Basque-Spanish Language Contact

An empirical study on word order in interrogatives

Doctoral thesis for obtaining the academic degree Doctor of Philosophy (Dr. phil.)

submitted by
Dold, Simon

at the

Universität Konstanz

Faculty of Humanities
Department of Linguistics

Konstanz, 2018
Date of the oral examination: May 29, 2018

First Referee: Prof. Dr. Georg A. Kaiser
Second Referee: Dr. Ricardo Etxepare
Third Referee: Prof. Dr. Tanja Kupisch
Acknowledgment

First of all, I want to thank my first supervisor and Doktorvater Georg Kaiser for his continued support in this project, from the very start to its completion. I greatly appreciate your trust in me and I am very grateful that you reawakened a dormant interest in the Basque culture and language in me, and that you encouraged me to take a professional interest in it. Thank you for the freedom and flexibility you offered me in my research, writing, and teaching. I am very grateful to my second supervisor, Ricardo Etxepare. Thank you Ricardo for your support and patience guiding me through this project. It was a great opportunity to work with you in IKER in Bayonne. Every one of our conversations was filled with inspiring ideas as well as vital input. Sharing your broad theoretical knowledge in all fields of linguistics helped me find my way in this project. Thank you both! Thank you, Tanja, for agreeing to be on my committee.

To my amazing colleagues, who I am very grateful to have worked with. Thank you Carmen, Katharina, and Marieke for supporting me, cheering me up, discussing my work, and always having time for a short chat or a cup of coffee. Thank you, Michael and Janina, for your support and the many fruitful discussions about different subjects concerning my work. Thank you, Niklas, for your help with the statistics, R, and the great times we shared training for our various sporting events that we needed to clear our heads from all this studying. A special thank also goes to the staff at IKER in Bayonne. Your help with the French university’s bureaucracy was invaluable. To my fellow investigators in Bayonne, especially Aritz, Maitena and Urtzi, thank you for a great year with you all in Bayonne!

I want to thank the 181 participants from all over Spain who were willing to fill out my questionnaires. Without their help and support, this work would not have been possible.

A big thank you goes to all my friends who have been here for me throughout this entire time. I thank each and every one of you for your continued incredible support, motivation, and – when necessary – encouragement. A special thanks goes to the Bodensee underwater rugby team for sharing my passion for this great sport. To you Heike, thank you for your extraordinary help and support, which even included airmail care and motivation packages! Thank you for your company, your great hospitality, and simply for being an amazing friend. I wish you the very best for your new beginnings.

My greatest thanks goes to my family and my wonderful wife-to-be Shanti. Thank you my dear parents, sisters (-in-law), and brothers (-in-law)! You were always there for me, and your interest in what I was doing helped keep me going through the difficult times. I am so grateful for your support and help that, in so many different ways, allows me to go on. Shanti, thank you so much for believing in me!! Without your constant love and support, this work would have never been completed. Quite literally, as I owe the readability of this work to your proofreading. Thank you for being such an amazing and wonderful woman!
Abstract

Basque and Spanish are two typological very different languages. Nevertheless, they show surprisingly similar word order restrictions in wh-interrogatives. Both languages exhibit obligatory wh-fronting in an unmarked wh-interrogative and they usually require wh-phrase-verb adjacency, which can only be interrupted by clitics or some select adverbs. The two languages differ, however, regard to two types of interrogatives: complex and multiple wh-questions. In this work, I will mainly discuss the first type. In Spanish, complex wh-sentences allow an intervening constituent between the wh-phrase and the verb, interrupting the otherwise obligatory adjacency. It seems that Spanish offers the possibility of analyzing these complex wh-phrases as clitic-left dislocations. In Basque, this is never possible, because the wh-phrase-verb adjacency has to be retained, no matter the complexity of the wh-phrase.

In this thesis, I examine differences between monolingual Spanish and bilingual Basque-Spanish speakers regarding the acceptability of Spanish complex wh-questions without wh-phrase-verb adjacency. In two studies conducted in 2016, I observed that bilingual speakers accept such wh-questions with an intervening constituent much less than monolinguals. I propose that this difference has two main sources: a different underlying structure in wh-questions in the two languages despite the superficial resemblance and a general preference for common structures, as well as with a lack of input during bilingual language acquisition.

I elaborate a theoretical model based on Cable’s (2010) Q(uestion)-particle (QP) and QP-movement approach. In this approach, the variation in question formation is explained by a different core parameter setting of the Q-particle, namely that Spanish is a Q-projection language and Basque a Q-adjunction language. According to Cable (2010), wh-fronting is not the result of a special relationship between the wh-phrase and a higher C-head, but of one between the Q-particle and a C-head. The wh-phrase only appears sentence initial because it is ‘dragged’ along by this Q-particle, which takes the phrase containing the wh-word as a complement. This is the case in Spanish. Basque differs from Spanish in that the Q-particle does not take the phrase containing the wh-word as a complement, but is only adjoined to it. This allows the Q-particle-to move to the C-head alone, without taking the whole wh-phrase with it. The reason that Basque nevertheless exhibit wh-fronting lies in an independent property: a preverbal focus feature that has to be satisfy by an adequate phrase. In a positive wh-question, the wh-phrase normally suffices for the job and therefore moves to the preverbal position.

An important observation in a bilingual language contact situation is the following: if one language has possibilities A and B to analyze a certain structure and a second language only allows option A, bilingual speakers tend to primarily use option A in both languages, especially during acquisition. Applied to the case at hand, option A would be the analysis of complex wh-phrases with wh-phrase-verb adjacency and option B the alternative analysis as clitic-left dislocations. In Spanish, both options are available, in Basque only the first. Therefore, the child that acquires Basque and Spanish intrinsically prefers the option with adjacency and, in addition, the surrounding adults probably do too. This means, the child simply does not hear enough positive evidence to establish the exception rule for the complex wh-phrases, in contrast to monolingual children, who do receive enough evidence.

The study presented in this work is another example that an intense language contact can influence the involved languages on a level deeper than just the lexical. It shows that syntactic structures can be affected, even in typological very distinct languages like Basque and Spanish.
Zusammenfassung


Allgemein ist zu beobachten, dass bilinguale Sprecher*innen in beiden Sprachen, wenn möglich, ähnliche Konstruktionen benutzen. Das heißt, wenn eine Sprache zwei Optionen A und B zur Analyse einer Konstruktion ermöglicht, die zweite Sprache jedoch nur A erlaubt, dann tendieren bilinguale Sprecher*innen dazu, in beiden Sprachen vorwiegend Option A zu benutzen. In dem betrachteten kon-
kreten Fall entspricht A der Konstruktion mit Adjazenz zwischen der *wh*-Phrase und dem Verb und B der alternativen Konstruktion, der die Analyse der komplexen *wh*-Phrase als klitische Linksdisklokation zu Grunde liegt. Im Spanischen sind beide Optionen möglich, im Baskischen nur die erste. Dies plausibilisiert die Beobachtung, das bilingual aufwachsende Kinder die Adjazenzanalyse bevorzugen. Die erhobenen Daten legen nahe, dass diese Tendenz auch bei erwachsenen bilingualen Sprecher*innen existiert. Dies wiederum wirkt sich insofern auf den Spracherwerb aus, als dass bilinguale Kinder die entsprechende Analyse nicht lernen.

Die in dieser Arbeit präsentierte Studie ist ein weiterer Beleg dafür, dass intensiver Sprachkontakt die beteiligten Sprachen auf einer strukturellen Ebene jenseits des Lexikons beeinflussen kann. Sie zeigt, dass syntaktische Strukturen auch bei typologisch sehr unterschiedlichen Sprachen beeinflusst werden können.
Table of Contents

List of Abbreviations iv
List of Figures vi
List of Tables vii

Chapter 1  Introduction 1
  1.1 What this is all about ........................................... 1
  1.2 Roadmap ......................................................... 5
  1.3 Conventions for glosses and translations ...................... 6

Chapter 2  Basque-Spanish language contact – a geographical and historical introduction 8
  2.1 A brief note on geographies and naming ........................ 9
  2.2 A brief note on Basque dialects ................................ 10
  2.3 A brief note on the history of Basque-Spanish contact ........ 11

Chapter 3  Word order, questions, and focus in Spanish 15
  3.1 On Spanish word-order ........................................ 15
    3.1.1 Focus and Topic – a short description ...................... 17
    3.1.2 CLLD, CLRD and HTLD .................................. 19
    3.1.3 Focus-Fronting ............................................. 25
    3.1.4 External merge and internal merge ....................... 27
  3.2 On Spanish wh-questions ...................................... 30
    3.2.1 wh-questions – main characteristics ....................... 31
    3.2.2 Wh-phrase-verb adjacency – exceptions ................... 36
      3.2.2.1 Adverbs ............................................... 36
      3.2.2.2 wh-phrases ........................................... 37
      3.2.2.3 Complex wh-phrases ................................ 39
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1.3 Contact phenomena in Basque and Spanish in the Basque speaking area</td>
<td>131</td>
</tr>
<tr>
<td>6.1.4 whSV(X) – a modern phenomenon?</td>
<td>136</td>
</tr>
<tr>
<td>6.2 On the role of language acquisition in a language contact situation</td>
<td>137</td>
</tr>
<tr>
<td>6.2.1 Some general remarks on acquisition</td>
<td>137</td>
</tr>
<tr>
<td>6.2.2 wh-interrogatives in Basque-Spanish bilingual acquisition</td>
<td>139</td>
</tr>
<tr>
<td>Chapter 7 Empirical study on wh-interrogatives</td>
<td>146</td>
</tr>
<tr>
<td>7.1 Methodology</td>
<td>146</td>
</tr>
<tr>
<td>7.2 SolQ – the participants</td>
<td>150</td>
</tr>
<tr>
<td>7.3 Questionnaires</td>
<td>154</td>
</tr>
<tr>
<td>7.3.1 Pilot-Questionnaire – PilotQ</td>
<td>155</td>
</tr>
<tr>
<td>7.3.2 Questionnaire 1 – Q1</td>
<td>156</td>
</tr>
<tr>
<td>7.3.3 Questionnaire 2 – Q2</td>
<td>157</td>
</tr>
<tr>
<td>7.4 Results</td>
<td>160</td>
</tr>
<tr>
<td>7.4.1 Results – Q1</td>
<td>162</td>
</tr>
<tr>
<td>7.4.2 Results – Q2</td>
<td>166</td>
</tr>
<tr>
<td>7.4.3 Analysis Q1 + Q2</td>
<td>171</td>
</tr>
<tr>
<td>Chapter 8 Complex wh-phrases revisited</td>
<td>177</td>
</tr>
<tr>
<td>Chapter 9 Conclusion and outlook</td>
<td>190</td>
</tr>
<tr>
<td>Bibliography</td>
<td>195</td>
</tr>
<tr>
<td>List of Appendices</td>
<td>207</td>
</tr>
</tbody>
</table>
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>First person</td>
</tr>
<tr>
<td>2</td>
<td>Second person</td>
</tr>
<tr>
<td>2L1</td>
<td>Bilinguales, simultaneous BLA</td>
</tr>
<tr>
<td>3</td>
<td>Third person</td>
</tr>
<tr>
<td>A</td>
<td>Accusative marker in Spanish transitive sentences</td>
</tr>
<tr>
<td>vP</td>
<td>little-vPhrase</td>
</tr>
<tr>
<td>ABS</td>
<td>Absolutive</td>
</tr>
<tr>
<td>ALL</td>
<td>Allative</td>
</tr>
<tr>
<td>All-BA</td>
<td>All speakers from the Basque speaking area, bilinguals and monolinguals</td>
</tr>
<tr>
<td>Ara</td>
<td>Araba</td>
</tr>
<tr>
<td>AspP</td>
<td>Aspect Phrase</td>
</tr>
<tr>
<td>ATop</td>
<td>Aboutness Topic</td>
</tr>
<tr>
<td>AUX</td>
<td>Auxiliary</td>
</tr>
<tr>
<td>Biz</td>
<td>Bizkaia, Province</td>
</tr>
<tr>
<td>BLA</td>
<td>Bilingual Language Acquisition</td>
</tr>
<tr>
<td>Bq.</td>
<td>Basque</td>
</tr>
<tr>
<td>BS-ce</td>
<td>Bilingual speakers, sequential BLA Basque - Spanish, child early</td>
</tr>
<tr>
<td>BS-cl</td>
<td>Bilingual speakers, sequential BLA Basque - Spanish, child late</td>
</tr>
<tr>
<td>CAE</td>
<td>Comunidad Autónoma de Euskadi</td>
</tr>
<tr>
<td>Cast.</td>
<td>Castilian</td>
</tr>
<tr>
<td>CdE</td>
<td>Corpus del Español</td>
</tr>
<tr>
<td>CFoc</td>
<td>Contrastive Focus</td>
</tr>
<tr>
<td>CI</td>
<td>Clitic</td>
</tr>
<tr>
<td>CLD</td>
<td>Doubled by a Clitic</td>
</tr>
<tr>
<td>CLLD</td>
<td>Clitic Left Dislocation</td>
</tr>
<tr>
<td>CLRD</td>
<td>Clitic Right Dislocation</td>
</tr>
<tr>
<td>CORPES</td>
<td>Corpus del Español del Siglo XXI</td>
</tr>
<tr>
<td>COSER</td>
<td>Corpus Oral y Sonoro del Español Rural</td>
</tr>
<tr>
<td>CP</td>
<td>Complementizer Phrase</td>
</tr>
<tr>
<td>CREA</td>
<td>Corpus de Referencia del Español Actual</td>
</tr>
<tr>
<td>d-linked</td>
<td>Discourse-linked</td>
</tr>
<tr>
<td>DAT</td>
<td>Dativ</td>
</tr>
<tr>
<td>DO</td>
<td>Direct Object</td>
</tr>
<tr>
<td>DOM</td>
<td>Differential Object Marking</td>
</tr>
<tr>
<td>DP</td>
<td>Determiner Phrase</td>
</tr>
<tr>
<td>DS</td>
<td>Discourse Shell</td>
</tr>
<tr>
<td>EPP</td>
<td>Extended Projection Principle</td>
</tr>
<tr>
<td>ERG</td>
<td>Ergative</td>
</tr>
<tr>
<td>FF</td>
<td>Focus Fronting</td>
</tr>
<tr>
<td>FinP</td>
<td>Finiteness Phrase</td>
</tr>
<tr>
<td>FocP</td>
<td>Focus Phrase</td>
</tr>
<tr>
<td>ForP</td>
<td>Force Phrase</td>
</tr>
<tr>
<td>Fr.</td>
<td>French</td>
</tr>
<tr>
<td>FUT</td>
<td>Futur</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive</td>
</tr>
<tr>
<td>Germ.</td>
<td>Germanic</td>
</tr>
<tr>
<td>Gip</td>
<td>Gipuzkoa</td>
</tr>
<tr>
<td>HTLD</td>
<td>Hanging Topic Left Dislocation</td>
</tr>
<tr>
<td>IFoc</td>
<td>Information Focus</td>
</tr>
<tr>
<td>IMPF</td>
<td>Imperfective</td>
</tr>
<tr>
<td>INF</td>
<td>Infinitiv</td>
</tr>
<tr>
<td>INSTR</td>
<td>Instrumental</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>English Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>IO</td>
<td>Indirect Object</td>
</tr>
<tr>
<td>It.</td>
<td>Italian</td>
</tr>
<tr>
<td>KWT</td>
<td>Kruskal-Wallice Rank Sum Test</td>
</tr>
<tr>
<td>L1</td>
<td>First Language</td>
</tr>
<tr>
<td>L2</td>
<td>Second Language</td>
</tr>
<tr>
<td>LD</td>
<td>Left dislocation</td>
</tr>
<tr>
<td>lmer</td>
<td>Linear Mixed Effects Regression Model</td>
</tr>
<tr>
<td>M-BA</td>
<td>Monolinguals, Basque speaking area</td>
</tr>
<tr>
<td>M-SP</td>
<td>Monolinguals, other parts of Spain</td>
</tr>
<tr>
<td>M-SP/BA</td>
<td>All monolinguals</td>
</tr>
<tr>
<td>med</td>
<td>Median</td>
</tr>
<tr>
<td>medi_{x}</td>
<td>Median_{item_{x}}</td>
</tr>
<tr>
<td>Nav</td>
<td>Navarre</td>
</tr>
<tr>
<td>NEG</td>
<td>Negation</td>
</tr>
<tr>
<td>NOM</td>
<td>Nominalizer</td>
</tr>
<tr>
<td>NPI</td>
<td>Negative Polarity Item</td>
</tr>
<tr>
<td>NSL</td>
<td>Null Subject Language</td>
</tr>
<tr>
<td>NSR</td>
<td>Nuclear-Stress-Rule</td>
</tr>
<tr>
<td>O</td>
<td>Object</td>
</tr>
<tr>
<td>PART</td>
<td>Participle</td>
</tr>
<tr>
<td>PF</td>
<td>Phonological Form</td>
</tr>
<tr>
<td>PilotQ</td>
<td>Pilot-Questionnaire</td>
</tr>
<tr>
<td>PL</td>
<td>Plural</td>
</tr>
<tr>
<td>PolIP</td>
<td>Polarity Phrase</td>
</tr>
<tr>
<td>POT</td>
<td>Potential, hypothetical</td>
</tr>
<tr>
<td>PP</td>
<td>Prepositional Phrase</td>
</tr>
<tr>
<td>PRTT</td>
<td>Partitive</td>
</tr>
<tr>
<td>Ptg.</td>
<td>Portuguese</td>
</tr>
<tr>
<td>Q</td>
<td>Q(uestion)-Particle / Q-Particle</td>
</tr>
<tr>
<td>Q1</td>
<td>Questionnaire 1</td>
</tr>
<tr>
<td>Q2</td>
<td>Questionnaire 2</td>
</tr>
<tr>
<td>QP</td>
<td>Question-Particle Phrase</td>
</tr>
<tr>
<td>QPIC</td>
<td>QP-Intervention Condition</td>
</tr>
<tr>
<td>REFL</td>
<td>Reflexiv pronoun</td>
</tr>
<tr>
<td>RL</td>
<td>Romance languages</td>
</tr>
<tr>
<td>Rom.</td>
<td>Romanian</td>
</tr>
<tr>
<td>RTop</td>
<td>Referential Topic</td>
</tr>
<tr>
<td>S</td>
<td>Subject</td>
</tr>
<tr>
<td>SB-cl</td>
<td>Bilingual speakers, sequential BLA Spanish - Basque, child late</td>
</tr>
<tr>
<td>SCO</td>
<td>Strong-Cross-Over effect</td>
</tr>
<tr>
<td>SG</td>
<td>Singular</td>
</tr>
<tr>
<td>SoLQ</td>
<td>Sociolinguistic Questionnaire</td>
</tr>
<tr>
<td>SP-ce</td>
<td>Bilingual speakers, sequential BLA Spanish - Basque, child early</td>
</tr>
<tr>
<td>Sp.</td>
<td>Spanish</td>
</tr>
<tr>
<td>SVI</td>
<td>Subject-Verb-Inversion</td>
</tr>
<tr>
<td>TopP</td>
<td>Topic Phrase</td>
</tr>
<tr>
<td>TP</td>
<td>Tense Phrase</td>
</tr>
<tr>
<td>V</td>
<td>Verb</td>
</tr>
<tr>
<td>VPiSH</td>
<td>VP-internal Subject Hypothesis</td>
</tr>
<tr>
<td>WCO</td>
<td>Weak-Cross-Over effect</td>
</tr>
<tr>
<td>XP_{wh}</td>
<td>XP containing the wh-word</td>
</tr>
<tr>
<td>YLB</td>
<td>Young Labourdin Basque</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Euskal Herria in Europe and its provinces ........................................... 9

6.1 Types of bilingualism by age and sequence of acquisition ..................... 138

7.1 Item_2.03 – as in Q2 ........................................................................ 149
7.2 Gender across speaker groups ......................................................... 151
7.3 Total age distribution ................................................................. 152
7.4 Age across speaker groups .......................................................... 153
7.5 Degrees of education ................................................................. 153
7.6 Degrees of education across speaker groups .................................... 153
7.7 Daily use of languages in percentage ............................................. 153
7.8 Language(s) spoken at home ........................................................ 154
7.9 Language(s) spoken with the family ............................................. 154
7.10 Comfort .................................................................................. 154
7.11 Overall participation – Q1 ......................................................... 156
7.12 Overall participation – Q2 .......................................................... 158
7.13 Item_2.03 – as in Q2 .................................................................... 158
7.14 Boxplot – Example .................................................................... 161
7.15 Item_1.09 – Acceptability across all speaker groups ..................... 162
7.16 Item_1.12 – Acceptability across all speaker groups .................... 162
7.17 Item_1.02 – Acceptability across all speaker groups .................... 163
7.18 Item_1.07 – Acceptability across all speaker groups .................... 163
7.19 Item_1.01 – Acceptability across all speaker groups .................... 163
7.20 Item_1.03 – Acceptability across all speaker groups .................... 163
7.21 Item_1.04 – Acceptability across all speaker groups .................... 164
7.22 Item_1.05 – Acceptability across all speaker groups .................... 164
7.23 Item_1.08 – Acceptability across all speaker groups .................... 164
7.24 Item_1.09 – Acceptability across all speaker groups .................... 164
7.25 Item_1.10 – Acceptability across all speaker groups .................... 164
7.26 Item_1.11 – Acceptability across all speaker groups .................... 164
7.27 Item_1.06 – Acceptability across all speaker groups .................... 165
7.28 Item_1.12 – Acceptability across all speaker groups .................... 165
7.29 Item_2.04 – Acceptability .......................................................... 167
7.30 Item_2.07 – Acceptability .......................................................... 168
7.31 Item_2.05 – Acceptability across all speaker groups .................... 169
7.32 Item_2.03 – Acceptability across all speaker groups .................... 170
7.33 Item_2.02 – Acceptability across all speaker groups .................... 171
7.34 Item_2.08 – Acceptability across all speaker groups .................... 171

8.1 Item_2.04 – Acceptability: All-BA vs. M-SP ................................. 181
8.2 Item_2.07 – Acceptability: All-BA vs. M-SP ................................. 181
List of Tables

3.1 Interpretive types of focus and topic (Cruschina 2012: 12) ................................. 18
5.1 Paradigm for Basque wh-based indefinites ......................................................... 110
6.1 Borrowing scale according to Thomason & Kaufman ................................. 126
6.2 Percentages of wh-interrogatives with simple and heavy wh-phrases .......... 143
7.1 Total number of participants per speaker group ................................................. 151
7.2 Lmer – Summary all groups for questions without adjacency ................ 172
7.3 Overview medians ......................................................................................... 173
7.4 pKWT values for [±clitic]-contrast across all speaker groups ................... 174
7.5 Lmer – summary acceptability and mean difference of DO vs. IO wh-phrase across all
speaker groups. .............................................................................................. 175
7.6 Lmer – [±human] and [±insertion] ................................................................. 175
8.1 Lmer – Summary all groups for questions without adjacency ................... 183
8.2 Overview medians ......................................................................................... 184
8.3 pKWT-values for [±clitic]-contrast for the monolingual speakers .......... 186
8.4 Lmer – summary acceptability of DO vs. IO wh-phrase for M-SP/BA and bilinguals 187
8.5 Lmer – [±human, ±insertion] across bilingual speakers ......................... 188
8.6 pKWT-values for [±clitic]-contrast across bilingual speaker ...................... 189
CHAPTER 1

Introduction

1.1 What this is all about

The Basque-Spanish language contact is as old as Spanish itself. From a linguistic point of view, it is special in several aspects and distinct to all other contact situations on the Iberian Peninsula. This is mainly due to three reasons. The first is that Spanish, also known as Castilian, emerges around Burgos, right next to the traditional Basque-speaking area. That means, from the very first appearance of Spanish, the two languages are very close geographically and even overlapping to some extent. This however is not a unique situation, as Spanish is also surrounded by other emerging languages, like Asturian in the west. The second reason is Basque itself. It is the only pre-Indo-European language still spoken in Western Europe, being one of the oldest languages spoken there. Although 2000 years of intense contact with Latin and Romance languages has replaced much of the ancient vocabulary, the core grammatical structures of the language are still very different from all the Romance languages surrounding it. Several very distinct features like basic subject-object-verb (SOV) word order, a heavily agglutinating structure, and ergative case marking clearly distinguish it from these other languages. Basque does, however, also border on Gascon in the east and originally also on Aragonese in the southeast. This is where the third reason for the uniqueness lies. Spanish has become the official language of the Spanish state and therefore the major contact language with Basque for the last 600 - 800 years. In that aspect it differs, for example, from French. Although French is also the only official language in the French part of the Basque-speaking area, it has only become a relevant linguistic factor for contact in the last roughly 130 years. Before that, the main language of contact north of the Pyrenees was, for centuries, Gascon. All these three reasons together highlight the singularity of the Basque-Spanish contact situation from a linguistic point of view.
In the following work, I will explore and examine a possible result of this contact situation. I join a long tradition of investigation and literature about this topic. In earlier work on Basque, the main focus was often on the lexicon and on possible relationships to other languages. In the last decades however, more studies about consequences that go beyond the lexical influence have been conducted. An outstanding work is surely Mitxelena’s *Fonética histórica vasca* on Basque phonetics and its development in the last 2000 years. More recently, a lot of work on the mutual syntactic influences was done by a multitude of authors, among many others Duguine & Irurtzun (2014), Etchebarne (2012), Fernández & Rezac (2016), Gómez Seibane (2012), Irurtzun (2018), Kintana & Meisel (2006), Landa (2000), Oñederra (2004), Rodríguez-Ordóñez (2017), and Urrutia Cárdenas (1988). The combination of a large bilingual group of speakers and the structurally very distinct languages makes this contact situation a field of work par excellence.

The main focus of this present work lies on interrogative structures which are surprisingly similar in Basque and Spanish. Both languages exhibit obligatory *wh*-fronting in an unmarked *wh*-interrogative. Furthermore, they normally require *wh*-phrase-verb adjacency which can only be interrupted by clitics or some few adverbs. Examples (1.1) and (1.2) illustrate unmarked or *out-of-the-blue* *wh*-interrogatives in Spanish and Basque, respectively.

(1.1) a. ¿Quién escribió ese libro?  
who wrote.3sg that book  
   a’. *¿Quién ese libro escribió?  
who that book wrote.3sg  
   ‘Who wrote that book?’

   b. ¿A quién le dio Pedro ese dinero?  
who CL.10 gave.3sg Peter that money
   b’. *¿A quién Pedro le dio ese dinero?  
who Peter CL.10 gave.3sg that money  
   ‘Who did Peter give that money to?’

Contreras (1999: 1939)

(1.2) a. Nor ikusi du Jonek?  
who seen AUX John.ERG
   a’. *Nor Jonek ikusi du?  
who John.ERG seen AUX  
   ‘Who has John seen?’

Ortiz de Urbina (1995: 100)
1.1 What this is all about

b. Nori azaldu zion Jonek atzo here erabakia?
   who-DAT explain AUX Jon.ERG yesterday his decision

b’. *Nori Jonek azaldu zion atzo here erabakia?
   who-DAT Jon.ERG explain AUX yesterday his decision

‘To whom did John explain his decision yesterday?’

Etxepare & Ortiz de Urbina (2003: 459)

The observation that Spanish exhibits wh-fronting is not surprising, since it is a very common phenomenon in the Romance and Germanic languages. In Basque, on the other hand, it is strange, as most of the languages worldwide that exhibit the basic word order SOV are wh-in-situ-languages (Aldai 2011, Russel 1978).

Although Basque and Spanish wh-questions seem to be very similar, I will argue that the underlying structure is very different. I claim that the superficial similarities are derived from two very distinct underlying movements. In order to develop this hypothesis, I will argue on the basis of Cable’s (2010) Q(uestion)-particle (QP) and QP-movement. To summarize his argument, he claims that wh-fronting is not the result of a special relationship between the wh-phrase and a higher C-head, but of a Q-particle and a C-head. The wh-phrase only appears sentence initial because it is ‘dragged’ along by this Q-particle, which takes the phrase containing the wh-word as a complement. This would be the case for Spanish, English, and German, for example. Basque would differ from Spanish in that the Q-particle does not take the phrase containing the wh-word as a complement, but is only adjoint to it. This gives the Q-particle greater freedom and allows it to move to the C-head alone, without ‘dragging’ the whole wh-phrase along. The reason that Basque nevertheless exhibit wh-fronting lies in an independent property: a preverbal focus feature that has to be satisfy by an adequate phrase. In a positive wh-question, the wh-phrase normally suffice for the job. However, as I will elaborate in Chapter 5, other constituents can occupy this preverbal position, preventing the movement of the wh-phrase to this very position. One such element is the negative particle ez ‘not’.

This underlying structural distinction accounts for the two major differences in question formation in Basque and Spanish: multiple wh-questions and complex wh-questions. The former type will be used mainly to support the hypothesis that will be brought forward in this work and will be examined in more detail in the respective chapters. The latter type, however, will be the main subject of investigation. A complex wh-question is a question with a complex or heavy wh-phrase. They are mostly composed of the wh-word cuál(es) de ‘which of’ or quién(es) de ‘who of’. In Spanish, in the case of a realized lexical subject, there seems to exist two possibilities concerning the word order in these types of interrogatives. The first is the standard constituent order wh-phrase-verb-subject, which always con-
stitutes an acceptable structure. The second, however, is the constituent order wh-phrase-subject-verb, which is normally not allowed in Spanish wh-questions. Examples for both possibilities are given in (1.3).

(1.3) a. ¿A cuál de sus profesores de la escuela primaria le dedicó Juan un poema?  
   A which of his teachers of the school primary Cl.10 dedicated.3sg Juan a poem
   ‘Which of his primary school teachers did John dedicate a poem to?’

b. ¿A cuál de estas chicas tu hermana (la) había visitado en Sicilia?  
   A which of the girls your sister Cl.10 had.3sg visited in Sicily
   ‘Which of the girls had your sister visited in Sicily?’

adapted from Ordoñez (1998b: 346)

In Basque, however, only the first possibility exists. That is to say, the wh-phrase-verb adjacency is always obligatory, no matter the complexity of the wh-phrase, as can be observed in example (1.4).

(1.4) a. Zein herritan bizi zen Jon lehenago?  
   which town.in live aux Jon before
   ‘In which town did Jon live before?’

b. * Zein herritan Jon bizi zen lehenago?

Ortiz de Urbina (1989: 213/14)

Two empirical studies, conducted by myself, have shown that Basque-Spanish bilingual speakers differ in the acceptance of Spanish complex wh-questions without wh-phrase-verb adjacency from monolingual Spanish speaker. The former group seems to accept them significantly less than the latter. Therefore, due to the structural difference, I will advance the following hypothesis: Basque-Spanish bilingual speakers do not accept, or at least to a significantly lesser extend, complex wh-questions that do not exhibit wh-phrase-verb adjacency, compared to monolingual Spanish speakers. In other words, the bilinguals strongly prefer the word order wh-phrase-verb-subject, which is the common structure in both languages, and reject the alternative order wh-phrase-subject-verb, which is an exclusively Spanish structure. Spanish monolingual speakers, however, do accept the second word order much more readily. It is my hypothesis that these observed differences are due to a different analysis of these complex wh-phrases. The monolingual speakers analyze this type of wh-phrases as clitic-left dislocations (CLLD) and with that have the possibility to insert the subject between the wh-phrase and the verb. The bilinguals’ analysis, however, is the same as that of a standard wh-phrase, which means that for them, wh-phrase-verb adjacency is also obligatory in these cases. The theory put forward through my work is that the reasons for this difference in the analysis are the following: on the one hand, it happens
due to differences in input during the acquisition process. On the other hand, it is a result of a general observation that if one language has possibilities A and B to analyze a certain structure and a second language only allows option A, bilingual speakers tend to primarily use option A in both languages.

1.2 Roadmap

The work is organized as follows: in Chapter 2, I will give a brief overview of the history of the Basque-Spanish language contact. To understand today's language situation in the Basque-speaking area, it is my opinion that a minimal introduction to its history is appropriate and necessary.

The main work is divided into three basic parts: a theoretical analysis of the structures, an introduction to language contact and language acquisition, and finally the presentation of my own empirical studies and their discussion.

The first part begins in Chapter 3 with a presentation and analysis of the main phenomena in Spanish. Section 3.1 will give a general introduction to Spanish word order, the notion of focus and topic (3.1.1), left and right dislocations (3.1.2), and focus-fronting (3.1.3). Section 3.2 will provide a detailed introduction to Spanish wh-questions, taking a closer look at the possible exceptions to the otherwise obligatory wh-phrase-verb adjacency in Section 3.2.2, and wh-multiple questions in Section 3.2.3. Finally, in Section 3.3, I will present and discuss two different approaches to the analysis of wh-questions, namely Rizzi's (1997) Cartographic Approach and a more recent approach by Kempchinsky (2013). The chapter will conclude with a proposition concerning the analysis of complex wh-phrases as CLLDed constituents in Section 3.3.3.

Chapter 4 provides a similar introduction to these phenomena in Basque. Section 4.1 presents an overview of Basque word order constraints in general, introducing the prominent preverbal focus position called galdegaia in traditional Basque grammar. Section 4.1.1 and 4.1.2 then give a detailed description of wh-questions in Basque, with special attention given to negative, multiple, and complex wh-phrases. In addition, Section 4.1.3 will present pied-piping, a widespread phenomenon in Basque. In the case of Basque, presentation of the constructions is followed by an analysis of the underlying syntactic structures, presenting and discussing the two main approaches to Basque wh-questions, namely the Cartographic and the Nuclear-Stress-Rule approach. Finally, the chapter concludes with an alternative proposition for the Basque left peripheral structure based on work by Haddican (2004).

In Chapter 5, I will introduce Cable’s (2010) Grammar of Q. After presenting the main concepts and properties of this alternative approach to wh-questions I will then apply them to Spanish in Section 5.1.1 and Basque in Section 5.1.2. It will become clear that Spanish and Basque differ decidedly in
1.3 Conventions for glosses and translations

In general, I have maintained the glosses and translations of the original sources cited throughout the work. For the sake of uniformity, however, some changes were made. In the glosses of Spanish sentences, the person is normally marked. Only third person singular present is not. Object clitic pronouns are marked with ‘Cl.do’ or ‘Cl.io’. Other clitics are only marked with ‘Cl’. The Spanish marker for [+human] direct objects is glossed as ‘A’.

The Basque glosses are held as simple as possible and I have normally only included the minimum information needed. The nominal cases are all marked except the absolutive, which is the unmarked version. Auxiliaries are normally only glossed as ‘aux’. The Basque participle verb forms are unmarked for present, but marked for imperfect and future.
For all examples, sentence focus or contrastive focus is marked with small caps, and ‘#’ signifies an infelicitous sentence in that context. The translations are normally taken from the original source. If the original text did not provide any, it is my own. The same goes for in-text citations in Spanish and translations in the footnotes.
CHAPTER 2

Basque-Spanish language contact –
a geographical and historical introduction

From the moment of man’s arrival to Europe until the present day, multiple contact situations have arisen between different languages throughout the entire continent. One of these zones of intense and interesting language contact is located in the western part of the Pyrenees. From south-western France, crossing the Pyrenees, entering the Iberian Peninsula and reaching as far as the Cantabrian mountains and the vineyards of La Rioja there exists an ancient and exceptional contact zone. This is due to a local linguistic factor, namely, the language of the euskaldunak ‘the ones who speak Basque’. Basque has been a constant factor throughout the last 2500 - 3000 years north and south of the Pyrenees, while the surrounding languages have undergone many profound changes. Basque is very different from all of these. It is not only the last pre-Indo-European language left in Western Europe but also encompasses a wide range of outstanding linguistic features. Outstanding at least when compared to the vast majority of modern European languages (Mitxelena 1977a: 18). To name just a few of the most famous characteristics: it is an ergative language with an agglutinating structure, the basic word order is subject-object-verb (SOV), it has a very rich verbal inflection with direct object, indirect object, and subject agreement.

The survival of Basque during these eventful two millennia is surprising and remarkable. Thus, to understand today’s linguistic situation in this area, it is necessary to give at least a brief historical summary indicating and explaining the most influential political and social events. A language’s flourishing and/or decline is always closely connected to such events. A language never flourishes or disappears for linguistic reasons but primarily for social or political reasons (Echenique Elizondo 2010: 247, Thomason 2010). Either because the speakers see an advantage or disadvantage in speaking their own language, or because they are obligated to speak a specific language. The former situation was
given during the Roman period. The Romans never forced anybody to learn Latin, but the social advantages and opportunities to ascend in Roman society multiplied if you did (Penny 2008: 25). A case of coercion occurred during General Franco’s dictatorship, when very strict laws prohibited the public use of any language besides Spanish. Often only a few historical events suffice to shape the destiny of a language.

Before I give an overview of the history of Basque and Spanish, I will briefly establish some standards concerning geographies and naming for this work. This is necessary, because in the melting pot of half a dozen languages, the names of villages, cities, regions and states easily get mixed up. In addition, I will say some few words about Basque dialects in general. The following section is far from being an exhaustive and detailed analysis of the intense and long Basque-Spanish contact. For a much more detailed description, the reader is referred to Luis Mitxelena’s work in general and for an excellent summary to Trask’s The History of Basque.

2.1 A brief note on geographies and naming

The Basque speaking area is traditionally divided into seven provinces, of which three are north of the Pyrenees in France and four south of them, in Spain. Traditionally, the seven Basque provinces are subsumed under the name Euskadi or Euskal Herria. All provinces and political divisions are shown in Figure 2.1.

The northern Basque provinces are Lapurdi (Sp. Labort, Fr. Labourd), Zuberoa (Sp. Sola, Fr. Soule) and Low Navarre (Bq. Nafarroa Behera, Fr. Basse Navarre). Collectively, these are also known as Iparralde (ipar ‘north’ + alde ‘part’). On the Iberian Peninsula, the four provinces can be divided politically

![Image of Basque provinces]

Figure 2.1: Euskal Herria in Europe and its provinces (Zorion 2017: www.wikipedia.org)
and historically into the western and the eastern Basque Country. The former consists of the three provinces Gipuzkoa (Sp. Guipúzcoa, (Gip)), Bizkaia (Sp. Viscaya, (Biz)) and Araba (Sp. Álava, (Ara)) which are currently united politically in the Basque Autonomous Region (Comunidad Autónoma Vasca (CAV)) with the capital Vitoria-Gasteiz. The eastern part is the Chartered Community of Navarre (Sp. Comunidad Foral de Navarra, Bq. Nafarroako Foru Komunitatea), normally referred to as Navarre (Sp. Navarra, Bq. Nafarroa, (Nav)), with its capital Pamplona (Bq. Iruña).

Following Trask’s (1997) example, I establish the following designations for the regions and cities in this work. The Basque names of the provinces will be used in this work except for Nafarroa and Nafarroa Beherea, for which I will use the more common names Navarre and Low Navarre. The major cities in the Basque Country will retain the Spanish or French names, that is, Bilbao, San Sebastian, Bayonne, Biarritz, and Pamplona. Vitoria is cited by its official form Vitoria-Gasteiz. All other towns will be cited by their Basque names.

2.2 A brief note on Basque dialects

It is often said that a Basque speaker from Bilbao and one from Maule have serious difficulties understanding one another when speaking their vernacular Basque dialects. Mitxelena (1982a: 10) wrote: “Son muy distintos esos dialectos entre sí? La respuesta dependerá del punto de vista: del que pregunta, del que constesta o del de ambos”.1 The truth is that the Basque language does split up into various dialects, but as Mitxelena (1977a: 16) said, this is more the rule than the exception in any linguistic dominion. Especially in a society where oral transmission was and still is the main medium of passing on cultural knowledge.

The first detailed and extensive work about Basque dialectology was done by Louis-Lucien Bonaparte, nephew of the famous Napoleon Bonaparte. With his outstanding work “Carte des sept provinces basque montrant la délimitation actuelle de l’euscara et sa division en dialectes, sous-dialectes et variétés”, published in London around 1869, he laid the foundation for all further work in that field (Mitxelena 1977a). He divided the language into eight different dialects, differentiating between four literary ones, namely Bizkaian, Gipuzkoan, Lapurdian and Zuberoan, and four non-literary ones, Southern and Northern High-Navarrese and Western and Eastern Low-Navarrese (Zuazi 1995: 16)

Nevertheless, all Basque dialects “are overwhelmingly congruent in their fundamentals, and differ chiefly in vocabulary and in a few rather low-level phonological rules” (Trask 1997: 5). In Mitxelena

---

1 Are these dialects very different from each other? The answer depends on the point of view of the one who is asking, the one who is answering, or of both.
words: “para un comparatista […], los dialectos vascos son, podría decirse, desesperantemente uniformes”\(^2\) (Mitxelena 1977a: 16). This uniformity is probably caused by a relatively homogenous Basque koiné between the sixth and eleventh century. A strong argument in favor of such a koiné was stated by Mitxelena. He said that all Basque dialects not only share a lot of archaisms but also developed the same innovations on nearly all linguistic levels, especially in phonetics (Mitxelena 1977a: 16). Nevertheless, when doing scientific research on Basque, the existence of these dialects should be taken into consideration.

### 2.3 A brief note on the history of Basque-Spanish contact

“Two thousand years of intense contact with Latin and Romance have had a profound effect upon the Basque lexicon, though not so profound as to have obliterated its distinctively non-Romance (and non-Indo-European) character” (Trask 1997: 249). This quote succinctly summarizes the influence that Spanish has had upon Basque. Since the arrival of the Roman troops in 201 BC and the ensuing near complete latinization and romanization of the Iberian Peninsula, the two languages have been in intense contact. First with Latin, followed by the Proto-Romance languages, and finally with modern Spanish. Spanish, or Castilian, starts as a remote vulgar Latin spoken in the mountains around Burgos. The ‘birthplace’ of Spanish therefore lies right next to the Basque-speaking area. A fact that made many scholars discuss an alleged Basque influence on early Spanish, like the Spanish canonical vowel system of only five vowels, the change from Latin word initial /f/ to /v/, and several items in the Spanish lexicon (Baroja 1945, Echenique Elizondo 2005, Mitxelena 1977b, Penny 2002, Trask 1998). Further evidence of at least the regional proximity of Basque and Spanish are the well known Glosas Emilianenses, the Emilian Glosses, found in the monastery San Millán de la Cogolla in modern La Rioja. Written around the turn of the millennium, they contain the oldest written sentences in Spanish and also the oldest known entire sentences in Basque.

The development of early Spanish is influenced by the rule of the Germanic Kings from their capital Toledo and then, with the Arabic invasion in 711 AD, heavily by the Arabic culture. The Arabic language has given Spanish its second biggest word contingent with about 4000 lexemes with an Arabic etymology (Neumann-Holzschuh 2012, Penny 2002). With the Arabic invasion, the 700 year-long Reconquista ‘the retaking of the Iberian Peninsula’ begins, at the end of which the Castilians from Burgos establish the Kingdom of Spain with Spanish as its official language. They then set out to conquer

---
\(^2\) for a comparatist […], the Basque dialects are, one could say, hopelessly uniform
the world, quite literally, taking their language with them. Exactly as Antonio Nebrijas wrote in the first written Spanish grammar in 1492: "Siempre la lengua fue compañera del imperio."\(^3\) The Basques take part in most of the Spanish explorations, voyages, and trade because they are excellent sailors and traders. Interestingly, throughout all that time, Basque stays primarily a spoken language, used among its native speakers, and rarely takes hold in new parts. This is interpreted by many scholars as a sign that, since early on, many Basque speakers were bilingual (Pusch & Kabatek, 2011). The first ever published book in Basque, *Lingua vasconum primitiae* by Bernad Etxepare, only appears in 1545.

The standardization of Spanish receives an important thrust when, in 1713, the Spanish *Real Academia Española* ‘Royal Spanish Academy’ is founded. An institution, the Basque language has to wait for until 1918, when finally the *Euskaltzaindia* ‘Basque Language Academy’ is founded. Its primary goal is to create a standard Basque in order to prevent the further diversification of the Basque dialects. This project was only finished in 1964, when *Euskara Batua* ‘Unified Basque’ was presented by the Academy.

Since its ‘scientific discovery’ by Wilhelm von Humboldt at the beginning of the 18th century, Basque has fascinated and attracted many scholars, linguists, and anthropologists to conduct research and write about it (Trask, 1997). One of the main goals was to find languages related to Basque. Until now, however, no living language has been found that provides enough evidence for a cognation with Basque (Trask, 1997).

The 20th century brings many ups and downs for the relation between Spanish and Basque. First, a new Basque nationalism under Sabino Arana. He not only forms the Basque National Party (*Euskal Alderdi Jeltzalea/Partido Nacionalista Vasca* (PNV)), but also creates a whole new national identity. He introduces a new name for the seven Basque provinces, *Euskadi*, the Basque national flag, the *Ikurriña*, and the Basque national anthem, *Gora ta Gora*. His ideas about a politically independent *Euskadi* are widely received and are an important factor for his success (Garcia de Cortázar & Lorenzo Espinosa, 1988). He promotes the language and emphasize its uniqueness. After the civil war (1936-1939), all other languages except Spanish are prohibited in the public life by the military dictatorship under General Franco. This prohibition results in a strong decline in the use and transmission of the Basque language, especially in the bigger cities. It is only during the sixties that the language policy is loosened a little and the first *ikastolas* ‘Basque schools’ are opened. In the late seventies, after going through nearly four decades of military dictatorship, Spain finally enters the circle of European democracies with a peaceful transition, heralding a new era of progressive linguistic politics. The Spanish constitution is probably one of the most liberal regarding language policy. It states as follows:

\(^3\) The language has always been the companion of the empire.
Preamble of the Spanish constitution:

La Nación española, deseando establecer la justicia, la libertad y la seguridad y promover el bien de cuantos la integran, en uso de su soberanía, proclama su voluntad de: […]

Protect a todos los Españoles y pueblos de España en el ejercicio de los derechos humanos, sus culturas y tradiciones, lenguas e instituciones.

§ 3:

1. El Castellano es la lengua española oficial del Estado. Todos los españoles tienen el deber de conocerla y el derecho de usarla.
2. Las demás lenguas españolas serán también oficiales en las respectivas Comunidades Autónomas de acuerdo con sus Estatutos.
3. La riqueza de las distintas modalidades lingüísticas de España es un patrimonio cultural que será objeto de especial respeto y protección.

The Basque regional government has made and still makes sweeping use of their constitutional rights. Since the passing of the autonomous statute in 1979, Basque is the co-official language in the Comunidad Autónoma de Euskadi (CAE) ‘Autonomous Community of the Basque country’ and in the north of Navarre. Especially in the CAE, Basque is heavily promoted and the number of speakers is steadily increasing. The latest survey shows the percentage of Basque speakers in the CAE in the last 25 years increase from 9.8% to almost 34% of the population. In Navarre, the numbers are also growing (from 9.5% to 12.9%). In the northern French part of the Basque country, the number of speakers has decreased from 26.4% to 20.5%. The overall trend can be seen clearly in the total number of speakers that has increased from around 530,000 in 1991 to 750,000 in 2016. Especially among younger people, the use of Basque has increased immensely in the last 25 years.4

In summary, the Basque language has survived many radical changes throughout the last 2500 years and is still spoken by many people. The reasons for its survival despite all the changes around are multiple: the remoteness of the mountains, the peculiarity of the language and its ability to adapt and integrate new features, the social solidarity of its people in difficult time and their willingness to learn other languages, as well as a strong cultural identity through the language (Mitxelena 1977a: 17-3

4 All data is taken from the VI Encuesta Sociolingüística 2006.
2.3 A brief note on the history of Basque-Spanish contact

Spanish, on the other hand, has developed from a small and remote language to one of the most widely spoken languages in the world. The combination of military conquest, economical power, and high prestige of the culture and the language brought the Spanish language to where it is today. The Basque-Spanish language contact is probably one of the oldest in Western Europe and therefore of special interest for linguists. How does such a long contact situation reflect in the languages, especially if they are typological so different? What influences can be found beyond the lexicon? Is the contact intense enough to find structural and grammatical influences? The Spanish-Basque contact situation provides an outstanding opportunity to examine more closely the effects of such long and intense language contact.
CHAPTER 3

Word order, questions, and focus in Spanish

Basque and Spanish are two very different languages from a typological point of view, the basic word order being one of the most prominent distinctions. In Basque, it is subject-object-verb (SOV), whereas in Spanish, it’s subject-verb-object (SVO). In wh-questions, however, both languages behave surprisingly similar. Both exhibit wh-fronting and normally obligatory wh-phrase-verb adjacency. Interestingly, the same holds for focus fronting constructions in both languages. Additionally in Spanish, other so called left dislocations exist beside wh- and focus-fronting. In the following two chapters, I will present the main characteristics of these constructions in both languages. In Section 3.1, I will start with the main properties of all non-interrogative types of dislocations in Spanish, In Section 3.2, I will present Spanish wh-questions, followed by a syntactical approach to both phenomena in Section 3.3. In Chapter 4 I will do the same for wh-questions and focus-fronting constructions in Basque.

3.1 On Spanish word-order

Spanish has a wide range of possible word orders in a declarative transitive sentence. If we have a look at the six logically possible permutations of subject (S), object (O) and verb (V), five of them can be found in Spanish. It shares this property with other romance languages like European Portuguese or, under some restrictions, also with Italian (see, among others Olarrea 2012: 603, Costa 2004: 3, López 2009: 129). Spanish allows SVO (3.1a), VSO (3.1b), VOS (3.1c), OSV (3.1d), OVS (3.1e), excluding only the SOV (3.1f) word order.

This phenomena is often referred to as subject-verb-inversion (SVI). Nevertheless, as we will see later on, the term is somewhat misleading, since there does not occur a real “inversion” in the sense of a mutual change via left- and rightward movement. As I will show later, the idea is that both constituents move leftward, but stop at different positions, restrained by an adjacency condition. I will develop a detailed analysis in Section 3.3.
3.1 On Spanish word-order

(3.1) a. Pedro comió una manzana.
   Peter eat.3sg a apple
   'Peter ate an apple.'

   b. Comió Pedro una manzana.
   c. Comió una manzana Pedro.
   d. Una manzana, Pedro la comió.
   e. Una manzana comió Pedro.
   f. * Pedro una manzana comió.

These sentences do not differ in their propositional content, but do have different informational and
prosodic patterns. (3.1a) is considered the unmarked, basic word order in Spanish, a fact that is is not
totally uncontroversial, as I will explain later on. The standard word order of a language is derived from
the most natural answer to the question ‘What happened?’. Because from a perspective of discourse,
in such an answer all elements yield new information. As a result, such a sentence is called an all-focus
Costa 2004: 15, Ordoñez 1997: Ch.2). I will come back on the exception to this basic word order in
Section 3.3.1.

The other possible word orders have a special informational or prosodic pattern. I will adopt López'
definition of the different dislocation phenomena, namely Clitic Left Dislocation (CLLD), Clitic Right
Dislocation (CLRD), Hanging Topic Left Dislocation (HTLD)⁷ and Focus Fronting (FF). The following
examples illustrate the four dislocations providing the different word orders.

(3.2) a. Las flores las compré ayer.
   the flowers Cl.acc bought yesterday
   'The flowers, I bought them yesterday.' (CLLD)

   b. El bibliotecario los encontró ayer, los libros.
   the librarian Cl.acc.pl found yesterday, the books
   'The librarian found them yesterday, the books.' (CLRD)

   c. Las rosas, me encantan esas flores.
   the roses, Cl.dat love.3pl these flowers
   'Roses, I love those flowers.' (HTLD)

   d. ESTAS FLORES quiere María.
   those flowers wants Maria
   'María wants THESE FLOWERS.' (FF)⁸

adapted from Olarrea (2012: 610)

---

⁶ I will consider only transitive sentences. For other sentences with psych verbs or unaccusative verbs, see Gutiérrez-Bravo (2002)
⁷ some authors refer to it simply as Left Dislocation (LD), e.g. Zagona (2002)
⁸ following the syntactic tradition, I will use small caps to indicate Focus
3.1 On Spanish word-order

We can see in these four examples how the different dislocations vary. The next two sections are dedicated to give a descriptive overview, presenting the principal properties and distinctions of these constructions. However, since the terms focus and topic are widely and quite differently used by different authors, I will first give a short introduction to the terminology I will follow throughout my work. With the terminology settled, I will present the three types of dislocations, namely CLLD, CLRD, and HTLD in Section 3.1.2, followed by the main properties of FF constructions in Section 3.1.3.

3.1.1. Focus and Topic – a short description

Focus and Topic are two notions in the linguistic literature that have received a lot of attention. The terminology, definitions, and descriptions vary dependent on the author’s view. Sometimes we find different terms for the same construction and sometimes different constructions labelled with the same term (Cruschina 2012: 9). To avoid confusion, I will follow in this work mostly the terms and definitions established and used in Cruschina (2012). Thus, I will give a short introduction to his terminology.

Cruschina (2012) and subsequent work distinguishes three different notions: **Topic**, **information focus** (IFoc) and **contrastive focus** (CFoc). Furthermore, he defines two kinds of topics, **aboutness topic** (ATop) and **referential topic** (RTop). The first one, as the name already implicates, tells us “what the sentence is about” (cf. Frascarelli & Hinterhölzl 2007, Reinhart 1981). It is, “the constituent representing the theme of the predication” (Cruschina 2012: 11). In Romance languages, ATops occur normally in pre-verbal position. Since we have only one theme per sentence, there is always exactly one single ATop. RTops, on the other hand, can appear pre- and post-verbally and are not limited in their number, but by our ability of processing language (Cruschina 2012: 19). RTops are “constituent[s] within a sentence conveying shared and anaphoric information with respect to the previous discourse and the mental state of the interlocutors” (Cruschina 2012: 11). That is, RTops are optional because the corresponding clitic would be enough to establish the necessary reading. In addition, RTops can function as ATops in a given sentence. In general, topics in most Romance languages are dislocated and co-referential with a resumptive clitic pronoun (Cruschina 2012: 20–21). I will come back to the notions of RTop and ATop at the end of Section 3.1.2.

The other two notions, namely contrastive and information focus, Cruschina describes as follows:

The novelty/new information feature on IFoc indicates that the assertive part of the sentence, that is the focus of the sentence, must be interpreted as innovative and the most informative, in the sense that it contributes new and relevant information to the universe of discourse […]. The contrastive feature of CFoc instead indicates that the assertion corresponds to denying or correcting a previous innovative assertion or presupposition that the speaker does not share.

Cruschina (2012: 14)
This means, following Gundel & Fretheim (2004) and subsequent work, Cruschina assumes two parameters for the contrast givenness-newness, namely relational givenness-newness and referential givenness-newness. The former “refers to the speaker’s assessment of the relations between the elements of a sentence in a given discourse, so that the topic identifies what the sentence is about and the focus is what is predicated about the topic” (Cruschina 2012: 10, Gundel & Fretheim 2004: 177) whereas the latter “indicates the relation between a linguistic expression and the mental state and knowledge of the speaker or the hearer” (Cruschina 2012: 10, Gundel, Hedberg & Zacharski 1993, Lambrecht 1994). Thus, using the two parameters, we can distinguish IFoc and CFoc. IFoc is “the constituent conveying the information predicated about the topic that is also referentially new” and CFoc is “the relational predicate of a sentence that need not be new in reference to the knowledge and attention state of the speaker or the addressee” (Cruschina 2012: 11). The following table summarizes the previously defined notions:

<table>
<thead>
<tr>
<th>Referentially New</th>
<th>Relationally New</th>
<th>Relationally Given</th>
</tr>
</thead>
<tbody>
<tr>
<td>IFoc, CFoc</td>
<td>Relationally New</td>
<td>ATop</td>
</tr>
<tr>
<td>CFoc</td>
<td>Referentially Given</td>
<td>RTop</td>
</tr>
</tbody>
</table>

Every sentence has an IFoc, although the part ‘carrying’ it can differ. It may be the whole sentence, the predicate or only a single constituent as shown in (3.3) - (3.5), respectively. Following Bosque & Gutiérrez-Rexach (2011), Casielles Suaréz (2001), Cruschina (2012), López (2009), and Zubizarreta (1998, 1999a), among others, I assume that IFoc normally appears post-verbally in Spanish.

(3.3) Context: What happened?
[Pepínllegó tarde]_sentence focus
‘Pepin arrived late.’

(3.4) Context: What did Pepin do?
Pepín [llegó tarde]_predicate focus

(3.5) Context: Who arrived late?
Llegó tarde [Pepín]_argument focus

adapted from Bosque & Gutiérrez-Rexach (2011: 680)

CFoc, in contrast to IFoc, does not appear obligatorily in every sentence, but “often coincides with the constituent offering new information in a relational sense […]”, but it need not be referentially new” (Cruschina 2012: 12). That is, CFoc is used to update or correct an already available information in
the discourse universe. The referential set can be an explicit mentioning of alternatives as in (3.7) or a known limited set of candidates as in (3.6) (Cruschina 2012: 90).

(3.6) 
A: Marie arrived late yesterday.
B: No te enteras. [Pepín]$_{CFoc}$ llegó tarde.
You don’t get it. Pepín arrived late.

(3.7) 
A: Who did you just say arrived late yesterday, Pepin or Carlos?
B: [Pepín]$_{CFoc}$ llegó tarde.
It was Pepín who arrived late.

adapted from Bosque & Gutiérrez- Rexach (2011: 682)

In Spanish, we have a clear distribution regarding the two foci. As already mentioned, IFoc appears only post-verbally, whereas CFoc has to appear pre-verbally. It even needs to be in direct adjacency to the verb. That is, nothing is allowed to intervene between the constituent that bears the contrastive focus and the verb.9 López (2009) argues convincingly that there are no post-verbal contrastive foci. He states that even if the reading is contrastive, the focalized constituent always co-appears with some other left-peripheral element, often a NPI like no (see López 2009: Ch. 2.4). In Basque the situation is simpler. As I will show in Section 4.1 there is only one position for focalized constituents in Basque, namely the preverbal one. Therefore, all instances of focus and focus fronting that I will present in Section 3.1.3 will be CFoc if not otherwise indicated.

After having established the necessary terminology, I will come to the main characteristics of the different left-dislocations in Spanish.

3.1.2. CLLD, CLRD and HTLD

We have seen in (3.2a) and (3.2b) that dislocated objects are doubled by a clitic. CLLD and CLRD have various properties in common. They are both iterative, they can occur in root and embedded sentences, the presence of a clitic is obligatory and in the case of more than one dislocated element of the same type, their relative order is free. Nevertheless, the two dislocations cannot be used in exactly the same context. As described in López (2009) and Villalba (2000), CLRD contrast with CLLD in the possibility to establish a relationship with the antecedents. In general, such relation can be that of a subset, a superset, a set-membership or a part/whole. In (3.8a), the antecedent is the superset and an answer with a CLRD is infelicitous, meanwhile an answer with a CLLD is perfectly felicitous, as seen

---

9 Spanish differs in this property from other Romance languages like Italian or Sicilian, which do allow IFco pre-verbally and do not have the adjacency restriction regarding CFoc. Interestingly Cruschina states that in the case of a pre-verbal IFoc, the adjacency restriction applies again (Cruschina 2009: 24, 2012)
in (3.8b). The same happens with an antecedent which is a subset (see (3.9)), where a CLRD results in an infelicitous answer, too.10 (for a detailed analysis, see chapter 2 & 4 in (López 2009)).

(3.8) **Context:** What did you do with the furnitures?

a. # Las traje por la mañana, las mesas pero las traje por la tarde, las sillas.
   ‘The tables I brought in the morning, but the chairs I brought in the evening.’

b. Las mesas las traje por la mañana pero las sillas las traje por la tarde.
   ‘The tables I brought in the morning, but the chairs I brought in the evening.’

(3.9) **Context:** So, what do we do with the chairs?

a. # Te dije que los dejamos en casa, los muebles.
   ‘I told you that we leave the furniture at home.’

b. Los muebles te dije que los dejamos en casa.
   ‘I told you that we leave the furniture at home’

adapted from López (2009: 42-43)

A further difference is the strong intonational break that separates the CLRDed constituent from the rest of the sentence, similar to the dislocated constituent in a HTLDed construction (Olarrea 2012: 615). Apart from these differences, both dislocation types behave the same. However, since CLRD does not concern the phenomena of *wh*-movement, I won’t describe them in more detail but will have a look at the differences between HTLD and CLLD.

At first appearance, HTLD and CLLD appear to be similar constructions. However, in a more detailed analysis, they contrast in some very important aspects. First, they differ in the kind of constituent that can be dislocated and the relation with the resumptive element. In CLLD, any phrasal element can be CLLD (3.10), whereas only a DP can be a HTLD occupying the sentence initial position (3.11a). Any other phrase yields an ungrammatical sentence (3.11b) (Bosque & Gutiérrez- Rexach 2011: 686, Cinque 1990, López 2009: 3-6, Olarrea 2012: 611, Zagona 2002: 220-224).

10 The original sentences in 3.8 and 3.9 are in Catalan, but as noted by various authors (among others López 2009, Villalba 2000, Vallduví 1992), Spanish behaves very similar to Catalan in these cases.
Second, a CLLDed constituent must always be associated either with a clitic (3.10) or with an empty pronominal pro (3.12a). In the former case, the CLLD and the clitic must hold an agreement relation. Furthermore, a CLLD can never be associated with or doubled by a tonic pronoun or a phrase (ephitet), as seen in (3.12b) and (3.12d), respectively\(^\text{11}\) (Bosque & Gutiérrez-Rexach 2011: 687, Olarrea 2012: 612). Spanish has neither partitive, locative or subject clitics. Therefore, the agreement relation between the CLLDed constituent and the matrix sentence is set by an empty partitive, locative or subject pronominal clitic pro. Other romance languages like French, Catalan or Italian do display these sort of clitics in CLLD constructions (Bosque & Gutiérrez- Rexach 2011: 687, López 2009, Vallduví 1992). HTLDs, on the other hand, can be resumed by either a tonic pronoun (3.11a), a clitic (3.13a) or a phrase (ephitet) (3.13b). In addition, as shown in (3.13c), no agreement relation is needed.

3.1 On Spanish word-order

d. *A María nunca vi a esa chica tan enfadada. A María never see.1sg this girl so irritated

(3.13) a. María, hace tiempo que no la veo. María, make.3sg time that neg CLACC see.1sg
‘María, I haven’t seen her in a long time.’
b. María, hace tiempo que no veo a esa muchacha. María, make.3sg time that neg see.1sg A this girl
‘María, I haven’t seen that girl in a long time.’
c. El ordenador, yo odio estas máquinas infernales. the computer, I hate.1sg those machines evil
‘The computer, I hate those evil machines.’

adapted from López (2009) and Olarrea (2012)

Third, the two constructions differ in their asymmetric occurrence in root and embedded sentences. While CLLD can occur in both contexts, HTLD surface almost exclusively in “absolute first position” (examples in Olarrea 2012: 612).12 In addition, a HTLD can be preceded by “topicalizing expressions” like en cuanto a ‘as for’, hablando de ‘speaking of’, or por lo que afecta ‘regarding’ (3.14a). These kind of topicalizations are excluded with CLLD (3.14b)13 (Bosque & Gutiérrez-Rexach 2011: 686, Olarrea 2012: 612, Zagona 2002: 223).

(3.14) a. En cuanto a Antxon, él nunca va a terminar su capítulo. as for Antxon, he never will.3sg to finish.1NF his chapter
‘As for Anton, he will never finish his chapter.’
b. * Te he dicho que en cuanto a Juan lo vi ayer. you have.1sg said that as for Juan CLACC see.1sg yesterday
‘I have told you that regarding John, I saw him yesterday.’

Olarrea (2012: 612)

Fourth, HTLD and CLLD contrast in their sensitivity to island violations. It is commonly agreed that HTLD are insensitive to any island, be it weak or strong. Examples (3.15a), (3.15b) and (3.15c) show an anaphorically related element inside a Complex Noun Phrase, a Relative Clause and a Coordinated Island, respectively.

12 López (2009) states that the “ungrammaticality of HTLD in subordinated clauses is not uniform, ranging from total unacceptability to mild ungrammaticality”. I will adopt here his point of view, that in general, “HTLD in subordinate clauses gives rise to some degree of ungrammaticality”, which will suffice, as it is to no further importance to my work.
13 The example is an embedded sentence, since it would be difficult to distinguish the CLLD from the HTLD in a matrix sentence. Hence, in order to show the ungrammaticality of an insertion of the above mentioned type, a embedded CLLDed constituent serves the purpose.
3.1 On Spanish word-order

(3.15)  
a. En cuanto a este trabajo, no puedo aceptar la idea de que ya lo han conseguido.
‘As for that job, I can’t accept the idea that they have already gotten it.’
b. Hablando de “Freaks”, un amigo que ha visto esa película me ha dicho que es magnífica.
‘As for Freaks, a friend who has seen that movie has told me that it is great.’
c. Por lo que se refiere a ese libro, te tomas un par de días de descanso y seguro que lo acabas.
‘As for the book, you take a couple of days off and surely you will finish it.’

(Olarrea 2012: 614)

The case is not that clear with CLLD. Whereas some linguists claim that CLLD are sensitive to strong but not to weak island (Cinque 1990, Olarrea 2012, Zagona 2002), López argues that CLLD are also sensitive to weak island. He points out that in former approaches only dislocated definite object DPs were considered. They “are easily extractable, whether as a CLLD or a wh-phrase” (López 2009: 5), which is shown in (3.16) and (3.17), respectively.

(3.16)  
a. Este libro me arrepiento de habérmelo leído.
‘This book I regret having read it.’
b. Este libro, me pregunto cuándo le leeremos.
‘The book, I wonder when we will read them.’

(3.17)  
a. ¿Qué libros te arrepientes de haber leído?
‘Which books do you regret having read?’
b. ¿Qué libros te preguntas cuándo leeremos?
‘Which books do you ask yourself when we will read them?’

adapted from López (2009)

The sentences with CLLD do not sound better or worse than the interrogatives with wh-movement. Therefore (López 2009: 223) concludes that ‘the specialness of CLLD in skipping weak islands is a myth’. He suggests that once these definite object DPs are excluded, CLLD is also sensitive to weak islands. He proves it by replacing the CLLDed constituent with a nonspecific object. The following sentences show that extracting a adjective out of an wh-island (3.18a) or an indefinite direct object out
of a factive island (3.18b) yields very marginal or even ungrammatical results. Therefore, I follow López suggesting, that CLLD are sensitive to weak island.

(3.18) a. * Inteligente, me preguntan si lo es.
    intelligent CLACC ask.3pl if Cl. is
    ‘If he is intelligent, they ask.’

b. ?? Algunas historias, me parece haberlas olvidado.
    some stories Cl seem.3sg have.3sg CLACC forgotten
    ‘Some stories it seems that I have forgotten’

The fifth difference concerns multiple fronting. Various elements can easily be CLLDed in basically every desired order (3.19). Multiple HTLDs however are only possible if a conjoined phrase is present (3.20). Nevertheless, both dislocations can occur at the same time. In that case, the HTLD generally precedes the CLLD as shown in (3.21) (Olarrea 2012: 614).

(3.19) a. Ese libro, a Pedro, no se lo dio nadie.
    that book to Pedro neg CL.DAT CLACC give.3sg nobody
    ‘Nobody gave Peter that book.’

b. A Pedro, ese libro, no se lo dio nadie.

(3.20) a. * Juan, el libro, él no lo ha comprado.
    Juan the book he neg CLACC have.3sg bought
    ‘John, the book, he hasn’t bought it.’

b. En cuanto al dictador y al pueblo, éste repudia a aquél.
    as for the dictator and for the people this hates a this
    ‘As for the dictator and the people, the former hates the latter.’

(3.21) a. En cuanto a Juan, esa carta se la escribió Pedro.
    As for Juan, this letter CL.DAT CLACC wrote.3ps Pedro
    ‘As for Juan, Peter wrote him this letter.’

b. * Esa carta, en cuanto a Juan, se la escribió Pedro.

Zagona (2002: 223) and Olarrea (2012: 614)

---

14 The same accounts for a copula complement, as shown in López (2009: 223). In (i), he shows that a CLLDed copula complement is possible and (ii.) and (iii.), a wh-island and factive island, respectively, yield very marginal results.

i. Un buen estudiante, Juán dice que Carlos lo puede ser.
   A good student Juán says that Carlos Cl can.3sg be
   ‘Juan says that Carlos can be a good student’

ii. ?? Un buen estudiante me pregunta cuándo lo seré.
    A good student refl ask.1sg when Cl be.1sg

iii. ?? Un buen estudiante me arrepiento de haberlo sido.
    A good student refl regret.1sg have Cl been

Finally, let us recall the classification by Cruschina (2012: 12) in Section 3.1.1. He distinguishes two kinds of topics, namely Aboutness Topics (ATop) and Referential Topic (RTop). The formerly presented characteristics have shown that HTLDed elements can be associated with the notion of ATop whereas CLLDed constituents corresponds to the notion of RTop. 16

In summary, we have seen that CLLD and HTLD contrast in very important aspects with regard to how they relate to antecedents, potential resumptive elements and islands. The last contrast that remains to be elaborated for now is whether these dislocations are externally or internally merged. I will try to answer this question in Section 3.1.4. But first I will present the last dislocation construction of interest for my work, apart from wh-movement, namely focus fronting (FF).

3.1.3. Focus-Fronting

At first sight, focus fronting (FF) differs from CLLD in two main aspects. First, FFed constituents never get doubled by a resumptive element (3.22a). Second, it demands obligatory adjacency of the verb and the fronted constituent, unless the focalized element is the subject itself, illustrated in (3.22b) - (3.22d) (Olarrea 2012, Zagona 2002, López 2009, Bosque & Gutiérrez-Rexach 2011: 692/693).

(3.22) a. ESTAS FLORES (‘las) quiere María
these flowers Cl.acc wants María
‘It’s these flowers that María wants.’
b. ESTAS FLORES quiere María.
these flowers wants María
‘It’s these flowers that María wants.’
c. *ESTAS FLORES María quiere.
d. María me regaló la botella de vino.
María Cl.dat gave.3ps the bottle of wine
‘MARÍA gave me the bottle of wine’

Olarrea (2012: 616) and Zubizarreta (1998: 125)

Every type of maximal projection can be FFed, except VPs (Bosque & Gutiérrez-Rexach 2011: 693). Arregi (2003b) presents a test for Spanish that can clearly separate CLLD from FF. Namely, the former can always be followed directly by the particle sí (que) (3.23a), whereas the latter cannot (3.23b).

(3.23) a. LAS JUDÍAS SÍ ME LAS HE COMIDO.
the beans indeed Cl. Cl.acc have.1sg eaten
‘I have, in fact, eaten the beans.’

16 Assuming that RTop can occur pre- and post-verbally, CLRDeed elements can also be defined as RTop.
b. *las judías sí me he comido.
the beans indeed Cl. have.1sg eaten

López (2009: 8)

There are several other characteristics that distinguish LDs and FF. For instance, it is not possible to have more than one FFed element in one sentence (3.24a). However, nothing prevents a FFed constituent to follow another LDed element (3.24b). The other way round, by contrast, is not possible (3.24c) (Kempchinsky 2013: 312).

(3.24)  
a. *en la tienda una moto compró Pepe.
in the shop a motorcycle bought.3ps Pepe
‘Pepe bought a motorcycle in the shop.’

b. Tu compañero, muy poca atención pone en todo esto, me parece a mí. your friend, very little attention puts in this all Cl.dat seems to me
‘Your friend, it seems to me, puts very little attention into this.’

c. *una tarea a los estudiantes les di, no un examen. a homework to the students Cl.to gave.1sg not an exam
‘I gave a homework assignment to the students, not an exam.’

adapted from Bosque & Gutiérrez-Rexach (2011: 694) and Kempchinsky (2013: 312)

FF can occur in matrix as well as in subordinate clauses (3.25). However, FF is sensitive to strong and weak island restrictions. That is, it cannot be fronted out of a relative clause (3.26a), an adjunct island (3.26b), a complex subject island (3.26c) or a wh-island (3.26d).

(3.25)  
Me parece que eso exactamente quiere él. Cl.dat seems that this exactly wants he
‘It seems to me that he wants exactly this.’

Bosque & Gutiérrez- Rexach (2011: 694)

(3.26)  
a. *a Pedro conocemos la mujer que traicionó.
a Pedro know.1pl the woman who cheated.3ps
‘We know the woman who cheated on Pedro.’

b. *a Pedro terminemos la tarea antes de llamar.
a Pedro finish.1pl the task before to call.inf
‘We finish the work before calling Pedro.’

c. *a Pedro sorprendió a todo el mundo que María haya invitado. a Pedro surprised.3sg to all the world that María have.3sg invited
‘That María had invited Pedro surprised everybody.’

Zubizarreta (1999a: 4240)

There are some exceptions with regard to this point. FF in embedded clauses is restricted to whether it is a presupposed or assertive one. For a more detailed description see Bosque & Gutiérrez- Rexach (2011: 694) and Zagona (2002: 252).
3.1 On Spanish word-order

d. *CANSADO me pregunta si está.
   tired Cl. ask.3sg if is.3sg
   ‘He asks me if he is tired.’

One last important property is the complementary distribution with fronted wh-phrases. That
means, FFeD elements cannot cooccur with a fronted wh-phrase, as shown in (3.27a). The examples
also show that neither changing the order of the fronted constituents (3.27b) nor leaving the wh-phrase
in-situ (3.27c) yields a grammatical sentence.

(3.27)  a. *¿LAS MANZANAS cuándo compraron?
   the apples when bought.3pl
   ‘When did they bought the apples?’

   b. *¿Cuándo LAS MANZANAS compraron?

   c. *¿A LA OFICINA diste las llaves a quién?
   to the office gave2sg the keys to whom
   ‘You gave the keys to the office to whom?’
   adapted from López (2009: 124)

Zagona (2002: 251)

After having given a descriptive outline of these dislocation phenomena in Spanish, I will turn now
to another interesting question. I will discuss if these kinds of dislocated constituents are externally or
internally merged.

3.1.4. External merge and internal merge

The difference between an internally and an externally merged element is that the former interacts on
various levels with the rest of the sentence whereas the latter enjoys greater freedom in the structure
up to complete encapsulation. That is, an internally merged element α leaves “some copy of α” (López
2009: 214) in the place where it originally merged. If an element has left such a copy or trace, its
presence is detectable. Reconstruction and binding effects can attest those copies directly, whereas
externally merged elements do not exhibit such effects. (Olarrea 2012, López 2009: 214). HTLDs belong
to the latter group (López 2009). Already Hernanz & Brucart (1987) and Rivero (1980) analyze HTLD as
externally merged in a topic position or topic phrase. By assuming this type of merge, several crucial
properties follow directly. A direct grammatical link between the LDed constituent and its coreferential
element does not exist. It is only weakly bonded to the the latter, which can be overt, and “there is no
connectivity between this element and the dislocated phrase” (Olarrea 2012: 617). They appear almost
exclusively in root sentences in very first position as plain DPs (Zagona 2002: 225). In addition, by
assuming that the HTLDs are merged in the left periphery, it is easy to explain why they are insensitive
to strong and weak island (López 2009, Olarrea 2012, Zagona 2002).
The picture is not as clear with the CLLDs. On the one hand, they show clear properties of an externally merged constituent. For example, they do not trigger obligatory adjacency with the verb. They also have certain recursiveness patterns, like the arbitrary order in the case of several CLLDs. In addition, as pointed out by Olarrea (2012) and Zagona (2002), CLLDs neither induce Weak Cross Over effects (WCO) (3.28b) nor do they license parasitic gaps (3.28a). That is to say, some crucial characteristics of a clear internal merge with subsequent A′-movement do not figure in CLLDed constructions.

(3.28) a. A Teresa_k su_k madre la_k quiere mucho
   Teresa her mother CLACC loves much
   ‘Teresa, her mother loves her a lot.’

b. A una candidata el jefe la descartó sin entrevistar *(la).
   a candidate the boss CLACC rejected.3sg without interview *(her)
   ‘One candidate, the boss rejected her without interviewing her.’

Suñer (2006)

These properties have led various authors (among others Contreras 1991, Hernanz & Brucart 1987, Zagona 2002) to claim that “CLLDs are not a result of movement but rather externally merged in a left peripheral sentential adjunct position.” (Olarrea 2012: 618).

On the other hand, CLLDs are sensitive to strong islands and show typical reconstruction effects of A′-movement (López 2009, Suñer 2003, Zubizarreta 1998). As a result, López (2009) and Villalba (2000), among others, claim that CLLDs are internally merged and then moved to the left periphery. Following López (2009), I will shortly present his argumentation in favor of this position.

López’ argumentation is based on a comparison between FF, CLLD and HTLD, showing that the latter contrasts in all important aspects with the former two. Thus, he demonstrates a clear concordance between CLLDs and other A′-movements like FF or wh-movement. In order to support his hypothesis, López refutes the traditionally given arguments for external merge in a left adjoint position. He describes in detail various independent pieces of evidences for a movement hypothesis, i.e. case and agreement, resumptives, strong-cross-over (SCO, principle C-violation) and WCO effects, variable binding, subextraction, islands, floating quantifier and parasitic gaps. I will resume here only the three most controversial points, namely WCO, weak island sensitivity, and licensing of parasitic gaps.\footnote{For a detailed discussion of all the points, see López (2009: Chapter 6.2).}

According to López, WCO should not be used as a test for movement, because “there are instances of clear movement chains that are not sensitive to WCO and instances of resumptive chains that are
sensitive to WCO” (López 2009: 214). WCO, therefore, does not tell us anything about movement or lack thereof. Furthermore, as already described in Section 3.1.2, he argues convincingly that CLLD are sensitive to weak island, matching FF or wh-movement. López also shows that under the right conditions wh-movement (3.29b) as well as FFs (3.29a) can license parasitic gaps. Hence, CLLDs do not contrast in these properties with neither of them.19

(3.29) a. Los libros ha guardado sin leer.
   the books have.3sg put away without read.INF
   ‘It was the books that he put away without reading’

b. ¿Qué ha guardado sin leer?
   what have.3sg put away without read.INF
   ‘What did he put away without reading?’

c. Los libros, los ha guardado sin leer.
   the books CLacc have.3sg put away without read.INF
   ‘The books he put away without reading’

adapted from López (2009: 225-226)

Having presented López’s (2009) arguments, I will follow his assumption that CLLD are internally merged and subsequently moved to the left periphery, concurring with FF and wh-phrases. HTLD, on the other hand, are externally merged.

---

19 López (2009: 225) points out that the verb in the adjunct phrase can alter the acceptability. However, a sentence which sounds bad with CLLD (i.) does it with wh-movement or FF, too ((iii.) and (ii.), respectively).

i. * A Juan, lo he buscado durante meses sin encontrar.
   A Juan, CLacc have.1sg searched during month without find.INF

ii. * A Juan he buscado durante meses sin encontrar.

iii. * ¿A quién has buscado durante meses sin encontrar?
   A who have.2sg searched during month without find.INF
3.2 On Spanish *wh*-questions

In the previous section, I introduced different left-dislocations in Spanish. The most prominent left-dislocated structure, however, is still missing, namely the *wh*-interrogative. *wh*-questions belong to the so-called non-elliptic questions\(^{20}\) (Contreras 1999: Ch. 31.1).

The non-elliptic questions can be divided in three main groups: total or polar questions (3.30a), disjunctive questions (3.30b), and partial or *wh*-questions (3.31). The last group can be furthermore differentiated into bare (3.31a) and complex (3.31b) *wh*-questions, echo questions (3.31c) and multiple questions (3.31d).

\[(3.30)\]
\[\begin{align*}
\text{a. } & \text{¿Tienes amigos?} \\
& \text{have.2sg friends} \\
& \text{‘Do you have friends?’}
\end{align*}\]
\[\begin{align*}
\text{b. } & \text{¿Trabajas o estudias?} \\
& \text{word.2sg or study.2sg} \\
& \text{‘Do you work or do you study?’}
\end{align*}\]

\[(3.31)\]
\[\begin{align*}
\text{a. } & \text{¿Dónde trabajas?} \\
& \text{where work.2sg} \\
& \text{‘Where do you work?’}
\end{align*}\]
\[\begin{align*}
\text{b. } & \text{¿Qué libros de física has leído?} \\
& \text{what books of physics have.2sg read} \\
& \text{‘What kind of books about physics have you read?’}
\end{align*}\]
\[\begin{align*}
\text{c. } & \text{¿Lo viste dónde?} \\
& \text{Cl.do saw.2sg where?} \\
& \text{‘You saw him where?’}
\end{align*}\]
\[\begin{align*}
\text{d. } & \text{¿Quién conoce a quién?} \\
& \text{who know.3sg to who} \\
& \text{‘Who knows whom?’}
\end{align*}\]

Contreras (1999: 1934, 1937, 1941, 1942)

Furthermore, these questions differ in their relation to the context. We differentiate two kind of questions in this regard: d(iscourse)-linked and non-d(iscourse)-linked. The latter are also called ‘out-of-the-blue-questions’, because they do not need any special context to be interpretable. The former, on the contrary, do need a special context. We will see later that both question types behave differently in terms of word order and other characteristics. I will describe the d-linked questions in more detail

\(^{20}\) The opposite elliptic questions differ from the non-elliptic firstly in that they are heavily context bound. That is, they are only interpretable in a very special context. And secondly that they can consist of any word or phrase as long as it is not a finite verb form or contains a such (Contreras 1999: 1934).
in Section 3.2.2.3 when talking about complex wh-phrases. In addition, apart from the direct question, all the former mentioned interrogative constructions can appear also in indirect questions. These will be dealt with in Section 3.2.3.2.

Yes/no-questions are very common and the main type of non-wh-questions (Bosque & Gutiérrez-Rexach 2011: 711). The focus, however, will be on wh-questions and the different subtypes mentioned. In Section 3.2.1, a general introduction to wh-questions is given. In Section 3.2.2, exceptions to the adjacency rule will be discussed, before looking at multiple, echo, and indirect questions in Section 3.2.3. Finally, in Section 3.2.4, I will look briefly at dialectal variation in question formation in Caribbean Spanish.

3.2.1. wh-questions – main characteristics

Questions of the wh-type are among the most frequent subjects in modern linguistics. They are known as wh-questions\(^{21}\), content or constituent questions, partial questions or pronominal questions. Their main characteristic is the appearance of a wh-pronoun like ‘who’, ‘what’ or ‘where’. They main wh-words in Spanish are quién(es) ‘who’, qué ‘what’, cuándo ‘when’, dónde ‘where’, cómo ‘how’, por qué ‘why’\(^{22}\), cuál(es) ‘which’, cuánto(s) ‘how many/much’\(^{23}\). The position of the wh-word or wh-phrase in a normal wh-question is sentence initially directly followed by the verbal complex. The following examples show the wh-phrase in the function of a subject (3.32a), a direct object (DO) (3.32b), an indirect object (IO) (3.32c), a locative (3.32d), a temporal (3.32e), and a modal adverbial phrase (3.32f).

(3.32) a. ¿Quién escribió ese libro?  
‘Who wrote that book?’

b. ¿Qué compró Pedro?  
‘What did Peter buy?’

c. ¿A quién le dio Pedro ese dinero?  
‘Who did Peter give that money to?’

\(^{21}\) They are called wh-questions, because in English, their majority starts with ‘wh-’. In the Spanish literature, they are often called frases qu-

\(^{22}\) Por qué ‘why’ has two different readings. The standard causal reading of the English ‘why’, also called por qué simple ‘the simple why’ and the equivalent to ‘for what’ or ‘what for’, also called por qué complejo ‘complex why’. It is called ‘complex’ because it is seen as a combination of the preposition por ‘for’ and the wh-word qué ‘what’. The complex why behaves in every sense like the other wh-words whereas the simple does not. If it is not otherwise mentioned, por qué is a ‘simple why’.

\(^{23}\) From now on, by using the term wh-phrases, I include the cases where the phrase consists of only one wh-word as well as more complex phrases like for example ‘which of the X’.
Spanish belongs to the so-called *wh*-fronting languages. Languages generally can be separated into three types in regard to question forming. Spanish, English, German and most Indo-European languages belong to the first type, the *wh*-fronting or *wh*-moving languages. That means, the *wh*-element is moved to the left and starts the sentence. Mandarin Chinese, the Dravidian languages like Tamil or Telugu, and many other languages worldwide belong to the so-called *wh*-in-situ languages, the second type. The *wh*-phrase does not move to the front, but stays in-situ in its canonical position inside the sentence. The third type is a blend, where both question forming strategies are available. The respective use depends on different factors, such as formal or colloquial, or stylistic reasons, among others. French, for example, belongs to this latter group (among others, Barbosa 2001, Ž. Bošković 2002, Kayne 1975, Kayne & Pollock 1978, 2001).

Apart from the fronting of the *wh*-phrase, the second main characteristic of Spanish *wh*-questions is the obligatory *wh*-phrase-verb adjacency. As mentioned, the verbal complex normally has to immediately follow the *wh*-phrase. In (3.33) we see that the absence of the *wh*-phrase-verb adjacency yields ungrammatical sentences. As the examples show, the type of element intervening between the *wh*-phrase and the verb generally does not matter (Contreras 1999, Francom 2012, Goodall 1993, Ordoñez 1998b, Torrego 1984).

(3.33) a. *¿Dónde *tus* amigos viven?  
   ‘Where do your friends live?’

---

The following example from Cheng (2003: 103) in Mandarin Chinese shows that *shenme* ‘what’ stays in-situ.

i. Hufei mai-le shenme  
   Hufei buy:PERF what  
   ‘What did Hufeibuy?’
Spanish shares this adjacency requirement with many wh-fronting languages like English, German, and other Romance languages. Nevertheless, Spanish exhibits two main differences to the first two languages. First, unlike in English and German, in the case of a verbal complex consisting of either the auxiliar ser ‘be’ or haber ‘have’ plus a participle, both have to follow the wh-phrase together, as shown in (3.34) and (3.35), respectively. Second, this restriction also applies to indirect wh-sentences. I will come back to them in Section 3.2.3.2.

(3.34)  
\[\text{a. } \text{¿Qué ha organizado la gente?}
\]  
\[\text{What have organized the people}
\]  
\[\text{‘What have the people organized?’}
\]

\[\text{b. * ¿Qué ha la gente organizado?}
\]

(3.35)  
\[\text{a. ¿Por quién fue organizada la reunión?}
\]  
\[\text{by who was organized the meeting}
\]  
\[\text{‘By whom was the meeting organized?’}
\]

\[\text{b. * ¿Por quién fue la reunión organizada?}
\]  

adapted from Contreras (1999: 1939-1940)
In the case of a verbal sequence with more than one verb, “only the first verb of the sequence must necessarily precede the subject. Thus, with modals, aspectuals, and motion verbs the subject may precede the main verb”, as shown in (3.36) (Torrego 1984: 105).

(3.36) a. ¿Con quién podrá Juan ir a Nueva York?
   with who can John go to New York
   ‘With whom will John be able to go to New York?’
   b. ¿Qué viene Juan a hacer aquí?
   what come John to do here
   ‘What does John come to do here?’

Torrego (1984: 105)

I have presented already in Section 3.1 the notion of syntactic islands and the restrictions concerning extraction out of them in the cases of CLLD and FF. Also wh-phrases underly restrictions from where they can be extracted. They are sensitive to weak and strong islands. That means, wh-movement is not possible out of adjunct islands (3.37a), complex NPs (3.37b), coordinations structures (3.37c), subjects (3.37d), or wh-islands (3.37e).

(3.37) a. * ¿A quién, habló José con Irma [después de ver t₁]?
   a who spoke.3sg Joseph with Irma after of see.inf
   ‘Who did Joseph speak with Irma after seeing?’
   b. * ¿Qué defendió Juan [la propuesta de que se venda t₁]?
   What defended.3sg John the proposal of that Cl. sell.3sg
   ‘What did John defend the proposal that be sold?’
   c. * ¿Qué tocas [el piano y t₁]?
   What play.2sg the piano and
   ‘What do you play the piano and?’
   d. * De qué sabe Juan que [una botella t₁] se cayo de la mesa?
   Of what know.3sg John that a bottle Cl. fall.3sg off the table
   ‘What does John know that a bottle of fell off the table’
   e. * ¿Qué se pregunta Juan dónde fue María a comprar t₁?
   what Cl. ask.3sg John where went.3sg Marie to buy.inf
   ‘What does John wonder where Marie went to buy?’

Francom (2012: 540)

However, the picture is not as clear as it seems (see among others, Francom 2012: 540, Suñer 1991, Torrego 1984: Ch.3, Bosque & Gutiérrez- Rexach 2011: 450-455). With regard to wh-islands, Torrego (1984) states that “it is possible to wh-move from within indirect questions”, as long as “none of the elements that are wh-moved out of the embedded question are base-generated under the VP” (Torrego
That is to say, it is possible to extract out of indirect questions, if the indirect question does not exhibit obligatory adjacency between the *wh*-phrase and the verb, or if the extracted constituent is the subject, as in (3.38a). In (3.39a) and (3.39b), the questions differ in grammaticality regarding the facultative and obligatory adjacency, respectively. Furthermore, Spanish *wh*-questions are not submitted to the *that-trace effect* as their English counterparts are (3.40).

(3.38) a. ¿Quién no sabes qué es en esta empresa?
   who not know.2sg what is in this firm
   ‘Who don’t you know that is in this firm?’

(3.39) a. ¿Qué dices que no te explicas por qué Juan se habrá comprado?
   what say.2sg that not Cl. explain.2sg why John Cl. have.3sg bought
   ‘What do you say that you don’t understand why John will have bought?’

   b. * ¿Qué no sabes quién es en esta empresa?
   what not know.2sg who is in this firm
   ‘What don’t you know is who in this firm?’

Torrego (1984: 114-115)

(3.40) a. ¿Quién cree que Juan [que tío votó por Clinton]?
   who believe3sg John that Cl.ti voted3sg for Clinton
   ‘Who does John believe voted for Clinton?’

   b. * Who does John believe that voted for Clinton?

adapted from Francom (2012: 540)

*Wh*-interrogatives can co-occur with other LDs, except FF. That is, a CLLDed constituent may appear preceding the *wh*-phrase, but not following it, as shown in (3.41) (cf. Francom 2012: 538, Suñer 1994: 350-351, Zagona 2002: 251).

(3.41) a. Las acelgas, ¿quién las detesta?
   the spinach who Cl.do detests
   ‘Spinach, who detests it?’

   b. * ¿Quién las acelgas las detesta?

   c. A tu novia, ¿qué aun no le contaste?
   to your girlfriend what yet not Cl.to tell.2sg
   ‘What haven’t you told your girlfriend yet?’

   d. * ¿Qué a tu novia aun no le contaste?

Suñer (1994: 350-351)

A further restriction concerns the co-occurrence of *wh*-phrases and FFEd constituents. It is not possible to have both fronted in the same sentence, regardless of the order, as already seen in (3.27a).
and (3.27b), here repeated as (3.42a) and (3.42b) (Bosque & Gutiérrez-Rexach 2011: 695, Zagona 2002: 251).

(3.42)  a. *¿LAS MANZANAS cuándo compraron?
   the apples when bought.3PL
   ‘When did they buy the apples?’

   b. *¿Cuándo LAS MANZANAS compraron?

Zagona (2002: 251)

In all the examples presented so far, the wh-phrase has been always adjacent to the verb. Now, as a general rule, this adjacency is obligatory in all wh-interrogatives in Spanish, even in embedded interrogatives, as I will show in Section 3.2.3.2. However, we can find three major exceptions. First, some specific adverbs and clitics can always surface between the two constituents. Second, some few wh-phrases like por qué ‘why’ allow intervening constituents and third, complex or heavy wh-phrases also seem to allow different elements to appear between the wh-phrase and the verbal complex. In the following section, I will describe these exceptions.

3.2.2. Wh-phrase-verb adjacency – exceptions

The three main exceptions concerning the obligatory wh-phrase-verb adjacency are structured as follows: Adverbs and clitics in Section 3.2.2.1, anomalous wh-phrases in Section 3.2.2.2, and complex wh-phrases in Section 3.2.2.3.

3.2.2.1. Adverbs

The only elements that theoretically can always intervene are NPI (3.43a), clitics (3.43b) and some adverbs. In the latter case, adverbs of frequency (3.43c) and the two modal adverbs ya ‘already’ (3.43d) and todavía ‘still’ (3.43e) can surface between the wh-phrase and the verb (Contreras 1999, Francom 2012, Goodall 1993, Ordoñez 1998b, Suñer 1994, Torrego 1984).

(3.43)  a. ¿Qué lugar nunca visitas?
   what place never visit.2sg
   ‘Which place do you never visit?’

   b. ¿A quién le dio Pedro el dinero?
   A who            Cl.10 gave.3sg Peter the money
   ‘Who did Peter give the money to?’

   c. ¿A quién siempre le das regalos?
   A who always Cl.10 give.2sg presents
   ‘To whom do you give always presents?’
3.2 On Spanish \textit{wh}-questions

\begin{itemize}
  \item[d.] ¿Quién ya te conoce?
    \begin{tabular}{p{0.5\textwidth}}
      who already Cl. know.3sg
      ‘Who already knows you?’
    \end{tabular}
  \\
  \item[e.] ¿De quién todavía te acuerdas?
    \begin{tabular}{p{0.5\textwidth}}
      of who still Cl. remember.2sg
      ‘Whom do you still remember?’
    \end{tabular}
\end{itemize}

adapted from Contreras (1999: 1939-1940)

The fact that these adverbs can always intervene between the \textit{wh}-phrase and the verbal complex, is a core argument in Suñer’s (1994) discussion about verb-movement in Spanish \textit{wh}-questions. She argues that these adverbs are IP-internal and since they can precede the verb in \textit{wh}-questions, no verb movement to $C^0$ has taken place. Therefore, Spanish does differ from standard V2 languages like English or German, in which no adverb whatsoever is allowed to intervene between the \textit{wh}-phrase and the verb (Suñer 1994). This observation will be of interest in Section 3.3.

3.2.2.2. \textit{wh}-phrases

In Spanish, we find especially two \textit{wh}-phrases which allow the intervention of constituents, namely the adjunct \textit{wh}-phrases \textit{por qué} ‘why’ and \textit{cómo que} ‘how come’ (Bosque & Gutiérrez-Rexach 2011: 712, Contreras 1999: 1940, Torrego 1984: 106, Fontana 1993: 207-208).

\begin{align*}
  (3.44) & \quad \begin{align*}
  a. & \quad \text{¿Por qué el tren se retrasa?} \\
  b. & \quad \text{¿Por qué se retrasa el tren?}
\end{align*} \\
& \quad \begin{align*}
  \text{why the train Cl. is.late.3sg}
\end{align*} \\
& \quad \text{‘Why is the train late?’}
\end{align*}

Bosque & Gutiérrez-Rexach (2011: 712)

\begin{align*}
  (3.45) & \quad \begin{align*}
  a. & \quad \text{¿Cómo que Pedro no quiere trabajar?} \\
  b. & \quad \text{¿Cómo que no quiere trabajar Pedro?}
\end{align*} \\
& \quad \begin{align*}
  \text{how that NPI want.3sg work.1NF Peter}
\end{align*} \\
& \quad \text{‘How come Peter does not want to work?’}
\end{align*}

Contreras (1999: 1940)

Various authors have shown that ‘why’ behaves differently in many languages. Rizzi (2001: 293) describes various differences of Italian \textit{perché} ‘why’, when compared to other \textit{wh}-words. It require \textit{wh}-phrase-verb adjacency, allowing certain adverbs (3.46a) or arguments (3.46b) to surface between the two. In addition, \textit{perché} can co-occur with FF, which is impossible with other \textit{wh}-phrases (3.46c). Also in French, \textit{pourquoi} ‘why’ behaves differently in several cases. Rizzi (1990) for example shows that
pourquoi is not allowed in-situ. Furthermore in French, the DP inversion construction is not possible with pourquoi (see (3.47a)) (among others, de Cornulier 1974: 139-140, Rizzi 1990). In Spanish multiple wh-questions, por qué ‘why’ cannot appear in a sentence final position (3.47b) (Bosque & Gutiérrez-Rexach 2011: 515, Contreras 1999: 1943).

(3.46)  

a. Perché Gianni è venuto?  
   why Gianni has come  
   ‘Why has Gianni come’

b. Perché (i tuo amici) già hanno finito il lavoro?  
   why (your friend) already have finished the work  
   ‘Why has (your friend) already finished the work?’

c. Perché questo avremmo dovuto dirgli, non qualcos’altro?  
   why this should have said to him not other thing  
   ‘Why should we have said this to him, not something else.’

Rizzi (2001: 293-294)

(3.47)  

a. *Pourquoi est partie Marie?  
   Why is left Marie  
   ‘Why has Marie leave’

b. *¿Quién va a clase por qué?  
   who go.3sg to class why  
   ‘Who goes to the school why?’

Drijkoningen & Kampers-Manhe (2008: 191)

Contreras (1999: 1943)

In addition, if the wh-phrase por qué ‘why’ co-occurs with a CLLed constituent, the former can precede the latter or vice versa, as illustrated in (3.48) (Suñer 1994: 351).

(3.48)  

a. A Paco, ¿por qué ya no lo aguanta nadie?  
   A Paco why already not Cl.do can stand nobody  
   ‘Why can’t anybody stand Paco already?’

b. ¿Por qué a Paco ya no lo aguanta nadie?

Suñer (1994: 350-351)

I will return to the special characteristics and a possible structural explanation of the behavior of these two types of wh-interrogatives in Section 3.3.2.
3.2.2.3. Complex wh-phrases

Suñer (1994) and Torrego (1984) and others have pointed out that the obligatory adjacency between the wh-phrase and the verbal complex can be violated if the wh-phrase is non-argumental. This would mean, (3.49a), (3.50a), (3.51a) and (3.52a) should be acceptable, which they are not. Ordoñez, however, claims “that the distinction about the obligatoriness of the inversion does not cut across the argumental – non-argumental distinction so much as the one between complex and non-complex wh-elements” (Ordoñez 1998b: 340). The b.- variants in the examples show a clear increase in the acceptance, for both non-argumental ((3.49) - (3.51)) and argumental (3.52) - (3.53) complex wh-words.

(3.49) a. *¿Cuándo tus hermanas se fueron? when your sisters Cl. left.3pl
   ‘When did your sisters leave?’
b. ¿En qué momento tus hermanas se fueron? in what moment your sister Cl. left.3pl
   ‘In which moment did you sisters leave?’

(3.50) a. *¿Cómo tu hermana se cayó de la cama? how your sister Cl. fall.3sg from the bed
   ‘How did your sister fall from the bed?’
b. ¿De qué manera tu hermana se cayó de la cama? in what way your sister Cl. fall.3sg from the bed
   ‘In which way did your sister fall from the bed?’

(3.51) a. *¿Dónde tus amigos se divorciaron? where your friends Cl. got divorced
   ‘Where did your friends get divorced?’
b. ¿En qué lugar tus amigos se divorciaron? in what place your friends Cl. got divorced
   ‘In which place did your friends get divorced?’

(3.52) a. *¿A quién tu hermana visitó? a who your sister visited.3sg
   ‘Who did your sister visit?’
b. ¿A cuál de estas chicas tu hermana (la) había visitado en Sicilia? a which of the girls your sister Cl.do had.3sg visited in Sicily
   ‘Which of the girls had your sister visited in Sicily?’

(3.53) a. *¿A quién Juan le dedicó un poema?
   ¿A cuál de sus profesores de la escuela primaria Juan le dedicó un poema?
   ‘Which of his primary school teachers did John dedicate a poem to?’

adapted from Ordoñez (1998b: 346)
3.2 On Spanish wh-questions

It is not only Ordoñez who shows that complex wh-sentences are less strict in the otherwise obligatory wh-phrase-verb adjacency. Bosque & Gutiérrez-Rexach (2011: 449), Contreras (1999: 1940), Olarrea (1998: 91ff), and Zubizarreta (2001: 184), among others, present the same observations. Interrogatives with heavy or complex wh-phrases allow the subject to appear preverbally. Complex wh-phrases are mostly constructed with quién(es) de ‘who of’, cuál(es) de ‘which of’ and cuánto(s) de ‘how many of’ and have a partitive reading. The same applies to qué/cuál(es)/cuánto(s) + NP ‘who, which, how many + NP’. Interestingly, while the former group can be separated as shown in (3.54), this is not possible with the latter group, as illustrated in (3.55). That is to say, in the first constructions pied-piping is facultative, whereas it is obligatory in the second.25

(3.54) a. i) ¿Quién de tus amigos es diputado?
   who of your friends is member of the parliament
   ‘Which of your friends is a member of the parliament?’
   ii) ¿Quién es diputado de tus amigos?

b. i) ¿Cuáles de esos poemas te gustan?
    which of these poems Cl. like.3pl
    ‘Which of these poems do you like?’
   ii) ¿Cuáles te gustan de esos poemas?

c. i) ¿Cuántos de tus libros vendiste?
    How many of your books sold.2sg
    ‘How many of your books did you sell?’
   ii) ¿Cuántos vendiste de tus libros?

(3.55) a. i) ¿Qué libros de física has leído?
   which books of physics have.2sg read
   ‘Which books about physics have you read?’
   ii) * ¿Qué has leído libros de física?

b. i) ¿Cuáles poemas de Neruda conoces?
    which poems of Neruda know.2sg
    ‘Which of Nerudas poems do you know?’
   ii) * ¿Cuáles conoces poemas de Neruda?

c. i) ¿Cuánto dinero te dieron?
    how much money Cl. gave.3pl
    ‘How much money did they give you?’
   ii) * ¿Cuánto te dieron dinero?

25 Loosely speaking, the term pied-piping refers to the ability of certain elements, like wh-words or quantifiers, to ‘drag along’ the rest of the constituent. Cable writes that “the term ‘pied-piping’ describes cases, where an operation targeting the feature of a particular lexical item applies to a phrase properly containing the maximal projection of that item.” (Cable 2012: 6) I will come back to pied-piping in Section 4.1 and Chapter 5.
3.2 On Spanish wh-questions

Complex wh-questions are often characterized as being \(d\)\(\text{(discourse)}\)-linked (among others, Bosque & Gutiérrez-Rexach 2011, López 2009, Olarrea 1998, Pesetsky 1987). D-linked wh-phrases “must refer to members of a set that both the speaker and hearer have in mind, and are in that sense linked to the discourse” (Olarrea 1998: 103). They must be “\(\text{vinculadas a un discurso previo en la conciencia lingüística del hablante o del oyente}\)” (Bosque & Gutiérrez-Rexach 2011: 449). In order to tell them apart, Olarrea (1998) cites a test by Pesetsky (1987) to distinguish d-linked from non-d-linked wh-phrases. In the case of the former, it is impossible to insert a phrase like the hell or on earth (‘what the hell...’ vs. ‘which book the hell...’). Olarrea confirms a similar behavior with the corresponding expression leches (literally ‘milks’) for his dialect, as shown in (3.56a) and (3.56b). The same seems to be true for sentences with CLLDed constituents (3.56c).

\[
\text{(3.56) a. } \text{¿Dónde leches compró Juan el regalo para María?} \\
\quad \text{where on earth bought John the present for Mary} \\
\quad \text{‘Where on earth did John buy the present for Mary?’}
\]

\[
\text{b. * ¿En cuál de estas tiendas leches Juan compró el regalo para María?} \\
\quad \text{in which of these shops on earth John bought the present for Mary} \\
\quad \text{‘In which of these shops did John buy the present for Mary?’}
\]

\[
\text{c. * A Juan, leches, no le he visto en años.} \\
\quad \text{A John, on earth no Cl. have seen in years} \\
\quad \text{‘I haven’t seen John in years.’}
\]

Olarrea (1998: 103)

In accordance with Olarrea (1998) and Pesetsky (1987), I assume that heavy wh-sentences in the above form are always d-linked. An immediate consequence is the observation made by Bosque & Gutiérrez-Rexach (2011: 449). They state that d-linked wh-phrases can precede adverbs that in turn have to precede non-d-linked wh-phrases. Thus, d-linked wh-phrases apparently move to a higher position in the left-peripheral structure than non-d-linked. The latter have to immediately dominate the projection of the finite verb, whereas d-linked wh-phrases occur in a higher position. That means, I follow Ordoñez (1998b) and Ordoñez & Olarrea (2006) and adopt the idea “that complex wh-words in the preceding examples in Spanish [...] are left dislocated. Therefore, [...] it becomes less surprising that pre-verbal subjects can intervene between the complex wh-word and the verb.” (Ordoñez 1998b: 347). The behavior of CLLDed constituents and complex wh-phrases is very similar. Both constructions are d-linked\(^{27}\) and seem to allow intervening elements between themselves and the verbal complex.

\(^{26}\) (they) have to be linked to a previous discourse in the linguistic consciousness of the speaker or the hearer. (\textit{own translation})

\(^{27}\) In Section 3.1.2, I have described CLLDed constituents as RTops, which means d-linked, as explained in Section 3.1.1.
I will argue in Section 3.3.3 that complex \(wh\)-phrases are very similar to CLLDs. I have already argued, following Bosque & Gutiérrez- Rexach (2011) and López (2009), among others, that they are always d-linked. However, López (2009: 120) mentions two critical points in the comparison with CLLDe elements. First, the non-obligatory clitic doubling in the interrogatives and, second, differences in reconstruction effects between CLLDs and complex \(wh\)-phrases. The first point is somewhat weakened by the example (3.52b), here repeated as (3.57), from Ordoñez (1998b: 347). It seems that a clitic appears to be at least optional. Additionally, López himself mentions that “in other Romance varieties it is possible to have a clitic – Romanian (Dobrovie-Sorin 1994), Rioplatense Spanish.” (López 2009: 120).

(3.57) ¿A cuál de estas chicas tu hermana (la) había visitado en Sicilia?
A which of the girls your sister Cl.do had.3sg visited in Sicily

‘Which of the girls had your sister visited in Sicily?’

Rizzi (1997) points out a similar dichotomy observed with negative quantifiers in Italian. He shows that complex negative quantifiers can involve topicalization whereas simple negative quantifiers cannot. In Olarrea (1998: 73-74) and Ordoñez & Olarrea (2006: 93), they claim the same effect for Spanish. The relevant examples are shown in (3.58)

(3.58) a. A ninguno de estos chicos, María lo conoce bien.
A none of these boys María Cl.do know.3sg well

‘Maria does not know some of these boys well.’

b. * A nadie lo he visto.
A nobody Cl.do have.1sg seen

‘I have seen nobody.’

Ordoñez & Olarrea (2006: 93)

The second point, the reconstruction effect, is more complicated to explain. In Spanish, reconstruction in sentences with a CLLD depends on the position of the subject. If the bindee is in the CLLDe constituent and the binder in the subject phrase, but both originate in the same sentence, binding is only possible if the subject is pre-verbal (cf. (3.59)). Interestingly, topicalized PPs can reconstruct to their original position inside the VP, as shown in (3.60). The following examples from Zubizarreta (1998) illustrate the difference.

(3.59) a. A su propio hijo, ningún padre, lo quiere castigar.
A his own child no father Cl.do wants punish.INF

‘No father wants to punish his own child.’
3.2 On Spanish wh-questions

b. * A su propio hijo no lo quiere castigar ningún padre.
   A his own child no Cl.do wants punish.Inf no father

(3.60) a. Sobre su silla, cada niño puso un libro.
   on his chair each child put.3sg a book
   ‘Each child put a book on his chair.’

b. Sobre su silla puso cada niño un libro.


With fronted complex wh-phrases, reconstruction to the initial merge position is always possible, no matter the position of the subject. For both of the following questions, the answer ‘The youngest one’ is felicitous. The examples are originally in Catalan, but the claim also holds for the Spanish translation.

(3.61) a. ¿A cuál de sus hijos tiene que acompañar cada madre mañana?
   A which of her sons has that accompany each mother tomorrow
   ‘Which of her sons must each mother accompany tomorrow?’

b. ¿A cuál de sus hijos ninguna madre tiene que acompañar mañana?
   A which of her sons NPI mother has that accompany tomorrow
   ‘Which of her sons must no mother accompany tomorrow?’

translated from López (2009: 120)

However, since reconstruction with other CLLDed elements such as PPs is possible, it is safe to assume that the restriction of reconstruction on some CLLDs is due to a different, independent reason.

A final difference can be observed in the case of multiple CLLDs and multiple complex wh-phrases. As we have seen, it is possible to have multiple CLLDs but not multiple fronted wh-phrases. In this case too, one can assume that the distinct behavior is due to an independent restriction, this time of the wh-phrase itself. It prevents the multiple fronting of complex wh-phrases. It would be the same effect that prevents multiple fronting of single wh-words, as we have see in Section 3.2.1. I will return to this observation in Section 3.3.3.

In summary, we have seen that complex wh-questions resemble CLLDs in many aspects. The structural similarities will be discussed in more detail in Section 3.3. In the following section, I will have a closer look at multiple and indirect wh-questions, as well as a short review of some dialectal variation.
3.2 On Spanish \textit{wh}-questions

3.2.3. More on questions in Spanish

3.2.3.1. Multiple \textit{wh}-questions

In Spanish multiple \textit{wh}-questions, only one of the \textit{wh}-phrases has to be fronted. In fact, fronting more than one results in an ungrammatical sentence (see (3.62a) and (3.62b)). It is possible, however, to front more than one \textit{wh}-phrase if they are coordinated. Nevertheless, the coordinated fronting is subjected to some restrictions and is only grammatical in the following cases: the coordinates are adverbiales (3.63a) or the coordinates are either both parts of the subject (3.64a), the direct object (3.64b), or the indirect object (3.64c). It is not possible to mix different arguments or adverbiales ((3.65a) and (3.65b)).

(3.62)  
\begin{enumerate}[a.]
\item ¿Quién compró qué?
  \begin{tabular}{l}
  who \ \\
  bought.3sg what
\end{tabular}
‘Who bought what?’
\item * ¿Quién qué compró?
\end{enumerate}

Bosque & Gutiérrez-Rexach (2011: 51)

(3.63)  
\begin{enumerate}[a.]
\item ¿Cuándo y cómo piensas pagar ese dinero?
  \begin{tabular}{l}
  when \ \\
  how \ \\
  think.2sg pay.inf that money
\end{tabular}
‘When and how do you think to pay that money’
\item ¿Cuántos estudiantes y cuántos profesores fueron a la fiesta?
  \begin{tabular}{l}
  how many students \ \\
  and \ \\
  how many teachers \ \\
  went.3pl to the party?
\end{tabular}
‘How many students and how many teachers went to the party?’
\end{enumerate}

(3.64)  
\begin{enumerate}[a.]
\item ¿Qué estudiantes y qué profesores conoces?
  \begin{tabular}{l}
  what students \ \\
  and \ \\
  what teachers \ \\
  know.2sg
\end{tabular}
‘Which students and which teachers do you know?’
\item ¿A qué editoriales y a qué periódicos les mandaste el manuscrito?
  \begin{tabular}{l}
  to what publishers \ \\
  and \ \\
  to what newspapers \ \\
  Cl.10 sent.2sg \ \\
  the manuscript
\end{tabular}
‘To which publishers and to which newspaper did you sent the manuscript?’
\end{enumerate}

(3.65)  
\begin{enumerate}[a.]
\item * ¿Quién y cómo vive allí?
  \begin{tabular}{l}
  who \ \\
  and \ \\
  how \ \\
  live.3sg here
\end{tabular}
‘Who lives here and how?’
\item * ¿Cuándo y a quién viste?
  \begin{tabular}{l}
  when \ \\
  and \ \\
  a who \ \\
  saw.2sg
\end{tabular}
‘Who did you see when?’
\end{enumerate}

Contreras (1999: 1942-1943)

In non-coordinated multiple \textit{wh}-interrogatives, the order of the \textit{wh}-phrases is also restricted depending on their function. In a mix of adverbia and argument \textit{wh}-phrases, the latter normally appears
at the end of the sentence (3.66).²⁸ In the case of two or three argument wh-phrases, the order is normally the canonical order S-DO-IO, but may alter²⁹ (see (3.67a) - (3.67c)) (Bosque & Gutiérrez- Rexach 2011: 512ff, Contreras 1999: 1942ff). Spanish multiple wh-questions are not constraint by superiority effects as English is. In the latter, a question like *What did who buy?* is ungrammatical, because the direct object precedes the subject wh-word.

(3.66) ¿Cuándo viene quién?
when comes who
‘Who comes when?’

(3.67) a. ¿Quién conoce a quién?
who knows who
‘Who knows whom?’
b. ¿Qué le diste a quién?
what Cl.10 gave.2sg to whom
‘What did you gave to whom?’
c. ¿Quién le dio qué a quién?
who Cl.10 gave.3sg what to whom
‘Who gave what to whom?’

Contreras (1999: 1942-1943)

At this point, I draw attention to the large similarity between wh-fronting and FF. Both constructions normally demand obligatory adjacency between the verb and the wh-phrase/FFed constituent and do not get doubled by a resumptive clitic, contrary to CLLD or HTLD. They are both subject to similar restrictions concerning the extraction out of syntactic islands. It seems that they are in a complementary distribution regarding their appearance, but share the same distributions with other LD. In neither case is it possible to front two constituents of the same kind, that is to say, multiple wh-phrases or multiple FFed constituents are ungrammatical. Thus, whether to assume similar or even identical constructions and mechanisms behind both phenomena, or not, is an ongoing debate (see discussion and references cited in Bosque & Gutiérrez-Rexach 2011: 695ff, López 2009, Olarrea 1998: 67, Zagona 2002: 250ff).

²⁸ The adverbial wh-phrases can also appear sentence final, but with varying acceptability. As we have seen in (3.47b), *por qué* ‘why’ is not accepted, whereas dónde ‘where’, cuándo ‘when’ and other adverbial wh-phrases are (Contreras 1999: 1942).

²⁹ Bosque & Gutiérrez-Rexach (2011: 514) say that the non-canonical orders IO-S or DO-S is not accepted by all speakers but vary depending on the dialect. That is, (i) would be ungrammatical for them, but not for all speakers.

i. ¿Qué dijo quién?
what said.3sg who
‘Who said what?’
3.2 On Spanish wh-questions

3.2.3.2. Indirect questions

Indirect wh-sentences in Spanish behave very similarly to their direct counterparts. They also exhibit wh-fronting and obligatory adjacency between the wh-phrase and the verb (Contreras 1999: 1947ff, Francom 2012: 538, Zagona 2002: 547, Torrego 1984: 104-105). They sometimes can be introduced by que ‘that’. Although not all matrix verbs allow this variation (Contreras 1999: 1949)30. Examples (3.68) - (3.69) show the main characteristics of embedded wh-questions.

(3.68) Obligatory wh-phrase-verb adjacency

a. No sabían qué querían esos dos.
   not knew.3pl what wanted.3pl those two
   ‘They didn’t know what those two wanted.’

b. * No sabían qué esos dos querían.

(3.69) Introduction with que

a. Me pregunté que dónde vivía.
   Cl. asked.3sg that where lived.3sg
   ‘I asked myself where they lived.’

b. * No sabía que cuándo terminaría la carrera.
   not knew.3sg that when finished.3sg the studies
   ‘He/she didn’t know when the studies would finish.’

Therefore, Spanish differs from English, German oandr other wh-fronting languages in that the obligatory adjacency applies to embedded wh-questions as well. It is a combination that is not very common among the wh-fronting languages of the world (Emonds 1976, Francom 2012: 538).

Also multiple wh-questions can occur in embedded interrogatives. As with simple embedded questions, the matrix verb must be compatible with an indirect question. The multiple embedded interrogatives are restricted in the same way as their direct counterparts.

In the next section, I will present some dialectal variations, concentrating on Caribbean Spanish, which shows some major differences in their properties.

3.2.4. Dialectal variation

Caribbean Spanish, mainly spoken in Cuba, the Dominican Republic, and Puerto Rico, differs significantly form peninsular Spanish. Various authors have shown that subjects may intervene between

30 Contreras (1999: 1949) states that this is probably due to the inability of verbs like saber ‘know’ to report a direct quotation.
the wh-phrase and the verb (see e.g., Contreras 1999, Goodall 1993, 2004, forthcoming, 2010, Gutiérrez-Bravo 2008, Lantolf 1980, Lipsky 1977, Ordoñez & Olarrea 2006, Suñer 1994, Toribio 1993). Although this phenomena is widely attested, there is an ongoing debate on the kind of subjects that can intervene. Some authors have claimed that the phenomena is restricted to personal pronouns, especially tú ‘you’, while others go as far as to accept all kinds of full subject DPs (for a detailed discussion see Ordoñez & Olarrea 2006: 67ff, Zimmermann forthcoming). That is, questions like the following from Ordoñez & Olarrea (2006: 60/61) and Toribio (1993: 26) seem to be acceptable for most speakers of Caribbean Spanish.

(3.70)  

(a) ¿Qué tú comes?  
what you eat.2sg
‘What do you eat?’

(b) ¿Dónde yo he dejado los espejuelos?  
where I have.1sg left the glasses
‘Where have I left my glasses?’

(c) ¿Qué ellos trajeron a la fiesta?  
what they brought.3pl to the party
‘What did they bring to the party?’

(d) ¿Cuánto un medico consume en un mes?  
how much a doctor uses in one month
‘How much does a doctor use in one month?’

In order to explain the non-adjacency, especially with personal pronouns, Ordoñez & Olarrea (2006) claim that the latter have developed into weak pronouns according to the classification of Cardinalletti & Starke (1999). Therefore, they can appear pre-verbally. Goodall (2010) on the other hand attributes the differences in the acceptance of the various intervening subjects to an “interplay of syntactic properties of the language and processing considerations, especially the idea that processing a wh-dependency and a discourse referent both make use of the same set of limited shared resources.” (Goodall 2010: 245). Since I will work with Peninsula Spanish and a possible influence of Basque, I will not go into more details here but refer the reader to the ongoing debate in the literature. Zimmermann (forthcoming) relates the phenomena to a similar process in medieval French. He compares the Caribbean dialect with an intermediate step in the development from Old French, a NSL, to modern French, a non-NSL (for a detailed discussion see Zimmermann forthcoming).

Having presented the main characteristics and restrictions of wh-interrogatives in Spanish, I will now explain more detailed the syntactic structures underlying the different constructions.
3.3 On syntactic structures in Spanish *wh*-questions

When we talk about *wh*-questions in Spanish, the two main questions that have to be answered are the following: Why are *wh*-phrases fronted and where do they move to. Both questions have been explored and discussed widely by the linguistic community throughout the last decades. The first question of why *wh*-phrases move is answered by generative grammar and, more precisely, the minimalistic framework with the notion of feature probing and feature checking. Since Rizzi’s (1990) seminal work about *wh*-fronting, many scholars have followed his idea of a [+wh] feature in the CP triggering movement of the *wh*-phrase in order to fulfill the *wh*-Criterion. Others have claimed that prosody is enough to trigger movement (among others, Richards 2010, Zubizarreta 1998). I will follow Cable (2010), Kempchinsky (2013), and López (2009), and subsequent work, assuming that movement is triggered by feature checking/valuation.

More complex is the question where the *wh*-phrase and the verb move to. In English, for example, it is relatively clear that both the *wh*-phrase and the verb move to the CP-layer (among others, Rizzi 1990). However, in Spanish, the situation is not as transparent for various reason, one being that the position and status of preverbal subjects is not very clear. Some authors claim that the preverbal subject position is always an A′-position (e.g., Barbosa 2001, Casielles Suaréz 2001, Ordoñez 1998b, Zubizarreta 2012) whereas others argue for a mixed status dependent on the context and information structure (e.g. López 2009, Olarrea 2012, Suñer 2003). I will give a short overview on preverbal subjects in Spanish in Section 3.3.1. In Section 3.3.2 I will first present the most known approach to the left periphery, namely Rizzi’s cartographical project (Section 3.3.2.1), and then, in Section 3.3.2.2, introduce a more derivational approach used by Kempchinsky, López, and subsequent work.

3.3.1 Spanish pre-verbal subjects

At the beginning of this chapter, we have seen that the word order in Spanish is relatively free. In a transitive sentence, all logical combinations except SOV are possible. The basic word order, however, also called unmarked word order, is considered to be SVO, as in (3.1a), repeated as (3.71). The reason for this is that the word order in the most natural answer to the question ‘What happened?’ is SVO. Such an answer is also called an all-focus-sentence, because no element bears special focalization or contrast.

Context: What happened?

(3.71) Pedro comió una manzana.
Peter ate.3sg a apple
‘Peter ate an apple.’
Although the basic word order is SVO, there are certain circumstances in which the verb has to precede the subject even in an all-focus sentence. Olarrea (2012: 608) describes six cases in which the answer to the question ‘What happened?’ has the so-called non-canonical word order VS(O). They are all listed in (3.72) along with corresponding examples.

(3.72) A. when the subject is a bare NP
   a. Llegaron niños.
      arrived.3pl children
      ‘There arrived children.’
   a’.
   # Niños llegaron.

B. in certain infinitival and imperative constructions
   b. Al salir el sol ...
      to-the rise.INF the sun ...
      ‘As the sun rises …’
   b’.
   * Al sol salir ...
   b”.
   Que lo diga Pedro
   that CL do say.3sg Pedro
   ‘(I command) for Pedro to say it.’
   ch”’.
   # Que lo diga Pedro.

C. with unaccusatives like llegar ‘to arrive’ or crecer ‘to grow’
   c. Han crecido flores
      have grown flowers
      ‘(Flowers have grown.’
   c’.
   # Flores han crecido.

D. with SE passives and in middle voice constructions
   d. Se convocaron los premios Goya
      refl announced.3pl the awards Goya
      ‘The Goya awards were announced.’
   d’.
   # Los premios Goya se convocaron.

E. with verbs like faltar ‘to lack/to be missing’ or ocurrir ‘to happen’
   e. Ocurrieron demasiadas catástrofes.
      occurred.3pl too.many catastrophes
      ‘Too many catastrophes occurred.’
   e’.
   # Demasiadas catástrofes ocurrieron.

F. with psychological verbs like gustar ‘to like’, asustar ‘to frighten’, preocupar ‘to worry’
   f. Me gustan las películas de Billy Wilder
      refl like.3pl the movies of Billy Wilder
      ‘I like Billy Wilder’s movies’
   f’.
   # Las películas de Billy Wilder me gustan.

Olarrea (2012: 608)
3.3 On syntactic structures in Spanish wh-questions

In the examples we see that the second utterances are always infelicitous as a response to the out-of-the-blue-question ‘What happened?’ Preposing the subject is possible in all cases except (3.72B), although only in conjunction with a contrastive reading.

These facts reflect the still ongoing debate about the position of the subject in Spanish and its preverbal status. Since the introduction of the VP-internal subject hypothesis (VPiSH) by Koopman & Sportiche (1991), Kuroda (1988) and others, it has become general consensus that the subject is first merged internally within the VP-shells. Then the verb moves to the head of IP, leaving the subject in a post-verbal position, deriving the word order of sentences like (3.72). If the subject is merged VP internal and if it is possible to leave it there, it seems that a pre-verbal subject is not obligatory. Therefore, Alexiadou & Anagnostopoulou (1998), Barbosa (2001), Ordoñez (1998a), Ordoñez & Treviño (1999), and Zubizarreta (1998, 1999b), among others, assume that the A-position for subjects is VP-internal and that SpecIP is not an argument position but an A’-position, with the result that pre-verbal subjects would always be some kind of topicalized element. Some of these authors even claim that in the case of a projected CP, IP is Spec-less (among others, Barbosa 2001). I will come back to this point later on.

Prominent work on that behalf was done by Alexiadou & Anagnostopoulou (1998) for Greek, which is also a Null-Subject-Language (NSL). They show that preverbal subjects in Greek are always in a topic position. They apply the same argumentation to Spanish and other NSLs and claim the same status for the pre-verbal subjects. Barbosa (2001), López (2009), Ordoñez & Treviño (1999), and Zubizarreta (1999b) follow their line of work, assuming that SpecIP is an A’-position. Ordoñez & Treviño (1999) compare the congruence of preverbal subjects and left-dislocated objects in detail. They show convincingly that both behave very similar.

Suñer (2003), on the other hand, assumes that pre-verbal subjects can be topics, but not in an unmarked SVO sentence. She refutes Alexiadou & Anagnostopoulou’s (1998) interpretational facts and binding conditions for Spanish. Instead, she claims that the pre-verbal subject in SpecIP is in an A-position and that pre-verbal topicalized subjects move to a higher position in the left periphery. In order to explain the differences, she refers to and expand the specialization hypothesis by Cardinaletti (1997) and Cardinaletti & Roberts (1991). She assumes two different specifier positions in IP. The subject moves to the first or the second, depending on its type. Full DP subjects and heavy pronouns move to the upper specifier, weak pronouns to the lower. In this way, she explains some minor differences in the distribution of the various subject types.31

31 The differences can be shown by inserting a parenthetical between verb and subject as in (i.) and (ii.) (Suñer 2003: 351):
3.3 On syntactic structures in Spanish *wh*-questions

I will follow the line of Kempchinsky (2013), López (2009), and Olarrea (2012). They state that it is without questions that pre-verbal subjects can be dislocated. At the beginning of a discourse, however, or in the case of an all-focus sentence, the pre-verbal subject is not topicalized but A-moved, most likely to SpecIP.

In my work, I focus on *wh*-interrogatives and sentences that already contain some topicalized structures (e.g., FFed constituents or CLLD). This means that for the remaining work, if not otherwise indicated, subjects are in a topic position. In the following section, I want to explore the left periphery and possible positions of CLLD, FFed elements and, of course, *wh*-phrases.

### 3.3.2. The left periphery

#### 3.3.2.1. Rizzi’s rich left periphery

One of the most influential works on the left periphery is undoubtedly Rizzi’s (1997) “The Fine Structure of the Left Periphery”. He presents and analyses a complex functional field at the sentence’s left periphery, traditionally known as CP. He claims that it is richer and more complex than in former approaches, consisting of a whole batterie of different functional phrases. The main structure of this CP as presented in Rizzi (1997: 297) is shown in (3.73).

---

**i.** Juan/Él, a mi parecer, es muy simpático.  
Juan/He to me seem.inf is very nice  
’Juan/He, according to me, is very nice’

**ii.** Ello (‘a mi parecer) no sería malo estudiar.  
It to me seem.inf not is.3sg.conj. bad study.inf  
’It, according to me, wouldn’t hurt to study’
Rizzi’s rich left periphery

The left field is framed by the Force Phrase (ForceP) on the very left and the Finiteness Phrase (FinP) on the right. The former encodes the sentence type, meaning the information which expresses the clause type or the force of a proposition, e.g. a question, an assertion, an exclamation, etc. Rizzi states that FinP therewhile contains “the core IP-related characteristics that the C system expresses”. That means, FinP works as “link between the C-system and the Infl-system” (Fortuny 2008). ForceP and FinP are always projected, since they are essential for the sentence’ processing. Nevertheless, as López states, “Force and finiteness are usually fused in one functional head and they are split into two heads only when the sentence includes a topic (a CLLD), a fronted focus or a wh-phrase.” (López 2009: 105).

In Rizzi’s approach, Topic Phrase (TopP) and Focus Phrase (FocP) are both functional categories carrying respectively a [+top] and [+foc] feature, which need to be satisfied, or in a minimalistic sense, have to enter a probe/goal relation with some other element carrying the same feature. Therefore, if a constituent bears a [+foc] or [+top] feature, it moves leftwards to the specifier of the corresponding phrase. In this approach, complementizer such as que ‘that’ merge in Force\(^0\), whereas wh-phrases and focus phrases move to the same SpecFoc, since they appear only in complementary distribution. Rizzi assumes that the Force-Fin pair is always projected, contrary to the Topic-Focus pair, which is only projected when needed. The asterisk on the TopPs indicates a recursive structure, as more than one TopP can be adjoin there.

Rizzi’s complex left periphery is accepted and used by a wide range of Spanish linguists, from Zubizarreta’s work on the Spanish left periphery (e.g. “The Left Edge in the Spanish Clausal Structure”
3.3 On syntactic structures in Spanish wh-questions


The way Rizzi structures the left periphery accounts for a wide range of the already presented LD phenomena in Spanish. The higher TopP for example can host CLLDs and, due to its recursive properties, there is no problem in hosting more than one. FFed constituents move to SpecFoc. In this way, it is secured that CLLDs precede FFed constituents. In addition, it also accounts for the complementary distribution of FFed constituents and fronted wh-phrases, since both move to the same specifier, namely SpecFoc. Using the same structural position is also a plausible explanation for their similar behavior and properties.

In order to explain the constraints on the movement of the wh-phrase, Rizzi introduces the *wh-Criterion*. It states as followed:

\[(3.74) \text{ The } wh\text{-Criterion by Rizzi (1996: 64):} \]

\begin{enumerate}
  \item A wh-operator must be in spec-head configuration with \( X^0 \) \[+\text{wh}\]
  \item An \( X^0 \) \[+\text{wh}\] must be in a spec-head configuration with a wh-operator
\end{enumerate}

The wh-Criterion demands that a constituent in a SpecX bearing a [+wh] feature needs a second element bearing the same feature in its head \( X^0 \). Rizzi (1996: 64) states that “Wh-Criterion simply expresses the fact that at the appropriate level of representation interrogative operators must be in the spec of CPs that are interpreted as questions and, reciprocally, CPs interpreted as questions must have interrogative operators as specifiers”. That means that in a wh-interrogative, the wh-phrase, also called the wh-operator, bears a [+wh] feature and has to occupy the SpecCP position at some point in the derivation. In addition, he proposes that the verbal inflection \( I^0 \) carries also a instance of [+wh]. To support this assumption, he refers to natural languages which have a special verb morphology in interrogatives. Thus, in order to satisfy the wh-criterion, the wh-phrase has to move to SpecCP and the main verb to \( C^0 \). This is called I-to-C-movement. In the cartographic approach presented in (3.73), they both move to FocP to fulfill the wh-Criterion. In embedded questions, the bearer of the [+wh] feature is the matrix verb. Thus, verb movement to \( C^0 \) in the embedded sentence is no longer necessary. The wh-phrase can check its [+wh] feature against the matrix verb. This accounts nicely for the asymmetry between root and embedded questions in English and other Germanic languages.

However, the I-to-C-movement analysis falls short in explaining some of the core features of Spanish and other Romance languages’ interrogatives. First, the asymmetry between matrix and embedded sentences is not attested for. Furthermore, under certain circumstances, Spanish allows the subject to
intervene between the verb and the wh-phrase, as we have seen in Section 3.2. Rizzi’s (1997) approach has difficulties to explain these observations. In the case of por qué ‘why’, he proposes in Rizzi (2001) that these wh-phrases are not base-generated inside the VP/IP, but are directly merged in the left periphery. He argues that these kind of high adjunct wh-phrases do not need to check selectional or interpretive requirements, unlike argument or lower adverbial wh-phrases. The latter have to be first merged in the VP or IP area and then moved to the left periphery. However, this explanation does not work for complex wh-phrases.

In order to explain these observations, several authors have claimed that the verb does not move to the C-complex in Spanish wh-questions at all. The most detailed analysis is given by Suñer (1994). She supports the claim with a wide range of arguments, showing that the verb stays in IP. The core arguments are exactly the former mentioned exceptions of obligatory adjacency. That is, the differences in adverb placement, the inversion in embedded interrogatives, and the facultative wh-phrase-verb adjacency in some cases.\(^{32}\) I have given examples for all three arguments in Section 3.2.2. Barbosa (2001) and Suñer (1994), among others, assume that only the wh-phrase moves to the SpecIP and the verb stays in \(\text{I}^0\). Thus, the wh-Criterion can be satisfied inside the IP without the need of projecting any part of the CP.

A further difficulty with Rizzi’s cartographic approach is “that it puts into the syntactic component notions that are in fact discourse notions” and that “overall the cartographic approach seems to over-determine the discourse roles of constituents at the left edge of the clause, without really explaining them” (Kempchinsky 2013: 312-313). That is, labelling phrases with discourse connected notions like topic and focus can result in an overrating of the syntax. Especially if we “consider, for example, categories such as ‘scene-setting adverbials’, ‘list distribution topic’, etc., from Benincà and Poletto (2004), and similar analyses” (Kempchinsky 2013: 312). López (2009) also refers to the difficulty of having various topics. Rizzi’s TopP divides the sentence into a topic (SpecTop) and a comment (complement of Top\(^0\)). The problem López points out is that in the case of various topics, the topic-comment-structure isn’t clear anymore. López states “that altering the order of dislocates should lead to altering the topic-comment structure of the clause. However, no such altering is apparent, as far as I can tell” (López 2009: 123). A further point against the Topic-Focus structure made by López is the incompatibility of wh-phrases and FFed constituents. In Rizzi’s account, the prediction would be that a fronted focus should be compatible with an in-situ-wh-phrase, or vice versa. López shows that this prediction is not fulfilled.\(^{33}\) He concludes that it is not a structural restriction preventing the co-occurrence but a pragmatic one. That

\(^{32}\) For a detailed discussion see Suñer (1994) and Ordoñez & Olarrea (2006: Ch. 2)

\(^{33}\) Contrary to multiple wh-phrases, as we have seen in Section 3.2.1.
is, a sentence cannot be a correction, which is an assertion, and a question at the same time (López 2009: 124).

Rizzi’s (1997) rich left periphery is widely used despite the difficulties and some unresolved issues. The analysis is often used as starting point and base for a divergent theory and its influence is seen throughout the linguistic literature, especially, but not exclusively, in Romance languages. Cable (2010) also uses the Force-Fin system in his very different approach to wh-fronting, as I will show in Chapter 5. I will not follow Rizzi’s approach directly, but López’s (2009) and especially Kempchinsky’s (2013) work. They both assume a less detailed and complex left periphery than Rizzi, but adopt his Force-Fin system as the ‘frame’ of the CP-area. Furthermore, in Chapter 5, I will present a Q-particle based account of wh-interrogatives and combine it with Kempchinsky’s (2013) analysis. Finally, I will show that in the Q-particle based approach, no wh-Criterion is necessary anymore. In the next section, I will present Kempchinsky’s (2013) approach in more detail.

3.3.2.2. Kempchinsky’s approach to the left periphery

Kempchinsky (2013) presents a new approach based on the work by Emonds (2004), López (2009), and Uriagereka (1995), among others. Her analysis uses data from clitic placement in Asturian and other Western Iberian languages. In a nutshell, she derives the different landing positions of CLLDs and preverbal subjects on the one hand, and FFEd elements and wh-phrases on the other hand, from different clitic placement restrictions. In Asturian, the former yield obligatory enclisis whereas the latter proclisis. From these differences, she deduces that CLLDs and preverbal subjects have to be in a higher position in the left periphery than the FFed material and wh-questions. The former move to a so-called Discourse Shells (DS), introduced by Emonds (2004) and the latter, following López (2009), land in SpecFin. I will explain this analysis in more detail.

Kempchinsky follows López (2009) and assumes that a Force-Fin frame limits the CP area. But contrary to López (2009), she does not adopt the idea of multiple specifier in FinP in the case of several LDs, but falls back on the idea of Discourse Shells by Emonds. She supports her analysis with Fernandez-Rubiera’s (2009) data about clitic placement in Asturian, a Western Iberian language spoken in Asturias. Asturian shows two patterns of clitic placement, namely proclisis and enclisis. Most importantly, it displays enclisis in declarative sentences, which is excluded in Spanish and most other Romance languages. Enclisis is obligatory if the sentence is V-initial or has a preverbal subject or a CLLD, whereas proclisi-

---

34 For a detailed discussion on some more difficulties with the cartographical approach see Kempchinsky (2013) and López (2009)
3.3 On syntactic structures in Spanish wh-questions

sis is obligatory with fronted affective elements, FF and fronted wh-phrases (Kempchinsky 2013: 315, Fernandez-Rubiera 2009: Ch. 1). The following two examples from Asturian exhibit enclisis with a CLLD (3.75) and proclisis (3.76) with a wh-phrase.

(3.75) a. Eso sábeslo per ller les cartes.
   that know.2sg for read.inf the (tarot) cards
   ‘That you know it from reading the cards.’
   b. * Eso lo sábes per ller les cartes.

(3.76) a. Cómo t’atreves?
   how dare.2sg
   ‘How dare you?’
   b. * Cómo atreveste?

Kempchinsky (2013: 316)

Kempchinsky adopts a basic approach to enclisis “according to which enclisis occurs because a clitic cannot be the initial element in some XP identified by the phonology as an Intonational Phrase. [...] That is, if there is no constituent preceding the clitic within XP, then the V must precede it, yielding enclisis.” (Kempchinsky 2013: 316). This approach to the enclisis/proclisis leads to the conclusion that CLLDe ded elements and FFed constituents/wh-phrases cannot land in the same place inside the CP. The former must be higher up in the hierarchy than the latter. In order to account for the cliticization facts, Kempchinsky adopts the analysis by Raposo & Uriagereka (2003). They assume a functional projection FP above IP which is headed by an abstract clitic f. f can be [−syntactic] or [±syntactic] and in the latter case [±morphological]. The idea is that if f is [±morphological] “movement of some constituent to Spec,FP will yield proclisis; in the absence of such movement, verb fusion will result”, that is, enclisis.35 (Kempchinsky 2013: 317). Kempchinsky associates this FP with Rizzi’s (1997) FinP. This way, FF/wh-movement yields proclisis. Besides this, CLLDs must be located in a higher position, since they yield enclisis. If we do not want to fall back on Rizzi’s rich left periphery, we need “an approach according to which syntactic projections are not co-terminous with discourse notions” (Kempchinsky 2013: 317). Various authors who use Raposo & Uriagereka’s (2003) approach have noted that CLLDed constituents have to be in a separate intonational phrase above the FP (see among others, Ž. Bošković (2004) for Serbo-Croatian and Barbosa (1996, 2000) for European Portuguese). Then the question is what position this could be.

Kempchinsky (2013: Sec.4) introduces the notion of Discourse Shells (DS), first proposed by Emonds (2004). Emonds uses DS to host “iterative base-dislocations in languages such as English and French.

35 For a detailed analysis of the f-parameter see Raposo & Uriagereka (2003)
Discourse Shells are unspecified [category-neutral] projections which immediately dominate IPs\(^{36}\) (Kempchinsky 2013: 318). (3.77) illustrate the structure of such an DS.

![Diagram of DS](image)

adapted from Kempchinsky (2013: 318)

According to Kempchinsky, FinP is the relevant domain for cliticization restrictions in Asturian. That is, depending on the presence or absence of material in SpecFin, the structure will yield proclisis or enclisis, respectively. If we follow Kempchinsky “assuming that the only difference between Western Iberian and languages such as Spanish or Catalan, is the lack of [+morphological] \(f\), then by extension CLLD dislocates in Spanish are also located in Spec,X. [SpecDS]” (Kempchinsky 2013: 318). Kempchinsky assumes that the clitic plus its double (the constituent to be CLLDed) first move to Spec\(\nu\)P, then the clitic cliticizes itself to I. Finally the double moves to the DS. That is, \(f\), the clitic and the double are in an agreement relation, as shown in (3.78).

![Syntax diagram](image)

Kempchinsky (2013: 320)

The question remains why the double does not stay in SpecFinP. Kempchinsky (2013: 320) states that “it is this chain of Agree relations that makes it possible for the CLLD dislocate [...] to skip Spec,Fin – because this dislocate is in an Agree relationship with \(f\) in the head of Fin, via the clitic”. It would follow that CLLDed constituents and preverbal subjects in Spanish move to the DS whereas FF\(\bar{e}\)d elements and \(wh\)-phrases to SpecFin, yielding the following syntactical structure:

\(^{36}\) Emonds (2004) specify IPs as Discourse Projections. In this sense, “root clauses are the quintessential Discourse Projections: clauses for which the speaker claims that the proposition expressed by the clause is a "real event”— [...] clauses which the speaker asserts to be true.” (Kempchinsky 2013: 318). Kempchinsky differ in that she assumes the FinP to be the Discourse Projection, “taking Fin to be the locus of deixic linking to the speech context” (Kempchinsky 2013: 318).
If we accept the proposed analysis, the following question arises immediately: if the wh-phrase/FFed constituent move to SpecFin and the verb stays in I₀, why can the subject not move to SpecIP and thus intervene between both? Several authors give different answers to this question. López (2009: 112) rejects that possibility by stating that “Spec,T is busy with other matters”. Suñer (1994: 358) assumes “a strict version of the Extended Projection Principle (EPP), which demands that when the subject remains within the VP an expletive pro must occupy SpecIP.” Barbosa (2001: 36-37) suggests that “Fin is the highest head in the inflectional system, which is equivalent to suggesting that Fin selects a Spec-less IP or NegP. Taking INFL to be a cover term for set of inflectional heads, I include Fin in this set.” That is, she assumes that in the case of a wh-movement or FF SpecIP is not projected. I will follow Barbosa (2001) and assume that in the cases of a wh-phrase or a FFed constituent in SpecFinP, the I₀ projects a Spec-less IP. If the sentence structure requires an extended left periphery, as in the case of a CLLDed element but without FF or wh-movement, the subject can move to SpecFinP or stay in-situ.
3.3 On syntactic structures in Spanish wh-questions

in its post-verbal position. If the sentence is a ‘simple all-focus’ sentence, no extended CP is necessary and I\(^0\) does project an IP with a specifier, to which the subject can move.

To illustrate the established structure for the left periphery, I will give the derivation for a sentence with a simple wh-phrase or a FFed constituent in (3.80) and for a sentence with a CLLDed element in (3.81).

(3.80) simple wh-question/FF

a. ¿Qué compró Pedro en la tienda?
   what bought.3sg Peter in the shop
   ‘What did Peter buy in the shop?’

b. Un libro compró Pedro en la tienda.
   a book bought.3sg Pedro in the shop
   ‘Pedro bought a book in the shop.’

c. [Qué/Un libro] \(k\)

\(\text{wh-movement/ FF}\)

---

Barbosa (2001: 52) assumes that even in the case of a simple wh-question, no further CP structure is needed. She claims that the wh-phrase can move to SpecIP and check the wh-criterion against the highest Infl-head. This way, simple questions would be bare IPs. I think it makes sense to reduce the necessary structure as much as possible. Since I look especially at complex wh-phrases and claim that they are truly CLLDed phrases, I always need the extended CP structure. Nevertheless, for simple wh-questions, nothing would interfere with such an analysis.
In summary, I adopt Kempchinsky’s idea of Discourse Shells that host CLLDs or preverbal subjects, whereas FFed elements and wh-phrases move to SpecFin not allowing intervening material between themselves and the verb.

Having presented the main properties of the analysis, I will come back once more to complex wh-phrases and high adjunct wh-phrases like por qué.

3.3.3. Complex wh-phrases in the left periphery

The proposed analysis accounts for the core properties of CLLD, wh-movement, and FF, presented in Section 3.1 and Section 3.2. FFed constituents and wh-phrases move to SpecFin in order to check the necessary features.\(^{38}\) Both movements are mutually exclusive, as I adopt the widely accepted assumption of a single specifier position, contrary to López (2009). wh-phrases and FFed constituents have to be strictly verb adjacent, because FinP dominates IP directly and SpecIP is, as already mentioned, not

\(^{38}\) I will come back to feature checking and the kind of features in more detailed in Chapter 5.
available as landing site for other elements. CLLD move to the specifier of the Discourse Shell, resulting in the desired order CLLD < wh-phrase/FFed element. DS are iterative structures which account for the phenomena of multiple CLLDs. ForceP, encoding the force or type of the proposition, always dominates the complete left periphery.

We have seen in Section 3.2.2.3 that complex wh-phrases are less strict concerning the wh-phrase-verb adjacency. How can we account for this observation within the analysis? As already mentioned, Ordoñez (1998b) and Ordoñez & Olarrea (2006) claim that they are actually CLLDs. Data from Bosque & Gutiérrez-Rexach about adverb placement supports this claim. They observe that d-linked wh-phrase have to be in a higher position than non-d-linked, because the former are compatibly with a wider range of sentential adverbs, contrary to the latter (Bosque & Gutiérrez-Rexach 2011: 449). That is to say, in a d-linked questions like (3.82a), the introduction of the sentential adverb todavía ‘still’ sounds fine whereas in a non-d-linked question like (3.82b), the adverb probablemente “probably” does not.

(3.82)  a. ¿En qué lugares de la costa mediterránea todavía quedan playas sin rasca-cielos?
      ‘In which places at the mediterranean coast are still beaches left without skyscrapers?’
      b. *¿A quién probablemente ha llamado María?
      ’Whom has María probably called?’

Bosque & Gutiérrez-Rexach (2011: 499)

Assuming Kempchinsky’s model, there are only two possible positions available, namely SpecDS and SpecForce. SpecForce is always projected, whereas SpecDS is not. The question that arises is that if a complex wh-phrase would move to SpecForce, why should other wh-phrases not move there. Considering the ordering restrictions on wh-phrases and CLLDs, I conclude that movement to SpecForce can be excluded. That leaves us with the last option, namely the specifier of the DS. We have seen in example (3.52b), here repeated as (3.83a), that the subject can surface between the verb and the wh-phrase. If we assume that the complex wh-phrase moves to SpecDS, a preverbal subject can appear in a second DS directly preceding FinP. The structure is presented in (3.83b).
3.3 On syntactic structures in Spanish wh-questions

(3.83) a. ¿A cuál de estas chicas tu hermana (la) había visitado en Sicilia?
   ‘Which of the girls your sister had visited in Sicily?’

b. [A cuál de estas chicas]Aquáldeestáschicas
   [tu hermana]tuhermanalalabahadvisitado
   [había visitado]había visitado
   IP
   vP
   vP'
   VP
   CLLD
   CLLD
   DS1
   DS2

In order to satisfy the necessary feature relation, I assume that the wh-phrase stops in SpecFin and in a second step move higher up to DS. This would also account for the observed reconstruction effects, presented and discussed earlier in Section 3.2.2.3. Furthermore, I follow Ordoñez in that “complex wh-constituents are left dislocated and that the quantifier part is moved at LF in order to obtain the needed quantificational interpretation” (Ordoñez 1998b: 347), resulting in the following structure.

(3.84) ¿A cuál [DS t de las chicas] tu hermana la había visitado antes?
   ‘Which of these girls had your sister visited before?’

adapted from Ordoñez (1998b: 348)

Therefore, I adopt the idea that complex wh-phrases are actually CLLDs which first move to SpecFin and then continue to SpecDS.

The other prominent exception of wh-phrase-verb adjacency, namely adjunct wh-phrases like por qué ‘why’, can also be explained by this approach. I have already mentioned in Section 3.3.2.1 that
Rizzi (2001) proposes that these type of wh-adjuncts are first merged in the CP-area. He argues that contrary to argument or low wh-adjunct phrases, these types of high adjuncts do not need to check selectional requirements. That is, they are not merged inside the VP/IP. Several arguments, already presented in Section 3.2.2.2, support this approach. High adjunct wh-phrases cannot stay in-situ in wh-in-situ-sentences nor can they occur in a sentence final position in Spanish multiple wh-phrases. Therefore I will follow Rizzi (2001) and others, assuming that *por qué* is merged in the left periphery and hence doesn’t have to move there.

If we adopt this line of argument, the question of course arises why it should not be merged in SpecFinP. Bosque & Gutiérrez-Rexach (2011: 449) states that *por qué* is also a d-linked wh-phrase. Their main argument is the same behavior in combination with high sentential adverbs. In a wh-question with *por qué*, these type of adverbs can occur preverbally but following the wh-phrase. Thus I will argue along the same line as with the complex wh-phrases. D-linked wh-questions, and in particular *por qué* are merged above the FinP. In the approach presented above, this would be inside the DS.

In summary, I assume that wh-phrases and FFed elements move to SpecFin, CLLD and complex wh-phrases to a higher DS. In that way, the structure accounts for the observed and described ordering and movement phenomena. In the next chapter, I will present the second language involved in the contact situation, Basque.
CHAPTER 4

Word order, questions, and focus in Basque

Basque is often characterized as a discourse configurational language (e.g. by Irurtzun 2007, Kiss 1995). The reason is that, although Basque normally displays a free word order, there is one position that has always to be occupied, namely the preverbal position (Ortiz de Urbina 1995: 99-100, 2003b: 448). In traditional Basque grammar this position is called galdegaia, which means “the topic of the question, what the question is about” (from galde ‘ask’ and gai ‘topic’) (Arregi 2003c: 169). The expression galdegaia was first introduced and described by Altube (1929).

In the following chapter, I will present the main characteristics of topic, interrogative and focus constructions in Basque. We will see that the latter two behave very much the same way, that means they are governed by the same underlying restrictions and rules. Therefore, I will follow Etxepare & Ortiz de Urbina (2003) and describe both phenomena together. Section 4.1 will treat the general characteristics of word order, questions, and focus formation in Basque, followed by Section 4.2 with a detailed description of the left peripheral structure of these constructions.

4.1 On Basque word order and question formation

This section will give an overview of the main characteristics in Basque focus and wh-constructions. In Section 4.1.1, I will describe the general word order phenomena and the preverbal focus position, followed in Section 4.1.2 by a short description of the main properties of polar questions, embedded questions, complex wh-questions, and multiple wh-questions. Section 4.1.3 will present pied-piping, a very common construction in Basque wh-questions.

---

39 The two phenomena differ only in one specific negative construction. I will come back to this in more detail at the end of Section 4.1.
4.1.1. Galdegaia - one position for two phenomena

The neutral word order in Basque is SOV, which classifies it as a verb-final language (Ortiz de Urbina 2003b: 448, 1995: 99-100, Uriagereka 1992: 417). ‘Neutral’ refers to an interpretation in which no specific constituent is emphasized. It is often characterized as the answer to the question ‘What happened?’. As Ortiz de Urbina puts it, “[t]he claim that SOV is the basic word order tries to capture the oft-noted intuition that while the preverbal object in this order need not to be interpreted as emphatic, any other preverbal argument in any other order […] is usually construed as focalized” (Ortiz de Urbina 2003b: 454). “Any other order” refers literally to almost every possible constituent order in a declarative sentence. Basque allows a wide variety of possible word orders, depending on which constituent should be emphasized. In the following example, all the constituents in the brackets can be permuted. Every permutation results in a grammatical sentence.

(4.1) [Ene aitak] [amari] [gona gorria] [ekarri dio].
      my father.erg mother.dat skirt red.det bring aux
   ‘My father brought mother a red skirt.’

(Ortiz de Urbina 2003b: 448)

Although all permutations are possible, they have different interpretations and informational content depending mainly on the constituent that occupies the preverbal position, which is called preverbal focus position or galdegaia. The element occupying this preverbal position is interpreted as focalized element (Etxepare & Ortiz de Urbina 2003: 460, Ortiz de Urbina 1995: 100, Uriagereka 1999: 406, among others). It is nevertheless important to notice that apart from the ‘neutral’ SOV word order, there are a few other cases in which the preverbal constituent does not have to be focalized. In Ortiz de Urbina’s words, “all focalized constituents will occur in the position immediately preceding the verb, but not all constituents immediately preceding the verb are interpreted as focalized. They will be so interpreted if they are uttered with an emphatic tone, but not otherwise” (Ortiz de Urbina 1989: 215). That means, preverbal occurrence is a necessary but not sufficient condition for a focus interpretation. An example for such a non-focalized preverbal constituent is a preverbal subject in cooccurrence with a heavy object. The latter appears normally post-verbally, which does not automatically entail that the preverbal subject is focalized. An example is given in (4.2). For a detailed analysis, see Ortiz de Urbina (2003b: 452).

(4.2) Jonek esan du Mikelek erlojua galdu duela.
     JON.erg say AUX MikeL.erg watch lose AUX.comp
   ‘Jon said that Mikel lost the watch.’

Ortiz de Urbina (2003b: 452)
Thus, it seems that in Basque the preverbal position must always be occupied by some element. A strong argument in favor of this assumption is the fact that even when a synthetic verb itself is focalized and therefore appears in a sentence initial position, it always has to be preceded by the particle ba-.

(4.3a) shows that in the case of yes/no-questions with a synthetic verb, the particle ba- has to precede the auxiliary. This holds true even if it is preceded by some topological or scrambled material as in (4.3b) and (4.3c). The particle al marks the yes/no-question in these cases. The negative particle ez ‘not’ is also capable of ‘saving’ the verb in a sentence initial position (4.3d), similar to the positiv polar marker ba.

(4.3)  
a. Ba al dugu musikan euskal eskolarik?  
\texttt{BA AL have music.in Basque school.PRTT}  
‘Do we have any Basque school of music?’

b. Euskal doinuek ba al dute berezkotasunik?  
\texttt{Basque songs.erg ba al have peculiarity.PRTT}  
‘Do Basque songs have any peculiarity?’

c. Jon ek badakar hori.  
Jon.erg ba-brings that  
‘Jon is bringing that.’

d. Ez da etxea erori.  
not AUX house fallen  
‘The house has not fallen down.’

Etxepare & Ortiz de Urbina (2003: 469/518)

wh-phrases strongly resemble the focus fronting word order pattern and restrictions. That is, wh-phrases have to appear sentence initially, directly followed by the verbal complex (Etxepare & Ortiz de Urbina 2003, Ortiz de Urbina 1989, Uriagereka 1992). No constituent can intervene between the two elements, as illustrated in the examples (4.4) and (4.5) with a focalized constituent and a wh-phrase, respectively.

(4.4)  
a. \texttt{BERA-RI azalu zion Jon ek atzo here erabakia}  
him-DAT explain aux Jon.erg yesterday his decision  
‘John explained his decision to him yesterday’

b. *\texttt{BERA-RI Jon ek azalu zion atzo here erabakia}  
him-DAT Jon.erg explain aux yesterday his decision

(4.5)  
a. Nor ikusi du Jon ek?  
who seen aux John.erg

b. *Nor Jon ek ikusi du?  
who John.erg seen aux  
‘Who has John seen?’
c. Nori azaldu zion Jonek atzo here erabakia?
   who-DAT explain AUX Jon.ERG yesterday his decision
   
   d. * Nori Jonek azaldu zion atzo here erabakia?
      who-DAT Jon.ERG explain AUX yesterday his decision
      ‘To whom did John explain his decision yesterday?’

   Etxepare & Ortiz de Urbina (2003: 459) and Ortiz de Urbina (1995: 100)

   In both constructions, other constituents may appear to the left of the wh/focus-verb unit. However, they are always interpreted as topicalized or scrambled material and are normally separated by a pause indicated with a comma (Ortiz de Urbina 2003b: 455). “[W]hile it is possible to have a sentence with a focused element but without topics (as in [771a]) [our (4.6a)], it is not possible to find a sentences with marked topics and no focus.” (Ortiz de Urbina 2003b: 456). Etxepare & Ortiz de Urbina (2003: 459-460) also state that if there is any topicalized or scrambled material, some element has to be focalized. In (4.6c), for example, either the subject Jonek or the verb itself must bear a focus interpretation, since niri ‘to me’ is topicalized. Additionally Etxepare (1997) says that scrambled material forces a contrastive focus reading. (4.7a) shows that a foci preceded by scrambled material “only has a contrastive reading, as the infeliciteness of the conditional shows” (Etxepare 1997: 117).

   (4.6) a. NEUK esango diot.
      I.ERG say.FUT AUX
      ‘I will tell him.’

   b. Nik, ordea, esango diot.
      I.ERG on.the.other.hand say.FUT AUX
      ‘I, on the other hand, will tell him.’

   Ortiz de Urbina (2003b: 456)

   c. Niri, Jonek azaldu zidan.
      I-DAT Jon.ERG explain AUX
      ‘Jon explained that to me.’

   Etxepare & Ortiz de Urbina (2003: 460)

   (4.7) a. Miren ardoa Jonek oparitu dio, “inork oparitu badio behintzat
      Miren.DAT wine.ABS John.ERG offer AUX anybody offer if.AUX at least
      ‘It is John who offered wine to Mary, if anybody really did.’

   Etxepare (1997: 117)

40 Contrary to Spanish topics, Basque topics are not doubled by a clitic, because Basque has no clitics. It is a three-way pro-dop language, meaning that the subject, the direct and the indirect object are encoded in the finite verb form and can therefore be dropped. Furthermore, topics preceding the focus-verb unit are normally separated by a pause from the rest of the sentence. In the case of various topics, it would be a pause after each (Ortiz de Urbina 2003b: 455).
Although the *wh*/focus-verb adjacency is nearly always obligatory, there are a few constructions where it is not. We can find the violation of the obligatory adjacency with some particles, with certain *wh*-words and in negative sentences. In the first case, the intervening elements are a small set of particles like *ote* ‘perhaps’ (4.8a) and *omen* ‘apparently’ (4.8b) or occasionally parenthetical elements (4.8c) (Etxepare & Ortiz de Urbina 2003: 464-465, Ortiz de Urbina 1989: 213).

(4.8) a. Zergatik ote dago hainbeste tximeleta?
   why Q is so many butterfly
   ‘Why are there so many butterflies?’

b. Jonek omen daki hori
   Jon.erg apparently knows that
   ‘Apparently Jon knows that.’

c. Zein idazle, gaurko edo denbora bateko, iruditzen zaizu idaririk
   which writer today.rel or time other.rel seem.impf aux guide.prtt
   zuzenena hitz kontuan?
   correct.most.det word matter.in
   ‘Which writer, from today or from other times, seems to you the best guide in terms of words?’

Etxepare & Ortiz de Urbina (2003: 464-465)

The second case are interrogative clauses that contain the adverbial *wh*-phrases *zergatik* ‘why’ and *nolatan* ‘how/ how come’. As shown in Etxepare & Ortiz de Urbina (2003: 465) and Uriagereka (1999: 412), a sentence like (4.9a) is perfectly grammatical. Uriagereka (1999: 438) cites another example that illustrates the special behavior of *zergatik*. In (4.9b), *zergatik* ‘why’ cooccurs directly with a fronted focus in an indirect clause. This is a further case that exemplifies the general crosslinguistic specialness of ‘why’.

(4.9) a. Zergatik zaldunak herensugea hil zuen?
   why knight.the.erg dragon.the.abs killed aux
   ‘Why has the knight killed the dragon?’

b. Ez dakit zergatik honi eman behar diozun.
   not know.1sg why this.dat give must aux
   ‘I don’t know why to this one you must give it.’

Uriagereka (1999: 412/438)

---

41 The same set of particles can be “attached immediately to the left of the tensed verbal form (whether auxiliary or synthetic verb).” (Ortiz de Urbina 2003a: 316). Nothing else can intervene between the tensed verbal form and the particle.

42 This exception matches the observations made for Spanish described in Section 3.2.1 for the Spanish counterpart *por qué* ‘why’.
The last case are negative questions. They exhibit a different word order than their positiv counterparts. The negative particle ez ‘not’ has to directly precede the tensed verbal form. Together, they immediately follow the wh-word. Interestingly, the participle can, but does not have to, follow the former unit. It can be separated by any number of constituents and even appear sentence final (Etxepare & Ortiz de Urbina 2003: 503). Examples (4.10a) and (4.10b) illustrate both cases.

(4.10) a. Nork ez du ulertu esan dudana?
   who.erg not aux understand say aux.comp.det
   ‘Who did not understand what I said?’

b. Nork ez du esan dudana ulertu?
   who.erg not aux say aux.comp.det understand
   ‘Who did not understand what I said?’

adapted from Etxepare & Ortiz de Urbina (2003: 503)

In case of negative focalization, two positions are generally possible, marking the only major difference between focus- and wh-movement in Basque. The first possibility matches the just mentioned pattern of the wh-phrase preceding the unit of ez+tensed verbal form, like in (4.11a). However, foci can appear in a second position on the right of the negation+tensed verbal form, directly preceding the participle, as illustrated in (4.11b). Nevertheless, the focalized constituent is only acceptable in this latter position if it “appears in the position it would occupy in a ‘neutral’ word order” (Etxepare & Ortiz de Urbina 2003: 508).

    Jon.erg not aux petition sign
    ‘Jon did not sign the petition.’

b. Ez du Jonek eskaera sinatu.
   not aux Jon.erg petition sign
   ‘Jon did not sign the petition.’

Etxepare & Ortiz de Urbina (2003: 505,507)

This last variation in the word order is not very surprising having in mind that the negative particle ez has other core properties like dissolving the V1 restriction, among others. I will come back to negative sentences and their structure in Section 4.2.2. In Chapter 5, I will venture an explanation of this difference between wh- and focus-movement in this specific construction. Furthermore, for an extended description and discussion of negative questions and negative foci, the reader is referred to Etxepare & Ortiz de Urbina (2003), Haddican (2004), and references quoted therein.

To recap, non-adjacency between the wh-phrase/the focalized constituent and the verb generally produces an ungrammatical sentence.
4.1.2. More on questions and foci in Basque

Although my main focus will be on direct wh-interrogatives and focus constructions in Basque, I want to briefly summarize the main properties of polar questions in Section 4.1.2.1 and embedded questions and foci in Section 4.1.2.2. Both constructions will be important for the argumentation in Section 5.1.2. Section 4.1.2.3 and 4.1.2.4 will address complex wh-questions and multiple wh-questions, respectively. Complex wh-questions will be dealt with in more detail in Section 4.1.3, whereas multiple wh-questions will return in Chapter 5.

4.1.2.1. Polar questions

Basque polar questions “need not be signaled by any mark other than interrogative intonation” (Etxepare & Ortiz de Urbina 2003: 467), exemplified in (4.12a). However, they can be marked by overt morphology, namely by the verbal particle *al* (4.12b) or the suffix *-a* (4.12c). The former appears preceding the finit verb form and can occur in root as well as embedded sentences, whereas the latter occupies the same position as a complementizer and is restricted to direct questions.

(4.12)  

a. Jonek liburu hori irakurri du? 
   Jon.ERG book.ABS that read AUX 
   ‘Has Jon read that book’

b. Gure literaturak aurrerakada haundirik egin al du urte hauetan 
   our literature.ERG improvement great.FRTT make al AUX year these.in 
   ‘Has our literature made any great improvements in the last years?’

c. Oro egin eta desegin, ez dea lur huntako legea? 
   all do and undo not is.a earth this.of law 
   ‘To do and undo everthing: isn’t that the law of the world?’

Etxepare & Ortiz de Urbina (2003: 467)

It is also possible to mark them syntactically, particularly by verb fronting, as shown in (4.13a). Verb fronting can be combined with the particle *al* as well as the suffix *-a*, as the examples from Etxepare & Ortiz de Urbina (2003: 468) illustrate in (4.13b) and (4.13c), respectively.

(4.13)  

a. Erosi du Jonek liburu? 
   bought AUX John book 
   ‘Has John bought the book?’

Ortiz de Urbina (1994: 128)

43 see Ortiz de Urbina (1999: 18iff) for a different point of view.
b. Esango al zeniguke zerbait azkenik
   say.fut al aux something finally
   ‘Would you tell us something finally?’

c. Entzun duzia gure auzo Biarnesek berek nola goresten duten
   hear aux.a our neighbor Bearinese.erg themselves.erg how praise.impf aux.comp
   beren hizkuntza?
   their language
   ‘Have you heard how our neighbors from Béarn themselves praise their language?’

Etxepare & Ortiz de Urbina (2003: 469)

Similar to *wh*-phrases, the fronted verb in polar questions can be “preceded by any number of
topicalized/scrambled elements and a focalized constituent” (Etxepare & Ortiz de Urbina 2003: 468). However, if the fronted verb is synthetic, it hast to be preceded by the particle *ba*, shown in (4.3), here repeated as (4.14a). The same holds for the case with topicalized/scrambled elements preceding the verb (4.14b). If the preceding element is focalized, however, no such particle is necessary (4.14c) (Etxepare & Ortiz de Urbina 2003: 469).

(4.14) a. Ba al dugu musikan euskal eskolarik?
   *ba al* have music in Basque school.prtt
   ‘Do we have any Basque school of music?’

b. Euskal doinuk ba al dute berezkotasunik?
   Basque songs.erg *ba al* have peculiarity.prtt
   ‘Do Basque songs have any peculiarity?’

c. Kaltegarritzat al daukazue joera hori?
   harmful.as *al* have tendency that
   ‘Do you consider that tendency HARMFUL?’

Etxepare & Ortiz de Urbina (2003: 469)

In summary, polar questions can be marked exclusively by interrogative intonation or by syntactic
marking via verb fronting and/or the use of certain particles. I will now turn briefly to embedded
questions.

4.1.2.2. Embedded questions/foci

The obligatory *wh*/focus-verb adjacency also applies to embedded sentences and indirect questions,
resembling their Spanish counterparts (Etxepare & Ortiz de Urbina 2003: 474-484). The verb is ad-
ditionally marked with the complementizer *-n* and the *wh*-phrase can optionally be preceded by *ea,*
illustrated in (4.15a) and (4.15b)-(4.15c), respectively. The *wh*-phrase can be preceded by topicalized or
scrambled material.
4.1 On Basque word order and question formation

(4.15)  

a. Nor etorriko d-en bihar galdetu duzu.  
who come AUX-en tomorrow ask AUX  
‘You have asked who will come tomorrow.’  
Ortiz de Urbina (1992: 297)

b. Galdetu du ea nola egingo duda-n.  
ask AUX ea how do.FUT AUX-n  
‘He has asked how I will do it.’

c. Ea nork agindu du-en hori jakin nahi nuke.  
ea who.ERG order AUX-en that know want AUX  
‘I would like to know who has ordered that.’  
Ortiz de Urbina (1999: 180)

Embedded foci can be found in tensed complement clauses and intransitive subject clauses. The foci have to be positioned directly to the left of the verbal complex, illustrated in (4.16) (Etxepare & Ortiz de Urbina 2003: 479, Ortiz de Urbina 1995: 115).

(4.16)  

a. Eta okerrena zen KARTAZALEARI botatzen zizkiola erru gustiak.  
and worst.DET was postman.DAT throw.IMPF AUX.that blame all.DET.PL  
‘And the worst thing was that she blamed the postman.’

b. esan didate Jon ikusi duela Mikelek  
tell AUX John see AUX.ia Mikel  
‘They told me that Mikel saw Jon’

Etxepare & Ortiz de Urbina (2003: 479) and Ortiz de Urbina (1995: 115)

It is possible to extract wh-words and foci from complement clauses of some verbs, especially ‘saying’ and ‘thinking’. The matrix verb has to immediately follow the wh-phrase/foci and the embedded clause from which is is extracted tends to be verb-initial. Etxepare & Ortiz de Urbina (2003: 485) state that “this effect is even perceivable where the displaced element occurs several sentences higher than the clause it enquires about or focalizes” (4.17c). (4.17a) and (4.17b) show an extracted wh-word and an extracted foci, respectively.

(4.17)  

a. Nor nahi duzue (zuek) etor dadin?  
who want AUX you.PL come AUX.COMP  
‘Who do you want to come’

b. XABIER nahi dugu (guk) etor dadin.  
Xabier want AUX we come AUX.COMP  
‘It is Xabier that we want to come.’

c. Nola esan du Jon este duela Peruk egin beharko litzazekeela?  
how say AUX Jon.ERG think AUX.COMP Peru.ERG make must.FUT AUX.COMP  
‘How did Jon say Peru thinks it should be made?’
3.2.1 that Spanish *wh*-questions are subject to the very same. It is not possible to extract a *wh*-phrase or foci out of a transitive subject clause, an adjunct clause, a complex NP, or a coordinated structure as seen in (4.18a) - (4.18d). The extraction out of an intransitive subject clause on the other hand is possible (4.18e).

(4.18) a. *Non/unibertsitate horretan estan du asko zutaz [onartua izateak]? where/university this.in sayimpe aux.a.lot you.about admitted being
   ‘Where/In that university does it say a lot about you being admitted?’

b. *Zer harritu zinen [entzun ondoren/entzun zenuenean]? what surprise aux. hear after/hear aux.comp.when
   ‘What where you surprised after hearing/when you heard?’

c. *Noren/Jonen izan ditu [etxeko teiltatuak] itoginak? whose/Jon.gen have aux house.rel roof.erg leaks.abs
   ‘The roof of Jon’s/Whose house has had leaks?’

d. *Zer/Bakailoa jango du lehengusinak [ eta txuleta]? What/Cod eat.fut aux cousin and cutlet
   ‘What/Cod will the cousin eat and cutlet’

e. Non/unibertsitate horretan iruditzen zaie [onartu dutela Jon]? where/university this.in seem.impf aux admit aux.comp Jon
   ‘Where/In that university does it seem to them they admitted Jon?’

Etxepare & Ortiz de Urbina (2003: 474/485/486) and Irurtzun (2007: 158)

In the case of *wh*-words embedded in complex NPs, relative clauses, and some adjunct clauses, Basque has an alternative strategy for fronting them, namely pied-piping. I will come back to this in Section 4.1.3.

4.1.2.3. Complex *wh*-phrases

In Section 3.2.1 I have shown that Spanish exhibits clear exceptions regarding the obligatory *wh*-phrase-verb adjacency with heavy or complex *wh*-phrases. That is to say, subjects can intervene between those *wh*-phrases and the verb.

In comparable cases in Basque, the adjacency cannot be violated by any constituent, no matter how complex the *wh*-phrase is.\(^{44}\) If the *wh*-phrase is embedded within another constituent, the whole constituent has to appear pre-verbally. This happens for example with question words inside noun

\(^{44}\) Except for the former mentioned cases of very few particles.
phrases. The examples in (4.19) show the obligatory adjacency of the complex wh-phrase with the verb.

\[(4.19)\]  
\begin{align*}  
a. \text{ Zein herritan bizi zen Jon lehenago?} & \quad \text{which town in live before} \\
\text{'In which town did Jon live before?'} & \\
\end{align*}

\[a'. \quad \text{Zein herritan Jon bizi zen lehenago?} \]

Ortiz de Urbina (1989: 213/14)

\[(4.20)\]  
\begin{align*}  
a. \text{ Neska horietako zein (zure arrebak) bisitatu zuen (zure arrebak) lehenago?} & \quad \text{Girl these of (your sister) visit before (your sister)?} \\
\text{'Which of the girls has visited your sister before?'} & \\
\end{align*}

\[b. \quad \text{ Bere lagunetako nori (Juanek) eramango dio (Juanek) whisky bottle bat his friends of who dat John erg bring fut aux 3sg John erg whisky bottle a (Juanek) festara (Juanek) gaur gauean (Juanek)? John erg party to John erg night today tonight} \]

\text{'Which of his friends will John bring a bottle of whisky to at the party tonight?'}

As indicated, the subject cannot appear between the complex wh-phrase bere lagunetako nori and the verbal complex eramango dio. In the next section, I will analyze multiple wh-questions and multiple foci.

4.1.2.4. Multiple wh-phrases/foci

In Basque, it is possible to have multiple wh-words or foci in the preverbal position. Stacked wh-words appear normally with a coordination marker like eta 'and', shown in (4.21).

\[(4.21)\]  
\begin{align*}  
\text{Non, norekin eta zertako eginen dugu gerla?} & \quad \text{where who with and why makre fut aux war} \\
\text{'Where, with whom and why will we make war?'} & \\
\end{align*}

Etxepare & Ortiz de Urbina (2003: 462/499)

Occasionally, wh-words can also be stacked without any marker. Etxepare & Ortiz de Urbina (2003: 461-462) say that in this case, the wh-phrases need to be in the neutral constituent order, that is, the order the constituents would be in a declarative sentence. In a sentence like (4.22a), therefore, the word order zer nork 'what who' would be ungrammatical. However, R. Etxepare (personal communication) says that this restriction is not valid for all wh-phrases. As we can see in (4.23), the simultaneous
fronting of nori “who.dat” and nork ‘who.erg’ is not subject to such restriction. It always holds that no non-wh constituent may occur between the stacked elements, as shown in (4.22b).

(4.22) a. Eta zure prestutasunaz... nork zer erranen du? and your virtue.on who say.fut aux ‘And who will say what on your virtue?’

b. * Nork horretaz/beraz/Joni zer erranen dio? Who.erg that.on/therefore/Jon.dat what say.fut aux ‘Who will say what to Jon/therefore/on that?’


Etxepare & Ortiz de Urbina (2003: 462/499)

We will see in Chapter 5 that zer exhibits further differences compared to other wh-words. I will come back to multiple stacked wh-words in that chapter as well.

Interestingly, the restriction concerning the linear order of the wh-phrases in multiple wh-constructions dissolves, if only one wh-phrase moves to the front and the other(s) stay in-situ. As a result, both sentences in (4.24) are acceptable, unlike the English counterpart of the second question.


b. Zer ekarri du nork? what bring aux who.erg ‘*What did who bring?’

(Etxepare & Ortiz de Urbina 2003: 499)

The restrictions on multiple foci are stricter. Generally, multiple fronted foci are not possible, as seen in (4.25a) and (4.25b) (Etxepare & Ortiz de Urbina 2003: 497). However, more than one element can be focalized “as a unit in preverbal position, in highly contrastive context.” (Etxepare & Ortiz de Urbina 2003: 498). The interpretation is similar to multiple wh-questions, but no coordination marker is needed. Nevertheless, the neutral word order has to be maintained, as seen in (4.26). With various stacked foci, no non-focalized constituent may occur between the single focalized elements, as seen in (4.26).
4.1 On Basque word order and question formation

(4.25)  

\[ \begin{align} 
\text{b. } & \neuk \text{ ardoa ekarri dut, ez Mikelek liburua.} \\
& \text{IINST.ERG wine.ABS bring AUX not Mikel.ERG book.ABS} \\
& \text{‘It was me that brought the wine, not Mikel the book.’} 
\end{align} \]

(4.26)  

\[ \begin{align} 
\text{b. } & \text{ Jonek badakar liburua.} \\
& \text{Jon.ERG brings book.} \\
& \text{‘It is Jon that is bringing the book.’} 
\end{align} \]

In summary, Basque offers different possibilities in order to express multiple wh-phrases. The important observation is that at least one wh-phrase has to be fronted in order to occupy the preverbal focus position. I will come back to these properties of multiple wh-fronting in Chapter 5.

In the next section, I will present another important feature of Basque wh-constructions, namely pied-piping.

4.1.3. Pied-Piping

Basque exhibits a further, widespread occurrence of wh/focus-fronting, namely pied-piping structures (Arregi 2003a, Etxepare & Ortiz de Urbina 2003, Ortiz de Urbina 1989). As already mentioned in Section 3.2.2.3, the term pied-piping refers to the ability of certain elements, like wh-words, focalized elements or quantifiers, to ‘drag along’ the rest of the constituent they are part of. Cable writes that “the term ‘pied-piping’ describes cases where an operation targeting the feature of a particular lexical item applies to a phrase properly containing the maximal projection of that item” (Cable 2012: 6). If a wh-phrase or focalized element originates inside an embedded clause, the entire clause moves to the front of the sentence into the preverbal focus position. The wh-phrase/foci now precedes the verbal complex inside the embedded clause and additionally ”the embedded clause itself will also appear immediately preceding the matrix verb” (Etxepare & Ortiz de Urbina 2003: 486). All kinds of complement clauses can be pied-piped. The following examples show pied-piping of phrases containing a wh-word ((4.27a)-(4.27b)) and a focalized constituent (4.27c). We can always see the obligatory adjacency between the pied-piped clause and the matrix verb.
(4.27) a. [Sein il banela Jonek] pentzaten dau Mirenek?
who.abs killed aux.that Jon.erg think aux Miren.erg
‘Who does Miren think Jon killed?’

Arregi (2003a: 126)

b. [Nor etor dadin] (*zuek) nahi duzue (zuek)?
who.abs come aux-n you.pl want aux you.pl
‘Who do you want to come?’

c. [Xabier etor dadin] (*guk) nahi dugu (guk).
Xabier come aux-n we want aux we
‘It is Xabier that we want to come.’

Etxepare & Ortiz de Urbina (2003: 487/488)

For some clauses, it is the only possibility to be fronted at all. As I have shown in Section 4.1.2.2, wh-phrases cannot be extracted out of relative clauses, complex NPs or adjunct clauses. However, all three types can be pied-piped, as long as the whole clause is fronted. Therefore, the a.-versions of the examples (4.28), (4.29), and (4.30) with pied-piping are fully grammatical, contrary to the b.-versions with extraction (Artiagoitia et al. 2003: 786, Ortiz de Urbina 1989: 249).

(4.28) Complex NP:

a. [Noren/Jonen etxeko teiltuak] izan ditu itoginak?
whose/Jon.gen house.rel roof.erg have aux leaks.abs
‘The roof of Jon’s/Whose house roof has had leaks?’

b. * Noren/Jonen izan ditu [etxeko teiltuak] itoginak?
whose/Jon.gen have aux house.rel roof.erg leaks.abs

Etxepare & Ortiz de Urbina (2003: 474)

(4.29) Relative clause:

a. [[Nork idatzi zuen] liburua] irakurri du Pellok?
who.erg wrote aux.comp book.abs read aux Peter.erg
‘The book that who wrote has Peter read?’

who read aux Peter.erg wrote aux.comp book.abs

Ortiz de Urbina (1989: 249)

(4.30) Adjunct clause:

a. [Zenbat gai suspensitzen denean] pasa daiteke hurrengo urtera?
how many subject fail.impf aux.when pass aux next course.to
‘When one fails how many subjects can one pass to the following course?’

b. * Zenbat pasa daiteke hurrengo urtera gai suspensitzen denean?
how many pass aux next course.to subject fail.impf aux.when

Etxepare & Ortiz de Urbina (2003: 490)
Another important characteristic of pied-piping structures in Basque is recursiveness. A clause can be pied-piped not only to the higher sentence immediately to the left, but even higher up. The verb-initial pattern occurs subsequently in all the ‘passed’ clauses. The pied-piped clause can even pied-pipe the intermediate clauses as well (Etxepare & Ortiz de Urbina 2003: 493). Both cases are shown in (4.31a) and (4.31b), respectively. It is further possible to pied-pipe adjuncts inside adjuncts, as shown in (4.31c)

(4.31)

a. Nork irabaziko duela esan du Jonek uste duela Mikelek?
   who.erg win.fut  auc.comp say  aux Jon.erg think aux.comp Mike.erg
   “That who will win did Jon say Mikel thinks?”

b. [Nor etorri dela uste duela Mikelek] esan dute horriek?
   who come auc.comp think aux.comp Mike.erg say aux those.erg
   “Who did those say Mikel thinks has come?”

c. [(Nor agertu denean) aldegia dutelako] hasseratu da Mikel?
   who show.up aux.when leave aux.because annoy aux Mike.erg
   “Because they left when who showed up did Mikel get annoyed?”

Etxepare & Ortiz de Urbina (2003: 493)

In summary, we have seen that pied-piping structures are widely spread but subject to some restrictions. If the pied-piped clause contains a wh-phrase or a foci, it must remain adjacent to the matrix verb. In addition, the wh-/focus-phrase has to be also clause-internally left-adjacent to the verbal complex. Furthermore it seems that in the cases where pied-piping as well as extraction are available (cf. (4.17a) vs. (4.27b) or (4.17b) vs. (4.27c)), Basque offers true optionality (Cable 2010: 167, Etxepare & Ortiz de Urbina 2003: 487). I will come back to pied-piping and its implications for the Q-particle theory by S. Cable in Chapter 5.

The next section will present the main approaches to Basque focal and wh-movement. Ortiz de Urbina’s I-to-C-movement and its further development by A. Irurtzun to a full cartographical approach, and K. Arregi’s Nuclear-Stress-Rule, which approaches the matter from a very different angle.

4.2 On syntactic structures in Basque wh-questions

We have seen in the previous chapter that the preverbal position is the main landing site for focalized elements and wh-phrases in Basque. In order to explain these facts, two main approaches have been developed. The first builds on the cartographic approach by L. Rizzi, presented already in Section 3.3. It is mainly described by Irurtzun (2007) and Ortiz de Urbina (1989, 1992, 1999). The second approach
is based on the so-called _Nuclear-Stress-Rule_ (NSR)\(^{45}\) introduced by Cinque and developed for Basque mainly by Arregi (2003c) and Elordieta (2001). The core difference between the two approaches is the following: in the former it is the focalized element that moves, namely to a focus position in the left periphery of the sentence. According to the NSR on the other hand, if the focalized constituent is not the most deeply embedded element, everything else has to move higher up in the hierarchy in order to ensure that it is.

Section 4.2.1 first gives a short introduction to both approaches followed by a brief discussion of the advantages of the cartographic approach. Section 4.2.2 will describe in more detail the left periphery in Basque sentences with _wh_- or focus-movement.

### 4.2.1. NSR-based and Cartographic approach

#### 4.2.1.1. Nuclear-Stress-Rule

The NSR-based approach stems mainly from Cinque’s (1993) seminal work on nuclear stress. He assumes that the most deeply embedded element in an utterance bears the nuclear stress. He formulates a focus-to-nuclear stress correlation saying: interpret focus where nuclear stress is. That means the most embedded element is focalized or, more precisely, the most embedded element has to be part of the focalized constituent, and can project its focal status to a higher node. Reinhart (2006) expresses it in her Focus Rule as followed: “The focus of IP is a(ny) constituent containing the main stress of IP, as determined by the stress-rule”. That is, “According to the NSR based theory of F-Structure, a sentence will not have an actual focus but ‘a set of possible foci’, the set of nodes that contain the nuclear stress” (Irurtzun 2007: 142). In (4.32), we can see that the most deeply embedded element _Judea_ wears the nuclear stress. Thus, it can project its focal stress to the DP (4.32b), to the whole VP (4.32c), or even to the entire sentence (4.32d).

(4.32)  
\begin{enumerate}
  \item Jesus preached to the people of \([\text{Judea}]_F\)
  \item Jesus preached to the [the people of \text{Judea}]_F.
  \item Jesus [preached to the people of \text{Judea}]_F
  \item [Jesus preached to the people of \text{Judea}]_F
\end{enumerate}

Irurtzun (2007: 141)

---

\(^{45}\) I will follow Irurtzun (2007: 140) and call this approach the NSR-based approach.
The main idea in the NSR-based approach is that if a non-focalized element is in the most deeply embedded position, a ‘nuclear stress avoiding’ movement must apply in order to ensure that the focalized constituent is the most deeply embedded instead. This ‘nuclear stress avoiding’ movement is called scrambling by Elordieta (2001) or left and right dislocations by Arregi (2003b). Arregi assumes an Aspect Phrase (AspP) between TP and vP. In the following sentences, the direct object is left or right dislocated, ensuring that in both cases the focalized subject is the most deeply embedded element. The derivations for (4.33a) and (4.33b) are given in (4.34a.) and (4.34b.), respectively.

(4.33)  
\[ \text{a. Mahaia Jonek hautsi du.} \]
\[ \text{table Jonek.erg broken aux} \]
\[ \text{‘Jon broke the table.’} \]
\[ \text{b. Jonek hautsi du mahaia.} \]
\[ \text{Jon.erg broken aux table} \]
\[ \text{‘Jon broke the table.’} \]

(4.34)  
\[ \text{a. left dislocation} \]
\[ \text{TP} \]
\[ \text{mahaia} \]
\[ \text{TP} \]
\[ \text{AspP} \]
\[ \text{T} \]
\[ \text{vP} \]
\[ \text{Asp} \]
\[ \text{hautsi du} \]
\[ \text{Jonek} \]
\[ \text{v} \]
\[ \text{VP} \]
\[ \text{t} \]
\[ \text{t} \]

\[ \text{b. right dislocation} \]
\[ \text{TP} \]
\[ \text{mahaia} \]
\[ \text{TP} \]
\[ \text{AspP} \]
\[ \text{T} \]
\[ \text{vP} \]
\[ \text{Asp} \]
\[ \text{hautsi du} \]
\[ \text{Jonek} \]
\[ \text{v} \]
\[ \text{VP} \]
\[ \text{t} \]
\[ \text{t} \]

Irurtzun (2007: 144)

The interesting result of such an approach is the fact that the focus-verb adjacency is purely coincidental and not at all a crucial structural feature in focus constructions. The adjacency is not mentioned explicitly in the NSR-based approach (Irurtzun 2007: 145). In summation, the three basic features of the NSR-based approach according to Irurtzun (2007: 145) are as follows:

46 Sometimes also called p(rosodic)-movement, (cf. among others, Zubizarreta 1998).
4.2 On syntactic structures in Basque wh-questions

(4.35) i. The focal element stays *in-situ*. Other material moves (if required by PF).
   ii. Movements of non-focal material can take place leftwards or rightwards.
   iii. Focus-Verb adjacency is accidental; a byproduct of the fact that the focal XP is the most
        embedded and the verb is the next element to its right.

Having seen the core properties of the NSR-based approach, we can now contrast the cartographic
approach for Basque focus and wh-constructions.

4.2.1.2. Cartographic approach

In the seminal work of Ortiz de Urbina (1989), FFed constituents and fronted wh-phrases in Basque
were analyzed to move to a left-peripheral CP (cf. among others Ormazabal, Uriagereka & Uribe-
split of the left periphery, presented already in Section 3.3.2 in (3.73), the landing site for focalized or
wh-elements has been FocP (cf. Irurtzun 2007, Ortiz de Urbina 1999). In this approach, Foc0 attracts
the focalized phrase to its specifier and the verb rises to the same head in order to satisfy a ‘focus-
criterion’. Ortiz de Urbina (1995) extends Rizzi’s (1996) ‘wh-criterion’ (3.74) to a more general ‘affective-
operator-criterion’ for Basque, which also includes focalized elements. In this way, he captures
the nearly identical behavior of wh-phrases and focalized elements. Therefore, the structure for (4.33b),
here repeated as (4.36a), would be derived as in (4.36c). Assuming a similar derivation for wh-questions,
the structure in (4.36c) can be used to present (4.36b)

(4.36) a. JONEK hautsi du mahaia.
    Jon.erg broken aux table
    ‘Jon broke the table.’

b. Nork hautsi du mahaia?
   who.erg broken aux table
   ‘Who has broken the table?’
This cartographic analysis presents a viable solution for the word order observed and described in the former sections. Either the focalized constituent or the wh-phrase move to the Specifier of FocP, attracted by a focus- or wh-feature, and the verb moves up to Foc\(^0\) to satisfy a ‘affective-operator-criterion’.

We have seen that the core difference between these two approaches is what moves in the derivation. The two analysis give very different answers to this question. In the cartographic approach, movement is motivated by a focus- or wh-feature feature which enters an agreement relation with a phrase in the left periphery, namely FocP. For the case of focalized as well as wh-elements, there are strong arguments in favor of movement, as for example island restrictions or mutual exclusiveness (Irurtzun 2007: 146). In the NSR-based approach, everything but the focalized element moves. That is, we have a lot of movement not motivated by any features or syntactical structures, but instead by avoiding the nuclear stress on the wrong element.

In the following section, I will present two empirical reasons that the cartographic approach is more suited to explain focus and wh-movement in Basque. The discussion is based mainly on Irurtzun (2007: 4.1.2.).

4.2.1.3. Cartographic vs. NSR-based approach

Irurtzun (2007: 4.1.2.) discusses several empirical reasons in favor of the cartographical approach in detail. He analyses movement patterns, long distance movement and pied-piping, islandhood, dialectal variation, focus in infinitival constructions, rightward focus, and fragmental answers. I will not repeat his detailed and extensive discussion in its full length. Instead, I will focalize in the following most
important contributions of his critical analysis: special movement patterns, the long distance movement with a close look at pied-piping constructions, and the islandhood.

The NSR-based approach would account for (4.37) by assuming that *Jon* has to be dislocated to the right. But there is no internal reason to motivate this movement. Since it is not the most deeply embedded element, no movement should be possible or necessary. Nevertheless, movement must take place. The cartographical approach on the other hand accounts directly for the observed word order. The direct object *mahaia* ‘table’ moves to SpecFocP and the verb follows to Foc\(^0\) (Irurtzun 2007: 148-150).

\[
\text{(4.37)} \quad \text{Mahaia hautsi du Jonk.}
\]
\[
\text{table broken aux Jon.erg}
\]
\[
\text{‘Jon broke the table.’}
\]

Irurtzun (2007: 147)

Further arguments in favor of the cartographical approach are long distance focus- or *wh*- movements and pied-piping constructions. Given a sentence like (4.38a), the derivation could be as follows: the subject *Jon* moves first to the CP-Shell of the embedded clause and then further on to the CP-layer of the matrix clause. This movement is also called successive cyclic. The derivation is represented in (4.38b).
4.2 On syntactic structures in Basque wh-questions

(4.38) a. Jon/Nor pentsatzen dut [CP ikusi zuela Mirenek.]
Jon/Who think aux seen aux-that Miren.erg
‘I think Miren saw Jon./Who do I think Miren saw?’

b.

In a sentence like (4.39a), the focalized element first moves to the embedded CP-layer, followed by the pied-piping of the complete embedded clause to the CP₁ of the matrix clause. This derivation is represented in (4.39b)
The cartographic approach directly accounts for the observed word order and pied-piping constructions. Arregi’s (2003) NSR-based approach relies on many more assumptions. In a sentence like (4.40a), similar to (4.38a) and (4.39a), Arregi assumes first a movement of the focalized element to the matrix’ vP-left-adjoint position (4.40b). In a second step, illustrated in (4.40c), the embedded clause moves to a right dislocated position adjoint to the TP, parallel to the simple sentences we have seen in Section 4.2.1.1. This way, eventually, the focalized object Jon is the most deeply embedded element in the derivation.
4.2 On syntactic structures in Basque wh-questions

(4.40) a. Jon pentsatzen dut [CP Mirenek t ikusi zuela]
Jon think AUX Miren.ERG seen AUX
‘I think that Miren saw Jon’

b. 

\[
\begin{array}{c}
\text{TP}_1 \\
\text{AspP} \\
\text{vP}_1 \\
\text{[pentsatzen dut]}_v \\
\text{Jon} \\
\text{vP}_1 \\
\text{pro} \\
\text{vP}_1 \\
\text{VP}_1 \\
\text{t}_v \\
\text{[Mirenek t\textsubscript{DP} ikusi zuela]}_{\text{CP}} \\
\text{t}_v \\
\end{array}
\]

adapted from Irurtzun (2007: 152)
4.2 On syntactic structures in Basque wh-questions

Irurtzun (2007: 153) points out several difficulties with this approach. The first concerns the object of the embedded clause. It is originally already merged in the most deeply embedded position. The movement to the vP-adjoint position is not motivated at all by the theory. The same problem holds for the other movements. Secondly, the grammaticality of such a sentence is not at all clear. Irurtzun (2007) would mark such constructions at the most with a ‘⁇’, Ortiz de Urbina (1989, 1999), and others, with a ‘*’. In any case, most speakers would prefer the word order in the matrix sentence as in (4.38a) and (4.39a) and not as in (4.40a). That is, they would prefer the order [verbal complex + subject] over [subject + verbal complex] (Etxepare & Ortiz de Urbina 2003: 458, Irurtzun 2007: 153).

A final argument against the NSR-based approach can be demonstrated, when we replace the silent pro in (4.40a) with an overt subject as in (4.41a). The NSR-based approach would posit another right dislocation, namely of the subject of the matrix clause to an TP-adjoint position. The movement is illustrated in (4.41b)

(4.41) a. Jon pentsatzen dut Aitorrek [CP Mirenek t ikusi zuela]
Jon think aux Aito.erg Miren.erg seen aux
‘Aitor thinks that Miren saw Jon’

b. adapted from Irurtzun (2007: 154)
In order to achieve the right word order, the ‘NSR-based theory of F-structure will have to impose an arbitrary order in the rightward dislocations so that the dislocation of the subject applies before the embedded clause adjoins to TP’ (Irurtzun 2007: 154). If this does not take place, the constituent order would be as in (4.42), which is ungrammatical.

(4.42) *Jon pentsaten dut [CP Miren[ek t i k usi zuela] Aitorrek
Jon think AUX Miren.ERG seen AUX Aitor.ERG
‘Aitor thinks that Miren saw Jon’

Irurtzun (2007: 155)

The same problems arise with the pied-piping constructions as seen in (4.39a). The object in the embedded clause has to move in order to leave the subject in the most embedded position. The subject of the matrix clause has to also move to a right dislocated position in order to establish the needed word order. Again, this latter movement is totally unmotivated within the theory (Irurtzun 2007: 157)

We have seen in Section 4.1.2.2 that Basque wh- and focus movement, as in Spanish, is subjected to island restrictions. These phrases cannot be extracted from adjunct or relative clauses, out of complex NPs, or coordinated structures (among others, Etxepare & Ortiz de Urbina 2003). In the cartographic approach, these facts find a compelling explanation. Wh-phrases and foci simply cannot pass out of the embedded structures, no matter what. Movement out of these islands is not possible for any constituent. The cartographic approach, therefore, offers an explanation for both wh- and focus movement. In the NSR-based approach, on the other hand, the focalized element does not move at all, “unless the violation is put in terms of extraction of non-focused elements due to p-movements” (Irurtzun 2007: 159). So the question remains how the violation comes along at all. In addition, the NSR-based approach does not account for wh-movement in terms of nuclear stress. That means, a different reason would have to be found in order to explain this. It is counterintuitive and conceptually not desirable, since both movements underly very similar restrictions (Irurtzun 2007: 160, Etxepare & Ortiz de Urbina 2003: 464). That is, “a unitary account of islands for focus and wh-phrases should be preferred instead of different explanations for data that vary just in the interrogative vs. focal interpretation of the moved element” (Irurtzun 2007: 160).

Irurtzun (2007) presents several more arguments in favor of the cartographic approach, namely dialectal variation, focus in infinitival constructions, rightward foci, and fragmental answers. For a detailed discussion of these further arguments the reader is referred to his work and references cited therein.
We can conclude that the cartographic approach “is much more adequate than the NSR-based approach when confronted with pied-piping data” and long distance movement (Irurtzun 2007: 157).

In conclusion, I will follow Irurtzun (2007) and Ortiz de Urbina (1989, 1999), among others, assuming that the wh-phrase as well as the focalized constituent move to a left-peripheral position inside a complex CP. The exact composition of this CP will be discussed the following section.

4.2.2. The Basque left periphery

Section 4.2.1.2 showed that Duguine & Irurtzun (2014), Irurtzun (2007), and Ortiz de Urbina (1989, 1999), among many others, establish wh- and focus movement in Basque as a movement to the CP. The wh-phrase moves to the specifier while the verbal complex follows to C0 in order to satisfy a wh-criterion. In Ortiz de Urbina (1999), he refines the analysis following Rizzi’s (1997) rich left periphery and assumes that they both move to a FocP in the middle of the extended CP-layer. The same holds for focus movement.

As already presented in Section 3.3.2, I do not assume Rizzi’s (1997) fully extended left periphery but the reduced model presented by Kempchinsky (2013). She proposes a more reduced CP than either L. Rizzi or A. Irurtzun. Her left periphery consists of [ForceP - DS - FinP], where DS stands for Discourse Shells, “a category-neutral projection proposed [...] to account for iterative base-dislocations” (Kempchinsky 2013: 318). CLLDs land and HTLDs merge in the specifier of this DS. Wh-phrases and FFed constituents however move to the SpecFinP. For Spanish we have seen that as soon as SpecFinP is occupied, SpecIP is not longer available as a landing site for the subject. In this section, I want to use this approach to explain the word order phenomena in Basque. In order to achieve an unified account of the movements, I will present Haddican’s (2004) approach to Basque sentence structure.

Haddican (2004) proposes an analysis based on Kayne’s (1994) grammar of antisymmetry and Cinque’s (1999) universal hierarchy of functional heads. He introduces a Polarity Phrase (PoLP), which dominates IP and hosts either the negative particle ez, the emphatic affirmative particle ba, or the verbal predicate (VP). He builds his argument around the fact that the word order in Basque depends on the sentence polarity. In a standard affirmative sentence, the lexical part of the verbal complex appears directly preceding the auxiliary, as we have already seen in (4.1). In standard negative sentences the negative particle ez directly precedes the auxiliary, illustrated in (4.43a), with the lexical verb not allowed to appear between ez and the auxiliary (4.43b). The lexical verb appears after the Neg-Aux unit
and can be separated from it by “any kind and number of constituents” (Laka 1991: 72), as the examples (4.44a) and (4.44b) show (Etxepare 2003, Haddican 2004, Laka 1991).

(4.43)  

a. Etxea ez da etorri.  
   house not aux fall  
   ‘The house has not fallen down.’

b. Etxea erori ez da.  
   house fall not aux

(4.44)  

a. Ez da etxea erori.  
   not aux house fall  
   ‘The house has not fallen down.’

b. Ez dio Irunek Iboni etxea eman.  
   not aux Irune.erg Ibon.dat house.abs given  
   ‘Irune hasn’t given the house to Ibon.’

Haddican observes that the VP in positive and the negative particle ez in negative declarative sentences share several core features. Neither can be separated from the auxiliary, except by very few “evidential, evaluative and speech act ‘particles’” (see Haddican 2004: 97 for a detailed analysis of these particles). Additionally, wh-phrases and FFed constituents have to be directly adjacent to the ez+aux or V+aux unit. He assumes a PolP right above IP whose specifier has to be always occupied, either by the negative particle ez or the VP. The sentence structure for (4.44a) would be along the line of (4.45a), whereas the same sentence without negation would be as in (4.45b).

(4.45)  

a. negative sentence structure  

b. positive sentence structure

---

47 The negative particle ez is merged further down in an NegP. He presents evidences for its movement to PolP based on the interaction with preverbal particles like omen, ei, bide, etc. For a detailed analysis see Haddican (2004: chap. 2).
4.2 On syntactic structures in Basque \textit{wh}-questions

Returning to the \textit{wh}/focus-movement, I follow Haddican (2004), Irurtzun (2007), Ortíz de Urbina (1995, 1999), and Uriagereka (1992), and others, assuming that Basque \textit{wh}-phrases and focalized constituents move to a left-peripheral position in an extended CP. In addition, I will adopt a CP-layer as in Kempchinsky (2013). As already presented in Section 3.3.2, the basic hypothesis is that the \textit{wh}-phrase and FFed element move to SpecFinP, whereas topicalized material move to a DS between ForceP and FinP.

I will adopt Haddican’s (2004) idea that in affirmative all-focus sentences the whole VP (or vP) moves to the left, resulting in the desired word order SOV. However, I will associate his PolP with my FinP, assuming that its specifier, being the preverbal focus position, must always be filled. In the case of an all-focus sentence, it is the whole VP (or vP) that moves there, in \textit{wh}-questions, the \textit{wh}-phrase, and in sentences with a focalized constituent, that constituent. In either of the two latter cases, if further constituents are topicalized, they land in the DSs. The verbal complex, or in the case of the presence of the negative particle \textit{ez} only the auxiliary, moves to Fin$^0$. The derivations are illustrated in (4.46) for an all-focus sentence, in (4.47) for a \textit{wh}-question, and in (4.48) for a phrase with a FFed constituent including a topicalized element in DS.

Peio.\text{\textit{erg}} present Ikea-in bought aux  
‘Peio bought the present in Ikea.’

b. 

\begin{center}
\begin{tikzpicture}
\node[ FinP ] {FinP}
child {node [vP] {vP}}
child {node [Fin'] {Fin'}}
child {node [IP] {IP}}
child {node [Peiok opari Ikean] {Peiok opari Ikean}}
child {node [eros du] {eros du}}
child {node [t_i] {t_i}}
child {node [t_j] {t_j}}
\end{tikzpicture}
\end{center}

\footnote{If one assumes the ‘\textit{wh}-criterion’, or Ortiz de Urbina’s (1995) ‘affective operator criterion’, the verbal complex must move further up to Fin$^0$ in order to satisfy the criterion. However, I will present a Q-based approach in the next chapter in which we don’t need these criterions anymore. Nevertheless, I will assume that the verbal complex moves to Fin$^0$. Here Basque contrasts with Spanish. As seen in Section 3.3.2.2 for Spanish, Fin$^0$ is occupied by \textit{f} and the verb cannot move there.}
4.2 On syntactic structures in Basque wh-questions

(4.47) a. Non erosi du Peiok oparia?
where bought aux present Peio.erg
‘Where did Peio buy the present?’

b. 

(4.48) a. oparia IKEAN erosi du Peiok.
present IKEA bought aux Peio.erg
‘The present, Peio bought it at IKEA.’

b. 

92
In summary, I have adopted an approach similar to the standard established by J. Ortiz de Urbina and others. That is, I assume a movement of the *wh*-phrase and the FFed constituent to the left periphery. However, I claim that this position is not SpecFocP but SpecFinP. Furthermore, I associate SpecFinP with Haddican’s SpecPolP, from which I conclude that SpecFinP has to always be filled by a constituent. In a *wh*-question or a focus sentence, it is occupied by the *wh*-phrase or the focalized element, respectively, and in an all-focus sentence by the whole VP.

In the next chapter, I will explore the reason for all these movements that we have seen so far, both in the Basque as well as the Spanish constructions. Furthermore, I will present an approach by S. Cable to explain *wh*- and focus movement.
Chapter 5

Q-particle and wh-movement

The introduction to Section 3.3 already addressed the two main questions in the ongoing debate on wh-interrogatives. The first question concerns the landing site of wh-phrases. As presented in Chapter 3 and 4, I argue for a movement to SpecFinP in both languages in the case of simple wh-phrases. Complex wh-phrases, on the other hand, move to a higher Discourse Shell (DS) in Spanish but not in Basque, where they move to the same position as simple wh-phrases.

The second question concerns the reason wh-phrases move to the left periphery at all. In Rizzi (1996: 64) and subsequent work, the cause of wh-movement is a [wh]-feature which triggers movement through a probe-goal agreement relation. That is, in order to satisfy its [-wh]-feature, FocP probes for a constituent with a [+wh]-feature. It establishes an agreement relation with the wh-phrase and attracts the latter to its Specifier. In order to fulfill the wh-criterion, the verbal head moves to Foc\(^0\). What follows is a different approach to this ongoing debate, namely Cable’s (2010) theory of Q(uestion)-Particles.

In (5.1), I will present Cable’s (2010) theory of Q-particle and its implication for Spanish and Basque. I will show that Spanish is of the same language type as English or Tlingit, concerning question formation, in Section 5.1.1. A Q-particle moves to the left periphery, taking the XP containing the wh-word along as complement. For Basque, on the other hand, I will argue in Section 5.1.2 that at its core it is similar to wh-in-situ-languages like Japanese or Korean. The Q-particle moves to a CP-head independently of the wh-word which stays in-situ. I will claim in Section 5.2 that the reason why Basque exhibits wh-fronting is caused by an independent movement in order to satisfy a focus condition and is not directly linked to the wh-phrase as such.
5.1 On Q-Particles – an introduction to the grammar of Q

The theory of Q-particle was, in a slightly different way, already introduced and developed by Watanabe (1992), Hagstrom (1998) and others. Cable (2008, 2010) and subsequent work, however, develop a model that includes not only \textit{wh-in-situ} languages, where question particles occur quite often, but also expands the theory to well-known and well-described \textit{wh-fronting} languages like English. He bases his analysis on the Na-Dene language \textit{Tlingit}, a language spoken mainly in southeastern Alaska. The reason he chose Tlingit lies in the fact that questions in that language are overtly marked by a particle, independent of the \textit{wh}-word and obligatorily present in all partial interrogatives, and in addition it exhibits \textit{wh}-fronting. This combination hardly occurs in any language in the world.

Cable assumes a Q-particle that is independent of the \textit{wh}-word and which enters in an agreement relation with a C_{[Q]} head. Due to this relation, the Q-particle moves, overtly or covertly, to the corresponding C-head. He adopts the idea that the \textit{wh}-word itself has no direct relation with the left-peripheral CP. Instead, he assumes that the XP containing the \textit{wh}-word is the sister of the Q-particle (Henceforward, I will use XP_{wh} for the ‘XP containing the \textit{wh}-word’). With that, it is guaranteed that the Q-particle always \textit{c-commands} the \textit{wh}-word. He singles out two possible configurations. Either the Q-particle takes the XP_{wh} as a complement, or the former adjoins to the latter. In the first configuration, the Q-particle projects a Q(uestion-Particle)-Phrase (QP), which then enters an agreement relation with a C_{[Q]} -head and moves overtly to its specifier, taking its complement XP along. This can be seen in (5.1). This configuration represents a \textit{wh}-fronting language. In the second configuration, shown in (5.2), the XP_{wh} takes the Q-particle as an adjunction, projecting the same type of XP. In this case however, though the Q-particle still moves to the higher SpecCP, it does not take the XP along with it. This structure would describe most of the \textit{wh-in-situ}-languages. The result of this theory is that \textit{wh}-fronting would be only a side effect, caused by the relation between the Q-particle and the XP_{wh}. \textit{wh}-fronting would not be caused by a special relation between the \textit{wh}-word and a C-head, but rather is a by-product of Q-movement.
5.1 On Q-Particles – an introduction to the grammar of Q

(5.1) wh-fronting as secondary effect of Q-movement in a “Q-projection language” - e.g. in Tlingit

\[
\begin{align*}
\text{Complementation} & \quad \text{QP}_1 \\
\text{XP} & \quad Q
\end{align*}
\]

...wh-word...

\[
\text{CP} \quad \text{QP}_1
\]

Agree

overt movement

(Cable 2010: 85)

(5.2) Q-movement in “Q-adjunction languages”, e.g. Japanese

\[
\begin{align*}
\text{Adjunction} & \quad \text{CP} \\
\text{XP} & \quad Q_1
\end{align*}
\]

...wh-word...

\[
\text{CP} \quad \text{Q}_1
\]

\[
\text{IP} \quad \text{C}_Q \quad \text{Q}_1
\]

Agree

overt movement

Cable (2010: 85), adapted from Hagstrom (1998)

Cable’s (2010) idea is based on two main assumptions. First, it is the Q-particle that enters an agreement relation with a C\textsubscript{Q}-head, not the wh-word. Second, it is the Q-particle that moves either to the left or right periphery, overtly or covertly, depending on the language configuration. The languages’ distinctions now come from different parameter settings in the relation between the Q-particle and the XP\textsubscript{wh} as well as the type of movement.

Cable differentiates between two types of languages. English, Tlingit, and Sinhala, where the XP\textsubscript{wh} is a complement to the Q-particle and both are immediately dominated by a QP, belong to the
first type, illustrated in (5.1). The Q-particle passes its [+Q] feature onto the QP, which enters an agreement relation with a C-head. Attracted by a probe by C, the whole QP moves to the left periphery and takes the XP_{wh} and the Q-particle along. In English and Tlingit the movement is overt, whereas in Sinhala it is covert. The latter case results in a superficial \textit{wh-in-situ}-language, which nevertheless has the underlying structure as in (5.1). The sole difference is that the QP moves covertly to the C-head. Cable calls the first type ‘Q-projection languages’. Languages like Japanese or Korean belong to the second type, illustrated in (5.2). They differ in that the Q-particle is adjoined to the XP_{wh} and the node immediately dominating both of them is of the same type as XP_{wh}. In this case, the Q-particle enjoys greater freedom and C[Q] attracts only the particle without its sister XP_{wh}. Thus, the \textit{wh}-word stays \textit{in-situ}. Cable calls this type ‘Q-adjunction languages’ (Cable 2010: 85-86). In order to classify the language type, Cable (2010: 102-103) proposes the following five parameters:

\begin{enumerate}
\item the nature of the relation between the XP containing the \textit{wh}-word and the Q-particle
\item the covert or overt movement of the Q-particle
\item the phonological realization of the Q-particle
\item the \textit{wh}/Q-particle agreement relation
\item the number of Q-particle in multiple \textit{wh}-questions.
\end{enumerate}

I have already broached (i) - (iii) in the short description above. I will nevertheless describe the five parameter in more detail, since I want to apply them to Spanish and Basque in the next section.

The first parameter has the two settings mentioned; The Q-particle can either take the XP_{wh} as a complement or the former can be adjoined to the latter. Each option has different implications for the \textit{wh}-phrase, which is either moving along with the QP with the former or staying \textit{in-situ} with the latter. The second parameters refers to the covert or overt movement of the Q-particle. That is, if Q moves before or after spell-out. I will assume, for both Basque and Spanish, that the movement is overt. The third parameter identifies the phonological realization or non-realization of the Q-particle. Languages like Japanese and Tlingit use a phonologically realized Q particle, like \textit{ka} and \textit{sa}, respectively, whereas English or Spanish do not. Their Q-particles are phonologically empty. This can also be the case when the Q-particle is inseparably fused with the \textit{wh}-word and cannot be analyzed alone.\footnote{There had been attempts to link the "wh-"morpheme in English to the Q-particle. See Cable (2010: 114-115) for a detailed analysis. The same argumentation can be adopted for Spanish questions, in which the "qu-"morpheme occurs with the same regularity. Nevertheless, I will follow Cable (2010) and assume a phonologically empty Q-particle for both Spanish and English.} I will argue that
Basque normally has a phonologically empty but independent Q-particle. For Spanish on the other hand the Q-particle is inseparably fused with the wh-phrase.

The fourth parameter is needed to explain differences regarding pied-piping structures. In order to implement it, Cable adopts the ideas from Kratzer & Shimoyama (2002) and Pesetsky & Torrego (2001) of a wh/Q-agreement relation in some languages. To classify languages with regard to this parameter, Cable distinguishes them by their behavior within pied-piping structures. He defines a limited pied-piping structure as a “structure where pied-piping past islands and pied piping past lexical categories is not permitted” and he refers to “limited pied-piping language [...] to languages where all pied-piping structures are instances of limited pied-piping.” (Cable 2010: 144/145). Languages like Japanese or Tlingit (5.4a) can, for example, pied-pipe relative clauses which contain the wh-word, in contrast to languages like German or English, which cannot (5.4b).

(5.4) a. [ [ Wáa kwligeyi ] xáat ] sá i tuwáa sigóo? how it.is.big.REL fish Q your at.it.is.happy
   Literally: ‘A fish that is how big do you want?’

b. * [dp A fish [cp that is how big ] ] do you want?
c. * [Which ] 1 did you read [dp t₁ book ] ?
d. * I wonder [ [np pictures of whom ] John bought ]?
e. * I wonder [ [ap proud of whom ] John was ] ?
f. * I wonder [ [vp eaten what ] John has ] ?

Cable (2010: 143/144)

The source for these restrictions lies in the setting of the forth parameter. It is either necessary to establish an agreement relation between the Q-particle and the wh-word or not. In the first case, lexical categories ((5.4d)-(5.4f)) or syntactic islands ((5.4b)-(5.4c)) can block this agreement (for detailed syntactic structures, see Cable 2010: 146-147). This happens in English, and, as I will show, also in Spanish. Both are so-called limited pied-piping languages. If no relation has to be established, obviously nothing can intervene and these pied-piping structures are grammatical. Tlingit, Japanese, and Basque belong to this second language type.

The fifth and last parameter accounts for intervention and superiority effects in multiple wh-questions. A language exhibits superiority effects if wh-phrases in a multiple wh-question have to appear in a fixed linear order to render a grammatical sentence. The order normally corresponds to the neu-
5.1 On Q-Particles – an introduction to the grammar of Q

total constituent order. Intervention effects arise, if a so-called offending operator\(^{56}\) cannot intervene between multiple wh-phrases. Both effects are demonstrated in (5.5) and (5.6).

(5.5) Superiority effects in English multiple wh-questions
a. Who bought what?
   b. * What did who buy?  
   
Cable (2010: 122/123)

(5.6) Intervention effects with German multiple wh-questions
a. * Wen hat niemand wo gesehen?
   whom has nobody where seen
   ‘Where did nobody see whom?’

b. Wen hat wo niemand gesehen?
   whom has where nobody seen
   ‘Where did nobody see whom?’

c. Wen hat Luise wo gesehen?
   who has Luise where seen
   ‘Where did Luise see whom?’

Beck (2006: 4)

Beck (2006), Cable (2010), and others, have noted an interesting connection between the two phenomena, which is given in (5.7).

(5.7) The Complementarity of Superiority Effects and Intervention Effects

In any language L, the in-situ wh-words of a multiple wh-question of L are subject to Intervention Effects if and only if the multiple wh-questions of L are not subject to Superiority Effects.

Cable (2010: 132)

That means, if any language exhibits intervention effects, it cannot show superiority effects, and vice versa. Cable attributes the dichotomy to the number of Q-particles present in a multiple wh-question. If it is only a single Q-particle, a language will exhibit intervention effects. Superiority effects arise, however, when each wh-phrase is accompanied by its own Q-particle (Cable 2010: chap.4.4). He presents evidence that in languages with phonologically realized Q-particles, this exact distinction can be observed. For example, Tlingit needs multiple Q-particles in multiple wh-questions and thus exhibits superiority effects. Navajo, on the other hand, only has a single Q-particle in multiple wh-questions.

\(^{56}\) Offending operators can be nominal quantifiers like ‘only, even, also’, etc., adverbial quantifiers like ‘always, often, never’, etc., or a negative particle (Beck 2006: 3-4).
and thus is subject to intervention effects (Cable 2010: 128). Accepting Cable’s (2010) theory, the logical conclusion is to assume the same differences for languages with phonologically empty Q-particle. Thus, two possible configurations can be analyzed as in (5.9) and (5.8).

(5.8) Multiple QPs Predicts Superiority Effects

Since all the wh-words are associated with QPs, the usual logic of ‘Attract Closest’ will predict the appearance of Superiority Effects in multiple wh-questions.

a. Superiority-Satisfying Structure

\[
\begin{align*}
[ [ Q_p \text{ Who Q }]_1 & \quad [ t_1 \text{ bought } Q_p \text{ what Q } ] \\
\uparrow & \\
\text{Attract Closest Satisfied}
\end{align*}
\]

b. Superiority-Violating Structure

\[
\begin{align*}
* [ [ Q_p \text{ What Q }]_1 & \quad [ \text{ did } Q_p \text{ who Q } \text{ buy } t_1 ] \\
\uparrow & \\
\text{Attract Closest Violated}
\end{align*}
\]

(5.9) Configuration Resulting in an Intervention Effect

\[
\begin{align*}
[ \ldots \text{ Q [ \ldots Offending Operator [ \ldots [ \text{ wh-word } \ldots ] ] ] } \\
\downarrow & \\
\text{No Q-particle}
\end{align*}
\]

The cause of the ungrammaticality of the last sentence is the intervening ‘offending operator’. That is, no matter what the relation is between the Q-particle and the wh-word, the former has to be the first focus-sensitive operator that c-commands the latter. This precondition is violated in the configuration in (5.9), where the first focus-sensitive operator is the ‘offending operator’ and not the Q-particle. Thus, the sentence is ungrammatical (Cable 2010: 127-128). Unfortunately, knowledge about the number of Q-particles does not help to determine if a language is a Q-adjunction or Q-projection language. Cable offers examples of either type in both groups. Some languages, like English, even exhibit both configurations, depending on if the wh-phrase is d-linked or not. In the following two sections I will discuss this concept in reference to Spanish and Basque, respectively, though we cannot go as far as to draw a conclusion about the type of relation between the XP_wh and the Q-particle. Nevertheless, we will see that Basque as well as Spanish confirm the generalization about the relation between intervention and superiority effects put forward in (5.7).
A last essential condition in Cable’s (2010) Q-based account is the so-called “QP-Intervention Condition” (QPIC) (Cable 2010: 57). It is valid for all languages, but serves mainly to distinguish between Q-adjunction and Q-projection languages, since only the latter project a QP.

(5.10) **QP-Intervention Condition**
A QP cannot intervene between a functional head F and a phrase selected by F.
(Such an intervening QP blocks the selectional relations between F and the lower phrase, as illustrated in the following.)

With the help of the QPIC, Cable explains, among other phenomena, restrictions on possessor extraction (5.11), wh-determiner extraction (5.12), and P(reposition)-Stranding (5.13). In all three cases, the QP cannot intervene between the dominating functional Category DP or PP (Cable 2010: 106).

(5.11) **No possessor extraction in wh-fronting languages**
   a. \([\text{dp, Whose book }]\) did you read \(t_1\)?
   b. *\([\text{Whose }]_1\) did you read \([\text{dp t}_1 \text{ book }]\)?

(5.12) **No wh-determiner extraction in wh-fronting languages**
   a. \([\text{dp, Which book }]\) did you read \(t_1\)?
   b. *\([\text{Which }]_1\) did you read \([\text{dp t}_1 \text{ book }]\)?

(5.13) **No P-Stranding in most of the wh-fronting languages**
   a. \(\text{¿[pp, Con quién ] has venido t}_1\?\)
      “With whom did you come?”
      with whom have.2sg come.part
   b. *\(\text{¿[ Quién ] has venido [pp con t}_1\]?\)
      “Who did you come with?”
      who have.2sg come.part with

(English examples from Cable (2010: 106), Spanish my own)
Nevertheless, Cable notes that exceptions can be found for all three constructions. The most famous one is probably P-Stranding in English. As we can see in (5.13b), the English translation is perfectly acceptable. In order to explain this phenomenon, Cable (2010: 111-112) argues that P is a lexical category in English. If P is not a functional head anymore, no QPIC is violated by P-Stranding.

Now that I have given a short introduction to the main characteristics of Cable’s (2010) ‘Grammar of Q’, I can advance the following hypothesis: Spanish resembles English and other 'standard' wh-fronting languages in that the Q-particle takes the XP containing the wh-phrase as complement and projects a QP, which then moves to a left peripheral position inside the CP-layer. The Spanish Q-particle is inseparably fused with the wh-word. Furthermore, it is a limited pied-piping language in the sense of Cable. Finally, Spanish does not exhibit intervention but superiority effects in multiple wh-questions. Basque, on the contrary, shows some crucial parallels with wh-in-situ-languages. That is, Basque has an independent but phonologically empty Q-particle which is adjoint to the XP containing the wh-phrase. The Q-particle enters in an agreement relation with a C-head and moves independently to the left. Basque is not a limited pied-piping language and exhibits intervention but not superiority effects in multiple wh-questions against a common assumption. Thus, I claim that Basque exhibits wh-fronting not due to QP-movement, as in Spanish, but due to the necessity to satisfy an independent focus feature in the preverbal position.

In the following two sections, I will provide evidence to support my hypothesis. I will start in Section 5.1.1 by showing that Spanish is a Q-projection language. Then, in Section 5.1.2, I will argue for Basque being a Q-adjunction language.

5.1.1. On the Q-particle in Spanish

We have seen in Chapter 3 that Spanish is a wh-fronting language that does not have a phonologically realized Q-particle. I assume that Spanish is a Q-projection language. That is, a phonologically empty Q takes the XP containing the wh-phrase as a complement, projecting a QP, as shown in (5.1). Then, QP enters an agreement relation with a C-head and moves up, taking the XP, along. wh-fronting is only a secondary effect of QP-movement.

If we recall the examples (5.11) - (5.13), here repeated as (5.14) - (5.16) in their Spanish version, we observe their ungrammaticality in the b.-versions.

(5.14) No possessor extraction in Old Spanish
a. [Cuyo prólogo] has leído?
   whose prologue have.2sg read
   ‘Whose prologue did you read?’

b. * [Cuyo]1 has leído [prólogo t1]?
   whose have.2sg read prologue
   ‘Whose prologue did you read?’

(5.15) No wh-determiner extraction

a. [Cuál libro] has leído?
   what book have.2sg read
   ‘Which book did you read?’

b. * [Cuál]1 has leído [libro t1]?
   what have.2sg read book
   ‘Which book have you read?’

(5.16) No P-Stranding

a. ¿[pp, Con quién] has venido t1?
   with whom have.2sg come.PART
   ‘With whom did you come?’

b. * ¿[Quién ]1 has venido [pp con t1 ]?
   who have.2sg come.PART with
   ‘Who did you come with?’

A construction with the prenominal possessor *cuyo ‘whose’ as in (5.14) was only possible in Old Spanish and is not available in Modern Spanish anymore (Bosque & Gutiérrez-Rexach 2011: 452). That is, Modern Spanish has lost the prenominal possessor which can also function as interrogatives. Nowadays, Spanish possessive interrogatives have to be expressed by a postnominal prepositional construction like *de quién ‘of whom’. Thus, the sentence in (5.14) would correctly be expressed as in (5.17) with a postnominal possessor.

(5.17) [El prologo de quién]1 has leído t1?
   the prologue of whom have.2sg read
   ‘Whose prologue did you read?’

The same accounts for sentences like (5.4d) and (5.11). They are possible in Spanish but only with postnominal possessors as shown in (5.18).

(5.18) a. ¿[np, El hermano de quién ] viene mañana t1?
   the brother of whom comes tomorrow
   ‘Whose brother comes tomorrow?’
5.1 On Q-Particles – an introduction to the grammar of Q

b. ¿[pp, De quién ] viene [ el hermano t₁ ] mañana?
   of whom comes the brother tomorrow
   ‘Whose brother comes tomorrow?’

The post-nominal possessors and the resulting possibility of so-called split-DP, as in (5.18b) are a common phenomenon in Romance but also other languages (Cable 2010: 169-170). Cable states that in these languages, either Q takes along the whole XP_{wh} (pied-piping) or only the wh-phrase itself, depending on the type of the XP.

The forth parameter concerns the Q/wh-agreement relation which in some languages has to be established and in others not. Cable uses structures of limited pied-piping in order to check for it. As the examples in (5.4) show, a language like English has to establish a agreement relation between the Q-particle and the wh-word. This agreement can be blocked by lexical phrases or syntactic islands. Comparing the Spanish structures in (5.19a) - (5.19d) with their English counterparts in (5.4b), (5.4c), (5.4e), and (5.4f), respectively, we see that the former sentences are also ungrammatical. Therefore I assume that Spanish is a limited-pied-piping language in the sense of Cable (2010: 144).

(5.19) a. * [dp La película [cp que ha estrenado qué director]] quieres ver?
   the movie that have.3sg premiered what director want.2sg watch
   ‘You want to watch the movie that was premiered by which director?’
   example adapted from Bosque & Gutiérrez- Rexach (2011: 454)

b. * ¿[ Cuáles ]₁ conoces [ dp, t₁ poemas de Neruda ]?
   which know.2sg poems of Neruda
   ‘Which of Neruda’s poems do you know?’
   Contreras (1999: 1938)

c. * Me preguntó [ [ap orgulloso de quién ] ha estado Juan. ]
   myself ask.1sg proud of whom has been Juan
   ‘I wonder who John has been proud of.’

d. * Me preguntó [ [vp comido qué ] ha Juan. ]
   myself ask.1sg eaten what has John
   ‘I wonder what John has eaten.’

Finally, I will look at Spanish multiple wh-questions. In Chapter 3 I have presented their main properties. In Spanish, it is obligatory to front one wh-phrase, whereas the others have to stay in-situ. In the case of several argument wh-phrases, Spanish exhibits no superiority effects. That is, it doesn’t matter which of the two wh-phrases is fronted and which stays in-situ, illustrated in (5.20).
5.1 On Q-Particles – an introduction to the grammar of Q

(5.20) a. Quién dijo qué?
     who said.3sg what
     ‘Who said what?’

b. Qué dijo quién?

Bosque & Gutiérrez-Rexach (2011: 514)

Nevertheless, in the case of argument and adjunct wh-phrases, superiority effects seem to exist. Although these effects are “probablemente marginales y no aparecen en ciertos dialectos”51 (Bosque & Gutiérrez-Rexach 2011: 515), the argument wh-phrase has to appear sentence final, as the examples (5.21) and (5.22) show.

(5.21) a. Cuándo compró Pepe qué?
     when bought.3sg Pepe what
     ‘What did Pep buy when?’

b. * Qué compró Pepe cuándo?

(5.22) a. Cómo has reparado qué coche?
     How have.2sg repaired which car
     ‘Which car have you repaired how?’

b. * Qué coche has reparado cómo?

Bosque & Gutiérrez-Rexach (2011: 515)

According to Cable’s (2010) definition in (5.7), Spanish shouldn’t exhibit any intervention effects, since it already shows, at least in some cases, superiority effects. As the sentences in (5.23a) - (5.23c) show, the prediction is correct.

(5.23) a. A quién no ha dado Juan qué?
     a who not have.3sg gave Juan what?
     ‘To whom didn’t Juan give what?’

b. Quién no ha comprado qué?
     who not have.3sg what
     ‘Who hasn’t bought what?’

c. Quién no vive en qué lugar?
     who not live in which place
     ‘Who doesn’t live in which place?’

51 They are probably marginal and do not appear in certain dialects.
To summarize, the Spanish Q-particle is phonologically empty, takes the $X_{P_{wh}}$ as complement and thus exhibit wh-fronting. The language can be classified as a *limited pied-piping* language in the sense of Cable (2010). That is, the Q-particle has to enter an agreement relation with the *wh*-word, which prevents pied-piping past lexical categories and past island, as shown in the examples. Finally, Spanish exhibits superiority effects to a certain degree and no intervention effects. I therefore follow Cable assuming that Spanish has a Q particle for every *wh*-phrase in a multiple *wh*-question. Spanish resemble most of the *wh*-fronting languages presented and analyzed in Cable (2010: chap. 4).

If we apply Cable’s (2010) approach to Spanish *wh*-sentences, we get the following derivations: the structure for a simple *wh*-phrase would be like (5.24b), for a complex like (5.25b), and for a high adjunct *wh*-phrase like (5.26b).

\begin{align*}
(5.24) \quad \text{a.} & \quad ¿Qué compró Pedro? \\
& \quad \text{what bought Peter?} \\
& \quad ¿Qué compró Pedro? \\
& \quad \text{‘What did Peter buy?’} \\

\text{b.}
\end{align*}
(5.25) a. ¿A cuál de estas chicas tu hermana (la) había visitado en Sicilia?
   ‘Which of the girls had your sister visited in Sicily?’

b.}

\[ \text{QP}_k \]

\[ \text{A cuál de estas chicas} \]

\[ \text{FinP} \]

\[ \text{[tu hermana]} \]

\[ \text{Fin} \]

\[ \text{Fin}^0 \]

\[ \text{IP} \]

\[ \text{[había visitado]} \]

\[ \text{vP} \]

\[ t_j \]

\[ t_k \] en Sicilia?
5.1 On Q-Particles – an introduction to the grammar of Q

(5.26) a. ¿Por qué el tren se retrasa todos los días?
    Why the train be late all the days
    ‘Why does the train always arrive late?’

b. We can see that the Q-based approach by Cable (2010) accounts for the observed structures in Spanish wh-interrogatives. In the following section, I will turn to Basque, discussing whether Basque is the wh-fronting language it seems to be.

5.1.2. On the Q-particle in Basque

We have already seen in Section 3.2 and 4.1 that Basque seems to share some core features with Spanish concerning wh-question formation. That is, both languages exhibit wh-fronting, normally obligatory wh-phrase-verb adjacency, distinct behavior with adverbial wh-phrases like zergatik/por qué ‘why’, wh-island restrictions, and others. Nevertheless, Basque shows some different characteristics with regard to yes/no and indirect questions, in the internal structure of the wh-words, in multiple wh-questions, as well as in pied-piping constructions. Although both languages share several superficial features, I will argue in favor of a different analysis for Basque wh-phrases, namely that Basque behaves in some
5.1 On Q-Particles – an introduction to the grammar of Q

crucial aspects more like the *wh-in-situ*-languages Japanese and Korean. In order to show that Basque has a different core *wh*-phrase structure than Spanish, I will again come back to Cable’s (2010) five main parameters, defined in (5.3), to establish the language type.

The first point in Cable’s (2010) list is the relation between the *wh*-phrase and the Q-particle. For Spanish I have argued that the Q-particle takes the XP<sub>*wh*</sub> as a complement, projecting a QP. For Basque, I will claim that the latter takes the former as an adjunct, allowing the Q-particle to move separately without taking the XP<sub>*wh*</sub> along. Cable (2010) shows the independence of the Q-particle in languages like Japanese or Korean by its behavior in *wh*-questions. The difficulty for Basque of course is that we cannot apply strictly the same arguments in direct *wh*-questions, since the Q-particle is not phonologically realized. Therefore, we have to use indirect evidences in *wh*-indefinite constructions. However, in order to present the evidences appropriately, I will first discuss the parameters (iii) - (v) in Cable’s (2010) approach (cf. (5.3)).

Basque has some strong independent characteristics that can motivate a Q-particle approach. First I will have a look at a potentially phonologically realized Q-particle in some constructions in Basque, followed by an analysis of the internal structure of Basque *wh*-words. Finally I will give a third argument for an independently moving Q-particle in Basque.

Looking at the standard Basque and several Basque dialects, we can find the morphological question markers *al* and/or -a in yes/no questions. The former one appears between the auxiliary and the participle (5.27), the latter is attached directly to the verb (5.28).

(5.27) Gure literaturak aurrerakada haundirik egin al du urte hauetan?
our literature.erg improvement great make *al* AUX year these.in
‘Has our literature made any great improvement in the last years?’

(5.28) Oro egin eta desegin, ez dea lur huntako legea?
all do and undo not is.a earth this.of law
‘To do and undo everything; isn’t that the law of this world?’

Etxepare & Ortiz de Urbina (2003: 467)

Also indirect questions exhibit morphological marking on the embedded verbal form.<sup>52</sup> The marker -n can appear in indirect yes/no as well as in indirect *wh*-questions ((5.29) and (5.30) respectively). In

---

<sup>52</sup> The extra marking is not always necessary, but only with the matrix verb *galdetu* ‘ask’ obligatory.
addition, indirect questions can be marked with the complementizer *ea*, directly preceding it (5.31).

(5.29)   Ez dakit zure laguna etorriko d-en. not know your friend come.FUT AUX-COMP
‘I don’t know whether your friend will come.’

(5.30)   Ez dakit hargatik nola daitekeen hori. not know however how can.be.comp that
‘I don’t know, however, how that can be.’

(5.31)   Bidriosek ea bazuen Institutuko berririk galdetu zion. Bidrios.ERG ea ba.AUX Insitute.REL knew.PRTT ask AUX
‘Bidrio asked him whether he had any news from the school.’

Hence, the existence of a Q-particle has grounds for support in Basque, although it is not phonologically realized in direct *wh*-questions. But if we have a look at the internal structure of *wh*-words in Basque, we will see another motivation for the assumption of an independent Q-particle.

The main *wh*-words in Basque are ‘nor’ ‘who’, *zer* ‘what’, *zein* ‘which’, *non* ‘where’, *noiz* ‘when’ and *nola* ‘how’. A closer look at Table 5.1 reveals a transparent derivational structure for *wh*-indefinites (Trask 2003: 154ff).53

Table 5.1: Paradigm for Basque *wh*-based indefinites

<table>
<thead>
<tr>
<th><em>wh</em>-word + particle</th>
<th>-bait</th>
<th>-i/-e-</th>
<th><em>edo</em>-</th>
<th>+ ere</th>
<th>+ Ø</th>
</tr>
</thead>
<tbody>
<tr>
<td>nor</td>
<td>somebody</td>
<td>nobody</td>
<td>anyone</td>
<td>whoever</td>
<td>who</td>
</tr>
<tr>
<td>zer</td>
<td>something</td>
<td>nothing</td>
<td>any</td>
<td>whatever</td>
<td>what</td>
</tr>
<tr>
<td>non</td>
<td>somewhere</td>
<td>nowhere</td>
<td>wherever</td>
<td>wherever</td>
<td>where</td>
</tr>
<tr>
<td>noiz</td>
<td>sometime</td>
<td>never</td>
<td>whenever</td>
<td>whatever</td>
<td>when</td>
</tr>
<tr>
<td>nola(z)</td>
<td>somehow</td>
<td>by no means</td>
<td>anyhow</td>
<td>how</td>
<td></td>
</tr>
</tbody>
</table>

The combinations of *wh*-words plus particle, e.g. coordination marker like *edo* ‘or’, deriving a universal or existential quantifier reading can be found regularly in the languages of the world. Jayaseelan

---

53 The same transparent derivations are possible with *nora* ‘where to’ and *nondik* ‘where from’. As the table shows already in the case of *nola*, not every possible combination exists. For example, there is no *zeinbait* or *nondikbait*. 

110
(2001) describes similar phenomena for Malayalam, Mandarin Chinese, Japanese, and others. In a nutshell, his analysis says that the *wh*-word itself is only a “variable with a range restriction determined by semantic features. [...] The connective (the coordination marker) is crucially involved in giving an interpretation to the variable, which (by itself) cannot be interpreted which needs the particle to get an interpretation” (Jayaseelan 2001: 75). The coordination marker, which has an interpretable feature F, takes the variable, which has an uninterpretable feature F, as complement and feature agreement applies. Transferring Jayaseelan’s (2001) approach to Cable’s (2010) Q-based theory, the relation between the marker and the *wh*-word looks very similar to the necessity for Q to c-command the *wh*-word.

A further argument that Q-particles in Basque are independent elements of the *wh*-words is the following construction. The *wh*-words nor `who’ and zein `which’ can appear in a non-interrogative function rendering a distributive quantifier reading, as illustrated in (5.32a)/(5.32b) and (5.33a)/(5.33b), respectively. No extra marking is necessary. The difference to the question ‘Who loves his country?’ is the word order, the question intonation, and the phonologically empty Q-particle. The contrast is shown in (5.32c) and (5.32b).³⁴

\[
\text{(5.32) a. } \text{Nori berea eman behar zaio.} \\
\quad \text{who.DAT his give need AUX} \\
\quad \text{‘Everybody shall get what he deserves.’} \\
\text{b. } \text{Nork bere herria maite du.} \\
\quad \text{who.ERG his country love AUX} \\
\quad \text{‘Everybody loves his country.’} \\
\text{c. } \text{Nork maite du bere herria?} \\
\quad \text{who.ERG love AUX his country} \\
\quad \text{‘Who does his country love?’}
\]

\[
\text{(5.33) a. } \text{Zein bere tokian jarri zen} \\
\quad \text{which his place.LOC positioned AUX} \\
\quad \text{‘Everyone, took his, place’} \\
\text{b. } \text{Zein bere etxean gelditzea onena da.} \\
\quad \text{which his house.LOC stay.NOM-the best-the is} \\
\quad \text{‘The best is that everybody stays home.’}
\]

Etxepare (2002)

³⁴ The distributive use of *wh*-words is subject to some restrictions. For example, this kind of reading is not possible with all kinds of *wh*-words. zer ‘what’ cannot be used in this way. A sentence like (i.) is ungrammatical (Etxepare, p.c.)

\[
\text{i. } \text{Mundu honetan zerk bere saria du} \\
\quad \text{world this-in what.ERG its reward has} \\
\quad \text{‘In this world, each thing has its corresponding reward.’}
\]

This difference between ‘who/which’ and ‘what’ can be found in several other languages too. In Japanese as well as in Spanish, similar constructions are not possible with the latter *wh*-word. A further restriction is the obligatory presents of a pronoun in the constituent following the *wh*-word. Without the pronoun, no distributive reading is possible, as illustrated in (ii.) (Etxepare 2002: 236).
5.1 On Q-Particles – an introduction to the grammar of Q

This observation gives evidence that the Q-particle is not fused with the wh-word in Basque. That is, the wh-word itself does not have an interpretation of its own, but needs an other element to get it. This is in line with Jayaseelan’s (2001) and Cable’s (2010) interpretation.

A final argument in favor of an independent, empty Q-particle is the only construction, in which focus- and wh-movement really differ in Basque, namely the IP-internal position of wh-phrases and FFed constituents in negative sentences. We have seen in Chapter 4 that both constructions share almost all properties, except one. In Section 4.1.1, the examples in (4.11), here repeated as (5.34), show the two possible positions for a focalized constituent in negative sentences. Examples (5.35), on the other hand, show the same sentences with a wh-phrase instead of a focalized element. As indicated, only (5.35a) is grammatical.

(5.34)

   Jon.erg not aux petition sign
   ‘Jon did not sign the petition.’

b. Ez du eskaera Jon. erg sinatu.
   not aux petition Jon.erg sign
   ‘Jon did not sign the petition.’

(5.35)

a. Nork ez du eskaera sinatu?
   who.erg not aux petition sign
   ‘Who did not sign the petition?’

b. * Ez du eskaera nork sinatu?
   not aux petition who.erg sign
   ‘Who did not sign the petition?’

adapted from Etxepare & Ortiz de Urbina (2003: 507)

Sentence (5.34b) shows convincingly that occupying the preverbal focus position with the negative particle ez is obviously enough to satisfy a focus feature in Basque. Consequently, the reason for the ungrammaticality of (5.35b) must be due to something else. I assume it is due to an intervention effect, produced by the intervening negative particle ez between the Q-particle in the left periphery and the post-verbal wh-phrase. As stated in (5.9), intervention effects arise if the Q-particle is not the first focus-sensitive operator c-commanding the wh-word. In (5.35b), an intervening ‘offending operator’ prevents the c-command (Cable 2010: 99). Therefore it seems that the problem in these construction is not the

ii. * Nork liburu bat erosir du.
    who.erg book one bought aux
    ‘Intended meaning: Everybody bought a book.’
absence of the wh-fronting, but the intervention of the negative particle ez between the Q-particle in the left periphery and the wh-word in-situ.

In summary, we have seen that Basque exhibit a phonologically realized Q-particle in some interrogative constructions, that it possesses a very transparent derivation built on the wh-words, a non-interrogative use of wh-words, and clear intervention effects. I think it is legitimate to assume a independent, phonologically empty Q-particle for Basque in all wh-interrogatives in Basque.

In order to establish if Basque is a limited pied-piping language, I will have a closer look at pied-piping structures in Basque. In Section 4.1.3, I have presented their main properties. The language makes wide use of them, including cases in which pied-piping is the only possibility to front a wh-phrase (cf. (4.28) - (4.30)). We also recall the definition of a limited pied-piping language according to S. Cable. A languages in which every “limited pied-piping structure is a pied-piping structure where pied-piping past islands and pied-piping past lexical categories is not permitted” (Cable 2010: 144). English and Spanish are limited pied-piping languages, as seen in the previous sections. Basque, nevertheless, allows both pied-piping past lexical categories and past islands. In (5.36a) - (5.36c), a wh-phrase is pied-piped past a NP, a VP, and a relative clause, respectively.

(5.36) a. \([\text{NP Jonen etxeko zerk }] \text{izan ditu itoginak?} \]
   \( \text{Jon.en GEN house.REL what.ERG} \text{ have AUX leaks} \)
   ‘What of Jon’s house had got leaks?’

   Etxepare & Ortiz de Urbina (2003: 474)

b. \([\text{VP Nork idatzi ]-a da liburu hori?} \]
   \( \text{who.ERG written.ABS AUX book this} \)
   ‘By whom is this book written?’

   Heck (2008: 134)

c. \([\text{[Nork idatzi zu-en ]} \text{liburu ] irakurri du Pellok?} \]
   \( \text{who.ERG write AUX-en book read AUX Pello.ERG} \)
   ‘Peter has read the book that who wrote?’

   Ortiz de Urbina (1989: 249)

In these cases, especially the first sentence in (5.36a) is of major interest. Both Cable (2010) and Heck (2008) observe that pied-piping of sentences like ‘[Whose doctor] is Dave?’ is fine, whereas in ‘*[Doctor of whom] is Dave?’ it is not. Cable associates this behavior to the possibility to establish a Q/wh-agreement relation in the first but not in the second sentence. The reason being that in the latter clause a lexical phrase, namely an NP, blocks the agreement. In the former case, however, the wh-word has
moved to a left-peripheral position inside the pied-piped clause and is therefore again reachable for the Q-particle. The idea is schematized in (5.37a.) and (5.37b.).

\[(5.37) \quad \begin{align*}
(a.) & \quad \text{No Q/wh-agreement passed islands or lexical phrases} \\
(b.) & \quad \text{pied-piping of a subordinate possible, if wh-word is in SpecYP}
\end{align*}\]

\[\text{adapted from Cable (2010: 149/155)}\]

Cable (2010) and Heck (2008) have only looked at sentences like (5.38a) and (5.38b). In these two counterparts, it seems that pied-piping across CPs is not possible in Basque. Cable explains it along the line of (5.37), saying that the IP blocks the necessary Q-wh-agreement relation.

\[(5.38) \quad \begin{align*}
(a.) & \quad [c_p \text{ Nor}_1 [i_p \text{ joango dela } t_1 ] ]_2 \ \text{esan du } \ \text{Jonek } t_2? \\
& \quad \text{who } \text{go } \text{aux } \text{said aux } \text{Jon.erg} \\
& \quad \text{‘Who did John say will go’}
\end{align*}\]

\[b. \quad [c_p [i_p \text{ Joango dela nor }] ]_2 \ \text{esan du } \ \text{Jonek } t_2? \\
& \quad \text{go } \text{aux who } \text{said aux } \text{Jon.erg}
\]

\[\text{adapted from Cable (2010: 164)}\]

I claim, on the other hand, that the reason (5.38b) is ungrammatical lies in the structure of the subordinate clause before it is pied-piped to the left. If the wh-phrase were to stay in-situ in the subordinate even without pied-piping it would make the sentence ungrammatical, as shown in (5.39b). The wh-word must be fronted, even inside the subordinated clause (5.39a).

\[(5.39) \quad \begin{align*}
(a.) & \quad \text{Jonek esan al du zer erosi duen Maitek dendan?} \\
& \quad \text{Jon.erg said al aux what buy aux Maite.erg shop-in} \\
& \quad \text{‘What did Maite buy in the shop said John?’}
\end{align*}\]

\[b. \quad [c_p \text{ Jon.erg said al aux Maite.erg what shop-in buy aux }]
\]

\[\text{adapted from Cable (2010: 149/155)}\]
5.1 On Q-Particles – an introduction to the grammar of Q

Furthermore, having in mind sentences like (5.36a), we can clearly see that even with a lexical phrase intervening between the Q-particle and the wh-word, the sentence is still fine. I therefore hypothesize that Basque is not a limited pied-piping language. That is, the Q-particle and the wh-word do not have to establish an agreement relation. In this sense, Basque behaves like Japanese.

The fifth parameter is the behavior in multiple wh-questions. As already described in detail in Section 4.1.2.4, Basque multiple wh-structures are not as limited as their Spanish counterparts. They can exhibit multiple or single wh-fronting. In the latter case, Basque allows an almost free choice as to which wh-phrase is fronted, as example (5.40) illustrates.

(5.40)  
a. Nork ekarri du zer?  
    who.erg bring aux what  
    ‘Who brought what?’

b. Zer ekarri du nork?  
    what bring aux who.erg  
    ‘*What did who bring?’

(Etxepare & Ortiz de Urbina 2003: 499)

Concerning multiple fronting, Basque is subject to some restrictions already presented in Section 4.1.2.4. They concern linear ordering as well as the types of wh-phrases which can be fronted. However, in general it is a possibility, as can be seen in the example (5.41).

(5.41)  

a. Nori nork eraman dio liburua?  
    who.dat who.erg gave aux book  
    ‘Who gave the book to whom?’

b. Nork nori eraman dio liburua?  
    who.erg who.dat gave aux book

(Etxepare & Ortiz de Urbina (2003: 462/499)

According to Cable, Basque should therefore exhibit intervention effects, since it does not show superiority effects. The following example illustrates that this prediction is met.

(5.42)  

a. * Nork ez du zer eros?  
    who.erg not aux what bought  
    ‘Who didn’t buy what?’

b. Nork zer ez du eros?  
    who.erg what not aux bought
In a negative multiple question, the negative particle *ez* intervenes between the Q-particle in the left periphery and the *wh*-word *in-situ*. Thus, both *wh*-words have to move to the front in order to be first c-commanded by the Q-particle.

From this I conclude that Basque has only one Q-particle in multiple *wh*-questions, giving rise to intervention but not to superiority effects.

I will now return to the first two parameters: whether the Q-particle in Basque takes the XP\_wh as a complement or not, and if the Q-particle moves to the left periphery overtly or covertly. In order to answer the latter question, the most significant example is (5.34a) compared to (5.35b), here repeated as (5.43a) and (5.43b), respectively.

(5.43) a. *Ez du eskaera nork sinatu?*  
not AUX petition who.ERG sign  
‘Who did not sign the petition?’

b. *Ez du eskaera Jonek sinatu.*  
not AUX petition Jon.ERG sign  
‘It is not Jon who signed the petition.’

I argued that the reason for the ungrammaticality of (5.43a) is the intervention of the negative particle *ez*, a focus-sensitive operator, between the *wh*-word in its base position and the Q-particle occupying a C-head in the left periphery. A similar effect can be observed in multiple *wh*-fronting, as shown in the contrast between (5.42a) and (5.42b). If the Q-particle would not move overtly to a C-head, (5.43a) as well as (5.42a) should be grammatical.

The most decisive parameter however is the first, namely the relation between the Q-particle and the XP\_wh. As we have seen in (5.1) and (5.2), the former can either take the XP\_wh as a complement projecting a QP or adjoin to it. Thus, in the latter case, the Q-particle moves independently to the corresponding C-head whereas in the former it takes the XP\_wh along. Cable (2010) shows the independence of the Q-particle in languages like Japanese and Korean through its behavior and superficial position in the clause. The Q-particle moves to the end of the sentence, while the *wh*-phrase stays *in-situ*. In addition, the QPIC does not effect the structure, since Q does not project a QP that could interfere between functional phrases (Cable 2010: 90). The difficulty that arises for Basque is that we cannot apply the same arguments, because the Q-particle is not phonologically realized. Thus we cannot see where it moves. Hence, we have to find indirect evidences that the Q-particle in Basque behaves similar to its equivalent in Japanese and Korean.
5.1 ON Q-PARTICLES – AN INTRODUCTION TO THE GRAMMAR OF Q

A possible way to uncover these indirect evidences could be looking at \textit{wh}-indefinites. In Japanese as well as Korean, Q-particles are used with \textit{wh}-indefinites. Both types are illustrated in (5.44a) and (5.44b), a \textit{wh}-interrogative and a sentence with a \textit{wh}-indefinite, respectively. We see that the same particle -\textit{ka} is used in both constructions.

\begin{enumerate}
\item[(5.44) a.] John-ga nani-o kaimasita ka?
John.nom what.acc bought.polite Q
‘What did John buy?’
\item[(5.44) b.] Taroo-wa doko-ka-e itta.
Taro.top where-Q-to went
‘Taro went somewhere.’
\end{enumerate}

If we recall the QPIC in (5.10), sentence (5.44b) should be ungrammatical if Japanese was an Q-projection language, because the QP would intervene between the two functional phrases (PP) and XP\textsubscript{wh} (DP). However, Japanese is a Q-adjunction language and therefore Q does not project its own QP. This structure is represented in (5.45).

\begin{enumerate}
\item[(5.45)] Structures where Q appears between P and its complement in Japanese
\end{enumerate}

\begin{center}
\begin{tikzpicture}

\node {PP} child {node {DP} child {node {DP} child {node {Q} child {node ...\textit{wh}-word...}}}};
\end{tikzpicture}
\end{center}

The same holds for sentences in which Q can appear between the possessor and the possessed in Japanese and Korean, illustrated in (5.46). Both kind of structures are not possible in a Q-projection language like Spanish or English, as already shown in Section 5.1 and 5.1.1.

\begin{enumerate}
\item[(5.46)] Taroo-wa [ dare-ka-no oniisan ]-ni atta.
Taro.top who-Q-gen brother-DAT met
‘Taro met someone’s older brother.’
\end{enumerate}
In Table 5.1, we saw the transparent derivational structure for wh-indefinites in Basque. Assuming a similar structure for all the derivations, the following sentences can give evidences for analogous behavior in Basque and Japanese.

\[(5.47)\]

\[
a. \text{Jon erik non-bait-era etorri da.} \\
\text{Jon.erg where-bait-to go aux} \\
\text{‘John went somewhere’}
\]

\[
b. \text{Jon erik nor-bait-en etxea bisitatu du.} \\
\text{Jon.erg who-bait.gen visit aux} \\
\text{‘John visited somebody’s house.’}
\]

In (5.47a), the affixal particle -bait, which contributes existential quantification, intervenes between the postposition -(e)ra ‘to’ and its complement. In (5.47b), it appears between the possessor and the possessed. In these kind of structures therefore, Basque resembles Japanese and Korean, both Q-adjunction languages.

Concluding this section, we have now seen some evidence for an independent Q-particle in Basque. The sporadic phonological realization of Q-particles, the transparent paradigm of wh-words in combination with different particles, and the distributive reading of wh-words allow us to assume an independent and phonologically empty Q-particle in Basque. It is not inherently fused with the wh-word, as in Spanish or English, but operates independently. The restrictions in combination with the negative particle ez in single as well as in multiple wh-questions show convincingly that the Q-particle moves overtly to a C-head. Finally, based on the behavior of the particle -bait we can conclude that the Q-particle only adjoins to the XP\text{wh}, allowing the former to appear, for example, between P and its complement. Thus, one can conclude that Basque is a Q-adjunction language.

If we accept this conclusion, the remaining question that has to be answered is clear. If the wh-fronting in Basque is not due to QP-movement, then why does the wh-phrase always move to the preverbal position? I shall endeavor to answer this question in the last section of this chapter.

### 5.2 On an alternative interpretation of Basque wh-movement

As I have already mentioned in Section 4.1, the preverbal position, called galdegaia in traditional Basque grammar, must always be occupied. We have seen that Basque offers several possibilities to avoid an empty preverbal position. In an all-focus sentence, the whole VP/VP moves to this position, as detailed in (4.46). In negative or positively emphasized clauses, the negative particle ez or the affirmative marker
5.2 On an alternative interpretation of Basque wh-movement

*ba*- have to appear in the preverbal position. Finally, *wh*-phrases or focalized constituents have to occupy the *galdegaia*.

Having seen the arguments for an independent Q-particle in Basque and the clear parallels in the behavior of *wh*-phrases and focus constructions, my hypothesis is the following: *wh*-phrases in Basque do not move due to the Q-movement, but in order to satisfy the focus requirement in the preverbal position. If we assume that the responsible feature for the *wh*-fronting is an independent [+foc], a *wh*/Q-agreement is not necessary, nor does the Q-particle need to take XP<sub>wh</sub> as complement in order to display *wh*-fronting. Instead, the former can be adjoint to the latter. The phonologically unrealized Q-particle moves to a higher C-head, whereas XP<sub>wh</sub> moves to the preverbal focus position to satisfy an independent [foc]-feature. This assumption accounts for many of the properties we have observed. XP<sub>wh</sub> moves to the preverbal position, so no other constituents can intervene between the two elements. In multiple *wh*-questions, fronting one *wh*-phrase is enough to satisfy the [foc]-feature. In negative sentences, the reason that *wh*-phrases has to move higher up than the negative particle can be explained by the otherwise arising intervention effects. In order to prevent the negative particle interfering between Q and the *wh*-phrase, the latter has to move past the negative particle. To elucidate my hypothesis, I shall now break it down in more detail.

According to my hypothesis, the structure in (5.48) would be the starting point for the derivation. YP here stands for a obligatorily occupied preverbal phrase. I will detail its exact nature further down.

(5.48)

---

The preverbal position probes for an element capable of satisfying the [-foc]-feature in its c-commanding area. Either a specific element in vP/VP bears such a feature or, in an all-focus sentence, the whole vP. In either case, the feature bearing element moves. In the presence of a negative particle, as detailed in Section 4.2.2, I assume that NegP is already merged above IP and therefore can do the job (among
5.2 On an alternative interpretation of Basque wh-movement

others, cf. Haddican 2004, Ortiz de Urbina 1989). In a nutshell, Basque sentences may never be pure IPs, but always need some kind of left peripheral structure dominating IP.

In Chapter 4.2.2, I concluded that the preverbal focus position is SpecFinP, associating it with Haddican’s (2004) PolP, a phrase that has always to be occupied by some element. That would mean, the structure in (5.48) should be modified as in (5.49).

(5.49)

If the wh-phrase moves to SpecFinP, then the Q-particle has to move to an even higher head, because it must c-command the wh-phrase. Adapting a reduced left periphery with only ForceP-(DS)-FinP, ForceP offers itself as a natural landing site. The optional Discourse Shell (DS), as detailed in Chapter 3 and Chapter 4, can host topics and other scrambled material. Bringing it all together, I propose a structure as in (5.50).

(5.50)
Following the structure in (5.50), the corresponding syntactical tree for (4.47), here repeated as (5.51a), would be (5.51b).

\[(5.51)\]
\[
\begin{align*}
a. & \quad \text{Non erosi du Peiok oparia?} \\
& \quad \text{where bought aux Peio.erg present} \\
& \quad \text{‘Where did Peio buy the present?’}
\end{align*}
\]

b. 

\[
\text{The structure in (5.50) can also explain multiple wh-fronting in multiple wh-questions. We have seen that Basque multiple wh-phrases are not subject to superiority effects. I therefore postulate that only one single wh-phrase moves to SpecFinP in order to satisfy the focus-feature of Fin\textsuperscript{0}, while the other wh-phrases, if fronted at all, move to a DS. As we know, DSs are recursive and interchangeable and the absence of superiority effects show clearly that the linear order of the wh-phrases does not matter. That means, a sentence with multiple wh-fronting like (5.52a) would have the syntactic structure as in (5.52c). The structure for (5.52b) is similar, just with the two wh-phrases having switched places. In any}
\]

121
case, not all *wh*-phrases have to move. All but one can stay *in-situ*. This only turns ungrammatical if the multiple *wh*-question is negative. Then all *wh*-phrases have to move.

(5.52) a. Nork nori ematen dio liburua?
   who.erg who.dat give aux book
   ‘Who gives the book to whom?’

b. Nori nork ematen dio liburua?
   who.dat who.erg give aux book

c. ForceP
   Force^0
   Q_j
   DS
   Nor_i
   FinP
   nork_i
   Fin^0 [+foc]
   [ematen dio]_j
   t_j
   vP
   DP
   t_q
   DP
   t_i
   IP
   v
   VP
   t_j
   liburua
   t_j

As just mentioned, in negative sentences, all *wh*-phrases have to move. The same accounts for a single *wh*-question. That is, the *wh*-word still has to move to SpecFinP, though the reason is different. It must move in order to prevent that the negative particle *ez*, instead of the Q-particle, is the first focus-sensitive operator c-commanding the *wh*-word. (5.53b) shows the structure for a sentence like (5.35a), here repeated as (5.53a).

(5.53) a. Nork ez du eskaera sinatu?
   who.erg not aux petition sign
   ‘Who did not sign the petition?’
In summary, I have presented evidence and arguments for a different approach to Basque wh-movement. Basque wh-phrases move in two possible configurations to the left periphery. First, to satisfy a focus-feature hosted in FinP to guaranty that the preverbal focus position in Basque is occupied. Second, in the case that the negative particle ez already occupies that position and therefore satisfies the preverbal-focus feature, the wh-phrase has to move in order to prevent intervention effects between the negative particle and the wh-phrase.

An interesting consequence of this approach is that Basque would actually be a wh-in-situ-language, since the wh-phrase do not move because of a Q-feature or anything directly related to the wh-question. It moves out of a different motivation, namely a focus-feature that has to be satisfy. The wh-phrase is normally just the most suitable constituent to do the job. This should not come as a surprise, bearing in mind that most of the verb-final languages in the world have a wh-in-situ-structure, rather than wh-fronting (see among others, Aldai 2011, Dryer 1992, 2005, Duguine & Irurtzun 2014, Russel 1978).
In the previous three chapters, I have elaborated and developed the theoretical background for interrogative constructions in Basque and Spanish. I have compared the two systems and shown that the superficial resemblance as well as the differences, are derived from two different underlying processes. However, as I have mentioned already in the introduction, I will argue that despite these differences, Basque-Spanish bilingual speakers are influenced by this superficial similarity. As I will outline below, my hypothesis is that bilingual speakers adopt Basque’s overall stricter adjacency rules in the complex wh-constructions in Spanish. In other words, although monolingual speakers accept an intervention, for example, of a lexical subject between the heavy wh-phrase and the verb in Spanish, Basque-Spanish bilingual speakers do not, or at least do so to a lesser extent. Before I will present and discuss the outcome of my empirical research in Chapter 7, I want to give a short introduction to contact-induced change and the driving force for such contact phenomena: bilingual language acquisition or L2 acquisition. I will present an approach to the notion of contact-induced change in Section 6.1, mainly relying on work by Heine & Kuteva (2010) and Thomason (2010). In Section 6.2, I will discuss support for my hypothesis in the main approaches to bilingual language acquisition.

6.1 On language contact

If we have a look at a map of the world’s languages, Thomason’s following statement is easy to believe: “Language contact is the norm, not the exception. We would have a right to be astonished if we found any language whose speakers had successfully avoided contact with all other languages for periods longer than one or two hundred years” (Thomason 2001: 10). In other words: “language contact is everywhere” (Thomason 2001: 8). That means, bi- or multilingualism is most likely the norm and not the exception. This is probably the reason why language contact has become more attractive, especially
over the last three decades, when searching for the reasons for language change (Thomason 2010: 31). These sources for an external change can for example be grammatical calquing, that means copying of grammatical structures from one language to another, or replica grammaticalization. Apart from these external reason, language change can also be driven by internal reasons, like simplification. If the source of external change is contact, it is called contact-induced. The main difficulty now is to separate the two sources and evaluate if an observed change is internally or externally driven. An additional difficulty is the fact that they most likely go often hand in hand (Heine & Kuteva 2010). Thomason (2010: 32) contrasts internal and external change defining that “[c]ontact is a source of linguistic change if it is less likely that a particular change would have happened outside a specific contact situation.” That means, before trying to find external influences for a language change, internal reasons should always be considered first. Thomason (2010) defines six main conditions for language change and evaluates them regarding their role in internal and external language change. In the following sections, I will present these conditions and concepts and apply them to the Basque-Spanish contact situation.

6.1.1. What is language contact and how does it happen?

Language contact is everywhere, especially if we count dialectal-normative language situations as well. Nevertheless, Thomason (2010) points out a crucial prerequisite for speaking of a language contact situation.

The mere juxtaposition of two speakers of different languages, or two texts in different languages, is too trivial to count: unless the speakers or the texts interact in some way, there can be no transfer of linguistic features in either direction. Only when there is some interaction does the possibility of a contact explanation for synchronic variation or diachronic change arise.

Thomason (2010: 31-32)

Language change happens because of the daily use and interaction between two or more languages embodied by their speakers (Matras 2010, Thomason 2010). Since Thomason & Kaufman’s (1988) seminal work, many scholars have defended the claim that “language contact may have a wide range of implications for the languages involved, and it may affect virtually any component of language structure” (Heine & Kuteva 2010: 86). That is, phonological or morphological elements as well as syntactic or semantical structures can be transferred from the model to the replica language (Hickey 2010: 14).

Thomason (2001) singles out six main predictors that have to be considered and can influence the outcome of language contact depending on their degree of realization. They are divided into three social and three linguistic predictors. The first three are intensity of contact, presence or absence of imperfect

---

55 Model vs. replica language is the terminology used by Heine & Kuteva. Others call them source/donor and receiving/target language, respectively (Heine & Kuteva 2010: 86). I will use the first terms, model vs. replica language.
6.1 On language contact

learners, and speaker’s attitude. The linguistic predictors are universal markedness of the transferred feature, the overall integration into the linguistic system and the typological distance between the languages involved. I will briefly present the six conditions here.

Out of the three social predictors, the most important factor for transfer is the intensity of contact. “The intensity of contact is typically derived from things like the duration of contact and the level of bilingualism in the receiving-language speech community” Thomason (2010: 37). Thomason & Kaufman (1988) establish a so-called borrowing scale, indicating the possibly transferred items and structures depending on the intensity of the contact. It is illustrated in Table 6.1.

Table 6.1: Borrowing scale according to Thomason & Kaufman (1988: chap. 4)

<table>
<thead>
<tr>
<th>Casual Contact</th>
<th>Category 1: content words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Category 2: function words, minor phonological features, lexical semantic features</td>
</tr>
<tr>
<td></td>
<td>Category 3: adpositions, derivational suffixes, phonemes</td>
</tr>
<tr>
<td></td>
<td>Category 4: word order, distinctive features in phonology, inflectional morphology</td>
</tr>
<tr>
<td>Intense contact</td>
<td>Category 5: significant typological disruption, phonetic changes</td>
</tr>
</tbody>
</table>

That means, the longer and more intense a contact situation is, the more likely it is that structural features get transferred. An important secondary factor is the level of bilingualism in the speech community. The higher the rate and fluency in both languages, the easier structural features are transferred (Thomason 2010: 37).

Thomason’s second predictor is the number of imperfect learners. Imperfect learners acquire the replica language later in life, usually meaning after puberty, and never speak it with native-like competence. The result is often that they copy phonological, syntactical, and lexical items, structures, or features to the replica language. If a whole community changes to the replica language, we speak of language shift. We will see further down that for language changes in Basque imperfect learning could play a role, in contrast to changes that affect Spanish within Basque-speaking areas.

The last social predictor in Thomason’s (2001) approach is the attitude of the speakers towards the languages spoken in the contact scenario. As Thomason notes, this is “a very vague notion, but it is difficult to make it precise – and equally difficult, unfortunately, to prove that it has affected the course of language history” (Thomason 2010: 38). Despite the difficulty to determine the speaker’s attitude, Thomason names a strong indicator for a positive attitude towards a language. She states that “[t]he clearest examples of speakers’ attitudes as a cause of change therefore come from cases of deliberate change” Thomason (2010: 38). I will explain later how the Basque speaking community
6.1 On language contact

makes a deliberate effort to spread Basque, increase the number of speakers, and create incentives to learn the language.

Depending on the degree of realization of the three social predictors, the likelihood of structural borrowing increases. A long and intense contact of two languages with many bilingual speakers, including imperfect learners, combined with high prestige of the replica language certainly would be a good basis for contact-induced changes.

I will now turn to the three linguistic predictors universal markedness, integration in the linguistic system, and typological distance. I will start with the last predictor, since it is the most obvious one to evaluate and relatively independent regarding particular phenomena. For the other two, I will first have to give more details about the phenomena I want to analyze before regarding the possibility of contact-induced influence.

Typological distance can "prove a barrier to contact-induced change. Not an absolute barrier: under circumstances of intense contact, any linguistic feature can be transferred to any other language" (Thomason 2010: 41). Basque and Spanish are two typologically very different languages. Therefore, the intensity of contact is again very important as indicated with the borrowing scale in (6.1). Interestingly, in the case of wh-phrases, Basque and Spanish behave very similarly, at least on the surface. I will argue later that it is exactly this superficial similarity that allows learners to reinterpret the Spanish structures. In other words, although Basque and Spanish are typologically very different, in this particular phenomenon, they behave surprisingly similar.

Universal markedness "is ultimately connected with ease of learning. Universally marked features are believed to be those that are harder to learn; unmarked features should be easier to learn" (Thomason 2010: 43). Therefore, the idea is that universally unmarked features are more easily transferred to the replica language than marked features. Concerning wh-interrogatives, the fronting of the wh-phrase seems to be the marked version, compared to wh-in-situ-structures. This is probably due to an overall principle of computational economy (movement is less costly), that is, avoidance of syntactic movement (cf. Duguine & Irurtzun 2014: e25-e26, and references cited therein). Thomason (2010: 43) says that universal markedness "plays an important role in shift-induced interference […]. But in contact-induced changes that do not involve imperfect learning, markedness is likely to be of considerably less importance, and may not play any role at all". Since the number of imperfect learners of Spanish is negligible, I follow Thomason’s argument and assume no important impact of the markedness or unmarkedness of the features under examination.
The third predictor, overall integration, follows a similar notion. Thomason (2010: 45) says that “[c]ontact-induced changes are much more common in loosely integrated linguistic subsystem than in subsystems which - like the inflectional morphology - are characterized by sets of interconnected forms organized into paradigms”. In other words, highly integrated features like inflectional morphology are very unlikely to be affected by contact-induced change. Especially compared to less integrated features like word order or informational structure (Gómez Seibane 2012). However, if the contact is intense and prolonged, transfer of highly integrated features is possible. Although the basic word order is different, both languages use word order permutations in order to express informational content, like focus. These similar means to express the informational content of a sentence through word order variation could support a possible contact-induced change.

In summary, large typological distance, universal markedness and high integration would more likely prevent contact-induced change.

In the next section, I will describe the sociolinguistic situation in the Basque speaking area and in Section 6.1.3, give several examples of contact-induced changes in both languages.

6.1.2. The socio-linguistic situation in the Basque speaking area

This section will give a brief overview of the sociolinguistic situation in the Basque speaking area. Since 1991, the governments of the Comunidad Autónoma Vasca and the Government of Navarra together with the Office Public de la Langue Basque have already carried out six large surveys, in 5-year intervals, on the sociolinguistic situation in the Basque speaking area. I will not present all of the data, but just some core statistics and numbers which concern the above described social predictors. The entire survey can be accessed on the official government website. All data I present in the following section is taken from the VI Encuesta Sociolingüistica 2006 if not indicated otherwise. The sorting parameter for the data is either the origin of the speakers or the age. For the purpose of the survey, origin is separated into the Comunidad Autónoma de Euskadi (CAE), Navarra, and the Northern Basque country on the French side of the border. The CAE is further subdivided into its three provinces Araba, Bizkaia, and Gipuzkoa, while the French Basque Country is subdivided into the three provinces Zuberoa, Lapurdi and Low Navarre. The age groups are 16-24, 25-34, 35-49, 50-64, and ≥ 65.

56 For the northern part of the Basque speaking area, data is only available from 1996 on.
57 http://www.euskadi.eus/gobierno-vasco/politica-linguistica
Due to the regularity of surveys, we have a very good picture of the sociolinguistic development over the last 25 years. Probably the most prominent statistic is the overall increase of Basque speakers from 528,500 (22.3%) in 1991 to 751,500 (28.4%) in 2016. The number of passive Basque speakers has also increased, from 7.7% to 16.4% in the same time frame. Although the overall percentage of speakers has increased notably, the actual number of speakers in the northern Basque country has decreased. We must bear in mind that the legal position of Basque is a very different one in Spain and France. In Spain, it is a co-official language, whereas the central government in France hardly acknowledges that there are different languages spoken within its borders. This is why, in the southern Basque country, the language is much more promoted and therefore more widespread than in the northern part. However, since I investigated a possible influence on Spanish, I will concentrate on the southern part of the Basque speaking area.

In the previous section, I have introduced the three social predictors by Thomason (2001), namely intensity of contact, number of imperfect learners, and attitude towards the languages. I will review these three predictors for the Basque-Spanish language contact scenario now in more detail supported by numbers and statistics from the *VI Encuesta Sociolingüística 2006*.

**Intensity** If we look at the Basque-Spanish contact situation, we can definitely confirm a long and intense contact. If we consider the Latin and the Proto-Romance languages, especially early Castilian, the period of contact expands over more than 2000 years. Nowadays, every Basque speaker in Spain, no matter their proficiency in Basque, is perfectly fluent in Spanish, too. The high number of bilingual speakers can be traced back to the *Reconquista*, as mentioned already in Chapter 2. The Spanish influence already increased noticeable during the industrialization in the 19th century. However, it was during the Franco regime that Spanish was established as the only official language and every Basque speaker had to learn it. In any case, we are facing at least 100 years of widespread bilingualism, preceded by nearly 2000 years of partial bilingualism (Etxebarría Arostegui 2007: 39, Schlaak 2014: chap.3.3.2).\(^{58}\) Hence, both criteria for intensity named by Thomason (2010: 37), namely duration of contact and level of bilingualism are applicable in this scenario and increase the likelihood of contact-induced language changes. We will see in Section 6.1.3 that we can indeed find many traces of this long and intense contact in both languages.

\(^{58}\) Obviously, this counts only for Basque speakers. Especially since the 19th century, many people from outside the Basque speaking area moved there in the course of the heavy industrialization. These newcomers did not speak Basque at all.
6.1 On language contact

**Imperfect Learning**  This factor is of special interest in the case of Basque. In the course of the progressive language policy implemented in the 1980s, many of today’s Basque speakers learned the language as adults. They are called *euskaldun berri* ‘new Basques’. In the age group of 16-34 years, around half of the speakers did not acquire Basque at home, but in school. Of course, we have to differentiate these speakers in more detail, depending on when exactly they learned or acquired Basque. I will come back to that point in Section 6.2. Nevertheless, the group of *euskaldun berri* becomes more and more relevant in the Basque speaking community. Rodríguez-Ordóñez (2017) described their possible influence in a recent study on DOM in Basque. However, when looking into possible influences on today’s Spanish spoken in the Basque country, imperfect learners of Spanish play almost no role. We can assume, without hesitation, that a huge majority of bilingual speakers have native competence in Spanish. Only in the group of older speakers, that is 65 years or older, do we find a significant number of people (42%) who express themselves better in Basque, whereas 24% are equally at ease in both languages, and only 33.8% prefer Spanish. As for the younger speakers, 20-25% prefer Basque, 19.5%-35.9% express themselves with equal ease in both languages, and 56.7%-39.8% do better in Spanish/French (Gobierno Vasco, Gobierno de Navarra & Office Public de la Langue Basque 2016: 10).

**Attitude towards Basque/Spanish**  As already mentioned, it is very difficult to determine the attitude of speakers toward their respective languages. Nevertheless, the Basque example allows us to draw some conclusions. Thomason (2010: 38) states that “[t]he clearest examples of speaker’s attitudes as a cause of change therefore come from cases of deliberate change”. Since the 1960s with the first release of *Euskara Batua* and the active language politics since the 1980s, language planing can be rated as a very strong indicator of the positive attitude towards Basque. For example, the high number of children attending Ikastolas, Basque speaking schools, even if their parents do not speak Basque or come from outside the Basque speaking area. Nowadays, these make up about 50% of the Basque speakers in the group of 16-34 year olds. The effort of the Euskaltzaindia to publish new dictionaries with technical and specialized vocabulary confirms a widespread political and social interest in the advancement of the Basque language. This is an indication of the high esteem the language is held in, at least in the Basque speaking areas of Spain.

The Basque-Spanish contact scenario is a perfect constellation to investigate contact-induced language change. It offers a longterm and intense contact with a high level of bilingualism as well as speakers who are conscious of the languages and have an overall positive attitude towards Basque and its promotion (Gobierno Vasco, Gobierno de Navarra & Office Public de la Langue Basque 2016: 31-
Furthermore, we find a significant number of imperfect Basque learners, but very few imperfect Spanish learners. That means that the most important predictors for contact-induced change all point towards possible structural transfer. The likelihood of a contact-induced change increases if there is already evidence for other changes in the same contact scenario. In the next section, several examples for both diachronic and synchronic changes in the Basque-Spanish contact situation will show that both languages are in constant and productive exchange.

### 6.1.3. Contact phenomena in Basque and Spanish in the Basque speaking area

Over the last decades, numerous theoretical and empirical studies about Basque-Romance language contact have been published (Camus Bergareche & Seibane 2012). From Baroja’s (1945) *Materiales para una historia de la lengua vasca en su relación con la latina*, via various works by Mitxelena (1959, 1964, 1966, 1974, 1977a, 1982b, 1984), to more recent work by Duguine & Irurtzun (2014), Etchebarne (2012), Etxebarria Arostegui (2007), and Rodríguez-Ordóñez (2017), to name only a few.

As can be expected in a contact situation of such long duration, the biggest influence on Basque can be seen in the lexicon. It is estimated that the proportion of words derived from Latin-Romance source outnumber the native Basques by now (Trask 1997: 249). We can also find some influences of Basque in the Spanish vocabulary, the most prominent example being *Sp. izquierda* < *Bq. ezker* ‘left’. However, as Trask (1997: chap. 6.13) argues convincingly, these influences on the Spanish lexicon are much rarer than has often been claimed. I will not go into detail about the vocabulary here.\(^59\) My interest lies in possible structural or grammatical influences on the two languages.

The important point is that we can already find structural influences on Basque at a relatively early stage of contact. Etchebarne (2012) details three major changes in the structure of Basque which were very likely induced by contact with Romance languages: the appearance of the direct article, the spreading of periphrastic verb structures, and diversification of verbal plural marking. Etchebarne discusses the hypothesis whether these language changes could also have been internally driven. All three characteristics are not uncommon in the languages of the world. It would not be too surprising if Basque developed them by itself, without the influence of the Romance languages. Nevertheless, Etchebarne argues convincingly that the time and way of the change strongly favor a Romance influence. I will briefly summarize his argumentation for the emergence of the articles.\(^60\)

---

\(^59\) The reader is referred to the work by Baroja (1945), Mitxelena (1977a, 1995), and Trask (1997, 1998), among others.

\(^60\) For a full discussion of all three phenomena, the reader is referred to Etchebarne (2012).
6.1 On language contact

Basque epitaphs from the first to the third century, found in Aquitaine, do not bear any trace of definite articles. This provides a strong evidence that the language did not have them back then. In the case of Latin, we know for certain that definite articles do not exist. That means, it was also an innovation in the Romance languages. According to Carlier & De Mulder (2007), they arose due to four main reasons: (i) the disappearance of the Latin case system,\(^6\) (ii) the change from a synthetic to an analytic language,\(^6\) (iii) the alteration of the sentence constituent order, and (iv) the necessity to mark the theme (Etchebarne 2012: 394). Basque did not undergo the same grammatical changes as the Romance languages in regard to these four steps. Hence, according to Etchebarne, it is highly possible that the development of the definite article was initiated and influenced by the innovation in the Romance languages surrounding it. He argues in a similar fashion for the emerging of periphrastic verb structures and diversification of verbal plural marking.

The Basque influence on the development of early Spanish is still a subject of discussion. As I have detailed in Chapter 2, Castilian emerges originally bordering the Basque speaking area at the time. The two areas probably overlapped to some extent, as can be seen by the notes found in the monasteries in that area (Trask 1997: 42-43). Because of this closeness, some linguists accredited certain phonological characteristics of Spanish to a Basque influence. Among them, the partial phonological change from Latin word-initial /f/ to Castilian Spanish /h/ and the so-called canonical vowel system of only five vowels. The simple vowel system is still a core property of present-day Basque and the absence of the phonem /f/ is verified for that time period (Mitxelena 1977b). A very similar development in Gascon, the language bordering Basque on the east side of its territory, could be seen as support for these hypotheses. In Gascon, the phonological change from word-initial /f/ to /h/ is complete. According to Echenique Elizondo (2005: 71-72), the divergence of the vowel systems is at least highly probable. If a language with a rich vowel system, like Latin, comes into contact with a canonical one, it is precisely the latter’s simplicity that can be the base for a simplification of the vowel system. Nevertheless, the Basque influence on Spanish in its creation was limited.

The investigation of diachronic contact-induced language change is always difficult. In the case of Basque, the scarceness of written testimony before the 16th century makes it even more so. Fortunately, investigations of synchronic contact-induced changes have increased in the last decades, especially of Basque influences on Spanish. Synchronic influences of Spanish on Basque, however, are not so extensively studied yet.

---

\(^6\) See for example Käser (2014) for an exhaustive description on the change from the rich Latin case system to a practically non-existing one in most of the RL (except in Rumanian and some remnants in the pronominal system of other RL).

\(^6\) Latin is seen as a synthetic language with a high morpheme-per-word ratio, whereas the Romance languages have developed a much lower morpheme-per-word ratio.
Interestingly, it seems that the Spanish spoken in the Basque country has had its known peculiarities already in the Siglo de Oro. In theaters, the typical Spanish Bizkaian dialect was used as humorous speech and to depict simple people from the country (Camus Bergareche & Seibane 2012: 5). Nowadays, we can still find some distinctive traits in the speech of the people living there. How much it is due to Basque influence cannot always be determined with certainty. We can find lexical borrowings, phonological and morphological influences, and syntactic variation. I will not go into more detail about lexical, morphological, or phonological influences except for one phenomena. For a full description of these traits, the reader is referred to Camus Bergareche & Seibane (2012), Etxebarria Arostegui (2007), Etxebarria (2004), Oñederra (2002, 2004), and Urrutia Cárdenas (1995), and references cited there.

The only phonological influence I want to address here is the phenomenon called yeísmo. Yeísmo is the “neutralization of the lateral palatal /ʎ/ and the central [palatal] /ʝ/ […] in favor of the non-lateral pronunciation” (Etxebarria Arostegui 2007: 42). Yeísmo emerged in Spanish in the second half of the 17th century and started to spread slowly throughout Spain and Hispanoamerica (Alonso 1961: 159). However, in the Basque speaking area, the neutralization spread much slower and is still not ubiquitous. This “slower de-lateralization of the palatal could, in principle, be attributed to the influence of the Basque language, in which laterals preceded by /i/ (vowel or semivowel) palatalize. This is a productive phonological process in many geographical varieties of [the] Basque language” (Etxebarria Arostegui 2007: 42). Although the neutralization process is definitely gaining ground, especially among younger speakers, the older generation still maintains the distinction (Camus Bergareche & Seibane 2012: 11). I quote this example here because it shows a further important trait of language contact, namely delay or acceleration. In a contact situation where both languages display certain contrastive features (phonological or morpho-syntactic), neutralizations, disappearance, and emergence of phenomena can be delayed or favored/accelerated. Heine & Kuteva (2006: ch. 2, 2010: 94-97) discuss several examples for acceleration and delay with language contact as the possible source. As an example for Spanish acting as a propelling force on Basque structures, they quote Trask (1998: 320). He describes a grammaticalization process that denotes the change in the use of question words from a pure interrogative marker to the markers of relative clauses. Heine & Kuteva and Trask point out that this process appears to be largely restricted to the Indo-European languages. “[I]f attested elsewhere, this is likely to be in languages that have had some history of contact with European languages” (Heine & Kuteva 2010: 95). An example for the opposite case, language contact as a delaying force, seems to be the slower spreading of yeísmo in Spanish in the area where it is in contact with Basque.
Two very similar studies on word order and information structure were carried out by Gómez Seibane (2012) and Irurtzun (2018). Both authors describe the often encountered OV word order in Basque Spanish. That is, sentences like (6.1a), (6.1b) and (6.1c) are not unusual to hear.

(6.1) 

a. Tres sobresalientes tiene Juan.  
three outstanding has Juan  
‘Juan has three outstandings’

b. Con el mogollón, la policía vino.  
with the crowd, the police came  
‘The police came with the crowd.’

c. Una sidra me ponen?  
one cider Cl.DAT put  
‘Would you serve me a cider?’

Irurtzun (2018)

Although both constructions occur only in certain contexts and have very special pragmatic and discourse functions, they sound questionnable or simply ungrammatical for speakers coming from other regions of Spain (Irurtzun 2018). Especially in the example (6.1c), “it seems that the fronting strategy employed in polite commands in the Spanish of the Basque Country may derive from transfer [...] from Basque” (Irurtzun 2018). These observations would be in line with the comment in Gómez Seibane (2012: 15), citing Lozano (2006), where she says that “la interfaz sintáctico-pragmática y, en especial, las propiedades relacionadas con tópicos y focos resultan particularmente sensibles a la transferencia interlingüística (cross-linguistic transfer)”63 (cf. also Serratrice, Sorace & Paoli 2004).

Not only word order and information structure are subject to contact-induced change. Etxebarria Arostegui (2007), Oñederra (2004), and Urrutia Cardenas (2002), among others, give more examples of contact phenomena in Spanish. The most known phenomenon is the use of clitics, which I want to present briefly. There are three main phenomena concerning object clitic pronouns. The first is the widespread use of the indirect clitic pronoun _le_ instead of the direct _lo/la_ for human antecedence and is called _leísmo_. Although _leísmo_ is not at all restricted to the Basque speaking area, it is especially strong there, meaning that masculine _lo/los_ as well as feminine _la/las_ pronouns are replaced by _le/les_ (Camus Bergareche & Seibane 2012: 13, Rodríguez-Ordóñez 2017: 327). The second is the strong tendency to elide the direct object clitic pronoun _lo/la_ in the cases of non-animated direct objects. Sentences like ¿Me compras? ‘Do you buy it?’ instead of ¿Me lo compras? are often used. This elision is probably

---

63 The syntax-pragmatic interface and, especially, the characteristics related to topics and foci result especially sensible to cross-linguistic transfer.
derived from the Basque’s three-way pro-drop property which allows the speaker to drop not just the subject as in Spanish, but also the direct and indirect object. The third phenomenon is the use of the indirect object clitic pronoun le/les in order to reduplicate definite or specific animated direct objects in cases different from ‘standard’ Spanish (Camus Bergareche & Seibane 2012: 13-14, Landa 2000). The following example demonstrates such a case.

(6.2)  Le estoy buscando a Juan, pero no le encuentro.
   Cl.1o am  looking  for Juan  but  not Cl.1o find.1SG
   ‘I’m looking for Juan but can’t find him.’

Landa (1993: ex. 31)

Landa (2000: 209) says that le in these contexts is closer to agreement morphology than to gender marking. A possible reason for this doubling could be found in the Basque agreement morphology. In other words, the clitic le is perceived and interpreted similar to the verbal absolutive agreement in Basque, that is to say, as a DO agreement marker (Camus Bergareche & Seibane 2012, Landa 1993, 2000). I will return to this last phenomena explicitly in Chapter 8. All these observations show that Basque exerts an influence on Spanish in the Basque-speaking area.

The synchronic influences on Basque, on the other hand, are much less examined. For the purpose here, I want to focus on a special feature, investigated in Fernández & Rezac (2016) and Rodríguez-Ordóñez (2017), among others, namely Differential Object Marking DOM. It refers to special case marking of a subset of direct objects. In Spanish, if the DO is [+human], the marker a is used to introduce the DO, as seen in (6.3a). The DOM marker a is never used with a [-human] direct object (6.3b).

(6.3)  a.  Esconde *(a)  Juan.
      hide  dom Juan
      ‘He/She hides Juan/the book’

      b.  Esconde (*al)  libro.
          hide  dom the book
          ‘He/She hides the book’

Rodríguez-Ordóñez (2017: 326-327)

Basque is traditionally a non-DOM language. Fernández & Rezac (2016) and Rodríguez-Ordóñez (2017), however, have observed that in some Basque dialects, for example in Gernika Basque, the speakers also mark a [+human] DO with the IO verb morphology. That means, these speakers use the dative suffix -ri and the corresponding auxiliary instead of the non-marked absolutive and the auxiliary for
transitive verbs. As the examples (6.4) and (6.5) show, the dative marking is facultative with a [+human] DO, but ungrammatical for non-human DO, respectively.

(6.4)  
a.  Ba Jone-ri1 atrapatz-o1.  
  So JON.DAT catch.3SG.DAT  
  ‘So it caught Jon.’  
b.  Ba Jon atrapa de-u.  
  SO JON.ABS catch AUX.3SG  
  ‘So it caught Jon.’

(6.5)  
a.  Nik etxe-i1 ikusi d1o-t  
  I.ERG house.ABS see AUX.3SG.ABS-1SG.DAT  
  ‘I have seen the house.’  
b.  *Nik etxie-ri1 ikusi do-ts1-ta-t  
  I.ERG house.DAT see AUX.3SG.ABS-3SG.DAT-1SG.ERG  
  ‘I have seen the house’

Rodríguez-Ordóñez (2017: 329)

The analysis put forward convincingly by Rodríguez-Ordóñez is that "Basque DOM is involved in a replica grammaticalization process induced by Spanish borrowed verbs and null objects of Basque, leading to a possible convergence between the two systems" (Rodríguez-Ordóñez 2017: 346).

To summarize, the described language situation bears all characteristics of a long-term and intense contact situation so that contact-induced changes are to be expected. I have presented several synchronic and diachronic changes in both Basque and Spanish which can be attributed to the extended contact between the two languages. With that in mind, let us now have a closer look at the presented phenomenon of non-adjacency in wh-interrogatives in Spanish and Basque.

6.1.4. whSV(X) – a modern phenomenon?

As described at length in Chapters 3 and 4, Basque and Spanish differ in the word order in wh-interrogatives with heavy wh-phrases. In Spanish, a lexical subject may intervene between the heavy wh-phrase and the verb. In Basque, however, this is never allowed. As mentioned already in the beginning of this chapter, my investigation suggests that bilingual Spanish-Basque speakers accept these interventions less than monolingual Spanish speakers. My hypothesis is that this is the result of the Basque-Spanish language contact.

If we assume an external, contact-induced influence, we have to first look at the phenomena in the model language to determine whether it is old or not (Thomason 2001: 93-94). Unfortunately, there is
hardly any work on diachronic development of word order within wh-interrogatives in Spanish. One of the few is Lapesa’s (1992) "La interpolación del sujeto en las oraciones interrogativas." He writes that these kind of insertions had been mentioned for the first time in an article by Navarro Tomás in 1929 (Lapesa 1992: 548). It seems that the phenomenon of subject intervention started in the Spanish speaking communities in central and southern America. The time of emergence in Andalusia and other parts of Spain, however, is not clear at all (Lapesa 1992: 548). Despite these difficulties, most linguists nowadays assume that is it a relatively recent phenomenon, which has spread to the most parts of the Spanish speaking areas.

This is where Basque starts to interfere. In Section 6.1 and 6.1.3, I have described the accelerating and decelerating effects of language contact on internal language change. A case of acceleration is the emerging use of interrogatives as markers of relative clauses. An example of delayed effect is the much slower spreading of the phonological phenomenon called yeísmo. Since Basque and Spanish both have a very similar structure in wh-interrogatives, the bilingual speakers prefer using the same structures in both languages. In other words, Basque speakers copy the overall very similar structure in wh-questions from Basque to Spanish, using the wh-phrase-verb adjacency no matter the length of the wh-phrase.

If we assume a decelerating influence on the Spanish structures, the question is, which speakers ‘slow down’ the spreading. Is it mainly the bilingual population or, as in the case of the yeísmo, a more general characteristic of the whole Basque-speaking area? If it is the bilingual speakers, do we find significant differences between sequential or simultaneous bilinguals speakers?

Before I will answer these questions using my own study on interrogatives, I will briefly detail some important aspects of language acquisition and how it could possibly interfere with contact-induced change and delay in language change.

6.2 On the role of language acquisition in a language contact situation

6.2.1. Some general remarks on acquisition

In the ongoing discussion about contact-induced language change, the role of bilingual language acquisition (BLA) is an important subject. The question of wether bilingual aquisition can be responsible of change or not is widely discussed. The first position, claiming that bilingual speakers have no influence on language change is supported, among others, by Meisel, Elsig & Rinke (2013) and Poplack & Levey (2010). In a nutshell, they say that contact-induced change only happens through imperfect L2 learner who impose their grammar on the new L1 speakers. The other position is supported, among others,
by Duguine & Irurtzun (2014), Kehoe (2015), Thomason (2011), and Yip & Matthews (2007). Especially Thomason (2011) argues that Meisel, Elsig & Rinke’s (2013) arguments fall short on explaining certain changes described by other scholars. A cogent example of contact-induced change by bilingual learners is Duguine & Irurtzun’s (2014) work about the emergence of *wh-in-situ*-structures in Iparralde, in the French Basque country. I will follow the latter group, assuming that under certain circumstances a contact-induced change may be triggered by BLA.

In the term BLA, various gradings of bilingualism are included. In order to differentiate them, I will give a brief overview of the main types. In the investigation on language acquisition, two main factors characterize the acquisition type, namely the onset of acquisition and the order or sequence of acquisition (among others, Montrul 2008: 17, Meisel 2012: 393). The following diagram shows the four main types.

As we can see, the first distinction is simultaneous versus sequential bilingual acquisition. We speak of the former if the child acquires both languages from birth on. For example because the parents use their respective native language. The child hears both languages to a more or less equal degree. Investigations have shown that children can differentiate both languages from very early on building two separate grammars for the respective languages (among others, Meisel 2012: 393, Paradis & Genesee 1996: 2, Yip & Matthews 2007: 2). Furthermore, 2L1 learners “subsequently proceed through the same developmental sequences as monolinguals in their respective languages, and […] attain native compe-
tence in each of their languages” (Meisel 2012: 393). The latter type, sequential bilingualism, must be divided further into two subtypes, namely early and late sequential acquisition. Apparently, the ease and speed of acquisition decreases to the point that after a certain age, no native competence can be gained anymore. Nowadays, linguists agree that the reason for these sensitive phases lies in neuropsychological changes (among others, Meisel, Elsig & Rinke 2013, Montrul 2008, Müller, Kupisch, et al. 2011). Following the cited works, I assume two steps of sequential bilingualism. The first is early child bilingual acquisition (ecL2) with an age of onset of around 4 years, and the late child bilingual acquisition (lcL2) with an age of onset of around 6 years. In Chapter 7, I will use this classification in order to divide the participants into the respective speaker groups.

In the next section, I will have a more detailed look into the explicit bilingual acquisition of Basque-Spanish bilinguals and a possible reason for a contact-induced language influence.

6.2.2. *wh*-interrogatives in Basque-Spanish bilingual acquisition

As mentioned earlier, the developmental sequences of language learning is surprisingly uniform for 2L1, ecL2, and monolingual children. In fact, the order of acquisition of constructions does not alter between children, merely the time they need for to acquire it (cf. Kintana & Meisel 2006, Meisel 2011, Montrul 2008, Müller, Kupisch, et al. 2011, Yip & Matthews 2007). Studies on the acquisition of *wh*-interrogatives in Basque-Spanish bilingual children by Barreña (1994) and Kintana & Meisel (2006) have shown this for Basque and Spanish. Kintana & Meisel, for example, write that their “investigation confirms the finding by previous research, according to which the implementation of finiteness into developing grammars precedes productive use of *wh*-questions” in both languages (Kintana & Meisel 2006: 571). Barreña (1994: 482) notes the same in his investigation. That means, children always ‘activate’ the IP-layer before starting to use questions productively. This holds for Basque as well as Spanish. However, Barreña (1994: 484) as well as Kintana & Meisel (2006: 571/572) describe a crucial difference in the acquisition of Basque compared to Spanish.

> [It has become evident that, in spite of considerable similarities between Basque and Spanish with respect to possible word order in *wh*-constructions, the two children studied longitudinally develop, from very early on, distinct syntactic systems for the two languages.]

Kintana & Meisel (2006: 571)

---

64 I will count the Finit-Phrase as the extended part of the IP, similar to Barbosa (2001). This kind of the extended IP is supported by the assumption that in both languages no verb movement to a higher C-head takes place. As detailed in Section 3.3.2 and 4.2.2, the verbal complex normally stays in $I^0$. 

---

139
As soon as children who acquire Basque start to use wh-questions productively, they also start to use several other structures associated with the preverbal focus position. Two bilingual children Mikel and Peru, from the study conducted by Kintana & Meisel (2006), for example, start to use wh-questions productively around the age of 2:00 and 2:05, respectively. That means, Mikel starts to use wh-phrases other than ‘where’ and Peru other than ‘what’ together with a verb form other than ‘is’. Around the same time, they begin to place the negative particle ez pre-verbally. The example for the first productively used wh-question are given in (6.6a) and (6.7a), the examples for the first pre-verbal negation in (6.6b) and (6.7b).

\[(6.6)\]
\[\begin{align*}
\text{a.} & \quad \text{Hemen zer dago?} \\
& \quad \text{here what be.3sg} \\
& \quad \text{‘What (do we) have here?’ (Mikel 2;00;00)} \\
\text{b.} & \quad \text{Hau ez da ibiltzen} \\
& \quad \text{this not aux going} \\
& \quad \text{‘This is not working.’ (Mikel 1;11;06)}
\end{align*}\]

\[(6.7)\]
\[\begin{align*}
\text{a.} & \quad \text{Tatoga(k) zer egin du?} \\
& \quad \text{tractor what do aux} \\
& \quad \text{‘What has the tractor done?’ (Peru 2;05,12)} \\
\text{b.} & \quad \text{Hori ez daukat handia?} \\
& \quad \text{this not has big} \\
& \quad \text{‘Doesn’t (she) have a big one there?’ (Peru 2;05,12)}
\end{align*}\]

Kintana & Meisel (2006: 567,569)

In addition, around the same time, the children start to position elements pre-verbally in order to focalize them, as can be seen in the examples (6.8a) and (6.9a). Especially (6.9a) indicates a clear focalization, since subject pronouns are normally dropped in Basque. Finally, they also start to use polar questions with a target-like word order, as illustrated in (6.8b) and (6.7b) (Kintana & Meisel 2006: 572).

\[(6.8)\]
\[\begin{align*}
\text{a.} & \quad \text{han dago halakoa} \\
& \quad \text{there be.3sg so} \\
& \quad \text{‘There it is like that.’ (Mikel 2;00;00)} \\
\text{b.} & \quad \text{Apurtu egin da hori?} \\
& \quad \text{break make aux this} \\
& \quad \text{‘Has that one been broken?’ (Mikel 2;00,00)}
\end{align*}\]

\[(6.9)\]
\[\begin{align*}
\text{a.} & \quad \text{Hori nik egingo dut} \\
& \quad \text{that L.ERG do.FUT aux} \\
& \quad \text{‘I will do that.’ (Peru 2;05,12)}
\end{align*}\]

Kintana & Meisel (2006: 567,569)
The data suggests that all these are movements to the same position and set off by the same trigger, for example the necessity to always occupy the preverbal position. This could explain the simultaneous emergence of all these different phenomena.

In the Spanish data, on the other hand, this simultaneous emergence of different but interlinked structures does not take place, even though the preverbal position SpecFin is also the primary landing site for wh-phrases during acquisition. An interesting observation in this context, however, is that the children in both languages “commit virtually no word order errors” and with that behave very different from, for example, children acquiring English who “in a later phase, fail to place the subject post-verbally in wh-questions with clause-initial wh-words” (Kintana & Meisel 2006: 571). English wh-movement entails I-C-movement of the verb in order to establish the necessary adjacency between wh-phrase and verb since the subject moves always to SpecIP. For a correct production of wh-interrogatives, the CP-layer has to be activated. This is not necessary in Basque and Spanish, in which the wh-phrases move only to the IP-layer.

If my prediction is correct and bilingual and monolingual speakers differ in the acceptance of the non-adjacency with heavy wh-phrases, how could we explain this fact with the presented data of the acquisition in both languages?

Two important properties of the input could play a decisive role here, namely quantity of input and input ambiguity. Both can influence the acquisition process (among others, Duguine & Irurtzun 2014, Kehoe 2015, Kupisch 2007, Yip & Matthews 2007). Quantity of input is related by several linguists to dominance. Kehoe (2015: 153) for example writes that “[t]he language that the child hears and uses the most is typically his dominant language”. For the ease of discussion, I will assume that a child who acquires Basque and Spanish simultaneously, receives a more or less equal input between 40-60% in each language, at least until they enter kindergarten. In the case of the sequential bilinguals, I assume that the first language continues being the language with the higher input due to the normally higher amount of hours at home than in the kindergarten, at least until primary school.

Input ambiguity refers to superficial structures that could be interpreted in different ways, depending on the underlaying structure. A good example for this is an emerging wh-in-situ-strategy in young speakers of Labourdin Basque in the French part of the Basque-speaking area described in Duguine & Irurtzun (2014). Contrary to Spanish and Basque, French offers the possibility of wh-in-situ-interrogatives (cf., among many others, Z. Bošković 2000, Cheng & Rooryck 2000, Duguine & Irurtzun 2014). Duguine & Irurtzun (2014) have observed that the speakers of young Labourdin Basque (YLB) start to use wh-in-situ-interrogatives similar to the French structures. These newly arisen structures in YLB are governed by the same restrictions as the French wh-in-situ-questions. Duguine & Irurtzun (2014)
give the large amount of input ambiguity as one of the main causes for the emergence of this new structures. This ambiguity arises due to Basque’s three-way pro-drop properties. Many Basque questions resemble \textit{wh-in-situ}-questions on the surface. The example in (6.10) illustrate such a ambiguity. In both sentences, the verb and the object \textit{wh}-phrases superficially appear to be in the canonical order, since Basque is a verb-final language. This means the structure could be analyzed equally well as \textit{wh}-movement to the preverbal position or as a \textit{wh-in-situ}-question. The linear order on the surface allows for both options. In Basque speech, we find a lot of such ambivalent structures due to its three-way pro-drop property (Duguine & Irurtzun 2014: section 3.1).

(6.10) a. Zer eman duzu?  
what give AUX.2SG.ERG.3SG.ABS  
‘What did you give?’

b. Nori eman dizkiozu?  
who.DAT give AUX.2SG.ERG.3SG.DAT.3PL.ABS  
‘When will you come?’

Duguine & Irurtzun (2014: e13)

Therefore, children learning Basque and French are apparently inclined to interpret certain interrogatives as \textit{wh-in-situ}-questions, similar to French. A further important factor, according to the authors, is the overall preference for ‘movementless’ operations. That is to say, if the child has a more ‘economical’ alternative to movement available, like \textit{wh-in-situ}-constructions, it is inclined to use it. (Duguine & Irurtzun 2014: e18-e21).\footnote{Another mayor reason for this new \textit{wh-in-situ}-phenomenon, according to the authors, is the sociolinguistic change from primarily Basque-French sequential BLA to French-Basque sequential BLA. For a detailed discussion about this newly arisen \textit{wh-in-situ}-structures, the reader is referred to Duguine & Irurtzun (2014).} Similar causes are assumed by Strik & Teresa Pérez-Leroux (2011) for French-Dutch bilinguals. Also these children prefer the more ‘economic’ movement of leaving the \textit{wh}-phrase \textit{in-situ}. Especially in the case of the French-Basque speakers of YLB, we see that these changes can remain and be integrated into their grammar.

In the case of the Basque-Spanish bilinguals however, the difference in questions formation, at least superficially, is not as big as between French and Basque. That means, every simple or complex \textit{wh}-question a child hears or produces in Basque normally fits into the question formation pattern of Spanish. In other words, Spanish question formation presents a superset of question formation in Basque.\footnote{This even holds for the exceptions with the adjunct \textit{wh}-words like Sp. \textit{por qué} Bq. \textit{zerkatik} ‘why’.} The only major difference in question formation are complex \textit{wh}-questions and multiple \textit{wh}-questions. I will return to the latter type in more detail in Chapter 8. In the former type of questions, Spanish offers two possibilities. The first is the standard word order in Spanish \textit{wh}-questions, namely
On the role of language acquisition in a language contact situation

the *wh*-phrase-verb adjacency. The second is the non-adjacency of the *wh*-phrase and the verb. Basque, on the other hand, only allows the first configuration of question formation. Several studies, Müller & Hulk (2001) among others, have shown that in such a setting, the following tends to happen:

Once language A allows for more than one grammatical analysis from the child’s perspective and language B contains positive evidence for one of those possible analyses, language A is likely to be influenced by language B.

Müller & Hulk (2001: 19)

If we apply this statement to the situation at hand, A would correspond to Spanish with the two possible structures in complex *wh*-questions, and B to Basque with the ‘stricter’ but less complex word order constraint. That means, the simultaneous and sequential Basque-Spanish bilingual children would be influenced in their decision of which structure to use in Spanish questions by Basque. They would tend towards the *wh*-phrase-verb adjacency. The question is now, why these children do not seem to acquire the second option in Spanish. At this point, the quantity of input becomes a deciding factor.

If we compare the quantity of input, that is the number of *wh*-questions with simple and heavy *wh*-phrases actually heard by the child, we can see a clear distribution. I have consulted four corpora: three written corpora, *Corpus de referencia del español actual (CREA)*, *Corpus del Español del Siglo XXI (CORPES)*, and the *Corpus Del Español (CdE)*, and one oral corpus *Corpus Oral y Sonoro del Español Rural (COSER)*. The percentages of simple and complex *wh*-phrases from the Spanish subcorpora are illustrated in Table 6.2. The complex *wh*-sentences represent all questions of this type, not taking into consideration differences in the *wh*-phrase-verb adjacency.

<table>
<thead>
<tr>
<th>Corpora</th>
<th>simple <em>wh</em>-phrases</th>
<th>heavy <em>wh</em>-phrases</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
<td>#</td>
</tr>
<tr>
<td>CREA</td>
<td>104,645</td>
<td>99,4%</td>
<td>734</td>
</tr>
<tr>
<td>CORPES</td>
<td>136,018</td>
<td>99,4%</td>
<td>881</td>
</tr>
<tr>
<td>COSER</td>
<td>16,620</td>
<td>99,97%</td>
<td>6</td>
</tr>
<tr>
<td>Corpus del Español</td>
<td>3,138,459</td>
<td>99,6%</td>
<td>14786</td>
</tr>
</tbody>
</table>

CREA, CORPES and CdE are written corpora and therefore do not represent the relation in oral speech with total accuracy. The COSER is a much smaller corpus, but represents oral data. As the data from COSER already indicates, we can assume with near certainty that the percentage of heavy *wh*-phrases in oral speech is lower than in written texts, possibly due to their complexity. In addi-

---

57 The data is only for Spanish. However, a similar distribution in Basque seems probable.
tion, the non-adjacency is only a optional structure and, in Spanish, most likely also a relatively recent phenomenon. Presumably therefore, the share of complex wh-questions without wh-phrase-verb adjacency is very low, even in a monolingual Spanish environment. If we assume a more or less equal exposure to both languages in the bilingual cases, the number of any form of a complex wh-sentence is even lower for each language. Nevertheless, for the monolingual children at least, the amount of input seems to be enough to acquire both possibilities. According to Yang (2016), it is exactly the scarcity of the input that allows children to establish the necessary rules. He assumes that a threshold exists, up to which exceptions are tolerated. His Tolerance Principle states that if the number of items that the child hears crosses this threshold, the rule is rejected (Yang 2016: 10). He furthermore continues that children acquiring a language only generalize a certain rule if enough positive evidence is collected. His Principle of Sufficiency says that a rule is only established in the first place if the following holds: the actual number of items following the rule less the items that contradict the rule have to stay below a certain threshold again (Yang 2016: 182). The important point here is that Yang’s (2016) Principle of Sufficiency exactly predicts that only few examples are necessary to establish the correct rules. Applied to our example here, it means the following: the monolingual child that acquires Spanish probably hears very few complex wh-sentences in its first two or three years. Out of these already few examples, only a small share has a non-adjacent structure. Nevertheless, the amount is enough to establish the necessary rule that non-adjacency is facultative in these constructions. The bilingual child, however, only hears half the amount of complex wh-sentences in Spanish compared to the monolingual child. As a consequence, the amount of structures without adjacency is very low. Thus, according to the Principle of Sufficiency, this very small amount does not give enough positive evidence to establish an independent rule for complex wh-sentences. Therefore, “the absence of evidence will simply hold the learner back to stay conservative” Yang (2016: 220). Conservative here means that the children always apply the well established, and less complex, obligatory adjacency rule present in Basque, which is anyway compatible with most of what the bilingual children hear in Spanish. Therefore, two factors come together in the case of bilingual children. First, the bilingual child hears an insignificant number of complex wh-questions without adjacency and, second, the child anyway tends to use the structure that is common to both languages and less complex, namely the one with wh-phrase-verb adjacency. If we consider this behavior under the aspect of internal language change in Spanish, this ‘conservatism’ by the bilingual speakers, concerning the analysis of complex wh-phrase, could be interpreted as ‘resistance’ to the change.

68 For a complete and exhaustive explanation on the two principles and their mathematical derivation and justification, the reader is encouraged to have a closer look at Yang (2016).
What would such a hypothesis predict? First, the study should show differences between 2L1 and child early/late Basque-Spanish sequential learners on the one hand and Spanish-Basque sequential on the other hand. The reason being that the former have already set the preverbal focus parameter for wh-phrases in Basque and, due to the very similar superficial structure, tend to set the same for Spanish wh-questions. We have seen earlier, in the data form the studies conducted by Kintana & Meisel (2006), that the productive use, and with that the implementation, happens early in both languages. Spanish-Basque sequential bilingual learners however, have a relatively well established Spanish structure when they start their exposure to Basque at age three or five, respectively (Kintana & Meisel 2006). This last point is corroborated by other studies on wh-questions. Slabakova (2016: 125) writes, citing among others Cole et al. (2000), that children acquiring a wh-in-situ-language leave the wh-phrases in-situ from early on. Whereas in the case the acquisition of wh-moving languages, they move them to the beginning of the ‘utterance’\(^{69}\): “This distinction suggests that the wh-movement parameter, which regulates whether wh-words move to the beginning of the sentence or not, is set rather early” Slabakova (2016: 125).

A further indication for a different interpretation of the preverbal focus position could be other focus-sensitive structures in which the first group differs from the second. I haven’t pursued this question in my study in depth. In Section 6.1.3, however, I have mentioned the works by Gómez Seibane (2012), Irurtzun (2018), and Urrutia Cárdenas (1995), who suggest that Basque bilingual speakers indeed use Basque-like structures in Spanish in other focus-sensitive environments.

In summary, we have seen that the overall setting of the Basque-Spanish contact situation offers a perfect ground to expect contact-induced changes and influences. I have presented my hypothesis that simultaneous Basque-Spanish and cL2 Basque-Spanish bilinguals apply the Basque preverbal focus position to all wh-structures in both languages, contrary to the other sequential bilinguals and the monolingual speakers. The hypothesis is supported by the circumstances during the acquisition process, like small quantity of input, input ambiguity, and the overall tendencies of bilinguals to stick to one common structure. This could account for the slower spread of a possible language change that is observed in most of the other Spanish speaking regions with respect to the non-adjacency in complex wh-questions.

In the next chapter, I will present my study on the acceptance of interrogative structures by mono- and bilingual speakers in and outside the Basque-speaking area.

\(^{69}\) Slabakova (2016: 125) gives examples such as where go? or what hit?. These utterances are not complete sentences yet, but they show that the object or adverbal wh-phrases have clearly been moved.
CHAPTER 7

Empirical study on \textit{wh}-interrogatives

In order to verify my hypothesis, I carried out an empirical study on the acceptability of non-\textit{wh}-phrase-verb adjacency in \textit{wh}-interrogatives. The study consisted, in total, of four questionnaires. First I launched a pilot-test (Pilot-Questionnaire, \textit{PilotQ}), then a sociolinguistic questionnaire to gather a stable group of participants (Sociolinguistic Questionnaire, \textit{SoLQ}), and finally two questionnaires to test the relevant structures (Questionnaire 1 and 2, \textit{Q1}, \textit{Q2}). The questionnaires were carried out between December 2015 and October 2016.

The methodology and the structure of the questionnaires will be addressed in Section 7.1, followed by a detailed description of the participants and their sociolinguistic background in Section 7.2. Section 7.3 will describe the questionnaires in detail and finally, in Section 7.4, I will present the results before discussing them in Chapter 8.

7.1 Methodology

When linguists decide to collect data, one has to choose between indirect or direct collection. One crucial question is whether we need to carry out interviews to get spoken data, or if it is enough to rely on written data. Both methods have their advantages. The first method provides real-time data. A clear advantage is the ability to react to and interact with the participants. They can be asked to confirm a certain grammatical structure and or provide more detail. The disadvantage is that depending on the complexity of the phenomenon you are investigating, it is difficult to have sufficient incidents occurring in natural speech during an interview. In addition, you never get ‘negative’ evidence or ungrammatical structures in an interview, since the participant obviously only uses ‘correct’ sentences. The interviewer can ask about a specific construction, but a speaker will not produce a structure that does not occur in his or her grammar. In addition, interviews need much more time to conduct and to transcribe.
Finally, it is much more difficult to reach a larger number of people. The second method, written data, has the advantage that you can reach many more speakers in a shorter time. Especially nowadays, the internet provides a perfect instrument to find participants and carry out a written study. Questions can be answered at home and there is no need to go somewhere in particular or send the filled out questionnaire back to you. A written questionnaire offers a way to ask about very specific constructions and control the parameters involved. The disadvantage is that it is not possible to interact with the participants. There is always the possibility that the participants did not understand the task fully. You cannot quickly change a word or element in the sentence to see if it would alter the speaker’s judgement. In addition, in order to get viable data, you have to insert enough filler phrases (Schütze & Sprouse 2013: 14). They have two tasks. First, they prevent the participants from guessing the relevant questions and deducting what the study as a whole is about. Second, they “also serve as anchors, to remind subjects [i.e. participants] of the range of potential goodness and badness” (Schütze 2016: 190).

For every 10 items you are interested in, between 10-20 fillers should be added to the questionnaire. This condition lengthens the questionnaire considerably. I will show, however, that they can be useful, beyond the aspects mentioned above.

For my purpose, the written, internet-based questionnaire was the method of choice for two main reasons. First, I wanted to reach as many speakers as possible and have the possibility to contact them again easily. The objective was to be able to argue on the basis of statistically significant differences and the more data I had, the more reliable they would be. The second reason is the type of examined constructions. Heavy wh-phrases are very particular constructions, which do not occur very often in oral speech. It was therefore more convenient to create them artificially. Therefore I decided to use an internet-based questionnaire to perform a grammatical judgement test. In the following segment I will describe the method, implementation, and execution of the questionnaires.

All questionnaires were carried out with the internet-based program ‘SoSci-Survey’70. The program is freely available for non-profit scientific research. It is a powerful tool which allows many configurations and manual adjustments via HTML-programming. For example, it is possible to create a participant data base including names, e-mail addresses, and division into special groups, and it allows for sending of personalized links for questionnaires. It offers several pre-designed question types and can easily randomize the questions for every participant. In addition, the program has very clear and detailed documentation, which helps to design and conduct good questionnaires.

70 www.soscisurvey.de
7.1 Methodology

The structure of the four questionnaires was always adapted to the necessity and objective. The SoLQ consisted mainly of multiple choice questions in order to categorize the participants depending on their sociolinguistic background. Q1 and Q2 were composed of 31 and 20 questions, respectively. They were both grammatical judgment tests. The first PilotQ was a mixture of both types.

I designed the SoLQ with two main objectives. First, all participants should be classified according to their sociolinguistic profile. After requesting certain information, the program automatically divides the participant into speaker groups depending on the answers given and according to the configurations defined by myself. For example, if a participant declared that he/she learned Spanish and Basque at home, he/she was assigned to the 2L1 group. If they declared having learned Basque at home and Spanish in primary school, they were assigned to the Basque-Spanish child-late group, etc. The second reason for using an indirect, internet based survey, was that I wanted to gather a stable group of participants in order to have the same population of speakers for all my further research. For this purpose, at the end of the SoLQ, the participants were asked to provide their e-mail address. With that, they gave their consent to save their data and to receive further invitations to my questionnaires. They had to confirm an automatically generated e-mail. Only then their data was saved in the participant data base. The stored data for each participant consisted of the name, the e-mail address and the assigned group. All further data was saved anonymously. This was done to protect the anonymity of the participants, as I cannot connect any of the data with the names of the participants.

The participant’s recruitment happened mainly by personal diffusion of the web-link for the SoLQ via e-mail and Facebook. After compiling the SoLQ online, it was enough to send a web-link to the potential participants. With that, they could access the questionnaire online and fill out the requested information.

Q1 and Q2 were designed very similar to one another with just some minor adjustments to the layout in the second. Due to the saved email addresses, I could send the link directly to the assigned participants. Furthermore, I sent a second e-mail some days later to remind those participants who had not yet answered the questionnaires. While the questionnaires were online, I could see who had already finished, who opened the link but hadn’t completed the survey, who received the e-mail but had not yet accessed the questionnaires, and who did not receive the mail at all. In this way, I was aware of the progress at all times.

The questions’ structure was the same for all questions in the two questionnaires Q1 and Q2. An example is shown in Figure 7.1.\footnote{The internal coding for the items is as follows: In Item_1.01 or Item_2.03, the first number before the dot indicates the questionnaire, Q1 or Q2, respectively. The second number indicates the item/question. Thus, Item_2.03 is the third item in the second questionnaire. All items are listed in Appendix B and C.}
7.1 Methodology

As can be seen in the example, the item is introduced by a short context to establish the discourse and give the necessary information. Since heavy wh-phrases are, as I have argued, mostly analyzed as d-linked, it is necessary to establish this d-linkededness through context. The sentence that was to be judged is in bold and clearly recognizable. The last segment is the instruction and the actual scale. I decided to use a running scale instead of a five or seven point scale in order to capture more subtle differences. The two extremes of the scale were me parece muy raro ‘it sounds strange to me’ and me parece natural ‘it sound natural to me’. The participants had to use the indicator to give their judgment. The qualification was saved as a number between 1-101. The data could then be downloaded as .csv-file to import it to the relevant statistical program. I used R, a free open-source software for statistical computing and graphics.

The process was the same for questionnaires Q1 and Q2. The participants used the email link to access the questionnaire. The first page was a short introduction that explained the task. Then the program shuffled the questions so that every participant had to answer the questions in a different order. This way, I avoided possible effects of tiredness or reluctance towards the end. On the last page the participants were thanked for their help and participation. The introductory texts and all questions are listed in Appendix B and C.

In the following section, I will give a detailed picture of the participants and their sociolinguistic background before describing the questionnaires in the subsequent section.

Figure 7.1: Item_2.03 – as in Q2

72 www.r-project.org
7.2 SolQ – the participants

SolQ established a clear and useful sociolinguistic classification of the participants into distinct groups. The participants had to answer several questions about their socio-linguistic background. At the end of the questionnaire, they could insert their e-mail address and confirm their willingness to participate in future questionnaires. 181 participants subscribed in this way. However, my data analysis is based on only 141 persons for two reasons. First, 38 persons answered neither Q1 nor Q2. Second, two additional speakers had to be removed due to extreme irregularities in their answers. They judged several verified correct sentences as wrong, or wrong sentences as correct. One can conclude from this behavior that they simply did not read the sentences or the instructions carefully enough. The final set of 141 participants became the base for all further analysis.

Depending on the answers to the questions about their sociolinguistic background, the participants are divided into the two main groups bilingual Basque-Spanish and monolingual Spanish speakers. The latter are further separated into monolinguales originating from the Basque speaking area (M-BA) and monolinguales from other parts of Spain (M-SP). The former group, the bilinguals, is divided into five subgroups. These derive from the types of bilingual acquisition presented in Figure 6.1 in Section 6.2. The first are the simultaneous Basque-Spanish bilinguales (2L1). The second and third group are the sequential bilinguales who started with Basque and added Spanish at four (Basque-Spanish, child-early - BS-ce) or six years of age (Basque-Spanish, child-late - BS-cl), respectively. The forth and fifth group are the sequential bilinguals who started with Spanish and acquired Basque at four (Spanish-Basque, child-early - SB-ce) or six years of age (Spanish-Basque, child-late - SB-cl), respectively.

In order to get the necessary information about the age of onset in the two languages, I asked about the moment they began to acquire or to learn the respective languages. The possible answers were ‘at home’, ‘in kindergarten’, ‘in primary school’, ‘in secondary school’, ‘after the age of 18’, and ‘other’. ‘Kindergarten’ corresponds to the second age step in Figure 6.1, that is, from three or four years onwards, and ‘primary school’ to the third age step, that is, from 5 or 6 years onwards. ‘Secondary school’ and ‘after the age of 18’ were classified as adult second language acquisition. Thus, depending on the combination of the answers, the participants were divided into the seven speaker groups. In Table 7.1, the overall distribution and number of speakers per group are shown.

We can see in the table that the two groups BS-ce and SB-cl are very small. In Q1, only one speaker of the latter group finished the questionnaire and from the former group only five speakers in both Q1 and Q2 did. Due to this small number in both questionnaires, I decided not to include them in the
Table 7.1: Total number of participants per speaker group

<table>
<thead>
<tr>
<th>Speaker Group</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simultaneous bilingual acquisition</strong></td>
<td></td>
</tr>
<tr>
<td>2L1 Basque - Spanish, simultaneous</td>
<td>40</td>
</tr>
<tr>
<td><strong>Sequential bilingual acquisition</strong></td>
<td></td>
</tr>
<tr>
<td>BS-ce Basque - Spanish, child early</td>
<td>6</td>
</tr>
<tr>
<td>BS-cl Basque - Spanish, child late</td>
<td>24</td>
</tr>
<tr>
<td>SB-ce Spanish - Basque, child early</td>
<td>10</td>
</tr>
<tr>
<td>SB-cl Spanish - Basque, child late</td>
<td>2</td>
</tr>
<tr>
<td><strong>Bilingual - total</strong></td>
<td>82</td>
</tr>
<tr>
<td><strong>Monolingual acquisition</strong></td>
<td></td>
</tr>
<tr>
<td>M-BA Spanish monolinguals, Basque speaking area</td>
<td>26</td>
</tr>
<tr>
<td>M-SP Spanish monolinguals, other parts of Spain</td>
<td>33</td>
</tr>
<tr>
<td><strong>Monolingual - total</strong></td>
<td>59</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>141</td>
</tr>
</tbody>
</table>

statistical analysis as their number is statistically not representable. Nevertheless, I will include them in the description of the participants in this ongoing section.

In total, 78 of the participants were women, 43 men, and 20 did not indicate their gender. The distribution per group is shown in Figure 7.2.

A further important parameter is the place of origin. The vast majorities of the bilingual and monolingual speakers from the Basque speaking area are from Gipuzkoa (75/82 and 21/27, respectively).
7.2 SolQ – the participants

The remaining seven speakers in the former and the six in the latter come from Bizkaia, Araba or Navarre.\textsuperscript{73}

This distribution is highly welcome. In order to minimize dialectal influences in the bilingual groups, it is important that they speak the same or at least a similar dialect. As I have explained in Chapter 2, this has to be taken into special consideration with the strong dialectal variety we find in Basque. The remaining monolinguals are widely spread over the whole country. The three regions or cities where most speakers come from are Catalonia (2), Barcelona (5), and Madrid (10).\textsuperscript{74}

I requested additional information from the participants about their age and their degree of education. As can be seen in Figure 7.3, the age of the participants ranges from 15 to over 60.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Figure_7.3.png}
\caption{Total age distribution}
\end{figure}

The distribution represents both the generation of speakers before and after 1980. This was an important point in time in the history of the Basque country, as explained in Chapter 2. One third of the bilingual speakers are born before 1980 (28/82). That is, they grew up in a Spanish state that was not supportive of Basque language. The remaining speakers grew up in an Basque-friendly environment. Thus, from the 1980s on, the number of children who learned Basque in school increased and the \textit{Euskara Batua} introduced a degree of standardization. In Figure 7.4, we see the distribution of the age in the speaker groups.

\textsuperscript{73} The overall distribution:
- 2L1: Araba (1), Bizkaia (1), Gipuzkoa (38)
- BS-ce: Gipuzkoa (6)
- BS-cl: Bizkaia (2), Gipuzkoa (22)
- SB-ce: Araba (1), Bizkaia (1), Navarra (1), Gipuzkoa (7)
- SB-cl: Gipuzkoa (2)
- M-BA: Navarra (2), Bizkaia (4), Gipuzkoa (20)

\textsuperscript{74} The other regions are Andalucía, Burgos, Cáceres, Ciudad Real, Córdoba, Galicia, Girona, Logroño, Madrid, Málaga, Murcia, Santa Cruz de Moya, Sevilla, and Zaragoza. For each of these regions, always one speaker named it as his or her origin.
The overall distribution of the second parameter, namely ‘Degree of Education’, is shown in Figure 7.5. As can be seen, the majority of the participants have a higher education, i.e. they hold an academic title or attended a technical college.

In the case of the bilingual speaker, I was interested in the daily use of their languages. I asked them to indicate, as a percentage, the use of each of their languages during a normal day (Figure 7.7). Closely connected to that, I asked them about the language(s) spoken at home and with their families (Figure 7.8 and Figure 7.9, respectively). Additionally, I wanted to know in which language they feel more comfortable (Figure 7.10).
In Figure 7.10, we can see in the case of the 2L1 that the average language use is almost the same for Basque and Spanish. In addition, about 50% of them feel equally comfortable in both languages, whereas the other half is divided almost equally between the two languages. The slight majority speaks either Spanish or Basque at home and with the family. However, the use of both languages is very common. The BS-cl prefer Basque and use it more at home. However, also in this group, the equal usage of both languages is widely confirmed by the participants. The rest of the bilinguals clearly prefer Spanish. The two other languages mentioned in the M-SP group were Catalan and Galician.

In addition to all the former mentioned parameters, the participants were asked about the native languages and place of origin of their parents.

All questions from the SolQ can be looked up in Appendix A. In the next section, I will present and describe the three questionnaires I carried out.

7.3 Questionnaires

In total, I conducted three online-surveys. The first questionnaire was a pilot-test (PilotQ) in order to get a clearer picture on the subject of investigation. It was carried out between December 2015 and January 2016. The second (Q1) and third (Q2) questionnaire were carried out in July and October 2016,
respectively. I will briefly talk about the pilot-test in Section 7.3.1 and then describe Q1 and Q2 in Section 7.3.2 and Section 7.3.3, respectively.

7.3.1. Pilot-Questionnaire – PilotQ

The pilot-test had three objectives. First, I used it to evaluate which sentences were interesting to look into. Second, I wanted to know if the obtained sociolinguistic information was sufficient to divide the participants into the intended groups and subgroups. Third, I wanted to test the layout and the overall comprehension. I will give a short summary of the insights and ideas the PilotQ gave me for Q1 and Q2, both in terms of structure and content. A full list of all the items and filler phrases is shown in Appendix D.

For the PilotQ, I extracted the items from CREA and adjusted them slightly. The questionnaire consisted of over 80 questions, including direct and indirect questions, filler phrases, and all the sociolinguistic questions described in the former section. Furthermore, the participants could add additional comments to each question, which gave me some very important feedback. The total number of participants was 31. Based on the experience and the insights gained through the PilotQ, I adapted the subsequent surveys in terms of structure as well as content.

The structural changes consisted of adapting the layout, increasing the proportion of filler phrases, adjusting the overall length, and refining the process of acquiring participants. The last two points go hand in hand. One of the reasons I established a stable base of participants was the insight and the feedback from the participants that the questionnaire had been too long. The stable group of participants helped me avoid that problem, as I could send several short instead of one long survey. In my experience, participants are more inclined to start a survey when they know that it only takes five to ten minutes to finish it. The layout was adapted slightly, especially to indicate more clearly the different poles and the task for the participants. Finally, the proportion of filler-phrases was increased to exceed the number of investigated items. With regard to the content, the changes included limiting myself to direct questions and overall fewer types of wh-phrases. I also added a short context to every questions and used sentences especially constructed for my purpose. The reason for leaving out the indirect wh-interrogatives and limiting myself to very few types of wh-phrases was mainly due to the theoretical underpinning. As elaborated in Chapter 5, the hope is to show that the bilingual and monolingual speakers differ especially in interrogatives with complex wh-phrases. Therefore, I concentrated on these type of wh-phrases, especially in Q2. For the PilotQ, I extracted the questions from CREA. However, I had to adapt them and simplify them in order to prevent a distraction of the participants.
due to unusual words or names. In the next survey, I decided to create the items myself. This way, I had much more control over all the parameters involved. In addition, as suggested by several participants in the feedback, I decided to add a short description the relevant context to every item in the subsequent questionnaires.

Overall, the PilotQ fulfilled its purpose, resulting in important insights and ideas for adjustments, which I could consider and implement in the next questionnaires.

### 7.3.2. Questionnaire 1 – Q1

The first full scale questionnaire was carried out in July 2016. I sent the first e-mail with the invitation and the personalized link to every participant subscribed in the database on July, 22nd. A few days later, I sent another reminder to all participants who had not yet responded. In total, 131 of the 181 subscribed participants finished the questionnaire, which is a return rate of slightly over 70%. Figure 7.11 shows the overall distribution of the 131 participation divided by the speaker groups.

![Figure 7.11: Overall participation – Q1](image)

Q1 consisted of 32 questions of which 20 were filler phrases. The twelve items were four interrogatives with *por qué* ‘why’, four with *a cuál de* ‘which of’, and four with *a quién de* ‘who of’. The full list of all items can be looked up in Appendix B. I decided to use these three types of wh-phrases for the following reasons. In the case of the first type, the wh-phrase *por qué* ‘why’, I wanted to verify the results from the PilotQ. They seemed to indicate that *por qué* in combination with any pronominal subject sounds deviant to the participants. This would contrast with the observation in other Spanish dialects. In addition, I wanted to see if the participants accept adverbial insertions between the wh-phrase and the verb, like *casi* ‘almost’ or *también* ‘also’. Both of them are attested in CREA and discussed in the literature, for example in Bosque & Gutiérrez-Rexach (2011) and Contreras (1999). Therefore, I had two
wh-interrogatives with a personal pronoun inserted between the wh-word por qué ‘why’ and the verb, namely nosotros ‘we’ and tú ‘you’, and two with one of the former mentioned adverbs each.

The second and third types were complex wh-phrases with cuál(es) de ‘which of’ and quién(es) de ‘who(m) of’. The insertions between the wh-phrase and the verb were either an adverb or a subject. The adverbs were of two different kinds. The adverbs of the first type were the two high sentential adverbs casi ‘almost’ and frecuentemente ‘frequently’. It had been noted in the literature that sentences with one of these adverbs as insertion between the wh-phrase and the verb should be acceptable (Bosque & Gutiérrez- Rexach 2011, Suñer 1994). The second group of low sentential adverbs, represented by probablemente ‘probably’ and también ‘also’, on the contrary, should not. The intervening subjects were either lexical or pronominal. The former could be a full DP like tu hermana ‘your sister’ or just names like María. The pronominal insertions were singular and plural second person tú ‘you’ and vosotros ‘you (pl.)’, respectively, and first person plural nosotros ‘we’. The function of the complex wh-phrases differed between subject (3x) and direct objects (5x). All 12 sentences and their characteristics are listed in Appendix B.2.

The filler phrases were a mixture of wh- and polar questions, as well as declarative and exclamatory sentences. Furthermore, some of the sentences were clearly ungrammatical due to agreement or word order errors. The full list of filler phrases can be looked up in Appendix B.3.

I will discuss the results in more detail in Section 7.4 and will only provide a brief overview here. In a nutshell, the results were not as clear as hoped for. Mostly because of the choice of the sentence and partly because of the layout. Therefore, I tried to improve the former and adapt the latter, creating Q2.

7.3.3. Questionnaire 2 – Q2

The second questionnaire was launched on October, 7 in 2016. It consisted of 20 questions in total, of which eight were items and twelve were filler phrases, resulting in the same ration as in Q1. As with Q1, I sent a first invitation e-mail and a second reminder a few days later. In total, 124 of the 181 subscribed participants finished it. Comparing it to Q1, the response rate dropped slightly to around 68%. Comparing the single groups, we see that for every group the return rate decreased by one or two speakers. Except for the group of the SB-cl in which the number increased by one person. Figure 7.12 gives an overview of the 124 participants, again divided by the speaker groups.

Q2 differs in terms of layout in two points from Q1. First, every question was now prefaced by a short introductory context, including all filler phrases. Second, the instruction was repeated after every
question. That is, I put the following request right above the judging scale: Por favor, indica cómo de natural te parece esta frase en el contexto indicado.\textsuperscript{75} An example is given in Figure 7.1, here repeated as Figure 7.13.

75 Please indicate how natural the sentence seems to you in the indicated context.

In terms of content, there were two basic changes. First, I removed one type of \textit{wh}-phrases and second, I focused more on the parallelism between CLLDed elements and heavy \textit{wh}-phrases. The reduction resulted in the two remaining \textit{wh}-phrases \textit{por qué} ‘why’ and \textit{cuál(es)} de ‘which of’. The main focus, however lay on the second type. I only added one item with \textit{por qué} in order to verify the acceptance rate in combination with an expressed pronominal subject tú ‘you (sg.)’ and \textit{wh}-phrase-verb adjacency.
The other seven items were variations with the heavy *wh*-phrase *cuál(es) de* ‘which of’. The reason I limited myself to the latter is the following: a core property of CLLDeD constituents that I presented in Chapter 3 is that these dislocated elements are always d-linked. A complex *wh*-question with *cuál(es) de* ‘which of’ as *wh*-phrase offers the same d-linkedness. It has a strong inherent referential reading (Bosque & Gutiérrez-Rexach 2011: 449). In the *Nueva gramática de la lengua española. Manual* it says the following:

> Se emplea cuál para pedir que se precise la referencia de alguna expresión nominal introducida en el discurso precedente […] o bien para solicitar información acerca del elemento o los elementos que deben seleccionarse de algún conjunto, a veces expresos […], pero otras tácitos.76


In choosing to use only these kind of *wh*-phrases, I assured the same level of d-linkedness as is the case with CLLDeD elements. In order to account for another core property of CLLD, namely the clitic itself, five of the seven complex *wh*-phrases were doubled by the corresponding DO or IO clitic pronoun. In the following example (7.1), representing Item_2.03, the d-linkedness is established by the previous context and the *wh*-phrase is doubled by a clitic.

(7.1) Item_2.03:

**CONTEXT:** Your are looking with your girlfriend at a photo of some Italian friends who live in Sicily. She knows that one of your friends had visited one of the Italians and she asks you:

¿A cuál de estas chicas tu amiga la/le había visitado en Sicilia?

‘Which of these girls had your friend Cl.do had visited in Sicily?’

A further characteristic of the clitic had to be taken into account. As described in Section 6.1.3, Spanish speakers in the Basque country are strong *leístas*. That is, for [+human] DO, they strongly prefer the use of the IO clitic le/les instead the corresponding feminine or masculine DO clitic lo(s)/la(s) (Camus Bergareche & Seibane 2012: 13, Rodríguez-Ordóñez 2017: 327). However, Spanish speakers from other regions would use less *leísmo*. Therefore, I created two different versions of Q2: one for the Basque speaking area and one for the rest of Spain. In the Basque version, I took into consideration the *leismo* phenomena. If the DO *wh*-phrase was [-human], I used the corresponding standard DO clitic lo(s)/la(s), if it was [+human], the IO clitic le(s). An example of the former case is shown in (7.2), for the latter in (7.9).77

---

76 *Cuál* is used to ask to precise the reference of some nominal expression introduced previously in the discourse […] or to ask for information about the element or the elements which should be selected out of some set, sometimes explicitly […], sometimes implicitly known. (my own translation)

77 For the sentences with a different clitic, I will always display both versions in the same sentence in this work. The speakers, of course, saw either the standard version with the DO clitic or the *leísta* version with the IO clitic.
7.4 Results

The overall results from Q1 showed hardly any differences between the speaker groups, as I will describe in Section 7.4.1. The second questionnaire gave much more promising results, which I will detail in Section 7.4.2. In Section 7.4.3, I will combine the results of both questionnaires and expand the analysis.

I briefly want to recall the main properties of the questionnaires. The participants were asked to evaluate sentences with or without a short context on a running scale from 1 to 101. The more natural the sentences seemed to them, the higher the value should be on the scale. The questionnaires contained items and fillers. In the following analysis, I will only consider the items. I will, however, speak briefly on the fillers in Chapter 8. The speakers are separated into seven speaker groups, as shown in Table 7.1. Due to the small number of speakers in the two groups 'BS-ce' (in total: 6) and 'SB-cl' (in total: 2) in both questionnaires (5 and 1 in Q1, 5 and 2 in Q2, respectively), I will not include them in the analysis. Also the statistical results of the group of early sequential Spanish-Basque speaker (SB-ce) should be interpreted with caution, as there were only 10 participants in this category.
Both questionnaires have the same design. They have both a within-subject design, which means every participants had to judge every item. The main independent variable is the speaker group. Secondary independent variables can be divided into two groups, namely sociolinguistic and linguistic factors. The sociolinguistic factors are ‘age’, ‘place of origin’, ‘education’, ‘comfort’, ‘percentage of daily speech’ in each language, and ‘language spoken at home’. The linguistic factors are [±insertion], [±human] wh-phrase, DO or IO wh-phrase, and [±clitic]. All these factors are fixed effects. The two random effects are the participants and the items. That means, I understand my participants to be a random sample of all possible speakers with the same sociolinguistic profile and the items as a random sample of all sentences with a similar structure. Thus, I will be able to generalize to all possible speakers and all similar sentences. This information will be important in the statistical analysis and interpretation in this section.

I used two main statistical tools. First, the Kruskal-Wallice Rank Sum test to compare non-normally distributed samples. The corresponding function in R is called \texttt{kruskal.test(…)} (KWT) from the R-package ‘stats’ (version 3.4.3). Second, I calculated a linear-mixed-effects regression model with the formerly described fixed and random effects, allowing for random adjustments. P-values were calculated using the Satterthwaite approximation in the R-package ‘lmerTest’ (version 2.0-36). The corresponding \texttt{R}-function is called \texttt{lmer(…)} (lmer). The main statistical parameters for significant differences are the median\textsuperscript{78} (med) and the \textit{p-value}. If the latter is under 0.05, I assume a statistically significant difference between the compared parameters. In order to describe the data adequately, I will often use boxplots (BP). An example of how to interpret it is given in Figure 7.14.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{boxplot_example}
\caption{Boxplot – Example}
\end{figure}

\textsuperscript{78} My data is not normal distributed, therefore the mean value is not very meaningful. This is the reason why I will generally use the median (med) as an indicator for the acceptance rate. The median is the middle value, if all the single values are arranged according to size. It is therefore much less affected by outlier.
7.4 Results

One note of importance before the analysis: in the coming sections, I will speak often about the acceptance rate of an item. It is therefore necessary to establish a few standards. If the speakers indicate an acceptance value below 20%, I assume that they consider this sentence unnatural or very forced. If it is above 80% on the other hand, I assume that the speakers consider the sentence natural. Furthermore, acceptation rates of 20-40% will be denoted as ‘marginal’, rates of 40-60% ‘acceptable’, and rates of 50-80%, as ‘good’. Though this is a very broad categorization, it will suffice as the starting point.

7.4.1 Results – Q1

The first questionnaire shows very diverse answers. In 10 of the 12 items, the range of the acceptance rate in every speaker group extends from 1 to over 90, often even up to 101. The other two items show a strong rejection by all groups. Figure 7.15 exemplifies a case of a widespread acceptability range, spanning from totally unacceptable to natural. Figure 7.16 shows a case of near unanimous rejection by all groups.

The ten items can be further divided in two groups according to their overall acceptance. The first consists of Item_1.02 and Item_1.07, which both have an overall median around 50% (\(\text{med}_{1.02} = 46\), \(\text{med}_{1.07} = 50\)). That means, more than half of the participants rated these questions as at least ‘good’, over one quarter even as ‘natural’. These distributions are shown in Figure 7.17 and Figure 7.18.
The second group consists of the items 1.01, 1.03 - 1.05, and 1.08 - 1.11, shown in the following eight figures (Figure 7.19 - Figure 7.26). They received overall median values between 15% and 36%. The median per speaker group lies around or below 50% for almost all items. Nevertheless, there is no clear tendency discernible within the groups due to the widely scattered values.
The remaining two items, namely Item_1.06 and Item_1.12, were judged very low by almost all speakers. The overall median values are 4 and 3 and the third quantile is 13 and 10, respectively. That
means, almost everybody reject these two questions as unnatural. The two graphics in Figure 7.27 and Figure 7.28 show the corresponding distributions. They differ from the former items in that no judgment is higher than 40%, except very few outliers.

![Figure 7.27: Item_1.06](image1.png) ![Figure 7.28: Item_1.12](image2.png)

Acceptance rate across all speaker groups

In general, there are no statistically significant differences between the speaker groups, apart from the two items 1.05 and 1.11. That is of special interest, since those two questions are the only items comparable to the items in Q2. As can be seen in (7.3) and (7.4), both are wh-questions with the DO wh-phrase and a lexical subject interfering between the heavy wh-phrase and the verb. The only difference is that the wh-phrases are not doubled by a clitic, while in Q2 they always are.

(7.3) Item_1.05:

**CONTEXT:** A friend of yours, who likes your sister, asks you about the classes she wants to enroll this summer.

¿A cuál de los cursos de cocina tu hermana se apuntará este verano?

to which of the classes of kitchen your sister Cl enroll this summer

‘To which of the cooking classes will your sister enroll this summer?’

(7.4) Item_1.11:

**CONTEXT:** One roommate to the other about Maria, who lives with them.

¿A quién de sus dos amantes María ha invitado a cenar esta noche?

A Which of her two lovers María has invited to dine this evening

‘Which of her two lovers has María invited to dinner this evening?’

The first question shows significant differences between the 2L1+BS-cl and M-SP. Running the KWT returns a p-value of 0.0206 when comparing the two groups. That means the 2L1 and BS-cl judge the questions significantly less natural (med_{1.05} = 17) than monolingual Spanish speakers from outside the Basque country (med_{1.05} = 24). The second question does not show significant difference, it only
approaches significance when comparing 2L1 and M-SP ($p_{KWT} = 0.0857$), or all speakers from the Basque speaking area (All-BA) and M-SP ($p_{KWT} = 0.0798$). Although it is only a single item that indicates a significant difference while the other only approaches significance, I will detail in the next section how questions with similar structures show the same tendencies.

The overall results from Q1 are not very transparent. They do not indicate clear tendencies neither within the groups themselves nor between groups. I will now turn to Q2, which shows much more promising results.

7.4.2. Results – Q2

The second questionnaire was shorter and tested less parameters. The items can be divided into 3 groups, of which the first is just one item: a wh-interrogative with por qué ‘why’, wh-word-verb adjacency and a realized pronominal subject. The sentence is shown in (7.5).

(7.5) Item_2.01:

\[CONTEXT: \text{You know that your colleague from work arrived late at a conference the day before. Now you want to know the reason.}\]

¿Por qué viniste tú tarde ayer para la conferencia?

‘Why were you late for the conference yesterday?’

The overall median is $\text{med}_{2.01} = 24$ and the overall 3rd Quantile is under 40. That means, most of the speakers in all groups find that question forced or unnatural.

The remaining six items are all complex wh-interrogatives with the wh-phrase cuál(es) de ‘which of’ that is doubled by a clitic. I divided them into two groups, depending if wh-phrase-verb adjacency applies or not. The two items 2.04 and 2.07 exhibit adjacency, whereas the items 2.02, 2.03, 2.05, and 2.08 do not. The insertions in the latter cases are always lexical subjects. The complex wh-phrase is the DO in all interrogatives except Item_2.02, in which it is the IO. An overview of all parameters is given in Appendix C.2.

I will first describe the data for the two sentences in the first group. Both wh-phrases function as DO that differ with respect to the feature [+human], as can be seen in (7.6) and (7.7), respectively.
(7.6) Item_2.04:

**CONTEXT:** Maria and Ana are talking about Maria’s son Juan and his journey to England last summer. They have family there. Ana, who knows the whole family very well, asks:

¿A cuál de las primas que viven en Inglaterra le visitó Pablo para mejorar su inglés durante el verano pasado?

‘Which of his cousins, that live in England, did Pablo visit last summer to improve his English?’

(7.7) Item_2.07:

**CONTEXT:** Isabel’s parents have prepared a list with all the songs she knows how to play on the piano. At her grandparent’s birthday, they see the list and ask her mother:

¿Cuáles de estas canciones maravillosas va a tocar Isabela esta noche?

‘Which of these wonderful songs will Isabella play tonight?’

As discussed in Section 7.3.3, if the DO is [+human], the clitic used differs. The participants from the Basque speaking area were presented the IO clitic le, whereas the speakers from the rest of Spain received the DO clitic la (cf. (7.6)). In the case of a [-human] DO, the standard DO clitic las was used for all speakers (cf. (7.7)). Both sentences exhibit an interesting distribution between all speakers from the Basque speaking area (All-BA) and the M-SP. Figure 7.29 and Figure 7.30 show the acceptability rates for all speaker groups on the left and for the two groups All-BA and M-SP on the right.

![Acceptance rate for Item_2.04](7.29a) All speaker groups

![Acceptance rate for Item_2.04](7.29b) All-BA vs. M-SP

Figure 7.29: Acceptance rate for Item_2.04
M-SP rate both sentences almost identical (M-SP: med$_{I_2.04} = 24$, med$_{I_2.07} = 26.5$, $p_{KW} = 0.99$), whereas the other speakers differ significantly between the two items (All-BA: med$_{I_2.04} = 61$, med$_{I_2.07} = 12$, $p_{KW} = 3.25 \cdot 10^{-8}$). Looking at the two questions separately, both show a significant differences between the M-SP and 2L1. As can be seen in Figure 7.29a, the 2L1 speakers show a high acceptance rate for Item$_{2.04}$ (med$_{I_2.04} = 68$, 3rd Quantile$_{I_2.04} = 95.25$), and a very low rate for Item$_{2.07}$ (med$_{I_1.07} = 5$, 3rd Quantile$_{I_1.07} = 32$). The M-SP, however, show exactly the opposite acceptability. In Item$_{2.04}$, they accept the sentence significantly less than the 2L1 ($p_{KW} = 0.004$), whereas in Item$_{2.07}$, they do so significantly more ($p_{KW} = 0.03$). This is, however, the only significant difference observable. The two groups represent the two ends of the acceptability range in both questions. We can corroborate this by looking at the medians for the other speaker groups, especially the M-BA. The distribution in both items is surprisingly similar, in the way that the M-BA always lie almost exactly in the middle of the 2L1 and M-SP. In Item$_{2.04}$, the M-BA’s median lies by 43 and in Item$_{2.07}$, the median of the M-BA is at 12.0. The last point to mention is that the total acceptance for Item$_{2.07}$ is very low in all groups. The overall median lies at 15, and the 3rd Quantile only at 45.

In the second group of heavy wh-sentences with cuál de ‘which of’, all questions have the intervening subject between the wh-phrase and the verb in common. The wh-phrase functions as DO in four cases and as IO in Item$_{2.02}$. Again, the complete properties can be consulted in Appendix C.2.

Only Item$_{2.05}$ is rejected by the vast majority of all speakers (med$_{I_2.05} = 7$, 3rd Quantile$_{I_2.05} = 22$). It is the only item in this second group with a [-human] DO as wh-phrase which is doubled by the DO clitic la. The sentence is shown in (7.8) and its distribution for all speaker groups in Figure 7.31.
Item_2.05:

**CONTEXT:** The morning of her wedding just before the ceremony, Elena, is getting ready with the help of two friends. The two pass to the other side of the room where the flowers lay prepared for the bride and her maids of honor. One of them asks:

¿Cuál de estas flores Elena llevará en el pelo durante la ceremonia?

‘Which of these flowers will Elena wear in her hair during the ceremony?’

---

![Figure 7.31: Item_2.05 – Acceptance rate across speaker groups](image)

Item_2.03, which can be seen in (7.9), also has an overall low acceptance rate. (med$_{2.03} = 19$, 3$^{rd}$ Quantile$_{2.03} = 54$). The distribution is shown in Figure 7.32. The wh-phrase is a [+human] DO and the intervening subject is a full DP. It is a very similar construction as Item_2.08. We will see, however, that the latter item exhibits much clearer difference in the acceptance rate.

---

Item_2.03:

**CONTEXT:** You are looking with your girlfriend at a photo of some Italian friends who live in Sicily. She knows that one of your friends had visited one of the Italian girls and she asks you:

¿A cuál de estas chicas tu amiga la/le había visitado en Sicilia?

‘Which of these girls had your friend visited in Sicily?’
The two remaining items 2.02 (7.10) and 2.08 (7.11) show the clearest contrast between groups. If we have a look at their distribution in Figure 7.33 and Figure 7.34, we can already see the different acceptability rates.

(7.10) Item_2.02:

**Context:** At Juan’s wedding, his friends have prepared a game in order to see how well his new wife knows him. They ask her questions about his childhood and teenage years. For example:

¿A cuál de sus profesores de primaria Juan le dedicó un poema?

to which of his teachers of primary Juan Cl. dedicated a poem

‘Which of his primary school teachers did Juan dedicate a poem to?’

(7.11) Item_2.08:

**Context:** Yesterday, several swimmers were late for the medical examination. Since the doctor didn’t have much time, he only saw a few of them. The following day, the trainer wants to know:

¿A cuáles de las que llegaron tarde el médico las/les examinó durante más de cinco minutos?

‘Whom of you that came late did the doctor examine for more than five minutes?’
The differences can be seen especially between the M-SP on the one side and the 2L1 and BS-cl on the other side (Item_2.02: $p_{KWT} = 0.01$, Item_2.08: $p_{KWT} = 1.59 \times 10^{-5}$). Also, the difference between the combined monolingual speaker and the bilinguals is significant for these two items (Item_2.02: $p_{KWT} = 0.033$, Item_2.08: $p_{KWT} = 0.00063$).

I have now presented the data from the two questionnaires separately. The following section will combine the two as much as is possible.

### 7.4.3. Analysis Q1 + Q2

Most of the items in Q2 differ significantly in their structure from the items in Q1. Nevertheless, as already mentioned in Section 7.4.1, the two items Item_1.05 and Item_1.11 ((7.3) and (7.4), respectively) are very similar to the items from Q2. Both are complex wh-questions with the subject intervening between the wh-phrase and the verb. The first item has a [-human] DO wh-phrase, whereas the second has a [+human] DO wh-phrase. The only structural difference with the items in Q2 is the absence of a clitic doubling the complex wh-phrase. The basis for further analysis will therefore be two items from Q1 (1.05 and 1.11) and six items from Q2 (2.02 - 2.05, 2.07 - 2.08). These eight items will now be analyzed according to the sociolinguistic and linguistic factors presented in the introduction of this section.

In order to get a general overview for the items with an insertion ([+insertion]), Table 7.2 summarizes the results of a linear-mixed-effects regression model (lmer) to determine significant differences between the speaker groups.

---

The number of asterisks indicate the level of significance.
We see that, on average, the M-SP show the highest acceptance rates for the items with insertions ($\bar{\Omega}_{M-SP} = 45.1$). They differ significantly from all other speaker groups except from M-BA and SB-ce. Nevertheless, even compared to these groups, the M-SP rate the items an average of 7% and 11% higher, respectively. The largest difference, however, exists between the bilinguals and the M-SP. As can be seen in Table 7.2, the average differences vary between 12.7% and 15.3%. The M-SP and the 2L1 define the two extremes of the acceptability judgments. The 2L1, on average, show the lowest acceptance for these constructions, being followed closely by the BS-cl. Furthermore, all the bilinguals combined show a big difference compared to both the M-SP as well as all the monolinguals together (M-SP/BA). The M-BA also provide some very interesting data. They do neither differ significantly from the M-SP nor from the bilingual groups, but lie almost directly between these two. This can be seen by looking at the average difference between the M-BA and the bilinguals on the one hand, and the M-SP and the bilinguals on the other. The first value is almost exactly half of the second. A final observation is that even if all speakers from the Basque speaking area are taken together and then compared to the M-SP, the difference is still significant. The M-SP still rate the questions with insertions about 12% higher than the All-BA.

The results from the Lmer-model presented in Table 7.2 only takes the speaker groups as fixed effect into consideration. In order to account for possible influences from other sociolinguistic or linguistic factors, one needs to look at further tests. The sociolinguistic factors that could have had an influence as fixed effects are ‘age’, ‘education’, ‘comfort in the languages’, and ‘percentage of the daily language...
use’. The latter two effects were of course only applied to bilingual speakers. None of these mentioned factors, however, showed any significant differences between the groups. This would indicate that the only significant sociolinguistic factor is the speaker group.

The linguistic factors used in the analysis were the following:

\[(7.12)\]
\[i) \ [+\text{clitic}]: \text{wh-phrase doubled by a clitic or not}\]
\[ii) \ DO \text{ or IO } \text{wh-phrase}\]
\[iii) \ [+\text{insertion}]: \text{all cases of insertions are the subject}\]
\[iv) \ [+\text{human}] \text{DO } \text{wh-phrase}\]

The medians for all items across all the speaker groups divided by the linguistic factors are shown in Table 7.3.\(^\text{86}\) The medians confirm the overall differences between the speaker groups already shown in Table 7.2. Of further interest, however, is to now look at each speaker group and determine the influence of each of the factors as well as some combinations thereof.

Table 7.3: Median of all relevant items divided by linguistic factors

<table>
<thead>
<tr>
<th>Group/Item</th>
<th>[+clitic]</th>
<th>[-clitic]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[+insertion]</td>
<td>[-insertion]</td>
</tr>
<tr>
<td>2L1</td>
<td>42.5  12  37.5  6</td>
<td>68  5</td>
</tr>
</tbody>
</table>

In order to analyze the first factor, it is necessary to distinguish between the questions with and without \text{wh-phrase-verb adjacency}. As can be seen in the table, I have not included the combination [-insertion, -clitic] in the questionnaire. This combination would be the canonical \text{wh-interrogative word order} and one can therefore presume that it is well accepted by all speakers. If we would compare such a canonical sentence with a high acceptance rate to the rates given for the cases with a clitic, I deduct that the difference between the presence or absence of a clitic would be significant. The differences between the speaker groups were already analyzed in detail in the previous section. In a nutshell, the

---

\(^\text{86}\) The second linguistic factor is not indicated in the table, because only Item_2.02 has an IO \text{wh-phrase}. All others \text{wh-phrases} are DOs.
7.4 Results

Results are that All-BA and the M-SP differ significantly in Item_2.04, but not in Item_2.07 ($p_{2.04,kwt} = 0.005$, $p_{2.07,kwt} = 0.097$). In addition, the M-SP rate both items almost identically ($\text{med}_{2.04} = 24$, $\text{med}_{2.07} = 26.5$, $p_{kwt} = 0.99$), whereas the All-BA show a significant difference in their acceptance in the two cases ($\text{med}_{2.04} = 61$, $\text{med}_{2.07} = 12$, $p_{kwt} = 3.25 \cdot 10^{-8}$). Since the only distinction between these two questions and a standard complex wh-question is the presence of the clitic, it is not too far fetched to assume that such a question would be rated, by all speakers, significantly higher without a clitic. The difference should be especially big for the M-SP. Therefore I deduct that for complex wh-phrases that exhibit wh-phrase-verb adjacency, the presence or absence of the clitic is a significant factor.

For the questions without wh-phrase-verb adjacency, the situation changes. Table 7.4\textsuperscript{81} gives an overview across all speaker groups of the differences between questions with and without a clitic. We can clearly see that in the wh-questions with an insertion, the presence or absence of a clitic has a significant influence on the acceptance rate for all monolinguals and especially the M-SP. All the bilinguals combined also show significant difference concerning [+clitic] in the first configuration [+insertion,+human], although each group alone does not.

<table>
<thead>
<tr>
<th>Group</th>
<th>Item_2.08 $p_{kwt}$</th>
<th>Item_1.11 $p_{kwt}$</th>
<th>Item_2.05 $p_{kwt}$</th>
<th>Item_1.05 $p_{kwt}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2L1</td>
<td>0.084</td>
<td>0.361</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BS-cl</td>
<td>0.362</td>
<td>0.282</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2L1 + BS-cl</td>
<td>0.055</td>
<td>0.168</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SB-ce</td>
<td>0.205</td>
<td>0.099</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M-BA</td>
<td>0.196</td>
<td>0.022 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M-SP</td>
<td>0.004 **</td>
<td>0.007 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biling.</td>
<td>0.022 *</td>
<td>0.084</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M-SP/BA</td>
<td>0.004 **</td>
<td>0.0004 ***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The analysis with the lmer-test for the second factor, a DO and IO wh-phrase, is shown in Table 7.5. The data indicates that the DO-IO-variation only is highly significance for the M-BA. For all other groups, the p-value indicates a non-significance.

The fourth linguistic factor has the highest impact on the acceptability. In other words, the variation between a [+human] object and a [-human] object as wh-phrase influences the acceptance of the

\textsuperscript{81} The number of asterisks indicate the level of significance and ‘.’ means ‘approaching significance’.
Table 7.5: Lmer – summary acceptability and mean difference of DO vs. IO wh-phrase across all speaker groups.

<table>
<thead>
<tr>
<th>Group</th>
<th>Plmer</th>
<th>Lv. of Signif.</th>
<th>Ø DO</th>
<th>Ø diff. DO vs. IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>2L1</td>
<td>0.485</td>
<td>—</td>
<td>32.2</td>
<td>+14.8</td>
</tr>
<tr>
<td>BS-cl</td>
<td>0.319</td>
<td>—</td>
<td>34.9</td>
<td>+16.2</td>
</tr>
<tr>
<td>2L1 + BS-cl</td>
<td>0.105</td>
<td>—</td>
<td>33.3</td>
<td>+15.3</td>
</tr>
<tr>
<td>SB-ce</td>
<td>0.437</td>
<td>—</td>
<td>32.9</td>
<td>+41.7</td>
</tr>
<tr>
<td>M-BA</td>
<td>0.005</td>
<td>**</td>
<td>42.0</td>
<td>+20.8</td>
</tr>
<tr>
<td>M-SP</td>
<td>0.209</td>
<td>—</td>
<td>52.5</td>
<td>+14.6</td>
</tr>
<tr>
<td>Bilinguals</td>
<td>0.429</td>
<td>—</td>
<td>33.2</td>
<td>+19.0</td>
</tr>
<tr>
<td>M-SP/BA</td>
<td>0.523</td>
<td>—</td>
<td>47.9</td>
<td>+17.3</td>
</tr>
</tbody>
</table>

questions significantly. The two sentences without insertion, Item_2.04 and Item_2.07, only differ parametrically in the [±human] factor in the DO wh-phrase. As already discussed earlier, the sentence with a [+human] DO wh-phrase is rated significantly higher by All-BA, but not by the M-SP. The latter group rates both questions equally low. The difference between the two groups disappears when looking at the second item with the [-human] DO wh-phrase. It follows that in the case of wh-phrase-verb adjacency combined with an object clitic, the factor [±human] has a significant influence for the All-BA, but not for the M-SP.

The questions with insertion indicate a different distribution. If we include ‘insertion’ and ‘human’ as fixed effects, the lmer-test returns the results shown in Table 7.6.

Table 7.6: Linear-mixed-effects regression model – Summary for [±human, ±insertion]

<table>
<thead>
<tr>
<th>Groups</th>
<th>[+human] [+insertion]</th>
<th>[-human] [+insertion]</th>
<th>Difference</th>
<th>Plmer</th>
</tr>
</thead>
<tbody>
<tr>
<td>2L1</td>
<td>37.6</td>
<td>14.7</td>
<td>-22.9</td>
<td>0.029 *</td>
</tr>
<tr>
<td>BS-cl</td>
<td>37.6</td>
<td>14.6</td>
<td>-23.0</td>
<td>0.029 *</td>
</tr>
<tr>
<td>2L1+BS-cl</td>
<td>37.6</td>
<td>17.2</td>
<td>-20.1</td>
<td>0.021 *</td>
</tr>
<tr>
<td>SB-ce</td>
<td>45.1</td>
<td>11.1</td>
<td>-34.0</td>
<td>0.077 .</td>
</tr>
<tr>
<td>M-BA</td>
<td>45.2</td>
<td>21.9</td>
<td>-23.3</td>
<td>0.057 .</td>
</tr>
<tr>
<td>M-SP</td>
<td>51.4</td>
<td>32.3</td>
<td>-19.1</td>
<td>0.189</td>
</tr>
<tr>
<td>Bilinguals</td>
<td>38.7</td>
<td>15.9</td>
<td>-22.8</td>
<td>0.029 *</td>
</tr>
<tr>
<td>M-SP/BA</td>
<td>48.7</td>
<td>27.8</td>
<td>-20.9</td>
<td>0.093 .</td>
</tr>
</tbody>
</table>

The p-values and the relative differences show that questions with an insertion are much less accepted if the wh-phrase is a [-human] object. The acceptability ratings of all speakers except the M-SP are affected by the type of object, at least to a certain degree. In other words, if the wh-phrase-verb adjacency is interrupted by a subject, the sentence is accepted more easily when the wh-phrase is a
[±human] object. On the other hand, if the wh-phrase is a [-human] object, the speakers do not differ significantly when altering between the presence or absence of an insertion. That means, for all groups, the deciding factor is [±human], but only if an insertion is present.

In summary, the data reveals clear differences between several of the speaker groups. Especially the monolinguals from outside the Basque-speaking area accept the questions with insertions significantly more than the bilinguals. Apart from the ‘speaker group’, the other sociolinguistic factors do not seem to have any influence. The linguistic factors, in the contrary, do significantly influence the acceptability. The main factors, ‘[±human] DO’, ‘[±clitic]’, ‘DO or IO wh-phrase’, and ‘[±insertion]’, all have a significant influence on all or at least some speaker groups. In the next chapter, I will discuss these results in more detail, as well as connect them to the theoretical findings from the earlier chapters.
CHAPTER 8

Complex wh-phrases revisited

In the preceding chapters, I have advanced the hypothesis that simultaneous and early bilingual Basque-Spanish speakers differ from other Spanish speakers in their acceptance of non-wh-phrase-verb adjacency in complex wh-interrogatives. I postulated that due to the combination of superficial resemblance and distinct underlying structures, the former group analyzes Spanish interrogatives the same way they analyze them in Basque. This means they do not accept insertion between any kind of wh-phrases and the verb, no matter the complexity.\(^{82}\) I argued that, in Spanish, wh-phrases normally move to SpecFinP as a consequence of Q-movement. Complex wh-phrases, however, differ in that they first move to DS, a higher structural position in the left periphery. With that, they behave in the same way as CLLDe-d constituents. In a second movement, the Q-particle then moves to ForceP, taking along the wh-phrase.

In Basque however, in a standard affirmative wh-interrogative, the wh-phrase always moves to the preverbal focus position. It does not matter if the wh-phrase is simple or complex. As explained in detail in Section 5.2, the landing position of the wh-phrase can change however, if some other element, for example the negative particle ez, occupies this position. In these cases, the wh-phrase moves up higher in the hierarchy to DS. I argued that the reason for this movement is not Q-movement, but a necessity to avoid intervention effects between the Q-particle and the negative particle. These differences between Basque and Spanish derive from two contrasting causes for wh-movement. As elaborated in Section 5.1.1 and 5.1.2, Spanish is a Q-projection and Basque a Q-adjunction language, by Cable’s (2010) definition. Spanish wh-movement is therefore a simple ‘by-product’ of Q-movement, whereas in Basque, the wh-phrase moves independently of the Q-movement in order to satisfy a focus feature. If some other element already occupies this preverbal focus position, the wh-phrase only has to move if an intervention effect otherwise arises. In the case of complex wh-phrases in Spanish, the first movement

\(^{82}\) As said in Chapter 3 and 4, adverbial wh-phrases like Sp. por qué, Bq. zergatik 'why' present an exception to this observation. With these wh-phrases, insertions are allowed in both languages.
to the DS is also independent of the Q-movement and only the second movement to ForceP is caused by Q-movement. The result of this structural difference is that, in these last cases, other constituents can surface between the complex wh-phrase and the verb. That is not possible in Basque.

Returning to the hypothesis stated above, it relies on the fact that almost all other interrogative structures have the same superficial word order in the two languages. Both normally exhibit wh-phrase-verb adjacency with the same exceptions in the case of some adjunct wh-phrases. In Basque as well as in Spanish negative wh-interrogatives, the negative particle Bq. ez, Sp. no ‘not’ appears pre-verbally, directly following the wh-phrase (cf. for Spanish: Contreras 1999: 1939, for Basque: Etxepare & Ortiz de Urbina 2003: 503). In indirect questions, the adjacency rule also normally applies, and both languages exhibit obligatory successive cyclicity (cf. for Spanish: Torrego 1984, for Basque: Etxepare & Ortiz de Urbina 2003: 484-494). There are only two major exceptions in wh-interrogatives in which Spanish and Basque differ significantly in their linear constituent order. These are multiple wh-questions and complex wh-interrogatives, as elaborated in detail in the Section 3.2 and 4.1.2. I claim, however, that the distinct underlying structures only influence Spanish complex wh-questions and not multiple wh-questions. The data presented supports this claim.

In Section 5.1.1 and 5.1.2, I have shown that multiple wh-questions in Spanish exhibit superiority effects but no intervention effects and in Basque vice versa. That means, that in Spanish the linear order of the multiple wh-words is significant, whereas in Basque it is not. Other focus sensitive operators, however, like the negative particle ez, no ‘not’, can occupy a position between the multiple wh-phrases in Spanish, but not in Basque. Both languages therefore fulfill the prediction in (5.7) made by Cable (2010) that superiority and intervention effects are mutually exclusive. Examples (8.1) - (8.4) illustrate these relations.

(8.1) a. Cuándo compró Pepe qué?
    when bought.3sg Pepe what
    ‘What did Pep buy when?’

b. *Qué compró Pepe cuándo?

(8.2) A quién no ha dado Juan qué?
    a who not have.3sg gave Juan what?
    ‘To whom didn’t Juan give what?’

(8.3) a. Nork ekarri du zer?
    who.ERG bring AUX what

b. Zer ekarri du nork?
    what bring AUX who.ERG
    ‘*Who brought what?’
According to Cable (2010), the explanation for these differences lies in the different number of Q-particles, as detailed in Section 5.1. Spanish does not exhibit intervention effects because each wh-phrase comes along with its own Q-particle. One Q-particle is fronted, which satisfies the respective [Q] feature in the C-head. Thus, no other Q-particle has to be fronted and as a result, multiple wh-movement is not allowed. In Basque meanwhile, only one Q-particle is generated no matter how many wh-phrases are present, and this is why Basque is sensitive to intervention effects. Therefore, in the case of an intervening operator like the negative particle ez, all wh-phrases have to be fronted. In this way it is assured that the first focus sensible operator c-commanding the wh-phrases is the Q-particle and not the negative particle. In an affirmative multiple wh-phrase however, the multiple fronting is facultative, since the fronting does not depend on Q-movement. Only one wh-phrase has to be fronted to satisfy the focus feature in the preverbal focus position. Hence, the cause of the different behavior in multiple wh-questions is an independent structural distinction, namely the different number of Q-particles.

Turning to the data, especially Filler_1.24 is of interest here. The sentence is shown in (8.5a) along with a similar construction extracted from CORPES XXI in (8.5b). The three valid variations of a similar Basque question from Etxepare & Ortiz de Urbina (2003: 500) are shown in (8.5c).

\[(8.4)\]

a. * Nork ez du zer eros?  
   who.erg not aux what bought  
   ‘Who didn’t buy what?’

b. Nork zer ez du eros?  
   who.erg what not aux bought

\[(8.5)\]

a. Filler_1.24:
   * ¿Dónde dio Carmen a quién qué?  
     where gave Carmen to who what  
     ‘To whom did Carmen give what where?’

b. ¿Quién dice qué a quién?  
   who says what to who  
   ‘Who says what to whom?’

c. i. Nork ekarri dio zer nor?  
   ii. Nork ekarri dio nori zer?  
      who.erg brought aux who.to what  
      ‘Who brought what to whom’
   iii. Zer ekarri dio nork nori?

Although the questions are not exactly the same, they illustrate the argument. In the Basque examples we can see that the order of the wh-phrases is free. In Spanish, however, it is not. The postverbal
wh-phrases have to be in the canonical constituent order. If a bilingual Basque-Spanish speaker would also be influenced in other wh-interrogative structures, a difference in the acceptance in these multiple wh-questions would be a strong evidence. However, looking at the data, the respective items do not show any significant difference between any speaker groups. For example, comparing the acceptability of multiple questions using the lmer test returns no significant differences (M-SP vs. 2L1+BS-cl: \( p_{\text{Lmer}} = 0.91 \), monolinguals vs. bilinguals: \( p_{\text{Lmer}} = 0.17 \)). The data shows the same results for all other wh-interrogatives, apart from the complex wh-questions. If we have a brief look at the five items with the wh-phrase *por qué*, we get a similar result. The first four items in Q1 and the first item in Q2 are wh-interrogatives with *por qué*. Running the lmer test returns also that the ratings do not differ significantly (for example, monolinguals vs. bilinguals: \( p_{\text{Lmer}} = 0.826 \)). The same non-significance also holds for the other wh-questions among the filler phrases. I conclude from this data, that the only interrogative structures in which the speaker’s acceptability differs significantly are the ones with a complex wh-phrase. However, reality is never as uniform as theory would suggest, and instead offers some more subtle differences. These will be discussed subsequently.

In order to create a clear framework, I will first discuss the results of the complex wh-phrases that exhibit wh-phrase-verb adjacency and venture an explanation for the observed variation. Then, I will briefly recall the results for the questions without adjacency, and follow with a possible explanation.

The two items exhibiting wh-phrase-verb adjacency are 2.04 and 2.07, here repeated in (8.6) and (8.7). As discussed in Section 7.4.2, the only parametrical difference between the two sentences is the type of DO wh-phrase. In Item_2.04, it is [+human], in Item_2.07 [-human]. As discussed in detail in Section 6.1.3 and 7.3.3, the questionnaires differed in this, depending on which group it was for. For the M-SP, I used the standard DO clitics *la(s)* in both sentences. For the All-BA, I used the IO clitic *le* for the [+human] DO to account for the widespread leísmo in the Basque-speaking area and the standard DO clitic *las* for the [-human] DO wh-phrase.83

(8.6) Item_2.04:

**Context:** Maria and Ana are talking about Maria’s son Juan and his journey to England last summer. They have family there. Ana, who knows the whole family very well, asks:

¿A cuál de las primas que viven en Inglaterra la/le visitó Pablo para mejorar su inglés durante el verano pasado?

‘Which of his cousins, that live in England, did Pablo visit last summer to improve his English?’

---

83 For the sentences with a different clitic, I will always display both versions in the same sentence in this work. The speakers, of course, saw either the standard version with the DO clitic or the leísta version with the IO clitic.
Item_2.07:

**Context:** Isabel’s parents have prepared a list with all the songs she knows how to play on the piano. At her grandparent’s birthday, they see the list and ask her mother:

¿Cuáles de estas canciones maravillosas va a tocar Isabela esta noche?  
Which of these beautiful songs will Isabela play tonight?

We have already seen in Section 7.4.3 that the All-BA show a conspicuous distribution compared to the M-SP. The acceptability rates are also repeated here in Figure 8.1 and 8.2.

![Figure 8.1: Item_2.04](image1.png)  
**Acceptance rate for the items with wh-phrase-verb adjacency, divided into All-BA and M-SP**

![Figure 8.2: Item_2.07](image2.png)

The interesting observation here is that the M-SP rate both sentences almost identically. Their acceptance is, for the vast majority of speakers, under 50%. I deduct from this that, for M-SP, the [±human] DO dichotomy does not influence their acceptability. The question of course is why the All-BA differ so significantly from the M-SP in the first, but not in the second item. In Section 6.1.3 I presented, among others, the following differences in the use of clitics in the Spanish of the Basque-speaking area: the use of the IO clitic pronoun le as an agreement marker for a [+human], definite DO. I propose that this particular clitic use is the cause for the observed difference in the two sentences. In Item_2.04, the DO wh-phrase is [+human] and definite, a combination that favors a ‘non-canonical’ redoubling of the DO with an IO clitic pronoun to express agreement. In Item_2.07, on the other hand, the DO is [-human] and therefore does not need this kind of agreement. Thus, the clitic doubling sound deviant for most of the All-BA. We can even refine the analysis by having a closer look at the distribution within this group. The graphics in Figure 8.3 and 8.4 show that the M-BA and the bilingual speakers also differ in their acceptance rates.
It is obvious that the second question sounds at least marginal to most speakers. In the first question, however, the M-BA range lies almost exactly midways between the bilinguals and the M-SP (M-SP: med$_{I_2.04}$ = 24, M-BA: med$_{I_2.04}$ = 43, bilinguals: med$_{I_2.04}$ = 61.5). This distribution allows the following theoretical proposal: The M-BA are, up to a certain degree, influenced by the bilingual speakers’ use of the IO clitic pronoun. Some of the monolingual speakers apparently adapt or at least accept the particular use of the IO clitic le in the case of a [+human], definite DO as an agreement marker. The majority of the M-SP do not. Furthermore, the distribution in Item 2.07 supports the hypothesis that the doubling is bound to a definite, [+human] DO and is not triggered by a [-human] DO. In this case, the bilinguals and the M-BA even give a significantly lower acceptability judgment than the M-SP. Furthermore, if we once again assume that the standard word order wh-phrase-verb-subject would be acceptable for all Spanish speakers, we can identify two significant factors that influence the acceptability in wh-questions with wh-phrase-verb adjacency. First, the general absence or presence of a clitic that doubles the complex wh-phrase and, second, in the case of the All-BA, the type of clitic. With this in mind, I will now turn to the items without wh-phrase-verb adjacency.

Let us recall the hypothesis stated at the beginning of this chapter. Simultaneous and sequential Basque-Spanish bilingual speakers express a lower grade of acceptability for complex wh-questions without wh-phrase-verb adjacency than monolingual speakers or sequential Spanish-Basque bilinguals. In other words, in the case of the six complex wh-interrogative from Q1 and Q2, the first two groups should show significantly lower acceptance rates than the other speakers. We have seen in Section 7.4.3 in Table 7.2, here repeated as Table 8.1, that the data supports this hypothesis. Especially the mono-
linguals from outside the Basque-speaking area accept insertions in complex wh-interrogatives more readily than simultaneous and sequential Basque-Spanish bilingual speakers.

Table 8.1: Linear-mixed-effects regression model – Summary of the wh-questions without wh-phrase-verb adjacency for all speaker groups

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Ø acceptability group 1</th>
<th>Ø diff. in accept. group 2 vs. group 1</th>
<th>p_Lmer</th>
<th>Lvl. of Signif.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-SP</td>
<td>2L1</td>
<td>45.1</td>
<td>-15.3</td>
<td>0.0003</td>
<td>**</td>
</tr>
<tr>
<td>M-SP</td>
<td>BS-cl</td>
<td>45.1</td>
<td>-13.2</td>
<td>0.0012</td>
<td>*</td>
</tr>
<tr>
<td>M-SP</td>
<td>2L1+BS-cl</td>
<td>45.1</td>
<td>-14.5</td>
<td>0.0007</td>
<td>***</td>
</tr>
<tr>
<td>M-SP</td>
<td>SB-ce</td>
<td>45.1</td>
<td>-11.1</td>
<td>0.1168</td>
<td>—</td>
</tr>
<tr>
<td>M-SP</td>
<td>Biling.</td>
<td>45.1</td>
<td>-12.7</td>
<td>0.0021</td>
<td>**</td>
</tr>
<tr>
<td>M-SP</td>
<td>M-BA</td>
<td>45.1</td>
<td>-7.7</td>
<td>0.1351</td>
<td>—</td>
</tr>
<tr>
<td>M-SP</td>
<td>All-BA</td>
<td>45.1</td>
<td>-12.4</td>
<td>0.0019</td>
<td>**</td>
</tr>
<tr>
<td>M-BA</td>
<td>SB-ce</td>
<td>37.4</td>
<td>-3.3</td>
<td>0.6406</td>
<td>—</td>
</tr>
<tr>
<td>M-BA</td>
<td>2L1+BS-cl</td>
<td>37.4</td>
<td>-6.7</td>
<td>0.1367</td>
<td>—</td>
</tr>
<tr>
<td>M-BA</td>
<td>Biling.</td>
<td>37.4</td>
<td>-6.3</td>
<td>0.1560</td>
<td>—</td>
</tr>
<tr>
<td>M-SP/BA</td>
<td>2L1+BS-cl</td>
<td>41.8</td>
<td>-11.1</td>
<td>0.0043</td>
<td>**</td>
</tr>
<tr>
<td>M-SP/BA</td>
<td>Biling.</td>
<td>41.8</td>
<td>-10.6</td>
<td>0.0023</td>
<td>**</td>
</tr>
</tbody>
</table>

There are two interesting observations here. The first is the average difference in the acceptance rate of the M-BA. In the case with non-wh-phrase-verb adjacency, the acceptance rate also lies between that of the M-SP and the bilinguals, as it does for the previously discussed sentences with adjacency. The average difference between the M-BA and the bilinguals is almost exactly half of the average difference between the M-SP and the bilinguals (Ø-diff. M-BA vs. bilinguals: 6.3; Ø-diff. M-BA vs. M-SP: 7.7). In other words, the M-BA seems to accept insertions already more readily the bilingual speakers, but still less than the M-SP.

The second observation is the non-significant difference comparing the SB-ce to the other groups. Looking at the medians in Table 8.2, the SB-ce show similar behavior to the M-BA, just slightly more distinctive.84 These two observations support the hypothesis that the acceptance of such structures is already in an advanced state outside the Basque-speaking area, whereas within, the monolinguals and the Spanish-Basque sequential bilinguals are only beginning to accept these constructions, in keeping with the general trend in the Spanish-speaking community. The other bilinguals however, most likely due to a Basque influence, neither use nor accept these structures and with that slow down the overall spread of this phenomenon in the Basque-speaking area. This can be seen to be a further example of a

---

84 In addition, we have to bear in mind that the SB-ce is the smallest speaker group with only 10 participants.
slower spread of an internal language change due to language contact. The examples of Basque-Spanish language contact phenomena given in Section 6.1.3 show that there have been other, similar processes.

I have shown in the previous chapter that while the overall distribution does support my hypothesis, a more detailed analysis reveals more subtle differences. We have seen that several of the linguistic factors have a significant influence on the acceptability, though to varying degrees. The more influential factors are [+human] DO wh-phrase, [+clitic], and [+insertion], and to a lesser degree also the variation between a DO and IO wh-phrase. The complete list of all medians of all eight items examined is shown in Table 7.3, and is here again repeated as Table 8.2.

<table>
<thead>
<tr>
<th>Group/Item</th>
<th>[+clitic]</th>
<th>[-clitic]</th>
<th>[+human]</th>
<th>[-human]</th>
<th>[+insertion]</th>
<th>[-insertion]</th>
<th>[+human]</th>
<th>[-human]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2L1</td>
<td>42.5</td>
<td>12</td>
<td>37.5</td>
<td>6</td>
<td>68</td>
<td>5</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>BS-cl</td>
<td>37</td>
<td>23</td>
<td>37</td>
<td>10</td>
<td>51</td>
<td>20</td>
<td>25</td>
<td>19.5</td>
</tr>
<tr>
<td>2L1 + BS-cl</td>
<td>42</td>
<td>14</td>
<td>37</td>
<td>7</td>
<td>61</td>
<td>14</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>SB-ce</td>
<td>79</td>
<td>12</td>
<td>57</td>
<td>4</td>
<td>81</td>
<td>10</td>
<td>16.5</td>
<td>8.5</td>
</tr>
<tr>
<td>M-BA</td>
<td>81</td>
<td>26</td>
<td>27</td>
<td>5</td>
<td>43</td>
<td>12</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>M-SP</td>
<td>70</td>
<td>22.5</td>
<td>83</td>
<td>9.5</td>
<td>24</td>
<td>26.5</td>
<td>38</td>
<td>24</td>
</tr>
<tr>
<td>Biling.</td>
<td>46.5</td>
<td>12.5</td>
<td>40.5</td>
<td>6</td>
<td>61.5</td>
<td>13</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>M-SP/BA</td>
<td>76</td>
<td>26</td>
<td>72</td>
<td>8</td>
<td>32</td>
<td>15</td>
<td>25</td>
<td>19</td>
</tr>
</tbody>
</table>

The interesting question now is how these observed differences can be explained on a syntactical level and how they fit into the theoretical model presented in Chapter 5. I will venture the following hypothesis: First, monolingual Spanish speakers only analyze a complex wh-phrase as a CLLDed constituent if an additional element, for example a preverbal subject, is also moved into the left periphery. In other words, the CP-layer is only activated if it becomes necessary due to the presence of more than one fronted constituent.\(^{85}\) That means, complex wh-phrases generally behave like simple wh-phrases and move to SpecFinP. In the case of a further element appearing in the left periphery however, the complex wh-phrase first moves to a higher DS, similar to a CLLD, and then further on to ForceP, as described in Chapter 3. The important point here is that if there is no additional LD, there is also no reason to activate the higher CP-layer, the complex wh-phrase only moves to the ‘standard’ wh-landing site in SpecFinP, and the clitic doubling of the wh-phrase sounds deviant.\(^{86}\) Second, the high variance

---

\(^{85}\) In the model described in Chapter 3, FinP counts as an extension of the IP.

\(^{86}\) Barbosa (2001) also proposes interrogative structures where the CP-layer is only projected when needed and not in the case of simple wh-questions, as mentioned in footnote 37.
in the acceptability ratings within the speaker groups support the hypothesis of an ongoing internal language change in Spanish. The general spreading, however, advances at varying speed in the different speaker communities. It has advanced the furthest outside the Basque-speaking area and the least within the group of the Basque-Spanish bilingual speakers. The third point is that simultaneous and sequential Basque-Spanish bilingual speakers never interpret Spanish complex \textit{wh}-phrases as CLLDs but apply the ‘Basque model’ to all \textit{wh}-phrases. That means all types of \textit{wh}-phrases are always moved to SpecFinP. Lastly, due to the ‘resistance’ or ‘conservatism’ of the bilingual speakers, cf. Section 6.2.2, the M-BA are not as susceptible to the change as the M-SP. The four claims are summarized in (8.8).

(8.8) i. M-SP only analyze complex \textit{wh}-phrases as CLLDed constituents, if an additional element is LDed, appearing between the \textit{wh}-phrase and the verb.

ii. The reanalysis of these complex \textit{wh}-phrases as CLLDed elements is an ongoing internal language change in Spanish.

iii. Simultaneous and sequential Basque-Spanish bilinguals never interpret complex \textit{wh}-phrases as CLLDed constituents.

iv. The ‘resistance’ of the bilingual speaker slows down the overall spreading of the phenomenon in the Basque speaking area.

What evidences can be given from the data to support these hypothesis? The first claim (8.8i) is corroborated by several observations. First of all, the striking difference in the overall acceptability rating by the monolinguals when comparing the complex \textit{wh}-questions with and without \textit{wh}-phrase-verb adjacency. If the monolingual speakers would analyze every complex \textit{wh}-phrase as a CLLDed element, clitic doubling should be obligatory, since it is one of the main properties of CLLD. In other words, the presence of a clitic in a complex \textit{wh}-sentence with \textit{wh}-phrase-verb adjacency should be acceptable for monolingual speakers. The data shows, however, that clitic doubling in these cases decreases the sentence’s acceptability drastically for most of the monolingual speakers, as can be seen in Figure 8.3 and 8.4. Therefore it seems that the monolingual speakers do not usually analyze a complex \textit{wh}-sentence as a CLLD in such interrogatives. That is why they perceive the clitic doubling in these questions as ungrammatical, as they would in the case of simple \textit{wh}-questions. The deciding factor now is that the situation changes for the monolingual speakers if the question exhibit non-\textit{wh}-phrase-verb adjacency. In this case, the presence or absence of the clitic does have a significant influence on the acceptability. As I have listed in Table 7.4, here repeated partly as Table 8.3, the M-SP’s acceptability of the corresponding question pairs differ significantly.
If the *wh*-phrase is a [+human] DO, the acceptability drops significantly without a realized clitic. If it is [-human], in contrast, the question without the clitic has a higher acceptance rate. A possible cause for this difference could be that [-human] DOs in Spanish are generally not doubled by a clitic, as already explained in Section 7.3.3 (Suñer 1988: 396). In other words, a clitic-doubled [-human] DO sounds extremely deviant to the monolinguals, because these types of DO are never clitic-doubled. Furthermore, if we compare the items 2.08 and 2.05 (cf. Table 8.3), we see that the crucial parameter is the clitic and not the insertion. Item_2.08 has a [+human] DO *wh*-phrase doubled by a clitic and Item_2.05 a [-human] DO doubled by a clitic. The acceptability differs significantly between these two questions (\(p_{\text{KWT}} = 1.01 \cdot 10^{-7}\)). That means, the factor [+human] shows a highly significant difference among monolinguals. In the case that a DO is not doubled by a clitic (Item_1.11 vs. Item_1.05), the [+human] dichotomy has no significant influence on the acceptability of the sentences (\(p_{\text{KWT}} = 0.698\)). That means, the presence or absence of the clitic only alters the acceptability if the DO *wh*-phrase is [+human]. A further cause for these differences could be the second hypothesis stated in (8.8ii). The apparent variation in the rating by the M-SP of items that are structurally almost identical supports the hypothesis that the evolution towards full parallelism between complex *wh*-phrases and CLLDed constituents is still in process. If we look at similar processes in other syntactic phenomena, like *leísmo* or DOM, the change started with [+specific, +human] objects and is now advancing towards less marked objects (cf. among others, Rodríguez-Ordóñez 2017). Further evidence in the data for an ongoing change in the interpretation of complex *wh*-phrases is the ratings of Item_2.03 and Item_2.08 (\(\text{medi}_{2.03} = 22.5, \text{medi}_{2.08} = 83\)). As can be seen in (8.9) and (8.10), the two questions have an almost identical structure, with the sole difference being that the *wh*-phrase is singular in the first and plural in the second.

---

87 The doubling is only obligatory in the case of a real CLLDed DO or if the doubled DO is a pronoun (Suñer 1988: 395-397).
Nevertheless, the M-SP’s acceptance vary significantly ($p_{KWT} = 0.0001$). This could indicate that their reanalysis of these structures is not yet fully established.

(8.9) Item 2.03:

**Context:** Your are looking with your girlfriend at a photo of some Italian friends who live in Sicily. She knows that one of your friends had visited one of the Italian girls and she asks you:

¿A cuál de estas chicas tu amiga la/le había visitado en Sicilia?

‘Which of these girls had your friend visited in Sicily?’

(8.10) Item 2.08:

**Context:** Yesterday, several swimmers were late for the medical examination. Since the doctor didn’t have much time, he only accepted to see a few of them. The following day, the trainer wants to know:

¿A cuáles de las que llegaron tarde el médico las/les examinó durante más de cinco minutos?

‘Whom of you that came late did the doctor examine for more than five minutes?’

Additional support for a change in progress can be found in the ratings by the M-BA. Their acceptance rates lie above the bilingual speakers but below the M-SP in the majority of the items. This can be seen in Table 8.1 as well as in the single medians of each item in Table 8.2. Furthermore, the M-BA seem to be the only speaker group that is influenced by the DO-IO-dichotomy, as can be seen in Table 7.5, here repeated partly as Table 8.4.

<table>
<thead>
<tr>
<th>Group</th>
<th>$p_{Lmer}$</th>
<th>Lv. of Signif.</th>
<th>Ø DO</th>
<th>Ø diff. DO vs. IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-BA</td>
<td>0.005</td>
<td>**</td>
<td>42.0</td>
<td>+20.8</td>
</tr>
<tr>
<td>M-SP</td>
<td>0.209</td>
<td>—</td>
<td>52.5</td>
<td>+14.6</td>
</tr>
<tr>
<td>Bilinguals</td>
<td>0.429</td>
<td>—</td>
<td>33.2</td>
<td>+19.0</td>
</tr>
</tbody>
</table>

Neither the M-SP nor the bilingual speakers show this difference. Again, the ratings of the M-BA lie exactly in the middle between the other two groups. The M-BA accept insertions much more readily if the *wh*-phrase is the IO. For the other two groups, the acceptance is either much lower or much higher. In other words, the DO-IO dichotomy is not significant for the M-SP and the bilinguals. The former anyway accept them more and the latter generally reject them. Final evidence for a change in progress is also the behavior of the SB-ce. As already mentioned, their behavior ranges widely, in the sense that they sometimes show similar behavior as the monolinguals, sometimes they behave as
the other bilinguals do. I therefore assume that the M-BA and SB-ce are already influenced by the
general change in the whole Spanish-speaking community, but the change is slowed down by the other
bilingual speakers who never analyze complex wh-phrases as CLLDed constituents. This brings me to
the third claim.

In (8.8iii), the claim is as follows: simultaneous and sequential Basque-Spanish bilinguals never
analyze complex wh-phrases as CLLDed constituents. The strongest evidence for this is the overall low
acceptance of all complex wh-questions without wh-phrase-verb adjacency by this group of speakers.
They generally reject such sentences no matter the type or function of the wh-phrase, or the absence or
presence of a clitic. If we once more have a look at Table 8.2, we see the median does not pass 50% for
any item with insertion. Although the overall acceptance is very low, we have nevertheless seen that
these speakers are also influenced by the [±human, ±insertion]. The relevant data from Section 7.4.3 is
repeated in Table 8.5.

Table 8.5: Lmer model – Summary for the [±human, ±insertion] distinction across bilingual speaker groups

<table>
<thead>
<tr>
<th>Groups</th>
<th>[+human] [+insertion]</th>
<th>[-human] [+insertion]</th>
<th>Difference</th>
<th>pLmer</th>
</tr>
</thead>
<tbody>
<tr>
<td>2L1</td>
<td>37.6</td>
<td>14.7</td>
<td>22.9</td>
<td>0.029 *</td>
</tr>
<tr>
<td>BS-cl</td>
<td>37.6</td>
<td>14.6</td>
<td>23.0</td>
<td>0.029 *</td>
</tr>
<tr>
<td>2L1+BS-cl</td>
<td>37.6</td>
<td>17.2</td>
<td>20.1</td>
<td>0.021 *</td>
</tr>
</tbody>
</table>

The 2L1+BS-cl almost unanimously reject the cases of a [-human] object wh-phrase without wh-
phrase-verb adjacency. Only in case of a [+human] object wh-phrase, is the acceptance slightly higher.
For the majority, however, also these sentences are marginal to unacceptable. We can also see in Ta-
ble 7.4, here repeated partly as Table 8.6, that the absence or presence of a clitic does not influence the
acceptability significantly. In this point, the bilinguals differ clearly from the monolingual speakers.

For the 2L1 and BS-cl, a [+human] object wh-phrase seems to be the only factor that slightly im-
proves the acceptance. The variation with this kind of wh-phrase is perhaps the first sign that the
change is also starting to spread to these speakers. If the phenomenon indeed starts with [+human,
+specific] DOs, it would behave along the same lines as other changes, like DOM or leísmo, as already
detailed in Section 6.1.3. However, the overall acceptability is low, no matter the configurations of
the linguistic factors. I therefore assume that the majority of these speakers analyze every wh-phrase,
complex or not, in the same way: it moves to the preverbal focus position and no insertion whatsoever
is accepted.
Table 8.6: \( p_{\text{KWT}} \)-values for [+clitic]-contrast across bilingual speaker

<table>
<thead>
<tr>
<th>Group</th>
<th>[-human]</th>
<th>[-human]</th>
<th>( p_{\text{KWT}} )</th>
<th>[-human]</th>
<th>[-human]</th>
<th>( p_{\text{KWT}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>median</td>
<td>median</td>
<td>( p_{\text{KWT}} )</td>
<td>median</td>
<td>median</td>
<td>( p_{\text{KWT}} )</td>
</tr>
<tr>
<td>2L1</td>
<td>37.5</td>
<td>20</td>
<td>0.084</td>
<td>6</td>
<td>14</td>
<td>0.361</td>
</tr>
<tr>
<td>BS-cl</td>
<td>37</td>
<td>25</td>
<td>0.362</td>
<td>10</td>
<td>19.5</td>
<td>0.282</td>
</tr>
<tr>
<td>2L1 + BS-cl</td>
<td>37</td>
<td>20</td>
<td>0.055</td>
<td>7</td>
<td>17</td>
<td>0.168</td>
</tr>
</tbody>
</table>

In summary, the data convincingly shows that differences in the acceptability of non-\( wh \)-phrase-verb adjacency exist between monolingual and bilingual speaker. They vary according to the type of \( wh \)-phrase and the absence or presence of a clitic doubling it. The majority of the simultaneous and sequential Basque-Spanish bilingual speakers reject any insertions and analyze complex \( wh \)-phrases the same as simple \( wh \)-phrases, possibly influenced by parallel structures in Basque. The monolingual speakers from outside the Basque-speaking area, however, distinguish between \( wh \)-interrogatives with and without \( wh \)-phrase-verb adjacency. In the bilinguals’ case, a general analysis of a complex \( wh \)-phrase as a CLLDed constituent is rather unlikely since the acceptability by the majority of the speakers is not very high, if the \( wh \)-phrase is doubled by a clitic. For the monolinguals outside the Basque speaking area however, two facts support the theory of a general analysis of a complex \( wh \)-phrase as CLLDed constituent. First, the M-SP strongly favor the version with a clitic that doubles the \( wh \)-phrase, and second, compared to the other speakers, they have an overall high acceptance of such constructions. It therefore seems that the interpretation as CLLDed element is bound to the additional presence of a second preverbal, dislocated element and is strongly favored if the \( wh \)-phrase is a [+human] object. In the case of the M-BA and the sequential Spanish-Basque bilinguals, the acceptability rates lie between the other two groups, which is a sign for a possible ongoing change in the Spanish-speaking community.
Conclusion and outlook

The study presented and discussed in this work attests clear differences in the acceptability of complex \textit{wh}-sentences between bilingual speakers of the Basque-speaking area and monolingual speakers from other parts of Spain. The collected data confirms that if the \textit{wh}-question does not exhibit \textit{wh}-phrase-verb adjacency, M-SP accept it much more readily than bilingual speakers. In other words, if the complex \textit{wh}-phrase is combined with an additional preverbal constituent, M-SP can apparently analyze both preverbal elements as left-dislocated. Strong evidence for this analysis of a complex \textit{wh}-phrase as a CLLDe constituent is the clear preference for a clitic-doubled \textit{wh}-phrase in these cases. If the \textit{wh}-phrase is not doubled by a clitic, the acceptability drops significantly. This observation confirms the hypothesis stated in Chapter 3 and 5 about the alternative analysis of complex \textit{wh}-phrases as CLLDe constituents, proposed by Bosque & Gutiérrez- Rexach (2011), Ordoñez (1998b), and Ordoñez & Olarrea (2006). The bilinguals, however, do not accept these kind of interventions or, at least, do so to a significantly lesser extend. It seem that they always prefer a \textit{wh}-phrase-verb-subject order, no matter the complexity of the \textit{wh}-phrase. Furthermore, the absence or presence of a clitic doubling the \textit{wh}-phrase, does not significantly influence the acceptability. Both observations, the high acceptability by the M-SP and the low acceptability by the bilinguals, especially hold for [+human] DO \textit{wh}-phrases. So far, the data seems to support the hypothesis that the bilingual speakers do not analyze these complex \textit{wh}-phrases as CLLDe constituents, contrary to the monolinguals.

We have seen, however, that the distribution changes as soon as the \textit{wh}-phrase is either a [-human] DO or if the subject appears post-verbally, which results in the standard word order for \textit{wh}-questions. For the first case, the acceptability rates drop significantly and remain low for all speaker groups. Although the M-SP are still more inclined to accept such constructions than the bilinguals, the overall ratings indicate a general rejection of these questions. I therefore concluded that the availability of the alternative analysis of complex object \textit{wh}-phrases as CLLDs depends on the type of object. It is,
however, an interesting observation that the line, where the new analysis is possible, is again drawn along the [+human] separation, similar to other syntactic phenomena like DOM or leísmo, as I have briefly presented in Section 6.1.3.

The second case, the standard word order with a clitic-doubled complex *wh*-phrase, shows another surprising variation among the speaker groups. First, the M-SP apparently do not like a complex *wh*-phrase doubled by a clitic if the sentence otherwise exhibits the standard word order of a *wh*-question. This is a strong argument against a general reanalysis of complex *wh*-phrases as CLLDed constituents. In this case, it also does not matter for the M-SP if the *wh*-phrase is [+human] or not. Their ratings differ significantly from the acceptability rates given by the other speakers. M-BA as well as all bilingual speakers are influenced in their ratings by the factor [+human]. They accept questions with a [+human] DO *wh*-phrase doubled by a clitic much more readily than the M-SP. In fact, especially bilingual speakers did not seem to be disturbed by the clitic. With a [-human] DO *wh*-phrase however, the acceptability rate drops drastically with the presence of a clitic doubling the *wh*-phrase. This observed contrast supports the claim, made among others by Camus Bergareche & Seibane (2012: 13-14) and Landa (2000), that the IO clitic pronoun *le(s)* is used as a kind of agreement marker for the DO, if the DO is [+definite, +human]. Both conditions are met in the complex *wh*-phrases at hand. M-BA show a similar acceptance, although not to the same extent as the bilinguals. The data gives reason to believe that the M-BA can be positioned in the middle between the bilingual speakers and the M-SP, regarding the acceptance rates in these type of questions.

I have advanced the hypothesis that the cause for these observed differences in the acceptance rate could be the following: differences in the input during the acquisition process and an overall tendencies of bilingual speakers to use similar structures in both their languages. In Section 6.2.2, I argued that the quantity of input for bilingual speakers is just not enough to acquire the alternative analysis of complex *wh*-questions. The lacking input mostly happens because children acquiring two languages only receive half the input per language of monolingual children. In other words, they only hear half the amount of the anyway rare complex *wh*-sentences in Spanish. In addition, the adults presumably also use fewer complex *wh*-sentences without *wh*-phrase-verb adjacency. The reason being that they too are subject to the generalization put forward by Müller & Hulk (2001: 19) and others: if one language has possibilities A and B to analyze a certain structure and a second language contains only positive evidence for option A, bilingual speakers tend to primarily use option A in both languages. Applied to our case of the complex *wh*-questions, the bilingual children prefer using the option available in both languages which is the standard word order with *wh*-phrase-verb adjacency. The same counts for the adult bilingual speakers, reducing the input of the non-adjacency structures even more. The
Conclusion and outlook

Monolingual speakers, however, are not influenced by another language and therefore do not show these kind of crosslinguistic preferences. The bilinguals’ tendency to use only the standard word order can be interpreted as a ‘resistance’ to a possible internal language change in Spanish.

The theoretical framework put forward in Chapter 5, based on Cable (2010) and subsequent work, is a possible explanation for the subtle differences in Basque and Spanish wh-question formation. I have shown that a parametrical difference between Basque and Spanish, regarding the core properties of the Q-particle, accounts for the differences in complex wh-interrogatives and multiple wh-questions. Spanish can be classified as a Q-projection language and Basque as a Q-adjunction language. This classification entails the different behavior of the Q-particle. In Spanish, the Q-particle takes the XP that contains the wh-phrase (XP_{wh}) as a complement and projects a Q-particle-phrase (QP). The QP then overtly moves to a higher C-head, ‘dragging’ the wh-phrase along. In Basque, on the other hand, the Q-particle is only adjoint to the XP_{wh}. It therefore enjoys greater freedom and can move to a higher C-head in the left periphery without taking the XP_{wh} along. We have seen that the reason that Basque nevertheless exhibits wh-fronting, lies in an independent feature of the language: the preverbal focus position that always has to be occupied by an appropriate element. This is the cause for the obligatory wh-phrase-verb adjacency in all Basque wh-questions, no matter the complexity of the wh-phrase.

The Q-particle approach can also account for another difference between Basque and Spanish question formation. I have argued in Chapter 5 that the observed variation in multiple wh-questions regarding the complementary occurrence of superiority and intervention effects can be explained by the different number of Q-particles. Basque seems to always have just one Q-particle independent of the number of wh-words, whereas Spanish generates a Q-particle for each of them. It is my conviction, that, in general, the Q-particle approach to wh-questions enables us to explain many crosslinguistic differences in question formation.

On the basis of the work done in this thesis, further interesting questions remain to be answered. In the surveys, I placed the focus mainly on complex wh-questions of the cuál(es) de ‘which of’ type. Some authors, however, claim that other types of complex wh-phrases are also subject to the described phenomenon. In Section 3.2.2.3, I have presented several such examples from Ordoñez (1998b: 346), such as wh-phrases like en qué momento ‘in which moment’, en qué lugar ‘in which place’, and de qué manera/en qué medida ‘in which way’. These kind of wh-phrases clearly resemble more the por qué adjunct type and not the object wh-phrases examined in this work. However, with the corresponding simple wh-words cuándo ‘when’, dónde ‘where’, and cómo ‘how’, non-wh-phrase-verb adjacency does not seem to be accepted (among others, Bosque & Gutiérrez- Rexach 2011: 449, Contreras 1999: 1940,
Conclusion and outlook

Olarrea 1998: 91ff, Zubizarreta 2001: 184, Ordoñez 1998b). This suggests that the complexity could indeed be a decided factor with these wh-phrases.

Closely connected to these complex wh-phrases are other linguistic factors. This work has examined mainly the [±human] DO/IO wh-phrase with and without wh-phrase-verb adjacency. Additional factors that could influence the acceptability are possibly [±specific], [±definite], or the use of subject wh-phrases. The first two factors could be of special importance if the wh-phrase is not a complement, but an adjunct. It is more difficult, however, to compare them to the standard CLLDs, because Spanish does not have locative clitics as, for example, Catalan or French do. The [±clitic] contrast, which turned out to be significant in the study at hand, cannot be tested with wh-phrases other than IO and DO phrases. This is also the difficulty with subject wh-phrases. A further observation that would be worth examining closer is the contrast regarding the factor [±human] object, which was found significant in the surveys within all parameter variation and across all speaker groups. It could be interesting to test a larger monolingual speaker population outside the Basque-speaking area to differentiate further contrasts and other possible factors influencing the acceptability of such sentences. The examination of more factors would definitely allow us to draw a clearer picture of the restrictions and generalizations that underlie these complex wh-questions in Spanish.

Having in mind the parametrical differences in the Q-particle approach, a further interesting line of investigation could be to examine multiple wh-questions in more detail. I have touched on this topic briefly in Chapter 8, but a more detailed and exhaustive study could analyze a possible crosslinguistic influence on the effects of superiority and intervention. In Chapter 5 I have shown that the two languages differ in this regard. Basque exhibits intervention effects but not superiority effects, and in Spanish vice versa. The question would then be if, for example, a weakening of superiority or intervention effects in the respective other language could be found. This would suggest that the mutual influence on wh-structures is even stronger than proposed in this work.

Although the general topic of language acquisition is only tangential to this work and touched upon in Section 6.2, it is my hope to have added further evidence that crosslinguistic influence happens every day and effects the day-to-day language of speakers in bilingual communities. Section 6.2.2 briefly introduced the work by Yang (2016). In his book, he approaches the question of acquisition in a mathematical and quantitative way. His Tolerance Principle and Principle of Sufficiency demand an explicit quantification that allows the computations of the respective thresholds. It would be interesting to have a more detailed and differentiated analysis, especially of the smaller oral corpus COSER, as well as the huge written Corpus Del Español. Having a more in-depth analysis would provide the
necessary numbers and verify the assumptions quantitatively. This could further the discussion about
the importance of input quantity.

Finally, it only remains to say that it has been an honor to have added a hopefully interesting and
insightful element to the ever growing body of work in the fascinating field of cross-linguistic research.
Bibliography


Davies, Mark. 2016. Corpus Del Español: 2 Billion Words: 21 Countries. URL: http://www.corpusdelespanol.org/web-dial/ (visited on 02/05/2018).


Goodall, Grant. forthcoming. Syntactic Satiation and the Inversion Effect in English and Spanish Wh-Questions. San Diego.


Schlaak, Claudia. 2014. Das zweigeteilte Baskenland - Sprachkontakt, Sprachvariation und regionale Identität in Frankreich und Spanien. Walter de Gruyter GmbH.


Suñer, Margarita. 2003. The Lexical Preverbal Subject in a Romance Null Subject Language - Where Are Thou?*. In: Rafael Núñez-Cedeño, Luis López & Richard Cameron (eds.), A Romance Perspective on Language Knowledge and Use. Selected Papers from the 31th Linguistic Symposium on Romance Languages (LSRL), Chicago, 19-22 April 2001. 341-357. Amsterdam: Benjamins.


# List of Appendices

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Questionnaire</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix A</td>
<td>Sociolinguistic Questionnaire – SolQ</td>
<td>208</td>
</tr>
<tr>
<td>Appendix B</td>
<td>Questionnaire 1 – Q1</td>
<td>212</td>
</tr>
<tr>
<td></td>
<td>B.1 Introductory text</td>
<td>212</td>
</tr>
<tr>
<td></td>
<td>B.2 Items</td>
<td>213</td>
</tr>
<tr>
<td></td>
<td>B.3 Filler phrases</td>
<td>217</td>
</tr>
<tr>
<td>Appendix C</td>
<td>Questionnaire 2 – Q2</td>
<td>220</td>
</tr>
<tr>
<td></td>
<td>C.1 Introductory text</td>
<td>220</td>
</tr>
<tr>
<td></td>
<td>C.2 Items</td>
<td>221</td>
</tr>
<tr>
<td></td>
<td>C.3 Filler phrases</td>
<td>224</td>
</tr>
<tr>
<td>Appendix D</td>
<td>Pilot-Questionnaire – PilotQ</td>
<td>227</td>
</tr>
<tr>
<td></td>
<td>D.1 Items</td>
<td>227</td>
</tr>
<tr>
<td></td>
<td>D.2 Filler phrases</td>
<td>233</td>
</tr>
</tbody>
</table>
APPENDIX A

Sociolinguistic Questionnaire – SolQ

The original questionnaires were completely in Spanish. However, in order to ease the reading, I translated contexts, introductory texts and the sociolinguistic questions into English. Only the items and fillers phrases are in the original language, and then glossed.

1. \textit{Gender}

   \textit{I am} ...
   \begin{itemize}
   \item female
   \item male
   \end{itemize}

2. \textit{Age}

   \textit{How old are you?}
   \begin{itemize}
   \item under 15
   \item 15 - 19
   \item 20 - 24
   \item 25 - 29
   \item 30 - 34
   \item 35 - 39
   \item 40 - 44
   \item 45 - 49
   \item 50 - 59
   \item 60 or older
   \end{itemize}

3. \textit{Education}

   \textit{Please indicate your highest degree.}
   \begin{itemize}
   \item pupil
   \item High School
   \item Apprenticeship
   \item Bachillerato
   \item Technical College
   \item University
   \item Other
   \end{itemize}
### Origin

**Where are you from?**
- Alava
- Guipúzcoa
- Vizcaya
- Navarra
- other part of Spain
- other

**Where is your mother from?**
- Alava
- Guipúzcoa
- Vizcaya
- Navarra
- other part of Spain
- other

**Where is your father from?**
- Alava
- Guipúzcoa
- Vizcaya
- Navarra
- other part of Spain
- other

### Native Language

**Which is your native language? (Multiple selection is possible)**
- Spanish
- Basque
- others: 

**Which is your mother’s native language? (Multiple selection is possible)**
- Spanish
- Basque
- others: 

**Which is your father’s native language? (Multiple selection is possible)**
- Spanish
- Basque
- others: 
### Age of Onset

*When did you learn Basque?*
- at home
- in kindergarten
- in primary school
- in secondary school
- after the age of 18
- other: 

*When did you learn Spanish?*
- at home
- in kindergarten
- in primary school
- in secondary school
- after the age of 18
- other: 

### Comfort

*In which language do you feel more comfortable?*
- Spanish
- Basque
- equally in both
- others: 

### Spoken language at home

*Which language do you normally speak at home with your parents? (Multiple selection is possible)*
- Spanish
- Basque
- Spanish and Basque
- others: 

*If you do not live with your parents anymore: Which language do you normally speak at home with your family? (Multiple selection is possible)*
- Spanish
- Basque
- Spanish and Basque
- others: 

210
**Daily language use**

*Please indicate the languages you use on a normal day as a percentage?*

*For example, if you only speak Spanish, you write 100 after “Spanish”,
if you speak half Spanish and the other half Basque, you write 50 in both cases,
if you speak one or several other languages, please indicate which they are and the percentage of use.*

- Spanish: ________
- Basque: ________
- Other language(s)
  - Which: ________
  - Percentage: ________

<table>
<thead>
<tr>
<th>Spanish</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basque</td>
<td>%</td>
</tr>
<tr>
<td>Other language(s)</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B

Questionnaire 1 – Q1

The original questionnaires were completely in Spanish. However, in order to ease the reading, I translated contexts and introductory texts into English. The items and fillers phrases are in the original language, and then glossed.

B.1 Introductory text

Hello;

In the following, different sentences will be presented to you. Please read each one aloud with a normal intonation and decide how natural they sound to you in your Spanish. Some of the sentences are accompanied by a short text, indicating their context.

You can classify the sentences on a scale between “it sounds very strange to me” (if the sentences seems forced or unnatural) and “it seems natural to me” (if it is a sentence you would use without problems in your daily language). Please, don’t think too much about it, simply follow your intuition.

Before we start with the sentences, the program asks you to indicate your relative use of your language(s) during a normal day.

Thank you
Simon
## B.2 Items

| Item | Context: For the third time, a printer has broken in the office. You walk up to the boss’s office and ask him:  
¿Por qué nosotros estamos usando estas máquinas viejas?  
why we are using these machines old  
‘Why are we using these old machines?’ |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.01</td>
<td><strong>Function of wh-phrase</strong></td>
</tr>
<tr>
<td></td>
<td><strong>wh-phrase CLDed</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Insertion</strong></td>
</tr>
</tbody>
</table>
| 1.02 | Context: The first day of your vacation with some friends you want to go to the beach. One of them doesn’t want to go. You ask him:  
¿Por qué tú te quieres quedar en casa a diferencia de los demás?  
why you CL. want 2sg stay at home contrary to the others  
‘Why do you want to stay at home unlike the others?’ |
|      | **Function of wh-phrase** | Adjunct |
|      | **wh-phrase CLDed** | no |
|      | **Insertion** | Subject pronoun tú ‘you’ |
| 1.03 | Context: You hear through a friend that Juan had just escaped being run over by a car. You ask that friend:  
¿Por qué casi chocó Juan con el coche en plena calle?  
why almost crash Juan with the car on plain road  
‘Why did Juan almost crash with the car on the open road?’ |
|      | **Function of wh-phrase** | Adjunct |
|      | **wh-phrase CLDed** | no |
|      | **Insertion** | Adverb casi ‘almost’ |
| Item | CONTEXT: Your friend is already playing in three bands. He tells you that he also wants to play with another band. You ask him: |  |
|------|---------------------------------------------------------------------------------------------------------------------------------|  |
| 1.04 | ¿Por qué también quieres tocar con este grupo de música? 'Why do you want to play in this band as well?' | Functions of wh-phrase: Adjunct wh-phrase CLDed: no Insertion: Adverb también ‘also’ |

| 1.05 | CONTEXT: A friend of yours, who likes your sister, asks you about the classes she wants to enroll in this summer. |  |
|------|---------------------------------------------------------------------------------------------------------------------------------|  |
|       | ¿A cuál de los cursos de cocina tu hermana se apuntará este verano? 'In which of the cooking classes will your sister enroll this summer?' | Functions of wh-phrase: Direct object wh-phrase CLDed: no Insertion: Subject DP tu hermana ‘your sister’ |

| 1.06 | CONTEXT: You recommended some science-fiction books to a friend because he was looking for a birthday present for Juan. Some time later, you see him again and ask: |  |
|------|---------------------------------------------------------------------------------------------------------------------------------|  |
|       | ¿Cuál de los libros recomendados tú has regalado a Juan por su cumple al final? 'Which of the recommended books did you give Juan for his birthday in the end?' | Functions of wh-phrase: Direct Object wh-phrase CLDed: no Insertion: Subject pronoun tú ‘you’ |
| Item | **Context**: You are talking about the upcoming elections.  
¿Cuál de los nuevos partidos democráticos probablemente saldrá ganando en las elecciones?  
'Which of the new democratic parties will probably end up winning the elections?' |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Function of wh-phrase</strong></td>
<td>Subject</td>
</tr>
<tr>
<td><strong>wh-phrase CLDed</strong></td>
<td>no</td>
</tr>
<tr>
<td><strong>Insertion</strong></td>
<td>Adverb <em>probablemente</em> ‘probably’</td>
</tr>
</tbody>
</table>

| Item | **Context**: You are choosing the movies for a movie night with your friend Maria this weekend.  
¿Cuál de las películas probablemente quiere ver María este fin de semana?  
'What movies will Maria probably want to watch this weekend?' |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Function of wh-phrase</strong></td>
<td>Direct Object</td>
</tr>
<tr>
<td><strong>wh-phrase CLDed</strong></td>
<td>no</td>
</tr>
<tr>
<td><strong>Insertion</strong></td>
<td>Adverb <em>probablemente</em> ‘probably’</td>
</tr>
</tbody>
</table>

| Item | **Context**: Several pupils have failed the exam in philosophy. Now you find out that one or several of them have also failed their language exam the previous week. You ask the language teacher:  
¿Quién de los alumnos también ha suspendido el examen de lengua la semana pasada?  
'Which one of the pupils also failed the language exam last week?' |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Function of wh-phrase</strong></td>
<td>Subject</td>
</tr>
<tr>
<td><strong>wh-phrase CLDed</strong></td>
<td>no</td>
</tr>
<tr>
<td><strong>Insertion</strong></td>
<td>Adverb <em>también</em> ‘also’</td>
</tr>
</tbody>
</table>
### 1.10

**CONTEXT:** –

¿Quién de los alumnos frecuentemente olvida sus deberes en casa?

‘Which one of the pupils forgets his/her homework at home frequently?’

<table>
<thead>
<tr>
<th>Function of wh-phrase</th>
<th>wh-phrase CLDed</th>
<th>Insertion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>no</td>
<td>Adverb frecuentemente ‘frequently’</td>
</tr>
</tbody>
</table>

### 1.11

**CONTEXT:** One roommate to the other about Maria, who lives with them.

¿A quién de sus dos amantes María ha invitado a cenar esta noche?

‘Which of her two lovers has Maria invited to dinner this evening?’

<table>
<thead>
<tr>
<th>Function of wh-phrase</th>
<th>wh-phrase CLDed</th>
<th>Insertion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Object</td>
<td>no</td>
<td>Subject María</td>
</tr>
</tbody>
</table>

### 1.12

**CONTEXT:** You go back to your village along with your brother. Your girlfriend asks you:

¿A quién de los viejos amigos vosotros veréis la semana que viene cuando volvéis al pueblo?

‘Who of the old friends will you see next week when you go back to the village?’

<table>
<thead>
<tr>
<th>Function of wh-phrase</th>
<th>wh-phrase CLDed</th>
<th>Insertion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Object</td>
<td>no</td>
<td>Subject pronoun nosotros ‘we’</td>
</tr>
</tbody>
</table>
# B.3 Filler phrases

| Filler | CONTEXT: Pedro is very occupied in the garage since hours. You ask his girlfriend:  
¿Qué Pedro hace?  
what Peter do  
‘What is Peter doing?’ |
|---|---|
| 1.13 | CONTEXT: Two friends speak about another friend and his girlfriend and about what he should do on their anniversary.  
Pablo debería sorprender a su novia con un regalo.  
Pablo should surprise his girlfriend with a present  
‘Pablo should surprise his girlfriend with a present.’ |
| 1.14 | CONTEXT: –  
¡No me vengas con tonterías!  
not Cl. come with nonsense  
‘Don’t tell me nonsense!’ |
| 1.15 | CONTEXT: –  
Los estudiantes se sorprendían por lo difícil que era obtener los puntos necesarios.  
The students Cl. surprised about Cl. difficult that was get the points necessary  
‘The students were surprised at how difficult it was to get the necessary points.’ |
| 1.16 | CONTEXT: –  
Antes de volver a casa, María pasó por la tienda de su tía.  
Before of return to home, Maria passed by the shop of her aunt  
‘Before returning home, Maria passed by her aunt’s shop.’ |
| 1.17 | CONTEXT: You just heard that your nephew has been born.  
¡Trae el champán, que tenemos que celebrar la buena noticia!  
bring the champagne that have.1pt. that celebrate the good news  
‘Bring the champagne! We have to celebrate this good news!’ |
<table>
<thead>
<tr>
<th>1.19</th>
<th><strong>CONTEXT:</strong> –</th>
</tr>
</thead>
<tbody>
<tr>
<td>¿Cuándo viene adónde quién? when comes.3sg where whom</td>
<td></td>
</tr>
<tr>
<td>‘Who came when where?’</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1.20</th>
<th><strong>CONTEXT:</strong> –</th>
</tr>
</thead>
<tbody>
<tr>
<td>María expresó su punto de vista sencillamente. Maria expressed her point of view simply</td>
<td></td>
</tr>
<tr>
<td>‘Maria simply expressed her point of view.’</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1.21</th>
<th><strong>CONTEXT:</strong> Where is Pepe and what is he doing?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pepe está en casa con su novia cenando. Pepe is in house with his girlfriend dining</td>
<td></td>
</tr>
<tr>
<td>‘Pepe is at home with his girlfriend having dinner.’</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1.22</th>
<th><strong>CONTEXT:</strong> What is Berta doing?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berta está por todos lados buscando sus llaves. Berta is everywhere searching her keys</td>
<td></td>
</tr>
<tr>
<td>‘Berta is searching for her keys everywhere.’</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1.23</th>
<th><strong>CONTEXT:</strong> –</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuestros amigos siempre traían dulces cuando nos visitaban. our friends always brought sweets when us visited.3pl</td>
<td></td>
</tr>
<tr>
<td>‘Our friends always brought sweets when they visited us.’</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1.24</th>
<th><strong>CONTEXT:</strong> –</th>
</tr>
</thead>
<tbody>
<tr>
<td>¿Dónde dio Carmen a quién qué? where gave Carmen to who what</td>
<td></td>
</tr>
<tr>
<td>‘To whom did Carmen give what where?’</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1.25</th>
<th><strong>CONTEXT:</strong> –</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pepe ha insistido en comer temprano. Pepe has insisted on eat early</td>
<td></td>
</tr>
<tr>
<td>‘Pepe has insisted on eating early.’</td>
<td></td>
</tr>
</tbody>
</table>
| 1.26 | **CONTEXT:** –
|      | ¡No quiero que vengas! |
|      | not want.1sg that come.2sg |
|      | ‘I don’t want you to come!’ |

| 1.27 | **CONTEXT:** You, about the concert of your choir the other day
|      | Nosotros cantabais muchas canciones conocidas. |
|      | we sang.2pl a lot of songs known |
|      | ‘We sang a lot of well-known songs.’ |

| 1.28 | **CONTEXT:** –
|      | Nadie puede descansar cuando el vecino toca el piano. |
|      | Nobody can rest when the neighbor plays the piano |
|      | ‘Nobody can rest when the neighbor plays the piano.’ |

| 1.29 | **CONTEXT:** –
|      | Juan ha el diario leído durante el desayuno. |
|      | Juan has the newspaper read during the breakfast |
|      | ‘Juan read the newspaper during breakfast.’ |

| 1.30 | **CONTEXT:** –
|      | Marta cantará después de que yo salí. |
|      | Marta sing.3sg.fut after that I left |
|      | ‘Marta sang after I left.’ |

| 1.31 | **CONTEXT:** Juan’s opinion about Maria.
|      | Juan cree María ser inteligente |
|      | Juan thinks Maria be intelligent |
|      | ‘Juan thinks that Maria is intelligent.’ |

| 1.32 | **CONTEXT:** Your friend asks if Maria will also come to have dinner tonight. |
|      | María también vendrá esta noche a cenar. |
|      | María also come.3sg.fut this night to dine |
|      | ‘María will also come for dinner tonight.’ |
APPENDIX C

Questionnaire 2 – Q2

The original questionnaires were completely in Spanish. However, in order to ease the reading, I translated contexts and introductory texts into English. The items and fillers phrases are in the original language, and then glossed.

As mentioned in Chapter 7, I adapted the sentences slightly to the regions of the speakers. For the speakers in the Basque speaking area I adapted for the leísmo phenomena, in contrast to the rest of the Spanish speakers. If I used two different versions of the sentences, both clitics are given in the same sentence, since this is the only difference between the two versions.

C.1 Introductory text

Hello;

In the following, different sentences with a short introductory context will be presented to you. Please indicate on the bar below the sentence how natural it seems to you. From “it sounds very strange to me” on the left (if the sentences seems wrong or forced) to “it seems natural to me” on the right (if it is a sentence you would use without problems in your everyday language).

Please judge the sentences intuitively, that is, according to your first impression. In order to judge them, it can be helpful to read them out loud with a normal intonation.

Thank you!
Simon
## C.2 Items

<table>
<thead>
<tr>
<th>Item</th>
<th>CONTEXT: You know that your colleague from work arrived late at a conference the day before. Now you want to know the reason.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.01</td>
<td>¿Por qué viniste tú tarde ayer para la conferencia?</td>
</tr>
<tr>
<td></td>
<td>'Why were you late for the conference yesterday?'</td>
</tr>
<tr>
<td></td>
<td><strong>Function of wh-phrase</strong></td>
</tr>
<tr>
<td></td>
<td>wh-phrase CLDed</td>
</tr>
<tr>
<td></td>
<td>Insertion</td>
</tr>
</tbody>
</table>

| 2.02 | CONTEXT: At Juan’s wedding, his friends have prepared a game in order to see how well his new wife knows him. They ask her questions about his childhood and teenage years. For example: |
|      | ¿A cuál de sus profesores de primaria Juan le dedicó un poema? |
|      | ‘Which of his primary school teachers did Juan dedicate a poem to? ’ |
|      | **Function of wh-phrase** | Indirect object |
|      | Type of wh-phrase | [+human] |
|      | wh-phrase CLDed | IO pronoun le |
|      | Insertion | Subject Juan |

<p>| 2.03 | CONTEXT: You are looking with your girlfriend at a photo of some Italian friends who live in Sicily. She knows that one of your friends had visited one of the Italian girls and she asks you: |
|      | ¿A cuál de estas chicas tu amiga la/le había visitado en Sicilia? |
|      | ‘Which of these girls had your friend visited in Sicily?’ |
|      | <strong>Function of wh-phrase</strong> | Direct object |
|      | Type of wh-phrase | [+human] |
|      | wh-phrase CLDed | DO pronoun la/le |
|      | Insertion | Subject DP tu amiga ‘your friend’ |</p>
<table>
<thead>
<tr>
<th>Item</th>
<th>CONTEXT: Maria and Ana are talking about Maria’s son Juan and his journey to England last summer. They have family there. Ana, who knows the whole family very well, asks:</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.04</td>
<td>¿A cuál de las primas que viven en Inglaterra la/le visitó Pablo para mejorar su inglés durante el verano pasado?</td>
</tr>
<tr>
<td></td>
<td>Which of the cousins that live in England did Pablo visit last summer to improve his English?</td>
</tr>
<tr>
<td></td>
<td>'Which of his cousins, that live in England, did Pablo visit last summer to improve his English?'</td>
</tr>
<tr>
<td>Function of wh-phrase</td>
<td>Direct object</td>
</tr>
<tr>
<td>Type of wh-phrase</td>
<td>[+human]</td>
</tr>
<tr>
<td>wh-phrase CLDed</td>
<td>DO pronoun la/le</td>
</tr>
<tr>
<td>Insertion</td>
<td>none</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Context</th>
<th>The morning of her wedding just before the ceremony, Elena, is getting ready with the help of two friends. The two pass to the other side of the room where the flowers lay prepared for the bride and her maids of honor. One of them asks:</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.05</td>
<td>¿Cuál de estas flores Elena la llevará en el pelo durante la ceremonia?</td>
</tr>
<tr>
<td></td>
<td>Which of these flowers will Elena wear in her hair during the ceremony?</td>
</tr>
<tr>
<td>Function of wh-phrase</td>
<td>Direct object</td>
</tr>
<tr>
<td>Type of wh-phrase</td>
<td>[-human]</td>
</tr>
<tr>
<td>wh-phrase CLDed</td>
<td>DO pronoun la</td>
</tr>
<tr>
<td>Insertion</td>
<td>Subject Elena</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Context</th>
<th>You didn’t go to school this morning. Your classmate tells you in the afternoon that the language teacher called some of the students to her office. You ask her:</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.06</td>
<td>¿A cuál de las alumnas de nuestra clase el director las/les llamó a su oficina?</td>
</tr>
<tr>
<td></td>
<td>Which of the students of our class the director called to his office</td>
</tr>
<tr>
<td>Function of wh-phrase</td>
<td>Direct object</td>
</tr>
<tr>
<td>Type of wh-phrase</td>
<td>[+human]</td>
</tr>
<tr>
<td>wh-phrase CLDed</td>
<td>DO pronoun las/les</td>
</tr>
<tr>
<td>Insertion</td>
<td>Subject DP el director 'the director'</td>
</tr>
</tbody>
</table>
**Item 2.07**

**CONTEXT:** Isabel’s parents have prepared a list with all the songs she knows how to play on the piano. At her grandparent’s birthday, they see the list and ask her mother:

¿Cuáles de estas canciones maravillosas va a tocar Isabela esta noche?

‘Which of these wonderful songs will Isabela play tonight?’

<table>
<thead>
<tr>
<th>Function of wh-phrase</th>
<th>Direct object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of wh-phrase</td>
<td>[-human]</td>
</tr>
<tr>
<td>wh-phrase CLDed</td>
<td>DO pronoun las</td>
</tr>
<tr>
<td>Insertion</td>
<td>none</td>
</tr>
</tbody>
</table>

**Item 2.08**

**CONTEXT:** Yesterday, several swimmers were late for the medical examination. Since the doctor didn’t have much time, he only saw a few of them. The following day, the trainer wants to know:

¿A cuáles de las que llegaron tarde el médico examinó durante más de cinco minutos?

‘Whom of you that came late did the doctor examine for more than five minutes?’

<table>
<thead>
<tr>
<th>Function of wh-phrase</th>
<th>Direct object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of wh-phrase</td>
<td>[+human]</td>
</tr>
<tr>
<td>wh-phrase CLDed</td>
<td>DO pronoun las/les</td>
</tr>
<tr>
<td>Insertion</td>
<td>Subject DP el médico ‘the doctor’</td>
</tr>
</tbody>
</table>
## C.3 Filler phrases

| Filler | CONTEXT: On a war ship, the captain has ordered several sailors to clean the deck as a punishment for lacking discipline. A scientist who accompanies the voyage asks the cook:  
¿A cuál de los marineros ha el capitán elegido para limpiar la cubierta?  
‘Which of the sailors has the captain chosen to clean the deck?’ |
|---|---|
| 2.09 | CONTEXT: You are talking with a friend about Carmen’s new work. The friend tells you that Carmen isn’t very happy.  
Lamento Carmen no esté contenta con su puesto nuevo en la empresa.  
‘I am sorry to hear that Carmen isn’t happy with her new job at the company.’ |
| 2.10 | CONTEXT: The mother asks the father what their son Juan is doing. He says that he is probably tidying up his room. After listening for a moment without hearing anything from the child’s room, she asks:  
¿Está Juan limpiando realmente su cuarto o está jugando a videojuegos?  
‘Is Juan really tidying up his room or is he playing video games?’ |
| 2.11 | CONTEXT: Pedro: Who did you give a book to? To Maria or to Juan?  
Alberto: “A María le di un libro ayer.  
‘Alberto: “I gave a book to Maria yesterday.”’ |
| 2.12 | CONTEXT: Pepe knows that Miguel has sent Joana a birthday present, but doesn’t know what it is. He asks a friend of Joana’s:  
¿Qué a Joana le envió Miguel?  
‘What did Miguel send to Joana?’ |
| 2.13 | CONTEXT: You ask François’s roommate: “Do you know the Marseillaise?” She responds:  
La Marsellesa, la canta en la ducha François a menudo.  
‘The Marseillaise, François sings it often in the shower.’ |
2.15  *CONTEXT:* In literature class you are talking about Cervantes. The teacher wants to know why he wrote his works. Someone raises his hand and says:

Yo recuerdo por qué Cervantes escribió el Quijote y María recuerda por qué escribió las Novelas Ejemplares.

‘I remember why he wrote the Quijote and Maria remembers why he wrote the *Novelas Ejemplares*’

2.16  *CONTEXT:* Anna: “In order to clarify the situation concerning the pools and the training schedule, Maria, the city’s swimming club director has to talk to the responsible minister. Everything should happen very quickly, because the swimmers have to start their training as soon as possible.”

María: “¿Con quién tiene que hablar quién?”

‘Maria: “Who has to speak with whom?”’

2.17  *CONTEXT:* José has been taken into custody because of a bank robbery. After further investigations, there are more arrests.

El amigo que le acusó a José también fue encarcelado.

‘The friend who accused Juan also got arrested.’

2.18  *CONTEXT:* During a police investigation, the police asks you about the suspect who you saw after the crime. You repeat the story that the man told you and the police wants to know if he told it to someone else. You reply:

Contó el cuento solo a nosotros.

‘He only told the story to us.’

2.19  *CONTEXT:* Pepe: “Did you see Carmen and Pedro at the university this morning?”

Alberto: “A Carmen le he visto, pero a Pedro no le he visto desde hace tiempo.”

‘Carmen I saw, but Pedro I haven’t seen in a long time.’
2.20  

**CONTEXT:** A while ago, your friend Mario had asked you where he could buy nice flowers and you recommended several florist to him. Some time later, another friend show you a picture of the flowers he had bought. You ask her:

¿En qué floristería las compró Mario finalmente estas flores hermosas para su novia?

‘In which florist did Mario end up buying these beautiful flowers for his girlfriend?’
## Appendix D

### Pilot-Questionnaire – PilotQ

The original questionnaires were completely in Spanish. However, in order to ease the reading, I translated the context of the questions into English. The items and fillers phrases are in the original language, and then glossed.

### D.1 Items

<table>
<thead>
<tr>
<th>Item</th>
<th>Original Question</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>¿Quién que no fuera un mentiroso podría poner en cuestión lo dicho?</td>
<td>'Who that is not an idiot would question the stated?'</td>
</tr>
<tr>
<td>02</td>
<td>¿Quién podría poner en cuestión lo dicho que no fuera un mentiroso?</td>
<td>'Who that is not an idiot would question the stated?'</td>
</tr>
<tr>
<td>03</td>
<td>¿Quién en Madrid quiere visitar el museo?</td>
<td>'Who wants to visit the museum in Madrid?'</td>
</tr>
<tr>
<td>04</td>
<td>¿Quién en el senado podría saber más de este tema?</td>
<td>'Who in the senate could know more about this subject'</td>
</tr>
<tr>
<td>05</td>
<td>Le preguntó quién en este edificio tenía un gato.</td>
<td>'He asked him who had a cat in this house.'</td>
</tr>
<tr>
<td>06</td>
<td>Siempre me pregunta quién de entre ellos será la próxima estrella.</td>
<td>'I always ask myself who of the two will be the next star.'</td>
</tr>
<tr>
<td>07</td>
<td>¿A quién nosotros hemos expulsado del grupo?</td>
<td>'Whom have we expelled from the group?'</td>
</tr>
<tr>
<td>08</td>
<td>¿A quién este cielo azul engaña?</td>
<td>'Whom will this blue sky deceive?'</td>
</tr>
<tr>
<td>09</td>
<td>¿A quién de las tres va a ofrecer el chico su ayuda?</td>
<td>'Which of the three will the boy offer his help to?'</td>
</tr>
<tr>
<td>Item</td>
<td>Spanish</td>
<td>English</td>
</tr>
<tr>
<td>------</td>
<td>------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>10</td>
<td>¿A quién de las tres el chicho va a ofrecer su ayuda?</td>
<td>'Which of the three will the boy offer his help to?'</td>
</tr>
<tr>
<td></td>
<td>'¿A quién de las tres el chicho va a ofrecer su ayuda?'</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>¿Sabes a quién la mina pertenecía?</td>
<td>'You know who owned the mine?'</td>
</tr>
<tr>
<td></td>
<td>'¿Sabes a quién la mina pertenecía?'</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>El padre preguntó a quién él votará cuando tenga 18 años.</td>
<td>'The father ask who he would vote for once he was 18.'</td>
</tr>
<tr>
<td></td>
<td>'El padre preguntó a quién él votará cuando tenga 18 años.'</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>¿A cuál de los cursos usted se apunta?</td>
<td>'Which of the classes will you sign up for?'</td>
</tr>
<tr>
<td></td>
<td>'¿A cuál de los cursos usted se apunta?'</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>¿A cuál de estos artistas la obra representada en aquel cuadro pertenece entonces?</td>
<td>'Whom of these artists does the work represented in the frame belong to?'</td>
</tr>
<tr>
<td></td>
<td>'¿A cuál de estos artistas la obra representada en aquel cuadro pertenece entonces?'</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Se detenía a preguntar a cuál correspondían los precios mencionados en el escaparate.</td>
<td>'He/She stopped in order to ask to what the prices in the shopping window corresponded.'</td>
</tr>
<tr>
<td></td>
<td>'Se detenía a preguntar a cuál correspondían los precios mencionados en el escaparate.'</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Se detenía a preguntar a cuál los precios mencionados correspondían en el escaparate.</td>
<td>'He/She stopped in order to ask to what the prices in the shopping window corresponded.'</td>
</tr>
<tr>
<td></td>
<td>Se detenía a preguntar a cuál los precios mencionados correspondían en el escaparate.</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>¿Con cuál de sus primos se lleva mejor Juan?</td>
<td>'With which of his cousins does Juan get along the best?'</td>
</tr>
<tr>
<td></td>
<td>'¿Con cuál de sus primos se lleva mejor Juan?'</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>¿Con cuál de sus primos Juan se lleva mejor?</td>
<td>'With which of his cousins does Juan get along the best?'</td>
</tr>
<tr>
<td></td>
<td>'¿Con cuál de sus primos Juan se lleva mejor?'</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>El periodista no sabia a cuál de los festivales de cine asistió el critico como espectador.</td>
<td>'The journalist did not know which film festival the critic attended as spectator.'</td>
</tr>
<tr>
<td></td>
<td>'El periodista no sabia a cuál de los festivales de cine asistió el critico como espectador.'</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>El periodista no sabia a cuál de los festivales de cine el critico asistió como espectador</td>
<td>'The journalist did not know which film festival the critic attended as spectator.'</td>
</tr>
<tr>
<td></td>
<td>El periodista no sabia a cuál de los festivales de cine el critico asistió como espectador</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>¿Dónde esta larga carretera acaba?</td>
<td>'Where does this long road end?'</td>
</tr>
<tr>
<td></td>
<td>'¿Dónde esta larga carretera acaba?'</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Pablo echaba vistazos al estudiante, preguntándose dónde había visto él aquella cara.</td>
<td>'Pablo was looking at the student, asking himself where he had seen his/her face before.'</td>
</tr>
<tr>
<td>Item</td>
<td>Sentence</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
<td></td>
</tr>
</tbody>
</table>
| 23   | Pablo echaba vistazos al estudiante, preguntándose dónde él había visto aquella cara.  
‘Pablo was looking at the student, asking himself where he had seen his/her face before.’ |
| 24   | Querían saber dónde los peregrinos se refugiaban para mandarles comida.  
‘They wanted to know where the pilgrims had found shelter in order to send them food.’ |
| 25   | ¿Dónde usted lo usa?  
‘Where do you use it?’ |
| 26   | Nadie sabía de dónde los generales aliados habían tomado la fuerza para resistir el ataque.  
‘Nobody knew where the allied generals had taken the force from to resist the attack.’ |
| 27   | ¿Dónde sería posible tal manifestación de descontento en Europa?  
‘Where would such a demonstration of discontentment be possible in Europe?’ |
| 28   | ¿Dónde en Europa sería posible tal manifestación de descontento?  
‘Where would such a demonstration of discontentment be possible in Europe?’ |
| 29   | ¿Cuándo esta película aburrida acabará ya por una vez?  
‘When will this boring movie end for heavens sake’ |
| 30   | ¿Cuándo tu has escrito esa carta?  
‘When did you write this letter?’ |
| 31   | Los padres preguntaron cuándo los hijos pensaban irse.  
‘The parents ask when the children were planning to leave?’ |
| 32   | ¿Cuándo la Justicia lo va a convertir en sentencia firme?  
‘When will the court turn it into a binding sentence?’ |
| 33   | El misterio de saber cuándo la sopa está hecha es echando un huevo.  
‘The mystery in knowing when the soup is done is putting an egg in.’ |
| 34   | Solo un químico podrá decirte cuándo el experimento está terminado.  
‘Only a chemist can say when the experiment is complete.’ |
| 35   | ¿Desde cuándo a una reina se llama como a la dueña de una pensión?  
‘Since when do you call a queen as if you were calling the owner of a guest house.’ |
| 36   | ¿Cuándo en concreto debe iniciarse el tratamiento?  
‘When exactly should the treatment start?’ |
<table>
<thead>
<tr>
<th>Item</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
</table>
| 37   | Las instrucciones no decían cuándo exactamente había que tomar la medicina.  
      'The instructions did not say exactly when you should take the medicine.' |  |
| 39   | ¿Cómo los novios se conocieron?  
      'How did the couple meet?'  
      *Answer: At a party.* |  |
| 40   | Yo quiero saber cómo ella lo sabe.  
      'I want to know how she knowns.'  
      *Answer: Her friend told her* |  |
| 41   | ¿Cómo ella iba a obtener la anulación?  
      'How would she get the cancellation?'  
      *Answer: She would have to present herself before the authorities and ask for it.* |  |
| 42   | Le pregunté cómo él se concentra ahora para leer un libro.  
      'I ask him how he could concentrate on reading a book now.'  
      *Answer: Putting on headphones to cut out the noise.* |  |
| 43   | ¿Cómo esas dos personas se enfrentaron con el problema de la crisis económica?  
      'How did these two people treat the problem of the economic crisis?'  
      *Answer: One emigrated and the other one started to study again.* |  |
| 44   | Algunos vecinos se preguntan cómo Pedro puede vivir a este ritmo con una renta tan pequeña.  
      'Some neighbors ask themselves how Pedro can live at this rhythm with such a small income.'  
      *Answer: It seems that he inherited a small fortune from his uncle.* |  |
| 45   | El policía no dijo cómo los servicios secretos se habían enterado del plan.  
      'The policeman did not say how the secret services had found out about the plan.'  
      *Answer: Reading messages and listening in on telephone conversations.* |  |
| 46   | ¿Cómo a partir de estos datos deducen ustedes un cambio de la opinión pública?  
      'How do you conclude a change in the public opinion from this data?'  
      *Answer: Comparing them to older data.* |  |
<table>
<thead>
<tr>
<th>Item</th>
<th>Translation</th>
</tr>
</thead>
</table>
| 47   | Uno se pregunta cómo, caso de aceptar su teoría, la materia ha podido cambiar desde un gas hasta un sólido.  
"If you accept his/her theory, you may wonder how the matter could change from gas to solide."  
*Answer: Just because of its very peculiar chemical characteristics.* |
| 48   | Uno se pregunta cómo la materia ha podido cambiar desde un gas hasta un sólido, caso de aceptar su teoría.  
"If you accept his/her theory, you may wonder how the matter could change from gas to solide."  
*Answer: Just because of its very peculiar chemical characteristics* |
| 49   | ¿Por qué ibas a advertirme?  
"Why would you warn me?" |
| 50   | ¿Por qué una cosa así habrá sucedido en un día como hoy?  
"Why has such a thing happened on a day like today?" |
| 51   | Da pie a preguntarse por qué el accidente ocurrió ayer.  
"This invites you to ask yourself why the accident happened yesterday." |
| 52   | Me preguntó por qué yo sabía de aquel encuentro.  
"I ask myself why I knew about that meeting." |
| 53   | ¿Por qué vosotros hacéis estas preguntas?  
"Why do you ask these questions?" |
| 54   | No sabía por qué tú te negabas tan tercamente a enviarme a un colegio.  
"I didn’t know why you so stubbornly refused to send me to school." |
| 55   | ¿Por qué esta mujer va de blanco?  
"Why does this woman wear white?" |
| 56   | Martín tiene que decir por qué San Sebastian le parece más bonita que Madrid.  
"Martin has to say why San Sebastian seems more beautiful to him than Madrid." |
| 57   | ¿Por qué a mí toda idea importante se me ha ocurrido caminando por el bosque?  
"Why has every good idea occurred to me while I was walking through the forest?" |
| 58   | ¿Por qué más cosas preguntas?  
"Why do you ask more things?" |
<table>
<thead>
<tr>
<th>Item</th>
<th>Text</th>
</tr>
</thead>
</table>
| 59   | Yo no sé por qué a muchos grupos les asusta el cambio social.  
     'I don’t know why so many groups are afraid of social change.' |
| 60   | No voy a perder mi tiempo preguntando por qué a ti te parece extremamente importante tanto la teoría como la práctica.  
     'I don’t want to loose time asking why the practical as well as the theoretical is so important to you.' |
| 61   | ¿Por qué en cierto momento surgen teorías similares en lugares muy alejados?  
     'How come similar theories develop at a certain time in places far away from each other?' |
| 62   | No pude menos que decir por qué en el mundo las cosas pasan así.  
     'I couldn’t help but say why things happen like that in the world.' |
D.2 Filler phrases

<table>
<thead>
<tr>
<th>Filler</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>63</td>
<td>¿Qué Pedro hace?</td>
<td>'What does Pedro do?'</td>
</tr>
<tr>
<td></td>
<td>'What does Pedro do?'</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>¿A quién el libro regaló María ayer?</td>
<td>'Whom did Maria give the book to yesterday?'</td>
</tr>
<tr>
<td></td>
<td>'Whom did Maria give the book to yesterday?'</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>¿Cuándo el coche devuelve Juan a su tío?</td>
<td>'When will Juan give the car back to his uncle?'</td>
</tr>
<tr>
<td></td>
<td>'When will Juan give the car back to his uncle?'</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>¿Por qué estos libros ha comprado él?</td>
<td>'Why has he bought these books?'</td>
</tr>
<tr>
<td></td>
<td>'Why has he bought these books?'</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>¿Qué pregunta Pedro María ha comprado para ella?</td>
<td>'Pedro asks what María has bought for her.'</td>
</tr>
<tr>
<td></td>
<td>'Pedro asks what María has bought for her.'</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>¿Dónde a Paula viste tú ayer?</td>
<td>'Where have you seen Paula yesterday?'</td>
</tr>
<tr>
<td></td>
<td>'Where have you seen Paula yesterday?'</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>¿Cómo a Sancho podía escapar un asunto tan escandaloso?</td>
<td>'How could such a scandal escape Sancho?'</td>
</tr>
<tr>
<td></td>
<td>'How could such a scandal escape Sancho?'</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>Encantaría Pablo sorprender a su novia con un regalo.</td>
<td>'Pablo would love to surprise his girlfriend with a present.'</td>
</tr>
<tr>
<td></td>
<td>'Pablo would love to surprise his girlfriend with a present.'</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>¡No me vengas con tonterías!</td>
<td>'Don’t tell me nonsense!'</td>
</tr>
<tr>
<td></td>
<td>'Don’t tell me nonsense!'</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>Los estudiantes se sorprendían por lo difícil que era obtener los puntos necesarios.</td>
<td>'The students were surprised at how difficult it was to get the necessary points.'</td>
</tr>
<tr>
<td></td>
<td>'The students were surprised at how difficult it was to get the necessary points.'</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>Antes de volver a casa, María pasó por la tienda de su tía.</td>
<td>'Before she returned home, María went by her aunt’s shop.'</td>
</tr>
<tr>
<td></td>
<td>'Before she returned home, María went by her aunt’s shop.'</td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>María también vendrá esta noche a cenar.</td>
<td>'María will come for dinner too.'</td>
</tr>
<tr>
<td></td>
<td>'María will come for dinner too.'</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>¡Trape el champán, que tenemos que celebrar la buena noticia!</td>
<td>'Bring the champagne! We have to celebrate this good news!'</td>
</tr>
<tr>
<td></td>
<td>'Bring the champagne! We have to celebrate this good news!'</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>Pablo cree dónde está María.</td>
<td>'Pablo thinks he knows where Maria is.'</td>
</tr>
<tr>
<td></td>
<td>'Pablo thinks he knows where Maria is.'</td>
<td></td>
</tr>
<tr>
<td>Filler</td>
<td>¿Dónde Carmen da a quién qué?</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'To whom does Carmen give what where?'</td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>Está mañana he en la carretera de nuevo encontrado un billete de 50.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'This morning I again found a 50 dollar bill on the street.'</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>Todos los años a Rodrigo Carmen visita por Navidades.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'Every year, Carmen visits Rodrigo at Christmas.'</td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>¡No mentiras digas!</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'Don’t tell lies!'</td>
<td></td>
</tr>
</tbody>
</table>