright (ed.), *International Encyclopedia of the Social & Behavioral Sciences*, vol 1, 185-188. Oxford: Elsevier.
- Lindahl, Filippa. 2017. *Extraction from relative clauses in Swedish*. Gothenburg: University of Gothenburg. (Doctoral dissertation.)
- Lobo, Maria. 2002. On the structural position of non-peripheral adjunct clauses. *Journal of Portuguese Linguistics* 1(1). 83-118. DOI: <http://doi.org/10.5334/jpl.49>.
- Longobardi. 1985. Connectedness and island constraints. In Jacqueline Guéron, Hans-Georg Obenauer, Jean-Yves Pollock (eds.), *Grammatical Representation*, 169-85. Dordrecht: Foris Publications.
- Manzini, Maria Rita. 1992. *Locality: A theory and some of its empirical consequences*. Cambridge, MA: MIT Press.
- McCloskey, James. 2006. Resumption. In Martin Everaert and Henk van Riemsdijk (eds.), *Blackwell companion to syntax*, 94-115. Oxford: Blackwell.
- McCloskey, James, 2017. New thoughts on old questions: Resumption in Irish. In Jason Ostrove, Ruth Kramer, & Joseph Sabbagh (eds.), *Asking the right questions: Essays in honor of Sandra Chung*, 81-102. UC Santa Cruz.
- Merchant, Jason. 2001. *The syntax of silence*. Oxford: Oxford University Press.
- Michel, Dan & Grant Goodall. 2012. Finiteness and the nature of island constraints. In Nobu Goto, Koichi Otaki, Atsushi Sato & Kensuke Takita (eds.), *Proceedings of GLOW in Asia IX 2012: The Main Session*, 187-197. Japan: Mie University.

- Moro, Andrea. 2000. *Dynamic antisymmetry*. Cambridge, MA: MIT Press.
- Müller, Christiane. 2017. Extraction from adjunct islands in Swedish. *Norsk Lingvistisk Tidsskrift* 35. 67-85.
- Müller, Christiane. 2019. *Permeable islands A contrastive study of Swedish and English adjunct clause extractions*. Lund: Lund University. (Doctoral dissertation.)
- Müller, Gereon. 2010. On deriving CED effects from the PIC. *Linguistic Inquiry* 41(1). 35-82.
- Narita, Hiroki. 2011. *Phasing in Full Interpretation*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation.)
- Nunes, Jairo. 1995. The copy theory of movement and linearization of chains in the minimalist program. Maryland, MD: University of Maryland. (Doctoral dissertation.)
- Nunes, Jairo. 2001. Sideward movement. *Linguistic Inquiry* 31. 303-344.
- Nunes, Jairo & Juan Uriagereka. 2000. Cyclicity and extraction domain. *Syntax* 3. 20-43.
- Oseki, Yohei. 2015. Eliminating Pair-Merge. In Ulrike Steindl, Thomas Borer, Huilin Fang, Alfredo García Pardo, Peter Guekguezian, Brian Hsu, Charlie O'Hara, and Iris Chuoying Ouyang (eds.), *Proceedings of the 32nd West Coast Conference on Formal Linguistics*, 303-312. Somerville, MA: Cascadia Proceedings Project.
- Pesetsky, David. 1987. *Wh*-in-situ: Movement and unselective binding. In Eric J. Reuland & Alice ter Meulen (eds.), *The representation of (in)definiteness*, 98-129. Cambridge, MA: MIT Press.
- Phillips, Colin. 2006. The real-time status of island phenomena. *Language* 82(4). 795-823. DOI: <https://doi.org/10.1353/lan.2006.0217>.
- Phillips, Colin. 2011. Some arguments and non-arguments for reductionist accounts of syntactic phenomena. *Language and Cognitive Processes* 26. 1-32.
- Phillips, Colin. 2013. On the nature of island constraints I: Language processing and reductionist accounts. In *Experimental Syntax and Island Effects*, eds. Jon Sprouse and Norbert Hornstein. Cambridge: Cambridge University Press, 64-108.
- Polinsky, Maria, Lauren Eby Clemens, Adam Milton Morgan, Ming Xiang, and Dustin Heestand. 2013. Resumption in English. In Jon Sprouse and Norbert Hornstein (eds.), *Experimental Syntax and Island Effects*, 341-360. Cambridge: Cambridge University Press.
- Postal, Paul M. 2001. Parasitic and pseudo parasitic gaps. In Peter Culicover & Paul M. Postal (eds.), *Parasitic Gaps*, 253-313. Cambridge, MA: MIT Press.
- Prince, Ellen. 1990. Syntax and discourse: a look at resumptive pronouns. In Kira Hall, Jean-Pierre Koenig, Michael Meachan, Sondra Reinman & Laurel A. Sutton (eds.), *Proceedings of the 16th Berkeley Linguistics Society (BLS 16)*, 482-498. Somerville,

- MA: Cascadilla Press. DOI: <http://dx.doi.org/10.3765/bls.v16i0.1719>.
- Rackowski, Andrea & Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36(4). 565-599. DOI: <https://doi.org/10.1162/002438905774464368>.
- Ramchand, Gillian. 2008. *Verb meaning and the lexicon: a first- phase syntax*. Cambridge: Cambridge University Press.
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.
- Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of Grammar. Handbook in Generative Syntax*, 281-337. Dordrecht: Kluwer.
- Rizzi, Luigi. 2001. On the position “Int(errogative)” in the left periphery of the clause. In Guglielmo Cinque & Giampaolo Salvi (eds.), *Current studies in Italian syntax: Essays offered to Lorenzo Renzi*, 287-296. Amsterdam: Elsevier.
- Rizzi, Luigi. 2004. Locality and the left periphery. In Adriana Belletti (ed.), *Structures and beyond: The cartography of syntactic structures*, vol. 3, 223-251. Oxford: Oxford University Press.
- Rizzi, Luigi. 2013. Locality. *Lingua* 130. 169-186.
- Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).
- Ross, John Robert. 1969. Guess who? In R. I. Binnick, Alice Davison, Georgia M. Green & J. L. Morgan (eds.), *Papers from the Fifth Regional Meeting of the Chicago Linguistics Society*. Chicago: Chicago Linguistics Society.
- Rottman, Isaac & Masaya Yoshida. 2013. Sluicing, Idioms and Island Repair. *Linguistic Inquiry*. 44(4). 651-668.
- Sabel, Joachim. 2002. A minimalist analysis of syntactic islands. *The Linguistic Review* 19. 271-315.
- Sawada, Miyuki & Richard K. Larson. 2004. Presupposition & root transforms in adjunct clauses. In Matthew Wolf & Keir Moulton (eds.), *Proceedings of North East Linguistic Society* 34. 517-528. Amherst, MA: GLSA.
- Schütze, Carson. 2016. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Berlin: Language Science Press.
- Schweikert, Walter. 2004. *The order of prepositional phrases in the structure of the clause*. Venice: Università Ca' Foscari Venezia. (Doctoral dissertation).
- Selkirk, Lisa. 2005. Comments on intonational phrasing. In Sonia Frota, Marina Vigario &

- Maria Joao Freitas (eds.), *Prosodies*, 11-58. Berlin: Mouton de Gruyter.
- Sheehan, Michelle. 2013. The resuscitation of the CED. In Seda Kan, Claire Moore-Cantwell & Robert Staubs (eds.), *Proceedings of the 40th Annual Meeting of the North East Linguistic Society*, 135-150. Amherst: GSLA.
- Shlonsky, Ur. 1992. Resumptive pronouns as a last resort. *Linguistic Inquiry* 23. 443–468.
- Snyder, William. 2000. An experimental investigation of syntactic satiation effects. *Linguistic Inquiry* 31(3). 575-582.
- Sportiche, Dominique. 2018. Somber prospects for Late Merger. *Linguistic Inquiry*, 1-18.
- Sprouse, Jon. 2007. *A program for experimental syntax: Finding the relationship between acceptability and grammatical knowledge*. College Park, MD: University of Maryland. (Doctoral dissertation.)
- Sprouse, Jon & Diogo Almeida. 2012. Assessing the reliability of textbook data in syntax: Adger's core syntax 1. *Journal of Linguistics* 48. 609–652. DOI: <http://dx.doi.org/10.1017/S0022226712000011>
- Sprouse, Jon, Ivano Caponigro, Ciro Greco, & Carlo Cecchetto. 2016. Experimental syntax and the cross linguistic variation of island effects in English and Italian. *Natural Language and Linguistic Theory* 34. 307-344.
- Sprouse, Jon, Matthew W. Wagers & Colin Phillips. 2012. A test of the relation between working memory capacity and syntactic island effects. *Language* 88. 82-123.
- Starke, Michal. 2001. *Move dissolves into merge: A theory of locality*. Geneva: University of Geneva (Doctoral dissertation.)
- Stepanov, Artur. 2000. The timing of adjunction. In Masako Hirotani, Andries Coetzee, Nancy Hall & Ji-Yung Kim (eds.), *NELS 30: Proceedings of the North East Linguistic Society*, 597-612. Amherst, MA: GLSA.
- Stepanov, Artur. 2001. Late adjunction and minimalist phrase structure. *Syntax* 4. 94-125.
- Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domains. *Syntax* 10(1). 80–126.
- Storoshenko, Dennis Ryan. 2016. Fantasy islands? Testing semantic constraints on *wh*-extraction. In Lindsay Hrac (ed.), *Proceedings of the 2016 annual conference of the Canadian Linguistic Association*.
- Stowell, Tim. 1981. *Origins of Phrase Structure*. Cambridge, MA: Massachusetts Institute of Technology (Doctoral dissertation.)
- Svenonius, Peter. 2010. Spatial P in English. In Guglielmo Cinque and Luigi Rizzi (eds.), *Mapping spatial PPs: The cartography of syntactic structures, vol.6*, 127–161. New York: Oxford University Press.

- Szabolcsi, Anna. 2006. Strong vs. weak islands. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, 479-531. Oxford: Blackwell.
- Szabolcsi, Anna & Marcel den Dikken. 2003. Islands. In Lisa Cheng & Rint Sybesma (eds.), *The Second Glot International State-of-the-Article Book: The Latest in Linguistics*, 213-240. Berlin, Boston: De Gruyter Mouton. DOI: <https://doi.org/10.1515/9783110890952.213>.
- Szabolcsi, Anna & Terje Lohndal. 2017. Strong vs. Weak Islands. In Martin Everaert & Henk van Riemsdijk (eds.), *The Wiley Blackwell Companion to Syntax, Second Edition*, 1-51. New York: John Wiley and Sons Inc.
- Szabolcsi, Anna & Frans Zwarts. 1993. Weak islands and an algebraic semantics of scope taking. *Natural Language Semantics* 1(3). 235-284.
- Takahashi, Daiko. 1994. *Minimality of movement*. Storrs: University of Connecticut. (Doctoral dissertation.)
- Tanaka, Misako. 2015. *Scoping out of adjuncts. Evidence for the parallelism between QR and wh-movement*. London: University College London. (Doctoral thesis.)
- Torrego, Esther. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15. 103-129.
- Truswell, Robert. 2007a. Extraction from adjuncts and the structure of events. *Lingua* 117. 1355-77.
- Truswell, Robert. 2007b. Tense, events and extraction from adjuncts. In Malcolm Elliott (ed.), *Proceedings of the 43rd Chicago Linguistics Society*, 136-151. Chicago: Chicago Linguistic Society.
- Truswell, Robert. 2011. *Events, phrases and questions*. Oxford: Oxford University Press.
- Tucker, Matthew A., Ali Idrissi, Jon Sprouse, & Diogo Almeida. 2019. Resumption Ameliorates Different Islands Differentially: Acceptability Data from Modern Standard Arabic. In Amel Khalfaoui & Matthew A. Tucker (eds.), *Perspectives on Arabic Linguistics* 30, 159-193. Stony Brook, NY: John Benjamins.
- Uriagereka, Juan 1999. Multiple Spell-Out. In Samuel David Epstein & Norbert Hornstein (eds.), *Working minimalism*, 251-282. Cambridge, MA: MIT Press.
- Uriagereka, Juan. 2011. *Spell-Out and the Minimalist Program*. Oxford: Oxford University Press.
- Valmala, Vidal. 2009. On the position of central adverbial clauses. *Anuario del Seminario de Filología Vasca* 43(1-2). 951-970.
- Vendler, Zeno. 1957. Verbs and times. *Philosophical Review* 66. 143-60.
- Wagers, Matthew W., & Colin Phillips. 2009. Multiple dependencies and the role of the

grammar in real-time comprehension. *Journal of Linguistics* 45. 395–433.

Warren, Tessa & Edward Gibson. 2002. The influence of referential processing on sentence complexity. *Cognition* 85. 79–112. DOI: [http://dx.doi.org/10.1016/S0010-0277\(02\)00087-2](http://dx.doi.org/10.1016/S0010-0277(02)00087-2)

Wiklund, Anna-Lena. 2007. *The syntax of tenselessness: Tense/mood/aspect-agreeing infinitivals* (Studies in Generative Grammar 92). Berlin/New York: Mouton de Gruyter.

Appendix A

Materials of Experiment A

DOPO:

- 1) Quali dolci ha mangiato Luca dopo aver rubato?
- 2) Quali dolci Luca ha mangiato dopo averli rubati?
- 3) Quali dolci Luca li ha mangiati dopo aver rubato?
- 4) Quali dolci Luca li ha mangiati dopo averli rubati?
- 5) Quali dolci Sara dice che Luca ha mangiato dopo aver rubato?
- 6) Quali dolci Sara dice che Luca ha mangiato dopo averli rubati?
- 7) Quali dolci Sara dice che Luca li ha mangiati dopo aver rubato?
- 8) Quali dolci Sara dice che Luca li ha mangiati dopo averli rubati?
- 9) Quali dolci Luca è scappato dopo aver rubato?
- 10) Quali dolci Luca è scappato dopo averli rubati?
- 11) Quali dolci Sara dice che Luca è scappato dopo aver rubato?
- 12) Quali dolci Sara dice che Luca è scappato dopo averli rubati?

SENZA:

- 1) Quale ragazzo Silvia ha guardato senza salutare?
- 2) Quale ragazzo Silvia ha guardato senza salutarlo?
- 3) Quale ragazzo Silvia lo ha guardato senza salutare?
- 4) Quale ragazzo Silvia lo ha guardato senza salutarlo?
- 5) Quale ragazzo tutti dicono che Silvia ha guardato senza salutare?
- 6) Quale ragazzo tutti dicono che Silvia ha guardato senza salutarlo?
- 7) Quale ragazzo tutti dicono che Silvia lo ha guardato senza salutare?
- 8) Quale ragazzo tutti dicono che Silvia lo ha guardato senza salutarlo?
- 9) Quale ragazzo Silvia è partita/sparita senza salutare?
- 10) Quale ragazzo Silvia è partita/sparita senza salutarlo?

- 11) Quale ragazzo tutti dicono che Silvia è partita/sparita senza salutare?
- 12) Quale ragazzo tutti dicono che Silvia è partita/sparita senza salutarlo?

PRIMA:

- 1) Quale libro Anna ha cercato per mesi prima di trovare?
- 2) Quale libro Anna ha cercato per mesi prima di trovarlo?
- 3) Quale libro Anna lo ha cercato per mesi prima di trovare?
- 4) Quale libro Anna lo ha cercato per mesi prima di trovarlo?
- 5) Quale libro Eva dice che Anna ha cercato per mesi prima di trovare?
- 6) Quale libro Eva dice che Anna ha cercato per mesi prima di trovarlo?
- 7) Quale libro Eva dice che Anna lo ha cercato per mesi prima di trovare?
- 8) Quale libro Eva dice che Anna lo ha cercato per mesi prima di trovarlo?
- 9) Quale libro Anna se ne è andata/si è preoccupata prima di trovare?
- 10) Quale libro Anna se ne è andata/si è preoccupata prima di trovarlo?
- 11) Quale libro Eva dice che Anna se ne è andata/si è preoccupata prima di trovare?
- 12) Quale libro Eva dice che Anna se ne è andata/si è preoccupata prima di trovarlo?

FILLER (grammatical):

- 1) Elisa ha mangiato un panino a pranzo.
- 2) Noemi è andata in vacanza a Parigi per una settimana.
- 3) Chi ha visto Maria dal dottore?
- 4) Come ha fatto Sara ad aggiustare il lampadario?
- 5) Quando avrai l'esame della patente?
- 6) Le ha regalato un anello di diamanti!
- 7) Andrò in Africa in autunno.
- 8) Silvia ha incontrato Lara in giardino questa mattina.
- 9) Hai mai bevuto questo vino?
- 10) Non mi avevano mai detto che nuotare dopo pranzo fosse pericoloso.
- 11) Cosa pensi che abbiano comprato i ragazzi al supermercato?

- 12) Luca ha detto al nonno che è stato bocciato all'esame di francese.
- 13) Tutti credono di aver visto Gianni e Maria che si baciavano
- 14) Hai mai pensato che Franco sarebbe un attore perfetto?
- 15) Quando pensi di andare in vacanza quest'anno?
- 16) Roberto sostiene che Emma si trasferirà in Costa Rica.
- 17) Lisa ti ha confessato di aver venduto la collana che le avevi regalato?
- 18) A quante persone hai detto che Piero vuole farsi prete?

FILLER (ungrammatical):

- 19) Ada e Cristina è andate al parco questa mattina presto.
- 20) Cosa hai mangiato un panino con prosciutto e?
- 21) Anna sarebbe partita per la Cina stamattina.
- 22) Ho saputo niente.
- 23) Come hai fatto quale test?
- 24) Cesare se ne sono andati di casa.
- 25) Fiona ha fatto piangere il.
- 26) Dove hai comprato quel vestito nel negozio?
- 27) Tutti ha detto che saresti bellissimo vestito così.
- 28) Perché non pensi a nessuno tranne che?
- 29) Di chi sostieni che essere meno intelligente sarebbe una vergogna?
- 30) Maria mi ha giurato che ha saputo niente da te.
- 31) Quale parte del film ti sei addormentata guardando?
- 32) Avresti mai pensato che Gianni regalasse?
- 33) E chi tutti i ragazzi hanno guardato Sara e?
- 34) Quanti bambini non dicevi che volevi da giovane?
- 35) Sara e Giacomo hanno pensato che tu fossi partita senza.
- 36) Laura ha preparato la torta e ha messa in frigorifero tutta la notte.

Appendix B

Materials of Experiment B

DOPO:

- 1) Quali dolci ha mangiato Luca dopo aver rubato?
- 2) Quali dolci ha mangiato Luca dopo che ha rubato?
- 3) Quali dolci Luca è scappato dopo aver rubato?
- 4) Quali dolci Luca è scappato dopo che ha rubato?
- 5) Quale esame Elisa ha registrato dopo aver superato?
- 6) Quale esame Elisa ha registrato dopo che aveva superato?
- 7) Quale esame Elisa ha festeggiato dopo aver superato?
- 8) Quale esame Elisa ha festeggiato dopo che aveva superato?

SENZA:

- 1) Quale ragazzo Silvia ha guardato senza salutare?
- 2) Quale ragazzo Silvia ha guardato senza che salutasse?
- 3) Quale ragazzo Silvia è partita senza salutare?
- 4) Quale ragazzo Silvia è partita senza che salutasse?
- 5) Quale film Silvia ha recensito senza aver visto?
- 6) Quale film Silvia ha recensito senza che vedesse?
- 7) Quale film Silvia è uscita senza aver visto?
- 8) Quale film Silvia è uscita senza che vedesse?
- 9) Quale mistero Andrea ha indagato senza risolvere?
- 10) Quale mistero Andrea ha indagato senza che risolvesse?
- 11) Quale mistero Andrea si è preoccupato senza risolvere?
- 12) Quale mistero Andrea si è preoccupato senza che risolvesse?

PRIMA:

- 1) Quale quaderno Anna ha cercato per mesi prima di trovare?
- 2) Quale quaderno Anna ha cercato per mesi prima che trovasse?
- 3) Quale quaderno Anna ha riflettuto prima di trovare?
- 4) Quale quaderno Anna ha riflettuto prima che trovasse?
- 5) Quale libro Gianni ha guardato a lungo prima di comprare?
- 6) Quale libro Gianni ha guardato a lungo prima che comprasse?
- 7) Quale libro Gianni è passato in banca prima di comprare?
- 8) Quale libro Gianni è passato in banca prima che comprasse?

Appendix C

Materials of Experiment C

PER

- (1)
 - a. Quale libro Sergio ha acquistato per leggere?
 - b. Quale libro Sergio ha acquistato per leggerlo?
 - c. Quale libro Sergio è tornato a casa per leggere?
 - d. Quale libro Sergio è tornato a casa per leggerlo?
- (2)
 - a. Quale giornale Simone ha comprato per leggere?
 - b. Quale giornale Simone ha comprato per leggerlo?
 - c. Quale giornale Simone è tornato a casa per leggere?
 - d. Quale giornale Simone è tornato a casa per leggerlo?
- (3)
 - a. Quale città Maria ha esplorato per conoscere?
 - b. Quale città Maria ha esplorato per conoscerla?
 - c. Quale città Maria ha preso il treno per visitare?
 - d. Quale città Maria ha preso il treno per visitarla?
- (4)
 - a. Quale città Anna ha visitato per conoscere?
 - b. Quale città Anna ha visitato per conoscerla?
 - c. Quale città Anna ha preso il treno per vedere?
 - d. Quale città Anna ha preso il treno per vederla?

SENZA

- (1)
 - a. Quale ragazza Marco ha guardato senza salutare?
 - b. Quale ragazza Marco ha guardato senza salutarla?
 - c. Quale ragazza Marco è partito senza salutare?
 - d. Quale ragazza Marco è partito senza salutarla?
- (2)
 - a. Quale ragazzo Elisa ha guardato senza salutare?
 - b. Quale ragazzo Elisa ha guardato senza salutarlo?
 - c. Quale ragazzo Elisa è partita senza salutare?
 - d. Quale ragazzo Elisa è partita senza salutarlo?
- (3)
 - a. Quali ragazze Elena ha guardato senza salutare?
 - b. Quali ragazze Elena ha guardato senza salutarle?

- c. Quali ragazze Elena è scappata senza salutare?
- d. Quali ragazze Elena è scappata senza salutarle?
- (4) a. Quali ragazzi hai guardato senza salutare?
- b. Quali ragazzi hai guardato senza salutarli?
- c. Quali ragazzi sei partita senza salutare?
- d. Quali ragazzi sei partita senza salutarli?

PRIMA

- (1) a. Quale esame Giulia temeva prima di superare?
- b. Quale esame Giulia temeva prima di superarlo?
- c. Quale esame Giulia si è preoccupata prima di superare?
- d. Quale esame Giulia si è preoccupata prima di superarlo?
- (2) a. Quale esame Paolo temeva prima di superare?
- b. Quale esame Paolo temeva prima di superarlo?
- c. Quale esame Paolo si è preoccupato prima di superare?
- d. Quale esame Paolo si è preoccupato prima di superarlo?
- (3) a. Quale libro Gianni ha cercato a lungo prima di trovare?
- b. Quale libro Gianni ha cercato a lungo prima di trovarlo?
- c. Quale libro Gianni ha riflettuto prima di trovare?
- d. Quale libro Gianni ha riflettuto prima di trovarlo?
- (4) a. Quali fotografie Francesco ha cercato a lungo prima di trovare?
- b. Quali fotografie Francesco ha cercato a lungo prima di trovarle?
- c. Quali fotografie Francesco ha riflettuto prima di trovare?
- d. Quali fotografie Francesco ha riflettuto prima di trovarle?

DOPO

- (1) a. Quali dolci Luca ha mangiato dopo aver rubato?
- b. Quali dolci Luca ha mangiato dopo averli rubati?
- c. Quali dolci Luca è scappato dopo aver rubato?
- d. Quali dolci Luca è scappato dopo averli rubati?
- (2) a. Quale dolce Noemi ha mangiato dopo aver rubato?
- b. Quale dolce Noemi ha mangiato dopo averlo rubato?
- c. Quale dolce Noemi è scappata dopo aver rubato?

- d. Quale dolce Noemi è scappata dopo averlo rubato?
- (3) a. Quale ragazzo Nora ha salutato dopo aver baciato?
b. Quale ragazzo Nora ha salutato dopo averlo baciato?
c. Quale ragazzo Nora se ne è andata dopo aver baciato?
d. Quale ragazzo Nora se ne è andata dopo averlo baciato?
- (4) a. Quale ragazza Piero ha salutato dopo aver baciato?
b. Quale ragazza Piero ha salutato dopo averla baciata?
c. Quale ragazza Piero è corso via dopo aver baciato?
d. Quale ragazza Piero è corso via dopo averla baciata?

FINCHÈ

- (1) a. Quale film ti interessava finché non hai guardato?
b. Quale film ti interessava finché non l'hai guardato?
c. Quale film hai insistito finché non ho guardato?
d. Quale film hai insistito finché non l'ho guardato?
- (2) a. Quale città ti interessava finché non hai visitato?
b. Quale città ti interessava finché non l'hai visitata?
c. Quale città hai insistito finché non ho visitato?
d. Quale città hai insistito finché non l'ho visitata?
- (3) a. Quale libro ti interessava finché non hai letto?
b. Quale libro ti interessava finché non l'hai letto?
c. Quale libro hai insistito finché non ho letto?
d. Quale libro hai insistito finché non l'ho letto?
- (4) a. Quale corso ti interessava finché non hai seguito?
b. Quale corso ti interessava finché non l'hai seguito?
c. Quale corso hai insistito finché non ho seguito?
d. Quale corso hai insistito finché non l'ho seguito?

PERCHÈ

- (1) a. Quali scarpe hai sistemato perché avevi rovinato?
b. Quali scarpe hai sistemato perché le avevi rovinate?
c. Quali scarpe ti sei arrabbiata perché ho rovinato?
d. Quali scarpe ti sei arrabbiata perché le ho rovinate?

- (2)
 - a. Quale libro hai sistemato perché avevi rovinato?
 - b. Quale libro hai sistemato perché lo avevi rovinato?
 - c. Quale libro ti sei arrabbiato perché ho rovinato?
 - d. Quale libro ti sei arrabbiato perché l'ho rovinato?
- (3)
 - a. Quale quadro hai sistemato perché avevi rovinato?
 - b. Quale quadro hai sistemato perché lo avevi rovinato?
 - c. Quale quadro ti sei arrabbiato perché ho rovinato?
 - d. Quale quadro ti sei arrabbiato perché l'ho rovinato?
- (4)
 - a. Quale borsa hai sistemato perché avevi rovinato?
 - b. Quale borsa hai sistemato perché l'avevi rovinata?
 - c. Quale borsa ti sei arrabbiata perché ho rovinato?
 - d. Quale borsa ti sei arrabbiata perché l'ho rovinata?

Studente: Chiara Dal Farra

matricola: 835035

Dottorato: Lingue, culture e società moderne e Scienze del Linguaggio

Ciclo: 32°

Titolo della tesi: To be or not to be an island The status of adjuncts

Abstract:

In questa tesi indago la possibilità di estrazione dagli aggiunti preposizionali in italiano, sia da un punto di vista teorico che sperimentale. Dimostro che a volte l'estrazione da questi domini è permessa, sia in italiano che cross-linguisticamente, e che è influenzata da vari fattori quali la posizione dell'aggiunto nella struttura e il livello di integrazione sintattica con la frase principale. Dimostro che a seconda degli aggiunti si ottengono risultati diversi nella valutazione dell'estrazione. Propongo quindi che non si possano considerare una classe sintattica uniforme. Propongo un'analisi dell'estrazione basata su (i) punto di attaccamento dell'aggiunto nella struttura e (ii) possibilità di formare un macroevento. Il primo punto è il criterio che ci permette anche di spiegare la differenza tra gli aggiunti, e serve nella creazione dei macroeventi, possibile solo con aggiunti bassi, ma non con quelli alti. Propongo che gli aggiunti trasparenti vengano rianalizzati nella struttura come dei complementi. Infine affronto la variabilità tra i parlanti, dimostrando come l'abilità di formare un unico evento possa spiegare questo dato.

In this thesis I focus on the possibility of extraction from prepositional adjuncts in Italian, from both a theoretical and an experimental point of view. I show that extraction from these domains is sometimes licensed, both in Italian and cross-linguistically, and that it is deeply influenced by many factors, such as the position of the adjunct and its syntactic integration with the main clause. I show that different adjuncts behave differently with respect to extraction phenomena. I claim that they are not a uniform syntactic class. I propose an analysis of extraction, based on: (i) the point of merger of adjuncts in the structure, and (ii) the possibility of forming a macroevent. I claim that the former is a crucial criterion able to explain the different behaviours of adjuncts, and that it is also needed for the creation of a macroevent. I propose that transparent adjuncts are reanalysed as complements. I discuss also inter-speaker variability, and show that the creation of a single macroevent is able to account for these data.

Firma dello studente

Chiara Dal Farra