Complementizers on edge

On the boundaries between syntax and pragmatics in Ibero-Romance

Anna Kocher

Open Romance Linguistics 1

language science press

Open Romance Linguistics

In this series:

1. Kocher, Anna. Complementizers on edge: On the boundaries between syntax and pragmatics in Ibero-Romance.

Complementizers on edge

On the boundaries between syntax and pragmatics in Ibero-Romance

Anna Kocher

language science press

Anna Kocher. 2022. *Complementizers on edge: On the boundaries between syntax and pragmatics in Ibero-Romance* (Open Romance Linguistics 1). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog/book/350

ISBN: 978-3-96110-391-1 (Digital)
978-3-98554-049-5 (Hardcover)

DOI: 10.5281/zenodo.6907848
Source code available from www.github.com/langsci/350
Errata: paperhive.org/documents/remote?type=langsci&id=350

Cover and concept of design: Ulrike Harbort
Proofreading: Amir Ghorbanpour, Elliott Pearl, Fahad Almalki, Gerald Delahunty, Hossep Dolatian, Irene Fally, Jean Nitzke, Jeroen van de Weijer, Lachlan Mackenzie, Neneng Sri, Sean Stalley
Fonts: Libertinus, Arimo, DejaVu Sans Mono
Typesetting software: XeLaTeX

Language Science Press
xHain
Grünberger Str. 16
10243 Berlin, Germany
http://langsci-press.org

Storage and cataloguing done by FU Berlin

Contents

Acknowledgments

This book is a matured and reworked version of my PhD dissertation. Among all the people who inspired me, encouraged me and asked me hard questions during my graduate years, my PhD advisor Eva-Maria Remberger deserves a special mention here and, of course, my most sincere gratitude. I also thank my editors, in particular Anne Wolfsgruber, for their support, and the reviewers for their insightful comments. I furthermore thank Felix Kopecky, Sebastian Nordhoff and the community proofreaders at Language Science Press for their excellent and thorough work. Finally, I thank António Variações and Leonard Cohen who kept me company and, without knowing it, wrote the soundtrack for this book.

1 Introduction

This monograph explores the question of how the interface between syntax and pragmatics is organized in Ibero-Romance. The principal aim is to investigate where the boundary between these two components of grammar lies, through the use of tools from generative syntax and formal pragmatics.

The empirical focus is on what I will call root-clause complementizer constructions.[1] With this term I refer to cases in which the complementizer *que* shows an unexpected behavior and appears in apparently unembedded contexts. The constructions prove particularly interesting with respect to the question I address in this book. Complementizers are traditionally considered to serve the syntactic function of marking a sentence as embedded. However, in the contexts that I investigate, they do not carry out this function; instead their presence has an impact on the pragmatics of the sentence that contains them.

A complementizer can generally be defined as the element that identifies a sentential complement and that has the function of subordinating a dependent sentence (see for instance Noonan 2007, Kehayov & Boye 2016). This means that we typically expect to encounter complementizers only in embedded contexts. The data in (1) and (2) show that based on this definition the Ibero-Romance word *que* qualifies as a complementizer. The sentences following *que* are all sentential arguments dependent on the matrix clause verb. In (1) the sentences introduced by *que* function as the object of the verb 'say' and in (2), as the subject of the copula verb 'be'.

(1) a. Catalan
La Joana diu que la Maria és lingüista.
the Joana say.3SG.PRS that the Maria be.3SG.PRS linguist

b. Portuguese
A Joana diz que a Maria é linguista.
the Joana say.3SG.PRS that the Maria be.3SG.PRS linguist

[1] In this book, I use the term *construction* in a theory-neutral way to refer to a pairing of a productive abstract structure with a compositional meaning.

c. Juana dice que María es lingüista.
Juana say.3SG.PRS that Maria be.3SG.PRS linguist
'Jo/uana says that Maria is a linguist.'

(2) a. Catalan
Que la Maria sigui lingüista és fantàstic.
that the Maria be.3SG.SBJV.PRS linguist be.3SG.PRS fantastic

b. Portuguese
Que a Maria seja linguista é fantástico.
that the Maria be.3SG.SBJV.PRS linguist be.3SG.PRS fantastic

c. Spanish
Que María sea lingüista es fantástico.
that Maria be.3SG.SBJV.PRS linguist be.3SG.PRS fantastic
'That Maria is a linguist is fantastic.'

The data in (3) and (4) show that the element *que* is furthermore present in modifiers. In (3) it introduces a relative sentence that modifies a DP.[2] The examples in (4) show that *que* can co-occur with adverbs when introducing sentential adjuncts.[3]

(3) a. Catalan
La Joana parla d' un llibre que està llegint.
the Joana talk.3SG.PRS of a book that AUX.3SG.PROG.PRS reading

b. Portuguese
A Joana está a falar dum livro que está
the Joana AUX.3SG.PROG.PRS to talk of a book that AUX.3SG.PROG.PRS
a ler.
to read

c. Spanish
Juana habla de un libro que esta leyendo.
Juana talk.3SG.PRS of a book that AUX.3SG.PROG.PRS reading
'Jo/uana is talking about a book that she is reading.'

[2]There are opposing views regarding whether *que* is best analyzed as a relative operator or a complementizer in certain relative sentences. On Spanish, see for instance Rivero (1982), who considers it a relative operator, and Arregi (1998), Brucart (1992) who consider it a complementizer.

[3]The presence of the complementizer in sentential adjuncts is subject to adverb-, language- and register-dependent variation.

(4) a. Catalan
Has de netejar el teu dormitori abans que
have.2SG.PRS to clean the your room before that
arribi jo.
arrive.1SG.SBJV.PRS I

b. Portuguese
Tens de limpar o teu quarto antes que eu
have.2SG.PRS to clean the your room before that I
chegar.
arrive.1SG.SBJV.FUT

c. Spanish
Has de limpiar tu cuarto antes que llegue
have.2SG.PRS to clean your room before that arrive.1SG.SBJV.PRS
yo.
I
'You have to clean your room before I arrive.'

The phenomena that are the central concern of the present monograph pose a problem for a traditional notion of complementizer. In these constructions *que* does not seem to exhibit its typical function of identifying a complement clause but appears in root contexts without being syntactically embedded. The presence of the complementizer, contrary to the examples above, is not obligatory in these sentences. It is not required strictly for syntactic reasons but the presence of the complementizer has an impact on the interpretation of the sentence that it introduces.

(5) a. Catalan
(Que) la Maria ve a la festa.
QUE the Maria come.3SG.PRS to the party

b. Portuguese
(Que) a Maria vai à festa.
QUE the Maria go.3SG.PRS to the party

c. Spanish
(Que) María viene a la fiesta.
QUE Maria come.3SG.PRS to the party
'[reportative:] Maria is coming to the party.'

One interpretation of the examples in (5) is that they constitute reported speech. This interpretation is usually supported by the contextual presence of

a *verbum dicendi*. In the sentences in (6–8) the presence of *que* has a different effect.

(6) a. Catalan
Certament (que) és un arros bó.
certainly QUE be.3SG.PRS a rice good

b. Portuguese
Certamente (que) é um arroz bom.
certainly QUE be.3SG.PRS a rice good

c. Spanish
Ciertamente (que) es un arroz bueno.
certainly QUE be.3SG.PRS a rice good
'Certainly, this is a great rice.'

(7) a. Catalan
Que bó (que) és aquest arròs!
how good QUE be.3SG.PRS this rice

b. Portuguese
Que bom (que) é este arroz!
how good QUE be.3SG.PRS this rice

c. Spanish
¡Qué bueno (que) es este arroz!
how good QUE be.3SG.PRS this rice
'How great this rice is!'

(8) a. Catalan
Sí (que) és un arròs bó.
VERUM QUE be.3SG.PRS a rice good

b. Spanish
Sí (que) es un arroz bueno.
VERUM QUE be.3SG.PRS a rice good
'This IS a great rice.'

In all of these examples the proposition introduced by *que* is marked as information the hearer should already know. In the constructions that give rise to this interpretation, the complementizer often co-occurs adjacent to other left-peripheral material. For instance in (6), the complementizer follows the epistemic adverb 'certainly', in (7) it follows the *wh*-expression of a *wh*-exclamative and in

(8), it follows the verum marker *sí*.[4] In these constructions the complementizer shows a different syntactic distribution than in reportatives, where these same expressions are preceded rather than followed by *que*.

The examples presented so far suggest that there are at least three different functions expressed by *que*: It functions as a typical subordinating complementizer in (1), (2), (3) and (4), it marks a sentence as reported speech in (5) and indicates that the content of the sentence constitutes known information in (6), (7) and (8). This could lead to the conclusion that we are dealing with three different lexical items: One *que* that is a subordinator and two *que*s that are pragmatic markers, one of which gives rise to a reportative interpretation while the other imposes on the hearer a commitment to the proposition in the scope of *que*. While this is doubtless a valid line of reasoning, this book will present arguments for precisely the opposite view. In the following chapters, I will present evidence in support of an analysis that does not propose multiple lexical items with dedicated functional meanings, but rather assumes only one lexical item with an underspecified meaning.

1.1 Aims, scope and motivation

1.1.1 Aims

The central aims of this book are to determine where the boundaries between syntax and pragmatics lie, how these components of grammar interact and how the interaction is most adequately modeled within a formal approach to grammar. These questions will be addressed in the light of root clause complementizer constructions. A further goal is therefore also to investigate the structure of these constructions and develop an analysis that can account for their empirical distribution. The results of this investigation provide information on the interaction of *que* with other left-peripheral material and hence lead to a broader understanding of the left periphery in Ibero-Romance languages.

An additional goal is to gain insights into the properties of the complementizer itself. As stated above, the empirical focus of this book is those examples in which we encounter an item that looks like a complementizer but does not behave in accordance with the traditional notion of complementizer. In principle this leaves us with two choices. The first option would be to conclude that these are not complementizers and that there are multiple lexical items that are spelled

[4] European Portuguese makes use of a different structure to express verum. The different strategies are discussed in §3.3.5.

out as *que*, each encoding a specific meaning. The alternative option I propose is to assume that there is only one *que* and that it is the notion of complementizer that needs to be adjusted. In other words, *que* is a complementizer but, crucially, a complementizer is something other than what we thought it was. One of the objectives of this book is therefore to find empirical evidence in favor of this second option and to show that, a revised notion of complementizer allows us to account for apparently atypical behavior like that illustrated in (5–8). One empirical point that motivates the assumption that there is only one *que* is that there is no formal distinction that would suggest that there are multiple types of *que*. Independent of its function, it is always spelled out the same way. Furthermore, although it does surface at different points in the functional field, *que* is restricted to appearing in the left periphery of a sentence, the natural habitat of a complementizer. If the item had gained a different function as a pragmatic marker through a process of grammaticalization, for instance, greater syntactic mobility might be expected.

1.1.2 Scope

The examples I discuss in this book are taken from Catalan, Spanish and Portuguese. Most generalizations I present hold for European and non-European varieties of the latter two languages. The Portuguese data that stem from a corpus are identified as either European or Brazilian. Where relevant, comparisons are drawn with similar phenomena in other Romance and non-Romance languages. The examples are taken from a range of sources. Some evidence is drawn from corpus data. For Spanish, I consulted the Corpus del Español (henceforth CdE), making use of the contemporary portion of the 2001, 100 million token, Genre/Historical subcorpus and the entire 2016, 2 billion token, Web/Dialects subcorpus. For Portuguese, I consulted the equivalent Corpus do Português (henceforth CdP), again making use of the contemporary portion of the 2006, 45 million token, Genre/Historical subcorpus and the entire 2016, 1.1 billion token, Web/Dialects subcorpus. These corpora are useful in studying the phenomena under investigation in this book: They are typically employed in informal registers and the databases contain oral data (Genre/Historical subcorpus) and web data (Web/Dialects subcorpus) in which informal registers usually prevail. The corpora have the additional advantage of being annotated, which facilitates the query. For Catalan, there are no comparable annotated corpora that are publicly available. I mainly relied on the 2014 780 million token Catalan Web as Corpus (henceforth caWac). This corpus does not have an online interface but can be downloaded freely. The Catalan database also included a small self-compiled ebook corpus

(400,000 tokens) (henceforth ebook-cat). In addition to the corpus data, I elicited judgments on grammaticality and acceptability of constructed examples. My informants were predominantly native speakers of the European varieties of Portuguese and Spanish, and for Catalan, speakers of Central and Balearic Catalan. The experimental stimuli in Chapter 4 stem from the corpora listed above. The judgments in the experiments are elicited from native speakers mostly of Central Catalan and of the European variety of Spanish.

1.1.3 Motivation

There are a number of considerations that motivated the development of a new analysis to account for the phenomena under investigation, despite the fact that most of them have already been explored in the literature. First, a global goal of this monograph is to adopt a unified perspective and to focus on the shared properties of data that have previously only been examined separately. I attempt to achieve a broader empirical coverage than previous accounts and show how different phenomena are related on an underlying level. I can thus contribute new insights that deepen our understanding of the nature of the complementizer and its interaction at the left periphery, which in turn allows conclusions drawn regarding the central question of how syntax and pragmatics are related. The analyses presented in this book adopt a different modeling of the two relevant components of grammar. The most drastic difference is that my proposal does not rely on a neo-performative hypothesis, while most previous analyses have adopted versions of this approach. Neo-performative hypotheses propose that illocutionary forces and related pragmatic concepts are encoded syntactically, without treating them as deleted performative clauses as in the traditional performative hypothesis presented in Ross (1970). In the neo-performative hypotheses – the most influential of which is developed in Speas & Tenny (2003) – the performative structure is a part of the architecture of the clause envisioned as a functional domain above the CP. An overview of the central assumptions of (neo-)performative hypotheses as well as some criticism of these found in the literature is given in §1.4.

Further motivating factors behind my adoption of this position will emerge over the course of the following chapters, in particular in §2.1 and §3.1, where I review the analyses that have been put forward in the literature and point out potential limitations in some, though not all, of these proposals. I want to stress at this point that the weaknesses of the other accounts are by no means dramatic enough to warrant rejection out-of-hand. The motivation of this investigation is not to argue that the explanation provided in this book is the only valid one. It is instead an attempt to show that the phenomena can be accounted for with

a less inflated syntactic structure. A more limited amount of structure could be brought in as an argument in favor of economy. However, I refrain from presenting my analysis as the more economical alternative, primarily because I do adopt a fairly rich cartographic structure and therefore cannot claim that the assumed structure is minimal in any serious way. A second reason is that assuming less structure in the present account comes at a cost, namely, the attribution of a more dominant role to pragmatic mechanisms; whether this is truly more economical cannot currently be determined. Ultimately, in a theoretical discipline as rich in conceptual alternatives as linguistics, the choice between multiple suitable and convincing analyses comes down to personal preference to some degree. This fairly mundane factor has undeniably played a non-negligible role in developing the present proposal.

1.2 Theoretical background

In this section, I briefly summarize the theoretical background of my analysis. I outline my re-conception of what an Ibero-Romance complementizer constitutes and give a brief introduction to my main assumptions and the minimal adaptations to the cartographic approach to the left periphery, which provides the theoretical framework for my analysis.

1.2.1 The nature of *que*

In traditional conceptions of complementizers, the subordinating function is dominant (cf. the definition I cite at the beginning of this chapter). These conceptions, however, do not seem adequate to capture the behavior of Ibero-Romance *que* exemplified in (5–8). It is therefore necessary to redefine what an Ibero-Romance complementizer is in order to comply with one of the goals of the book: Namely, to maintain that there is only one lexical element *que* and that this element is in fact a complementizer. My proposal is influenced by what Bayer (2002, 2004) concludes about German *was* and its Bavarian cognate *wos*. This word is notoriously polyfunctional, as illustrated by the examples in (9). They show that *was* can take up nominal functions, for instance as an indefinite pronoun in (9a) and (9b), and a *wh*-pronoun in (9c).

(9) German

a. (Bayer 2002: 288: ex 23a)
Ich hab da was gesehen.
I AUX.1SG.PRF.PRS there WAS see.PTCP
'I have seen something.'

b. (Bayer 2002: 288: ex 20b)
Was auf dem Oberdeck saß, war deutsch
WAS on the upper-deck sit.3SG.IPFV.PST be.3SG.IPVF.PST German
und trank Sekt.
and drink.3SG.IPFV.PST sparkling wine
'The people that sat on the upper deck, were German and drank sparkling wine.'

c. (Bayer 2002: 288: ex 19b)
Was hast du gegessen?
WAS AUX.2SG.PRF.PRS you eat.PTCP
'What did you eat?'

Additionally, Austro-Bavarian *wos* appears in embedded contexts and can acquire a subordinating function, for instance in the relative sentence in (10), where *wos* is preceded by a relative pronoun.

(10) Austro-Bavarian (Bayer 2002: 290: ex 26a)
die Frau (die) wos am Eck Wiaschtln vakauft
the woman who WAS at.the corner sausage.PL sell.3SG.PRS
'The woman who sells sausages at the corner'

The examples in (11) suggest that *wos* can be a subordinating item and can simultaneously function as a *wh*-pronoun in (11a). This is supported by the fact that in this example, a doubly filled CP – which could suggest that the two functions are distributed across two items and which is grammatical otherwise (cf. 11c) – is not grammatical here (cf. 11b).

(11) Austro-Bavarian (Bayer 2004: 4: 9a-c)

a. I woaß, wos–a gern trinkt.
I know.1SG.PRS WAS-he preferably drink.3SG.PRS
'I know what he likes to drink.'

b. * I woaß, wos dass–a gern trinkt.
I know.1SG.PRS WAS that-he preferably drink.3SG.PRS

c. I woaß, wos fiar–a Bier dass-a gern trinkt.
I know.1SG.PRS WAS for-a Bier that-he preferably drink.3SG.PRS
'I know which beer he likes to drink.'

Bayer's solution to the puzzle is that *wa/os* is a maximally underspecified item which acquires its function contextually. My proposal for Ibero-Romance *que* is

very similar: I also consider it to be underspecified. In my analysis this is translated to mean that it carries an unvalued feature. The Ibero-Romance complementizer is therefore simply a lexical item with the form *que* that has an unvalued feature and that is merged in a position in the CP. Depending on its merge position, a different value and consequently a different functional meaning are acquired. In order to account for the data that are the core of the investigation, I assume that *que* is valued with a *subordinate* feature when it is merged in the highest projection of the left periphery and with an *attributive* feature when it is merged in the lowest projection of the left periphery. It will be shown over the course of this book that these features are not stipulated to account for the unembedded constructions exemplified in (5–8); on the contrary, they are the same features that *que* acquires in syntactically embedded contexts like (1–4). This of course strengthens the claim that there is only one *que*.

This re-conception of *que* could be put to use when accounting for the interrogative pronouns Catalan *què*, Portuguese *que/quê*, Spanish *qué*. One idea that could be developed on this basis would be that these are again expressions of an underspecified *que* which receives a focus feature which has consequences for its prosodic make-up but also for its interpretation.

1.2.2 Cartographic approach

The theoretical approach of this book is that of generative linguistics, which determines the aims, argumentation, analysis and diagnostics that are developed and employed. More precisely, the analysis is formulated within a cartographic approach, and I adopt the assumption that the complementizer phrase (CP) (Chomsky 1986) is split into a universal hierarchy of functional projections (Rizzi 1997). The functional projections populating the left periphery mediate the interface between syntactic structure, interpretation and prosody. One assumption of the cartographic project is that the interpretative and prosodic properties are directly read off the syntactic structure (see Belletti 2004, Bocci 2009). The proposal of a split CP is related to and inspired by similar ideas that motivated a splitting of phrases into hierarchically ordered projections in the nominal domain (Abney 1987, Cinque 1994, Longobardi 1996) and in the inflectional domain (Pollock 1989, Belletti 1990, Ouhalla 1991, Cinque 1999). The development of a richer and more articulate functional field above the IP is supported by empirical findings relating to word order restrictions on sentence peripheral material such as complementizers, interrogative pronouns, topics and foci. The particular implementation of the idea that I adopt was initially proposed by Rizzi (1997) and

was later further refined (see Rizzi 2001, 2004b, 2013 and references therein).[5] While the empirical focus was initially on Romance languages, support for the universal nature of the hierarchy also stems from cross-linguistic evidence, for instance from Germanic (Grewendorf 2002, Haegeman 2004, 2006, 2012), Japanese (Saito 2012, Belletti 2013), Semitic (Shlonsky 2000, 2014) and Niger-Congo languages (Aboh 2004, 2010, Torrence 2013). The full hierarchy of the functional heads is given in (12).

(12) [Force [Top* [Int [Top* [Foc [Mod* [Top* [Fin [IP]]]]]]]]]
(Rizzi 2013)

ForceP, at the left edge of the functional field, is established in Rizzi (1997) as a projection that encodes the clause type of a sentence. Subordinating complementizers are assumed to occupy the head of this projection, cf. (13a) and (13b). The examples show that a clitic left dislocated topic targeting a Top-position is only grammatical below the complementizer (as in 13a) but not above it as in (13b). This is in line with the idea that the complementizer occupies ForceP, the highest position in the left periphery that has no Top-projection above it. In Rizzi (1997) subordinating complementizers express the clause type of a sentence. In more recent publications, some authors have re-purposed ForceP, or decompositions thereof, as a projection mediating clause types and illocutionary force (Speas & Tenny 2003, Coniglio & Zegrean 2010, Corr 2016, among many others).

(13) Italian (Rizzi 1997: 288: ex 10a,b)

a. Credo [$_{\text{ForceP}}$ che] [$_{\text{TopP}}$ il tuo libro$_i$], loro lo$_i$
believe.1SG.PRS that the your book they CL.AKK
aprezzerebbero molto.
appreciate.3PL.COND much
'I believe that your book, they would appreciate it a lot.'

b. * Credo, il tuo libro$_i$, che loro lo$_i$
believe.1SG.PRS the your book that they CL.AKK
aprezzerebbero molto.
appreciate.3PL.COND much

FinP delimits the functional field at the lower edge. Rizzi (1997) proposes that this projection is related to finiteness, and is the host of the non-finite counterpart of the finite complementizer (cf. 14a and 14b). The examples apply the same

[5] Rizzi's account was influenced by previous works proposing multiple functional projections in the left periphery, for instance by Reinhart (1981), Uriagereka (1988, 1995), Brody (1990, 1995) and Culicover (1992).

diagnostic as above with Force. A clitic left dislocated topic is grammatical above the non-finite complementizer (14b) but not below it (14b), which is in keeping with the assumption that *di* occupies the lowest left-peripheral position. FinP has been proposed as the projection that is targeted by the finite verb in Germanic verb-second configurations (Roberts 2004).[6] An important observation in the context of this book is that in certain constructions a finite complementizer can also be merged in this position (see for instance Belletti 2009, 2013, Ledgeway 2005).

(14) Italian (Rizzi 1997: 288: ex 11a,b)

a. * Credo di il tuo $libro_i$ $aprezzarlo_i$ molto.
believe.1SG.PRS to the your book appreciate.CL.M.SG much

b. Credo $[_{TopP}$ il tuo $libro_i]$ $[_{FinP}$ di] $aprezzarlo_i$
believe.1SG.PRS the your book to appreciate.CL.M.SG
molto.
much
'I believe to appreciate your book a lot.'

IntP is postulated as the location in which the interrogative complementizer is merged. Evidence for this comes from Spanish embedded polar questions like (15), which permit the co-occurence of the finite and the interrogative complementizer and in which, crucially, the interrogative complementizer *si* follows *que*. IntP is furthermore the location of expressions like Italian *perché* 'why' (cf. Shlonsky & Soare 2011). Moreover, IntP has been proposed as the host of the complementizer in complementizer-initial polar questions in Sicilian (Cruschina 2012) and Catalan (Kocher 2017b) (but see §3.3.2 for my revised take on complementizer-initial polar questions).

(15) Spanish (Suñer 1994: 349: ex 30b)
Me preguntaron $[_{ForceP}$ que] $[_{IntP}$ si] tus amigos ya
CL.1SG ask.3PL.PRF.PST that whether your friends already
te visitaron en Granada.
CL.2SG visit.3PL.PRF.PST in Granada
'They asked me (that) whether your friends already visited you in Granada.'

FocP is the projection that hosts fronted foci and *wh*-pronouns (cf. 16). Much work has been dedicated to studying the structure, prosody and interpretation of

[6] For a recent account, see Lohnstein (2016) and Kocher (2018a) who propose that the German finite verb in verb second configurations targets the clause typing head MoodP.

foci within a cartographic framework (see for instance Belletti 2004, Cruschina et al. 2015, Cruschina & Remberger 2017b, Bianchi et al. 2016).

(16) Italian (adapted from Rizzi 2013: 203: ex 5)
Credo [$_{\text{ForceP}}$ che] [$_{\text{TopP}}$ a Gianni$_i$] [$_{\text{FocP}}$ IL MIO LIBRO] Piero
believe.1SG.PRS that to Gianni the my book Piero
gli$_i$ doverebbe dare.
CL.3SG should.3SG.COND give
'I believe that to Gianni, Piero should give MY BOOK.'

According to Rizzi (1997), while there is only one FocP per clause, multiple TopPs, hosting topics, are sandwiched between each of the other projections. This is motivated empirically by examples like (17) that illustrate the grammaticality of multiple clitic left dislocated topics in one sentence. Rizzi (1997) does not elaborate on whether these topic positions are distinct in any way. Frascarelli & Hinterhölzl (2007) propose that the different positions correlate with different interpretations.

(17) Italian (adapted from Rizzi 1997: 290: ex 21)
[$_{\text{TopP}}$ Il libro$_i$], [$_{\text{TopP}}$ a Gianni$_j$], glielo$_{i,j}$ darò senz'altro.
the book to Gianni CL.DAT.CL.AKK give.1SG.FUT for sure
'The book, to John, I'll give for sure.'

ModP is introduced as an additional projection in Rizzi (2004a) as the locus of high sentential modifiers (cf. 18). Some authors (cf. for instance Giorgi 2010: 84, van Gelderen 2011: 248) propose that ModP is split into an Evaluative, Evidential and Epistemic Phrase to accommodate the insights drawn from Cinque (1999), which show that modifiers expressing these meanings follow ordering restrictions.

(18) Italian (Giorgi 2010: 77: ex 38)
[$_{\text{FocP}}$ A PARIGI] [$_{\text{ModP}}$ probabilmente] Paolo è già
to Paris probably Paolo AUX.3SG.PRF.PST already
stato (non a Londra).
be.PTCP not to London
'To PARIS probably Paolo has already been (not to London).'

1.2.3 Adaptations

In this book, I assume the slightly adapted version in (19) of the original functional hierarchy presented in (12).

(19) [SubP [TopP* [IntP [TopP* [FocP [ModP* [TopP* [MoodP [TopP* [FinP [IP]]]]]]]]]]

The two adaptations are inspired by Haegeman (2006) (see also Roussou 2010). The first change is the replacement of ForceP by SubP, a functional projection that hosts subordinating conjunctions. According to Haegeman (2006: 1661), Sub is identified as the functional head that subordinates a clause and makes it available for selection. Contrary to Rizzi's ForceP, this function is independent of the sentential force of the clause. That these two functions can be expressed by independent items follows from data such as (15). In this example, the sentential force is directly encoded through the interrogative complementizer *si*, located in IntP. Sentential force is therefore clearly separate from the subordinating function linked to *que*, which occupies the highest projection of the left periphery, SubP.

The second change I make is the introduction of a clause typing head, not at the left edge, but in the lower section of the left periphery. I adopt (sentence) MoodP from Lohnstein (2016) as the functional projection responsible for clause typing. Notably, the position of MoodP is identical to the position in which Haegeman (2006) relocates ForceP, which overlaps in its function. I use the term Mood instead of Force in order to prevent the unintended conflation of this functional projection with a recent conception of ForceP which links it to illocutionary force (Speas & Tenny 2003, Coniglio & Zegrean 2010, Corr 2016, among many others).

1.3 The analysis in a nutshell

In this section, I provide a brief preview of the main points of the analysis, with theoretical and empirical support for this analysis outlined in the following chapters. There are three main assumptions that determine the analysis. First, that there is only one item *que* in the Ibero-Romance lexicon. Second, that this element is underspecified in the sense outlined in §1.2. And third, that the syntactic position in which the complementizer is merged has an impact on its meaning. The three assumptions are not independent from each other. The first assumption is founded on the fact that there is no formal distinction between the instances of *que* in the different constructions I investigate. The theoretical solution I propose is that there is only one underspecified lexical item. This means that different interpretations of *que* are not encoded lexically. They must nevertheless have an explanation. This is where the third assumption comes in, which states that there is a relation between the syntactic structure and the resulting interpretation. This book provides the empirical evidence for these assumptions.

The gist of the analysis is that the complementizer is valued with an interface feature in the position in which it is merged. This feature has an impact on the interpretation of the proposition in the scope of the complementizer. In order to account for the data that are central to this book, I propose that there are two positions that the complementizer is externally merged in: one at the lower edge of the left periphery (FinP) and one at the higher edge of the left periphery (SubP). Additionally, there is evidence for at least a third left-peripheral position (MoodP) where the complementizer can be externally merged in. Although complementizers merged in MoodP are not a central concern of this book, I briefly return to the core properties of the construction involving a complementizer merged in MoodP in §2.3.1 and §3.3.5. A scheme of the relevant functional projections and their corresponding features is given in Figure 1.1.

The complementizer in *que*-initial reportative examples like (20a) is analyzed as appearing in SubP, the highest projection of the left periphery, cf. (20b).

(20) Spanish

a. Que Juan viene.
QUE Juan come.3SG.PRS
'[reportative] Juan is coming.'

b. [$_{\text{SubP}}$ Que$_{\text{subordinate}}$ …[$_{\text{IP}}$ Juan viene.]]

The functional head provides a subordinate feature. The consequence is that sentences introduced by a complementizer valued with this feature are interpreted as subordinate. This feature is not postulated specifically to account for the phenomenon at hand but is assumed to be present in all (finite) subordinate sentences, cf. (21a) and (21b).

(21) Spanish

a. María ha dicho que Juan viene.
María AUX.3SG.PRF.PRS say.PTCP QUE Juan come.3SG.PRS
'María said that Juan is coming.'

b. [$_{\text{SubP}}$ …[$_{\text{IP}}$ María ha dicho [$_{\text{SubP}}$ que$_{\text{subordinate}}$ …[$_{\text{IP}}$ Juan viene.]]]]

While both sentences are introduced by a complementizer carrying a subordinate feature, the difference between the unembedded sentence in (20a) and the embedded sentence in (21a) is that the latter is selected by a matrix clause while the former remains unselected. My claim is that sentences like (20a) receive a reportative interpretation because a verb of saying can be pragmatically reconstructed from the context.

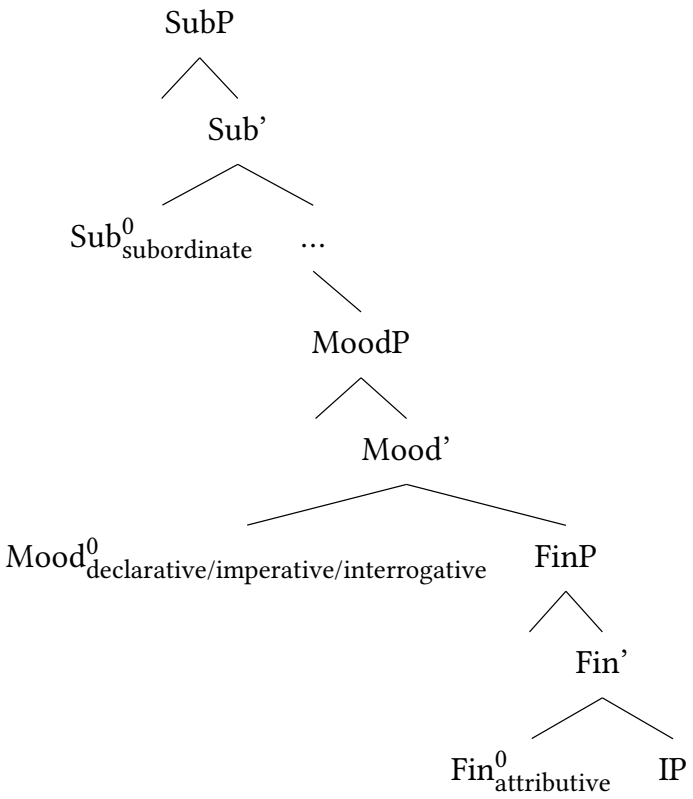


Figure 1.1: SubP and MoodP and their corresponding features

The complementizer in the other root-complementizer constructions examined here is analyzed as being merged in FinP. One example and a sketch of the corresponding analysis are given in (22a) and (22b).

(22) Spanish

a. Ciertamente que Juan viene.
certainly QUE Juan come.3SG.PRS
'Certainly, Juan is coming.'

b. [$_{\text{SubP}}$ …[$_{\text{ModP}}$ Ciertamente que$_{\text{attributive}}$ [$_{\text{FinP}}$ t [$_{\text{IP}}$ Juan viene.]]]]

FinP provides an attributive feature. Once again the feature has a consequence for the interpretation of the sentence. I propose that a commitment to the proposition in the scope of the attributive complementizer is attributed to the hearer. Again, this feature is not stipulated solely for the specific phenomenon under investigation here. It is influenced by de Cuba & MacDonald (2013) who assume a similar feature which they call *referential*, to account for the structural and interpretive difference between factive and non-factive complement clauses.

One final fact that requires further explanation is that attributive *que* surfaces at different points in the functional field in the different constructions. The proposal I put forward is that the surface positions are reached through head-to-head movement of the complementizer. The word order we observe in the constructions is predicted correctly if we assume that the movement of the complementizer is conditioned by its inability to cross a specifier that contains material that was externally merged in its current position.

1.4 Performative and neo-performative hypotheses

In recent years, neo-performative accounts have received considerable attention. Despite the popularity of a neo-performative vision, this book does not subscribe to this approach. This section outlines the central ideas of (neo)-performative hypotheses and some criticism from the literature.

One shared approach of neo-performative hypotheses is that certain pragmatic aspects are treated within syntax. The modern adaptations of this view of pragmatics are grounded in the classical performative hypothesis formulated in Ross (1970). Central to Ross's hypothesis is that every sentence is a performative utterance, in which the illocutionary force is directly encoded in the deep structure (DS) component of the transformational grammar (see also Katz & Fodor 1963, Sadock 1969, 1974). The illocutionary force is expressed by a performative verb that embeds the main clause. This performative verb is later deleted through *performative deletion* yielding the surface structure (SS) that we observe.

(23) a. [$_{\text{DS}}$ I tell you that I read Ross 1970.]
b. [$_{\text{SS}}$ ~~I tell you that~~ I read Ross 1970.] (via *performative deletion*)

(24) a. [$_{\text{DS}}$ I ask you whether Q you have read Ross 1970.]
b. [$_{\text{SS}}$ ~~I ask you whether Q~~ Have you read Ross 1970?]
(via *performative deletion* and *subject auxiliary inversion*)

The motivation for a performative hypothesis is to make a pragmatic theory of illocutionary force à la Austin (1961) obsolete by pushing the burden onto syntax and truth-conditional semantics. The following paragraphs go through some of the arguments cited in favor of Ross's hypothesis and contrast them with some of the criticism formulated in Levinson (1983: 246–276).

One of the arguments in favor of assuming a performative structure is the fact that first-person (25b) and second-person (25d) reflexives are licensed in contexts where their third-person counterparts are ungrammatical. Proponents of

the performative hypotheses explain this contrast syntactically: Reflexives have to be bound by an antecedent in their local domain, cf. (25a), (25c). They conclude that these data constitute evidence that the speaker and the addressee must be encoded syntactically within an implicit performative clause.

(25) a. Tom believed that the paper had been written by Ann and himself.
(Ross 1970: 226: ex 11b)

b. This paper was written by Ann and myself/him*self.
(adapted from Ross 1970: 228: ex 21a)

c. Herbert told Susan that people like herself are rare.
(Levinson 1983: 248: ex 33)

d. People like yourself/her*self are rare.
(adapted from Levinson 1983: 248: ex 34)

Levinson (1983) offers a different explanation. According to him, the licensing of speech-act-participant vs. non-speech-act-participant is a pragmatic rather than a syntactic issue. He states that *himself/herself* is only infelicitous at the beginning of a discourse but is felicitous in other contexts. An example of this can be found in (26).

(26) He [Zapp] sat down at the desk and opened the drawers. In the top right-hand one was an envelope addressed to himself.
(Zribi-Hertz 1989: 716: ex 65)

The reasoning behind the pragmatic explanation is that speaker and hearer are always active and can be addressed, whereas third-person participants need to be salient in order for it to be possible to refer to them using a reflexive, cf. also §2.5.

Another argument typically raised in favor of a performative structure is speech act adverbs like *frankly* in (27a). The idea is that they modify an implicit performative verb.

(27) a. Frankly, I don't care.

b. ~~I tell you~~ frankly ~~that~~ I don't care.

One problem for this argument is that speech act adverbs also appear in a syntactic location where they cannot be trivially analyzed as modifying the high performative clause. For instance, they can modify certain types of embedded clauses as in (28a). One attempt to rescue the performative hypothesis is to propose a second implicit performative clause preceding the embedded clause, which

is then in turn modified by the speech act adverb. This, however, derives the wrong meaning for the *because*-clause: Clearly "I tell you something because I tell you something else." is not the intuitive meaning of (28b).

(28) a. I voted for Labour because, frankly, I don't trust the Conservatives. (Levinson 1983: 262: ex 85)

b. I tell you that I voted for Labour because I tell you frankly I don't trust the Conservatives. (Levinson 1983: 262: ex 86)

Another issue for the speech-act-adverb argument is the fact that some of them only appear with explicit performatives (29a), hence the infelicity of (29b). This is not expected given the proposal put forward by proponents of the classical performative hypothesis.

(29) a. I hereby order you to polish your shoes. (Levinson 1983: 255: ex 53)

b. ? Hereby polish your shoes. (Levinson 1983: 255: ex 54)

In some cases, the adverb does not appear to modify the relevant implicit performative. (30a) is most adequately paraphrased by (30c) rather than (30b). This is an issue because the performative hypothesis assumes a one-to-one mapping between illocutionary force and performative verb. Therefore, for questions like (30a), the performative verb should be of asking (30b) rather than of answering (30c).

(30) a. Briefly, who do you think will win the gold medal? (Levinson 1983: 256: ex 60)

b. I ask you briefly, who do you think will win the gold medal? (Levinson 1983: 256: ex 61)

c. Tell me briefly, who do you think will win the gold medal? (Levinson 1983: 256: ex 62)

One core problem of the performative hypothesis is that it predicts that (31a) and (31b) have the same truth conditions.

(31) a. The world is flat. (Levinson 1983: 252: ex 42)

b. I stated to you that the world is flat. (Levinson 1983: 252: ex 43)

This does not do justice to the intuition that (31a) is a false statement about the actual round world we are inhabiting. In contrast, (31b) is true if in fact I made this statement. This is irrespective of whether the world is flat or not (cf. §2.5 where this fact is picked up again).

Another fundamental issue relates to the assumed direct mapping from clause type to illocutionary force via a performative verb. This results in problems when dealing with indirect speech acts. For instance, the interrogative in (32) can be interpreted as a question, a command or even a threat.

(32) Will you do your homework?

Finally, assuming an implicit performative syntactic structure that is interpreted semantically makes the prediction that every sentence should be assigned a truth value. However, not all meaningful sentences express statements that are either true or false. Obviously, questions, commands and exclamatives cannot be evaluated in this way.

These points – merely a selection was presented here, but for further details see Levinson (1983) – show that a classical performative hypothesis faces a number of serious issues. One attempt at rescuing the insights from Ross (1970) is in the recent developments of neo-performative hypotheses. In these, the classical implicit performative verb is replaced by abstract functional categories.[7] Neo-performative hypotheses encode aspects of pragmatics such as illocutionary forces in syntax without treating them as deleted performative clauses. In the neo-performative hypotheses, the performative structure is a part of the architecture of the clause envisioned as another functional domain above the CP (cf. for instance the contributions made by Benincà 2001, Garzonio 2004, Hill 2006, 2007b,a, Speas & Tenny 2003, Speas 2004, Tenny 2006, Poletto & Zanuttini 2003, Zanuttini 2008, Zanuttini et al. 2012, Krifka 2013, Haegeman 2014, Wiltschko 2014).

One very influential proposal is by Speas & Tenny (2003), who assume that pragmatic roles are encoded syntactically. They treat declarative, interrogative, imperative, subjunctive and quotative as the universal types of speech acts. In their system, these speech acts are modeled via different configurations of the pragmatic roles and the utterance content by following universal syntactic principles.

Speas & Tenny (2003) restrict the scope of their analysis to direct illocutionary forces, leaving aside the complications brought by indirect illocutionary forces

[7] But see the light performative hypothesis of Alcázar & Saltarelli (2014), which returns to a more classical version of the performative hypothesis with the difference that the implicit performative clause does not contain a lexical verb (Ross 1970) but a functional light verb.

as illustrated in example (32). Perhaps because they only focus on cases in which the clause type is mapped directly to an illocutionary force, they do not draw a terminological distinction between the clause types and their corresponding illocutionary forces. I follow this terminological imprecision when illustrating their proposal here. Speas & Tenny (2003) postulate an enriched revision of Rizzi's ForceP in the form of two projections above the CP called Speech Act Phrase (SAP) and Sentience Phrase (SenP). The SAP has three arguments: the pragmatic roles *speaker* and *hearer*, and the *utterance content*. The structure is illustrated in Figure 1.2. The structure of the SAP and the SenP are parallel to the vP shell. According to the authors, the lower projections can furthermore be iterated, which is indicated by the asterisk in the structures.

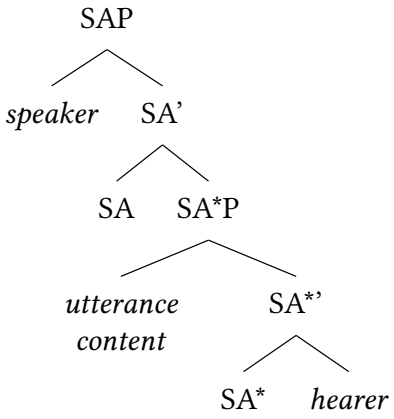


Figure 1.2: The structure of SAP (Speas & Tenny 2003: 320: ex 9)

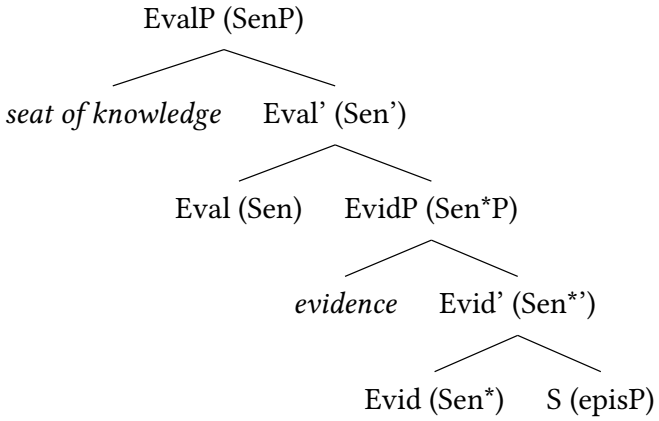


Figure 1.3: The structure of EvalP/SenP (Speas & Tenny 2003: 334: ex 34)

The third pragmatic role represented syntactically in Speas & Tenny (2003) is the *seat of knowledge* that encodes epistemic authority and evaluation of truth. It is located in SenP, the upper structural layer of the utterance content in the scope of SAP, cf. Figure 1.3. It consists of an EvaluativeP hosting *seat of knowledge* and EvidentialP hosting *evidence*. Both projections are adopted from Cinque (1990).

While in the classical performative hypothesis the illocutionary force is encoded directly through the semantic content of the performative verb, the head of the SAP is not considered to be a proper verb. Instead, in Speas & Tenny (2003), different illocutionary forces are derived syntactically through the interplay of two parameters. The first is a feature that marks the utterance content as + or –finite. The second parameter concerns the interaction of the pragmatic roles of *speaker*, *hearer* and *seat of knowledge*. In this model, *seat of knowledge* is controlled by the closest c-commanding pragmatic role. In the default configura-

tion, this role is the *speaker*. In a declarative clause, therefore, *speaker* and *seat of knowledge* coincide. This means that the speaker takes the epistemic authority and evaluates the truth of the utterance content. The basic structure in Figure 1.2 represents the configuration of a declarative.

Questions are derived through a movement operation in which the *hearer* is promoted to the specifier of the iterated lower SAP, cf. Figure 1.4. From this position the *hearer* controls the *seat of knowledge* and takes up epistemic authority.

A similar configuration is proposed for imperatives illustrated in Figure 1.5. The *hearer* also controls the *seat of knowledge* in this case. In order to achieve this, *hearer* is once again promoted to a higher position. The structure of imperatives differs from interrogatives in that the utterance content carries a –finite feature. Furthermore the orientation of the SA* projection dominating the *utterance content* is reversed in imperatives. However, no motivation for this is found in Speas & Tenny (2003).

The –finite equivalent of the declarative structure in Figure 1.2, illustrated in Figure 1.6, is the analysis Speas & Tenny (2003) assume for subjunctives.

Finally, quotatives are treated as declaratives in which the speaker is absent but is replaced by an expletive subject. Speas & Tenny (2003) do not offer a structure for this configuration in their article, but in their review of the proposal, Alcázar & Saltarelli (2014) do, cf. Figure 1.7.

Although I do not adopt the framework of Speas & Tenny (2003) for my analysis, it could nonetheless be useful for the analysis of reportative *que* constructions that are the topic of Chapter 2. However, additional stipulations are required. For one thing, marking the *speaker* as an expletive might run into problems since the *que*-initial reportative construction, for instance, also permits self-reports in which the actual speaker does in fact coincide with the speaker of the report.

Speas & Tenny's proposal, as well a similar proposal by Haegeman (2014), has been widely adopted by authors working on phenomena related to the interface between syntax and pragmatics. The proposal, however, faces a number of issues. Gärtner & Steinbach (2006) are critical of Speas & Tenny's claim that the structure they put forward in Figure 1.2 and its derivations in Figures 1.4–1.7 are universal; they also point out that there are other structures derivable by universal syntactic principles that are disregarded by the authors without a convincing explanation for their omission. The issue is not only of a theoretical nature: These other structures, Gärtner & Steinbach (2006) argue, would give rise to different illocutionary forces that are either not universal or do not exist at all.

Further problems are identified by Alcázar & Saltarelli (2014). They take issue with the fact that only *speaker*, *hearer* and *seat of knowledge* are treated as indexicals. *Speech location* and *speech time* are absent in Speas & Tenny's framework. It

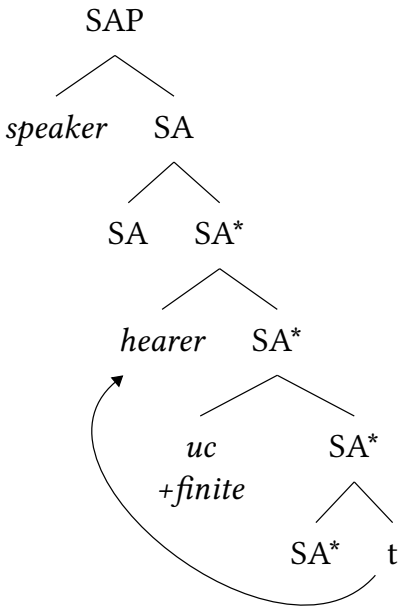


Figure 1.4: Interrogative (Speas & Tenny 2003: 321: ex 10)

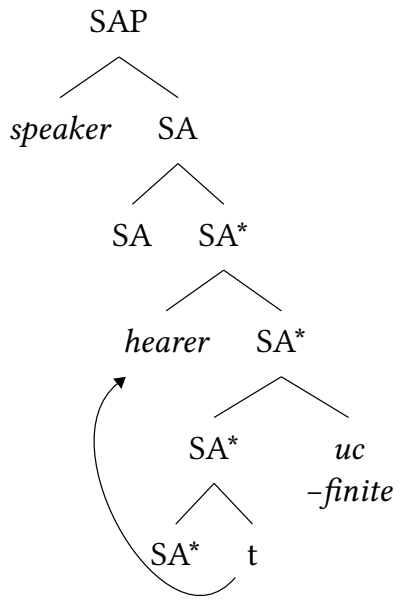


Figure 1.5: Imperative (Speas & Tenny 2003: 322: ex 11)

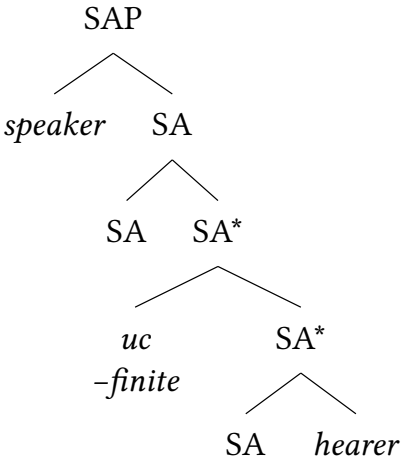


Figure 1.6: Subjunctive (Speas & Tenny 2003: 323: ex 13)

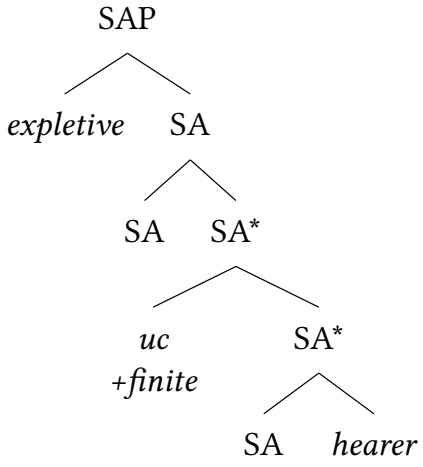


Figure 1.7: Quotative (Alcázar & Saltarelli 2014: 97: ex 24)

should be noted, however, that in later adaptations of Speas & Tenny (2003), this is accommodated. For instance, Corr (2016: 193–194) assumes with Sigurdsson (2010) that location and time are encoded in the syntactic structure. She proposes that EvidP hosts locative features, and EvalP hosts speech time features (see §2.1 and §3.1 for a more detailed discussion of Corr 2016).

Another aspect criticized by Alcázar & Saltarelli (2014) is that the two parameters assumed in Speas & Tenny (2003) are not sufficient to differentiate between different subtypes of certain illocutionary forces. They argue that the configuration proposed for questions, where the *hearer* controls the *seat of knowledge*, only works for genuine but not for rhetorical questions, where one would assume that the *speaker* controls the *seat of knowledge*. In their defense, it should be noted though that rhetorical questions fall into the category of utterances with an indirect illocutionary force and therefore do not fall into the scope of Speas & Tenny's analysis.

The main criticism put forward by Alcázar & Saltarelli (2014) is that subjunctives and quotatives are postulated as universal clause types. They conjecture that this happens purely out of a theoretical necessity, because the two parameters result in four possible configurations. The typological literature (for instance Sadock & Zwicky 1985, König & Siemund 2007), however, does not support this postulation. In fact, there is only agreement on the existence of three universal clause types: declaratives, interrogatives and imperatives.

1.5 Organization of the book

This book is structured as follows: Chapter 2 focuses on *que*-initial reportative sentences such as those illustrated in (5), in which *que* is merged in the top left projection of the split CP and is valued with a subordinate feature. In §2.1, I discuss the three main analyses advocated in the literature by Etxepare (2007, 2010, 2013), Demonte & Fernández Soriano (2014) and Corr (2016). In §2.2, I present my own analysis and compare it to the previous approaches, while in §2.3, I provide the empirical support for my analysis by focusing on the syntactic properties of the construction. In §2.4, I discuss some cross-linguistic differences and show that while the basic underlying syntactic principles are the same in all three languages, there is a pragmatic difference between *que*-initial reportatives in Portuguese on the one hand and Spanish and Catalan on the other hand. §2.5 focuses on the pragmatic requirements to felicitously utter a *que*-initial reportative. The chapter concludes with §2.6, in which I discuss how the analysis can be extended to non-reportative *que*-initial sentences.

Chapter 3 focuses on constructions like (6–8), in which *que* is merged at the right edge of the left periphery and receives an attributive value. In §3.1, I introduce the main analyses presented in the literature for the different constructions by Ambar (2003), Castroviejo (2006), Hernanz (2007), Prieto & Rigau (2007), Demonte & Fernández Soriano (2009), Corr (2016) and Cruschina & Remberger (2018). §3.2 describes the details of my own proposal for the constructions and compares it to the previous analyses. The syntactic properties that support my analysis are explored in depth in §3.3, in which I also discuss some cross-linguistic differences and offer explanations for the contrasts. The last section, §3.4, is dedicated to the pragmatical properties of the constructions.

Chapter 4 presents empirical studies that further explore the pragmatic contribution of attributive *que* in two constructions. The chapter begins with the methodological and statistical background laid out in §4.1 and §4.2. In §4.3, I discuss three of my studies, focusing on the interpretation of attributive *que* following epistemic and evidential modifiers. In §4.4, I present an experimental study investigating attributive *que* in Catalan polar questions. In the final section, §4.5, I reflect on the usefulness of exploratory empirical methods in generative linguistics. Finally, Chapter 5 concludes this book with a summary of the main points and results.

2 The syntax and pragmatics of reportatives

This chapter deals with the phenomenon for which the term insubordination was coined in functional literature (Evans 2007), namely embedded sentences that appear without a matrix clause. The empirical focus of this chapter is *que*-initial sentences that receive a reportative interpretation.

The example in (1) illustrates a typical case of what my analysis refers to as reportative *que*. In this example speaker G introduces a sentence with *que* to mark that he reiterates (part of) his previous utterance which the other speech participant, L, appears to not have fully understood or heard. In the discussion of this example, Gras (2016) says that the omission of *que* in this context, while not ungrammatical, would appear strange to the ears of a European Spanish speaker.

(1) Spanish (Gras 2016: 119: ex 6)
[Context: Three friends talking about the route to pick up a fourth friend.]

G: (bue)no ¿y ahora por dónde nos vamos a ir?
well and now for where us go.1PL.PRS to go

L: ¿adónde?
where

G: ¿que por dónde nos vamos ir?
QUE for where CL.REFL go.1PL.PRS go

'G: Well and now which way should we take? L: Where? G: I said which way should we take?'
(Val.Es.Co.[1] L.15.A.2: 103–107.)

The phenomenon of reportative *que* is at the heart of the question regarding how the boundaries between syntax and pragmatics are organized. While most previous accounts opted for a syntactic explanation when deriving the reportative interpretation, the central argument I develop here is that the reportative interpretation results from pragmatic rather than syntactic reconstruction. Strong

[1]Val.Es.Co. is a corpus of Spanish colloquial conversations (Briz & Grupo Val.Es.Co 2002).

empirical support for this claim comes from the fact that the requirements for syntactic reconstruction are not met in the contexts where *que*-initial reportatives are grammatical. However, there is a pragmatic condition that is found in all the contexts, namely that a verb of saying needs to be salient in order to felicitously utter a *que*-initial reportative.

In the generative syntactic literature, these *que*-initial sentences have sometimes been called *quotatives*; I use the term *reportative* instead, because they do not behave like prototypical direct quotations that repeat an expression verbatim (cf. Coulmas 1986: 2).[2] The example in (2) illustrates that the reportative sentence does not consist of the same words as the original sentence: In the reportative version, the speaker paraphrases and attenuates her initial statement.

(2) Catalan
A: Mentida.
lie
B: Què dius?
what say.2SG.PRS
A: Perdoni. Que no hi estic d'acord.
excuse.2SG.IMP QUE not CL.LOC agree.1SG.PRS
'A: That's a lie. B: What did you say? A: Sorry. [reportative:] I don't agree.' (ebook-cat)[3]

Another property that distinguishes *que*-initial reportatives from quotations is that deictic expressions typically undergo an origo switch. In direct quotations, the speaker that quotes an expression adopts the original speaker's perspective (cf. Coulmas 1986: 2). The origo switch that takes place in *que*-initial reportatives indicates that the deictic center is transferred from the original speaker to the external speaker who reports the sentence. The original sentence in (3a), in which *Juan* is the deictic center, is reported by the original hearer *Maria* in (3b). One result of the reporting is the change in the clitic pronoun from second to first person and the change in verbal morphology from first to third person. It is these adaptations, which are typical for reported but not quoted speech, that have led me to choose the term *reportative* for these constructions.

[2]There is discussion in the literature on different types of quotatives and to what extent the quoted expression must be matched verbatim (cf. Davidson 1968, Cappelen & Lepore 1997, 2003, 2005, Maldonada 1999, Abbott 2003, and references therein).

[3]ebook-cat: a small self-compiled e-book corpus (400,000 tokens).

(3) Catalan

a. [Juan$_j$:] T'$_m$ espero$_j$ a la porta.
Juan CL.2SG wait.1SG.PRS at the door
'[Juan:] I will wait for you at the door.'

b. [Maria$_m$:] Que m'$_m$ espera$_j$ a la porta.
Maria QUE CL.1SG wait.3SG.PRS at the door
'[Maria:] [reportative:] He will wait for me at the door.'

This chapter is structured as follows. In §2.1, I present the previous analyses of this construction by Demonte & Fernández Soriano (2014), Etxepare (2007, 2010, 2013) and Corr (2016). In §2.2, I introduce my own proposal, the gist of which is that the complementizer is analyzed as being merged in SubP, the highest projection of the split CP, where it is valued with the feature *subordinate*. This is conceived of as an interface feature that primarily has consequences for the interpretation of the sentences valued with this feature. My principal claim is that the CP of *que*-initial sentences does not differ from its embedded counterparts. In §2.2, I compare my own proposal to the previous accounts developed in the literature. One of the main differences is that my analysis does not allude to a hidden syntactic structure or elided material. I show that a simple and transparent structure is possible if we assume that a complementizer-initial sentence can remain syntactically unselected. Support for the present analysis is provided in §2.3, which shows that the complementizer in *que*-initial sentences surfaces in the same syntactic position as in their embedded, i.e. selected, counterparts. In §2.4, I show that the apparent syntactic differences between *que*-initial reportatives in Portuguese on the one hand and Spanish and Catalan on the other are not related to the phenomenon under discussion but are a reflex of a more general difference. This leads me to conclude that *que*-initial reportatives in all three languages can be treated with the same basic syntactic analysis. There is, however, a pragmatic difference: In Spanish and Catalan *que*-initial reportatives are felicitous if a host expression is salient in the general context; this means that the host expression is accessible or activated in some way. Crucially, the expression can have been, but does not have to have been, explicitly mentioned in the linguistic context. In Portuguese, however, the expression must be given, i.e. mentioned, in the linguistic context. In §2.5, I draw up a unified characterization, by proposing that the reportative interpretation results from pragmatic rather than syntactic reconstruction. Finally, this chapter focuses primarily on *que*-initial reportatives; in §2.6 I propose that in principle the analysis can be extended to account for other types of unembedded sentences.

2.1 Previous analyses

In this section I summarize how *que*-initial reportative sentences have been treated in the generative syntactic literature.[4] I focus on the analyses presented in Corr (2016), Etxepare (2007, 2010, 2013) and Demonte & Fernández Soriano (2014). The three analyses differ in their empirical coverage: Etxepare (2013) and Demonte & Fernández Soriano (2014) only look at Spanish data, while Corr's analysis extends to other standard varieties and dialects of Ibero-Romance and also accounts for the cross-linguistic variation. All three analyses treat *que* as a complementizer. But the authors each assume that it occupies a different syntactic position and fulfills a different function. It is treated as the head of ForceP in Demonte & Fernández Soriano (2014), the head of a LinkerP in Etxepare (2013) and the head of EvidP, one of the subheads of her split ForceP, in Corr (2016). The biggest difference is how each analysis accounts for the reportative interpretation. Etxepare (2013) and Demonte & Fernández Soriano (2014) propose silent verbs/nouns of saying. Corr (2016) presents a neo-performative hypothesis in which pragmatic roles and functions are encoded in a syntactic layer above the CP (for details on performative and neo-performative hypotheses see §1.4). The reportative interpretation does not arise through the meaning of an elided verb or noun of saying, but is encoded as a feature on the functional head in which the complementizer is merged.

2.1.1 Etxepare (2013) and Demonte & Fernández Soriano (2014)

Etxepare (2013), building on his previous work (2007, 2010), proposes that reportative-complementizer constructions are the visible part of a larger structure. The whole structure is a small clause in which an elided event noun of saying functions as the predicate. The author assumes a silent noun of saying – the silent equivalent of A SAYING – rather than a silent verb of saying because coordinated reported sentences can trigger plural agreement, cf. (4). In his theory, number is restricted to nominals, so he concludes that the silent predicate has to be a noun.

(4) Spanish
Que la lasaña estaba buena y que el vino estaba
QUE the lasagne be.3SG.IPFV.PST good and QUE the wine was
extraordinario resonaron en todo el restaurante
great resound.3PL.PRF.PST in all the restaurant
'The saying that the lasagne was good and a saying that the vine was great resounded in the whole restaurant.' (Etxepare 2010: 620: ex 59)

[4]There is also considerable discussion of the phenomenon in Spanish by authors relying on a functional grammar framework (see for instance Porroche Ballesteros 2000, Pons Bordería 2003, Gras 2010, 2016, Sansiñena et al. 2015).

The sentence introduced by *que* is analyzed as a ForceP that is the subject of the small clause. *Que* functions as the linker occupying the head of the dedicated LinkerP dominating the ForceP. Predicate inversion obtains and leaves the predicate noun in the specifier of LinkerP, see (5).

(5) [$_{\text{LinkerP}}$ [A SAYING] [$_{\text{Linker}}$ *que*] [$_{\text{ForceP}}$ […]]]
(adapted from Etxepare 2013: 98: ex 14, details omitted)

Demonte & Fernández Soriano (2014) revisit Etxepare's data and identify two distinct types of *que*. The first type is a proper complementizer selected by a silent verb of saying, whose properties mostly overlap with Etxepare's description. The second type, however, is not a complementizer in Demonte & Fernández Soriano's view, but a homophonous reportative evidential marker.

I will now examine the core data from Demonte & Fernández Soriano's study and propose that it is still possible to maintain the position that there is only one type of *que*: the complementizer. The second type, which Demonte & Fernández Soriano (2014) posit as an evidential marker, can be interpreted as a version of the former with different pragmatic restrictions because it is employed discourse-initially. In the following paragraphs, I will call this version of *que*-initial reportatives out-of-the-blue reportatives.

In what follows, I will show that the particular behavior that Demonte & Fernández Soriano (2014) observe and that motivates their postulation of a second type, can be explained by the context in which these *que*-initial utterances are found. They are discussed as cases of out-of-the-blue reportatives, meaning they appear in a context where there is no previous utterance that the *que*-initial reportative refers to. In general, in order for an utterance to work at the beginning of a conversation, it must be possible for the addressee to accommodate the information that is not explicitly stated. What a speaker can assume her hearer to accommodate depends on the shared common ground of the speech participants. Intuitively, more general information that can be assumed to be cultural, universal or otherwise common knowledge, is expected to be accommodated more easily than specific and context-dependent information.

While different types of clauses are allowed in a context in which the reported utterance is salient, Demonte & Fernández Soriano (2014) show that out-of-the-blue reportatives are restricted to declaratives. Starting a conversation with a reportative always requires some guess-work. A cooperative hearer is usually prepared to accommodate absent information when he is faced with an assertion, but making sense of a reported non-declarative clause is more difficult.

(6) Spanish
(Oye), que mañana no hay clase.
listen.2SG.IMP that tomorrow not there.be.3SG.PRS class
'Listen, there will be no class tomorrow (someone said/I just heard).'
(Demonte & Fernández Soriano 2014: 16: ex 12a)

To accept an out-of-the-blue reportative like (6), the hearer needs to accommodate that the speaker heard the statement that there will be no class the next day. The hearer can accommodate this reportative statement. He might conjecture that the source of the utterance is irrelevant. Based on shared knowledge, he might also conjecture who the source could have been. For instance, the hearer could conjecture that the speaker has talked to a classmate or the professor who informed her.

The situation is trickier with reported questions, imperatives and exclamatives. Even if not reportative, they must follow a number of requirements in order to be felicitously uttered in an out-of-the-blue context. Questions are typical conversation starters; but in true out-of-the-blue contexts only very general questions like *What's new?* or *How are you?* are felicitous. Questions like *How's your mum doing?*, that have a similar function, namely to initiate a conversation, are arguably not out-of-the-blue as they refer to previous knowledge of the hearer's mum's well-being or health. In any case, starting a conversation with *Hey, someone asked what's new?* is odd.

(7) Spanish (Demonte & Fernández Soriano 2014: 18, ex 16b)
Oye, ¿que hemos ganado la liga?
listen.2SG.IMP that AUX.1PL.PRF.PRS win.PTCP the league
Intended: 'Listen, have we won the league (I just heard)?'

A question like *Have we won the league?* requires a very specific type of common ground in order to be uttered felicitously at the beginning of a conversation: One has to imagine a setting in which the games of the specific league that is referred to are salient to such an extent that they can be talked about without further contextualization. A reported version of this question exemplified in (7) is judged infelicitous according to Demonte & Fernández Soriano (2014). A hearer confronted with an out-of-the-blue question like (7) would very likely be puzzled and uncertain about what is expected of him. *Que*-initial questions do not have the illocutionary force of a question but rather that of an assertion. This means that the speaker does not require an answer of the hearer. The hearer must then assume that the speaker intends to convey information with her statement.

However, retrieving information out of this out-of-the-blue *que*-initial question is not easy. In brief, given that there is no context that the hearer can rely on, a cooperative speaker is unlikely to use a *que*-initial question at the start of a conversation.

Regular, i.e. non-reported, exclamatives are adequate conversation starters just as questions are. In these contexts they usually express an emotion towards some property of the immediate context (for instance *What a beautiful day it is!*, *How nice to run into you!*). They are again awkward as a reported version for similar reasons that reported questions are awkward out-of-the-blue. A hearer would have trouble making sense of an utterance like (8) because he would expect a reported version of an exclamative only if the source and the original exclamative were salient. Reported exclamatives, just like questions, have the illocutionary force of assertions. A cooperative speaker, however, would not use a reported exclamative to convey information at the beginning of a conversation because crucial information necessary to understand the utterance, for instance who produced the original exclamative, is absent.

(8) Spanish (Demonte & Fernández Soriano 2014: 18: ex 16a)
Oye, que ¡qué bonito día hace!
listen.2SG.IMP that what nice day make.3SG.PRS
Intended: 'Listen, what nice day it is (I just heard)!'

Finally, imperatives can readily be used at the beginning of a conversation as well. Consider for instance a context where someone calls out an offender (*Put on your mask!*). In order to sensibly use the imperative in a reportative version, the original order needs to be salient. Again, a reported order like (9) is infelicitous as a conversation starter without a salient antecedent because more questions (*Who said that?*, *Why is this reported?*, *Does he/she expect me to put on a mask?*) are generated than are answered.

(9) Spanish
Oye, que te pongas la mascarilla.
listen.2SG.IMP that CL.REFL put on.2SG.SBJV.PRS the mask
Intended: 'Listen, put the mask on (I just heard).'

A second observation made by Demonte & Fernández Soriano (2014) is that out-of-the-blue reportatives are infelicitous with an explicit source. Starting a conversation with a reportative declarative can be felicitous, as I showed in the discussion around (6), but expressing its source as in (10a) is predicted to be infelicitous by Demonte & Fernández Soriano (2014).

(10) Spanish

a. # (Oye), el profe, que mañana no hay
listen.2SG.IMP the professor that tomorrow not there.be.3SG.PRS
clase.
class
'Listen, the professor was like there will be no class tomorrow.'

b. (Oye), acabo de encontrar nuestro profe. Y él,
listen.2SG.IMP end.1SG.PRS to meet our professor and he
que mañana no hay clase.
that tomorrow not there.be.3SG.PRS class
'Listen, I just ran into our professor. He was like there will be no class tomorrow.'

The example in (10b) shows that the same utterance is perfectly felicitous when the previous context makes the source, the professor, salient. This shows once again, that this property has to do more with the out-of-the-blue nature of the context than with the construction itself.

A further distinction drawn by Demonte & Fernández Soriano (2014) is that out-of-the-blue reportatives do not permit a speech participant to be the source of the reported utterance. For Demonte & Fernández Soriano (2014), (11) is infelicitous because the president cannot report his/her own declaration.

(11) Spanish (Demonte & Fernández Soriano 2014: 154, ex 15)

Ciudadanos, que se ha / que hemos
citizen.PL that CL.REFL AUX.3SG.PRF.PRS that AUX.1PL.PRF.PRS
declarado la guerra.
declare.PTCP the war
'Citizens, someone said that one has/we have declared war.'

Corr (2016), however, convincingly shows that this example is actually infelicitous because a reported sentence introduced by *que* is inappropriate in the formal register required in a situation where (11) could be uttered. This is demonstrated by the fact that, in an informal setting, such as the president making the same report to his/her significant other at home, *que* is felicitous, as in (12).

(12) Spanisch (Corr 2016: 154: ex 16)
Bill, que hemos declarado la guerra.
Bill that AUX.1PL.PRF.PRS declare.PTCP the war
'Bill, [I said] we've declared war.'

Demonte & Fernández Soriano (2014) also state that out-of-the-blue reportatives cannot be fragments, nor foreign words or onomatopeias. This can easily be explained on the basis of my general argument: In order to be a fragmented version of something, a full version needs to be salient. This requires a shared linguistic context that is not present at the beginning of a conversation. Likewise, starting a conversation with a foreign word or onomatopoetic expression is awkward since these require a context that makes them felicitous as a single-word utterance.

This brief discussion suggests that all these properties observed in Demonte & Fernández Soriano (2014) are related to the restrictions that limit the options at the start of a conversation. Thus, the fact that certain properties are not observed when *que*-initial reportatives are used out-of-the-blue can be explained as a result of the pragmatic requirements of discourse-initial statements. In my view, then, there is no need to assume two distinct types of *que*-initial reportative constructions. In agreement with Corr (2016), I assume that there is only one type of *que*-initial reportative construction. Postulating a distinct syntactic object with the function of an evidential marker in the sense of Demonte & Fernández Soriano (2014), in light of the review of the data above, does not appear necessary. The analysis proposed by Demonte & Fernández Soriano (2014: 39: ex 53), that I will discuss henceforth, is repeated in (13).

(13) (V) [ForceP [que ... [IP]]]

In the proposals by Etxepare (2013) and Demonte & Fernández Soriano (2014) *que* is analyzed as a complementizer that heads a clause that is subordinate to silent material. The silent material – a verb in Demonte & Fernández Soriano (2014) and a noun in Etxepare (2013) – contributes the reportative meaning.

Demonte & Fernández Soriano (2014) do not explicitly describe how the silencing of the verb is licensed. According to Etxepare (2013), the elision of the predicate of saying marks that its denotation is given in the common ground; however, no explanation is given as to how the predicate of saying entered the common ground. Since a previously mentioned predicate of saying is not required to render reportative complementizer constructions felicitous (see §2.5), it is not immediately evident how this issue is resolved in Etxepare's analysis.

2.1.2 Corr (2016)

The analysis proposed in Corr (2016) is based on the author's specific assumptions about the general structure of a sentence. She proposes the extension of the highest layers of the clausal structure that is illustrated in (14).

(14) [SAHigh [SALow [EvalP [EvidP [DeclP …]]]]]

She assumes a dedicated utterance domain above the CP (similar ideas have been developed in Benincà 2001, Garzonio 2004, Hill 2006, 2007a,b, Speas & Tenny 2003, Speas 2004, Tenny 2006, Poletto & Zanuttini 2003, Zanuttini 2008, Zanuttini et al. 2012, Krifka 2013, Haegeman 2014, Wiltschko 2014). This so-called Utterance Phrase (hence UP) is split into a high layer termed SAHigh (SpeechAct-high) and a low layer termed SALow (SpeechActlow). The higher layer is oriented toward utterance external aspects that are encoded through an activation feature. The lower level is oriented towards utterance-internal aspects that are encoded through a bonding feature. SALow is furthermore decomposed into projections dedicated to the addressee and the speaker. Corr's motivation for this structure is the observation that CP-external elements like vocatives and certain discourse markers co-occur and that they follow a hierarchical order (see also Moro 2003, Hill 2007b, 2013, 2014, Moreira 2013, Espinal 2013, de Carvalho 2013, Stavrou 2013, among others, on vocatives, and Munaro & Poletto 2009, Coniglio & Zegrean 2010, Poletto & Zanuttini 2010, Bayer & Obenauer 2011, Haegeman 2014, Bayer et al. 2015, Del Gobbo et al. 2015, among others, on discourse markers). Corr (2016) furthermore argues that the highest CP head ForceP (Rizzi 1997) is split into three projections. The higher two, EvalP (EvaluativeP) and EvidP (EvidentialP), are adopted from Cinque (1990) and find further support in Speas & Tenny (2003). DeclP (DeclarativeP) is adopted from Ledgeway (2012) and is associated with clause typing.

With regard to the construction under investigation here, Corr (2016) draws a distinction between reportative constructions where a potential matrix clause can be reconstructed from the context (Portuguese) and those where it cannot (Spanish, Catalan). Those with a reconstructible matrix clause are not a central concern in Corr (2016), but she proposes an analysis and treats them as cases of elision. One of Corr's examples, along with her corresponding analysis, is given in (15a) and (15b).

(15) Portuguese

a. (Corr 2016: 149: ex 6)
Rotsen, sabes o que me disseram?! Que a época
Rotsen know.2SG.PRS what CL.1SG tell.3PL.PRF.PST QUE the season
iria começar a 3 de Dezembro.
go.3SG.COND begin on 3 of December
'Rotsen, do you know what they told me?! That the season was going to begin on 3 December.'

b. (Corr 2016: 149: ex 7)
[CP1 ~~Disseram-me~~ [CP2 que a época iria começar a 3 de Dezembro]]

Corr (2016) only treats instances that lack a performative verb or performative clause in the contexts as true cases of reportative *que*, which she therefore considers to only be found in Spanish and Catalan. In these cases, *que* is analyzed as a dedicated evidential complementizer that is merged in the head of her Evidential Phrase. Her analysis of (16a) is given in (16b).

(16) Catalan

a. (Corr 2016: 161: ex 26)
Que quina pallisa que els van clavar.
QUOT what battering that they AUX.3PL.PRF.PST get
'[I said] what a battering they got.'

b. (Corr 2016: 188: ex 82)
[$_{\text{SAHigh}}$ [$_{\text{SALow}}$ [$_{\text{Eval}}$ [$_{\text{Evid}}$ QUE$_{\text{quot}}$ [$_{\text{Decl}}$ [$_{\text{Topic}}$ [$_{\text{Pol-int}}$ [$_{\text{Excl'}}$ quina pallissa [$_{\text{Excl}}$ que [$_{\text{Wh-int}}$ [$_{\text{Focus}}$ [$_{\text{Fin}}$ [$_{\text{IP}}$ els van clavar]]]]]]]]]]

The difference between Spanish and Catalan on the one hand and Portuguese on the other is explained in Corr (2016) as a case of feature scattering versus feature bundling. In Portuguese, there are three features bundled in only one Force head (as in Figure 2.2), while in Spanish and Catalan the features are scattered across three heads (EvaluativeP, EvidentialP and DeclarativeP) (as in Figure 2.1). Corr (2016) assumes that in the true cases of reportative *que* found in Spanish and Catalan the complementizer spells out only the evidential feature. Since the evaluative, evidential and declarative features are bundled in Portuguese, it is impossible to spell out only one of them. In Corr's account, this is the reason for the differences observed in the reportative constructions in Portuguese compared to Spanish and Catalan. Although it is not stated explicitly in Corr's (2016) book, this analysis might also be able to explain why non-declarative clause types are excluded in Portuguese *que*-initial reportatives (cf. §2.3.1) because the three features bundled on one head include the declarative feature.

With the central points of Corr's analysis in place, I now turn to potential issues in her approach. As stated above, at the core of the analysis is the idea that the Portuguese type of reportative construction differs from the Spanish/Catalan version in that the former requires a reconstructible matrix clause and the latter does not. One possible problem with this distinction is the fact that in Spanish and Catalan the construction can also appear in a context with an explicitly mentioned verb of saying. Crucially, these *que*-initial reportatives with a contextually

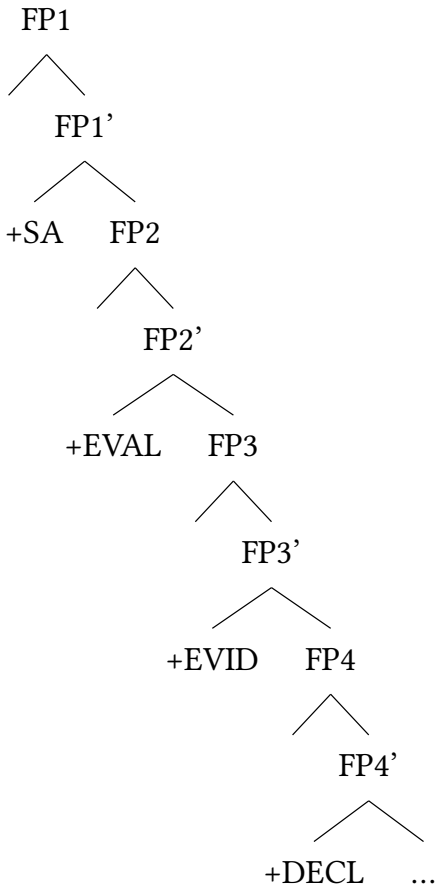


Figure 2.1: Spanish/Catalan (Corr 2016: 187: ex 81)

recoverable *verbum dicendi* have the same syntactic properties and surface in the same position as those without a recoverable *verbum dicendi* (cf. my take on this in §2.3 and §2.4). In the example in (17), a *verbum dicendi* can be recovered from the context. In the *que*-initial sentence the complementizer precedes a *wh*-phrase just as in (16a), an example with a non-reconstructible matrix clause that should be distinct based on Corr's analysis.

(17) Catalan
A: Què fa la Tecla? B: Què m' has preguntat?
what do.3SG.PRS the Tecla what CL.1SG AUX.2SG.PRF.PRS ask.PTCP
A: Que què fa la Tecla?
QUE what does the Tecla
'A: What does Tecla do? B: What did you say? A: [reportative:] What does Tecla do?' (ebook-cat)

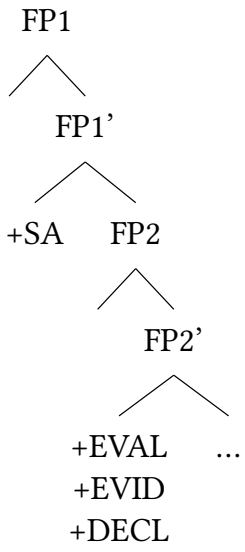


Figure 2.2: Portuguese (Corr 2016: 187: ex 80)

Another issue relates to the interpretation of *que*-initial sentences without a reconstructible matrix clause. According to Corr (2016) "Ibero-Romance quotative QUE constructions are reported speech clauses introduced by the item *que* which, crucially, do not rely on a retrievable *verbum dicendi* to be felicitous." (Corr 2016: 145). While the author assumes that *que*-initial reportatives receive a reportative interpretation, how this interpretation arises in her approach is not immediately evident. It is also not obvious to me whether the author considers *que* to be a reportative evidential marker or a complementizer. The first option could mean that *que* is perceived to be a lexical item although homophonous but still distinct from the default complementizer. In the second option, the reportative interpretation could be assumed to be encoded syntactically. With regard to the interpretation, the author states that the complementizer is merged in an evidential phrase and carries an evidential feature, indicating that the speaker has some sort of evidence for his/her statement. Corr (2016: 159–169) adopts the concept of a presentative force (cf. Déchaine et al. 2017) as the most basic type of illocutionary force that places a proposition in the common ground without committing to its truth. She states that the sentences headed by reportative *que* are presentative rather than asserted (cf. §2.5). However, she does not make explicit which mechanisms ensure that these presentatives with an evidential feature are interpreted as reported sentences.

2.2 Outline of the present analysis

In this section, I present an outline of my own account. The syntactic analysis that I adopt for *que*-initial reportatives builds on the observation that in these apparently unembedded sentences, *que* surfaces in the same location as it does in their embedded equivalents. As a consequence, I propose that *que*-initial sentences and their counterparts should be analyzed in the same way. In this view, the fact that *que* looks and behaves like a complementizer is not a mere coincidence: It is in fact a complementizer. Similar proposals have been made in the literature and have been reviewed in §2.1. The crucial difference is that, contrary to the other accounts, the analysis put forward here does not resort to a hidden performative structure nor to an elided matrix predicate.

The syntactic apparatus is very simple. This simplicity, however, comes at a price: It requires a new conception of what subordination means. The present analysis places the burden more on pragmatic mechanisms than previous analyses have. In a nutshell, the claim I put forward is that the syntactic structure does not encode anything other than that the sentence is subordinate. There are, however, pragmatic requirements for uttering a subordinate sentence. Its use is only felicitous if there is a salient linguistic expression that the sentence can be subordinate to. These pragmatic conditions are described in greater details in §2.5.

The basic syntactic idea is illustrated in the structure in Figure 2.3, which corresponds to the analysis I propose for (18). The complementizer is merged directly in the highest head of the split CP SubP where its underspecified feature is valued as *subordinate*.

(18) Catalan (adapted from Demonte & Fernández Soriano 2014: 34)
Que el Barça ha guanyat la Champeons.
QUE the Barcelona AUX.3SG.PRF.PRS win.PTCP the Championsleague
'[reportative:] Barcelona has won the Champions League.'

In what follows I present the reconceptualization of subordination that I have in mind. The central idea is that the subordinate feature assumed to be located in SubP primarily has consequences for the interpretation of a sentence. Thus, in my conception, *subordinate* is not a syntax-internal feature but an interface feature. A sentence can therefore be marked as subordinate without being selected by a matrix clause. This idea relies on a separation between the syntactic and semantic aspects of subordination. Syntactic subordination is defined as selection by a matrix clause. This means that a syntactically subordinate sentence depends on a matrix clause in the sense that it occupies the position of an argument or an adjunct within this matrix clause. A semantically subordinate sentence is simply

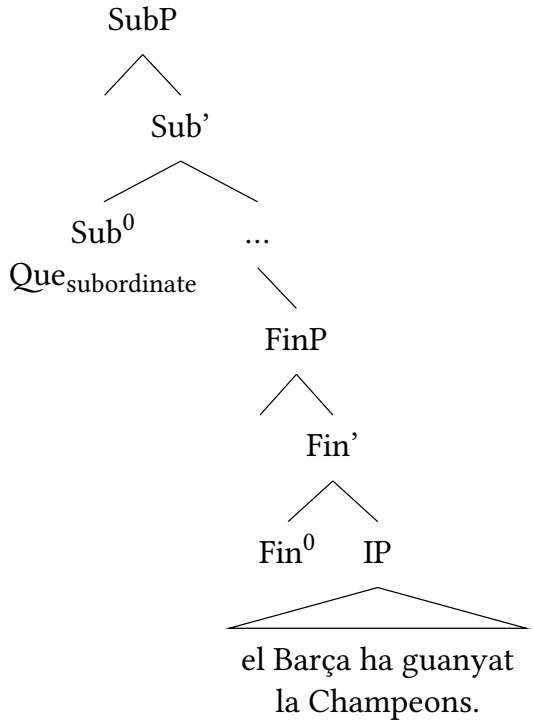


Figure 2.3: Analysis of (18)

interpreted as subordinate. While all syntactically selected subordinate sentences are at the same time semantically subordinate, i.e. interpreted as subordinates, the reverse entailment does not hold: Not all semantically subordinate sentences must be selected by a matrix clause. Applying this to the issue at hand, I propose that the *que*-initial reported sentences under investigation are unselected subordinate sentences. They are interpreted as subordinates but are not syntactically subordinate in the sense mentioned above. Henceforth, to maintain a consistent terminological distinction, I will use the term *embedded* to refer to syntactically selected subordinate sentences and *unembedded* to refer to syntactically unselected subordinate sentences.

The theoretical assumption outlined above can account for those unembedded sentences introduced by a complementizer that occur in contexts where its presence cannot be linked to a reconstructible matrix clause (cf. §2.3). Consequently, I argue that the complementizer is part of the structure for semantic and not for syntactic reasons. An additional empirical motivation is that we also find sentences that are syntactically subordinate but lack an overt complementizer, as in the examples in (21) below. This constitutes evidence for the assumption that a complementizer has primarily semantic functions since if its presence were required for purely syntactic reasons, then it should not be possible to omit it.

Another theoretical prerequisite for my analysis is a distinction between indirect speech and direct quotation. Cappelen & Lepore (1997) illustrate the difference with the examples in (19). If Alice uttered (19a), then (19b) is a direct quotation of (19a). It repeats the exact words she uttered. (19c) is an indirect quote or report. In this example, the indirect quote also repeats Alice's exact words. According to Cappelen & Lepore (1997), the difference boils down to the fact that (19b) is only true if Alice uttered the exact words that are presented as a quotation whereas (19c) is also true if she didn't.

(19) a. Life is difficult to understand.
b. Alice said "Life is difficult to understand".
c. Alice said that life is difficult to understand.
(Cappelen & Lepore 1997: 429: ex 1–3)

In written texts, direct quotations often receive a special orthographic marking like the quotation marks in (19b) that indicate the start and end of a quote, while indirect speech is not marked orthographically. There are reasons to believe that indirect speech, unlike direct quotation, is syntactically embedded. Typically, only indirect speech is introduced by a complementizer, cf. (19c). However, a complementizer is not always present in indirect speech constructions in English and German, because bridge verbs (cf. Erteschik 1973) such as *say* permit complementizer deletion in declarative complement clauses.[5] This leads to a superficial oral (yet normally not written) ambiguity between a direct quotation and an indirect speech reading, exemplified in (20).

(20) Alice said (that) life is hard.

Similar examples are illustrated for English in (21a) and German in (21b). There is no complementizer, but the examples can nevertheless receive an indirect speech interpretation. They are then treated as embedded sentences affected by complementizer deletion.

An important contrast between indirect speech and direct quotation is that they have different deictic centers (see also the discussion in the introduction to this chapter). In direct quotations it is the original speaker, *John* in (21a) and (21b), who said that someone has to leave, while in indirect quotations it is the external speaker, *Mary*, who reports that John said that someone has to leave.

[5] On complementizer deletion see for instance Erteschik (1973), Kayne (1981, 1984), Stowell (1981), Pesetsky (1995), Bošković & Lasnik (2003), Bianchi & Frascarelli (2017).

The first part of the examples in (21) can in principle be interpreted as either a direct quotation or an indirect speech report. However, in its most natural reading indicated by the indexes (see also Gutzmann & Stei 2011), it is an indirect speech report, in which the indexical pronoun *I* refers to the external speaker *Mary* and not to *John*. This reading is supported by the continuation, where once again *I* refers to *Mary* and she is the one who has to leave John's house. The intended interpretation is that Mary is unwelcome at John's house and therefore he tells her to leave.[6]

(21) a. [Mary$_m$:] John$_j$ said I$_m$ have to leave. Apparently I'$_m$ m not welcome at his$_j$ house.

b. German
[Maria$_m$:] Hans$_h$ hat gesagt, ich$_m$ muss gehen.
Maria Hans AUX.3SG.PRF.PRS say.PTCP I must.1SG.PRS go
Scheinbar bin ich$_m$ in seinem$_h$ Haus nicht willkommen.
apparently be.1SG.PRS I in his house not welcome

The parts of the first sentences in (21a), (21b) are not merely juxtaposed but must truly be syntactically subordinate in order to receive the indirect speech interpretation. Consequently, the reading is lost when the reported sentence is not embedded. In the examples in (22) the internal argument position of the verb *say* is filled by the pronoun *it*, meaning that the reported sentence cannot be syntactically embedded. A natural interpretation of this sentence is that of a direct quote with the indicated referents, rendering the continuation infelicitous.

(22) a. [Mary$_m$:] John$_j$ said it again: I$_j$ have to leave. #Apparently I$_m$ am not welcome at his$_j$ house.

b. German
[Maria$_m$:] Hans$_h$ hat es wieder gesagt: Ich$_h$
Maria Hans AUX.3SG.PRF.PRS it again say.PTCP I
muss gehen. #Scheinbar bin ich$_m$ in seinem$_h$ Haus
must.1SG.PRS go apparently be.1SG.PRS I in his house
nicht willkommen.
not welcome

[6]There is an alternative interpretation in which having Mary as an unwelcome guest leads to John's decision to leave his own house. In this scenario, a direct quote reading is possible and the continuations in (22a), (22b) and (24a) are felicitous.

For German, an indirect speech reading of (22b) is acceptable. Potentially, however, we are dealing with a different structure involving extraposition and complementizer deletion. Evidence for this hypothesis comes from the data below that show that the *dass*-initial embedded sentence is extraposed from the DP *die Drohung*. According to Bianchi & Frascarelli (2017) an equivalent structure is not possible in English.

(23) German
[Maria$_m$:] Hans$_h$ hat die Drohung wiederholt dass ich$_m$/$^*_h$
Maria Hans AUX.3SG.PRF.PRS the threat repeat.PTCP that I
gehen muss.
leave must.1SG.PRS
'[Maria:] Hans repeated his threat by saying that I have to leave.'

The Ibero-Romance languages under investigation exhibit a structure that is superficially equivalent to (21a) and (21b) and that also lacks a complementizer. This is exemplified for Spanish in (24a). Crucially, however, the interpretation as an indirect speech report with the external speaker as the deictic center, as in the German and English examples in (21), is not possible. It must be interpreted as a direct quotation: It is Juan who says that he himself has to leave. Therefore, in the interpretation indicated by the indexes, the continuation is infelicitous. In order to achieve an indirect speech interpretation, an overt complementizer is necessary (as in 24b). This shows that complementizer deletion is not an option in these contexts in Ibero-Romance.

(24) Spanish
a. [María$_m$:] Juan$_j$ ha dicho tengo$_{j/*m}$ que
María Juan AUX.3SG.PRF.PRS say.PTCP have.1SG.PRS that
irme$_{j/*m}$. #Aparentemente no estoy$_m$ bienvenida en su$_j$
go.CL.REFL apparently not be.1SG.PRS welcome.F.SG in his
casa.
house
'[María$_m$:] Juan$_j$ said I$_{j/*m}$ have to leave. #Apparently I$_m$ am not welcome at his$_j$ house.'
b. [María$_m$:] Juan$_j$ ha dicho que tengo$_{*j/m}$ que
María Juan AUX.3SG.PRF.PRS that say.PTCP have.1SG.PRS that
irme$_{*j/m}$. Aparentemente no estoy$_m$ bienvenida en su$_j$ casa.
go.CL.REFL apparently not be.1SG.PRS welcome.F.SG in his house
'[María$_m$:] Juan$_j$ said that I$_{*j/m}$ have to leave. Apparently I$_m$ am not welcome at his$_j$ house.'

What these facts demonstrate is that in the relevant contexts, the Ibero-Romance languages require an overt complementizer to interpret a sentence as subordinate where German and English do not. Conversely, this also means that the presence of a complementizer in German and English is not sufficient to identify a sentence as subordinate, since in contexts such as that in (21) the sentence is syntactically subordinate, hence selected by a matrix clause, but there is no overt marker. This might be one reason why in the right context, a complementizer heading a matrix sentence is enough to indicate that the sentence is indirect reported speech in Ibero-Romance languages, while in German and English more explicit strategies are called for, as in (25).

(25) a. Spanish
Juan: Tengo que irme.
María: ¿Eh?
Juan: Que tengo que irme.

b. John: I have to leave.
Mary: Huh?
John: *That I have to leave.

In the Spanish example in (25a), the sentence headed by *que* is understood as a report without the need for additional lexical material. In the equivalent version in English in (25b), *that I have to leave* on its own is not sufficient. A *that*-initial reported sentence is only acceptable in English when a verb of saying is given in the context:

(26) John: I have to leave.
Mary: What did you say?
John: That I have to leave.

Summing up, the main idea is that the presence of a complementizer has the same impact in embedded, i.e. selected, and unembedded, i.e. unselected, reported sentences. In both cases it ensures that the sentence following *que* is interpreted as a subordinate. Consequently, it is only logical that they should also surface in the same syntactic position. I analyze the complementizer as being merged in the highest projection in the split CP and valued with a subordinate feature. In the following sections I offer empirical support for this idea and show how this simple syntactic analysis can also account for more complex cases than that illustrated in (18).

I end this section with a brief comparison of the analyses of reportative *que* proposed by Etxepare (2013), Demonte & Fernández Soriano (2014) and Corr (2016)

(see §2.1) with my own analysis developed here. A summary of the main points is given in Table 2.1.

Table 2.1: Subordinate *que* in different analyses

	nature of *que*	location	reportative interpretation
Etxepare (2013)	COMP	LinkerP	silent noun of saying
D&F (2014)[a]	COMP	ForceP	silent verb of saying
Corr (2016)	evidential COMP	EvidP	feature in EvidP
present analysis	COMP	ForceP	pragmatic reconstruction

[a]Demonte & Fernández Soriano (2014)

With regard to the nature of *que*, my analysis is consistent with those of Etxepare (2013) and Demonte & Fernández Soriano (2014) in treating it as a complementizer. Corr (2016) unfortunately does not explicitly state what kind of element *que* constitutes and what the label *evidential complementizer* entails.

My analysis is also in agreement with Demonte & Fernández Soriano (2014) in assuming that the syntactic position in which *que* surfaces is ForceP. Arguably, Corr's analysis is also in agreement, since EvidP constitutes a subhead of her split ForceP. However, in Corr (2016), the function that my analysis proposes for *que* in reportatives, namely to mark a sentence as subordinate, is not associated with EvidP but with another subphrase: DeclP. The motivation in Corr (2016) for the splitting of ForceP into subheads is conceptual rather than empirical, and is driven by the goal of developing a system in which there is a one-to-one correspondence between a syntactic projection and an abstract pragmatic feature. This, however, makes it difficult to empirically test which of the two adjacent projections *que* actually occupies in reportatives, since Corr (2016) does not discuss any cases in which the two projections are filled at the same time. Etxepare (2013) proposes a small clause analysis: Although *que* is treated as a complementizer, its function differs from that proposed in Demonte & Fernández Soriano (2014) and from the function that I attribute to *que* in my analysis. This is made clear by the syntactic positions *que* occupies in Etxepare (2013). The complementizer has the function of establishing a link between a silent predicate and the *que*-initial sentence. It occupies a Linker Phrase. In the structure proposed by Etxepare (2013), LinkerP and ForceP are adjacent, so, as with Corr (2016), it is difficult to test which of the two projections the complementizer truly occupies.

Turning finally to the question of how the reportative interpretation arises, there are three different options presented. The first option is to assume silent

or elided material like Etxepare (2013) and Demonte & Fernández Soriano (2014). Etxepare (2013) adopts an approach that relies on elision, but how this elision is licensed is not immediately evident in his account. Syntactic elision relies on the possibility of reconstruction; however, not all contexts allow the reconstruction of the relevant material. The second option proposed by Corr (2016) is inspired by neo-performative hypotheses. The author assumes that there is a detailed structure above the CP that contains a dedicated evidential projection which, she proposes, hosts *que* in reportative constructions. In the current version of the analysis offered in Corr (2016), however, the relation between the evidential marker and the reportative interpretation does not seem completely transparent.

The third option, which I adopt in this book is presented in detail in §2.5. It is based on a reconceptualization of subordination which allows a reduced syntactic analysis to be proposed without the need to allude to silent material and maintaining the view that *que* is in fact a complementizer. The basic idea is that the only information encoded syntactically is that the sentence is subordinate. The incomplete information provided by the syntactic structure leads the hearer to look for possible matrix material in the context. The reportative interpretation therefore results from pragmatic rather than syntactic reconstruction. Unlike in previous accounts, the burden of deriving the interpretation is pushed toward pragmatics rather than syntax.

The analysis I lay out in this book has the advantage of wider empirical coverage. It does not need to make a distinction between reportatives with a recoverable versus a non-recoverable *verbum dicendi* that Corr (2016) proposes. Instead, my approach consists of a very simple syntactic analysis that builds on the parallels between the unembedded sentences and their embedded counterparts. They differ in that embedded sentences are selected by a matrix clause while unembedded sentences are unselected. As a consequence, the analysis does not need to postulate any additional structure or features in order to account for unembedded reportatives. This fact also gives my analysis a theoretical advantage: It works without assuming hidden syntactic layers. Analyses of similar phenomena often rely on a neo-performative approach, for instance the widely adopted structure developed in Speas & Tenny (2003) (adapted in Corr 2016). As reported in §1.4, these types of structures, however, are disputed in the literature (see for instance Gärtner & Steinbach 2006, Alcázar & Saltarelli 2014). The syntactic simplicity of my analysis is only possible because I grant a more dominant role to pragmatic mechanisms, as described in §2.5.

2.3 The syntax of *que*-initial reportatives

The central assumption of my analysis is that the complementizer in *que*-initial reportatives is no different from a "normal" complementizer that appears in syntactically subordinate reported speech. In this section, I discuss the syntactic evidence in favor of this assumption. §2.3.1 maps out the position of the complementizer relative to left-peripheral, i.e. CP-internal, material. §2.3.2 deals with its position relative to CP-external material. The empirical evidence laid out in these sections stems from Spanish and Catalan. §2.4 focuses on Portuguese and identifies the common and diverging properties.

2.3.1 Location within the left periphery

The analysis in §2.2 locates the complementizer in the construction under investigation at the left edge of the periphery.

(27) [SubP [TopP [IntP [TopP [ForceP [TopP [ModP [TopP [MoodP [TopP [FinP]]]]]]]]]]

(27) shows the cartographic structure of the split CP adopted in this book. There are two adaptations to the structure proposed by Rizzi (1997, 2004a, 2013) that are inspired by Haegeman (2004, 2006) and Lohnstein (2016) (but also Roussou 2010). The first change, most relevant for the present analysis, is the replacement of ForceP by SubP. In this structure, SubP is the dedicated functional projection that hosts subordinating conjunctions. The second change is MoodP in the lower section of the left periphery, adopted from Lohnstein (2016), and responsible for clause typing. I give more details on the cartographic structure and motivate these adaptations in §1.2.

In the analysis I propose for *que*-initial reportatives, *que* occupies the highest head SubP. It is therefore predicted that the complementizer should precede elements that occupy any of the other left-peripheral positions. This expectation is confirmed by the data in (28–33). In (28a) the left dislocated topic *la tinta*, which is resumed by a clitic pronoun, follows the complementizer. In Ibero-Romance, clitic left dislocated topics can occupy different positions within the left periphery. This is why Rizzi (1997) proposes that a topic position is sandwiched between each of the other left-peripheral projections.[7] Crucially, all the positions that clitic left dislocated topics can occupy are below SubP. The reverse word order illustrated in (28b), where a clitic left dislocated topic precedes the complementizer, is not grammatical.

[7] See Frascarelli & Hinterhölzl (2007) who propose that different topic projections result in different topic interpretations.

(28) Spanish

a. [$_{SubP}$ Que] [$_{TopP}$ la tinta$_i$] la$_i$ hemos de hacer
QUE the ink CL.AKK have.1PL.PRS to make
por semanas.
weekly
'(Somebody said) that we have to make the ink weekly.' (CdE)

b. * La tinta$_i$ que la$_i$ hemos de hacer por semanas.

As a side note, (28b) would be grammatical if *la tinta* were followed by an intonational break. This prosodic pattern is typical for hanging topics, which are usually analyzed in a CP-external position (cf. Bianchi & Frascarelli 2010). I briefly return to these CP-external hanging topics in §2.3.2.

In (29a), *que* precedes a polar question introduced by the interrogative complementizer *si*. The relative position and behavior of *si* are the reason behind the introduction of IntP as an additional left-peripheral head (cf. Rizzi 2001), assumed to be lower than SubP. The empirical facts support the proposed analysis. Once again, the reverse word order in which *que* follows *si* is ungrammatical, cf. (29b).

(29) Spanish

a. [$_{SubP}$ Que] [$_{IntP}$ si] eres feliz.
QUE if be.2SG.PRS happy
'[reportative:] Are you happy.'

b. * Si que eres feliz.

In (30a), the complementizer precedes a *wh*-pronoun. In (31a), it precedes the *wh*-expression of a *wh*-exclamative.[8] In accordance with their interpretation and the fact that they cannot co-occur with left dislocated foci, these elements have been analyzed as located in FocP, which is again assumed to be located below SubP. It is once more ungrammatical to reverse the order of the *wh*-pronoun and the complementizer (see 30b). The example in (31b) with the *wh*-expression followed by *que* in a *wh*-exclamative is ungrammatical with the relevant reportative interpretation. The sequence is grammatical, however, if it is interpreted as a *wh*-exclamative with a lower merged attributive complementizer (cf. §3 for my analysis of these constructions).

[8] On embedded exclamatives see Zanuttini & Portner (2003), Sæbø (2010) and Gutiérrez-Rexach & Andueza (2016).

(30) Catalan

a. A: Què fa la Tecla? B: Què m' has
what do.3SG.PRS the Tecla what CL.1SG AUX.2SG.PRF.PRS
preguntat? A: [$_{SubP}$ Que] [$_{FocP}$ què] fa la Tecla?
ask.PTCP QUE what do.3SG.PRS the Tecla
'A: What does Tecla do? B: What did you say? A: [reportative:] What does Tecla do?' (ebook-cat)

b. * Què que fa la Tecla?

(31) Spanish

a. El abogado, la secretaria y yo empezábamos a hablar. Muy
the lawyer the secretary and I start.1PL.IPFV.PST to talk very
amigable el abogado, solo sonrisas era. [$_{SubP}$ Que] [$_{FocP}$
amicable the lawyer only smiles be.3SG.IPFV.PST QUE
qué bonito] tenía el terreno.
how pretty have.1SG.IPFV.PST the terrain
'The lawyer, the secretary and I started to talk. The lawyer: very friendly and all smiles. [reportative:] How pretty I kept the terrain.' (CdE)

b. * Qué bonito que$_{subordinate}$ tenía el terreno.

Finally, the examples in (32a) and (33a) show that as predicted the epistemic adverb *seguramente* and the verum marker *sí* are also preceded by subordinate *que*. Their respective positions in the left periphery are explained in Kocher (2017a) and the constructions involving them are discussed extensively in §3. The second *que*-initial reportative in (33a) shows that multiple left-peripheral projections can also be occupied in *que*-initial reportatives. Here, *que* and *sí* are interrupted by the DP *Martha* which is analyzed as a topicalized subject moved to one of the TopP projections between SubP and ModP (see Kocher 2017a for similar data and analyses). Again, the reverse order, with the subordinate complementizer following the adverb or verum marker is not grammatical in the interpretation relevant here, cf. (32b), (33b). Just as in the example above, however, the word order is grammatical in a different interpretation where the complementizer carries an attributive feature and is merged low, which is the topic of §3.

(32) Spanish

a. ¿Hay mucha corrupción en el gobierno nacional
there.be.3SG.PRS a lot corruption in the government national

y popular? Los kirchneristas dirán que no. [$_{SubP}$ Que]
and popular the kirchnerista.PL say.3PL.COND that no QUE
[$_{ModP}$ seguramente] es una sola.
surely be.3SG.PRS one only
'Is there a lot of corruption in the national and popular government? The supporters of Kirchner would answer that there isn't. [reportative:] There is only one type of corruption.' (CdE)

b. * Seguramente que$_{subordinate}$ es una sola.

(33) Spanish

a. Por ustedes dijo también para luego agregar: [$_{SubP}$
for you.PL tell.3SG.PRF.PST also to afterwards add
Que] no se enreden: [$_{SubP}$ Que] [$_{TopP}$ Martha] [$_{MoodP}$
QUE not CL.REFL tangle.2PL.IMP QUE Martha
sí] va. [$_{SubP}$ Que] [$_{MoodP}$ sí] va porque
sí go.3SG.PRS QUE sí go.3SG.PRS because
ha demostrado que sí puede, con el
AUX.3SG.PRF.PRS demonstrate.PTCP that sí can.3SG.PRS with the
apoyo de todos.
help of all
'For you she also spoke and later added: [reportative:] Don't get all tangled up. [reportative:] Martha DOES go. [reportative:] She DOES come because she demonstrated that with everybody's help she CAN.' (CdE)

b. * Sí que$_{subordinate}$ va.

Multiple *que* constructions, in which a *que*-initial reported sentence contains another instance of *que*, are also possible, cf. (34). The second complementizer is valued with an attributive feature and observes strict adjacency to a *wh*-expression in the *wh*-exclamative in (34a), an evidential modifier in AdvC in (34b) and the verum marker *sí* in AffC in (34c). The structure and interpretation of sentences with a low-merged complementizer are dealt with in depth in §3. Regardless of the specific details, what the data in (34) demonstrate is that there is more than one position within the left periphery that can host a complementizer. A central argument of this work is that the position in which the complementizer is merged has an effect on the interpretation of the sentence. The claim is that while the higher position marks the sentence as subordinate, the lower position attributes to the hearer a commitment to the proposition in the scope of

the complementizer. In the present analysis, this is modeled by assuming that the different projections provide different interface features. Briefly, the lower complementizer is merged in FinP but does not remain in this position. Instead, it moves from head to head through the left periphery until it encounters an element in the specifier of a projection that blocks its movement.

(34) a. Catalan (Corr 2016: 161: ex 26)
[$_{SubP}$ Que] [$_{FocP}$ quina pallissa que$_i$] [$_{FinP}$ t$_i$] els van
QUE what battering QUE they AUX.3PL.PRF.PST
clavar.
get
'(I said) what a battering they got.'

b. Spanish
[$_{SubP}$ Que] [$_{ModP}$ claro que$_i$] [$_{FinP}$ t$_i$] le viene bien, que
QUE clear QUE CL.REFL come.3SG.PRS good QUE
qué alegría, que dónde.
what joy QUE where
'[reportative:] Of course it was no inconvenience. (She said) what a joy! (She said) where?' (CdE)

c. Catalan
[$_{SubP}$ Que] [$_{MoodP}$ sí que$_i$] [$_{FinP}$ t$_i$] heu vingut!
QUE SÍ QUE AUX.2PL.PRF.PRS come.PTCP
'[reportative:] You HAVE come.' (caWac)

My analysis proposes that *que*-initial reportatives are essentially subordinate sentences that are unselected and therefore lack a matrix clause. An expectation that follows from this is that the subordinate complementizer should surface in the same position in embedded and unembedded reportatives. The following paragraphs demonstrate that this expectation is borne out. There are two versions of embedded questions in Spanish and Catalan: One in which the complementizer is present and one it which it is absent (see the discussion of the examples (69) in §2.4 and see also Etxepare 2010, González i Planas 2014 and Corr 2016). Crucially, if the complementizer is present, as predicted, it surfaces in the same position as in unembedded reportatives (cf. 35–38). My proposal is that the syntactic structure of the embedded sentence CP in (35) is equivalent to the structure of the unembedded sentence CP in (30a). This is supported by the fact that *que* precedes the interrogative complementizer *si* in both cases.

(35) Spanish
Me acerqué a una oficina de turismo, pregunté
CL.REFL approach.1SG.PRF.PST to a tourist office ask.1SG.PRF.PST
[$_{SubP}$ que] [$_{IntP}$ si] estaba abierto el camino hacia
that whether be.3SG.IPFV.PST open the road to
Florencia.
Florence
'I went to a tourist office and asked (that) whether the road to Florence was open.' (CdE)

Example (36) parallels (30a). A *wh*-pronoun, analyzed in FocP, is preceded by the complementizer *que*.

(36) Spanish
Me encontré con un amigo y el amigo me
CL.REFL meet.1SG.PRF.PST with a friend and the friend CL.1SG
dijo [$_{SubP}$ que] [$_{FocP}$ qué] estaba haciendo.
say.3SG.PRF.PST that what be.3G.PROG.PRS doing
'I ran into a friend and the friend asked (that) what I was doing.' (CdE)

The examples in (37) and (38) show that syntactically embedded reportatives also permit the occurrence of multiple instances of *que*, just as their unembedded counterparts do in (34). The example in (37) demonstrates that the AdvC construction, like in the unembedded sentence in (34b), can be part of an embedded reportative and occupies a position lower than SubP. Example (38) shows that AffC also surfaces below the subordinating complementizer, similar to (34c). The example in (38) furthermore shows that a clitic left dislocated topic and an epistemic modifier can intervene between the high instance of *que* and the sequence *sí que*. The position of the topic and the modifier is also predicted by my analysis. The high *que* is analyzed in SubP while the lower one is merged in the lowest projection FinP before moving to the second-lowest MoodP. There are therefore a number of intervening projections available that can host further phrases (cf. the structure in 27).

(37) Catalan (Kocher 2017b: 35: ex 57a from ebook-cat)
A nivell de relació de parella us haig de dir [$_{SubP}$ que]
at level of relationship of couple you have.1SG.PRS to say that
[$_{ModP}$ evidentment que$_i$] [$_{FinP}$ t$_i$] les coses canvien.
evidently QUE the thing.PL change.3PL.PRS
'Concerning couples' relationships I have to tell you that evidently things change.'

(38) Catalan (Kocher 2017b: 35: ex 57c from ebook-cat)
Al museu, mirant un dibuix de Manolo Hugué ha
at the museum looking a drawing of Manolo Hugué AUX.3SG.PRF.PRS
conegut una dona d'ulls nets, profunds i amb una punta de
meet.PTCP a woman of eyes clean deep and with a hint of
malícia, i ha pensat [$_{SubP}$ que] [$_{ModP}$ potser] [$_{TopP}$
malice and AUX.3SG.PRF.PRS think.PTCP that maybe
d'aquesta dona$_j$] [$_{MoodP}$ sí que$_i$] [$_{FinP}$ t$_i$] se'n$_j$ podria
of this woman sí QUE CL.REFL.CL.PART can.3SG.COND
enamorar.
fall in love
'At the museum, while looking at a drawing of Manolo Hugué he met a woman with honest, deep eyes and with a hint of malice and thought that maybe with this woman he COULD fall in love.'

A characteristic that has been highlighted by Demonte & Fernández Soriano (2014) is the fact that what follows *que* in reportatives does not have to be a full sentence. This is illustrated in (39) repeated from (34b). At the end of the fragment, *que* is followed only by the *wh*-phrase *qué alegría* and the *wh*-pronoun *dónde* 'where' while the rest of the sentence is omitted.

(39) Spanish
[$_{SubP}$ Que] [$_{ModP}$ claro que$_i$] [$_{FinP}$ t$_i$] le viene bien, [$_{SubP}$
QUE clear QUE CL.REFL come.3SG.PRS good
que] [$_{FocP}$ qué alegría], [$_{SubP}$ que] [$_{FocP}$ dónde.]
QUE what joy QUE where
'[reportative:] Of course it was no inconvenience. (She said) what a joy! (She said) where?' (CdE)

The data below show that permitting fragmentation is not a special property of unembedded sentences, but can also be observed in syntactically subordinated indirect reportatives. In (40a) the reportative only consists of a *wh*-pronoun, and in (40b) it consists of the adverb *ahora* 'now'. This again supports my proposal that unembedded *que*-initial reportatives are unselected but essentially do not otherwise differ from their embedded counterparts.

(40) Spanish
a. Todos volvieron a preguntar que qué.
all ask.again3P.PRF.PST that what
'Everybody asked "what" again.' (CdE)

b. ¿Cuándo no hay condenación? ¡La Palabra de Dios
when not there.be.3SG.PRS damnation the word of God
dice que AHORA!
say.3SG.PRS that now
'When is the moment of no damnation? The word of God says that the moment is now.' (CdE)

Further support for the idea that the complementizer surfaces in the same position in *que*-initial reportatives as in their syntactically embedded counterparts stems from their behavior with respect to non-declarative sentences. Instances of *wh-* and polar questions were exemplified and discussed in (29a) and (30a). Another example is shown in (41).

(41) Spanish
Pues como te iba diciendo, ella me preguntó
well as CL.2SG go.1SG.IPFV.PST telling she CL.1SG ask.3SG.PRF.PST
por ti. [$_{SubP}$ Que] [$_{FocP}$ qué] tal te va, [$_{SubP}$ que] [$_{IntP}$
about you QUE how be.2SG.PRS you doing QUE
si] tienes trabajo y así.
whether 2SG.PRS work and so
'As I was telling you, she asked about you. [reportative:] How you are doing and do you have a job and so on.' (CdE)

What is important is that in both cases the complementizer is in a higher position than the *wh*-pronoun (in FocP) or the interrogative complementizer *si* (in IntP).

One substantial argument in favor of the assumption that *que* in unembedded sentences is truly a marker of subordination comes from imperatives. The examples in (42) show that in order to report an imperative, its structure and morphological makeup need to be adapted. Thus (42b), in which the imperative is directly headed by the subordination complementizer, is ungrammatical. In the grammatical version in (42a), the verbal morphology is no longer imperative but subjunctive. The verb furthermore occupies a higher position in the imperative in (42a) than in the reported version in (42c). This is shown by the position of the clitic pronoun, which is enclitic in the true imperative and proclitic in reported imperatives.

(42) Spanish
a. ¡Vete!
go.2SG.IMP.CL.REFL
'Go away!'

b. * ¡Que vete!
QUE go.2SG.IMP.CL.REFL
Intended '[reportative:] Go away!'

c. ¡Que te vayas!
QUE CL.REFL go.2SG.SBJV.PRS
'[reportative:] You should go!'

A comparison with syntactically embedded imperatives once again shows that they exhibit the same behavior (on embedded imperatives see Portner 2007, Kaufmann 2011, Kaufmann & Poschmann 2013). The example in (43b), in which the verb is inflected for imperative mood and is directly embedded under a verb expressing an order, is ungrammatical. The grammatical structure contains a verb inflected for subjunctive mood. The proclisis shows that the subjunctive verb surfaces lower in (43b) than the imperative verb in (42a).

(43) Spanish

a. Te ordeno que te vayas.
CL.2SG order.1SG.PRS that CL.REFL go.2SG.SBJV.PRS
'I order that you leave.'

b. * Te ordeno que vete.
CL.2SG order.1SG.PRS that go.3SG.IMP.CL.REFL
Intended: 'I order that you leave.'

The data in (42) and (43) are strong evidence in favor of the subordinate nature of *que*-initial reportatives and against the assumption that *que* is merely a pragmatic marker.

I now turn to a short digression on *que*-initial directives. These share a number of properties with reported imperatives but must nonetheless be kept separate. *Que*-initial directives encode commands directed at a third person who is not currently participating in the conversation. This type of directive is called a jussive. Examples of *que*-initial jussives are given in (44). They contrast with a (non-reported) command directed at a speech participant which is never introduced by the complementizer. The contrast between these *que*-initial directives and the other *que*-initial sentences I investigate in this book is also notable: With *que*-initial jussives, the omission of the complementizer results in ungrammaticality (see also Gras 2016: 117) which is not the case with the reportative and attributive *que* constructions. Note that there is also a *que*-less jussive in Ibero-Romance. However, it differs from jussives like those in (44) in more than the

mere presence or absence of *que* and shares the syntactic properties of second-person directives (see the discussion around example (48)).

(44) a. Spanish
Que se vaya.
QUE CL.REFL go.3SG.SBJV.PRS

b. Catalan
Que es vagi.
QUE CL.REFL go.3SG.SBJV.PRS

c. Que vá embora.
QUE go.3SG.SBJV.PRS away
'He/She should go away.'

From a formal point of view, the comparison illustrated in (45) between directives addressing a second and a third person is particularly revealing. First of all, the second- and third-person directives in these examples differ in their verbal inflection. The verb in the third-person directive in (45a) is inflected for subjunctive while the verb in the second-person directive in (45b) is inflected for imperative mood. Furthermore, the unstressed pronouns are proclitic in (45a) but enclitic in (45b). These structural differences have led authors to conclude that the verb reaches a higher position in imperatives than in third-person directives. Some (for instance Rivero 1994, Rivero & Terzi 1995, Demonte & Fernández Soriano 2009, Alcázar & Saltarelli 2014) propose that the imperative verb targets a left-peripheral position.

(45) Catalan

a. * (Que) es renti les mans.
QUE CL.REFL wash.3SG.SBJV.PRS the hands
'He/she should wash her/his hands!'

b. Renta't les mans.
wash.2SG.IMP.CL.REFL the hands
'Wash your hands!'

The examples in (46) moreover show that the third-person directives (cf. 46a, 47a) behave similarly to declaratives (cf. 46b), reported imperatives (cf. 42c) and embedded imperatives (cf. 43) in licensing preverbal negation and subjects, both of which are ungrammatical in imperatives (46c), (47b). This is cited as further evidence for the idea that an imperative verb reaches a higher position than a subjunctive or indicative verb.

(46) Spanish

a. Que él no lo compre.
QUE he not CL.AKK buy.3SG.SBJV.PRS
'He/She should not buy it.'

b. Tú no lo compras.
you not CL.AKK buy2SG.IND.PRS
'You don't buy it.'

c. (*Tú) (*No) Cómpralo tú.
you not buy.2SG.IMP.CL.AKK you

(47) Portuguese

a. Que ele não o compre.
QUE he not CL.AKK buy.3SG.SBJV.PRS
'He should not buy it.'

b. (*Tu) (*Não) compra-o.
you not buy.2SG.IMP-CL.AKK
'Buy it!'

These structural patterns hold in Spanish, Catalan and Portuguese.[9] As stated above, in addition to the *que*-initial jussives, there is also a grammatical *que*-less counterpart in all three languages. These behave on par with the second-person directives in triggering enclisis and postverbal subjects. They are illustrated in (48).

(48) a. Spanish
Váyase.
go.3SG.SBJV.PRS.CL.REFL

b. Catalan
Vagi-se.
go.3SG.SBJV.PRS-CL.REFL

[9] As an aside, it must be mentioned that in European Portuguese clitic placement in declaratives in general requires enclisis (although there are a number of properties that induce proclisis). This means that with respect to clitic placement, European Portuguese indicative verbs pattern with imperative and not subjunctive verbs. Some authors have argued that this is due to the fact that the finite indicative verb moves higher than the IP in Portuguese (cf. Raposo 2000, Galves et al. 2005). Famously, Brazilian Portuguese patterns differently and shows proclisis in declaratives, like Spanish and Catalan (cf. Galves et al. 2005). Unlike imperatives, declaratives do permit preverbal subjects in both Portuguese varieties. This suggests that while clitic placement is the same, the structure of declaratives and imperatives is nevertheless different.

c. Portuguese
Vá-se embora.
go.3SG.SBJV.PRS-CL.REFL away
'He/She should go away.'

To close, I want to mention a construction that is superficially identical to *que*-initial directives and that is employed to express optatives. These are defined in Grosz (2012) as sentences that express a wish, regret, hope or desire in the absence of an overt lexical item that encodes *wish*, *regret*, *hope* or *desire*. Ibero-Romance optatives require an initial complementizer when they are directed at a (formally marked) third person but also when they are directed at a second person, cf. (49a). An optative expressing a wish dedicated to a non-speech participant coincides with the form used when expressing a wish dedicated to a speech participant who is addressed with the honorific (cf. 49b). Sometimes the second-person optatives are employed as polite forms of an imperative (cf. Demonte & Fernández Soriano 2014, who even call them independent imperatives).

(49) Catalan
a. Que tinguis un dia magnífic.
QUE have.3SG.SBJV.PRS a day magnificent
'May you have a magnificent day!'
b. Que tingui un dia magnífic.
QUE have.3SG.SBJV.PRS a day magnificent
'[honorific:] May you have a magnificent day!'
'May he/she have a magnificent day!'

Summing up this section, I have shown that *que* precedes all other types of phrases that are assumed to be merged in the split CP. The complementizer must therefore be located at the left edge of the functional field. This analysis consequently predicts that *que* must follow linguistic material merged above the CP. In the following section, I discuss data that show that this prediction holds true.

2.3.2 Above the left periphery

This section discusses empirical evidence that confirms the prediction that the complementizer in *que*-initial reportatives follows CP-external material. First, I show that, as expected, *que* follows certain discourse markers, vocatives and hanging topics. I then discuss specific phrases that precede *que* and that have

been analyzed in the literature as expressing the original speaker of the reported sentence. I offer an alternative analysis that treats these phrases as frame setters.

Corr (2016) shows that discourse markers and vocatives are external to the left periphery and follow ordering restrictions (see also the syntactic analysis of vocatives for instance in Moro 2003, Hill 2007b, 2013, 2014, Moreira 2013, Espinal 2013, de Carvalho 2013, Stavrou 2013, and of discourse markers for instance in Munaro & Poletto 2009, Coniglio & Zegrean 2010, Poletto & Zanuttini 2010, Bayer & Obenauer 2011, Haegeman 2014, Bayer et al. 2015, Del Gobbo et al. 2015). This leads her to propose an extended speech act structure above the split CP (as in 50).[10]

(50) [$_{\text{SAHigh}}$ *discourse markers* [$_{\text{SALow}}$ *vocatives* [$_{\text{EvalP}}$ [$_{\text{EvidP}}$ [$_{\text{SpecEvid}}$ *source*] [$_{\text{EvidP'}}$ [$_{\text{Evid}^0}$ *que*$_{\textit{reportative}}$]] [$_{\text{DeclP}}$ …]]]]] (Corr 2016)

Corr (2016) proposes two speech act phrases above the split CP. According to her, the higher phrase hosts outward-oriented and discourse-activating markers like *oye* 'listen', while the lower one hosts inward-oriented and discourse-bonding markers like vocatives. As expected, the complementizer follows both types of expressions in *que*-initial reportatives.

The data in (51) once again confirm that the complementizer in unembedded reportatives surfaces at the left edge of the split CP, above other CP-internal material, but crucially below CP-external material. In particular, the example in (51a) shows that the discourse marker *oye*, analyzed as located in SAHigh in Corr (2016), can only precede but not follow *que* in reportatives. The same goes for the vocative *Irene* in (51b), analyzed as located in SALow in Corr (2016). Finally, the example in (51c) shows that a discourse marker and a vocative can co-occur and that they both precede *que*. This word order is predicted by my analysis. In (51c), in addition to the discourse marker and the vocative, there is a DP immediately above *que*. The DP *la peinadora* is interpreted as the source of the report; according to Corr (2016: 152) it encodes the "original interlocutor". She analyzes the phrase as being merged in the specifier of the projection occupied by the complementizer, hence the ungrammaticality of a complementizer preceding *la peinadora* in (51c).

(51) Spanish

a. (Corr 2016: 182: ex 66)
[$_{\text{SAHigh}}$ Oye], que (*oye) si has visto mis llaves.
DM QUE DM IF AUX.2SG.PRF.PRS see.PTCP my keys
'Hey, (I asked) have you seen my keys?'

[10] For a detailed discussion of Corr's analysis see §2.1.

b. Corr (2016: 182: ex 67)
[$_{SALow}$ Irene], que (*Irene) si has visto mis llaves.
VOC QUE VOC if AUX.2SG.PRF.PRS see.PTCP my keys
'Irene, (I asked) have you seen my keys?'

c. (Corr 2016: 182–183: ex 70)
[$_{SAHigh}$ Oiga], [$_{SALow}$ señora Marquesa], (*que) la peinadora que
DM VOC QUE the hairdresser QUE
no puede esperar.
not can.3SG.PRS wait
'Listen, Lady Marquis, the hairdresser says she cannot wait.'

In the example in (51c), the referent of *la peinadora* has a dual role: She is the original speaker but she is also interpreted as the agent of the action described in the proposition and thus coincides with the subject of the sentence. Still, in the version in (51c), *la peinadora* cannot be a dislocated and topicalized subject because if it were, it would target a position lower than *que* (cf. an equivalent version with a clitic left dislocated object in (28a)). The only possible analysis for this example, then, is that there is a null subject agreeing with the finite verb *esperar* 'wait'.

Analyzing *la peinadora* as a hanging topic is also not possible. While hanging topics can precede a reportative *que* (see §2.3.1), the sentence would have different properties. Hanging topics differ from CP-internal topics in that they are always resumed by a clitic, a full pronoun or an epithet. Furthermore, CP-internal topics are usually analyzed as being generated in an IP-internal position and moved to the left periphery,[11] while hanging topics are analyzed as being merged directly in the CP-external position (cf. Villa-García 2015: 165 for a concise comparison).

(52) Spanish
La peinadora$_i$, que esta impaciente$_i$ no puede esperar.
the hairdresser QUE this impatient not can.3SG.PRS wait
'[reportative:] The hairdresser, this impatient person cannot wait.'

The example in (52) illustrates the typical properties of hanging topics. *La peinadora* is followed by an intonational break indicated by the comma and is resumed by the co-referential epithet *esta impaciente*.[12]

[11] But cf. Villa-García (2015) for a special type of topic construction where the topic is analyzed as being merged directly in a CP-internal position.

[12] Some of the native speakers I consulted additionally accepted a version of (52) where *la peinadora* is sandwiched between two *ques*. Although further investigation is necessary, this is potentially a different phenomenon similar to the recomplementation analyzed by Villa-García (2015).

The phrases that precede *que* that have so far been analyzed as expressing an original speaker are sometimes also followed by an intonational break – although, according to some of my informants, the break is shorter than the break that follows the hanging topic in (52). These nevertheless do not qualify as hanging topics. Importantly, they cannot be resumed because they are not co-referential with an argument or an adjunct of the verb. This can be seen clearly in the example in (53).

(53) a. Catalan
La mare, que què t' has cregut.
the mother QUE what CL.REFL AUX.2SG.PRF.PRS think.PTCP
'[reportative:] The mother: What were you thinking.' (ebook-cat)

b. Brazilian Portuguese
E eu que não, que não.
and I QUE no NO
'[reportative] and I: no, no.' (CdP)

In previous accounts, the phrases that precede *que* have been used as supporting evidence for a special underlying performative (Demonte & Fernández Soriano 2014) or evidential (Corr 2016) structure in which these phrases are claimed to occupy a dedicated source or speaker position. In the following, however, I show that these preceding phrases do not always encode the original speaker or source of the reported sentence. And more importantly: even when the phrases can be interpreted as expressing an original speaker or source, they do not show any particular behavior that suggests that they have an especially local relation to the complementizer. My proposal is that these *que*-preceding phrases are frame setters. These are elements that set the frame in which the following expressions should be interpreted (Krifka 2008, cf. also Chafe 1976 and Jacobs 2001).

The *que*-preceding phrases can be DPs like in (53), which makes the conclusion that they express an original speaker appear plausible. However, *que* can also be preceded by PPs. This is illustrated by data taken from Etxepare (2013).

(54) Spanish (Etxepare 2013: 98: ex 11)
Oye, en Al-Jazeera, que Obama va a atacar Iran.
Hey in Al-Jazeera QUE Obama go.3SG.PRS to attack Iran
'Hey, there's a saying on Al Jazeera that Obama is going to attack Iran.'

In the example in (54), the phrase that precedes *que* is a locative PP. In Etxepare (2013) *en Al-Jazeera* is analyzed as an adverbial modifier of the elided predicate

of saying. I want to mention here that, although data like these are not elaborated on in Corr (2016), they could be reconciled with her analysis by assuming that phrases preceded by *que* do not express an agent (as she states at various points) but merely an information source. The examples in (55–58), however, pose a problem for her analysis. In (55), I have adapted Etxepare's example and present it in a context in which the locative PP *en Al-Jazeera* is contrasted with *en BBC*.

(55) Spanish
Ultimadamente uno no se puede fiar de las noticias: En
lately one not CL.REFL can.3SG.PRS trust in the news in
Al-Jazeera, que Obama va a atacar Iran. Pero en BBC, que
Al-Jazeera QUE Obama go.3SG.PRS to attack Iran but in BBC QUE
va a retirar sus tropas.
go.3SG.PRS to retreat his troops
'Lately you cannot trust the news: [reportative:] On Al Jazeera, Obama is going to attack Iran. [reportative:] But on the BBC, he is going to retreat his troops.'

In (56), in the same context, the temporal PPs *hoy por la mañana* and *por la tarde* are contrasted. These temporal PPs can hardly be taken to constitute the source of a report. Once again, these data do not contradict Etxepare's analysis but they are not expected under the analysis in Corr (2016).

(56) Spanish
Ultimadamente uno no se puede fiar de las noticias: Hoy
lately one not CL.REFL can.3SG.PRS trust in the news today
por la mañana, que Obama va a atacar Iran. Pero por la
at the morning QUE Obama go.3SG.PRS to attack Iran but at the
tarde, que va a retirar sus tropas.
afternoon QUE go.3SG.PRS to retreat his troops
'Lately you cannot trust the news: [reportative:] In the morning, Obama is going to attack Iran. [reportative:] But in the afternoon, he is going to retreat his troops.'

Moreover, (57) and (58) show that more than one phrase can precede reportative *que*. In (57), there are two adjacent PPs. Crucially, in this example the element that potentially expresses a source, i.e. *en Al-Jazeera*/*en BBC*, is not adjacent to *que*. In (58), there are DPs, *el reportero de Al-Jazeera* and *el reportero de BBC* which

qualify as sources, and they are followed by a temporal PP. Importantly, this sequence is not predicted by Corr's analysis since she proposes that the phrase encoding the source and *que* occupy the specifier and head of the same phrase (see Corr 2016: 183). The data show that a PP can intervene between a potential source or original speaker and *que*.

(57) Spanish
Ultimadamente uno no se puede fiar de las noticias: En
lately one not CL.REFL can.3SG.PRS trust in the news in
Al-Jazeera, hoy por la mañana, que Obama va a atacar Iran.
Al-Jazeera today at the morning QUE Obama go.3SG.PRS to attack Iran
Pero, en BBC, por la tarde, que va a retirar sus tropas.
but in BBC at the afternoon QUE go.3SG.PRS to retreat his troops
'Lately you cannot trust the news: [reportative:] On Al-Jazeera in the morning, Obama is going to attack Iran. [reportative:] But on the BBC in the afternoon, he is going to retreat his troops.'

(58) Spanish
Ultimadamente uno no se puede fiar de las noticias: El
lately one not CL.REFL can.3SG.PRS trust in the news the
reportero de Al-Jazeera, hoy por la mañana, que Obama va a
reporter of Al-Jazeera today at the morning QUE Obama go.3SG.PRS to
atacar Iran. Pero, el reportero de BBC, por la tarde, que va
attack Iran but the reporter of BBC at the afternoon QUE go.3SG.PRS
a retirar sus tropas.
to retreat his troops
'Lately you cannot trust the news: [reportative:] Al-Jazeera's reporter, in the morning: Obama is going to attack Iran. [reportative:] But BBC's reporter, in the afternoon: he is going to retreat his troops.'

In sum, the examples in (54–58) show that information source is not the only possible interpretation for the preceding phrase and that, even if a phrase appears to express an information source, it does not have to be adjacent to *que*, contrary to what we expect based on Corr (2016). As stated above, the data shown here do not contradict the analysis presented in Etxepare (2013), where the preceding phrases are treated as modifiers of an elided noun of saying. Since my aim is to present an analysis that does not need to assume elided material that could be modified, an alternative explanation is required. I characterize the phrases that precede *que* as instances of Krifka's frame setters (going back to similar

concepts presented in Chafe 1976 and Jacobs 2001). In what follows I discuss the evidence in favor of this idea focusing mainly on the interpretation of the phrases that precede *que*. Although a syntactic analysis of frame setters will not be fully fleshed out here, their high syntactic position is an additional indicator that the proposal is on the right track, since Benincà & Poletto (2004) showed that frame setters are always CP-external.

The first thing that Krifka (2008) establishes is that frame setters are not topics. In particular, he underlines the contrast between frame setters and aboutness topics, which, according to Jacobs (2001), have not been separated clearly in the literature. The sentences that are introduced by frame setters are not *about* the frame. Krifka (2008) illustrates this with the example in (59). In B's answer the topic of the sentences is Daimler-Chrysler and not Germany or America.

(59) A: How is business going for Daimler-Chrysler?
B: [In Germany]$_{\text{Frame}}$ the prospects are good, but [in America]$_{\text{Frame}}$ they are losing money. (Krifka 2008: 269: ex 48)

Although there are apparent similarities, Krifka (2008) argues that frame setters also do not constitute contrastive topics. In his theory, contrastive topics are defined as aboutness topics that contain a focus that generates alternative aboutness topics. Consequently, from a theoretical standpoint, the fact that frame setters are simply not what the sentence is about, is sufficient to not consider them contrastive topics in Krifka's conception. Contrastive topics and frame setters, however, also differ in their communicative functions. According to Krifka (2008), when a contrastive topic is uttered, the current common ground contains an expectation that information about a more comprehensive or distinct entity will be addressed in the conversation. Contrastive topics are used to show that the topic of the sentence diverges from this expectation. In (60), speaker A's question establishes *your sister* as a topic. This leads to the expectation of receiving information about her, in particular about whether or not she speaks Portuguese. Speaker B's answer introduces *my brother* as a topic that contrasts with the one introduced by speaker A and thereby shows that the expectations based on A's questions are not met.

(60) A: Does your sister speak Portuguese?
B: [My brother]$_{\text{Contrastive topic}}$ does. (adapted from Krifka 2008: 268: ex 46)

With frame setters, Krifka (2008) states that the current common ground also contains an expectation that more comprehensive or distinct information will be given. The frame setter indicates that the information is less comprehensive and

is just restricted to a particular dimension specified by the frame. Krifka (2008) illustrates this function with the example in (61). Speaker A's question leads to the expectation of an answer containing comprehensive information on John's general well-being. Speaker B's answer, however, provides information restricted to one particular area of well-being. This is possible because the predicate *fine* can be evaluated within different dimensions. John could also be doing *fine* financially, at work or with his love life, to name just a few. However, the frame restricts the information given about his well-being to the dimension of health.

(61) A: How is John?
B: [Healthwise]$_{\text{Frame}}$ he is fine. (Krifka 2008: 269: ex 47)

The *que*-initial phrases behave like frame setters. In the examples in (54) and (53a) repeated in (62a) and (62b), *en Al-Jazeera* and *la mare* are not what the sentence is about, i.e. they are not topics, but they establish a frame in which the proposition should be evaluated. The frame in (62a) gives us the origin of the statement and in (62b) the perspective or speaker that the statement is attributed to.

(62) a. Spanish (Etxepare 2013: 98: ex 11)
Oye, [en Al-Jazeera]$_{\text{Frame}}$, que Obama va a atacar Iran.
Hey in Al-Jazeera QUE Obama go.3SG.PRS to attack Iran
'Hey, there's a saying on Al Jazeera that Obama is going to attack Iran.'

b. Catalan
[La mare]$_{\text{Frame}}$, que què t' has cregut.
the mother QUE what CL.REFL AUX.2SG.PRF.PRS think.PTCP
'[reportative:] The mother: what were you thinking-' (ebook-cat)

Krifka (2008) furthermore observes that explicit frame setters always allude to alternative frames. This can be seen in (59) where the explicit frames *in Germany* and *in America* are contrasted. The examples in (55–57) show that the same is true for *que*-preceding adverbials. In (63) the two frames introduce alternative speakers.

(63) Catalan
[La mare]$_{\text{Frame}}$, que què t' has cregut. [Jo]$_{\text{Frame}}$,
the mother QUE what CL.REFL AUX.2SG.PRF.PRS think.PTCP I
que no volia tornar-ho a provar.
QUE not want.1SG.IPFV.PST do.again-CL.N to try
'[reportative:] The mother: what were you thinking. [reportative:] I: I didn't want to try it again.' (ebook-cat)

Finally, the interpretation of these frame setters is context dependent. Even material that is not mentioned explicitly can play a role. In (62a) the frame is interpreted as the origin of a reported statement and in (62b) as the speaker of a reported statement. Therefore, the denotation of a *verbum dicendi* clearly plays a role in the interpretation of the frame, even though in neither case is it part of the words that make up the current sentences.

To conclude, in this section I have discussed the syntactic properties of unembedded *que*-initial reportatives. I showed that in these sentences, *que* precedes CP-internal material but follows CP-external material. This is in line with the analysis that locates the complementizer in SubP, the highest projection of the split CP. The section has furthermore drawn parallels with syntactically subordinate indirect reportatives and has demonstrated that they exhibit the same behavior with respect to CP-internal peripheral material and non-declarative sentences and in permitting reportatives that are not full sentences. This corroborates my idea that unembedded *que*-initial sentences are structurally the same as their embedded counterparts, the only difference being that there is no matrix sentence. Finally, in the last part of the section, I have shown that phrases that precede *que* do not always express an original speaker or information source, which is predicted by Corr (2016), and have suggested an analysis of these phrases as frame setters.

2.4 Contrasting *que*-initial reportatives in Portuguese

The empirical evidence in the previous section was drawn from Spanish and Catalan. Portuguese also exhibits instances of *que*-initial sentences that are interpreted as reported speech. While they do differ in some aspects, I claim that they are still one and the same syntactic phenomenon as in Spanish and Catalan. In this section I outline my reasons for this claim. I show that the slightly different properties of Portuguese unembedded reportatives can be explained as resulting from independent syntactic properties and additional pragmatic requirements. The example in (64) illustrates a *que*-initial reportative in Portuguese. The sentence that follows the complementizer reports what the gynecologist said to her patient. This is a plausible interpretation because in the preceding context the act of talking to the gynecologist is made salient.

(64) European Portuguese
Fiz anestesia geral mas só lá fiquei uma
make.1SG.PRF.PST anesthetic general but only there stay.1SG.PRF.PST one
noite, porque falei logo com o meu ginec. Que não
night because talk.1SG.PRF.PST then with the my gynecologist QUE not

valia a pena lá ficar se depois não era
be.worthwhile.3SG.IPFV.PST there stay if afterwards not be.1SG.IPFV.PST
vista por mais nenhum médico.
seen by more any doctor
'I got a general anesthetic but I only stayed there for one night because I talked to my gynecologist. [reportative:] It was not worthwhile staying if in the end I wasn't seen by any other doctor.' (CdP)

(65) a. European Portuguese
O Sr. gosta de chafurdar na merda ideológica, filosófica,
you like.3SG.PRS to wallow in.the shit ideological philosophical
propagandística. Gosta de dizer que é muito bom,
propagandistic like.3SG.PRS to say that be.3SG.PRS very good
sabe muito bem, que possui uma viscosidade suave
taste.3SG.PRS very well QUE possess.3SG.PRS a viscosity soft
e um calor aconchegante. [$_{SubP}$ Que] [$_{ModP}$ claro que$_i$] [$_{FinP}$ t$_i$]
and a heat cozy QUE claro QUE
podia ser melhor, e será uma vez que
can.3SG.COND be better and be.3SG.FUT once that
estivermos todos afogados nela, e que um gajo até se
be.1PL.SBJV.FUT all drowned in her and QUE a guy even CL.REFL
habitua ao cheiro rapidamente.
habituate.3SG.PRS to.the smell quickly
'You like to wallow in ideological, philosophical and propagandistic shit. You like to say that it is very good, it tastes nice. [reportative:] It has a soft viscosity and a cozy warmth. [reportative:] Of course, it could be better and it will indeed be once we have all drowned in it. [reportative:] And one even gets used to the smell quickly.' (CdP)

b. Brazilian Portuguese
Perguntei se ele lembrava de mim e ele
ask.1SG.PRF.PST whether he remember.3SG.IPFV.PST of me and he
disse [$_{SubP}$ que] [$_{ModP}$ claro que$_i$] [$_{FinP}$ t$_i$]
say.3SG.PRF.PST that claro QUE
lembrava.
remember.3SG.IPFV.PST
'I asked whether he remembered me and he said that of course he remembered.' (CdP)

With regard to its syntactic properties, the complementizer in Portuguese unembedded reportatives occupies the same position as in their syntactically embedded counterparts. Furthermore, this position can also be identified with SubP, the highest position in the split CP. Evidence for this is given in (65). The examples illustrate that in embedded (65b) and unembedded (65a) reportatives alike the complementizer surfaces in a high position, preceding for instance AdvC, which I analyze as being in ModP. Importantly, this is the same behavior that we saw for Spanish and Catalan reportative sentences.

The examples in (66) show that Portuguese again behaves in the same way as Spanish and Catalan as far as sentence fragments are concerned. They are permitted in syntactically embedded (66b) and unembedded reportatives (66a).

(66) a. European Portuguese
Quando cheguei a casa perguntei à minha
when arrived.1SG.PRF.PST at home ask.1SG.PRF.PST to.the my
mãe se agora as mulheres bebiam cerveja no café. Que
other if now the woman.PL drank beer at.the coffeehouse QUE
sim. E muito.
yes and loads
'When I arrived at home I asked my mother whether nowadays women drank beer at the coffeehouses. [reportative:] yes: loads!' (CdP)

b. Brazilian Portuguese
40 alunos o total de 71,4% responderam que o conteúdo
40 student.PL the total of 71.4% answer.3PL.PRF.PST that the content
associado à aula prática facilita o aprendizado; 6
associated to.the class practical facilitat.3SG.PRS the learning 6
(10,7%) responderam que não e 10 (17,9%) disseram
(10.7%) answer.3PL.PRF.PST that no and 10 (17.9%) say.3PL.PRF.PST
que talvez.
that maybe
'40 students, a total of 71.4%, answered that the content of the practical class facilitated their learning; 6 (10.7%) answered that it didn't and 10 (17.9%) said maybe.' (CdP)

These data confirm that, with regard to their principal syntactic properties, unembedded reportatives in Spanish, Catalan and Portuguese are not distinct from one another. In all three languages, the complementizer occupies the same position, identified as SubP in the present analysis, in embedded and unembedded reportatives.

There are two respects, however, in which Portuguese unembedded reportatives do differ from the other languages. The first relates to reported questions. While Spanish and Catalan admit *wh*-pronouns and the interrogative complementizer below *que*, Portuguese does not. This is illustrated in (67a) for the *wh*-pronoun *onde* and in (67b) for the interrogative complementizer *se* (see also Corr 2016: 178: ex 61). Note that the culprit is not the unembedded question introduced by a subordinating element, as shown by example in (67b). A's repetition is different from the initial question: It is introduced by the interrogative complementizer *se*, marking the sentence as subordinate.

(67) Portuguese

a. A: Onde estás?
where be.2SG.PRS

B: O que é que disseste?
what be.3SG.PRS that say.2SG.PRF.PST

A: (*Que) onde estás.
QUE where be.2SG.PRS

'A: Where are you? B: What did you say? A: [reportative:] where are you.

b. (adapted from Corr 2016: 178: ex 61)
A: Vens?
come.2SG.PRS

B: O que é que disseste?
what be.3SG.PRS that say.2SG.PRF.PST

A: (*Que) se vens.
QUE if come.2SG.PRS

'A: Are you coming? B: What did you say? A: [reportative:] are you coming.'

What is important is that the inability of *que* to precede a *wh*-pronoun or the interrogative complementizer is not unique to unembedded sentences but also extends to syntactically embedded reportatives (see (68a) and (68b)).

(68) Portuguese

a. A Joana pergunta (*que) onde estás.
the Joana ask.3SG.PRS that where be.2SG.PRS
'Joana asks (that) where you are.'

b. A Joana pergunta (*que) se vens.
the Joana ask.3SG.PRS that if come.2SG.PRS
'Joana asks (that) if you are coming.'

This differs from what is observed in Catalan and Spanish, where the complementizer *que* can introduce *wh-* and polar interrogatives irrespective of whether the sentence is syntactically subordinate or not (cf. 29a, 30a and 35, 36).

As briefly mentioned in §2.3.1, Spanish and Catalan also permit embedded questions without an initial complementizer (69b), paralleling the Portuguese structure in (68a). In (69) the two structures are contrasted (see also González i Planas 2014 for more discussion on these contrasts, and also on embedded polar questions and exclamatives). The two sentences are not equivalent: (69a) is an embedded *wh*-question, while (69b) is a declarative in which the *wh*-pronoun is referential. This is demonstrated by the fact that (69b) can admit a continuation that spells out the referent, while the same continuation is infelicitous following (69a).

(69) Spanish

a. Juana repetió que quien vivia en la Rua da
Juana repeat.3SG.PRF.PST that who live.3SG.IPFV.PST in the Rua da
Saudade, #era Pereira.
Saudade is.3SG.IPFV.PST Pereira
'Juana repeated: who lived in Rua da Saudade: #it was Pereira.'

b. Juana repetió quien vivia en la Rua da
Juana repeat.3SG.PRF.PST who live.3SG.IPFV.PST in the Rua da
Saudade: era Pereira.
Saudade is.3SG.IPFV.PST Pereira
'Joana repeated who lived in Rua da Saudade: it was Pereira.'

The example in (70) shows that the structure is ambiguous in spoken Portuguese. It admits a continuation that spells out the referent but also a continuation that makes the question-reading prominent. It should be noted that in written text, this ambiguity is resolved through orthography. The question reading would be marked with symbols signaling direct speech. Similarly the Spanish version in (69b) would be presented differently in written form.

(70) Portuguese
A Joana repetiu quem morava na Rua da
the Juana repeat.3SG.PRF.PST who live.3SG.IPFV.PST in.the Rua da
Saudade:
Saudade

[continuation 1:] era Pereira.
is.3SG.IPFV.PST Pereira

[continuation 2:] mas ninguem o sabia.
but nobody it know.3SG.IPFV.PST

'Joana repeated who lived in Rua da Saudade: it was Pereira/but nobody knew it.'

The important empirical generalization is that Portuguese does not allow *que* to appear before a *wh*-pronoun or an interrogative complementizer in unembedded reportatives, but crucially also not in embedded reportatives, while Spanish and Catalan allow this ordering in both contexts. This once again supports my conclusion that *que*-initial reportatives are simply syntactically unselected reportatives with essentially the same properties as their embedded counterparts. What is fundamental to my argumentation is that these facts permit a uniform syntactic treatment of the Portuguese unembedded reportatives and the equivalent structures in the other two languages. The behavior of Portuguese *que*-initial reportatives is expected based on the proposal outlined in this book, which assumes the same structure for unembedded and embedded reported sentences. Indeed, it actually constitutes an argument in its favor. Within the general analyses proposed in this book, the syntactic contrasts between Portuguese on the one hand and Catalan and Spanish on the other, can be explained as a reflex of a difference in the feature specification of *wh*-pronouns and interrogative complementizers. I propose that the contrast boils down to the need to check a feature of a particular head. In Portuguese *wh*-pronouns and interrogative complementizers necessarily check the subordinate feature in SubP, while in Spanish and Catalan, there is one version where the feature is checked by the *wh*-pronoun and one version where it is not. If the feature is checked by the *wh*-pronoun, the result is the structure in (69b), but if the feature is not checked by the *wh*-pronoun, a complementizer is merged to check the feature which then results in the structure in (69a).

The second difference that distinguishes Portuguese *que*-initial reportatives from those in Spanish and Catalan relates to the contextual requirements that render these sentences felicitous. The central contrast is illustrated in (71).

(71) a. Spanish

A: No se oye bien.
not CL.REFL hear.3SG.PRS well

B: ¿Qué? ¿Eh?
what huh

A: Que no se oye bien.
QUE not CL.REFL hear3SG.PRS well

'A: You can't hear it well. B: Huh? A: (I said) you can't hear it well.'

b. Portuguese (adapted from Corr 2016: 148: ex 5)

A: Não se ouve bem.
not CL.REFL hear.3SG.PRS well

B: O que? Hein?
what huh

A: * Que não se ouve bem.
QUE not CL.REFL hear.3SG.PRS well

B': O que é que disseste?
what be.3SG.PRS that say.2SG.PRF.PST

A': Que não se ouve bem.
QUE not CL.REFL hear.3SG.PRS well

'A: You can't hear it well. B': What did you say? A': That you can't hear it well.'

While *que*-initial reportatives are felicitous without an explicit *verbum dicendi* in the context in Spanish and Catalan (71a), they are not in Portuguese where a *verbum dicendi* in the context is required (71b). Corr (2016) considers this difference substantial enough to propose two very different syntactic analyses for Portuguese and the latter two languages (cf. §2.1).

One empirical problem faced by Corr's proposal is that Spanish and Catalan also exhibit examples of the Portuguese type of reportatives, i.e. of *que*-initial sentences where a *verbum dicendi* can be recovered from the context, cf. (72).

(72) Spanish

A: No se oye bien.
not CL.REFL hear.3SG.PRS well

B: ¿Qué dices?
what say.2SG.PRS

A: Que no se oye bien.
QUE not CL.REFL hear.3SG.PRS well
'A: You can't hear it well. B: What did you say? A: That you can't hear it well.'

The important point is that the only difference between (71a) and (72) is that in the latter a verb of saying is previously mentioned in the context while in the former it is not. Crucially, no apparent syntactic differences are observed (cf. also the discussion of example (17) in §2.1). Corr (2016) unfortunately does not address how these facts are dealt with in her account, nor is it clear whether she assumes two different analyses for (71a) and (72).

Another aspect that makes the assumption of two different structures less convincing is that Portuguese unembedded reportatives do not exhibit any evident syntactic differences compared to the same construction in Spanish and Catalan. The only exception is that *wh*-pronouns and interrogative complementizers cannot follow *que*, which I showed is not unique to the phenomenon under investigation. On the contrary, the data presented in this section confirm that, just as in Spanish and Catalan, the complementizer behaves in the same way in syntactically embedded and unembedded sentences. Portuguese reportatives are therefore compatible with an analysis that assumes that the complementizer is merged in the highest CP position.

Given the absence of any real syntactic difference, I consider that the true contrast between Portuguese on the one hand and Spanish and Catalan on the other resides purely in the pragmatic conditions imposed on the context. These pragmatic conditions are the topic of §2.5. In a nutshell, the difference is that while all three languages require salient material that can function as a host for the subordinate sentence, Portuguese has a stronger requirement: The host material must be given in the context. This means that in Portuguese a verb of saying must have been mentioned previously. While givenness, i.e. a previous mention, is one way to render an expression salient, it is not the only way to achieve this. Extra-linguistic factors, for instance, can also come into play. Consequently, in this case, while a given expression is always salient, the reverse implication does not hold: A salient expression is not always given. This is an important fact that plays a role in the contrast between Portuguese on the one hand and Spanish and Catalan on the other. The empirical data suggest that in a conversational exchange such as that in (71a) a *verbum dicendi* is salient. This is sufficient to render a *que*-initial reportative sentence felicitous in Spanish and Catalan. In Portuguese, on the contrary, a conversational exchange is not enough: A *verbum dicendi* truly must be given in order to felicitously utter a *que*-initial reportative.

To conclude, in this section I have shown that Portuguese *que*-initial reportatives share crucial properties with their equivalents in Spanish and Catalan, and I have explored in detail the aspects in which they differ from each other. In line with the empirical facts, I argued that Portuguese reportative *que* constructions are syntactically not distinct from those in Spanish and Catalan. In all three languages they constitute sentences that are marked as subordinate by a complementizer merged in the highest CP projection SubP. Based on this, they are expected to behave in the same way as syntactically embedded reportatives. I have shown that this is indeed the case. Furthermore, I argued that the apparent syntactic differences between Spanish and Catalan on the one hand and Portuguese on the other are not unique to unembedded reportatives but are also found in embedded reportatives, which ultimately constitutes an argument in favor of the present analysis. Finally, I defended the claim that the true differences are related to pragmatics. In all three languages, there is a requirement for contextually salient material that can function as a host for the subordinate sentence. The difference is that in Portuguese, the material must be mentioned previously while in Spanish and Catalan, mere saliency of the relevant material is sufficient.

2.5 The pragmatics of *que*-initial reportatives

The previous sections have focused on the syntactic properties of unembedded reportative sentences. The empirical facts supported my claim that the syntactic behavior of *que* in these sentences does not differ from that of their syntactically embedded counterparts. I furthermore proposed that the data can be analyzed without needing to resort to an underlying performative or evidential syntactic structure to account for the reportative interpretation. The claim I make is that the only information that the syntactic structure provides is that the sentence is subordinate.[13] This section presents the pragmatic properties of *que*-initial reporatives and lays out how I assume that the reportative interpretation arises. I propose that this is not encoded in a hidden syntactic structure or a dedicated syntactic feature but instead results from the contexts in which the sentences appear. I will argue that when uttering a proposition introduced by *que*, the speaker asserts that the proposition is subordinate. This leads the hearer to infer that there must be a salient linguistic expression that can function as a host for the subordinate proposition.

[13] I am unaware of any study on the prosody of *que*-initial reportatives. However, potentially, intonation might play an additional role in encoding the meaning of these sentences.

As Corr (2016) observes, in *que*-initial reportatives the speaker does not assert the proposition corresponding to the sentence headed by the complementizer and hence the speaker is not committed to its truth. Corr illustrates this by the fact that with an unmarked declarative in (73a), the speaker cannot negate that she believes that the proposition is true, but in (73b), with a *que*-initial declarative, she can.

(73) Spanish

a. (adapted from Corr 2016: 158: ex 23)
Juana y Aique están casados, #pero no es verdad.
Juana and Aique be.3PL.PRS marry.PTCP but not be.3SG.PRS truth
'Juana and Aique are married, but it's not actually true.'

b. (adapted from Corr 2016: 158: ex 24)
Que Juana y Aique están casados, pero no es
QUE Juana and Aique be.3PL.PRS marry.PTCP but not be.3SG.PRS
verdad.
truth
'Someone said Juana and Aique are married, but it's not actually true.'

In this respect, *que*-initial reportatives behave just like indirect embedded reportatives. In (74) the embedded indirect reportative is felicitous even though it is followed by a continuation that clarifies that the speaker made a statement he does not consider to be true.

(74) [Galileio:] I said that the earth is flat. But I don't actually believe that it is.

Que-initial sentences do not have an independent illocutionary force (see Etxepare 2010). For instance, *que*-initial questions do not have the illocutionary force of a question but that of a declarative. The reported question in (75) does not require an answer from the hearer. Therefore a continuation where the hearer provides an answer appears odd. However, a reply from the hearer that does not answer the question but expresses his delight about the fact that someone asked about his well-being is perfectly acceptable.

(75) Spanish
María: Antes de salir he hablado con mi madre. Le
Maria before of leave AUX.1SG.PRF.PRS talk.PTCP with my mother CL.3SG
he dicho que estoy a punto de encontrarme
AUX.1SG.PRF.PRS tell.PTCP that be.1SG.PRS about to meet.CL.REFL
contigo. Y ella, que ¿qué tal Jorge?
with.you and she QUE how's Jorge

Jorge: #Estoy bien.
Jorge be.3SG.PRS fine

Jorge': ¡Qué amable!
Jorge how nice
'Maria: Before I left I talked to my mother. I told her I am about to meet with you. And she was like how is Jorge. Jorge: #I am fine. Jorge': How nice of her.'

Another concern in the literature has been the information status of the reported sentence. Etxepare (2010) and Corr (2016) claim that the original sentence, which is the basis of the report, needs to be traceable. Traceability in Etxepare (2010: 613) means that a reported *que*-sentence refers "to a contextually identified utterance preceding the report". Consequently, in this conception, a sentence is traceable if there is a past speech event that is salient, i.e. accessible to the speaker and the hearer, in which that sentence was first uttered. Etxepare (2010) illustrates this with the example in (76).

(76) Spanish (adapted from Etxepare 2010: 613: ex 34)
[Context: A and B share an office at a bank. B asks A about a particular transaction. A asks C in another office about the transaction. A tells B:]
Oye, que ya be.3SG.PRS hecho.
DM QUE already be.3SG.PRS done
'Hey, it's already done.'

The reportative sentence introduced by *que* is traceable in the context because it can be traced back to the previous speech event in which speaker A asked speaker C about the transaction.

There are, however, examples that show that the concept of traceability in its current formulation might be too strong. The examples in (77) and (78) show that it is possible to introduce future or even hypothetical utterances with *que*. In these cases, no previous uttering of the sentence could have taken place and thus it is not traceable in the sense of Etxepare (2010).

(77) Catalan
Avisa el comissari. Que ja pot venir.
notify.2SG.IMP the inspector QUE already can.3SG.PRS come
'Notify the inspector. [reportative:] He can already come.' (ebook-cat)

In (77) *que* introduces a future utterance. The speaker proposes the *que*-initial sentence to the hearer for him to utter at a future speech event. In (78) the *que*-initial sentence is hypothetical. The speaker thereby introduces a question that she anticipates before the hearer has even asked it.

(78) Catalan
És una cosa que no t' havia explicat mai,
be.3SG.PRS a thing that not CL.2SG AUX.1SG.IPFV.PST tell.PTCP ever
estimada. No vaig gosar dir-t'ho. Que per què
my love not AUX.1SG.PRF.PST have courage tell-CL.2SG.CL.N QUE why
ho vaig fer? Perquè jo no sóc el meu pare.
CL.N AUX.1SG.PRF.PST do because I not be.1SG.PRS the my father
'It is a thing that I never told you, my love. I didn't have the courage to tell you. [reportative:] why I did it. Because I am not my father.' (ebook-cat)

These data show that traceability, in its current definition, does not appear to be a requirement for *que*-initial sentences. Sentences can be introduced by *que* even before the "original" sentence was uttered and even in cases when it will never be uttered at all. This means that, contrary to what Etxepare (2010) and Corr (2016) assume, for a *que*-initial reportative to be felicitous, a previous speech event in which the original sentence was first uttered is not a necessary prerequisite. The data discussed above furthermore show that there are no clear conditions on the information status of the *que*-initial sentence: It can be given as in (76) but also new as in (77) and (78).

My own proposal does not focus on the information status of the *que*-initial sentence itself but on other contextual requirements. There are two important concepts already introduced in §2.4, which are key to the following discussion: Givenness and salience. Given is used here in the sense of Schwarzschild (1999) (but cf. also Rochemont 2016 for a recent comprehensive definition of givenness). Descriptively, it can be thought of as "previously mentioned" (Büring 2003). In a straightforward case, an expression is given if there is a previously uttered expression that is identical to it (*woman–woman*). However, an expression can also be characterized as given if there is a hyponym (*woman–human*), a co-referent expression (*a woman–the woman/she*), or a semantically vacuous expression (*someone*) that was mentioned previously.

Salience, in turn, refers to expressions that are prominent (Chiarcos et al. 2011), activated (Chafe 1976) or accessible (Ariel 1990) in a specific speech event. An expression can be salient because it was previously mentioned, i.e. is given. Givenness is however not a necessary requirement for saliency. An expression can also

be salient for other reasons, for instance because it constitutes general knowledge or because extra-linguistic factors make it prominent.

With these concepts in place, I will now show that a felicitous *que*-initial reportative requires a salient *verbum dicendi*. As an aside, this requirement also provides support for the argument that no reportative information is encoded in the syntactic structure of the *que*-initial sentence itself because if it were, the need for a contextually salient verb of saying would not be expected.

In the examples from (79) to (82), I exemplify cases of *que*-initial reportatives in different contexts to illustrate my claim. In (79) an entire matrix clause can be reconstructed from the context. The matrix clause from the previous sentence, *(la mare) va dir* '(the mother) said', is an adequate host for the following *que*-initial reportative sentence and contains the *verbum dicendi dir* 'say'.

(79) Catalan
Tenia ganes de començar o amb el rus o amb
want.1SG.IPFV.PST to start either with the Russian or with
l'arameu, però la mare va entrar a l'habitació i
the Aramenaic but the mother AUX.3SG.PRF.PST enter to the room and
va dir ni parlar-ne. Que ja
AUX.3SG.PRF.PST say not even talk.about-CL.PART QUE already
estava bé amb aquelles llengües que sabia.
be.3SG.IPFV.PST well with these languages that know.3SG.IPFV.PST
'I wanted to start with Russian or Aramaic but my mother came into my room and said: "That's out of the question". That the languages I already knew were enough.' (ebook-cat)

In (80) repeated from (71b), B's question introduces the host material. This example shows that there is no identity requirement: The host material is of a different clause type (interrogative vs. declarative) and has different verbal agreement (second vs. first person singular) than what is expected as a host for the unembedded sentence. Importantly, however, a *verbum dicendi* is involved.

(80) Portuguese (adapted from Corr 2016: 148: ex 5)
A: Não se ouve bem. B: O que é que
not CL.REFL hear.3SG.PRS well what be.3SG.PRS that
disseste? A: Que não se ouve bem.
say.2SG.PRF.PST QUE not CL.REFL hear.3SG.PRS well
'A: You can't hear it well. B: What did you say? A: That you can't hear it well.'

The example in (81) is different because no *verbum dicendi* is given. The relevant *que*-initial reportative is uttered within a conversational exchange. This seems to be sufficient to make unembedded reportatives felicitous in Spanish and Catalan. The conclusion within the present explanation must be that a *verbum dicendi* is salient in a conversational exchange.

(81) Spanish
Inf.a - Hoy es el día de salida. Inf.b - ¿Eh? Inf.a. - Que
today be.3SG.PRS the day of departure Hm? QUE
hoy es el día de salida.
today be.3SG.PRS the day of depature
'A: Today is the day of departure. B: Hm? A: [reportative:] Today is the day of departure.' (CdE)

Finally, in the example in (82) repeated from (78), the *que*-initial reportative appears in a context without a given *verbum dicendi* and without an active interlocutor. Just as in the case of (81), in such a context an unembedded reportative is only possible in Spanish and Catalan. In the fragment, there is only one speaker present. However, a hearer is addressed with *estimada*. This suggests that activating a hearer perspective is enough for a *verbum dicendi* to become salient.

(82) Catalan
És una cosa que no t' havia explicat mai,
be.3SG.PRS a thing that not CL.2SG AUX.1SG.IPFV.PST tell.PTCP ever
estimada. [...] Que per què ho vaig fer?
my love QUE why CL.N AUX.1SG.PRF.PST do
'It is a thing that I never told you, my love. [...] [reportative:] why I did it.' (ebook-cat)

The latter examples, in particular, show that an ellipsis analysis is not a useful means of accounting for unembedded reportatives because it would require syntactic reconstruction and the preconditions (deletion under identity) are not met. These examples are therefore a central piece of evidence illustrating the need for a distinction between syntactic reconstruction, required for ellipsis, and pragmatic reconstruction, required for the phenomenon at hand.

How the precise mechanism of pragmatic reconstruction works will not be explored in detail here. However, there are other phenomena that show that a theory of how meaning is picked up pragmatically is necessary. Hankamer & Sag (1976) illustrate that the interpretation of anaphoric pronouns in some cases

requires syntactic control (surface anaphora in their terminology), yet in other cases pragmatic control (deep anaphora in their terminology) is sufficient. A case of pragmatic control is illustrated in (83).

(83) [Scenario: Sag produces a cleaver and prepares to hack off his left hand]
Hankamer: ... he never actually does it. (Hankamer & Sag 1976: 392: ex 6b)

The referents for the pronouns *he* and *it* are not given in the linguistic context. The pronouns are therefore not controlled syntactically. The referents can be inferred from the general context because they are salient. According to Hankamer & Sag (1976), they are controlled pragmatically. Jacobson (2012) introduces the example in (84a), illustrating the same issue. The pronoun *he* is pragmatically inferred to be *Tony*, a referent who is salient but not given. In (84b), *it* is understood to mean *diving off the high diving board*, once again a salient yet not given expression.

(84) a. [Scenario: We are at a party, and a very obnoxious guy named Tony comes in. No one likes Tony, so no one talks to him all night, and no one mentions his name. He leaves, and I turn to you and say:]
Thanks goodness he left. (Jacobson 2012: 4: ex 13)

b. [Scenario: I know that for years my friend Chris has wanted to dive off the high diving board, but every time he gets up there he gets scared and climbs down. I see him on the high diving board one afternoon, and I turn to you and say:]
Poor Chris, I don't think he'll do it. (Jacobson 2012: 4: ex 14)

Jacobson (2012) makes a more radical proposal than Hankamer & Sag (1976). She suggests that the interpretation of anaphoric pronouns always involves pragmatically picking up a salient referent. In this sense, syntactic control is an illusion that appears to exist because one very efficient means of making things salient is to explicitly mention them. Notably, this is very close to what the present approach assumes for the reconstruction of the host material of the *que*-initial sentences. The minimal requirement to allow an unembedded sentence to be felicitously uttered is a salient expression that can function as its host. The expression can be salient in the general (extra-linguistic) context (cf. examples 81–82) or it can be made salient through an explicit mention (cf. examples 79, 80).

To conclude, in §2.3 and §2.4, I argued that there are no real syntactic differences between *que*-initial reportatives in Spanish and Catalan on the one hand and Portuguese on the other. I furthermore suggested that the behavior of *que* matches that of a typical complementizer, which goes against the assumption

that *que* is an evidential marker in these sentences. The conclusion I draw is that unembedded reportatives are essentially the same in all three languages and can be analyzed uniformly. This section has explored the pragmatics of *que*-initial reportatives. I argued that the minimal contextual requirement for a *que*-initial reportative is a salient *verbum dicendi* that can function as a host for the subordinate sentence. Portuguese differs from Spanish and Catalan in that a *verbum dicendi* needs to be given while in Spanish and Catalan givenness is not required.

2.6 Beyond reportatives

The focus of this section is to present further evidence for the analysis laid out over the course of this chapter by showing that there is no one-to-one-relation between a *que*-initial sentence and a reportative interpretation. On the contrary, I illustrate in this section that depending on the context a *que*-initial sentence can receive a variety of different interpretations. Since the approach developed here does not allude directly to any reportative meaning, this fact does not come as a surprise. Insubordination is not specialized for reports but it is a more pervasive phenomenon in Ibero-Romance languages. This section shows how these data can be accounted for in the present analysis. The central idea that my analysis is built upon is that the syntactic structure of unembedded sentences is essentially the same as that of their embedded counterparts. The complementizer surfaces in the highest position of the left periphery and receives a subordinate feature that has implications for the interpretation of the sentence: Namely that the sentence is semantically subordinate to a salient linguistic expression. There are two important empirical facts that support this analysis. Firstly, *que*-initial sentences show the same syntactic properties as their embedded counterparts. Secondly, they are only felicitous in contexts in which a plausible host is salient.

In what follows, I present a number of examples of *que*-initial sentences that are not interpreted as reports. The first set of data are sentences in contexts of mental predicates and as answers to questions. The second set are unembedded relative sentences.

In the examples in (85), a *que*-initial sentence appears in the context of a given mental predicate. These examples are very similar to the unembedded reported sentences presented in the previous sections, some of which also appeared in the context of a given verb of saying. In (85a) the juxtaposed *que*-initial sentences are understood to be subordinate to the mental predicate *mentalizar* ‘internalize’, which appears in the matrix clause of the first sentence of the fragment. In (85b), the unembedded sentence appears in the context of the given mental predicate

pensar 'think'. This case differs from that illustrated in (85a) in that the mental predicate is uttered by a different speaker than the one who utters the *que*-initial sentence.

(85) a. Brazilian Portuguese
Mentalize que você criou seu filho bem.
mentalize.2SG.IMP that you bring.up.3SG.PRF.PST your child well
Que você escolheu a dedo a sua escola. Que você sempre
QUE you handpick.3SG.PRF.PST the his school QUE you always
pode criar um laço de amizade com a professora.
can.3SG.PRS create a tie of friendship with the teacher
'Internalize that you brought your child up well. (Internalize) that you handpicked his school. (Internalize) that you can always create a bond of friendship with the teacher.' (CdP)

b. Spanish
Enc. ¿qué piensa de su trabajo? Inf. Que
interviewer what think.3SG.PRS of your work informant QUE
es agobiador.
be.3SG.PRS exhausting
'Interviewer: What do you think about your job? Informant: (That) it is exhausting.' (CdE)

The analysis proposed here does not explicitly allude to *verba dicendi*, so *que*-initial sentences in contexts of mental predicates like (85a) can be easily accommodated. The salient expression that is pragmatically reconstructed as the host of the subordinate sentence is simply a mental predicate.

Answers to *wh*-questions, as illustrated in (85b), are also typical contexts for *que*-initial sentences. Often the question contains a mental predicate like in (85b) or a verb of saying as in (86a). However, there are also examples like (86b), where neither is the case.

(86) a. Brazilian Portuguese
Então o período de escravidão no Brasil é amplo,
so the period of slavery in.the Brazil be.3SG.PRS ample
diversificado, mas o livro didático, durante muito tempo
diversified but the book didactic during much time
disse o quê? Que você era escravo,
say.3SG.PRF.PST the what QUE you be.3SG.IPFV.PST slave

propriedade de alguém, que apanhava, que
property of someone QUE take.a.beating.3SG.IPFV.PST QUE
era inferior, que aceitava a escravidão.
be.3SG.IPFV.PST inferior QUE accept.3SG.IPFV.PST the slavery
'So the period of slavery in Brazil was ample and diversified. Still, what did school books say for a long time? (That) you are a slave, somebody's property. (A slave that) took a beating, (that) was inferior and (that) accepted slavery.' (CdP)

b. Spanish
Cuáles son las falsas leyendas en la vida de George Sand? -
which be.3PL.PRS the false legend.PL in the life of George Sand
Que sus dos grandes amores fueran Alfred de Musset y
QUE her two big love.PL be.3PL.PRF.PST Alfred de Musset and
Chopin, cuando en realidad han contado poco.
Chopin when in reality AUX.3PL.PRF.PRS count.PTCP little
'What are the false legends of George Sand's life? - That her two big love affairs were with Alfred de Musset and Chopin when really they were of little importance.' (CdE)

These examples could be viewed as cases of fragmented answers. Many analyses of these types of reduced answers assume deletion under identity stating that parts of the answer can be elided because they are semantically identical to the content of the question.[14] Jacobson (2016), however, disagrees with these analyses. She shows that, contrary to what these types of analyses predict, short, fragmented and long, full answers are not semantically identical. One point of divergence is that in the short answer by B in (87), the speaker is committed to the belief that Jill is a mathematics professor. This makes the continuation in which the speaker withdraws the commitment infelicitous. With the long answer in B' the speaker offers the best she can do, but does not commit herself to the belief that Jill is a mathematics professor. Therefore the continuation is felicitous.

(87) A: Which mathematics professor left the party at midnight?
B: Jill. #Whether she is a mathematics professor I'm not sure.
B': Jill left the party at midnight. Whether she is a mathematics professor I'm not sure. (adapted from Jacobson 2016: 342: ex 14)

Similarly, the long and short answers in (88) are not identical. In the short answer the speaker is committed to the fact that what the book says about slavery

[14] On formal and semantic identity see Jacobson (2016: 353–356).

in Brazil is that the slaves are someone's property. In the long answer she is not committed. She might just offer her best guess by saying something the book says about slavery in general but not necessarily about the particular case of slavery in Brazil.

(88) Spanish
A: Qué dice el libro sobre esclavitud en Brasil?
What say.3SG.PRS the book about slavery in Brazil
B: Que los esclavos eran propiedad de alguien.
QUE the slave.PL be.3PL.IPFV.PST property of someone
B': El libro dice que los esclavos eran propiedad de alguien.
the book say.3SG.PRS that the slave.PL be.3PL.IPFV.PST property of someone
'A: What does the book say about slavery in Brazil? B: That slaves were someone's property. B': The book says that slaves were someone's property.' (CdE)

Based on these facts, Jacobson (2016) develops an analysis that accounts for short answers without the need to allude to silent material. Without going into the details of her analysis, one important prediction it makes is that a short answer can only be correctly interpreted when there is a salient question in the context. Importantly, this is very close to the felicity conditions I identified for *que*-initial sentences.

Finally, there are also unembedded relative clauses. One example is given in (89). The *que*-initial sentence is understood as a modifier that encodes yet another property of the NP *gent* 'people' from the preceding matrix sentence. The interpretation of the *que*-initial sentence is possible once again because there is a salient linguistic expression in the context that can function as a host for the subordinate sentence.

(89) Catalan (Kocher 2017b: 52: ex 103)
És gent que segur que has vist, però mai
be.3SG.PRS people that sure que AUX.2SG.PRF.PRS see.PTCP but never
has passat un cap de setmana amb ells. Que segur que
AUX.2SG.PRF.PRS pass.PTCP a weekend with them QUE sure QUE
conèixes però que mai te ha convitat a
know.2SG.PRS but that never CL.2SG AUX.3SG.PRF.PRS invite.PTCP to

casa seva.
house their
'These are people who surely you have met but never have passed a weekend with. Who surely you know but who never have invited you over to their house.'

The example in (89) shows that unembedded relatives share core properties that we have identified above. Here too, *que* shows the same syntactic behavior in embedded and unembedded sentences. The complementizer occupies the same high position in both cases. In both sentences it precedes *segur que*, which is analyzed as being in ModP. This is compatible with the analysis presented in this chapter that *que* occupies a high position in the left periphery, which is potentially also SubP in this case. The example also indicates that *que*-initial relatives, like the other types of unembedded sentences, require a salient linguistic expression that can function as a host for the subordinate sentence. In the case of relative clauses this is the modified noun phrase. In (89), a parallel structure consisting of an entire matrix clause and a relative clause can be recovered from the linguistic context. However, there are also examples where this is not the case. In (90a) and (90b) the *que*-initial sentences are understood as relative clauses modifying the noun phrases *ella* and *gente*, without a parallel relative structure in the context. These examples show that, just as with *que*-initial reportatives, a salient potential host expression is sufficient to render a *que*-initial sentence felicitous. This suggests that an analysis that does not allude to syntactic ellipsis, is not just tenable but actually required to account for these data.

(90) a. Spanish
Era extraño interrumpir sus labores y dedicarle
be.3SG.IPFV.PST strange interrupt his works and dedicate.CL.3SG
un minuto a ella. Que había pertenecido a un mundo
a minute to her QUE AUX.3SG.IPFV.PST pertain.PTCP to a world
tan diferente. Que había pertenecido, pero que ya
so different QUE AUX.3SG.IPFV.PST pertain.PTCP but QUE anymore
no pertenecía. ¡Que había dado un paso tan difícil de
not pertain.PTCP QUE AUX.3SG.IPFV.PST do.PTCP a step so hard to
explicar!
describe
'It was strange for him to interrupt his work and dedicate a minute to her. Who belonged to a world so different. Who used to belong there but didn't belong there anymore. Who took a step so difficult to explain.' (CdE)

b. Brazilian Portuguese
Há gente decente neste país. Que trabalha.
there.be.3SG.PRS people decent in.this country QUE work.3SG.PRS
Que não depende de benesses, padrinhos ou tutores. Que
QUE not depend.3SG.PRS on benefit.PL patron.PF or guardian.PL QUE
ouve e não aceita, mesmo em silêncio.
hear.3SG.PRS and not accept.3SG.PRS even in silence
'There are decent people in this country. Who work. Who don't depend on benefits, patrons or guardians. Who hear but don't accept even if only in silence.' (CdP)

A core argument that we are dealing with pragmatic rather than syntactic reconstructions comes from examples like that in (91).

(91) Catalan
Volver es una de las pel·lícules que vaig veure amb
Volver be.3SG.PRS one of the movies that AUX.1SG.PRF.PST see with
en Jordi. Que tota la gent ens recomanava però que
the Jordi QUE all the people CL.2PL recommend.3SG.IPFV.PST but QUE
no ens va agradar gaire.
not CL.2PL AUX.3G.PRF.PST please at all
'*Volver* is one of the movies that I saw with Jordi. (That) everybody recommended to us but (that) we didn't like at all.'

a. Everybody recommended a number of movies, one of them *Volver*, but we didn't like them at all.

b. Everybody recommended *Volver* but we didn't like it at all.

(91) is an example of an unembedded relative in the context of another relative clause. (91a) and (91b) paraphrase two different readings of the unembedded sentence. If the entire matrix clause were syntactically reconstructed we would expect a reading of the *que*-initial sentence like (91a). However, this is not the typical reading for (91). On the contrary, the *que*-initial sentence can be paraphrased by (91b), in which it functions as a relative clause modifying one salient expression: *Volver*. This shows that even when it is possible to syntactically reconstruct a matrix clause, it does not happen. This is strong evidence that something other than syntactic reconstruction is taking place, namely pragmatic reconstruction, as in my proposal.

This section showed that the central predictions following from the analysis developed for *que*-initial reports are also fulfilled by non-reportative *que*-initial sentences. This suggests that we are dealing with a relatively general phenomenon in which sentences are marked as subordinate and require a salient linguistic expression in order to be interpreted.

Finally, it is interesting to note that the translations of the examples discussed along the course of this section show that this phenomenon is not restricted to Ibero-Romance varieties. It appears to extend at least to English. At this stage, it would be premature to claim that we are dealing with the exact same phenomenon, since this would require a detailed comparison and thorough analysis. However, it does hint at the fact that my new conception of subordination and its implications is not only supported by Ibero-Romance but also by other languages.

2.7 Summary

This chapter explored *que*-initial reported sentences. I examined their properties and developed an analysis to account for them. As a new theoretical concept, I established a distinction between selected and unselected subordinate sentences. While both are interpreted as subordinate, only selected subordinate sentences are syntactically dependent on a matrix clause. This theoretical redefinition makes it possible to treat *que* as a normal complementizer, and makes it unnecessary to stipulate a new linguistic category to account for its atypical behavior. It furthermore permits an analysis that does not rely on the ellipsis of a matrix clause nor the assumption of a hidden performative structure. Moreover, I showed that the complementizer in unembedded *que*-initial sentences surfaces in the same position as the complementizer in embedded contexts, suggesting that these complementizers can be analyzed in the same way, namely as heads of the highest left-peripheral position, SubP, valued with the interface feature *subordinate*. The syntactic analysis is simple and, contrary to previous accounts, as mentioned, does not assume a hidden syntactic structure that contributes the special reportative meaning. In my approach, the interpretation results from pragmatic rather than syntactic reconstruction. This assumption is supported by the fact that the types of *que*-initial sentences discussed in this chapter require a salient linguistic expression that can function as a host for the subordinate sentence.

This chapter also presented a description of the pragmatic requirements. In order to formalize the generalizations in the future, a number of questions need to be addressed. One of them is the concept of pragmatic reconstruction. In order

to reach a useful formal definition further research is required to determine what it is that is reconstructed (salient expression, entire matrix clause, etc.). Furthermore, we need to investigate the questions of what being a host for a subordinate sentence means and what properties the salient host has to have to be identified as such.

The simplicity and generality of the analysis predicts that we should expect cases of unembedded sentences that have an interpretation other than that of a reportative. This is why, in the last section of this chapter, I turned to non-reportative unembedded sentences and showed that their properties can indeed be accounted for in principle by the same analysis developed for *que*-initial reportatives. In other words, they are also marked as subordinate and are only felicitous in a context where a host expression can be recovered from the context.

This chapter provided evidence for a clear distribution of labor between syntax and pragmatics. I argued that the information that is read off of the syntactic structure strictly pertains to this component of grammar. The additional information that has an effect on the interpretation of the sentence in its context is contributed by pragmatics.

3 The syntax and pragmatics of commitment-attribution

This chapter explores the syntactic behavior and discourse contribution of the complementizer in the constructions exemplified in (1).

(1) a. Spanish
Tranquilo, tío. Que no muerdo.
calm man QUE not bite.1SG.PRS
'Chill, man. I don't bite.'

b. Spanish
¿Que no muerdes?
QUE not bite.2SG.PRS
'You don't bite?'

c. Portuguese
Que chato que é.
how annoying QUE be.3SG.PRS
'How annoying this is.'

d. Catalan
Segur que son amics.
sure QUE be.3PL.PRS friend.PL
'Surely, they are friends.'

e. Catalan
Sí que son amics.
VERUM QUE be.3PL.PRS friend.PL
'They ARE friends.'

In (1a), the complementizer appears sentence-initially in a declarative and in a polar question in (1b). In (1c), the complementizer surfaces below the *wh*-expression of a *wh*-exclamative. In (1d), for which I use the term AdvC, it follows an epistemic modifier and in (1e), which I will call AffC, the complementizer follows the particle *sí* in a verum construction.

(2) Spanish (adapted from Gras 2016: 121: ex 9)[1]
[Context: Family conversation. B and C are married. They're discussing where to invest their money. Bancaja is a local bank in Valencia, Spain.]

C: antes de sacarlo de la Bancaja preguntaré si me dan
before to take.CL.M.SG of the Bancaja will ask if me give.3PL.PRS
más lo dejoo en la Bancaja […]
more it leave.1SG.PRS in the Bancaja

B: ¿la Bancaja? que no conocemos a nadie ahora te vas a dar
the Bancaja QUE not know.1PL.PRS ACC NOBODY now you go to give
de...
of

C: ¡que conozco yo al director!
QUE know.1SG.PRS I the manager

'C: Before I take it out from Bancaja I will ask them if they give me more interest if I leave it in Bancaja. B: Bancaja? We don't know anybody now you're going to fall flat on... C: I know the manager!'
(Val.Es.Co. VC.117.A.1: 2–15.)

In (2) two cases of *que*-initial declaratives are illustrated in conversational data. In this fragment, both speakers use *que*-sentences to introduce information that they think the other speech participant should have already been aware of. This becomes clear in this context where a couple discuss where to invest their money. In speaker B's *que*-initial sentence she says that *no conocemos a nadie* 'we don't know anyone'. Through the use of the first person plural, speaker B establishes a common perspective between herself and speaker C, and she states that they share the knowledge she is presenting. Speaker C reacts with another *que* sentence, saying that he does in fact know the manager of the bank. The omission of *que* in both of these sentences would not result in ungrammaticality, but it does have a clear effect on the meaning. Both speakers present information not as new, but as something each of them considers to be part of the shared common ground. The conversational effect of *que* is that the disagreement between the two speech participants is highlighted. Not only do they disagree on the facts: Speaker C believes they do not know anyone working at that particular bank whereas speaker B says he does know the manager. In addition, each of them also believes that the other should have known his/her belief to be true.

[1]The original example contained additional encoding relevant for conversational analyses but not for the current purpose.

Data like these are interesting in the light of the discussion regarding where the boundary lies between syntax and pragmatics. Once again the complementizer does not exhibit its prototypical function, but its presence has a distinct pragmatic effect.

In the various sections of this chapter, I will argue that in all of these constructions the complementizer is merged at the rightmost edge of the left periphery and reaches its surface position through head movement. The merge position is therefore distinct from the position that hosts subordinate *que* involved in *que*-initial reportatives in my analysis. Furthermore, I defend the idea that the low-merged complementizer makes the same contribution to the discourse in all the constructions. The effect is that a commitment to the proposition in its scope is ascribed to the hearer. This type of meaning has been termed *attributive* by Poschmann (2008), which is reflected in my decision to use the label *attributive que* (see §3.4 for a more detailed description of the meaning involved).

This chapter is organized as follows: In §3.1, I discuss the previous analyses that have been proposed in the literature for the constructions containing attributive *que*. In §3.2, I outline my own analysis and compare it to previous approaches. In §3.3, I report the empirical evidence behind the syntactic analysis and discuss further syntactic properties of the individual constructions. Finally, §3.4 focuses on the discourse contribution of attributive *que* and the different nuanced effects that it creates in the constructions under discussion.

3.1 Previous analyses

The principal idea of this chapter is that the different constructions exemplified in (1) all involve an instance of the complementizer merged in the same position and valued with the same feature. This is a novel insight: To date, a connection between all these constructions has not been identified in the literature; consequently, the analyses that I discuss in the present section do not account for the whole phenomenon at once but deal separately with the individual constructions. While I mention all of the most important analyses, I only provide the finer details of those that are most comparable in terms of theoretical assumptions with the analysis developed in this book. This section first discusses the accounts for *que*-initial declaratives with a non-reportative interpretation, followed by a review of how *que* in polar questions has been treated in the literature. Then I present the previous analyses of *que* in *wh*-exclamatives. Finally, I summarize the main proposal for *que* when it follows epistemic and evidential modifiers and for *que* in verum marked sentences.

3.1.1 Corr (2016)

The most extensive study of *que*-initial sentences with a non-reportative meaning is found in Corr (2016). The author distinguishes between two types: exclamative QUE and conjunctive QUE. The distinction is made on an interpretive and structural level. Exclamative QUE is illustrated in (3).

(3) a. Catalan (Corr 2016: 88: ex 9)
(Ai) que t' atrapo!
DM EXCL CL.2SG catch.1SG.PRS
'I'm coming to get you!'

b. Portuguese (Corr 2016: 88: ex 11)
Ai, quo o gato se me foi ao peixe!
DM EXCL the cat CL.REFL CL.1SG go.3SG.PRF.PST to the fish
'The cat went off after the fish!'

c. Spanish (Corr 2016: 92: ex 31)
¡Que hemos salido en la radio, oiga!
EXCL AUX.1PL.PRF.PRS go.PTCP in the radio DM
'We're on the radio, look!'

Similar data have been studied by Biezma (2008) on Spanish and Ledgeway (2012) on Romance in general. Otherwise, the use of a complementizer in these types of exclamatives has passed mostly without comment in the formal literature on Ibero-Romance, although similar phenomena found in other languages have been mentioned by some authors (cf. Sæbø 2005 on French, Schwabe 2006 on German and Delsing 2010 on Scandinavian languages).

Corr (2016) treats Ibero-Romance exclamative QUE sentences as exclamatives on the grounds that they are expressive: They convey the speaker's mental state or attitude with respect to the propositional content of the sentence, they are formally independent of interrogatives, they give rise to a degree interpretation and they are potentially also factive (on the central properties of exclamatives see also for instance Gutiérrez-Rexach 2001, Zanuttini & Portner 2003, Castroviejo 2006). To map out the position of exclamative QUE, the author relies again on her CP-external performative structure termed Utterance Phrase (UP) illustrated in (4) (repeated from (14) see §2.1 for a characterization of the UP).

(4) $[_{\text{MoodP}}\ [_{\text{SAlow}}\ [_{\text{EvalP}}\ [_{\text{EvidP}}\ [_{\text{DeclP}}\ \cdots]]]]]$

The data in (5) show that exclamative QUE follows discourse markers and vocatives.

(5) Catalan

a. (Corr 2016: 132: ex 118)
(Ai/ apa) que (*ai/ apa) em poso vermella.
DM DM EXCL DM DM CL.REFL put red
'Ohh/gosh, I've gone red!'

b. (Corr 2016: 134: ex 122)
(Amor) que (*amor) em poso vermella.
love EXCL love CL.REFL put.1SG.PRS red
'Darling, you make me blush!'

To account for the syntactic properties, Corr (2016) proposes the analysis in Figure 3.1. Exclamative QUE is assumed to be merged in $Eval^0$ where it picks up an evaluative feature and moves to SALow, in which it receives the speech act feature and takes on its performative function. EvaluativeP, according to Corr (2016), hosts certain types of performative particles like *mira* and *anda*, which express a speaker's attitude and "involve gradability: A key constitutive property of exclamatives" (Corr 2016: 136). Corr (2016) furthermore considers the fact that exclamative QUE is incompatible with the performative particles merged in this same position as evidence that they are in complementary distribution, cf. (6).[2]

(6) Portuguese (Corr 2016: 137: ex 132)
Ai, fofinha (*que) olha que estás com sorte!
DM cute.DM EXCL DM QUE be.2SG.PRS with luck
'Oh, darling, gosh aren't you lucky.'

The addition of a +SA, *speech act* (standing in for *performative* in Corr's 2016 account) feature, is justified on the basis that an exclamative QUE "performs an expression of one's attitude towards the proposition" (Corr 2016: 108).

Although all the Ibero-Romance varieties that Corr (2016) focuses on accept constructions involving exclamative QUE, the author claims that in European Portuguese they are restricted to declaratives. This is why she assumes that in this

[2]In my analysis, one possible explanation for the ungrammaticality of *que* above certain discourse markers in examples such as (6) is that we are dealing with two instances of attributive *que*. The grammatical word order in which *que* surfaces below *olha* leads to two possible conclusions: either *olha* is merged CP-externally, or, alternatively, it is merged in ModP and treated as an evaluative modifier.

language exclamative QUE is merged lower, in Decl^0 where it picks up the additional declarative feature (as in the structure in Figure 3.2).[3]

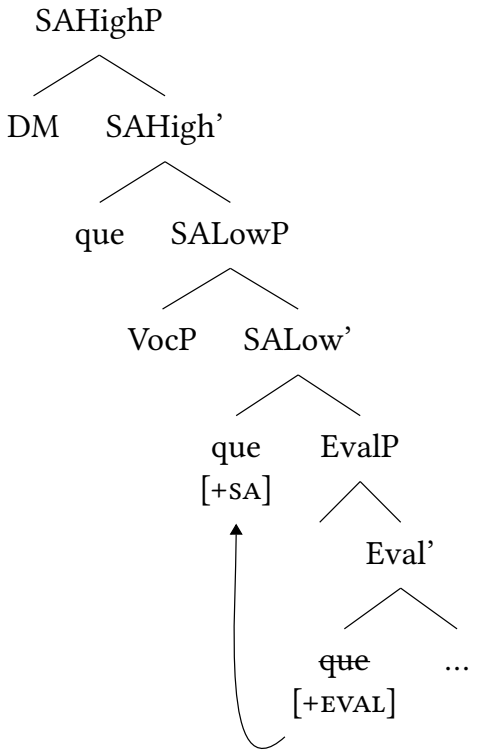


Figure 3.1: Exclamative QUE in Corr (2016: 137, ex 134)

The second type of non-reportative *que*-initial construction that Corr (2016) deals with involves a particle she calls conjunctive QUE. In the previous literature, it has been proposed that in these cases the complementizer establishes a causal (see Alarcos Llorach 1994, Porroche Ballesteros 2000, Peres & Mascarenhas 2006, Etxepare 2013, Wheeler et al. 1999, Cunha & Cintra 1984, Lobo 2003, Lopes 2012) or explicative (Colaço & Matos 2016) relation between a previous sentence and the sentence introduced by *que*. The use of an element homophonous with a complementizer as a clausal connective is also documented in the history of the Ibero-Romance languages (Martínez Marín 1978, Carrera de la Red 1982, Bartol Hernández 1988, Batllori et al. 2000, Batllori & Suñer 2005).

[3]The general assumption made in Corr (2016) is that many differences between European Portuguese and other Ibero-Romance varieties result from the fact that in the former the declarative, evidential and evaluative features are bundled on one functional head whereas they are scattered across three heads in the other varieties (see also §2.1). It is thus surprising that the author presents an analysis for exclamative QUE in European Portuguese that does not involve EvidP.

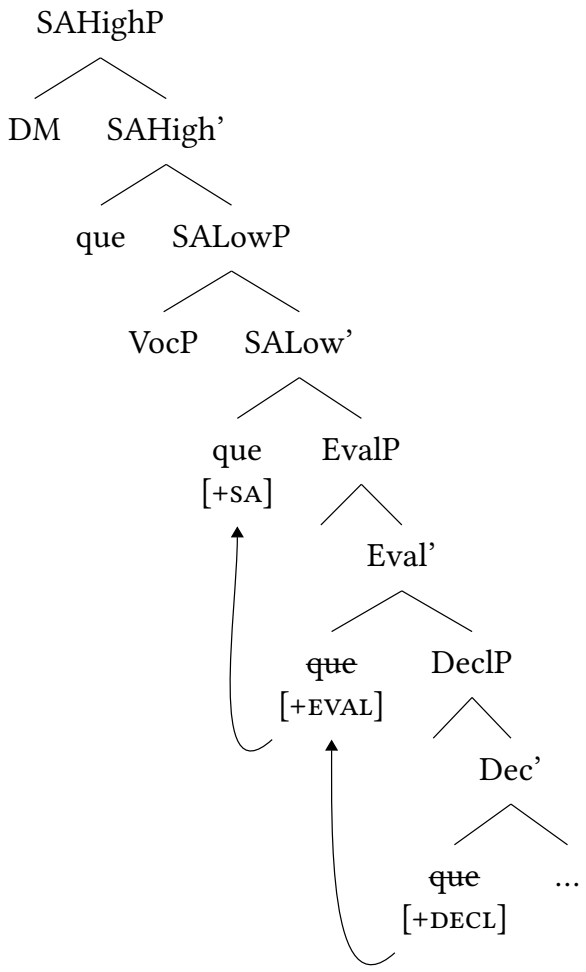


Figure 3.2: Exclamative QUE in European Portuguese in Corr (2016: 236, ex 112)

An example from contemporary Spanish is given in (7). Conjunctive QUE sentences often follow imperatives, but not always (see the example in (8)).

(7) Spanish (adapted from Corr 2016: 229: ex 90)
No me pises, que llevo chanclas.
not CL.1SG step.2SG.SBJV.PRS CONJ wear.1SG.PRS flipflop.PL
'Don't step on me, I'm wearing flipflops.'

Despite the focus on the causal function in the literature, Corr (2016) shows that conjunctive QUE is semantically not the same as a causal connective.[4] She illustrates this with the example in (8).

[4]For more arguments against a primary causal or explicative nature of these complementizers, see §3.4.

(8) Catalan (Corr 2016: 226: ex 84)
[Context: The addressee glances at some boarding passes on the speaker's desk and the speaker notices what the addressee is looking at.]

a. Que me 'n vaig de vacances.
QUE CL.REFL CL.PART go.1SG.PRS of holiday.PL
'I'm going on holiday.'

b. * Perquè/ car me' n vaig de vacances.
because for CL.REFL CL.PART go.1SG.PRS of holiday.PL

Conjunctive QUE, but not a true causal connective like *perquè/car* (cf. 8), is felicitous without a previous utterance and can serve to explain a non-linguistic situation (cf. 8a). According to Corr (2016), these two categories are distinct because conjunctive QUE introduces a syntactically independent clause, while the relevant causal connectives introduce a syntactically dependent clause. Corr (2016: 207) argues that the primary function of conjunctive QUE is not to establish a causal link but to maintain or improve the conversational flow.

Corr (2016) maps out the syntactic position against the backdrop of her UP. She uses the data in (9) to show that conjunctive QUE cannot co-occur with discourse markers or vocatives.

(9) Spanish

a. (Corr 2016: 235: ex 109)
¡Escúchame, (*oye) que (*oye) vamos a llegar tarde!
listen.IMP=CL.1SG DM CONJ DM go.1PL.PRS to arrive late
'Listen, *hey we're going to arrive late!'

b. (Corr 2016: 235: ex 110)
¡Escúchame, (*María) que (*María) vamos a llegar tarde!
listen.IMP=CL.1SG María CONJ María go.1PL.PRS to arrive late
'Listen, *María we're going to arrive late!'

Corr (2016) proposes the analysis given in Figure 3.3. Conjunctive QUE reaches SAHigh^0, the highest head in the UP. The complementizer is assumed to be merged in Decl^0, the lowest head of her split ForceP. On its way to SAHigh^0, it passes through Evid^0 and Eval^0. Conjunctive QUE is valued with the features hosted by the individual heads.[5] Corr (2016) assumes that the merger occurs in DeclP because conjunctive QUE is restricted to declaratives. The evidential feature picked up in Evid^0 is explained on the grounds that conjunctive QUE shows a

[5] As an alternative to the movement analysis, Corr (2016) also proposes a structure in which the features are valued simultaneously by a syncretic head (cf. Corr 2016: 236: ex 111).

parallel behavior to certain evidential complementizers merged in this projection. The evaluative feature in EvalP guarantees that the constructions are assertive and express a speaker's point of view. Finally the SA or performative feature ensures the performative nature of the constructions involving conjunctive QUE.

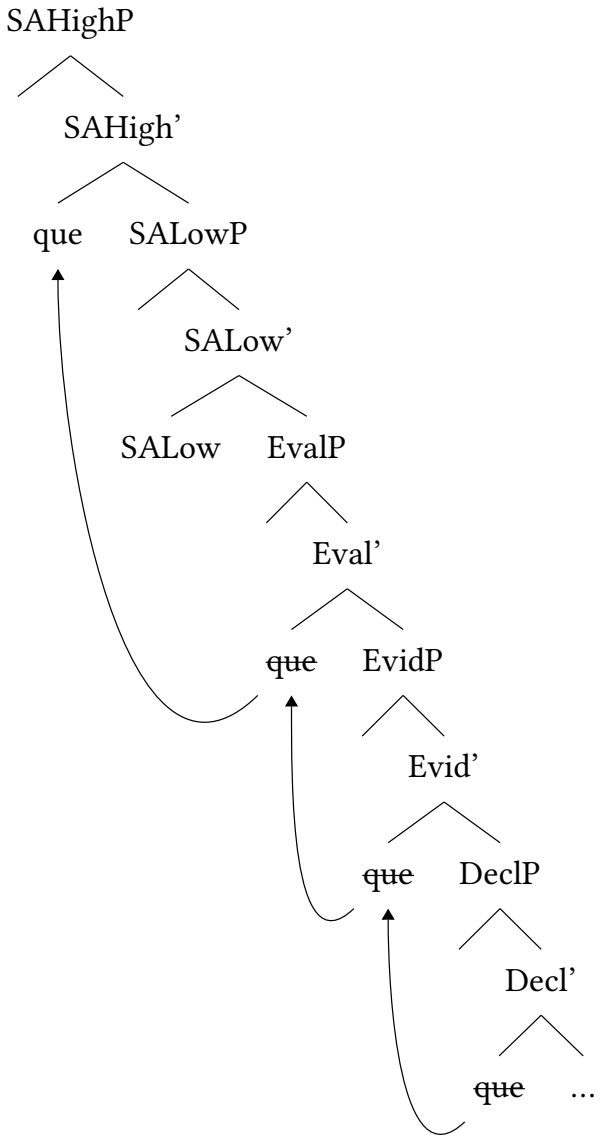


Figure 3.3: Head movement analysis of conjunctive QUE in Corr (2016: 236, ex 111)

The main issue I have with the analysis proposed in Corr (2016) is that there is no empirical motivation relating to word order restrictions to support the idea

that the complementizers reach a position in the CP-external UP. The explanation presented in Corr (2016) is based only on the theoretical assumptions made by the author when she relates certain interpretative properties of the constructions to the effects of the features in her assumed UP. This is a potential problem because, as in her analysis of reportative *que* discussed in §2.1, the question of how the abstract features give rise to the specific meaning is not answered satisfactorily. In principle, the empirical data are also compatible with an analysis that assumes that in both constructions *que* is merged in the highest CP head SubP (structurally equivalent to ForceP in the original hierarchy) and therefore below Corr's UP. The data in (5) show that exclamative QUE follows phrases like vocatives and discourse markers that are analyzed as being merged in the UP. These data are compatible with an analysis like that proposed by Corr (2016) that places the complementizer in the lowest UP head but they are also in line with my alternative analysis that locates them in the highest CP head. The data in (9) show that conjunctive QUE is incompatible with vocatives and discourse markers; the assumption that the complementizer here reaches a high UP projection lacks therefore compelling empirical support. However, the conclusion that the complementizer simply occupies the highest CP head is in line with the data. Pending further empirical evidence of movement to a CP-external projection, the position I adopt in the present monograph is that the complementizer in these contexts is merged in SubP, the top left projection of the split CP.

3.1.2 Prieto & Rigau (2007)

The syntactic behavior and the pragmatic effect of *que* in polar questions have so far been addressed almost exclusively as a feature of Catalan grammar (cf. Rigau 1984, Mascaró i Pons 1986, Cuenca 1997, Prieto 1997, 2002, Payrató 2002, Celdrán et al. 2005, Hernanz & Rigau 2006). However, I will show in §3.3.2 that it is also attested in Spanish. To the best of my knowledge, this has been widely disregarded in the literature. One exception is Hualde (1992), where it is mentioned briefly and characterized as a case of transfer from Catalan. To date, the most extensive studies of *que* in Catalan polar questions are those carried out by Rigau & Prieto (2005) and Prieto & Rigau (2007). This section summarizes the main points of their analysis.

One central finding in Prieto & Rigau (2007) is that the presence of *que* coincides with a falling question intonation. The authors furthermore suggest that there is dialectal variation with regard to the presence and absence of *que* in different pragmatic contexts. They argue that *que* is virtually unrestricted in Minorcan Catalan. All varieties accept the presence of *que* in anti-expectational

contexts, i.e. contexts in which the facts or the situation are in disagreement with the speaker's expectations. In these contexts, polar questions can be used to express the speaker's surprise or astonishment. An example is given in (10).

(10) Catalan (Prieto & Rigau 2007: 15: ex 30a)
Que vindràs a Barcelona? No em pensava pas que ens
QUE come.FUT.2SG to Barcelona not CL.1SG thought NEG that CL.1PL
acompanyessis.
accompany.SUBJ.2SG
'Are you coming to Barcelona? I didn't think you were coming with us.'

The anti-expectational nature of the context becomes evident from the statement that follows the *que*-initial polar question where the speaker explicitly asserts that the fact that the hearer is going to Barcelona was not part of her prior beliefs.

The complementizer is furthermore accepted in confirmatory questions in all dialects. According to Prieto & Rigau (2007), these are questions where the speaker expects an affirmative answer. In Catalan, confirmatory *que*-questions are often preceded by a question particle, cf. (11). According to the authors, the choice of the particle depends on the dialect.

(11) Catalan (Prieto & Rigau 2007: 17: 35 a-f)
Oi / Eh / Veritat / No / Fa / És ver que vindràs?
PARTICLE QUE come.FUT.2SG
'You're coming, aren't you?'

Castroviejo (2018) makes a compelling case that, at least in sentence-final positions like in (12), a subset of these particles are not synonymous but encode different meanings (see also §3.4).

(12) Catalan (Castroviejo 2018: ex 19)
T' has tallat els cabells, oi?/ eh?
CL.2SG AUX.3SG.PRF.PRS cut.PTCP the hair.PL OI EH
'You had your hair cut, right?/ huh?'

Other non-neutral polar questions that permit *que* are rhetorical questions, as in (13).[6]

[6] Although this is not central to the present discussion, in Kocher (2017b) I propose that questions like (13) are better characterized as hyperbolic rather than rhetorical questions. Hyperbolic and rhetorical questions both stand in for another utterance. However, a true rhetorical question stands in for an assertion, while these hyperbolic questions actually stand in for another polar question as an exaggerated version of it.

(13) Catalan (Prieto & Rigau 2007: 18: 36b)
Que et penses que tinc quatre mans, jo?
QUE you think.2SG.PRS that have.1SG.PRS four hand.PL I
'Do you think I have four hands?'

Prieto & Rigau (2007) state that in Northwestern, Central and Balearic Catalan, *que* is furthermore allowed in what they term polite polar questions. They are considered polite by the authors because the speaker uses them when they only require a low cost action by the hearer. A low cost action always implies that the speaker was certain that the hearer would answer positively.

(14) Catalan (Prieto & Rigau 2007: 4: ex 8a)
Que em deixes el teu apartament de la platja, aquest
QUE me leave.2SG.PRS the your apartment of the beach this
cap de setmana?
weekend
'Would you let me use your apartment by the beach this weekend?'

The use of *que* in (14) is felicitous only if the hearer has offered the apartment to the speaker previously. In a context where this is not the case, *que*-initial polar questions are not felicitous according to Prieto & Rigau (2007).

(15) Catalan (Prieto & Rigau 2007: 4: ex 9a)
Que puc fumar?
QUE can.1SG.PRS smoke
'Can I smoke?'

Similarly, (15) is only felicitous if the speaker can assume that her smoking is not going to bother the hearer but is not felicitous if she expects it will or if she has no expectations in this regard.

Prieto & Rigau (2007) adopt a cartographic approach and assume that *que* is merged at the lowest edge of the left periphery in FinP. They propose the structures in (16).

(16) Catalan
a. (Prieto & Rigau 2007: 25: ex 56a)
[$_{\text{ForceP}}$ [$_{\text{Operator}}$ Oi] [$_{\text{Force +confirmative interrogative}}$ [$_{\text{FinP}}$ que [$_{\text{IP}}$ en
OI QUE the
Pere no va a Barcelona?]]]]
Pere not go.2SG.PRS to Barcelona
'Pere isn't going to Barcelona, right?'

b. (Prieto & Rigau 2007: 25: ex 56b)
[$_{\text{ForceP}}$ [Operator] [$_{\text{Force}}$ +anti-expect./neutral interrogative [$_{\text{FinP}}$ que [$_{\text{IP}}$ no
QUE not
volies un collaret?]]]]
want.2SG.IPVF.PST a necklace
'Didn't you want a necklace?'

In their analysis, the non-neutral interpretation of *que*-initial polar questions is attributed to the presence of an interrogative operator in Force that is realized by prosodic means. *Que* is considered optional and does not contribute any meaning of its own.

I will now turn to my evaluation of this account. The prediction of the structural analysis proposed in Prieto & Rigau (2007) is that, given the low position of *que*, other left-peripheral material should precede rather than follow the complementizer.

(17) Catalan (Kocher 2017b: 49: ex 98a)
[Context: Marta finds a bag of oranges in the kitchen. She asks her roommate:]
Que les taronges$_j$ les$_j$ vas comprar tu?
QUE the orangePL CL.F.PL AUX.2SG.PRF.PST buy you
'The oranges, you bought them, didn't you?'

The data in (17) pose a problem for the analysis proposed in (16b) because in (17), a clitic left dislocated topic follows rather than precedes *que*. Furthermore, based on their analysis for particle questions in (16a), we would expect that certain phrases should be able to intervene between the particle and the complementizer. However, the data in (18) show that the particle and the complementizer must be adjacent (more data are discussed in §3.3.2).

(18) Catalan
* Oi en Jordi que l' has convidat tu?
OI the Jordi QUE CL.M.SG have.AUX.2SG.PRF.PRS invite.PTCP tu
Intended: 'You are the one that invited Jordi, right?'

To reconcile these data I will propose a revised analysis in §3.3.2. I adopt the idea from Prieto & Rigau (2007) that *que* is merged in FinP but in my analysis it does not remain in this position; instead, it moves through the left periphery and ends up in the head of the highest projection of the left periphery.

Turning to pragmatic considerations, Prieto & Rigau (2007) offer a detailed characterization of different contexts that license *que* in polar questions. Based on this characterization, I propose a generalization that, as will be seen in §3.4, ultimately makes it possible to assume that *que* has a uniform discourse contribution in all the different constructions. What all the contexts that license *que* in the majority of dialects have in common is that the speaker expects a positive answer from the hearer. In an anti-expectational context like (10), while the speaker's belief was the opposite, contextual evidence suggests that the answer is going to be positive (see Kocher 2017b and §3.4 for a revised definition of the notion of contextual evidence presented in Büring & Gunlogson 2000). In confirmatory contexts like (11), the speaker's belief itself makes her expect a positive answer. The polite polar questions in (14) and (15) can also be subsumed readily under the label of confirmatory questions because, as Prieto & Rigau (2007) state, *que* is only felicitous when the speaker has a hunch that the answer is going to be affirmative. Finally, rhetorical questions like (13) are more challenging because they do not have the illocutionary force of a question, hence the speaker does not necessarily expect an answer. However, I believe they can be accounted for if they are treated in the way that I propose for *que*-initial assertions, in the sense that a commitment to them is attributed to the hearer (cf. §3.4).[7]

3.1.3 Ambar (2003), Castroviejo (2006) and Demonte & Fernández Soriano (2009)

The analyses of *wh*-exclamatives involving *que* that I will discuss here are proposed by Demonte & Fernández Soriano (2009) for Spanish, Ambar (2003) for Portuguese and Castroviejo (2006) for Catalan (but cf. also Bosque 1984, Brucart 1993, Villalba 2008 and Gutiérrez-Rexach 2001). Relevant accounts of *wh*-exclamatives in other languages are put forward in Milner (1978), Radford (1982), Benincà (1996), Zanuttini & Portner (2003), Cruschina (2015), among others.

Castroviejo (2006) investigates *wh*-exclamatives with *que* in Catalan along with other types of exclamatives. In her proposal, the complementizer is characterized as semantically vacuous and its presence is deemed optional. Castroviejo (2006) does not adopt a split CP. She proposes the analysis in (19). The *wh*-phrase is merged vP internally and moves through the specifier of the TP to the specifier of the only CP projection in the structure. The complementizer is realized as the head of the same CP.

[7] Kocher (2017b) offers a more extensive discussion of Catalan biased polar questions. There, I propose a slightly different generalization that relies on the typology of question biases in Sudo (2013). My basic idea was that the presence of *que* is licensed when there is positive evidence in the context.

(19) Catalan (Castroviejo 2006: 50: ex 123b)
[$_{CP}$ [$_{SpecCP}$ Quins ingredients tan bons$_i$] [$_{C'}$ [$_{C^0}$ que] [$_{TP}$ t$_i$
which ingredient.PL so good QUE
té$_j$ [$_{vP}$ aquesta sopa t$_j$ t$_i$.]]]]
have.2SG.PRS this soup
'What great ingredients this soup has!'

Ambar (2003) offers an account of Portuguese *wh*-exclamatives. She adopts a structured left periphery; however, her proposed structure differs from the (minimally revised) structure based on Rizzi (1997) that I adopt here. XP is conceived of as a landing site for dislocated elements and can be considered parallel to a Rizzian TopP. WhP is an operator projection hosting *wh*-phrases. AssertiveP is projected when assertive properties are involved in the constructions and is also linked to a factive interpretation. EvaluativeP encodes the speaker's evaluation and hosts phrases that contain evaluative elements.

(20) [$_{XP}$ [$_{EvaluativeP}$ [$_{Evaluative'}$ [$_{AssertiveP}$ [$_{Assertive'}$ [$_{XP}$ [$_{WhP}$[$_{Wh'}$ [$_{FocusP}$ [$_{Focus'}$ [$_{XP}$ [$_{IP}$
(Ambar 2003: 211: ex 1)

The analysis that Ambar (2003) assumes for *wh*-exclamatives involves AssertiveP and EvaluativeP. The factive interpretation of *wh*-exclamatives (see §3.4) is attributed to the feature in AssertiveP that can be checked either by the *wh*-expression or, when present, by *que*. In both cases, the *wh*-expression is *wh*-moved from an IP-internal position passing through WhP, FocusP and in *que*-less *wh*-exclamatives also through AssertiveP. It ends up in EvaluativeP where it checks an evaluative feature.

(21) Portuguese (Ambar 2003: 238–239: ex 88–89)
[$_{EvaluativeP}$ que livro$_i$ [$_{Evaluative'}$ [$_{AssertiveP}$ [$_{Assertive'}$ que/t$_i$ [$_{XP}$ o João$_j$
what book QUE the João
[$_{WhP}$ t$_i$ [$_{Wh'}$ [$_{FocusP}$ t$_i$ [$_{Focus'}$ [$_{XP}$ [$_{IP}$ t$_j$ leu t$_i$]]]]]]]]]]
read
'What a book (that) John read!'

Demonte & Fernández Soriano (2009) propose an account of Spanish *wh*-exclamatives that relies on a split CP à la Rizzi (1997). The presence of *que* is considered to be optional. In their analysis, *que* is merged in FinP and the *wh*-expression in FocP (cf. 22).

(22) Spanish (Demonte & Fernández Soriano 2009: 33 : ex 19a, analysis added by the author)
[$_{FocP}$ ¡Qué rico] ... [$_{FinP}$ (que)] [$_{IP}$ está!]
how good QUE be.3SG.PRS
'How good this is!'

All these analyses are based on different theoretical assumptions, making it difficult to compare them. Only Demonte & Fernández Soriano (2009) assume a Rizzian style structure of the left periphery that makes it comparable to my own proposal. One potential issue with their analysis, however, is that it cannot account for the data in (23).

(23) Spanish
a. * Qué raro a Juana$_i$ que la$_i$ has invitado pero no a María.
how strange DOM Juana QUE CL.FS have.2S invited but not DOM María
b. A Juana$_i$ qué raro que la$_i$ has invitado pero no a María.
DOM Juana how strange QUE CL.F.SG AUX.2SG.PRF.PRS invite.PTCP but not DOM María
c. Qué raro que a Juana$_i$ la$_i$ has invitado pero no a María.
how strange QUE DOM Juana CL.F.SG AUX.2SG.PRF.PRS invite.PTCP but not DOM María
'How strange that you invited Juana but not Maria.'

The example in (23a) shows that a dislocated topic cannot intervene between the *wh*-expression and the complementizer, even though the account in Demonte & Fernández Soriano (2009) predicts that it should be able to. (23b) shows that the *wh*-expression can be preceded by a dislocated topic, which is in line with their analysis. However, a dislocated topic can also follow *que*, which is again not predicted by the analysis because the complementizer occupies the lowest position in the left periphery. In my revised analysis, which is similar to that proposed in Demonte & Fernández Soriano (2009), the problematic data are accounted for by assuming that the complementizer moves from its initial merge position in FinP through the left periphery and ends up adjacent to the *wh*-expression in the head of FocP.

3.1.4 Cruschina & Remberger (2017a, 2018)

The Ibero-Romance construction whereby a complementizer follows an epistemic, evidential or to a lesser extent, evaluative modifier has been mentioned by various authors in the literature (Martín Zorraquino 1998, Etxepare 1997, Hummel 2000, 2014, 2017, Gutiérrez-Rexach 2001, Freites Barros 2006, Hernanz & Rigau 2006, Ocampo 2006, Rodríguez Ramalle 2007, 2008, 2015, Gras 2010, Sansiñena et al. 2015). Apart from my own analysis (Kocher 2014, 2017a), the previous more extensive accounts of the construction focus primarily on other Romance languages and not on the three Ibero-Romance varieties under discussion here. Especially influential is the analysis by Hill (2007a) (see also the review of this work in Lupşa 2011) of this construction in Romanian that directly inspired the analysis in Cruschina (2013) for Italian and Sicilian. The analysis in Cruschina & Remberger (2017a, 2018) is also applied to the construction in Spanish along with other Romance languages. What these latter analyses have in common is that they adopt the neo-performative hypothesis by Speas & Tenny (2003) which postulates a functional field above the split CP mediating the interface between syntax and discourse (see §1.4). All the accounts assume that the complementizer is merged in the highest projection of the split CP, i.e. Force, which is structurally equivalent to my SubP. They assume that the modifier is located in a CP-external projection. In the structure in Figure 3.4, for example, I illustrate the analysis proposed in Cruschina & Remberger (2018), based on Speas & Tenny (2003).

One potential issue with an analysis along these lines are data such as those in (24).

(24) a. European Portuguese
Disse que certamente que iria ver logo
say.3SG.IPFV.PST that certainly QUE go.1SG.COND see soon
resultados.
results
'S/he said that certainly I would see results soon.' (CdP)

b. Spanish
Otra canción que claro que escuchamos todos y que
other song that claro QUE listen-to.1PL.PRS all and that
podría parecer muy buena, es "Realmente no estoy
can.3SG.COND seem very good be.3SG.PRS Realmente no estoy
tan solo".
tan solo
'Another song that clearly we all listened to and that could seem very good is "Realmente no estoy tan solo".' (CdE)

c. Catalan
I per això us hem preparat un article que
and therefore CL.2PL AUX.1PL.PRF.PRS prepare.PTCP an article that
segur que us serà útil un moment o altre.
sure QUE CL.2PL will-be useful one moment or other
'And therefore we have prepared an article for you that surely will be useful for you at some point or another.' (caWac)

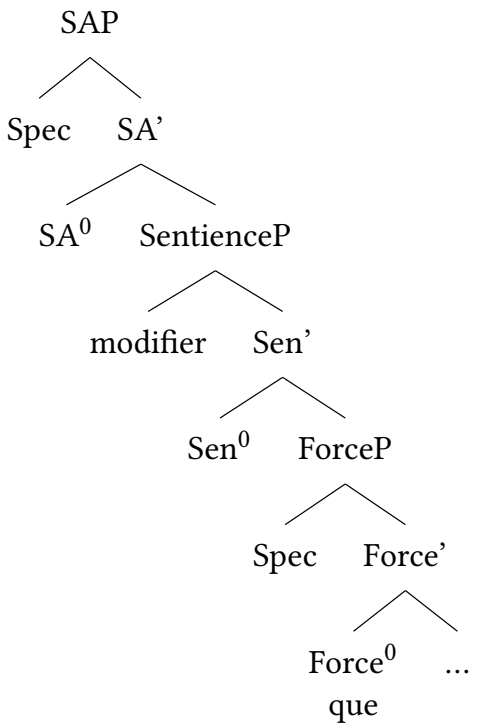


Figure 3.4: AdvC in Cruschina & Remberger (2018: 350: ex 23) (adapted)

These examples show that the construction is not restricted to root contexts but can also appear in embedded contexts, and in particular also in relative sentences (cf. an appositive relative in (24b) and a restrictive relative in (24c)). In this respect, the Ibero-Romance languages under investigation here appear to contrast with Italian, which according to Cruschina & Remberger (2018) only permits the construction in complements of verbs of saying like in (24a) but not in relatives. The analysis I propose for the construction assumes a surface position within the split CP and is therefore able to account for the data in (24).

3.1.5 Hernanz (2007)

The construction in which a complementizer follows the verum marker *sí* is attested in Spanish and Catalan but not in Portuguese. It has been explored in the literature, notably in Martins (2006, 2013), González Rodríguez (2008, 2009, 2016) Escandell-Vidal & Leonetti (2009b), Escandell-Vidal & Leonetti (2009a), Escandell-Vidal (2011), Rodríguez Molina (2014) and Villa-García & Rodríguez (2020a,b). The most widely adopted analysis of the verum construction in Spanish is developed in Hernanz (2007) (see also Batllori & Hernanz 2008 where the analysis is applied to diachronic data). The analysis has also been extended to Catalan in Batllori & Hernanz (2013).

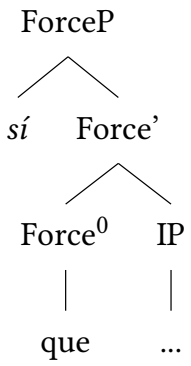


Figure 3.5: AffC in Hernanz (2007: 144: ex 87) (adapted)

Hernanz (2007) proposes the structure in Figure 3.5. The complementizer is merged as the head of ForceP and *sí* is assumed to be in its specifier.

Verum sentences do not require the presence of *que* in either language. Hernanz (2007) compares *sí que*-sentences, as in (25a), with their *que*-less equivalents, as in (25b).

(25) Spanish

a. (Hernanz 2007: 134: ex 3a)
Sí que ha llovido hoy.
yes that AUX.3SG.PRF.PRS rain.PTCP today
'It HAS indeed rained today.'

b. (Hernanz 2007: 134: ex 1a)
Sí ha llovido hoy.
yes AUX.3SG.PRF.PRS rain.PTCP today
'It HAS rained today.'

She notes that they differ in that the version with the complementizer emphasizes a proposition that has already been mentioned in the discourse. She adopts an idea from Etxepare (1997) and states that a proposition introduced by the verum marker requires a linguistic antecedent. Hernanz (2007) suggests that this function should be attributed to ForceP. Since this aspect of meaning, according to the author, does not occur when *que* is absent, she adopts a different syntactic analysis for *que*-less verum sentences, which is exemplified in (26).

(26) (adapted from Hernanz 2007: 129: ex 48)
[$_{\text{ForceP}}$ [$_{\text{TopicP}}$ [$_{\text{FocusP}}$ sí$_i$ [$_{\text{PolP}}$ t$_i$ [$_{\text{IP}}$ ...]]]]]

In this analysis *sí* starts out in a polarity position termed PolP that is sandwiched between FinP and IP. It is the same polarity position in which the sentence negation particle *no* is located (see Laka 1990). Hernanz (2007) argues that the *sí* in these contexts has focal properties, which is the reason behind her assumption that it moves to the left-peripheral FocP.

There are some empirical data that cannot be accounted for straightforwardly by the analysis proposed in Hernanz (2007), which ultimately leads me to argue in favor of a different account (see §3.3.5 and Kocher 2017a for a more detailed discussion). Hernanz's analysis for *sí que*-sentences fails to account for data such as in (27) which show that *sí que* is also attested in embedded sentences, see (27a) and (27b), and can be preceded by a clitic left dislocated topic, see (27c).

(27) a. Spanish (Kocher 2017a: 94: ex 32b from caWac)
En el bar de la Confederació General del Treball (CGT)
in the bar of the Confederació General del Treball (CGT)
confirman que sí que hay huelga.
confirm.3PL.PRS that VERUM QUE there.be.3SG.PRS strike
'In the bar of the CGT they confirm that there is a strike going on.'

b. Catalan (Kocher 2017b: 45: ex 88b from caWac)
A banda d'aquest dissentiment inicial, hi ha dos
at side of this disagreement initial there.be.3SG.PRS two party.PL
partits de l'oposició que sí que la
of the opposition that VERUM QUE it vote.3PL.FUT
votaran afirmativament.
affirmative
'Concerning this initial disagreement there are two parties of the opposition that WILL vote in favour of it.'

c. Spanish (Kocher 2017b: 94: ex 32a from caWac)
A López$_i$ sí que le$_i$ he visto agredir a
DOM López VERUM QUE CL.M.SG AUX.1SG.PRF.PRS see.PTCP attack at
dos de mis jugadores.
two of my player.PL
'I HAVE seen López attack two of my players.'

Finally, there is also a piece of data that is problematic for Hernanz's analysis for the *que*-less *sí*-sentences. Given that they are assumed to start out in PolP, examples like (28) are unexpected.[8]

(28) Spanish (Kocher 2017a: 94: 31a from CdE)
Eso sí no podía faltar en ninguna casa.
this VERUM no can.3SG.IPFV.PST miss in any house
'This could NOT be missing in any house.'

The example shows that *sí* can co-occur with the sentence negation particle *no*. This particle is assumed to occupy PolP, the same position in which Hernanz (2007) assumes *sí* to be originally merged. The alternative analysis I adopt and will be presented in §3.3.5 can account for these data. *Sí* always occupies the same left-peripheral position. The only difference between the two alternative means of expressing verum lies in the presence of *que*.

3.2 Outline of the present analysis

In this section I present my analysis that uniformly accounts for all the different constructions subsumed under the label of attributive *que*. This book strives to present a uniform analysis that is compositional on a structural and interpretive level whenever possible. In very simple terms, what I mean by this is that my goal is to develop an account in which, unless there is empirical evidence to the contrary, each element involved in the construction is merged where it always is and does whatever it always does.

The syntactic position of the complementizer in the relevant constructions will again be mapped out in the revised version of the split CP in (29) (repeated from (19)) that I assume for the present investigation. To repeat, the main differences between the split CP used here and the one developed by Rizzi are that I adopt

[8]These data contradict Hernanz (2007: 139: fn 8), where it is claimed that in these configurations *sí* cannot co-occur with negative particles.

SubP instead of ForceP and assume a lower projection termed MoodP which takes over the functions originally associated with ForceP (cf. §1.2).

(29) [SubP [TopP* [IntP [TopP* [FocP [ModP* [TopP* [MoodP [TopP* [FinP [IP]]]]]]]]]]

The point of departure of my analysis is the idea that in all the relevant constructions the presence of *que* has the same pragmatic impact (cf. §3.4). However, the word order suggests that the complementizer does not occupy the same position in all the constructions (see §3.3.1 to §3.3.5). One possible way of accounting for these facts, which will not be adopted here, would be to assume that a lexical item with a dedicated function is inserted directly into the different positions. This option could account for the shared meaning and explain why the complementizer surfaces in different positions, but it comes at the cost of potential overgeneralization, and additional motivations would be required to explain why the complementizer surfaces in these exact positions and in combination with these exact expressions.

I argue in this chapter for an alternative explanation, whereby attributive *que* is always merged in the same position in which it is valued with the interface feature responsible for its meaning. This explanation is in line with the general idea defended in this book that there is only one lexical item *que* whose meaning is determined by the syntactic position in which it is externally merged. This facilitates a unified account of the constructions involving attributive *que* and the *que*-initial reportatives discussed in Chapter 2.

The position hosting the attributive feature is at the right edge of the left periphery. This feature attracts the complementizer to its head. It does not remain in FinP but reaches its final surface position through head-to-head movement. It will be shown over the course of the following sections that this movement is restricted by a syntactic condition: The complementizer cannot cross a phrase that is externally merged in the left periphery. The structure in Figure 3.6 illustrates the basic ideas. In what follows, I briefly present theoretical support for each of the basic assumptions in the analysis. The detailed empirical basis is given in §3.3 and §3.4.

The assumption that the complementizer starts out in the lowest projection of the left periphery, FinP, has its origin in independent observations in the literature that this projection hosts finite complementizers (cf. for instance Belletti 2009, 2013, Ledgeway 2005). Moreover, some Romance dialects have morphologically distinct forms to express low and high merged complementizers (see e.g. Ledgeway 2005 on Southern Italian dialects and D'Alessandro & Di Felice 2015

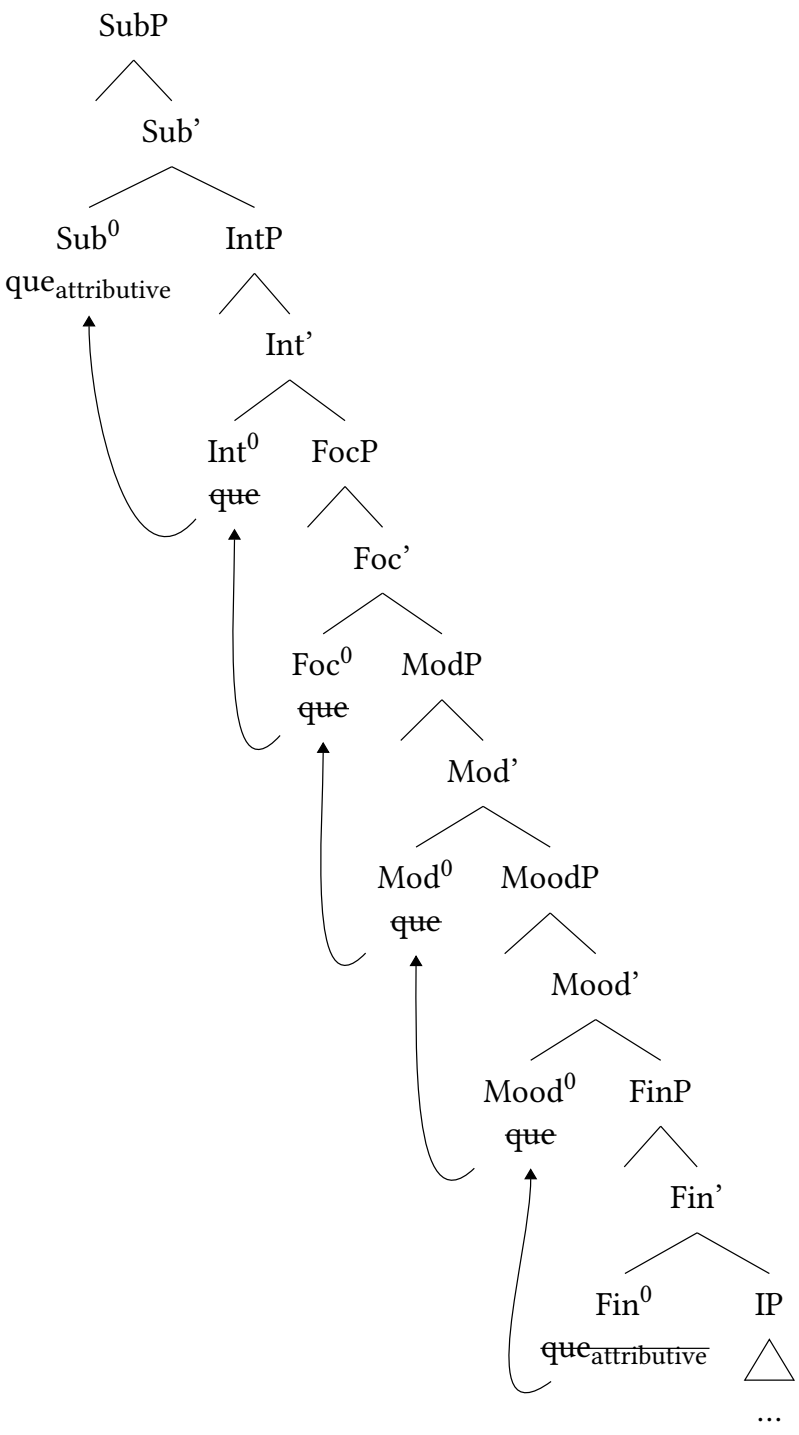


Figure 3.6: Complementizer movement

on Abruzzese). The data from the Corsican variety reported in Ledgeway (2012) are particularly interesting for the present investigation because sentences containing the different morphological forms have different readings. *Chì*, the complementizer merged in ForceP, introduces a declarative. An example is given in (30a) where it heads a sentence embedded under a verb of saying. In contrast, *chè* is merged in FinP and according to Ledgeway (2012) the proposition that it introduces receives an exclamative reading. At present, I am unable to determine whether what the author calls an exclamative reading is similar to the interpretation that is identified for low *que* in Ibero-Romance. An interesting parallel is that they are also able to introduce main clauses, illustrated in (30b).

(30) Corsican

a. (Ledgeway 2012: 175: ex i.a)
Dì a Caccara chì, à ott' ore sì Diu vole,
tell.IMP.2SG to Grandma that at eight hour.PL if God want.3SG.PRS
saremu in casa
we.shall.be.1PL in house
'Tell Grandma that, God willing, we shall be home by eight o'clock'

b. (Ledgeway 2012: 175: ex i.b)
Chè vo un caschete!
that you not fall.3PL.PRS
'Watch you don't fall.'

The empirical distribution of the Ibero-Romance complementizer in the relevant constructions (cf. §3.3) shows that it always occupies positions above FinP, including the position immediately above it. These facts are explained by assuming that the complementizer is base-generated below the lowest head that it surfaces in, namely FinP. In this position, the complementizer is valued with an interface feature that has an impact on the interpretation of the sentence in its scope. I propose a feature that I call *attributive*. My analysis is also inspired by de Cuba & MacDonald (2013). The authors show that complement clauses of factive and non-factive verbs have a different interpretation and also a different structure. Another analysis that assumes structural differences in the CP between different types of complement clauses was proposed before de Cuba & MacDonald (2013) by Haegeman (2004, 2006). She focuses on the contrasts between (31) and (32).

(31) * I haven't seen Mary since she probably left her job.
(Haegeman 2006: 1653: ex 2b)

(32) I won't be seeing Mary, since she probably will be leaving early today.
(Haegeman 2006: 1653: ex 3b)

Haegeman (2006) observes word-order restrictions in the left periphery of these sentences, illustrated here by the ungrammaticality of *probably*. She adopts a cartographic framework and proposes that adverbial clauses like (31) are integrated at the IP-level and have a more reduced structure than adverbial clauses like (32), which she considers to be adjoined to the host clause at a later stage in the derivation. The latter types start with a SubP and contain essentially the full set of projections that are also found in root clauses; the former type, integrated at the IP-level, only project a Sub and a Fin head.

De Cuba & MacDonald (2013) make empirical observations similar to those in Haegeman (2004, 2006) for the contrast between factive and non-factive complement clauses.

(33) a. (*non-referential*)
John thinks that this book Mary read.
(de Cuba & MacDonald 2013: 8: ex 10a)

b. (*referential*)
* John regrets that this book Mary read.
(de Cuba & MacDonald 2013: 8: ex 9a)

The proposal by de Cuba & MacDonald (2013) builds on a contrast in the interpretation of the two types of complement clauses. The complement clause of a non-factive verb like *think* is interpreted as a new proposition. In contrast, the complement clause of the factive verb *regret* is presented as part of the common ground. A property that follows from this contrast is that the propositional content of the complement of a factive verb remains constant under negation of the matrix verb, while the complement of a non-factive verb does not. In the negated version of (grammatical) (33b) *John doesn't regret that Mary read this book.*, it is still presupposed that Mary read this book. In the negated version of (33a), *John doesn't think that Mary read this book.*, the fact that Mary read this book is not presupposed.

According to de Cuba & MacDonald (2013), the two types of complement sentences differ not only in their interpretation but also in their structure. The examples in (33) show that while non-factive sentences admit a left dislocated topicalization of the object *this book* (cf. 33a), the same dislocation is ungrammatical in factive complements (cf. 33b). The authors draw the conclusion that the CP of a factive complement is smaller. In de Cuba & MacDonald (2013), Spanish examples are also discussed in order to demonstrate that the difference in size also

holds for this language. While a clitic left dislocated topic is grammatical in a non-factive complement clause (cf. 34a), it is ungrammatical in a factive complement clause (cf. 34b).

(34) a. Spanish (*non-referential*) (de Cuba & MacDonald 2013: 9–10: ex 11c)
Juan cree que ese libro$_i$ ya se lo$_i$
Juan believe.3SG.PRS that that book already CL.REFL CL.AKK
había leído.
AUX.3SG.IPFV.PST read.PTCP
'Juan believed that that book he had already read.'

b. Spanish (*referential*) (de Cuba & MacDonald 2013: 10: ex 12a)
* Sabía a Juan$_i$ qué le$_i$ había
know.3SG.IPFV.PST to Juan what him AUX.3SG.IPFV.PST
prometido el decano.
promise.PTCP the dean
'I knew what the dean had promised John.'

The critical example of a complement sentence under the factive verb *saber* 'to know' in (34b) does not contain a complementizer in standard Spanish. Some dialects, however, do permit the co-occurrence of a *wh*-pronoun and a complementizer below factive verbs as in (35), in which case the complementizer follows *qué*. This is an indication that the complementizer must be located lower than FocP.

(35) Peruvian Spanish
¿Sabes qué que le dice una madre a su hijo
know.2SG.PRS what QUE him tell.3SG.PRS a mother to her son
informático?
computer scientist
'You know what (that) a mother tells her computer scientist son?' (CdE)

De Cuba & MacDonald (2013) use the terms *referential* vs. *non-referential* to distinguish between the two types of complement sentences. The logic behind these terms is that a sentence introduced by a factive verb is considered by the authors to be referential in the sense that there is a sentence in the common ground that it refers to. The common ground can be understood as the set of propositions to which all the speech participants are committed. This means that by marking a proposition as part of the common ground, the speaker also claims that the hearer shares this commitment. In the constructions that are at the center

of the present chapter, it is precisely the attributive commitment to the hearer that is highlighted (see §3.4). I therefore choose to call the feature *attributive* rather than *referential* to reflect this fact. De Cuba & MacDonald (2013) do not adopt a cartographic approach. The larger structure of non-referential sentences is modeled by assuming that the CP is selected by a small cP, paralleling the vP-shell analysis (cf. Chomsky 1955, Larson 1988, 1990).

Villa-García (2015) proposes a cartographic adaptation of the insights presented in de Cuba & MacDonald (2013) that is reminiscent of the account provided in Haegeman (2004, 2006) for adverbial complements. He suggests that the difference in size is represented in the number (and nature) of functional heads projected in each type of complement clause. In his view, the left periphery of a factive complement clause consists only of a FocP and a FinP while a non-factive complement clause has an additional ForceP and a TopicP.

Table 3.1: Properties of referential and non-referential complements in de Cuba & MacDonald (2013) and Villa-García (2015)

	referential (~ factive)	non-referential (~ non factive)
interpretation	new proposition	common grounded
de Cuba & MacDonald (2013)	[cP $-ref$ [CP]]	[CP $+ref$]
Villa-García (2015)	[ForceP [TopicP [FocusP [FinP]]]]	[FocusP [FinP]]

De Cuba & MacDonald's goal is to capture the contrast in embedded sentences, but I propose that their insights can also be adapted to account for unembedded sentences, which are the main focus of this book. The novel aspect of my approach is therefore that the property of referentiality, understood in the sense of de Cuba & MacDonald (2013), is not restricted to embedded sentences but also plays a role in the interpretation of unembedded sentences. I furthermore suggest that the feature responsible for the interpretation is anchored within the cartographic structure. This means that *attributive* is not a feature linked to an entire CP, as *referential* is in de Cuba & MacDonald (2013), but a feature hosted by a single functional head that is part of the split CP. My proposal, in line with the discussion above, is that the feature is hosted in the lowest head of the left periphery, FinP. The presence of the feature requires the merger of a complementizer. As a consequence, the sentence in the scope of the complementizer is interpreted as a non-discourse new proposition. If the feature is not present,

no complementizer is merged in FinP and the sentence does not receive a non-discourse new interpretation. The exact pragmatic contribution of attributive *que* will be discussed in §3.4.

Another assumption that I make in my analysis is that the complementizer does not remain in FinP but moves through the left periphery. This idea is mainly based on the empirical facts given in §3.3, which show that the complementizer surfaces at different positions in the different constructions. It is by no means unprecedented, however, for a complementizer to move: Rizzi (1997), Poletto (2000), Roberts (2001), Ledgeway (2005) and Belletti (2009, 2013) all independently assume the merger of a complementizer in a lower CP position and movement to a higher one. In order to explain this movement, I adopt the idea presented in Belletti (2009, 2013), who states that complementizer movement always obtains in languages where the same C-element realizes the content of Fin and Force (equivalent to Sub in my terminology), which is the case in all three languages under investigation. How this insight should be modeled, perhaps via a feature-driven conception of movement, is left open for future research.

As a consequence of the complementizer movement, my analysis predicts that *que* can surface in any of the split CP heads. The idea that a complementizer can occupy positions other than ForceP or FinP is implicit in the proposals by Haegeman (2004, 2006) and Villa-García (2015) and finds further support in a number of different works (cf. for instance Roussou 2000, 2010, Gutiérrez-Rexach 2001, Brovetto 2002, Rodríguez Ramalle 2003, Demonte & Fernández Soriano 2009, Ledgeway 2005, Villa-García 2012a,b, 2015, Corr 2016).

The complementizer-movement analysis I propose predicts that *que* can in principle move all the way up to the left edge of the functional field, SubP.[9] However, the word order observed in the different constructions shows that the complementizer follows rather than precedes certain left-peripheral material, which would not be expected if attributive *que* always reached SubP. Furthermore, the data that will be provided in §3.3 show that the complementizer surfaces at different heights in the different constructions. This suggests that the movement is conditioned in some way. In light of the empirical facts, my generalization is that the movement is inhibited by base-generated material. In other words, a complementizer cannot move on to the next projection if the specifier of the projection that currently hosts the complementizer is occupied by an externally-merged phrase. As a consequence, I assume that all the elements that surface immediately above attributive *que* are externally merged in-situ. The reasoning behind

[9] In Corr (2016) it is suggested that the complementizer can even reach a position in her performative field above the CP.

this idea is given in §3.3. The following examples show the analyses I assume for the different constructions.

For attributive *que* in declaratives (36a) and polar questions (36b), since there is no material hindering it, the movement of complementizer is unrestricted and consequently *que* does in fact reach the highest projection SubP.

(36) a. Spanish
Tranquilo, tío. [$_{\text{SubP}}$ Que$_{\text{attributive}}$] …[$_{\text{FinP}}$ ~~que$_{\text{attributive}}$~~] [$_{\text{IP}}$ no
calm man QUE QUE not
muerdo.]
bite.1SG.PRS
'Chill, man. I don't bite.'

b. [$_{\text{SubP}}$ Que$_{\text{attributive}}$] …[$_{\text{FinP}}$ ~~que$_{\text{attributive}}$~~] [$_{\text{IP}}$ no muerdes?]
QUE QUE not bite.2SG.PRS
'You don't bite?'

In *wh*-exclamatives, the *wh*-expression is analyzed as being merged directly in FocP and the complementizer movement comes to a halt at the head of this phrase (cf. 37).

(37) Portuguese
[$_{\text{FocP}}$ Que chato que$_{\text{attributive}}$] …[$_{\text{FinP}}$ ~~que$_{\text{attributive}}$~~] [$_{\text{IP}}$ é.]
how annoying QUE QUE be.3SG.PRS
'How annoying this is.'

Similarly, epistemic and evidential modifiers in AdvC as well as the verum particle *sí* in AffC are assumed to be merged directly in the left-peripheral position in which they appear in the surface structure. The analyses I assume for these constructions are given in (38a) and (38b).

(38) a. Catalan
[$_{\text{ModP}}$ Segur que$_{\text{attributive}}$] …[$_{\text{FinP}}$ ~~que$_{\text{attributive}}$~~] [$_{\text{IP}}$ son
sure QUE QUE be.3PL.PRS
amics.]
friend.PL
'Surely, they are friends.'

b. [$_{\text{MoodP}}$ Sí que$_{\text{attributive}}$] …[$_{\text{FinP}}$ ~~que$_{\text{attributive}}$~~] [$_{\text{IP}}$ son
VERUM QUE QUE be.3PL.PRS
amics.]
friend.PL
'They ARE friends.'

To repeat the main idea, I assume that in (37–38), intervening material blocks the movement of the complementizer to the next phrase, which results in the surface word order observed in each of the constructions. The example in (39) is more complex because in addition to the evidential modifier preceding *que*, there is also a clitic left-dislocated topic below the attributive complementizer.

(39) Spanish
Claro que a Juan$_i$ lo$_i$ invitaron.
clear QUE to Juan CL.M.SG invite.3PL.PRF.PST
'Clearly, they invited Juan.'

The analysis for examples like (39) is illustrated in the structure in Figure 3.7. The complementizer starts out in Fin^0 where it is valued with the attributive feature. It then moves from head to head. The complementizer is able to cross the filled specifier of TopP because *a Juan* is a clitic left-dislocated topic that is moved from the IP to its current position. In turn, the complementizer cannot cross the filled specifier of ModP because *claro* is directly merged in the left periphery. This is why any further head movement of the complementizer is blocked at this point.

To conclude this section, I compare my own analysis to those discussed in §3.1. In Table 3.2, I summarize the main aspects of my analysis and in Table 3.3 the main aspects of the central analyses from the literature.

In my analysis, I treat *que* as a complementizer. In this book, a complementizer in Ibero-Romance is defined as an underspecified element of the form *que* that occupies a left-peripheral head position and that acquires its functional interpretation from a feature in the location in which the complementizer is merged (see §1.2). This definition allows us to maintain that in the current constructions we are dealing with the same lexical item as in unembedded *que*-initial reportatives and in any regular embedded sentences.

The different functions and syntactic behaviors follow from the assumption that the complementizer is merged in different syntactic projections where it receives its featural values. In the case of attributive *que*, we furthermore observe a greater syntactic mobility than in subordinate *que*. This is guaranteed by the low merge position, which enables upward movement through the left periphery. This movement is not possible in the case of subordinate *que* because it is already merged in the highest position of the functional field. Most previous analyses are in agreement with my assumption that *que* is a complementizer. However, some authors further state that this complementizer is different from the element that

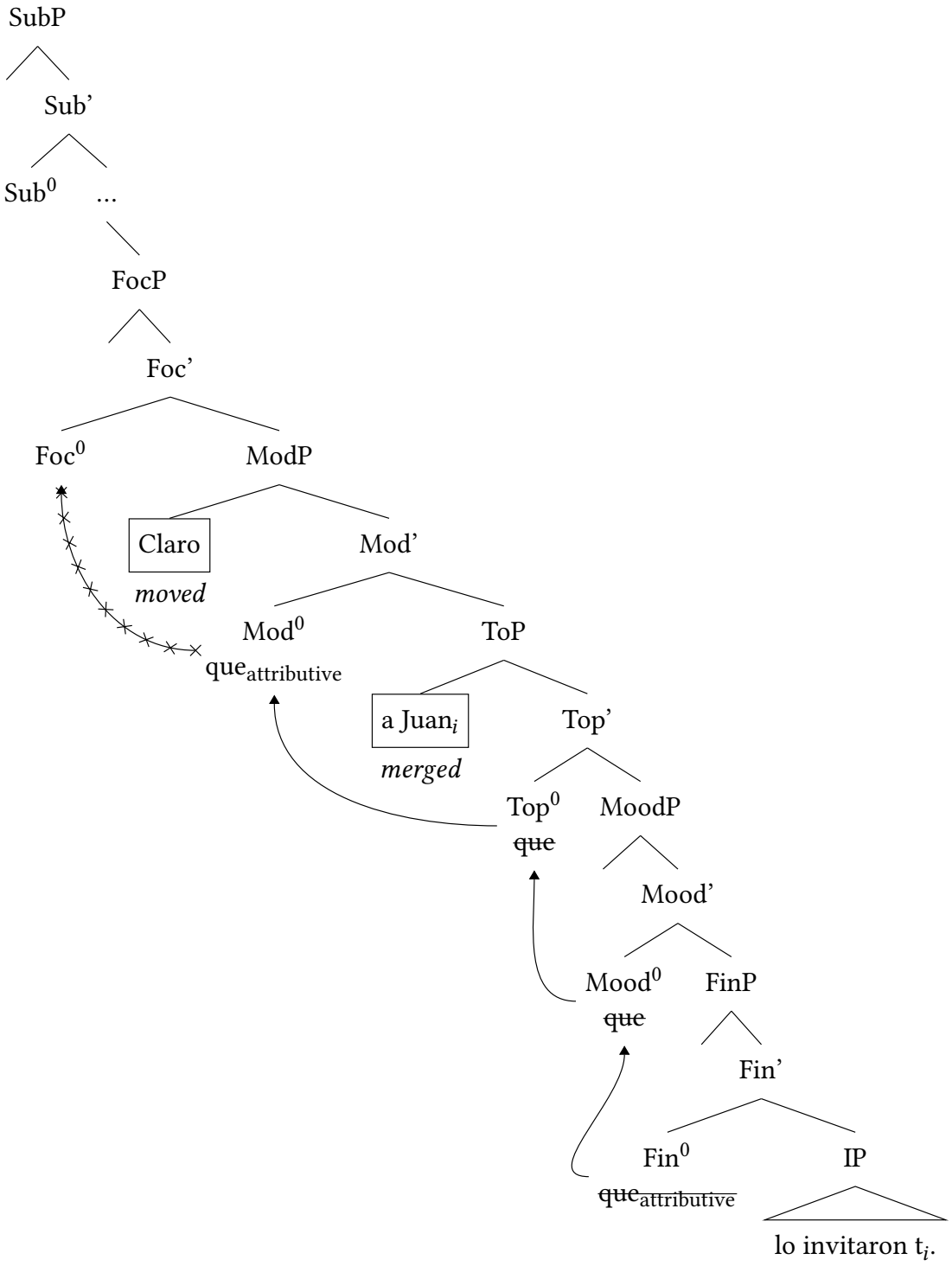


Figure 3.7: Analysis of (39); crossed out lines represent a potential movement that does not take place

Table 3.2: Attributive *que* in the present analysis

	que-initial declaratives	polar questions	*wh*-exclamatives	AdvC	AffC
nature of *que*	—— complementizer ——				
location	FinP>... >SubP	FinP>... >SubP	FinP>... >FocP	FinP>... >ModP	FinP>... >MoodP
interpretation	—— attributive feature in FinP ——				

Table 3.3: Attributive *que* in the previous analyses

	que-initial declaratives		polar questions
author	Corr (2016)		Prieto & Rigau (2007)
nature of *que*	— illocutionary complementizer —		complementizer
location	exclamative: EvalP>SAlowP	conjunctive: DeclP>EvidP> EvalP>SAHighP	FinP
interpretation	—— features in UP ——		operator in Force
	wh-exclamatives	AdvC	AffC
author	Demonte & Fernández Soriano (2014)	Cruschina & Remberger (2018)	Hernanz (2007)
nature of *que*	complementizer	complementizer	complementizer
location	FinP	ForceP	ForceP
interpretation	none	attributed to active performative structure	attributed to ForceP

introduces subordinate clauses. Corr (2016) considers it an illocutionary complementizer that is homophonous with other complementizers. Demonte & Fernández Soriano (2009) assume that the complementizer involved in *wh*-exclamatives is distinct from other homophonous elements of the form *que*; in their approach the difference is expressed in terms of different merge positions. Notably, this is very close to what I assume in this book.

In my analysis, the final landing site of the complementizer in the different constructions is derived straightforwardly by head movement that is inhibited by base-generated material. This means that the complementizer, while always merged in FinP, ends up at different heights within the left periphery in each construction. Corr (2016) also assumes a movement derivation for *que*-initial declaratives in her analysis. The complementizer, however, starts out in a higher position in one head of her split ForceP (DeclP, EvidP, EvalP) and reaches positions in her performative UP. As noted in §3.1, the motivation for this movement is mainly theoretical. My analysis in which *que* reaches SubP also captures the empirical data. Prieto & Rigau (2007) propose that *que* is merged in FinP in polar questions and do not assume further movement. I showed in §3.1 that the analysis provided in Prieto & Rigau (2007) fails to account for some critical data. My analysis, on the other hand, can account for these data because I assume that just as in declaratives, in the absence of intervening base-generated material, the complementizer reaches SubP in polar questions. In the account proposed for *wh*-exclamatives in Demonte & Fernández Soriano (2009), *que* is also analyzed as located in FinP. The empirical facts again pose some problems for this analysis, but support my own assumption that the complementizer movement stops in the head of FocP, adjacent to the *wh*-expression (see §3.1). The analyses for *que* with epistemic and evidential modifiers by Cruschina & Remberger (2018) and for *que* in verum constructions put forward by Hernanz (2007) both assume a very high position for the complementizer in ForceP. Both of these analyses, among others, are hard to reconcile with the data that show that the constructions can appear in embedded relative clauses (see §3.1). Again, my own analysis makes the correct predictions with respect to these data.

Finally, the backbone of my analysis is that in all the constructions under investigation in this chapter, the presence of the complementizer has the same basic effect on the interpretation of the sentence in its scope. In my account this common interpretation is linked to the interface feature *attributive*, with which the complementizer is valued. A more detailed characterization of the interpretation of the constructions is given in §3.4, but the basic idea is that the attributive complementizer attributes to the hearer a commitment to the proposition. Some of the previous analyses do not assume that *que* has any special meaning. Others represent the apparently different sorts of *que* in a similar way as I do here,

namely as different features in the syntactic structure. The accounts differ in whether they link these features to the complementizer or to other material or properties of the constructions. In Hernanz's analysis of verum constructions, an interpretative impact similar to factivity is attributed to the functional head that hosts the complementizer, which is ForceP in her analysis. Among the accounts for *wh*-exclamatives, Castroviejo (2006) and Demonte & Fernández Soriano (2009) do not claim that *que* gives rise to a special interpretation. In turn, Ambar (2003) states that the syntactic position that *que* is merged in guarantees a factive interpretation of the content of the *wh*-exclamatives. The feature responsible for this interpretation is also present and checked in her derivation of *que*-less *wh*-exclamatives. It therefore appears that there is no particular interpretive function attributed directly to *que* in Ambar (2003) either. In the analysis of *que*-initial sentences by Corr (2016), a combination of multiple UP-features checked by the complementizer gives rise to the interpretation. In Prieto & Rigau (2007), while *que* itself is not equipped with interpretive features and is considered optional, an operator in ForceP expressed through prosodic means is responsible for the different interpretations. In Cruschina & Remberger (2018), the special interpretation is also not linked to the position in which *que* is merged, but to the activation of the performative structure. Thus it appears that the authors connect the interpretive effect more strongly to the modifier, merged directly in the CP-external structure, than to the complementizer itself located in ForceP.

In §3.3, I outline the empirical evidence in support of my syntactic analysis. I deal with each construction individually and show how the predictions made by my analysis are confirmed by the word order that we observe.

3.3 The syntax of attributive *que*

The analysis proposed in §3.2 is based on three assumptions. First, the complementizer in the constructions involving attributive *que* is always merged in the same position: FinP. Second, the complementizer moves from head to head through the left periphery. Third, the movement of the complementizer is inhibited by a base-generated phrase in the specifier of the projection that the complementizer currently occupies.

The first two assumptions are based on theoretical and empirical considerations. A central observation, in this context, is that the pragmatic impact of the complementizer is the same in all the different constructions. A commitment to the proposition in the scope of attributive *que* is ascribed to the hearer (cf. §3.4). Since the complementizer surfaces in different positions in the constructions,

there are two options. The first option would be to assume one lexical item that is externally merged in different left-peripheral positions. The second would be to assume that the complementizer is always merged in one and the same projection and that the surface positions in the different constructions are reached through complementizer movement (see also Rizzi 1997, Poletto 2000, Roberts 2001, Ledgeway 2005). In this book, I argue in favor of the second option, which is a better fit in light of the general ideas proposed here.

The following sections contain the empirical support for my analysis. I map out the position in which the complementizer appears in the cartographic structure and show that its inability to cross a base-generated phrase explains the word order we observe in the different constructions. Moreover, I present further syntactic properties of the constructions and show how these can be accounted for in the present analysis. In §3.3.1, I focus on the syntactic properties of attributive *que*-initial declaratives and in §3.3.2 on attributive *que*-initial polar questions. The following sections are dedicated to the structure of attributive *que* when following certain left-peripheral phrases. In §3.3.3, I deal with *wh*-exclamatives, in §3.3.4 with *que* following epistemic and evidential modifiers (AdvC) and in §3.3.5 with *que* in verum sentences (AffC).

3.3.1 *Que*-initial declaratives

The basic idea of the attributive *que* analysis is that its movement is only blocked by externally merged material in the specifier of a projection. One prediction that follows from this is that if nothing intervenes, the complementizer can reach the highest projection of the functional field, i.e. SubP. In this section, I show that this is precisely where it surfaces in attributive *que*-initial declaratives. As a consequence, on a superficial level, these *que*-initial sentences show a structure parallel to that of *que*-initial reportatives. Ultimately, it is only the context that disambiguates between the two readings of the complementizer. The precise discourse contribution of attributive *que* is discussed in greater depth in §3.4. Examples (40) and (41) illustrate the difference between *que*-initial reportatives and attributive *que*-declaratives.

(40) Catalan
Pare: És dolent demanar a un fill que llegeixi un
father be.3SG.PRS bad demand.INF of a son that read.3SG.SBJV.PRS a
llibre? Mare: Que té nou anys.
book mother QUE have.3SG.PRS nine year.PL
'Father: Is it a bad thing to ask your son to read a book? Mother: He's only nine!' (ebook-cat)

(41) Catalan
Mare: Té nou anys. Pare: Eh? Mare: Que té nou anys.
mother have.3SG.PRS nine year.PL father huh mother QUE have.3SG.PRS nine year.PL
'Mother: S/he is nine years old. Father: Huh? Mother: [reportative:] S/he is nine years old.'

In (40) the reaction of the mother that their son is only nine years old does not contain new information for the father, who can be expected to be aware of the age of his own son. The mother's motive was therefore not to inform the father of the age of his son. The commitment to the proposition introduced by *que* is attributed to the hearer, the father. With her utterance, the mother communicates that she does not consider the books that the father gave to the son to be age appropriate. In (41), the same utterance by the mother results in a different reading. In this context, the father's reaction suggests that he did not understand the mother's utterance. The function of *que* at the beginning of the last sentence of this mini-dialog is therefore to mark that the sentence is a reported version of the previous statement.

The decision to analyze attributive *que* as located in SubP is not based on theoretical considerations alone. There are also empirical facts that lead to the same conclusion. First, attributive *que* precedes left-dislocated topics (cf. 42a). This is expected because moved phrases, as these topics are, do not constitute an obstacle for the head movement of the attributive complementizer.

(42) Spanish
a. ¡Pues has sido muy rápida! [$_{SubP}$ Que] [$_{TopP}$ el vídeo$_i$] lo$_i$ puso esta misma semana, jajaja.
well AUX.2SG.PRF.PRS be.PTCP very fast QUE the video CL.M.SG put.1SG.PRF.PST this same week hahaha
'Well, you've been very fast! The video, he just put it up this same week, haha.' (CdE)
b. Chicas, si son fashionistas, abran la mente! [$_{SubP}$ Que] [$_{TopP}$ la moda$_i$] la$_i$ hacen todos!
girl.PL if be.3PL.PRS fashionista.PL open.3PL.IMP the mind QUE the fashion CL.F.SG make.3PL.PRS everyone
'Girls, if you consider yourselves fashionistas, open your mind! Fashion is made by everyone.' (CdE)

A schematic derivation of the example in (42a) is given in Figure 3.8. For ease of exposition, the intermediate projections through which the complementizer moves are not displayed.

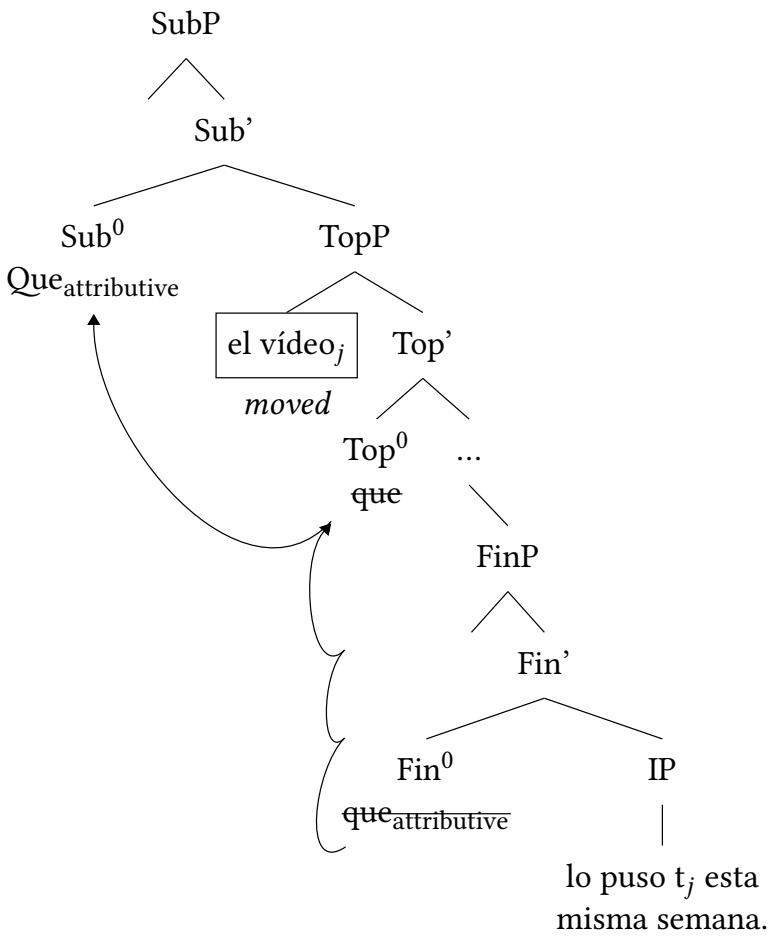


Figure 3.8: Analysis of (42a)

A further piece of evidence in favor of the assumption that attributive *que* reaches the highest position in the left periphery is that it is impossible to embed attributive *que*-initial declaratives (cf. also Corr 2016), hence the ungrammaticality of (43).

(43) Catalan
La mare va dir que (*que) té nou anys.
the mother AUX.3SG.PRF.PST say that QUE have3SG.PRS nine year.PL
'The mother said that (*that) s/he is nine years old.'

Notably, it is not because attributive *que* is disallowed in embedded contexts, as can be seen in (44) where AdvC is embedded (see also §3.3.3, §3.3.4 and §3.3.5 where further examples illustrating embedded attributive *que* constructions are discussed).

(44) European Portuguese
Só queria dizer que obviamente que, por falta de
just want.1SG.IPFV.PST say that obviously QUE for lack of
informação, podemos, por vezes, fazer dietas mais tontas.
information can.1PL.PRS at times do diet.PL more stupid
'I just wanted to say that obviously, for lack of information, at times we can end up doing stupider diets.' (CdP)

There are different theoretical alternatives to account for the fact that a sequence of two *que*s is disallowed in embedded contexts. One option is to assume that one of the complementizers, including its value, is deleted at PF. It is most likely that the complementizer with the attributive feature would be deleted, since the subordinate feature needs to be visible in syntactically selected sentences (cf. §2.2). A different way of accounting for the facts is to assume that one *que* can carry multiple values. In this case, the complementizer that is visible in the structure could at the same time be valued with the attributive feature, picked up in FinP, and the subordinate feature, picked up in SubP. Which of these alternatives proves more adequate is not a central concern here and is left aside for future research.

Attributive *que* can furthermore introduce the answer particles Catalan and Spanish *sí*, Portuguese *sim* and Catalan and Spanish *no*, Portuguese *não* (45).

(45) a. Brazilian Portuguese
– Solfieri, não é um conto, isso tudo? – Pelo inferno,
Solfieri not be.3SG.PRS a swindle this all for.the hell
que não!
QUE no
' – Solfieri, isn't all this a swindle? — Hell no!' (CdP)

b. Spanish
La multitud responde: Que sí, que sí.
the crowd answer.3SG.PRS QUE yes QUE yes
'The crowd answers: Yes. Yes.' (CdE)

c. Catalan
CARLES: Home no, tampoc no cal. PEP: Que sí, home.
Carles man no either not is necessary Pep QUE yes man
Que sí.
QUE yes
'Carles: Man, don't, it's not worth it either. Pep: Yes it is, man. Yes it is.' (caWac)

The effect is that the affirmation or negation is emphatic and in particular the fact that the hearer shares the commitment is stressed. This is in line with the general contribution of attributive *que* assumed in this book (cf. also §3.4). Concerning their syntactic structure, I adopt the idea that these answer particles are merged in a polarity position sandwiched between the lowest head of the left periphery, FinP, and the IP (cf. Laka 1990, Zanuttini 1997, Martins 2006, 2007, 2013, Hernanz 2007, Batllori & Hernanz 2008, among many others), for which, in line with Batllori & Hernanz (2008), I use the label Pol(arity)P.

(46) $[_{\text{SubP}}$ Que$_i]$ ...$[_{\text{FinP}}$ t$_i]$ $[_{\text{PolP}}$ sí/sim/no/não.]

There are no strong theoretical considerations behind this terminological choice and I am open to adopting other notations such as Laka's ΣP used by Martins (2013), if they prove more adequate.

In my analysis given in (46), I assume that attributive *que* moves to SubP in these cases too. This should be taken as a cautious proposal because further investigation into the size and internal make-up of the syntactic structure of short answers could show that some adjustment is necessary. The underlying assumption of the structure in (46) is that answer particles appear in a structure that projects a left periphery. One prediction that follows is that the answer particles should be preceded by left-peripheral material. Additionally, if attributive *que* is present, it should be subject to the same restrictions as are observed in full sentences. That this is indeed the case will be demonstrated in §3.3.4 and §3.3.5, where I will show that the answer particles are also compatible with AdvC and AffC.

Que-inial declaratives of the attributive type are attested in all three languages, although my (subjective) impression is that they are less frequent in Portuguese than in Spanish and Catalan. An even stronger contrast arises with respect to the occurrence of attributive *que* in polar questions, which is covered in §3.3.2. This is a very common construction in Catalan, but less frequent in Spanish and virtually unattested in Portuguese.

3.3.2 *Que*-initial polar questions

For *que* in polar questions, the same reasoning holds as for *que* in declaratives: The complementizer moves from head to head unrestricted and reaches the highest projection of the left periphery, if there is no externally-merged phrase in the specifier of any of the intermediate projections. Since this is generally the case in the *que*-initial polar questions under investigation, it follows that the position I assume for the complementizer here coincides with the position of *que* in the declaratives discussed in §3.3.1. The observed word order again results from the fact that the complementizer movement is not inhibited and thus reaches SubP in both *que*-initial declaratives and *que*-initial polar questions. This permits a unified account that makes use of the same mechanisms presented in §3.2 without requiring further stipulations. The relevant points are shown in Figure 3.9, where again the intermediate positions through which *que* passes are omitted in the structure.

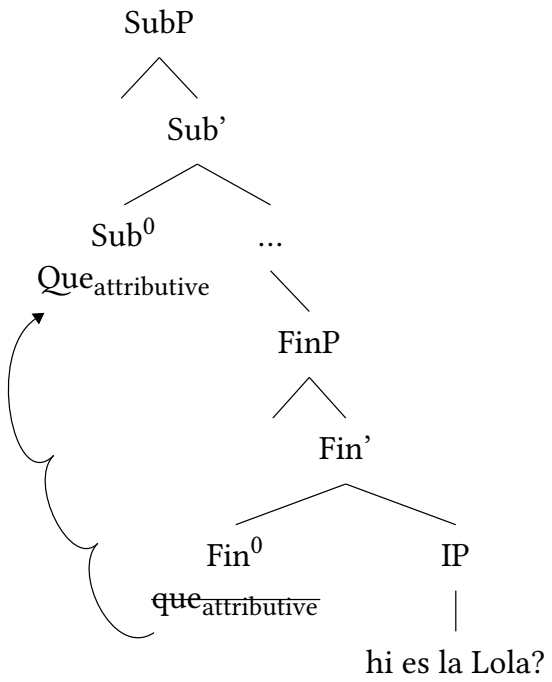


Figure 3.9: Analysis of a *que*-initial polar question

It is important to note here that I am not saying that the syntactic structure of interrogatives and declaratives is the same. There are syntactic and prosodic differences that suggest that the two clause types have different derivations; however, these differences are not a central concern for the present discussion. The

focus of this section is merely the syntactic behavior of attributive *que*. The argument I put forward is only that it behaves in essentially the same way in polar questions and declaratives.

There are no obvious differences between the syntactic properties of polar questions with *que* and those without. It therefore seems that *que* is not required for syntactic reasons. My proposal is instead that attributive *que* has a pragmatic function in polar questions too. These questions are not neutral: The speaker in fact uses them when she is biased toward a positive answer (cf. also Kocher 2017b). *Que* in polar questions is often employed when there is contextual evidence that makes the speaker suspect that the answer to the question is going to be "yes".

(47) Catalan
La Caterina va entrar i va córrer
the Caterina AUX.3SG.PRF.PST enter and AUX.3SG.PRF.PST run
cap al lavabo amb el paraigua que regalimava. –
in direction to the bathroom with the umbrella that drip.3SG.IPFV.PST
Que plou?
QUE rains
'Catarina entered and ran to bathroom with a dripping umbrella. – Is it raining?' (ebook-cat)

This can be seen in example (47). Here the dripping umbrella functions as (indirect) contextual evidence that it is raining. This leads the speaker to expect a positive answer to her question and makes the use of *que* acceptable. Further discussion of the interpretation of *que* in polar questions follows in §3.4, where I offer an explanation that allows a unified account of the discourse contribution across all constructions involving attributive *que*. The seemingly different effect in polar questions is shown to result from more general pragmatic differences between questions and assertions.

Que-initial polar questions have been mainly discussed in connection with Catalan. While many authors consider them disallowed in Spanish (for instance Mioto 2003, Prieto & Rigau 2007, Rodríguez Ramalle 2007, Etxepare 2008, Demonte & Fernández Soriano 2009, Gras 2010, González i Planas 2014, Villa-García 2015, Corr 2016), I did find attestations in corpora that point to their existence in this language.

(48) Madrileño Spanish
B: Son cinco bloques, con pista de tenis colectiva ¿no?
be.3PL.PRS five block of flats.PL with tennis court collective no

Entonces, ahí jugamos. A: ¿Que tenéis apartamento no? B: No, es un piso en un bloque.
then there play.1PL.PRS QUE have.2PL.PRS apartment no
no be.3SG.PRS a flat in a block of flats
'A: It's five blocks of flats with a shared tennis court. So, this is where we play. B: You have an apartment, right? B: No, it's a flat in a block.' (CdE)

The example in (48) can be characterized in a similar way to the Catalan example in (47). It also constitutes a biased polar question that encodes the speaker's suspicion that the answer is going to be positive. In the Spanish example, the bias of speaker A toward a positive answer is motivated by what speaker B is saying about her living situation. This makes speaker A suspect that she is living in an apartment. Therefore, with her question, speaker A intends to confirm her belief. In this case, the non-neutral nature of the polar questions is marked through the initial *que* but also through the question tag *no*.

Although examples such as (48) therefore suggest that attributive *que* is also licensed in Spanish polar questions, in what follows I will still rely mostly on Catalan examples to illustrate the core syntactic properties. I will return to Spanish and to issues of cross-linguistic variation at the end of the section.

Attributive *que* is restricted to polar questions and is not attested in *wh*-questions, hence the ungrammaticality of B's answer in (49a) when *que* is intended as attributive.

(49) a. Catalan
A: M' he d' anar, tinc un camí llarg a casa.
CL.1SG have.1SG.PRS to leave have.1SG.PRS a way long to home
B: (*$\text{Que}_{\text{attributive}}$) on vius?
QUE where live.2SG.PRS
'A: I have to leave, I have a long way home. B: (*Que) where do you live?'

A superficially equivalent version of the critical sentence is however perfectly acceptable with a subordinate complementizer giving rise to a reportative interpretation (cf. 50).

(50) Catalan
A: On vius?
where live.2SG.PRS
B: Què has dit?
what AUX.2SG.PRF.PRS say.PTCP

A: Que$_{\text{subordinate}}$ on vius.
QUE where live.2SG.PRS
'A: Where do you live? B: What did you say? A: [reportative:] Where do you live?'

The attributive and the subordinate-valued complementizers also result in different structural and interpretative properties within polar questions. In a polar question introduced by a subordinate *que*, there is an additional interrogative complementizer present merged below it (cf. 51). Furthermore, it does not have the illocutionary force nor the prosodic make-up of a question. This means that the speaker does not expect an answer. Therefore, the interjection "uf" is a natural reaction on the part of the hearer who thereby expresses his negative emotions towards his supervisor's insistence.

(51) Catalan
B: M' he trobat amb la teva directora de tesi l'
CL.1SG AUX.1SG.PRF.PST meet.PTCP with the your thesis supervisor the
altre dia i m' ha preguntat per tu.
other day and CL.1SG AUX.3SG.PRF.PRS ask.PTCP about you

A: Què t' ha preguntat?
what CL.2SG AUX.3SG.PRF.PRS ask.PTCP

B: Que si has acabat la tesi.
QUE if AUX.2SG.PRF.PRS finish.PTCP the thesis

A: Uf...
Ugh...

'A: I met your thesis supervisor the other day and she asked about you. B: What did she ask? A: Whether you'd finished your thesis. B: Ugh... '

The situation with an attributive *que*-initial polar question is different, as illustrated in (52). It appears without the additional interrogative complementizer. It furthermore has the illocutionary force of a question and, unless it is used as a rhetorical question, the speaker does expect an answer. Speaker A's statement that he now has time functions as evidence based on which speaker B can conjecture that the answer to her questions, whether he finished his thesis, will be positive. Therefore the use of *que* is felicitous.

(52) Catalan
A: Fem un cafe la setmana que ve? Ara
make.1PL.PRS a coffee the week that come.3SG.PRS now
tinc temps.
have.1SG.PRS time

B: Que has acabat la tesi?
QUE AUX.2SG.PRF.PRS finish.PTCP the thesis

A: Sí!
yes

'A: Should we have a coffee next week? I have time now. B: Have you finished your thesis? A: Yes!'

With regard to the word order, the analysis in Figure 3.9 predicts that the complementizer should be followed rather then preceded by left-peripheral material. Evidence for the high position of *que* in polar questions is provided by data like (53) (repeated from (17)), which show that the complementizer precedes a clitic left-dislocated topic.

(53) Catalan
[Context: Marta finds a bag of oranges in the kitchen. She asks her roommate:]

[$_{SubP}$ Que$_i$] [$_{TopP}$ les taronges$_j$ t$_i$] ...[$_{FinP}$ t$_i$] [$_{IP}$ les$_j$ vas
QUE the orange.PL CL.F.PL AUX.2SG.PRF.PRS
comprar tu?]
buy you

'The oranges, did you buy them?' (Kocher 2017b: 49: ex 98a)

As shown in §3.1, a topic can precede *que* in polar questions. However, these topics exhibit properties typical of hanging rather than clitic left-dislocated topics. They are followed by an intonational break and can be resumed not only by a clitic but also by a full pronoun, a DP or an epithet as the example in (54) shows.

(54) Catalan
[Context: Marta and Maria are at a party. Marta sees that Maria's colleague Jordi is also there. Marta had a fight with Jordi recently and is not pleased about his presence. She is also sure that nobody at the party knows him but Maria. She asks her:]

[$_{\alpha P}$ En Jordi$_i$,] [$_{SubP}$ que] [$_{TopP}$ aquest idiota$_i$ ...[$_{IP}$ l$_i$'
the Jordi QUE that idiot CL.M.SG
has convidat tu?]
AUX.2SG.PRF.PRS invite.PTCP you

'Did you invite that idiot Jordi?'

Additionally, according to my informants, it is not sufficient to have a co-referential epithet or DP: The example is in fact only grammatical with an additional co-referential clitic pronoun. These properties suggest that the topic that precedes *que* is merged outside of the core structure in a CP-external position. Therefore, examples like (54) do not constitute evidence against my analysis. In the structure in (54) *en Jordi* is merged directly in a CP-external position and *aquest idiota* is a co-referential clitic left-dislocated topic that is moved to the left periphery.

Catalan *que*-initial polar questions can furthermore be introduced by pragmatic particles such as *oi* and *eh*, which according to Prieto & Rigau (2007) are present when the speaker wants to achieve a confirmatory reading of the question. I briefly return to their pragmatic function in §3.4.

(55) Catalan
Oi que ens entenem?
OI QUE CL.2PL understand.2PL.PRS
'We understand each other, right?' (caWac)

Prieto & Rigau (2007) propose that these particles are merged in the specifier of ForceP, which structurally coincides with SubP in the cartographic structure assumed in this book. The authors assume that *que* is in FinP, hence at the lower edge of the left periphery. Based on this analysis, one would expect that material would be able to intervene between *oi* and *que*. The data in (56), however, suggest precisely the opposite. The ungrammaticality of a topic intervening between *oi* and *que* shows that these two words need to be adjacent to each other. I take this as evidence that they are located in the specifier and head of the same projection.

(56) Catalan
* Oi en Jordi que l' has convidat tu?
OI the Jordi QUE CL.M.SG AUX.2SG.PRF.PRS invite.PTCP tu
Intended: 'You are the one that invited Jordi, right?'

The example in (56) is consistent with the analysis that follows from the general assumption presented in this chapter. Just as in regular *que*-initial polar questions, the complementizer moves all the way up through the left-peripheral heads and ends up in SubP. Further support for the high position of *(oi) que* is given in example (57), which shows that it precedes a clitic left-dislocated topic.

(57) Catalan
[$_{\text{SubP}}$ Oi que$_i$] [$_{\text{TopP}}$ aquesta pregunta$_j$ t$_i$] ...[$_{\text{FinP}}$ t$_i$] [$_{\text{IP}}$ no se l$_j$'
OI QUE this question not CL.REFL CL.F.SG
havien fet mai?]
AUX.3PL.PRF.PRS made never
'They never asked themselves this question, right?' (caWac)

An additional piece of evidence that *que* reaches a high position in polar questions is that, just as in the case of *que*-initial declaratives discussed in §3.3.1, *que*-initial polar questions are not found in embedded contexts (cf. 58a, 58b). This is predicted by the present analysis because the two instances of *que* would compete for the same projection, SubP (see the discussion around (43) for some suggestions of how to account for this theoretically).

(58) Catalan

a. La Maria va preguntar que (si) (*que) l'
the Maria AUX.3SG.PRF.PST ask that whether QUE CL.3SG
he convidat jo.
AUX.1SG.PRF.PST invite.PTCP I
'Maria asked whether I invited him.'

b. La Maria va preguntar que (si) (*oi que) hi es
the Maria AUX.3SG.PRF.PST ask that whether OI QUE CL.LOC is
la Lola.
the Lola
'Maria asked whether Lola is there.'

As stated above, *que*-initial polar questions have, to date, primarily been discussed as a feature of Catalan grammar. To the best of my knowledge, apart from Hualde (1992), its existence in Spanish has so far been disregarded. The review of corpus data, however, shows that Spanish does have *que*-initial polar questions as well. Moreover, they have the same function, namely, they express that the speaker is biased towards a positive answer (cf. 48). Another example that illustrates this is given in (59).

(59) Chilean Spanish

– Creo que no se le pagan 15 millones
believe.1SG.PRS that not CL.REFL CL.DAT pay.3PL.PRS 15 million.PL
mensuales para que ande haciendo proselitismo político
monthly so that go.3SG.PRS do.PTCP.PRS proselytism politic
junto a Horst Golborne. – ¿Qué? – ¿Que no se llama
together to Horst Golborne what QUE not CL.REFL call.3SG.PRS
Horst? Ah perdón, me confundí...
Horst ah sorry CL.1SG confuse.1SG.PRF.PST
'– I think that he doesn't get 15 million a month to wander around doing political proselytism along with Horst Golborne. – What? – So he's not called Horst? Ah, sorry, I confused the name.' (CdE)

The speaker who utters the *que*-initial polar question takes her interlocutor's reaction to mean that she used the wrong name. She therefore expects the answer to her question *He is not called Horst?* to be affirmative, and consequently is biased towards a positive answer. The origin of the two examples moreover shows that *que*-initial polar questions in Spanish cannot be the result of Catalan influence, contra Hualde (1992: 2) where it is stated that "the use of *que* in questions when transferred to Spanish, is stereotypical of a Catalan background". This can hardly be the case given the attestations from varieties that are not in contact with Catalan (cf. (48) from Madrileño and (59) from Chilean Spanish).

The example in (60), in which *que* precedes a clitic left-dislocated topic, suggests that Spanish *que*-initial polar questions also have the same syntactic properties as their Catalan equivalents. Therefore, my analysis can be extended to Spanish.

(60) US-Spanish
[$_{SubP}$ ¿Que$_i$] [$_{TopP}$ la respuesta$_j$ t$_i$] [$_{FinP}$ t$_i$ [$_{IP}$ la$_i$ publicamos en
QUE the answer CL.F.SG publish.1PL.PRS in
periódicos de provincias?
newspaper.PL of province.PL
'Are we publishing the answer in provincial newspapers?' (CdE)

The data presented here suggest that there is no evidence for a systematic syntactic difference between Spanish and Catalan *que*-initial polar questions. The only difference is that Spanish does not appear to allow particles like *oi* or *eh*. My conclusion is therefore that both languages permit attributive *que* in polar questions.

Although a more extensive investigation is yet to be completed, a brief review of Spanish corpus data shows that attributive *que* is actually quite frequent (cf. Table 3.4). For the purpose of this corpus study, I relied on the data from the modern Spanish subportion of the 2001 version of the CdE, which are annotated for text type and country of origin. I used this smaller subcorpus of CdE for this inquiry because the large web-based corpus contains more orthographic irregularities. A particularly common misspelling is the omission of diacritics, leading to the loss of the formal distinction between the interrogative pronoun *qué* 'what' and the complementizer. Using these imperfect data would risk wrongly including a great proportion of *wh*-questions in the sample. In total there are 180 occurrences (total tokens in the subcorpus: 20.4 million) of *que*-initial polar questions in the subcorpus. In comparison, the Catalan caWac contains 3124 occurrences (total tokens in the corpus: 780 million). This means that the relative number

of occurrences in the CdE is actually larger than in the caWac. However, given the differences in the composition of the corpora, I am hesitant about drawing conclusions from this result. The CdE contains dialog data that favors the use of questions, as opposed to the web-based caWac which might be less likely to include questions. The majority of the occurrences in CdE stem from oral and fiction data. Unfortunately, there is no information on the number and proportion of tokens from the different dialectal varieties in the corpus, so the absolute numbers of occurrences presented in the Table cannot be systematically compared in this respect. What the data once again confirm, however, is that *que*-initial polar questions are unlikely to result from contact with Catalan, since there are attestations from various non-European varieties.

Table 3.4: Absolute numbers of *que*-initial polar questions in the 2001 contemporary subcorpus of CdE (N/A: not available)

	fiction	press	oral	total
Argentina	21	1	7	29
Bolivia	0	0	4	4
Chile	19	0	5	24
Colombia	2	3	1	6
Cuba	1	1	0	2
Gran Canaries	0	0	1	1
Guatemala	2	1	1	4
Honduras	0	1	0	1
Mexico	13	0	19	32
Paraguay	14	0	0	14
Peru	11	1	0	12
Puerto Rico	1	0	1	2
Spain	22	7	77	106
Venezuela	0	0	8	8
N/A	20	0	0	20
total	126	15	124	180

Finally, in Portuguese there are no cases of *que*-initial polar questions attested. This suggests that in this respect there is in fact a systematic syntactic contrast between Portuguese on the one hand and Spanish and Catalan on the other.

(61) a. Portuguese
A Joana pergunta (*que) se vens.
the Joana ask.3SG.PRS that if come.2SG.PRS

b. Spanish
Juana pregunta que si vienes.
Juana ask.3SG.PRS that if come.2SG.PRS

c. Catalan
La Joana pregunta que si vens.
the Joana ask.3SG.PRS that if come.2SG.PRS
'Jo/uana asks if you are coming.'

Corr (2016) offers an explanation for this discrepancy. The author hypothesizes that Portuguese *que* is more restricted than its cognates in Spanish and Catalan in that it is specialized for declaratives and cannot appear in other clause types. One of the reasons behind this hypothesis is the ungrammaticality of *que* in embedded and reported questions above an interrogative complementizer *se* in Portuguese (see 61a), which contrasts with the other two languages (see 61b, 61c; cf. §2.4 for further discussion on this contrast).

3.3.3 *Que* in *wh*-exclamatives

The present subsection deals with the syntactic properties of attributive *que* in *wh*-exclamatives attested in all three languages under investigation. The complementizer, when present, always appears adjacent to the *wh*-expression. I adopt the proposal made in Demonte & Fernández Soriano (2009) that the *wh*-expression is located in FocP. In the cartographic literature, this projection has been identified as the host of foci and *wh*-phrases (cf. Rizzi 1997). The derivation I assume for (62) (repeated from (22)) is given in Figure 3.10.

(62) Spanish (Demonte & Fernández Soriano 2009: 33 : ex 19a)
¡Qué rico que está!
how good QUE be.3SG.PRS
'How good this is!'

The intermediate positions through which the complementizer passes are again omitted in the present structure. Attributive *que* is as always assumed to be merged in FinP and moves from head to head until it reaches FocP, where the movement comes to a halt.

This analysis assumes that the *wh*-expression is merged in the left periphery rather than moved to it. This allows us to maintain the idea that the movement of the attributive complementizer is inhibited by one simple condition that disallows the crossing of base-generated material. In what follows, I will show that in addition to the theoretical plausibility of this account, there is also empirical evidence in favor of the assumption that the *wh*-expression is base-generated in the left periphery.

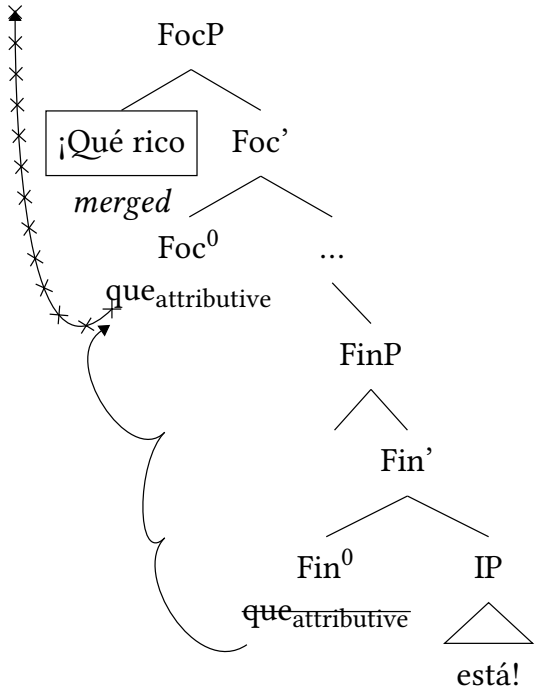


Figure 3.10: Analysis of (62); crossed out lines represent a potential movement that does not take place

If I adopt a derivation with a left-peripheral base position for the *wh*-expression, one crucial aspect that requires explanation is how the *wh*-expression ends up being interpreted as an argument or adjunct dependent on the IP-internal verb. A non-local theta- and even case-assignment as must be assumed here are not an unprecedented idea in the literature (see Bošković 2007, Villa-García 2015: 168–170, Saab 2015). A similar issue is encountered in Villa-García (2015). He analyzes recomplementation configurations such as that in (63), where the clitic left-dislocated topic *a los alumnos* is sandwiched between two complementizers. Assuming non-local case-assignment in these cases becomes necessary because Villa-García (2015) shows that the topics are generated directly in the left periphery.

(63) Spanish (Villa-García 2015: 18: ex 23)
Susi dice que a los alumnos, que les van a
Susi say.3SG,PRS that DAT the student.PL that CL.DAT.3PL go.3PL.PRS to
dar regalos.
give presents
'Susi says that they are going to give the students presents.'

In order to explain how these topics are assigned their case, Villa-García (2015) adopts the agreement mechanism described in Bošković (2007), building on the principle of *Greed* introduced in Chomsky (1993). The basic idea of Bošković's system is that the standard assumption that the *v* is the probe and the case-marked DP is the goal is reversed.[10] In this system, the DP is the probe and moves to a position from which it c-commands the goal *v*. The DP probes *v* to license its case. Villa-García (2015) applies this system to the recomplementation data. The derivation is even simpler in this case because according to the author the probe *a los alumnos* is base-generated in the left-peripheral position sandwiched between two complementizers. It is therefore never lower than its case-licensor (*v*) and no movement of the DP is required. The DP probe c-commands its goal *v* from its base-generated location and is therefore in a position to check off its case feature (see Villa-García 2015: 168–170).

Case assignment from a left-peripheral position, as proposed in Villa-García (2015), does not seem to be active in the construction under investigation. As can be seen in (64b), the left-peripheral *wh*-expressions cannot bear case marking when followed by attributive *que*.

(64) Spanish

a. Han dado un premio a una estudiante
AUX.3PL.PRF.PRS give.PTCP the prize DAT a student
inteligente.
intelligent
'They gave the prize to an intelligent student.'

b. * ¡A qué estudiante más inteligente que han
DAT which student more intelligent QUE AUX.3PL.PRF.PRS
dado el premio!
give.PTCP the prize

[10] In the minimalist program, syntactic agreement is perceived of as a matching relation between a probe and a goal (see Chomsky 2000), where matching can be defined as feature identity.

c. * ¡Qué estudiante más inteligente que han dado
which student more intelligent QUE AUX.3PL.PRF.PRS give.PTCP
el premio!
the prize
Intended: 'What an intelligent student they gave the prize to!'

In the declarative equivalent to the *wh*-exclamative in (64a), the object of the verb is introduced by the dative marker *a*. The corresponding *wh*-exclamative in (64b), in which the dative object is a left-peripheral *wh*-expression, is judged ungrammatical by my informants. Furthermore, example (64c) shows that an unmarked *wh*-expression is not grammatical either.[11]

These data contrast with the example in (65), which shows that in the absence of attributive *que* a dative-marked object can be a left-peripheral *wh*-expression.

(65) Spanish
¡A qué estudiante más inteligente (le) han dado
DAT which student more intelligent CL.DAT AUX.3PL.PRF.PRS give.PTCP
el premio!
the prize
'What an intelligent student they gave the prize to!'

This example provides crucial evidence for my idea, which I will return to below, that *wh*-exclamatives with and without *que* differ in that in the former case the *wh*-expression is base-generated in the left periphery, while in the latter, it reaches the surface position through movement.

The next question I address is how left-peripheral elements end up being interpreted as dependent on the verb.

(66) a. European Portuguese
Que coisa mais idiota que fazem aos animais.
what thing more stupid QUE do.3PL.PRS to.the animal.PL
'What stupid things people do to animals.' (CdP)

b. Catalan
Que malament que anem.
how badly QUE go.1PL.PRS
'How badly it is going for us.' (caWac)

[11] There appears to be a small degree of variation involved, as a minority of my informants judged examples like (64b) marginally acceptable. Additionally, some of the informants found (65) only grammatical with a dative clitic co-referent to the *wh*-expression.

The potentially more complex cases are those in which the *wh*-expression appears to be an argument of the verb, as in (66a). In line with Villa-García (2015), I adopt the idea that for theta-role assignment, clausematehood is a sufficiently local configuration. In other words, a verb can assign its theta-role not only to the elements in its argument positions, but also to elements that are externally merged in different positions, as long as they are contained in the same clause. Saab (2015) presents a formal account of long distance theta assignment. According to him, there are two central principles that must be met in theta assignment: locality and activity (see also Chomsky 2000, 2001). Saab (2015) states that a thematic head can assign a theta-role to a given argument if and only if the argument is active and local with respect to the thematic head (cf. Saab 2015: 2). Activity is conceived of as an unvalued K-feature at the point of derivation when the theta-role is assigned. The crucial point for the present argument is that in Saab's proposal locality is not based on merge. This contrasts with previous accounts, such as that proposed by Sheehan (2012) who states that theta-roles are assigned via internal or external merge with a thematic head. On the contrary, in Saab (2015), a local argument is simply defined as the closest argument to the thematic head.

The principles of activity and locality required for theta-role assignment in the system outlined in Saab (2015) are both met by the *wh*-expressions in *wh*-exclamatives like (66a). The object *os animais* is local but not active because it has been assigned its case by the preposition *a* and thus does not contain an unvalued K-feature. In turn, the object *que coisa mais idiota* is both active, i.e. not case-marked, and local in relation to the thematic head *fazem* because there is no other potential active argument closer to the thematic head (i.e. c-commanded by *que coisa mais idiota*).

There are also cases of *wh*-exclamatives such as (66b), in which the *wh*-expression appears to be a modifier of a verb. Non-local structural positions yet local interpretations of modifiers have also been addressed previously in the literature. It has been proposed independently that adverbial modifiers can be base-generated at the edge of the clause in which they are interpreted (Rizzi 1990: 46–51, Uriagereka 1988). In a similar vein, it is maintained by some authors that the counterpart of *why* in various languages is base-generated in its left-peripheral surface position, even though its interpretation is dependent on an element deeper in the structure (Hornstein 1995, Rizzi 1990, 2001, Shlonsky & Soare 2011).

Having discussed how the IP-dependent interpretation of the *wh*-expressions can be accounted for, I will now evaluate the predictions of the analysis outlined in Figure 3.10. I show that the word order of the *wh*-expression and the complementizer with respect to other left-peripheral material confirms the analysis. As

stated at the start of this section, *que* and the *wh*-expression must observe strict adjacency, meaning that no left-peripheral material can intervene, hence the ungrammaticality of (67). In the analysis, this is captured by postulating that they appear in the specifier and head of the same projection.

(67) Catalan (adapted from Kocher 2017b: 48: ex 95a)
* Que estrany al Jordi$_i$ que li$_i$ hagin trucat.
how strange DAT.the Jordi QUE CL.DAT AUX.3PL.SBJV.PRF.PRS call.PTCP
Intended: 'How strange that they called Jordi of all people.'

The analysis furthermore predicts that it should be possible for the sequence of *wh*-expression plus complementizer to be preceded and followed by a clitic left-dislocated topic. This prediction also holds true, as can be seen in (68a) and (68b).

(68) Catalan

a. (adapted from Kocher 2017b: 48: ex 95b)
[$_{\text{FocP}}$ Que estrany que$_i$] [$_{\text{TopP}}$ al Jordi$_j$ t$_i$] [$_{\text{FinP}}$ t$_j$] [$_{\text{IP}}$ li$_j$
how strange QUE DAT.the Jordi CL.DAT
hagin trucat.]
AUX.3PL.SBJV.PRF.PRS call.PTCP

b. (adapted from Kocher 2017b: 48: ex 95c)
[$_{\text{TopP}}$ Al Jordi$_i$,] [$_{\text{FocP}}$ que estrany que$_j$] [$_{\text{FinP}}$ t$_j$] [$_{\text{IP}}$ li$_i$
DAT.the Jordi how strange QUE CL.DAT
hagin trucat.]
AUX.3PL.SBJV.PRF.PRS call.PTCP
'How strange that they called Jordi of all people.'

Moreover, since the complementizer only moves up to FocP but not higher, *wh*-exclamatives with attributive *que* should appear in embedded contexts as well as in *que*-initial reportatives. Both of them contain the high complementizer merged directly in SubP which is valued with a subordinate feature in my analysis. Examples (69a) and (69b) confirm that this is indeed the case.

(69) Spanish

a. María dijo [$_{\text{SubP}}$ que$_{\text{subordinate}}$] [$_{\text{FocP}}$ qué suerte que$_i$]
María say.3SG.PRF.PST that what luck QUE
[$_{\text{FinP}}$ t$_i$] [$_{\text{IP}}$ no salió de fin de semana.]
not go-out.3SG.PRF.PST for weekend
'María said how lucky that she didn't go out on the weekend.' (CdE)

b. [$_{SubP}$ Que$_{subordinate}$] [$_{FocP}$ qué susto que$_i$] [$_{FinP}$ t$_i$] [$_{IP}$ no se
QUE what shock QUE not CL.REFL
aguante mis pataletas de treceañera que no puedo
bear.3SG.SBJV.PRS my teenage tantrums that not can.1SG.PRS
controlar.]
control
'[reportative:] What a shock that s/he cannot bear my tantrums that I cannot control.' (CdE)

Having established that the analysis makes the right predictions with respect to the word order in the left periphery, I now return to the idea that *wh*-expressions that precede *que* are merged in, rather than moved to, the left periphery. As mentioned above, this assumption is necessary in order to maintain the uniform analysis defended in this chapter. The relative position of attributive *que* with respect to other left-peripheral material follows straightforwardly from a simple restriction on an otherwise unconditioned complementizer movement, which disallows movement across material that is merged in the left periphery. I will now show that this is not just a theoretical necessity; on the contrary, it is also supported by empirical evidence.

In previous works on *wh*-exclamatives, the presence of *que* has been considered optional (see for instance Villalba 2003, Castroviejo 2006). The reasons for this is that the Ibero-Romance languages under investigation allow *wh*-exclamatives with and without a low complementizer.[12] The existence of parallel examples like (70a) and (70b) leads the authors to conclude that they have the same underlying structure and optionally allow the merger of a complementizer. The complementizer itself has been described as semantically vacuous (Castroviejo 2006).

(70) Catalan

a. Quin nas més gros que tens.
what nose more big QUE have.2SG.PRS

b. Quin nas més gros tens.
what nose more big have.2SG.PRS
'What a big nose you have.'

[12]There might be some cross-linguistic variation with respect to the possibility of omitting the complementizer. González i Planas (2010) states that the complementizer is obligatory in Catalan *wh*-exclamatives in many contexts where it can be omitted in Spanish.

Previous accounts furthermore assume that *wh*-exclamatives involve movement of the *wh*-expression to a left-peripheral position (see for instance Benincà 1996, Gutiérrez-Rexach 2001, Zanuttini & Portner 2003, Ambar 2003, Villalba 2003, 2008, Castroviejo 2006, Demonte & Fernández Soriano 2009, Gutiérrez-Rexach & Andueza 2011, Cruschina et al. 2015). In contrast, I believe that there are two different derivations for *wh*-exclamatives in Ibero-Romance. In one, the *wh*-expression is moved to the left periphery from an IP-internal position; and in the other, it is merged directly in the left periphery. The two examples in (70a) and (70b) are then not structurally equivalent. Furthermore, *que* only appears to be optional, but this is not in fact the case. The contrast between (71a) and (71b) repeated from (64b) and (65) shows that in a *wh*-exclamative with attributive *que*, a case-marked *wh*-expression is impossible. On the contrary, the same case-marked *wh*-expression is grammatical in *que*-less exclamatives.

(71) Spanish

a. * ¡A qué estudiante más inteligente que han
DAT which student more intelligent QUE AUX.3PL.PRF.PRS
dado el premio!
give.PTCP the prize
intended: 'What an intelligent student they gave the prize to!'

b. ¡A qué estudiante más inteligente (le) han
DAT which student more intelligent CL.DAT AUX.3PL.PRF.PRS
dado el premio!
give.PTCP the prize
'What an intelligent student they gave the prize to!'

Another key piece of evidence for two different derivations is the existence of *wh*-exclamatives like (72), where crucially, the complementizer cannot be omitted (cf. also González i Planas 2010)

(72) a. Catalan
Quina sort *(que) neva.
what luck QUE snow.3SG.PRS
'What luck that it is snowing!'

b. Portuguese
Que mau *(que) isto se tenha passado.
how bad QUE this CL.REFL AUX.3SG.SUBJ.PRF.PRS happen.PTCP
'How bad that this happened.'

c. Spanish
Qué lástima *(que) no vienes.
what pity QUE not come.2SG.PRS
'What a pity that you're not coming.'

At this point I have to briefly digress to acknowledge that there is an alternative view to the one maintained here, namely that these examples are not monoclausal exclamatives but biclausal clefts with an omitted copula. In this view, the structure of (72a) can be represented by (73), adapting an analysis of clefts proposed in Belletti (2009, 2013).

(73) $[_{\text{SubP1}} \ldots [_{\text{IP1}}]$ ~~es~~ $[_{\text{SubP2}} \ldots [_{\text{FocP}}$ Quina sorte que$_i$ $\ldots [_{\text{FinP}}$ t$_i$ $[_{\text{IP2}}$ neva.]]]]]

Although my claim is that cases such as (72) do in fact constitute exclamatives and I provide evidence for this below, the alternative view is not necessarily in complete disagreement with my general proposal. Adopting a cleft-analysis à la Belletti, firstly, places the complementizer in FinP and assumes movement from left-peripheral head to left-peripheral head that comes to a halt in FocP where the clefted phrase is merged. This parallels the syntactic analysis I assume for all attributive *que* constructions addressed in the present chapter. Secondly, it has been observed that clefts carry an existential presupposition (among many others, Büring & Kriz 2013). A sentence like *It is Fred she invited.* carries the presupposition *She invited someone.* This is very close to the meaning I identify for attributive *que*. I will not go into detail here. Note, however, that if, as is commonly assumed, what it means to presuppose a proposition is that this proposition is part of the speaker's and the hearer's common ground, then the presence of the existential presupposition can be reconciled with the meaning contribution I identify for attributive *que* (see §3.4). This could allow us to explain why we have the same resulting interpretation even though these could be viewed as embedded and not root sentences. The idea would be that a complementizer merged in FinP, irrespective of whether we are dealing with an embedded or unembedded context, picks up an attributive feature which has the described effect on the interpretation. That (72a) presupposes *It snows* then boils down to the simple fact that the feature linked to the FinP-merged complementizer as usual attributes a commitment to the proposition in its scope (in this case: *It snows*) to the hearer.

Returning now to the discussion at hand, the difference between the examples in (70) and in (72) is that in the former the *wh*-expression can be reconstructed in the core sentence, while in the latter it cannot. Consider the examples in which

the *wh*-expression appears to be syntactically dependent on the verb: It is a predicate dependent on a copula verb in (74), an argument in (70) and an adjunct in (75).

As a side note, a contrast also seems to be present in English between *wh*-exclamatives where the *wh*-expression appears to originate in the IP and *wh*-exclamatives where it does not. This can be seen in the translations given below the relevant examples: Only those where the *wh*-expression does not have a function in the IP (for instance (72) vs. (74a)) allow the presence of a complementizer below the *wh*-expression.

(74) a. Portuguese
Que chato que és.
how annoying QUE be.2SG.PRS

b. Que chato és.
how annoying be.2SG.PRS
'How annoying you are.'

(75) a. Spanish
Qué bien que nos ha ido.
how well QUE CL.1PL AUX.3SG.PRF.PRS go.PTCP

b. Qué bien nos ha ido.
how well CL.1PL AUX.3SG.PRF.PRS go.PTCP
'How well it went for us.'

In sum, while parallel examples like those illustrated in (70), (74) and (75) could lead to the conclusion that a derivation by movement is a tenable analysis, examples such as those in (72), where the *wh*-expression does not have a function in the IP, are not consistent with an account in which their left-peripheral position is derived by movement. I propose that the contrast is not between cases where the *wh*-expression has a function in the IP and those where it does not, but between *wh*-exclamatives with and without a complementizer. I suggest that these are two types of *wh*-exclamatives that are structurally distinct. In *wh*-exclamatives with *que* the *wh*-expression is merged in the left periphery (cf. Figure 3.12), while in those without *que*, the *wh*-expression is promoted to the left periphery through movement from an IP-internal position (cf. Figure 3.11).

Therefore, examples like (70a), (74a), (75a) and (72) have the same underlying structure in the CP, while examples like (70b), (74b) and (75b) differ structurally. Further motivation behind the assumption that there are distinct derivations comes from the position of subjects. At least in Spanish and Catalan, the complementizer-less *wh*-exclamatives do not allow a preverbal subject: The subject necessarily follows the verb (see 76). Castroviejo (2006) takes this as evidence

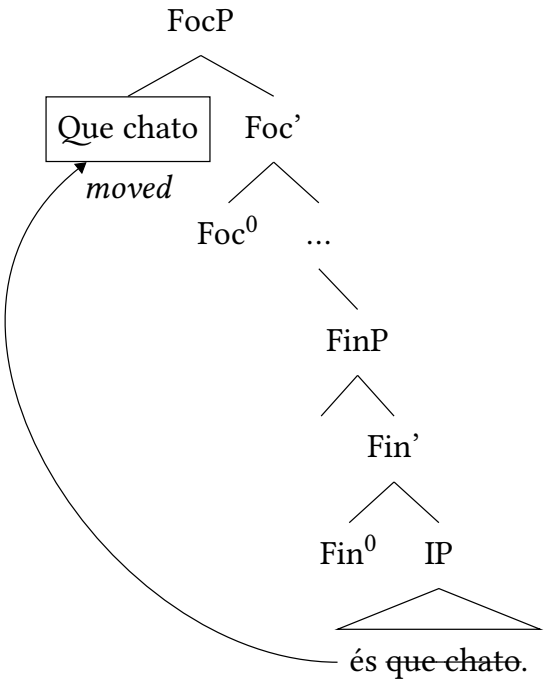


Figure 3.11: Analysis of (74b)

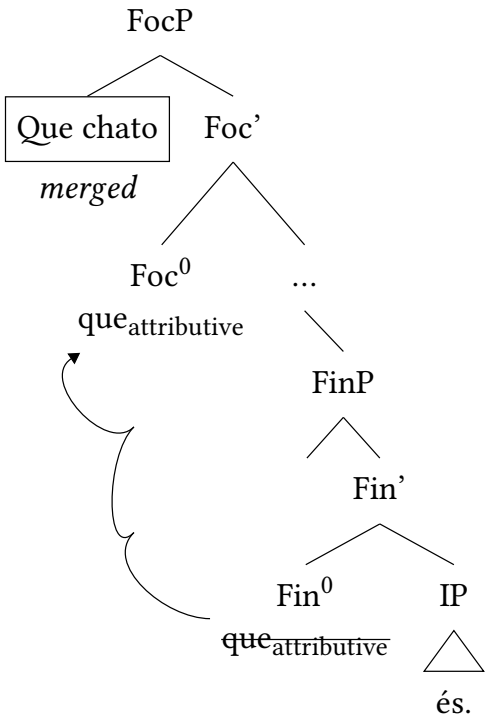


Figure 3.12: Analysis of (74a)

for the movement derivation of *wh*-exclamatives, suggesting that the moved *wh*-expression passes through a position at the edge of the IP which would otherwise be occupied by the preverbal subject.

(76) a. Catalan
Que blanca i bonica es presenta des d' allí la
how white and pretty CL.REFL present.3SG.PRS from there the
ciutat.
city
'How white and pretty the city looks from there.' (caWac)

b. Spanish
Qué suerte tenemos los lectores de tener quien nos
what luck have.1PL.PRS the reader.PL to have who CL.1PL
escriba.
write.3SG.SBJV.PRS
'What luck we readers have that someone writes to us.' (CdE)

Portuguese permits both word orders (cf. 77).

(77) Brazilian Portuguese

a. Que sorte tive eu do ter
what luck have.1SG.PRF.PST I to-CL.M.SG AUX.INF.PRF.PRS
conhecido.
know.PTCP
'What luck I had to get to have known him!' (CdP)

b. Que sorte eu tive de conseguir me despedir de
what luck I have.1SG.PRF.PST to manage CL.1SG say goodbye to
você.
you
'What luck I had to have managed to say goodbye to you.' (CdP)

All three languages allow preverbal subjects in *wh*-exclamatives with an attributive complementizer (as in 78). Crucially, this is the case irrespective of whether the *wh*-expression appears to have a function in the IP as in (78b) or does not as in (78a), (78c).

(78) a. Catalan
Que trist que tu no sàpigues ni escriure ni
how sad QUE you not know.2SG.SBJV.PST neither write nor
expressar -te correctament.
express CL.2SG correctly
'How sad that you know neither how to write nor how to express yourself correctly.' (caWac)

b. Brazilian and European Portuguese
Que parva (que) eu sou!
what fool QUE I be.1SG.PRS
'What a fool I am.' (CdP)

c. Spanish
Qué lástima que la vida sea tan corta y las
What pity QUE the life be.1SG.SUBJ.PRS so short and the
dictaduras tan largas.
dictatorships so long
'What a pity that life is so short and dictatorships are so long.' (CdE)

In the examples in (78a) and (78c), the verb is inflected for subjunctive and not indicative as in the rest of the examples. In both *wh*-exclamatives the *wh*-expression is an evaluative term. The choice of mood, however, is not consistent but varies to a large degree. For instance, example (78c) contrasts with (72c),

which contains the same evaluative term, *lástima* 'pity', but the verb is inflected for indicative. As the subjunctive is most commonly encountered in complement clauses, one might be led to the conclusion that these are elliptical structures ([$_{CP1}$ Qué lástima ~~(es)~~ [$_{CP1}$ que la vida sea tan corta.]]). However, as can be seen in the example in (79), the subjunctive is not restricted to complement clauses but also appears in main clauses. This is why, for now, I stick to the idea that, despite the subjunctive mood, these *wh*-exclamatives can also be analyzed as main clauses.

(79) Spanish
Ojalá nos dejen.
hopefully CL.1PL leave.3PL.SBJV.PRS
'Hopefully they will leave us alone.' (CdE)

One final argument in support of the move vs. merge distinction comes from data reported in Corr (2016). The author presents examples in which a complementizer precedes a *wh*-expression in a *wh*-exclamative such as (80).

(80) Galician (Corr 2016: 90: ex 24)
Hala, que que ben (*que) fala a irlandesa!
wow QUE what well QUE speak.3SG.PRS the Irish
'Wow, the Irish girl speaks so well!'

(81) [$_{SubP}$ Que$_{j,attributive}$] ... [$_{FocP}$ que ben$_i$ t$_j$] ... [$_{FinP}$ t$_j$][$_{IP}$ fala t$_i$ a irlandesa!]

Incidentally, in all the examples of this type discussed in Corr (2016), a complementizer below the *wh*-expression is ungrammatical and all examples are cases where the *wh*-expression can be reconstructed in the IP. This behavior is expected based on the assumptions made here. My analysis is given in (81). The *wh*-expression is merged in the IP and moves to FocP. The complementizer is merged in FinP and valued with an attributive feature. It moves from head to head, eventually reaching SubP. It is able to cross the *wh*-expression in the specifier of FocP because the moved phrase does not constitute an obstacle to the complementizer movement.

To conclude this section, I presented evidence in favor of my argument that the contrast between the *wh*-exclamatives with and without the complementizer boils down to a move vs. merge distinction. As predicted by the general idea maintained in this book, it is precisely the *wh*-expressions that are merged in the left periphery that are followed by a complementizer.

3.3.4 *Que* with epistemic and evidential modifiers

This section deals with attributive *que* when it appears adjacent to epistemic and evidential modifiers. I use the term AdvC to refer to this pattern. The analysis I propose is that the modifier is merged directly in the specifier of ModP. The complementizer starts out in FinP, where it receives the attributive feature and subsequently moves from head to head. The movement comes to a halt in the head of ModP because the complementizer cannot cross the base-generated modifier.

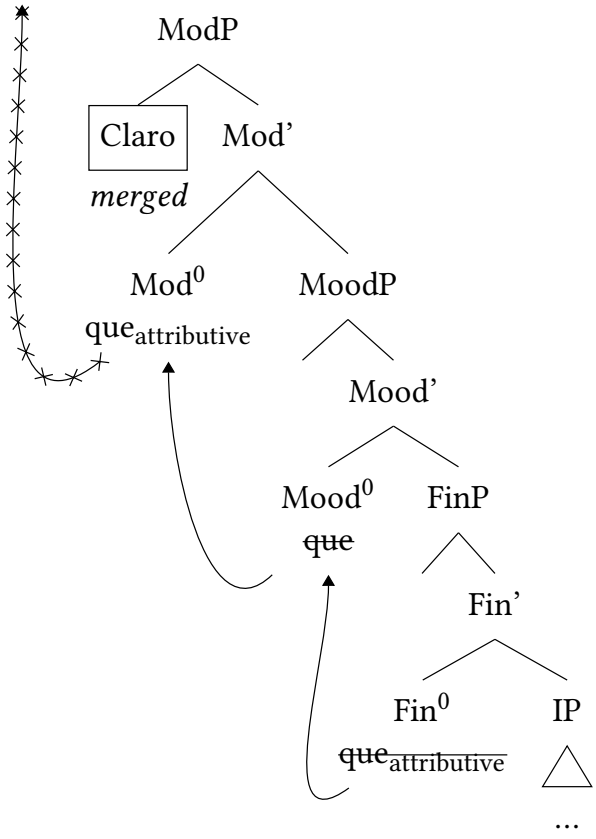


Figure 3.13: AdvC; crossed out lines represent a potential movement that does not take place

Independent of the construction at hand, the merger of these types of modifiers in the left periphery has been proposed previously: ModP was identified by Giorgi (2010) and van Gelderen (2011) as the position in which epistemic and evidential modifiers are base-generated.[13] Based on the universal word order restrictions of co-occurring modifiers described in Cinque (1999), Giorgi (2010) and

[13]The proposal that sentence adverbs are merged in the CP is different from the analysis in

van Gelderen (2011) propose that ModP consists of three sub-projections dedicated to evaluative, evidential and epistemic meanings. For ease of exposition, I still assume that there is just one functional projection, ModP, since co-occurring modifiers expressing these types of meanings are not a concern at present.

Although the current focus is on epistemic and evidential modifiers, since they are most frequent, there are also some attestations of evaluative adverbs in AdvC (cf. 82, 83). This is in line with what the proposal in Giorgi (2010) and van Gelderen (2011) would suggest. Note that these examples, as well as most of the ones I will discuss in this section, are grammatical when attributive *que* is absent. However, the attributive meaning is then lost (cf. §3.4 and §4).

(82) European Portuguese
[$_{\text{ModP}}$ Felizmente que] na Suiça se faz imensas
fortunately QUE in the Switzerland CL.REFL make.3SG.PRS many
coisas sem referendos.
thing.PL without referendum.PL
'Fortunately there are many things happening in Switzerland without a referendum.' (CdP)

(83) Spanish
[$_{\text{ModP}}$ Lamentablemente que] no va a ocurrir.
lamentably QUE not go.3SG.PRS to occur
'Unfortunately, it won't occur.' (CdE)

The idea introduced above allows a compositional analysis, in which the modifiers are merged in the same place irrespective of whether *que* is present or not.

Assuming that epistemic, evidential (and evaluative) modifiers are merged in, rather than moved to, the left periphery explains why attributive *que* follows precisely these types of modifiers but precedes other types. This can be seen in (84) in which the complementizer follows the evidential modifiers *obviamente* and *claro* but precedes *inmediatamente* and *rapidamente*. This word order is expected under the analysis here because the lower adverbs are preposed from an IP-internal position. As moved elements, they do not constitute an obstacle to complementizer movement.

Cinque (1999), who assumes that the projections dedicated to Mod are below FinP. The conception of ModP implicit in Giorgi (2010) and van Gelderen (2011) furthermore differs from that initially proposed by Rizzi (2004b), where ModP is considered to be the position targeted by preposed adverbs which, as he claims, are distinct from left-peripheral topicalized and focalized adverbs.

(84) a. Spanish
[$_{ModP}$ Obviamente que] inmediatamente le surgirá la duda.
obviously QUE immediately him arise.3SG.FUT the doubt
'Obviously, doubt will arise immediately.' (CdE)

b. European Portuguese
[$_{ModP}$ Claro que] rapidamente deixa de ser confortável fazer
clear QUE rapidly stop.3SG.PRS be comfortable make
chamadas com o tablet ao ouvido.
call.PL with the table at.the ear
'Clearly, it quickly stops being comfortable to make calls with a tablet at the ear.' (CdP)

As mentioned in §3.3.1, epistemic and evidential modifiers can appear in short emphatic affirmations and negations introduced by attributive *que* (cf. 45). In this case, the modifiers precede the complementizer, as expected. My analysis for these patterns adopts a polarity projection PolP sandwiched between FinP and IP.

(85) a. Catalan
M' havia venjat jo mai de ningú? [$_{ModP}$ Segur
CL.1SG AUX.3SG.IPFV.PST avenge.PTCP I never of no one sure
que$_i$] [$_{FinP}$ t$_i$] [$_{PolP}$ sí.]
QUE yes
'Had I never sought revenge from anyone? Sure I did.' (ebook-cat)

b. Spanish
Me preguntas si veo todo eso en tus ojos. [$_{ModP}$
CL.1SG ask.2SG.PRS if see.1SG.PRS all this in your eye.PL
Naturalmente que$_i$] [$_{FinP}$ t$_i$] [$_{PolP}$ sí], amada mía.
naturally QUE yes beloved mine
'You ask whether I see all of this in your eyes. Naturally I do, my beloved.' (CdE)

c. Brazilian Portuguese
Mas isso não significa que o veganismo é uma
but this not mean.3SG.PRS that the veganism be.3SG.PRS a
política a se abandonar. [$_{ModP}$ Claro que$_i$] [$_{FinP}$ t$_i$] [$_{PolP}$ não.]
politics to CL.REFL abandon clear QUE not
'But this does not mean that veganism is a policy to abandon. Of course not.' (CdP)

Since the modifier and the complementizer occupy the specifier and the head of the same projection, the present analysis predicts that it should not be grammatical for any left-peripheral material to intervene between them. The ungrammaticality of the intervening topic in (86a) shows that this prediction holds true. A further prediction is that AdvC can be followed and preceded by a topic since here it is merged in neither the lowest nor highest position of the left periphery. The grammaticality of (86b) with a preceding topic and (86c) with a topic following the complementizer shows that this is indeed the case.

(86) Catalan

a. [$_{\text{ModP}}$ *Segur] [$_{\text{TopP}}$ aquest llibre$_i$] que l'$_i$ ha
sure this book QUE CL.M.SG AUX.3SG.PRF.PRS
llegit.
read.PTCP

b. [$_{\text{TopP}}$ Aquest llibre$_i$] [$_{\text{ModP}}$ segur que] l'$_i$ ha
this book sure QUE CL.M.SG AUX.3SG.PRF.PRS
llegit.
read.PTCP

c. [$_{\text{ModP}}$ Segur que] [$_{\text{TopP}}$ aquest llibre$_i$] l'$_i$ ha
sure QUE this book CL.M.SG AUX.3SG.PRF.PRS
llegit.
read.PTCP
'Surely, he indeed has read this book.'

Furthermore, AdvC appears in embedded contexts. In (87a), it appears in the complement of a verb of saying, in (87b), it appears in an appositive relative clause and in (87c) in a restrictive relative clause. This behavior is also predicted by the analysis proposed for AdvC, and moreover constitutes evidence in favor of the assumption that there is more than one merge position accessible to the complementizer in Ibero-Romance.

(87) a. European Portuguese
Disse [$_{\text{SubP}}$ que] [$_{\text{ModP}}$ certamente que] iria ver
say.3SG.PRF.PST that certainly QUE go.1SG.COND see
logo resultados.
soon result.PL
'S/he said that certainly I would see results soon.' (CdP)

b. Spanish
Otra canción [$_{SubP}$ que] [$_{ModP}$ claro que] escuchamos todos y
other song that claro QUE listen-to.1PL.PRS all and
que podría parecer muy buena, es "Realmente no
that could seem.3SG.COND very good be.3SG.PRS Realmente no
estoy tan solo".
estoy tan solo
'Another song that clearly we all listened to and that could seem very good is "Realmente no estoy tan solo".' (CdE)

c. Catalan
És gent [$_{SubP}$ que] [$_{ModP}$ segur que] has vist pero
are people that sure QUE AUX.2SG.PRF.PRS see.PTCP but
mai has passat un cap de setmana amb ells.
never AUX.2SG.PRF.PRS spend.PTCP a weekend with them
'These are people that you surely met but with whom you never spent a whole weekend.' (caWac)

Moreover, AdvC can appear in *que*-initial reportatives as illustrated in example (88) repeated from (34b) in §2.3.1. This again supports the idea that the two instances of *que* have different functions and are merged in distinct positions.

(88) Spanish
[$_{SubP}$ Que] [$_{ModP}$ claro que$_i$] [$_{FinP}$ t$_i$] le viene bien, que qué
QUE clear QUE CL.3SG come.3SG.PRS good QUE what
alegría, que dónde.
joy QUE where
'[reportative:] Of course it was no inconvenience. (She said) what a joy! (She said) where?' (CdE)

Having presented the motivation for the analysis, I will now describe further properties of the construction; for a more complete characterization see Kocher (2014, 2017a, 2018b). One significant feature of AdvC is that it admits both underived modifiers (cf. 89a), which formally coincide with adjectives, and derived modifiers (cf. 89b), which take the adverbial derivational suffix *mente* (cf. also Cruschina & Remberger 2017a, 2018).

(89) Spanish
a. Cierto que la culpa no es suya.
certain QUE the fault not be.3SG.PRS his
'Certainly, it's not his fault.' (CdE)

b. Ciertamente que la culpa no es suya.
certainly QUE the fault not be.3SG.PRS his

In Kocher (2014, 2018b), I presented evidence for the claim that despite these distinct forms, the underlying structure is the same. One central finding outlined in Kocher (2014) is that the variation between the two different morphological forms depends on the language and the modifier (see also §4.3). I propose that they are all sentential adverbs. This is not a controversial assumption for examples like (89b), but requires further justification for examples like (89a). In Kocher (2014, 2017a), I showed that, contra Martín Zorraquino (1998), Freites Barros (2006), and Ocampo (2006), it is not tenable to assume that cases like (89a) result from the ellipsis of a copula construction (*~~Es~~ cierto que...*), which I term EsAdjC. First of all, Spanish, Catalan and Portuguese all allow underived modifiers to express adverbial functions (Hummel 2000, 2017). What is more, when the construction emerged in the 16th century (Kocher 2017a), the underived modifiers were in fact the dominant means of encoding adverbial functions (cf. Hummel 2017). A second argument is that, irrespective of whether the modifier is derived or not, they exhibit a parallel syntactic behavior, suggesting that the underlying structure must be the same. Finally, the ellipsis analysis is also not tenable from an interpretive point of view. It will be shown in §3.4 and §4.3 that the copula construction has a different interpretation than the AdvC construction, which is consistent regardless of the morphological form of the modifier.

AdvC displays further properties that show that it is not merely an elliptical version of a copula construction. First and foremost, derived adverbs would not be expected to be attested. Moreover, the analysis proposed here predicts that any epistemic and evidential modifier merged in ModP can appear in AdvC. Consequently, not only derived and underived adverbs should be allowed, but also epistemic and evidential adverbial locutions. The examples in (90a), (90b) and (90c) show that this prediction holds true.

(90) a. Spanish
Con certeza que crearemos productos muy interesantes en
with certainty QUE create.1PL.FUT product.PL very interesting in
esta nueva etapa.
this new phase
'Certainly, we will create very interesting products in this new phase.'
(CdE)

b. European Portuguese
Sem dúvida que nos Açores há espaço e
without doubt QUE in the Azores there.be.3SG.PRS space and
público para este tipo de eventos!
audience for this type of event.PL
'Without a doubt, there is a space and an audience for this type of event in the Azores.' (CdP)

c. Decerto que o condutor adormeceu.
of-certain QUE the driver fall-asleep.3SG.PRF.PST
'Certainly the driver fell asleep.' (CdP)

Additionally, the fact that only epistemic and evidential (and evaluative) modifiers are permitted constitutes an argument in favor of the non-elliptical nature of the structure. If this were merely a copula construction, other modifiers would be expected. For instance, *fundamental* in (91a) is found frequently in copula constructions but is ungrammatical in AdvC (see 91b).

(91) Brazilian Portuguese

a. É fundamental que a criança explore o seu
be.3SG.PRS essential that a child explore.3SG.PRS the its
mundo.
world
'It is essential that a child explores its world.' (CdP)

b. * Fundamentalmente/ Fundamental que a criança explore
essentially essential that a child explore.3SG.PRS
o seu mundo.
the its world
Intended: 'It is essential that a child explores its world.'

One additional argument is that the modifier in AdvC cannot be itself modified (compare (92a) and (92b); see also Cruschina & Remberger 2017a, 2018). This would once again not be expected if it were just a version of a copula construction.

(92) European Portuguese

a. É bem certo que as rivalidades se semeiam.
be.3SG.PRS good certain that the rivalry.PL CL.REFL spread.3PL.PRS
'It is pretty certain that the rivalries are spreading.' (CdP)

b. * Bem certamente/ certo que as rivalidades se
good certainly certain that the rivalry.PL CL.REFL
semeiam.
spread.3PL.PRS

Furthermore, AdvC is incompatible with morphological negation (cf. 93a and 93b), which further supports the idea that it is more than just an elliptical version of a copula construction. The same fact is also observed in Cruschina & Remberger (2017a, 2018).

(93) Catalan

a. És impossible que es belluguin tant sense
be.3SG.PRS impossible that CL.REFL move.3PL.SBJV.PRS so without
la meva voluntat.
the my will
'It is impossible that they move this much against my will.'
(ebook-cat)

b. * Impossiblement que es belluguin tant sense la
impossibly that CL.REFL fight.3PL.SBJV.PRS so without the
meva voluntat.
my will

With regard to its cross-linguistic distribution, AdvC is attested in all three languages, although there is some variation in the frequency of individual modifiers and whether derived or underived adverbs are preferred (cf. Kocher 2014, 2017a and §4.3).

To close this section, I discuss the development of a new adverb in Catalan that illustrates the productivity of AdvC. Recently, *esclar* (< *és clar*) has emerged as a new evidential adverb (see §4.3 where I interpret experimental results in light of this new adverb). Catalan displays a tendency to grammaticalize new words on the basis of univerbation. A prominent example is *sisplau* 'please', which results from the grammaticalization of conditional *si us plau* 'if you$_{2p}$ please' (see Alturo & Chodorowska 2009 on the grammaticalization of *sisplau*). The process features typical traits of grammaticalization, namely phonological ([siwsplaw] > [sisplaw]) and morphological reduction (loss of second person plural agreement), syntactic reanalysis (protasis of conditional clause > adverb) and semantic change (cf. Hopper & Traugott 2003). Similarly, the Catalan evidential adverb *esclar* resulted from univerbation, originating from EsAdjC (cf. 94).

(94) [$_{CP1}$ És clar [$_{CP2}$ (que) [$_{IP}$ ...]]]

At some point in the development, the parenthetical use of the sequence *és clar* enabled grammaticalization towards an adverb, see (95).

(95) Catalan

a. Jo, és clar, en això no sé què aconsellar-te.
I is clear in that not know.1SG.PRS what recommend-you
'I -that is clear- don't know what to recommend you in this case.' (caWac)

b. Jo no conec mai els meus clients. Ni ells a mi, és clar.
I not know.1SG.PRS ever the my client.PL nor they to me be.3SG.PRS clear
'I never know my clients, nor they know me - that is clear.' (caWac)

Typical features of grammaticalization can also be observed in the case of *esclar* such as phonological reduction ([eskla] > [əskla]) and syntactic reanalysis (clause > adverb).

(96) Catalan

a. Sola no hi he estat mai, esclar, però sé què vull dir.
alone not there AUX.1SG.PRF.PRS be.PTCP ever clearly but know.1SG.PRS what want.1SG.PRS say
'I've never been there on my own, evidently, but I know what it means.' (caWac)

b. No es tracta, esclar, de desmentir l'existència d'una organització social.
not CL.REFL treat.3SG.PRS clearly to deny the existence of a organization social
'Evidently, it's not about denying the existence of a social organization.' (caWac)

The examples in (96) illustrate the syntactic mobility of *esclar*, which parallels the syntactic mobility of epistemic and evidential adverbs.

(97) Catalan
¿Com era possible que una obra mestra [...] tingués un final feliç com aquell? -Esclar que per a Espert
how be.3SG.IPFV.PST possible that a masterpiece have.3SG.IPFV.PST a happy ending like that clearly QUE for E.

el suïcidi de la protagonista també es pot entendre
the suicide of the protagonist also CL.REFL can.3SG.PRS understand
com un final feliç.
as a happy ending.
'How was it possible that a masterpiece had a happy ending like that? - Evidently for Espert the suicide of the protagonist can also be interpreted as a happy ending.' (caWac)

One final piece of evidence demonstrating the adverb-hood of *esclar* is the fact that it can appear in AdvC (97). This development consequently constitutes compelling evidence for the compositional structure of the constructions that I argue for in this chapter. The reasoning is that, if these were fixed grammaticalized expressions, a productive extension of the construction to novel modifiers would not be expected.

3.3.5 *Que* in verum sentences

In this section I focus on a construction in which attributive *que* appears jointly with the verum marker *sí*. I will use the term AffC to refer to this pattern. The construction is not attested in European Portuguese, so the data illustrating it are drawn exclusively from Spanish and Catalan. Explanations for this cross-linguistic contrast will be addressed at the end of this section.

The analysis I adopt here is very similar to the analysis outlined above for AdvC and is illustrated in Figure 3.14. The assumption is once again that the complementizer is merged in FinP where it receives the attributive interface feature. The complementizer then moves to the next head in the hierarchy, the sentence mood head MoodP. The verum marker is directly merged in the specifier of this head. This idea is inspired by Lohnstein (2016), who argues that the verum interpretation results from stress on sentence mood. Consequently, since the verum marker *sí* is merged rather than moved, the complementizer cannot cross it, and therefore the surface word order in which *que* follows *sí* obtains.

Previous syntactic analyses of verum focus and other polarity related phenomena also propose a projection in the same area of the left periphery. The difference is that in these accounts the projection was specialized for polarity or verum (Laka 1990, Martins 2006, Hernanz 2007, Batllori & Hernanz 2008, Rodríguez Molina 2014, Villa-García & Rodríguez 2020a,b). Since the present analysis is grounded in Lohnstein's sentence mood theory of verum focus, which argues that the interpretive effect of verum results from focusing sentence mood, a dedicated projection for verum becomes obsolete.

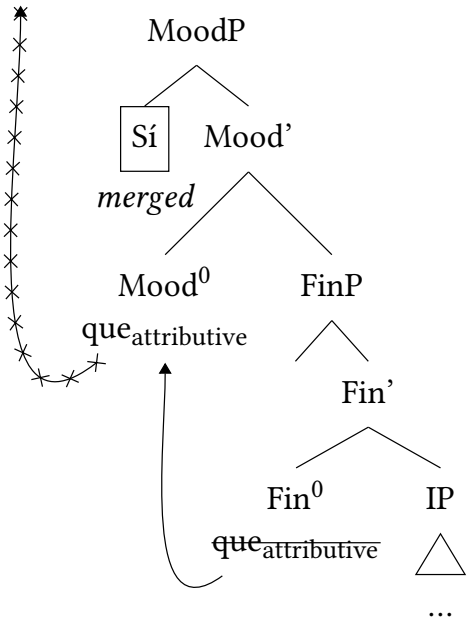


Figure 3.14: AffC; crossed out lines represent a potential movement that does not take place

Lohnstein (2016) uses German data to support his proposal. German is a V2 language, meaning that the finite verb always occupies the second position in a root declarative. This is often explained as the result of movement of the finite verb to a left-peripheral position (cf. Roberts 2001 and references therein). Lohnstein (2016) proposes that the position targeted is a sentence mood head that he calls MoodP. As stated above, Lohnstein's basic idea is that the effect of verum results from focus on sentence mood. A relation between sentence mood and verum has already been observed by Höhle (1992) who offers the first extensive study of verum focus. The main empirical evidence for Lohnstein's idea is given in (98). In German declaratives, verum is marked through stress on the finite verb as in (98a). In embedded declaratives, in which the finite verb is sentence-final, the stress does not fall on the verb, but on the complementizer (98b). Lohnstein (2016) takes this as evidence that the complementizer occupies MoodP, hence why the focus feature is expressed as stress on this element.

(98) German

a. (Lohnstein 2016: 1: ex 1a)
Karl HAT den Hund gefüttert.
Karl AUX.3SG.PRF.PRS the dog feed.PTCP
'Karl DID feed the dog.'

b. (Lohnstein 2016: 2: ex 1e)
Aber Maria glaubt dass Karl in Urlaub gefahren
but Maria believe.3SG.PRS that Karl in holiday drive.PTCP
ist.
AUX.3SG.PRF.PRS
'But Maria believes that Karl DID go on holiday.'

In Kocher (2017a, 2018a), I show that the analysis proposed by Lohnstein (2016) to account for the German data can be extended to account for Spanish. Given that Spanish and Catalan employ the same verum strategy, it is plausible to assume that the analysis can be extended to account for Catalan as well. The basic idea is that the underlying structure is the same for the two Ibero-Romance languages and German. I model this by assuming a focus feature in MoodP. The finite verb remains in the IP in Spanish and Catalan; the focus feature on sentence mood requires lexical material to be expressed, and *sí* is merged directly in the specifier of the projection to satisfy this requirement.

This analysis offers a straightforward explanation for why the verum marker *sí* does not occur in imperatives (cf. 101a vs. 101b). To construct this argument a short digression is required to account for my analysis of directives. The properties of the Ibero-Romance directives are illustrated in (99). The central contrast is between verb- and *que*-initial directives ((99a) repeated from (42a), (99c) repeated from (48a) in §2.3.1 vs. (99b) repeated from (42c) in §2.3.1). They differ in their verbal mood (imperative vs. subjunctive), the position of pronominal clitics (enclisis vs. proclisis) and subjects (post- vs. preverbal). A detailed discussion of their empirical properties can be found in §2.3.1.

(99) Spanish

a. ¡Vete!
go.IMP-CL.2SG
'Go away!'

b. ¡Que se vaya!
QUE CL.3SG go.3SG.SBJV.PRS
'He/She should go!'

c. ¡Váyase!
QUE CL.3SG go.3SG.SBJV.PRS
'He/She should go!'

In order to account for third-person *que*-initial directives in Ibero-Romance, I build on the claim from Demonte & Fernández Soriano (2009) and Rivero & Terzi

(1995) that the complementizer is merged in a left-peripheral position. A difference between Demonte & Fernández Soriano (2009) and the analysis I propose is the location in which *que* is assumed to be merged. Demonte & Fernández Soriano (2014) propose that *que* is merged in FinP to license an imperative feature. I, however, consider Lohnstein's clause typing projection MoodP, located immediately above FinP, to be the more appropriate candidate because FinP is not associated with clause typing in the present analysis. There is also empirical proof exemplified in (100) that shows that *que* can be followed by a clitic left dislocated topic that is assumed to be merged in a left-peripheral topic position. Crucially, this word order is correctly predicted by the structure proposed here, but would not be expected if the complementizer occupied the lowest projection of the left periphery FinP, which does not have a topic projection below it.

(100) Spanish
[$_{\text{MoodP}}$ Que [$_{\text{TopP}}$ los libros$_i$ [$_{\text{FinP}}$ [$_{\text{IP}}$ los$_i$ lean los
QUE the book.PL CL.3PL read.3PL.SBJV.PRS the
niños y no las revistas.]]]]
children and not the magazines
'The children must read the books, not the magazines.'

The basic idea of the analysis is that MoodP encodes the clause type of each sentence through a feature (see Lohnstein 2016 and Kocher 2018a). For the present purposes, I assume the minimal repertoire of a declarative, an interrogative and an imperative feature.[14] While the verbal mood of *que*-initial directives is subjunctive, the clause type is nevertheless imperative (cf. Portner 2004). In line with the empirical data, I suggest that imperative sentences in Ibero-Romance require that the feature is checked directly in MoodP. In *que*-less imperatives the imperative verb itself checks the feature. The data illustrated in (99) and discussed in greater detail in §2.3.1 show that the imperative verb occupies a higher position than indicative verbs in Catalan and Spanish and than subjunctive verbs in all three languages. Unstressed pronouns are enclitic and neither subjects nor negative particles can appear before the imperative verb.

To account for the merger of *que* in *que*-initial third person directives, I adopt the last resort explanation from Rivero & Terzi (1995) and Demonte & Fernández Soriano (2014). In regular imperatives, the verb has the necessary morphological features (=imperative inflection) to check the imperative feature. This triggers

[14]While these three clause types are generally accepted in the literature, some authors argue that there are additional types. See for instance Zanuttini & Portner (2003) who consider exclamatives as a fourth clause type.

the movement of the verb to the projection. In third-person *que*-directives, the subjunctive verb is not equipped with the necessary feature; its movement to the left periphery is therefore inhibited and the verb is unable to check the feature in Mood^0. Since imperatives require a local checking of the feature, as a last resort, the underspecified complementizer is merged directly in MoodP to satisfy this requirement.

Returning now to the issue of the incompatibility of *sí que* and imperatives, the main point is that directives require local checking of the sentence mood feature and that the imperative verb moves to Mood^0. What this means for the verum construction is that in imperatives, there is no need for the merger of *sí* in Mood because the focus feature can be expressed on the verb itself.

(101) Spanish (Kocher 2018a: 27: ex 49)

a. * ¡Sí que cógete una silla de una vez!
VERUM QUE grab.IMP-CL.2SG a chair for once

b. ¡CÓGETE una silla de una vez!
grab.IMP-CL.2SG a chair for once
'GRAB a chair, at once!'

The presence of *que* is not required for Spanish and Catalan verum sentences (cf. 102).

(102) Catalan
Sí (que) ve a la festa.
VERUM QUE come.3SG.PRS to the party
'S/he DOES come to the party.'

I discuss some interpretive differences between verum sentences with and without *que* in Spanish and Catalan in §3.4. With regard to the syntactic structure, the idea of the analysis presented here is again that *sí* is merged in the same place irrespective of whether *que* is present or not. This allows for a compositional analysis. In the literature focusing on *sí* and *sí que*, it was often proposed that *sí* without *que* is merged in the same IP-internal polarity position as sentential negation *no* (see Hernanz 2007, Villa-García & Rodríguez 2020a,b). The data in (103) constitute the principal evidence against this assumption, showing that *sí* cannot occupy this position because it can co-occur with a sentential negation particle.

(103) a. Spanish (Kocher 2017a: 94: ex 31a from CdE)
[$_{TopP}$ Eso] [$_{MoodP}$ sí] [$_{PolP}$ no podía faltar en
this VERUM not can3SG.IPFV.PST miss in
ninguna casa.]
any house
'This could NOT be missed in any house.'

b. Catalan
[$_{TopP}$ Això] [$_{MoodP}$ sí] [$_{PolP}$ no sé fins quan
this VERUM not know.1SG.PRS until when
durarà amb l' ajuntament que tenim.]
take.3SG.FUT with the city hall that have.1PL.PRS
'I do NOT know how long this will take with the city hall we have.' (caWac)

One prediction that follows from the analysis I have presented here is once again that it should not be possible for *sí* and *que* to be separated by other left-peripheral material. The example in (104) shows that this prediction holds true. A clitic left-dislocated topic cannot intervene between *sí* and *que*.

(104) Catalan
* Sí aquesta noia$_i$ que la$_i$ conec.
VERUM this girl QUE CL.F.SG know.1SG.PRS
Intended 'This girl I DO know.'

Because of its position within the left periphery, it should be possible for the sequence of *sí* and *que* to be both followed and preceded by topics. The examples in (105a) and (105b) show that this is indeed the case.

(105) a. Catalan
[$_{TopP}$ Aquesta noia$_i$] [$_{MoodP}$ sí que] la$_i$ conec.
this girl VERUM QUE CL.F.SG know.1SG.PRS
'This girl I DO know.'

b. [$_{MoodP}$ Sí que] [$_{TopP}$ aquesta noia$_i$] la$_i$ conec.
VERUM QUE this girl CL.F.SG know.1SG.PRS
'This girl I DO know.'

The analysis moreover predicts that *sí que* should follow modifiers merged in ModP, which is also what is found in the data.[15]

[15]There are some examples in which the reverse order obtains. Although a definite analysis of these cases is pending further investigation, it is notable that examples of *sí* following AdvC

(106) Catalan
[$_{ModP}$ Segurament] [$_{MoodP}$ sí que] és el millor dissenyat
surely VERUM QUE be.3SG.PRS the best design
a nivell funcional (o això crec jo).
at level functional or this believe.1SG.PRS I
'Surely this IS the best design in terms of functionality (or at least I think so).' (caWac)

Furthermore, AffC, like AdvC, can appear in embedded contexts. In (107a) and (107b), it follows a verb of saying. In (107c), it appears in an appositive relative clause. In (107b), a topic and an epistemic modifier are merged above AffC, demonstrating the availability of structural positions above it.

(107) a. Spanish
Ella le dijo [$_{SubP}$ que] [$_{MoodP}$ sí que] le
she CL.M.SG tell.3SG.PRF.PST that VERUM QUE CL.M.SG
extrañaba.
miss.3SG.IPFV.PST
'She told him that she DID miss him.' (CdE)

b. Catalan
[...] i ha pensat [$_{SubP}$ que] [$_{ModP}$ potser] [$_{TopP}$
and AUX.3SG.PRF.PRS think.PTCP that maybe
d' aquesta dona$_i$] [$_{MoodP}$ sí que] se 'n$_i$
of this woman VERUM QUE CL.REFL CL.PART
podria enamorar.
can.3SG.COND fall in love
'And he has thought that maybe it was true that he could fall in love with this woman.' (ebook-cat)

c. Les tecnologies no són dolentes, el [$_{SubP}$ que] [$_{MoodP}$
the technology.PL be.3PL.PRS not bad the that
sí que] pot ser perjudicial és l' ús que en
VERUM QUE can.3SG.PRS be prejudicial is the use that CL.PART
fem.
make.1PL.PRS
'Technologies themselves are not bad, what CAN be prejudicial is our use of them.' (caWac)

are restricted to cases involving *claro que*. There are indications that *claro que* is undergoing a process of grammaticalization and turning into a fixed expression (Kocher 2014, 2018b, see also §4.3). It might therefore no longer be analyzed compositionally by some speakers, which could be a potential explanation for these examples.

(108) illustrates an example of AffC in a *que*-initial reported sentence. These data show again that the complementizer involved in *que*-initial reportatives has a different function and is merged in a different position than the complementizer present in AffC.

(108) Spanish
A: No vas a venir, ¿pues no?
not go.2SG.PRS to come right

B: Sí que vengo.
VERUM QUE come.1SG.PRS

A: ¿Eh?
huh

B: [$_{SubP}$ Que] …[$_{MoodP}$ sí que] vengo.
QUE VERUM QUE come.1SG.PRS

'A: You're not coming, right? B: I AM coming. A: Huh? B: [reportative:] I AM coming.'

Finally, the data shown in (109), in which *sí que* appears in short emphatic affirmations and negations, constitute further evidence that when expressing verum, *sí* targets a left-peripheral position and is not merged in the lower PolP.

(109) a. Spanish
¡Eso [$_{MoodP}$ sí que$_i$] [$_{FinP}$ t$_i$] [$_{PolP}$ no!]
this VERUM QUE no
'This: no way!'(CdE)

b. Catalan
Ara [$_{MoodP}$ sí que$_i$] [$_{FinP}$ t$_i$] [$_{PolP}$ sí!]
now VERUM QUE yes
'Now: absolutely yes!' (caWac)

I will now turn to the question of cross-linguistic variation. As stated at the start of this section, AffC is not available in Portuguese. According to Martins (2013), there are two alternative strategies to express verum in European Portuguese: Verb reduplication and final or post-verbal *sim*. The verb reduplication strategy, in which the finite verb is doubled, is illustrated in (110).

(110) Portuguese (Martins 2013: 97: ex 1)
A: O João não comprou o carro.
the João not buy.3SG.PRF.PST the car

B: O João comprou o carro comprou.
the João buy.3SG.PRF.PST the car buy.3SG.PRF.PST
'A: João didn't buy the car. B: João DID buy the car.'

The analysis proposed in Martins (2013) is given in (111).

(111) $[_{\text{ToP}}$ [ele comprou$_i$ o carro]$_k$ $[_{\text{Top'}}$ $[_{\text{CP}}$ $[_{C'}$ $[_{C}$ comprou$_i$] ~~$[_{\Sigma P}$ ele$_j$ $[_{\Sigma'}$ comprou$_i$ $[_{\text{IP}}$ $[_{I'}$ comprou$_i$ $[_{\text{VP}}$ ele$_j$ comprou$_i$ o carro$]]]]]_k$~~]]]] (Martins 2013: 101: 8c)

Her proposal is that verb reduplication results from the phonetic realization of two copies of the verbal movement chain. The finite verb moves from the IP to ΣP, a polarity position sandwiched between IP and the left periphery and can thus be identified with PolP. The verb then moves to a CP position, the nature of which is not further specified in Martins (2013). As a final step in the derivation, she assumes remnant movement of the ΣP to the specifier of Top, which she analyzes as an example of IP-topicalization.

Martins (2013) draws a comparison between this and the Spanish and Catalan AffC construction. She proposes that the contrasting strategies result from the allowance vs. disallowance of verbal movement to Σ^0 and the (unspecified) C-position. In her account, this is why the verb reduplication that relies on movement of the verb to these positions is only available in European Portuguese, which allows verb movement of this sort, but not in the other two languages, which do not.

Although I will not offer a detailed account at this point, it seems that this analysis can be integrated fairly straightforwardly into the theory developed in this chapter, simply by proposing that the unspecified C-position in Martins (2013) can be identified with MoodP. It is not only involved in the Spanish and Catalan verum construction: As stated above, this clause-typing projection has been identified as the host of the finite verb in V2 languages and is also the landing site of the high-merged verbs in Ibero-Romance imperatives.

The second alternative strategy that European Portuguese employs according to Martins (2013) is one in which *sim* appears post-verbally. This is exemplified in (112).

(112) Portuguese (Martins 2013: 117: ex 57b)
Comprou sim.
buy.3SG.PRF.PST AFF
'Yes, he DID buy it.'

Final *sim* is not restricted to European Portuguese; its attestations in CdP actually show a greater preference in Brazilian than in European Portuguese. The pattern is furthermore attested in Spanish, and here too, to a larger degree in Latin American varieties than in European Spanish.

Martins (2013) proposes an analysis of final *sim* that makes use of the same mechanisms assumed for verb reduplication. She proposes that *sim* is directly merged in the C-position, the verb moves to Σ^0, and the whole ΣP is again moved to a left periphery topic position resulting in the observed word order in which the finite verb precedes *sim*.

I am cautious about adopting this analysis, because there are multiple attestations of the post-verbal *sim* pattern that cannot be accounted for. The examples in (113) show that *sí/sim* appears adjacent to the finite verb and crucially, precedes the arguments of the verb. This word order is not expected based on the analysis proposed by Martins (2013).

(113) a. Brazilian Portuguese
E ainda esqueceu-se de consultar um dicionário
and yet forget.3SG.PRF.PST-CL.REFL to consult a dictionary
para ver que a palavra presidenta existe sim na
to see that the word presidenta exist.3SG.PRS SIM in the
língua portuguesa.
language Portuguese
'And yet it was forgotten to consult a dictionary to see that the word 'presidenta' DOES exist in the Portuguese language.' (CdP)

b. Columbian Spanish
Escribir novelas tiene sí algo de riesgo.
write novels have.3SG.PRS SI something of risk
'Writing novels DOES carry some risks.' (CdE)

Final *sí/sim* is furthermore compatible with attributive *que*. In (114) this is illustrated by an example where it co-occurs with AdvC.

(114) a. Brazilian Portuguese
Claro que existe sim uma relação.
clear QUE exist.3SG.PRS SIM a relation
'Clearly, there DOES exist a relation.' (CdP)

b. Spanish
Claro que la creatividad debe sí girar entorno a la
clear QUE the creativity should.3SG.PRS SI revolve around of the
música.
music
'Clearly, the creativity SHOULD revolve around music.' (CdE)

There are data that suggest that an altogether different way of accounting for post-verbal *sim* might be called for. There are multiple instances of *sim* surfacing adjacent to different types of XPs. What *sim* appears to mark in these examples is a contrast with a salient (negative) alternative. In (115a) the days the young woman works are contrasted with those she does not work. In (115b) the contrast is between the father who is protective and the mother who is not.

(115) a. Brazilian Portuguese
A jovem trabalha dia sim, dia não.
the young-woman work.3SG.PRS day SIM day NÃO
'The young woman works every other day.' (CdP)

b. Angolan Portuguese
Os seus pais são muito protectores? O meu pai sim,
the your parents be.3PL.PRS very protective the my father SIM
a minha mãe não.
the my mother NÃO
'Are your parents very protective? My father is, my mother isn't.' (CdP)

I believe that the meaning of post-verbal *sim* can be captured in this way as well. The examples in (116) illustrate a contrast noted in Martins (2013) between an answer with *sim* as a regular answer particle and an answer containing post-verbal *sim*. While *sim* followed by a pause is used as an answer to a neutral, unbiased question (116a), in (116b) the final *sim* is used to rebut the salient negative alternative *não comprou* 'he didn't buy'.

(116) Portuguese

a. (Martins 2013: 117: ex 56a-b)
A: Ele comprou o carro vermelho?
he buy.3SG.PRF.PST the car red
B: Sim, comprou.
AFF buy.3SG.PRF.PST
'A: Did he buy the red car? B: Yes, he did.'

b. (Martins 2013: 117: ex 57a-b)
A: Ele não comprou o carro vermelho, (pois não)?
he not buy.3SG.PRF.PST the car red POIS not
B: Comprou sim.
buy.3SG.PRF.PST AFF
'A: He didn't buy the car, did he? B: Yes, he DID.'

Post-verbal *sim* and Spanish/Catalan *sí (que)* appear in overlapping contexts but I believe that they are not the same. Generally, speakers use verum for different purposes: It emphasizes the truth value of a proposition. As such, it can also be used to emphasize agreement with the hearer or to rebut a hearer's previous utterance. Before the seminal work of Höhle (1992), Watters (1979) introduced the term *polar focus* to refer to both of these functions. This type of focus is therefore functionally equivalent to what I call verum in this book. Watters (1979) furthermore introduces the term *counter-assertive focus* to refer to the sub-function of verum that serves to rebut a hearer's previous utterance. Spanish and Catalan *sí (que)* can express both functions. This is illustrated with Spanish data in (117). In (117a), speaker A states that the interlocutors disagreed in some aspect and speaker B emphatically stresses the truth of this statement. In (117b) the sentence preceding the one headed by AffC states that what follows contradicts the interlocutor's belief.

(117) Spanish

a. A: Recuerdo perfectamente que discrepamos en este
remember.1SG.PRS perfectly that disagree.1PL.PRS in this
aspecto. B: Sí que discrepamos.
aspect VERUM QUE disagree.1PL.PRS
'A: I remember perfectly that we disagreed in this respect. B: We DID in fact disagree.' (CdE)

b. No estoy de acuerdo con tí. Sí que hay
not be.1SG.PRS in agreement with you VERUM QUE there.be.3SG.PRS
sistemas de pensiones que son sostenibles.
systemPL of pension.PL that be.3PL.PRS sustainable
'I don't agree with you. There ARE systems of pensions that are sustainable.' (CdE)

It is precisely in counter-assertive contexts like (117b), where a negative version of the finite verb contained in the proposition is salient, that both verum and post-verbal *sim* are predicted to be felicitous. In both cases the effect is that the salient alternative is contrasted and rejected.

To conclude the discussion of the Portuguese data, I now turn to the properties of post-XP *sim* in general. The sequence XP-*sim* can occupy various positions in the syntactic structure, among them a high, potentially left periphery position, as in (118). In this position, it superficially parallels the verum structure in Spanish and Catalan.

(118) Brazilian Portuguese
Isto sim é ser feliz.
this SIM be.3SG.PRS be happy
'This is what being happy means.' (CdP)

Even more curious are the cases where *sim* is followed by a complementizer, giving rise to the sequence *sim que*.

(119) Brazilian Portuguese
Isso sim que é uma idéia completamente errada.
that SIM QUE be.3SG.PRS an idea completely wrong
'It is that idea that is completely wrong.' (CdP)

This could lead to the assumption that these are cases of AffC (Kocher 2019). However, I believe that these examples are more adequately captured as instances of post-XP-*sim* structures. They do not give rise to a verum interpretation of the sentence. As an anonymous reviewer suggested, the meaning of cases like these is best paraphrased not as stress on the truth of the proposition but as a contrast of the XP with a salient alternative. This is very much in line with what I propose above for post-XP *sim* in general. Additionally, if it were a verum structure, we would expect a similar distribution as in Spanish and Catalan. For one thing, we would expect to find *sim que* at the beginning of a sentence. However, this pattern is virtually absent in the corpus (CdP), whereas there are numerous examples of the sentence-initial sequence of a nominal XP followed by *sím que*.

Although a thorough analysis is pending, the sequence *sim que* in examples like (119) could be the result of an elision process.[16] Mioto & Lobo (2016) mention that in Brazilian but not European Portuguese the copula in an inverted cleft can be dropped, as shown in (120).

(120) a. European/Brazilian Portuguese (Mioto & Lobo 2016: 287: ex 42c)
João é que pescou esse peixe.
João be.3SG.PRS that fish.3SG.PRF.PST this fish
'It was João that fished this fish.'

b. Brazilian Portuguese (Mioto & Lobo 2016: 287: ex 43a)
João que pescou esse peixe.
João that fish.3SG.PRF.PST this fish
'It was João that fished this fish.'

[16] I am grateful to an anonymous reviewer who made this insightful proposal during the reviewing process of this book.

An explanation for the *sim que* examples in (119) is then that they result from a similar case of copula drop. What happens is that the XP is focused via an inverted cleft construction and additionally a contrast is marked with the particle *sim*. This "full" structure is illustrated in (121). As expected, it is more frequent in European Portuguese than in Brazilian Portuguese in the corpus data.

(121) a. European Portuguese
Isso sim é que é uma grande notícia.
that SIM be.3SG.PRS that be.3SG.PRS a big news
'That is what big news is.' (CdP)

b. Brazilian Portuguese
Isto sim é que é justiça.
this SIM be.3SG.PRS that be.3SG.PRS justice
'This is what justice means.' (CdP)

Finally, the sequence *sim que* emerges via the elision of the copula. Although this pattern is in fact more frequent in Brazilian Portuguese (122a), there are still examples in European Portugueses (122b) in the corpus.

(122) a. Brazilian Portuguese
Isso sim é que é capitalismo tardio.
that SIM be.3SG.PRS that be.3SG.PRS capitalism late
'That is what late capitalism is.' (CdP)

b. European Portuguese
Isso sim é que era trabalhar.
that SIM be.3SG.PRS that be.3SG.IPFV.PST work
'That is what work was.' (CdP)

To conclude the discussion of the syntactic properties of attributive *que*, I briefly summarize the major points. In the preceding sections, I provided the empirical evidence in support of my unified analysis of the different constructions involving attributive *que*. I showed that when the complementizer appears sentence-initially it is plausible to assume that it reaches the highest position in the left periphery SubP (see §3.3.1, §3.3.2). In *wh*-exclamatives, AdvC and AffC, the complementizer and the preceding material always observe strict adjacency, suggesting that the complementizer moves to the head of the respective projection. I studied the word order of attributive *que* with respect to other left-peripheral material. These data corroborated the idea that the movement of the complementizer is only inhibited by material merged in the left periphery but

can cross moved phrases. In §3.4, I will present a more concise characterization of the discourse contribution of attributive *que*. The main argument is that there is a shared meaning present in all the different constructions. Different interpretive effects depend on the clause type and the meaning of the elements with which attributive *que* co-occurs in each construction.

3.4 The pragmatics of attributive *que*

In this section I focus on the meaning of attributive *que*. All the constructions are attributive in the sense of Poschmann (2008). This means that the commitment expressed is attributed to someone other than the speaker. In the present case, the central function of attributive *que* is to ascribe to the hearer a commitment to the proposition in the scope of the complementizer. In other words, the proposition is presented as something that in the view of the speaker, the hearer should believe to be true (cf. also Gras 2016 for a similar idea). This is illustrated by the contrast between the examples in (123) and (124).

(123) Spanish

A: Cual es la falsa idea que tienen los
what be.3SG.PRS the false idea that have.3PL.PRS the
doctorandos al inicio de sus estudios?
PhD student.PL at.the beginning of their study.PL

B: Que seguramente acabarán su tesis a tiempo.
that surely finish.3PL.FUT their thesis on time

B': # Que seguro que acabarán su tesis a tiempo.
that sure QUE finish.3PL.FUT their thesis on time

'A: What is the false belief PhD students have at the beginning of their studies? B: That surely (#que) they will finish their thesis on time.'

In (123), version B' of the answer involving AdvC is not acceptable while version B, with the same epistemic modifier but no attributive *que*, is perfectly acceptable. Since attributive *que* marks the proposition as something that is considered to be true by the hearer, it is incompatible in the context of this example in which the proposition is declared as a false belief by interlocutor A. Therefore, the belief of *p* cannot be attributed to interlocutor A, the hearer because he has just stated that he does not consider *p* to be true. Version B of the answer is fine because the epistemic modifier is not attributive in itself.

(124) Spanish

A: Qué dicen los doctorandos al inicio de sus
what say.3PL.PRS the PhD students at.the beginning of their
estudios?
study.PL

B: Que seguramente acabarán su tesis a tiempo.
that surely finish.3PL.FUT their thesis on time

B': Que seguro que acabarán su tesis a tiempo.
That sure QUE finish.3PL.FUT their thesis on time

'A: What do PhD students say at the beginning of their studies? B: That surely they will finish their thesis on time.'

In the context presented in (124), both versions of the answer are acceptable because speaker A inquires about what PhD students say at the beginning of their studies, but does not declare or retract his commitment to the truth of the proposition.

In what follows, I show that this function of attributing the belief of the unmodified sentence, called the prejacent, to the hearer is consistent in all the constructions containing attributive *que*. There are different nuanced effects in the various constructions which, I argue, result from the clause type and the function of the other elements involved in them.

Before going into the detail of my account, I will briefly mention that similar observations have been made in the literature. Bianchi & Frascarelli (2017), for instance, note that the propositional content of certain embedded clauses, for instance the complements of factive verbs, is imposed on the common ground. They use the term *informative presupposition*, going back to Prince (1978).

Another concept that might seem relevant to the phenomena being discussed here is presupposition accommodation (Beaver 2001, Simons 2003, Beaver & Zeevat 2007, von Fintel 2008 and references therein). However, propositions that are added to the common ground by way of accommodation are usually not at issue in contrast to attributive *que*-initial propositions (cf. 126).

AnderBois et al. (2010, 2015) offer an analysis that seeks to capture the contrast between referential and appositive relative clauses. Similar to the observation by Bianchi & Frascarelli (2017), they propose that the propositional content of appositives is imposed on the common ground. Lohiniva (2017) links this function to the complementizer in the French constructions involving subordinating *bien que* 'although' and coins the term *impositive complementizer*. A core aspect of

the analysis by AnderBois et al. (2015) is that these propositions are not at issue. Among other things, one consequence of this information status is that the proposition imposed on the common ground cannot be denied by the hearer. This explains the oddness of (125b) as a reaction to (125a) because it directly denies the content of the appositive clause.

(125) a. His husband, who had prostate cancer, was being treated at the Dominican Hospital. (AnderBois et al. 2015: 116: ex 53a)

b. ?? No, he had lung cancer. (AnderBois et al. 2015: 116: ex 53b)

c. No, he was being treated at the Standford Hospital. (AnderBois et al. 2015: 116: ex 53c)

The same is not true for propositions headed by attributive *que*. They are at issue and thus can be denied. Therefore the reaction of B, rejecting the AdvC-proposition introduced by A, is perfectly acceptable.

(126) Brazilian Portuguese

A: Seguro que o Pedro chegou à hora.
sure QUE the Pedro arrive.3SG.PRF.PST on.the time

B: Não, chegou atrasado.
no arrive.3SG.PRF.PST late

'A: Surely Peter arrived on time. B: No, he arrived late.'

At present, this crucial contrast between presupposition accommodation or imposition such as in the phenomena explored by AnderBois et al. (2015) and the phenomena under investigation here prevents me from systematically relating this analysis to my own proposal. Future research will show whether and how these approaches can be reconciled.

As stated above, the shared meaning that I propose for all attributive *que* constructions is that the speaker attributes a commitment to the proposition to the hearer. In many cases, in an attributive *que* sentence the speaker revisits a proposition that has been asserted previously. An example of this type is given in (127). Speaker E asserts the same content that is then taken up again by speaker M in the proposition introduced by AdvC. It is used as a tool to emphasize that the commitment to the proposition is shared by both interlocutors.

(127) European Portuguese
E. - Isso é uma aposta e o governo pode sair
this be.3SG.PRS a gamble and the government can.3SG.PRS leave
perdendo. M. - Claro que é uma aposta.
losing clearly QUE be.3SG.PRS a gamble
'E. - This is a gamble and the government could end up losing. M. Clearly, this is indeed a gamble.' (CdP)

Attributive *que* is however not limited to propositions to which the hearer has expressed a previous commitment. It can also be a way for the speaker to persuade the hearer to accommodate a proposition. Attributive *que* can thus be used to speed up the conversation or to present propositions with potentially controversial content as if they were uncontroversial. The speaker, thereby, anticipates the hearer's commitment and leaves him less room to debate or reject it. This is what happens in (128), where AffC introduces a proposition that contains a potentially controversial opinion the speaker holds about the hearer.

(128) Catalan
No em valores prou. No em fas cas. A
not CL.1SG appreciate.2SG.PRS enough not CL.1SG make.2SG.PRS case to
ell sí que n' hi faries.
he VERUM QUE CL.PART CL.3SG make.2.SG.COND
'You don't appreciate me enough. You don't pay heed to me. To him, you DO pay heed.' (ebook-cat)

In other contexts, attributive *que* is used despite the fact that the hearer has explicitly expressed an opposing commitment. The effect is that the speaker communicates that she was under the impression that the proposition was part of the common ground, but the hearer's recent linguistic or extra-linguistic behavior contradicts the idea that the hearer was aware of the proposition. In these contexts, attributive *que* can serve to express the speaker's surprise regarding the hearer's ignorance or rejection of the proposition, as the speaker considered it to be obviously true.

A context in which the attributive *que*-initial sentence rebuts an opposing hearer-commitment is exemplified in (129). In this case, the hearer has explicitly stated that he is committed to the fact that there are six feet. María's use of attributive *que* twice in her last enunciation expresses her surprise about the hearer's incorrect belief and she emphatically insists on the truth that there are actually four feet and that, in her opinion, the hearer should have been aware of this fact.

(129) Spanish
"Oye, María. Yo creo que aquí hay seis pies." "¿Pero
listen María I believe.1SG.PRS that here be.3PL.PRS six foot.PL but
qué dices?" "Yo estoy convencido de que aquí hay seis
what say.2SG.PRS I be.1SG.PRS convinced of that here be.3PL.PRS six
pies." "¡Que no, hombre, que sólo hay cuatro!"
foot.PL QUE no man QUE only be.3PL.PRS four.
'"Listen, María. I think that here are six feet." "But what are you saying?"
"I am convinced that here are six feet." "Absolutely not, man! There are
only four!"' (CdE)

For the characterization of the pragmatic effect that the presence of attributive *que* has in the different constructions, I adopt the implementation presented in Malamud & Stephenson (2015) of a conversational scoreboard in the style of Lewis (1979). Their version constitutes a modification of the model proposed in Farkas & Bruce (2010) that builds on Hamblin (1971), Gunlogson (2003), Ginzburg (2012) and others, and is further developed in Roelofsen & Farkas (2015). Scoreboards permit a dynamic modeling of conversations. In these models, the speaker keeps track of information states. When making a conversational move, like asserting a proposition, certain aspects of the informational states on the scoreboard change. In the original version in Farkas & Bruce (2010), the conversational states consist of discourse commitments (DC_X) for each conversational participant X, which are sets of propositions that participant X is committed to. The concept of a common ground (CG) is adopted from Stalnaker (1978) and refers to sets of propositions that the participants share a commitment to. The CG can therefore be viewed as the intersection of the DC of all the contextually relevant participants. There is furthermore a table, which is a similar concept to *question under discussion* (Ginzburg 1996, Roberts 1996, Engdahl 2006). It contains an ordered stack of propositions or issues to be resolved in the conversation. There is also a projected CG, which is a set of potential future CGs given the possible resolutions of the top issue on the table. The central modification of Malamud & Stephenson (2015) is that there are projected versions not only of the CG but also of other parts of the scoreboard. Particularly relevant to the phenomenon under discussion are the projected DCs for each participant. This allows the speaker to give tentative commitments by adding a proposition to the projected rather than the current DC. The projected DC of the speaker, according to Malamud & Stephenson (2015), coincides roughly with the concept of contingent commitment in Gunlogson (2008). In Gunlogson's system, however, there is no correspondence with the projected hearer commitment. In Malamud & Stephenson

(2015), there is also a projected table, which allows an issue to be tentatively raised for resolution in future moves.

To exemplify how these models work, the conversational scoreboard of an assertion is represented in Table 3.5. In the system proposed by Malamud & Stephenson (2015), when a proposition is asserted, it is added, along with its presuppositions, to the top of the stack on the table, the speaker's DC and the projected CG. What this means is that when asserting a proposition, the speaker raises it as an issue she seeks to resolve, signals her commitment to its truth and proposes to add it to the CG.

Table 3.5: Conversational scoreboard when asserting p

current		projected	
CG	$\{\}$	CG*	$\{\{p\}\}$
$DC_{Speaker}$	$\{p\}$	$DC^*_{Speaker}$	$\{\{\}\}$
DC_{Hearer}	$\{\}$	DC^*_{Hearer}	$\{\{\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\rangle\rangle$

With these general ideas in place, I now turn to the constructions under investigation and show how the discourse contribution of attributive *que* can be modeled in a conversational scoreboard à la Malamud & Stephenson (2015). As stated at the beginning of this section, the function of *que* is attributive in the sense of Poschmann (2008), meaning that a commitment to the proposition is ascribed to the hearer. Translating this into the scoreboard model, I propose that in all the different constructions containing attributive *que*, the proposition p is added to the hearer's DC. The contribution of *que* is thus the same in all the constructions. What is different in polar questions is that the proposition is added to the projected DC of the hearer and not the current one as is the case in declaratives.

Table 3.6 illustrates the basic case of attributive *que*-initial declaratives. They have all the same conversational states as normal assertions. Therefore, the proposition is added to the speaker's DC, to the table and to the projected CG. The presence of attributive *que* means that the proposition is furthermore added to the hearer's DC. The resulting effect is that the speaker conveys that, in her view, the commitment to the proposition is shared by the hearer.

Attributive *que*-initial declaratives often appear juxtaposed with directive sentences. In these contexts, the sentence headed by *que* is sometimes interpreted as the reason or explanation for the previous assertion or command. This accounts

Table 3.6: Conversational scoreboard when asserting attributive *Que p*

current		projected	
CG	$\{\}$	CG*	$\{\{p\}\}$
$DC_{Speaker}$	$\{p\}$	$DC^*_{Speaker}$	$\{\{\}\}$
DC_{Hearer}	$\{p\}$	DC^*_{Hearer}	$\{\{\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\rangle\rangle$

for the assumption made by some authors in the literature that the function of *que* in these contexts is to express a causal relation (cf. Alarcos Llorach 1994, Porroche Ballesteros 2000, Peres & Mascarenhas 2006, Etxepare 2013, Wheeler et al. 1999, Cunha & Cintra 1984, Lobo 2003, Lopes 2012, Colaço & Matos 2016). However, I refrain from relating this causal meaning to the complementizer itself because a causality interpretation is available even in the absence of *que*. Even without a special marker, the interpretation that two juxtaposed sentences are causally related is available.

(130) a. Catalan
No menges pipas. (Que) fa castellà.
not eat.2SG.PRS sunflower seed.PL QUE make.3SG.PRS Castillian
'Don't eat sunflower seeds. That makes you (look) Spanish.'
(ebook-cat)

b. Portuguese
Keep calm. (Que) hoje é sexta-feira.
keep calm QUE today be.3SG.PRS Friday
'Keep calm. Today is Friday.'

In (130a) the second sentence, irrespective of whether *que* is present or not, can be interpreted as an explanation for why you should not eat sunflower seeds. Similarly, in (130b), both versions of the second sentence could be interpreted as a reason why you should keep calm. Therefore, the causal function is independent of attributive *que*. If attributive *que* is present in an example with a causal interpretation, nothing changes about *que*'s general function: The proposition introduced by the complementizer is presented as uncontroversial information that, according to the speaker, both speaker and hearer are committed to.

There are cases where the *que*-initial declarative follows a directive, in which a causal interpretation is not consistent. One such example is given in (131).

(131) Spanish
Salid. Que no os mataré.
leave.IMP.2PL QUE not CL.2PL kill.1SG.FUT
'Come out. I won't kill you.'

In this case, the causal interpretation seems to be lost. A direct causal relation between following the order by coming out of a hiding place and not getting killed cannot be constructed. It can also hardly be understood as a causal relation on a higher level, where not killing the hearers would be the reason for saying *Salid*.[17] The motivation for the second sentence is to reassure the hearers and make them believe that there is no imminent threat to their lives. By adding attributive *que* at the beginning of the sentence, the speaker anticipates the hearers' agreement with it, which makes it sound more persuasive and less threatening.

My proposal for AdvC-declaratives is that they involve the same conversational states as attributive *que*-initial declaratives, illustrated in Table 3.7. The only difference is that the speaker's DC contains the modified rather than the bare proposition. In turn, the unmodified proposition is added to the hearer's DC, the table and the projected CG. The reason for this discrepancy is that the speaker's commitment to the prejacent depends on the modifier. Different epistemic and evidential modifiers imply different degrees of commitment towards the truth of the proposition. Weak epistemic modals like *probably* allow the speaker to retract her commitment to the proposition that it modifies. Similarly, certain types of evidential modifiers like *clearly* imply that the speaker is committed to the truth of the proposition while others like *allegedly* do not.

Table 3.7: Conversational scoreboard when uttering *Claro que p*

current		projected	
CG	$\{\}$	CG*	$\{\{p\}\}$
$DC_{Speaker}$	$\{claro(p)\}$	$DC^*_{Speaker}$	$\{\{\}\}$
DC_{Hearer}	$\{p\}$	DC^*_{Hearer}	$\{\{\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\rangle\rangle$

To understand the effect of attributive *que* in combination with evidential and epistemic modifiers, it is useful to contrast AdvC with similar constructions that

[17] An explicative interpretation, however, whereby not killing the hearers could be the explanation for why they should come out, is possible.

involve the same epistemic and evidential modifiers but no attributive complementizer (see Kocher 2018b and §4.3 for experimental and corpus studies focusing on the three constructions; see also Cruschina & Remberger 2018 for a similar contrastive characterization of the three constructions). The examples in (132) illustrate the constructions for comparison here.

(132) a. Spanish
Obviamente/ Ciertamente que Pedro viene a la fiesta.
obviously/ certainly QUE Pedro come.3SG.PRS to the party
'Obviously/Certainly, Pedro will in fact come to the party.'

b. Catalan
Obviament/ Certament en Pere ve a la festa.
obviously/ certainly the Pere come.3SG.PRS to the party
'Pere will obviously/certainly come to the party.'

c. Portuguese
É óbvio/ certo que o Pedro vem à festa.
is obvious/ certain that the Pedro come.3SG.PRS to.the party
'It is obvious/certain that Pedro will come to the party.'

In (132a), the modifier appears in AdvC, containing attributive *que*. In (132b) it functions as a simple adverb and in (132c), which I call EsAdjC, the modifier appears as a predicative adjective in a copula construction.

The contrast that I consider central for my argument is the different readings of the modifiers that are rendered prominent. Von Fintel & Gillies (2011) introduce the term *bare epistemic modals* to refer to epistemic modals that permit different readings depending on their modal base. The modal bases are built on the information states of different speech participants and are grounded in knowledge, beliefs or evidence. The key point here is that with bare epistemic modals the evaluation expressed by the modal can be anchored to different deictic centers. The Ibero-Romance epistemic modifiers involved in the three constructions can be classified as bare epistemic modals. This is illustrated for the cognate of *probably* in (133). A statement like *Peter will probably come* can be interpreted from the perspective of different speech participants and thus anchored to different deictic centers.[18]

[18] A similar concept is "(inter)subjectivity" (Nuyts 2001, Traugott 2010, going back to Lyons 1977), which is also used to refer to the fact that certain constructions or expressions are anchored to different speech participants. See also Kocher (2018b) for an application of this concept to Spanish and Portuguese modifiers.

(133) a. (*deictic center=speaker*)
(In the view of what I know) Peter will probably come.
b. (*deictic center=hearer*)
(In the view of what you know) Peter will probably come.
c. (*deictic center=speaker and hearer*)
(In the view of what you and I know) Peter will probably come.
d. (*deictic center=a (contextually relevant) group or authority*)
(In the view of what is generally known) Peter will probably come.

The idea maintained in von Fintel & Gillies (2011) is that bare epistemic modals are ambiguous by design. This means that when uttering *probable*(*p*), the speaker puts into play all contextually relevant readings of *probable*(*p*) that can then be either taken up or rejected. This is illustrated in example (134), where Alex uses the bare epistemic modal *might*, and the two acceptable continuations that follow Alex's statement pick up different readings of the modal. In the reaction in (134a), Billy takes up the reading of the modal anchored to the speaker Alex, while in the reaction in (134b), he takes up and rejects a reading anchored to a perspective that involves Billy, the hearer, i.e. himself.

(134) Alex is aiding Billy in the search for her keys:
Alex: You might have left them in the car.
a. Billy: You're right. Let me check.
b. Billy: No, I still had them when we came into the house.
(von Fintel & Gillies 2011: 114–115: ex 12–14)

While von Fintel & Gillies (2011) are only concerned with different readings of epistemic modals, there is reason to believe that evidential evaluation can similarly be tied to different speech participants.[19] I therefore also propose that when the speaker utters *obvious*(*p*) all the contextually relevant readings of *obvious*(*p*) are put into play (cf. 135).

(135) a. (In the view of what I know) Peter will obviously come.
(*deictic center=speaker*)
b. (In the view of what you know) Peter will obviously come.
(*deictic center=hearer*)

[19] I make no claims here about the relation between epistemic modality and evidentiality, but see van der Auwera & Plungian (1998), de Haan (1999), Aikhenvald (2004), Palmer (2001), Rooryck (2001), Faller (2006), Diewald & Smirnova (2010), von Fintel & Gillies (2010), Boye (2012), Matthewson (2015) for some prominent ideas, and see Cornillie (2009) and Kocher (2014) for an overview of the different positions.

c. (In the view of what you and I know) Peter will obviously come.
(*deictic center=speaker and hearer*)

d. (In the view of what is generally known) Peter will obviously come.
(*deictic center=a (contextually relevant) group or authority*)

As stated above, I consider the epistemic and evidential modifiers involved in the relevant constructions to be bare in the sense of von Fintel & Gillies (2011), in that they can be anchored to different deictic centers. The ambiguity is retained when they appear as adverbs as in (132b), but specific readings are made prominent when they appear in either AdvC or EsAdjC. My proposal is that the prominent reading of the modifiers in AdvC can be linked to the discourse contribution made by attributive *que*. Since the proposition is presented as something the hearer is committed to, by way of an implicature the interpretation arises that he also shares the epistemic or evidential evaluation. This results in the reading of the modifier that is characterized in (133c) and (135c). In the simple adverbial construction, the commitment to p is not attributed to the hearer nor to any participant other than the speaker and therefore the implicature does not arise automatically. The adverbs do not have a prominent reading, meaning that different readings are available depending on the context. Finally, EsAdjC is also attributive. However, the commitment to the proposition is not attributed to the hearer, but to a general group of people or an authority. In this case, again, an implicature arises resulting in the evaluation being presented as shared by the speaker and a general group of people, i.e. the reading of the modifier exemplified in (133d) and (135d). It is important to stress that the readings are not set in stone. In fact, even if a construction makes a specific reading prominent, speakers are able to accommodate other readings. This is strongly suggested by the results of the experiment in §4.3, which indicate that the prominent reading of the modifiers can be overridden if contextual factors make other readings available.

To conclude this section on AdvC, I will briefly discuss the conversational scoreboards that I assume for pure adverbial modification and for EsAdjC and point out how they contrast with the conversational scoreboard assumed for AdvC. Table 3.8 shows the conversational scoreboard when uttering a proposition modified by an evidential adverb. It parallels the scoreboard assumed for AdvC. The speaker is committed to the modified proposition and the prejacent, i.e. the unmodified proposition, is put on the table and added to the projected CG. In the absence of attributive *que*, the attributive addition of the proposition to the hearer's DC is not part of the scoreboard.

Table 3.8: Conversational scoreboard when uttering *Claramente p*

current		projected	
CG	$\{\}$	CG*	$\{\{p\}\}$
$DC_{Speaker}$	$\{claro(p)\}$	$DC^*_{Speaker}$	$\{\{\}\}$
DC_{Hearer}	$\{\}$	DC^*_{Hearer}	$\{\{\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\rangle\rangle$

Table 3.9 shows the conversational scoreboard that I assume for EsAdjC, the equivalent biclausal copula construction in which a proposition is modified by a predicative adjective. I propose that these are also attributive constructions, with the commitment not ascribed to the hearer but to a general group of people or an authority. The states involved are much the same as in AdvC, illustrated in Table 3.7, except that the proposition is not added to the hearer's DC. Instead, I propose that the speaker also keeps track of a set of general discourse commitments ($DC_{General}$), to which the proposition is added in these constructions.

Table 3.9: Conversational scoreboard when uttering *Está claro p*

current		projected	
CG	$\{\}$	CG*	$\{\{p\}\}$
$DC_{Speaker}$	$\{claro(p)\}$	$DC^*_{Speaker}$	$\{\{\}\}$
DC_{Hearer}	$\{\}$	DC^*_{Hearer}	$\{\{\}\}$
$DC_{General}$	$\{p\}$	$DC^*_{General}$	$\{\{\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\rangle\rangle$

In what follows I characterize the discourse effects of attributive *que* in *wh*-exclamatives and verum sentences. Before going into detail, I should stress that the conversational scoreboards that I will shortly present might be subject to further refinement in the future. This is because the information states involved in verum sentences and exclamatives are more complex than in attributive *que*-initial declaratives, polar questions and AdvC. Further research is certainly required to be able to offer all encompassing analyses.

Exclamatives are said to be factive (cf. Grimshaw 1979, Zanuttini & Portner 2003, Castroviejo 2006). This means that the information they contain is presupposed to be true. They furthermore cannot be used as answers to a question (cf. 136a) but they can be confirmed (cf. 136b).

(136) Catalan

a. A: Com és d'alt en Pau? B1: Molt alt. B2: Fa 1.90m. B3:
how be.3SG.PRS of tall the Pau very tall makes 1.90m
#Que alt que és!
how tall QUE be.3SG.PRS
'A: How tall is Pau? B1: Very tall. B2: He's 1.90m tall. B3: #How tall he is!.'

b. A: Que alt que és en Pau! B1: #Fa 1.90m. B2: I tant!
how tall QUE be.3SG.PRS the Pau makes 1.90m and so
'A: How tall Pau is! B1: #He's 1.90m tall. B2: Indeed!'

According to Gunlogson (2003) and Castroviejo (2006), the speaker commits herself to the descriptive content of the sentence, but does not assert it. In Castroviejo (2006) it is stated that exclamatives denote the fact that an individual has the property expressed by the *wh*-expression to a high degree. The speaker furthermore expresses her attitude towards the degree and, according to Castroviejo (2006), it is precisely this information that is used to update the CG. The scoreboard I propose for a *que*-less *wh*-exclamative like (137b) repeated from (74b) is given in Table 3.10.

(137) Portuguese

a. Que chato que és.
how annoying QUE be.2SG.PRS

b. Que chato és.
how annoying be.2SG.PRS
'How annoying you are.'

Table 3.10: Conversational scoreboard when uttering *Qué x p*

current		projected	
CG	$\{p\}$	CG*	$\{\{\,\}\}$
$DC_{Speaker}$	$\{Excl(p)\}$	$DC^*_{Speaker}$	$\{\{\,\}\}$
DC_{Hearer}	$\{\,\}$	DC^*_{Hearer}	$\{\{\,\}\}$
Table	$\langle\,\rangle$	Table*	$\langle\langle\,\rangle\rangle$

The proposition and the exclamative import are added to the speaker's DC.[20] As stated above, exclamatives do not assert the prejacent, i.e. the underlying proposition. In the model, I represent this by not putting the proposition on the table nor into the projected CG. Exclamatives are factive and hence presuppose the truth of the prejacent. I propose to model this by imposing the proposition on the current CG.

The main argument of this section is that the discourse contribution of the attributive complementizer is the same in all the constructions it involves. Therefore, in *wh*-exclamatives with *que* like (137a) repeated from (74a), the consequence of the presence of *que* is again that the proposition is added to the hearer's DC. Apart from that, the scoreboard illustrated in Table 3.11 is the same as the one that I assume for *que*-less exclamatives.

Table 3.11: Conversational scoreboard when uttering *Qué x que p*

current		projected	
CG	$\{p\}$	CG*	$\{\{\}\}$
$DC_{Speaker}$	$\{Excl(p)\}$	$DC^*_{Speaker}$	$\{\{\}\}$
DC_{Hearer}	$\{p\}$	DC^*_{Hearer}	$\{\{\}\}$
Table	$\langle\,\rangle$	Table*	$\langle\langle\,\rangle\rangle$

The CG is by definition the set of all discourse commitments shared by the interlocutors. Consequently, attributive *que* imposes something that is already part of the meaning of a *wh*-exclamative even in the absence of *que*. This might explain why many authors consider *que* to be optional and to have no effect on the interpretation of exclamatives (see for instance Villalba 2003, Castroviejo 2006). In those *wh*-exclamatives that permit versions without the complementizer, what we expect is that the version with *que* should place more emphasis on the hearer's commitment.

Verum sentences are similar in certain respects. Unlike exclamatives, they are asserted and can function as answers to questions (cf. 138).

(138) Catalan
A: És alt en Pau? B: Sí (que) ho és!
be.3SG.PRS tall the Pau VERUM CL.N be.3SG.PRS
'A: Is Pau tall? B: He IS (indeed)!'

[20] The semantic derivation of *wh*-exclamatives is not a concern here, but cf. Zanuttini & Portner (2003), Castroviejo (2006), Villalba (2008), Gutiérrez-Rexach & Andueza (2011, 2016) and references therein.

In the model, this means that the speaker does put the proposition on the table. However, like exclamatives, verum sentences are factive and thus presuppose the truth of the prejacent. In my proposal this is modeled as an imposition of the proposition on the CG.

(139) Spanish

a. Pablo sí que es alto.
Pablo VERUM QUE be.3SG.PRS tall

b. Pablo sí es alto.
Pablo VERUM be.3SG.PRS tall
'Pablo IS tall.'

The scoreboard for a *que*-less verum sentence like (139b) is given in Table 3.12. The verum marked proposition is part of the speaker's DC. The prejacent is put on the table and imposed on the current CG.

Table 3.12: Conversational scoreboard when uttering *Sí p*

current		projected	
CG	$\{p\}$	CG*	$\{\{\,\}\}$
$DC_{Speaker}$	$\{verum(p)\}$	$DC^*_{Speaker}$	$\{\{\,\}\}$
DC_{Hearer}	$\{\,\}$	DC^*_{Hearer}	$\{\{\,\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\,\rangle\rangle$

Table 3.13 illustrates the conversational scoreboard that I assume for *sí que*-sentences like (139a). The presence of the attributive complementizer again highlights an aspect that is already present in the interpretation of the verum sentence by virtue of the fact that whatever is part of the CG is also part of the hearer's DC. Again, a consequence of this proposal is that sentences with *que* should be interpreted as being more insistent and emphatic.

The final part of this section is dedicated to the discourse contribution of attributive *que* in polar questions like (140).

(140) Catalan
Que tens pressa?
QUE have.2SG.PRS stress
'Are you stressed out?' (ebook-cat)

Table 3.13: Conversational scoreboard when uttering *Sí que p*

current		projected	
CG	$\{p\}$	CG*	$\{\{\,\}\}$
$DC_{Speaker}$	$\{verum(p)\}$	$DC^*_{Speaker}$	$\{\{\,\}\}$
DC_{Hearer}	$\{p\}$	DC^*_{Hearer}	$\{\{\,\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\,\rangle\rangle$

My proposal is that here too the function of attributive *que* is to attribute a commitment to a proposition to the hearer. Other than in declaratives, the proposition is not added directly to the current DC but to the projected DC of the hearer. Table 3.14 illustrates the conversational scoreboard proposed for neutral polar questions in Malamud & Stephenson (2015). The prejacent, i.e. the declarative equivalent of the polar question, is added to the top of the table.[21] The positive and negative versions of the prejacent are added to the projected CG. In the following moves, the answer adds the hearer's commitment to either p or $\neg p$ which will then become part of the CG.

Table 3.14: Conversational scoreboard when asking *p?*

current		projected	
CG	$\{\,\}$	CG*	$\{\{p\},\{\neg p\}\}$
$DC_{Speaker}$	$\{\,\}$	$DC^*_{Speaker}$	$\{\{\,\}\}$
DC_{Hearer}	$\{\,\}$	DC^*_{Hearer}	$\{\{\,\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\,\rangle\rangle$

For *que*-initial polar questions, I assume again that they contain all the same conversational states as neutral polar questions, along with an additional state that results from the presence of the attributive complementizer (cf. Table 3.15). This means that the prejacent is put on the table and the positive and negative versions are added to the projected CG. In §3.1, I proposed the generalization that *que* in polar questions signals that the speaker expects a positive answer from the hearer. I model this by proposing that the prejacent is added to the hearer's projected DC. By adding the prejacent to the projected rather than the current DC,

[21] For an explanation of why it is the positive version of the proposition that is put on the table, see the concept of *highlighting* introduced in Roelofsen & Farkas (2015).

the speaker shows that she is not absolutely sure that the hearer will answer positively, but she suspects it. If she were sure that the hearer was committed to the proposition, there would be no need to ask a question in the first place; instead the speaker could directly assert the bare or attributive *que*-initial proposition. What this means is that the contribution of the complementizer in *que*-initial declarative and *que*-initial polar questions only differs in that in the latter the attributive commitment is tentative whereas in the former it is presented as definite. Importantly, delaying the attribution to the projected DC is not proposed ad hoc here, but rather, as I showed, it follows from general properties of the information states involved in questions.

Table 3.15: Conversational scoreboard when asking *Que p?*

current		projected	
CG	$\{\}$	CG*	$\{\{p\},\{\neg p\}\}$
$DC_{Speaker}$	$\{\}$	$DC^*_{Speaker}$	$\{\{\}\}$
DC_{Hearer}	$\{\}$	DC^*_{Hearer}	$\{\{p\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\rangle\rangle$

In §3.1, I presented the different contexts in which *que*-initial polar questions appear, which were teased apart by Prieto & Rigau (2007). For polar questions in an anti-expectational context, the speaker held a previous belief that $\neg\ p$ was the case (see the scoreboard for an anti-expectational *que*-initial polar questions in Table 3.16), but contextual evidence contradicts her belief. This is illustrated in the example in (141) repeated from (10).

(141) Catalan (Prieto & Rigau 2007: 15)
Que vindràs a Barcelona? No em pensava pas que ens
QUE come.2SG.FUT to B. not me think1SG.IPFV.PST NEG that us
acompanyessis.
acompany.2SG.IPFV.PST
'Are you coming to Barcelona? I didn't think you were coming with us.'

I define contextual evidence as evidence accessible to the speaker in the current discourse situation or in a previous situation (Kocher 2017b).[22] Importantly, given the contextual evidence and despite the fact that it is in disagreement with

[22]This definition contrasts with that proposed by Büring & Gunlogson (2000) and adopted by Sudo (2013), where it is stated that the contextual evidence has to be mutually accessible to

her own previous belief, the speaker expects that the hearer will answer her question in the positive. This licenses the use of attributive *que* in these contexts, which again places the prejacent in the projected DC of the hearer.

Table 3.16: Conversational scoreboard when asking an anti-expectational *Que p?*

previous		current		projected	
CG	$\{\}$	CG	$\{\}$	CG*	$\{\{p\},\{\neg p\}\}$
$DC_{Speaker}$	$\{\neg p\}$	$DC_{Speaker}$	$\{\}$	$DC^*_{Speaker}$	$\{\{\}\}$
DC_{Hearer}	$\{\}$	DC_{Hearer}	$\{\}$	DC^*_{Hearer}	$\{\{p\}\}$
Table	$\langle\,\rangle$	Table	$\langle p\rangle$	Table*	$\langle\langle\,\rangle\rangle$

Confirmatory polar questions such as (142) present a different picture.

(142) Catalan (Kocher 2017b: 19: ex 40a)
[Context: Anna meets her friend Carles. He is tanned and seems relaxed. Anna asks:]

Que has estat de vacances?
QUE AU.2SG.PRF.PRS be.PTCP on vacation
'Have you been on vacation?'

In these cases, there is no contextual evidence that contradicts the speaker's previous belief; rather, the speaker herself suspects that p is the case and she asks the question in order to confirm this suspicion. Given her belief, regardless of what it is founded on – in the case of (142), indirect contextual evidence –, she expects that the hearer's answer will be positive. She uses attributive *que* to express the fact that she considers the prejacent to be part of the projected discourse commitment of the hearer.

To conclude the discussion on *que* in polar questions, I will now provide a characterization of the scoreboard of these questions when they contain the question particles *oi* or *eh*. For the following line of argument, one crucial point is that according to Prieto & Rigau (2007), these constructions only appear in confirmatory questions. Furthermore, they are not restricted to full questions but can also appear in reduced confirmatory tags (cf. 143).

the speaker and the hearer. In Kocher (2017b), I show that to make *que* in Catalan polar questions felicitous the evidence does not have to be mutually accessible as long as the hearer can accommodate the fact that the speaker is biased.

Table 3.17: Conversational scoreboard when asking a confirmatory *Que p?*

current		projected	
CG	$\{\}$	CG*	$\{\{p\},\{\neg p\}\}$
$DC_{Speaker}$	$\{\}$	$DC^*_{Speaker}$	$\{\{p\}\}$
DC_{Hearer}	$\{\}$	DC^*_{Hearer}	$\{\{p\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\,\rangle\rangle$

(143) a. Catalan
Sense fer res es fa més llarga l'
without do nothing CL.IMP make.3SG.PRS more long the
estona. Eh que sí?
time period EH QUE yes
'When doing nothing time grows longer. Right?' (caWac)

b. No tots els partits estan acusats d' haver -se
not all the party.PL be.3PL.PRS accused of AUX.INF.PRF.PRS CL.REFL
finançat il·legalment, oi que no?
finance.PTCP illegally OI QUE no
'Not all parties are accused of being funded illegally, right?' (caWac)

The particles also appear as final tags on declaratives where they function as requests for confirmation. Castroviejo (2018) argues that in these contexts, *eh* and *oi* have slightly different discourse contributions. With *oi?* the speaker double-checks the truth of the prejacent, while with *eh?* she requests that the hearer voices his commitment. While the particles are interchangeable in many contexts, they do not behave the same way when a confirmation of facts is requested, as in (144a): These sentences are infelicitous with *eh?* but felicitous with *oi?*. On the other hand, in requests for confirmation of opinions like in (144b) repeated from (12), both particles are allowed.

(144) Catalan

a. (Castroviejo 2018: ex 18)
La Terra és rodona, oi?/ #eh?
the earth be.3SG.PRS round OI EH
'The Earth is round, right/ #huh?'

b. (Castroviejo 2018: ex 19)
T' has tallat els cabells, oi?/ eh?
CL.2SG AUX.2SG.PRF.PRS cut.PTCP the hair OI EH
'You had your hair cut, right?/ huh?'

Castroviejo (2018) proposes the following conversational scoreboards to model the discourse contribution of the sentence-final particles. Table 3.18 shows the scoreboard assumed in Castroviejo (2018) for *p, oi?*.

It has all the information states of an assertion, i.e. the prejacent is added to the table and the projected CG. The contribution of the particle comes by way of adding the prejacent to the speaker's DC because the commitment to the truth of the proposition is at issue.

Table 3.18: Conversational scoreboard when uttering *p, oi?* adopted from Castroviejo (2018)

current		projected	
CG	$\{\}$	CG*	$\{\{p\}\}$
$DC_{Speaker}$	$\{\}$	$DC^*_{Speaker}$	$\{\{p\}\}$
DC_{Hearer}	$\{\}$	DC^*_{Hearer}	$\{\{\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\rangle\rangle$

Castroviejo (2018) proposes a different contribution for *eh*. The corresponding scoreboard is given in Table 3.19. It again contains all the information states of an assertion but unlike with *oi*, the prejacent is not added to the projected DC of the speaker. The commitment to the truth is not at issue, which is why *eh* is infelicitous with facts (cf. 144a). The proposition is attributed to the hearer's projected DC because *eh* is used when the speaker seeks a confirmation of the tentative commitment she attributes to the hearer.

The contrast between the two particles is lost when they introduce a *que*-initial polar question. In these, the presence of either particle coincides with a confirmatory reading.

(145) Catalan (Castroviejo 2018: ex 15)
Oi/ Eh que acabaràs la feina?
OI EH QUE finish.2SG.FUT the work
'You'll finish your work, right?'

Table 3.19: Conversational scoreboard when uttering *p, eh?* adopted from Castroviejo (2018)

current		projected	
CG	$\{\}$	CG*	$\{\{p\}\}$
$DC_{Speaker}$	$\{\}$	$DC^*_{Speaker}$	$\{\{\}\}$
DC_{Hearer}	$\{\}$	DC^*_{Hearer}	$\{\{p\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\rangle\rangle$

The loss of the different interpretations, in my view, constitutes evidence that the contribution made by the particles and the attributive complementizer is compositional. To show this, I propose the scoreboard in Table 3.20.

Table 3.20: Conversational scoreboard when uttering *Oi/Eh p?*

current		projected	
CG	$\{\}$	CG*	$\{\{p\}\}, \{\{\neg p\}\}$
$DC_{Speaker}$	$\{\}$	$DC^*_{Speaker}$	$\{\{p\}\}$
DC_{Hearer}	$\{\}$	DC^*_{Hearer}	$\{\{p\}\}$
Table	$\langle p \rangle$	Table*	$\langle\langle\rangle\rangle$

In both cases, the polar questions introduced by the particles have a confirmatory reading. My assumption is therefore that they have all the same information states as confirmatory *que*-initial polar questions. This means that just as in a regular polar question, the prejacent is added to the table and the positive and negative version of the prejacent are added to the projected CG. As proposed above, I conceive of a confirmatory reading as a tentative commitment on the part of the speaker. This is modeled as an addition of the prejacent to the speaker's projected DC. The contribution of attributive *que* is once again modeled as an attributive commitment to the hearer, i.e. the addition of the prejacent to the hearer's DC. Given this setup, the contribution of each of the particles does not give rise to any change in the information states. In the view of Castroviejo (2018), the contribution of *eh* is that the prejacent is added to the hearer's DC, which coincides with the contribution of *que*. In the alternative case with *oi*, its contribution that adds the prejacent to the speaker's projected DC is also already present because of the confirmatory reading of the question. As a result, the scoreboards for questions containing either of the particles are identical, which explains why the contrast

in their contribution following declaratives observed by Castroviejo (2018) is lost when they introduce *que*-initial polar questions.

To sum up, in this section I have modeled the pragmatic contribution of attributive *que* by making use of conversational scoreboards as proposed by Malamud & Stephenson (2015). My point of departure was that all the relevant constructions are attributive in the sense of Poschmann (2008). This means that a commitment to the prejacent is attributed to the hearer. It was proposed that this can be modeled in a system à la Malamud & Stephenson (2015) by adding the prejacent to the current DC of the hearer in assertions and to the projected DC of the hearer in polar questions. I applied this idea to the individual constructions and suggested that further pragmatic effects that arise are a result of the interplay between the attributive meaning and the properties of the other elements involved in the constructions.

3.5 Summary

This chapter has explored a variety of different constructions involving attributive *que*. My proposal was that in all the constructions the complementizer is merged in the lowest projection of the left periphery, in FinP, where it is valued with an attributive feature. The surface position of the complementizer in the different constructions is reached through head-to-head movement. In the different sections of the chapter I demonstrated empirically that this movement is conditioned by the presence of externally-merged material in a specifier. It was shown that this simple mechanism allows the correct predictions to be made with regard to word orders involving the attributive complementizer. The final section of this chapter was dedicated to the discourse contribution of attributive *que*. My proposal was that it attributes a commitment to the proposition to the hearer. Different pragmatic effects were explained as the result of an interplay between the attributive meaning and the other elements contained in the constructions.

Concerning the interaction of syntax and pragmatics, this chapter again provides evidence for a decoupling of the two components of grammar. The presence of an attributive feature as well as the distributional properties of a complementizer with this value are syntax internal and therefore rightly treated within its domain. The consequences of the attributive feature for the interpretation of the sentences as well as its interaction with other elements involved in them are not syntactic but context-dependent and therefore part of pragmatics.

4 Empirical investigation

In this chapter I present empirical studies that focus on the pragmatic effects of attributive *que*. Very little has been said to date regarding the interpretation of these effects, meaning that more systematic research has proved necessary. Since there are no previous experimental or corpus analyses that can be used as a basis for predictions, the empirical studies that will be described in the following sections approach the phenomena with an exploratory mindset and toolkit. The pragmatic effect of attributive *que* is strongly context-dependent. This instance of *que* is hardly ever obligatory and often co-occurs with other context-dependent expressions. Pairing corpus-based and experimental methods appears to be a suitable means of gaining an initial impression of the possible pragmatic effects involved in the constructions. Given these properties, it is necessary to take native speaker's judgments on the acceptability and felicitousness of the constructions into account. The approach I present in this chapter is also in keeping with the move towards empirical studies that is currently changing the face of theoretical linguistics. This dynamic and ever evolving field invites experimentation and creativity in order to test and implement new methodological approaches. Exploring empirical methods to investigate the types of meaning involved in attributive *que* will therefore hopefully also prove useful for future researchers.

This chapter starts with a general introduction to the empirical and statistical methodology employed. §4.1 presents the experimental method used. The experiments are a version of acceptability judgment experiments that differ from the traditional type in that the experimental stimuli are not constructed by the experimenter but are sampled from corpus data. In §4.2, I provide some background information on the statistical tools that I used to model the data. §4.3 presents three empirical studies focusing on the interpretation of *que* in AdvC constructions. The first study is a corpus analysis that compares the distribution of AdvC to two other constructions in Spanish and Portuguese. The other two studies are corpus-based experiments on the same construction, one on Spanish and one on Catalan. In §4.4, I present a corpus-based experiment that focuses on attributive *que* in Catalan polar questions and compares their acceptability to their *que*-less counterparts. The aim is again to draw conclusions about the interpretation of

que. The chapter concludes with §4.5, which provides a general discussion of how exploratory empirical research can inform our theoretic choices.

4.1 Corpus-based experiments

For the experiments I use a method that is inspired by the work described in Degen (2015). It differs from more traditional experimental approaches in that the stimuli are not constructed by the experimenter but are sampled from corpora. This means that the stimuli are heterogeneous by design. This design proves particularly useful in pragmatics research, because the stimuli are embedded in felicitous contexts. This reduces the risk of unintended outcomes provoked by accidental artifacts in the data. A further advantage of randomly sampling the stimuli is that a wide range of statistical tools can be employed to analyze the data, enabling us to learn from the data and find patterns in a way that would not have been possible if the stimuli had been constructed with a specific type of analysis in mind. Phenomena that are strongly context-dependent are often affected by multiple factors and more traditional experiments concentrate on only a few of these. This can lead to very clear results with respect to the effects of these few factors in the precise experimental contexts. It is however not always clear whether and how these results can be generalized to natural data. One risk is therefore that reducing the variation in the data creates results with a limited predictive power because the complexity of the phenomenon is not represented. Instead, corpus-based experimental methods acknowledge and invite the complexity of linguistic data.

Developing this method means moving away from the established methods, which does carry some risks, but offers the significant advantage of finding new approaches to the data and achieving novel results. In a way, this method combines the best of both worlds: Corpus-based research expands our knowledge of the phenomenon and experimental research enhances this knowledge by including the native speakers' perspectives. In the best case scenario, it provides us with fine grained information on a phenomenon that cannot be extracted from the corpora alone. We therefore gain a more complete picture of the issues under investigation.

In Degen (2015), all occurrences containing the relevant pattern she investigated were sampled and presented to native speakers of English, who were asked to judge how similar the stimulus was to a paraphrase provided by the experimenter. I use a different design. In my experiments a small random sample of the relevant patterns is drawn from a corpus, which is then carefully modified to

accommodate the questions that I am trying to answer in this study. The judgments that I elicit are not based on similarity to a paraphrase. The reason for this is that similarity judgments rely on the idea that the participants have a relatively clear understanding of the interpretation of the constructions, which did prove to be the case in Degen's study, but cannot be assumed to be true for the constructions involving attributive *que*. I elicit judgments on the interpretation of the constructions in a more indirect way. My primary interest is in studying how constructions containing attributive *que* are different from similar constructions that lack the complementizer.

I follow the example of Degen and use no fillers in my experiment. The stimuli are already very heterogeneous, so there is no risk of habituation due to monotonous input. Additionally, the tasks I designed are tailored specifically to elicit judgments on very precise aspects of the interpretation of the constructions. Finding fillers that can be matched with the same tasks is therefore nearly impossible. Finally, one of the functions of fillers in traditional experiments is to make sure that the participants are unaware of the goal of the investigation. Ensuring that the participants do not know the aims of the study is necessary in many areas of linguistic research, for instance when a normative bias could influence the judgments. However, in the area that I am investigating here, the participants are unlikely to have explicit knowledge about what does and what does not correspond to the norm. Designing the experiment without fillers might therefore mean that the participants become aware of the goal of the investigation. Although the results of my experiments indicate that this was not actually the case, in principle a conscious participant can be a valuable informant for the questions that I am trying to answer.

4.2 Statistical modeling

I follow Baayen (2008) in subscribing to a modern type of exploratory data analysis. In this approach we allow for the possibility that not all of the patterns in the data can be explained by an a priori formulated theory. Instead, data exploration is carried out with an open mind, inviting the possibility that there is more to find in the data than the theory can predict. Therefore, in the following studies, I pair tools and strategies from descriptive and hypothesis-testing approaches to see what we can learn from the data beyond what we may expect based on theory-driven hypotheses.

Conditional inference trees are classification trees that are based on binary recursive partition (see Hothorn et al. 2006, Strobl et al. 2009). They are robust

against data sparseness and make no assumptions about the distribution of the population from which the data was sampled. They can be used to detect high-order interactions and work well even if the predictors are highly correlated. All these aspects make them useful for analyses of linguistic data that often have a high degree of variability (see Tagliamonte & Baayen 2012).

Conditional inference trees, rather like regression models, predict the response variable based on a series of predictor variables. The prediction is based on binary recursive splits. Conditional inference trees test whether a predictor is associated with a response variable and choose the predictor variable that has the strongest association with this response variable. The dataset is then split into two subsets with significantly different distributions of the response variable. These steps are repeated on the subsets until there is no predictor variable left that is associated with the response variable at the level of statistical significance. The statistical significance is determined by permutation (see Tagliamonte & Baayen 2012, Levshina 2015).

The example tree in Figure 4.1 as well as all the other conditional inference trees in this book are modeled using the `party` package in `R` (Hothorn et al. 2022). In Figure 4.1 the choice of the Dutch past perfect auxiliary (*hebben* 'to have', *zijn* 'to be' or 'both', i.e. 'have' and 'be') is predicted by the number of synonyms (VerbalSynsets) and the regularity of the lexical verb (Regularity). Both predictors are significant. The strongest association is with the number of synonyms that each verb has. This means that splitting the dataset into two groups based on the number of synonyms creates two datasets in which the distribution of the auxiliaries is significantly different. In the left branch the number of verbal synonyms is 6 or higher. The right branch contains the complementary set with fewer than 6 verbal synonyms. Within the subset of the verbs with high numbers of synonyms, there is a further split on regularity. This means that within this group, the choice of the auxiliary is significantly different depending on whether the verb is regular or not.

Random forests are a useful tool to measure the importance of the predictors in tree models. A random forest is computed on a large number of conditional inference trees, each of which is calculated based on a randomly generated subset of the data. The importance of the predictors is once again determined through permutation (cf. Breiman 2001). Figure 4.2 is based on a random forest calculated for the same model as the conditional inference tree in Figure 4.1. The Figure shows the relative importance of the two predictors in the model. The number of verbal synonyms is identified as the most important predictor. This is also evident from the conditional inference tree in Figure 4.1. The dot plot in Figure 4.2 additionally shows that its importance is far greater than that of the second predictor.

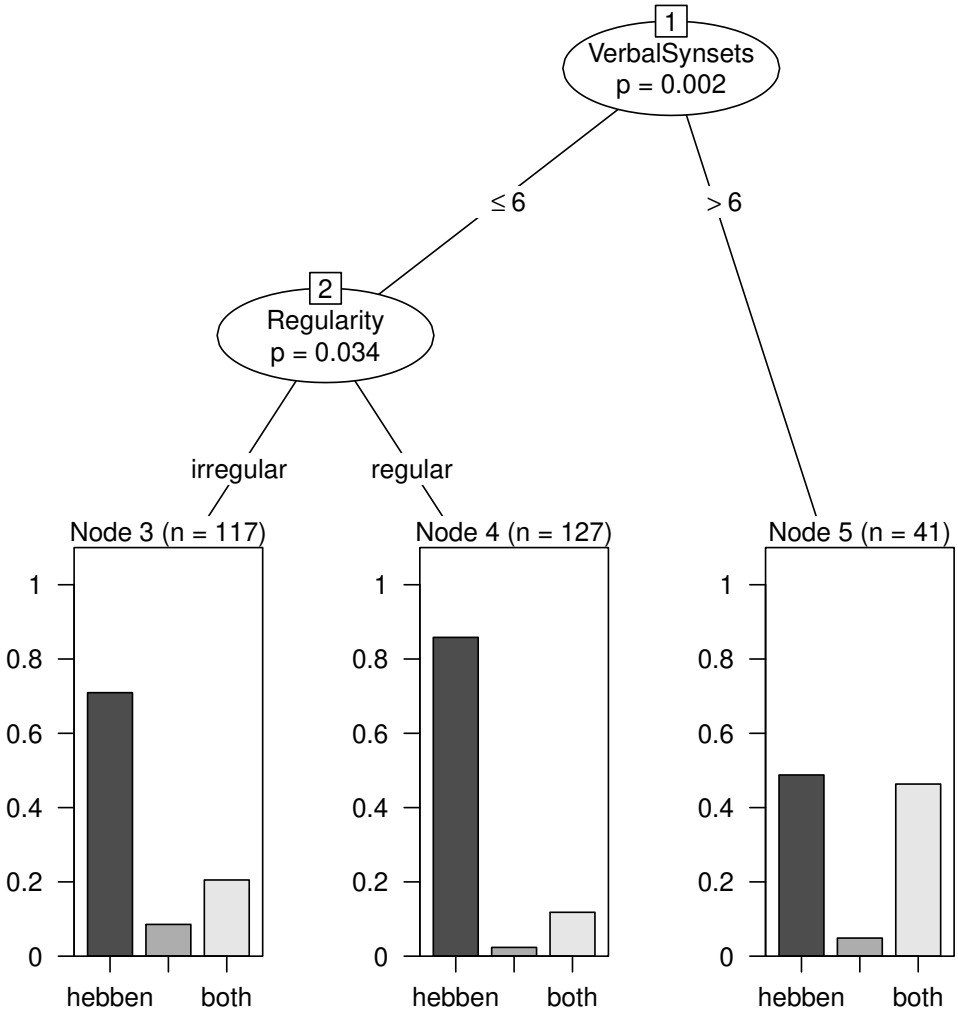


Figure 4.1: Conditional inference tree; The tree was fit to the built-in dataset `auxiliaries` from the `languageR` package (Baayen & Shafaei-Bajestan 2019). It shows the influence of the number of synonyms (VerbalSynsets) and the regularity of the lexical verb (Regularity) on the choice of the auxiliary in the Dutch perfect tense.

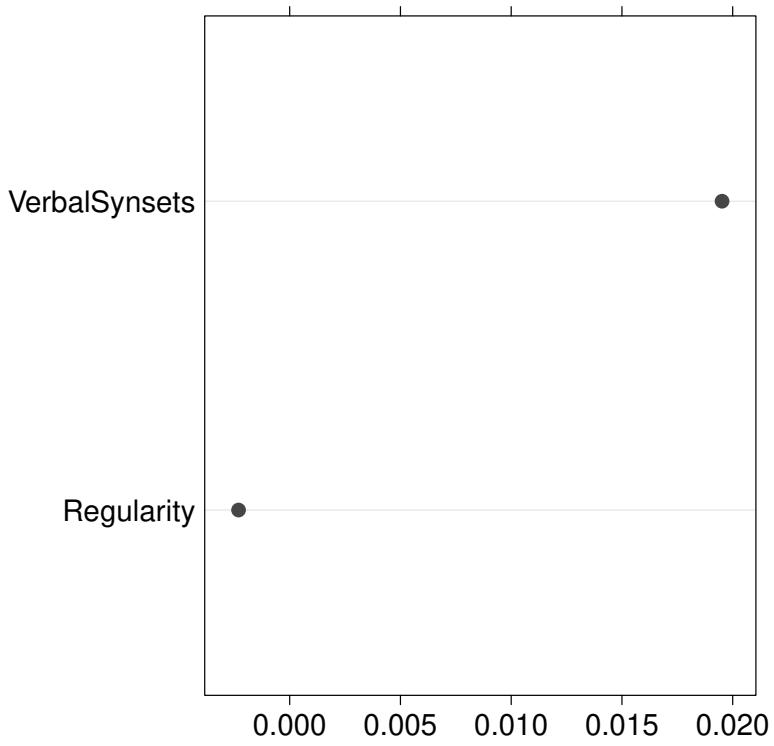


Figure 4.2: Dot plot of variable importance

4.3 Deictic centers of epistemic and evidential modifiers

In this section, I discuss three empirical studies, all of which focus on the interpretation of epistemic and evidential modifiers in different constructions. In principle, the studies are intended to investigate the function of *que* in AdvC. To do this, I have designed corpus studies and experiments that compare AdvC to two similar constructions, all of which contain epistemic and evidential modifiers. The three constructions are given in (1).

(1) Spanish

a. (AdvC)
Seguro que Juan viene.

b. (Adv)
Seguramente Juan viene.

c. (EsAdjC)
Es seguro que Juan viene.
'Sure QUE / Surely / It is sure that John will come.'

I am interested in determining the different readings that the modifiers receive in these constructions (see also §3.3.4 and §3.4 for a characterization of the constructions and their interpretation). Generally epistemic modals and evidentials express an evaluation of a proposition. This evaluation can be interpreted from the perspective of different deictic centers made up either of individual speech participants or of a joint perspective between multiple speech participants. The function that I propose for attributive *que* is that it attributes a commitment to the proposition to the hearer. A consequence is that the presence of *que* in AdvC usually establishes a shared perspective between the speaker and the interlocutor. I proposed in §3.4 that by way of an implicature, the reading of the modifiers in AdvC is also centered on this shared perspective. In EsAdjC the commitment to the propositions is added to the set of general discourse commitments. Again through an implicature, the perspective, on which the evaluation is centered should therefore be that of the speaker and some contextually relevant authority or group of people. The evaluation presented in Adv is unmarked with respect to which perspective it is centered on. Therefore, whichever perspective is most plausible in the context is adopted.

The first approach to the phenomenon is a corpus analysis that compares the distribution of the constructions in Spanish and Portuguese in different text types. The assumption is that different text types tend to invite different perspectives. For instance, in oral texts, which often include dialog sections, it is easier

to refer to or present assumptions about the interlocutor's perspective than it is in academic texts, in which the interlocutor is often not present. One expectation is therefore that AdvC, which centers the evaluation of the modifier on the speaker and the interlocutors, should be more frequent in informal and oral than in formal texts.

The second approach is composed of two experiments that build on the insights of the corpus study. Both elicit acceptability judgments and investigate the effect of the factors CONSTRUCTION, DEICTIC CENTER and MODIFIER on the acceptability of the stimuli. The first experiment deals with Spanish and the second with Catalan.

As stated above, I adopt an exploratory approach in the design and the statistical analysis of the following studies, since it offers the potential to learn from the data and draw novel conclusions that could not have been formulated a priori.

4.3.1 Corpus study on Spanish and Portuguese

In this section, I describe an exploratory corpus analysis through which I aim to determine the factors that influence the distribution of the three constructions. In this study I compare Spanish and Portuguese data. These languages were chosen because there are two very similar publicly available corpora, namely the 2006 Genre/Historical *Corpus do Português* (CdP) and the 2001 Genre/Historical *Corpus del Español* (CdE). They have a comparably moderate size but unlike large contemporary web-crawled corpora, they have the advantage of consisting of different text types. TEXT TYPE is a variable in the study that proved to be a significant predictor. A subset of the data has been used in a previous study reported in Kocher (2018b). The results presented here are novel, because the database and the statistical approach are new. This study also functions as preparation for the following experiments. The implications of the results form the basis for the formulation of the hypotheses that drive the development of the experiments.

4.3.1.1 Corpus

I used the 20th century subcorpus of the CdE and the CdP. Both corpora are publicly available and can be queried through an online interface. They cover data from the 12th century up to the 20th century and contain data from different dialects of both languages. The Spanish 20th century subcorpus makes up a quarter of the full corpus and the Portuguese one a little under half of it. These two subcorpora were chosen because they are comparable in their composition and their size (cf. Table 4.1). Sample size was considered as a potential predictor, but showed no significant effect.

Table 4.1: Number of tokens in million words per TEXT TYPE in CdE and CdP

	Text type				
	academic	press	oral	fiction	total
Corpus del Español	5.1	5.1	5.1	5.1	20.4
Corpus do Português	5.9	6.6	2.2	6	20.7

4.3.1.2 Sampling

I collected the sample through the online interface and then exported it to R (R Core Team 2013) in order to add annotations and perform analyses. The first step was to determine the most frequent epistemic and evidential modifiers in the three constructions in the corpora. These turned out to be the following: *cierto*/*certo*,[1] *claro*, *evidente*, *natural*,[2] *probable* and *seguro*. I then selected all the occurrences of the modifiers in the three constructions in sentence-initial position. The corpora are not annotated for sentence position; instead, this was approximated by querying all occurrences of the relevant pattern following a punctuation mark that indicates the end of a sentence. For AdvC, I included all cases of modifiers that end in *-mente* and those that do not (cf. Table 4.2). For EsAdjC, I included versions with copulas *ser* and *estar* (cf. Table 4.3). Neither of these variables showed a significant effect in the modeling.

Table 4.2: Distribution of derived and underived adverbs per modifier in AdvC

	cierto		*claro*		*evidente*		*natural*		*probable*		*seguro*	
	sp	pt	sp	pt	sp	pt	sp	pt	sp	pt	sp	pt
-mente	14	8	0	0	18	32	30	36	2	0	11	3
-0	34	9	576	306	0	11	2	4	0	0	82	0

[1] In the following I will use the Spanish cognate to refer to all the modifiers.

[2] For a discussion of the evidential nature of *natural* in Spanish and Portuguese, see Kocher (2014).

Table 4.3: Distribution of *ser* and *estar* per modifier in EsAdjC

	cierto		*claro*		*evidente*		*natural*		*probable*		*seguro*	
	sp	pt	sp	pt	sp	pt	sp	pt	sp	pt	sp	pt
estar	0	8	34	17	0	0	0	0	0	0	0	0
ser	126	81	24	190	64	96	11	31	85	25	8	1

4.3.1.3 Description of the data and predictors

The conditional inference tree that I will describe in the following section models the influence of the MODIFIER, the TEXT TYPE and the LANGUAGE on the distribution of the three CONSTRUCTIONS. In this section I describe the distribution of these variables in the sample.

Table 4.4: CONSTRUCTION by LANGUAGE

	Adv	AdvC	EsAdjC	total
pt	511	409	449	1369
sp	839	769	352	1960
total	1350	1178	801	3329

Table 4.4 shows the occurrence of each CONSTRUCTION per LANGUAGE. In general there are more propositions containing an epistemic or evidential modifier in Spanish than in Portuguese. The difference in frequency between the CONSTRUCTIONS is strongest in Adv and AdvC, which both occur far more frequently in Spanish than in Portuguese. The differences between the CONSTRUCTIONS are small in Portuguese, but are relatively large in Spanish. The variable LANGUAGE has only two levels and does not take dialectal variation into account. In the process of model fitting and selection, I considered dialectal variation as a possible predictor. Each data point in the corpus contains information about the country of origin of the author or speaker. For Portuguese, only two varieties (European and Brazilian Portuguese) are part of the corpus. For Spanish there is a larger range of countries of origin represented. I grouped these into seven dialectal varieties in accordance with Hualde et al. (2009) resulting in Andean, Canarian, Caribbean, Central American, Chilean, European, Mexican and Rioplatense Spanish. The variable was significant, but the increase in accuracy was

minimal (2%) and the resulting patterns were not straightforwardly interpretable. I hence chose to use a more simple and insightful model containing the variable LANGUAGE with only two values, Spanish and Portuguese.

Table 4.5: CONSTRUCTION by MODIFIER

Modifier	*cierto*		*claro*		*evidente*		*natural*		*probable*		*seguro*	
	sp	pt	sp	pt	sp	pt	sp	pt	sp	pt	sp	pt
Adv	96	91	13	13	136	106	175	144	189	139	230	18
AdvC	48	17	576	306	18	43	32	40	2	0	93	3
EsAdjC	126	89	58	207	64	96	11	31	85	25	8	1
	270	197	647	526	218	245	218	215	276	164	133	22
total	467		1173		463		433		440		353	

Table 4.5 shows the frequency of each CONSTRUCTION per MODIFIER for each LANGUAGE. The strongest contrast is between *seguro* and *probable*, which are both rarer in Portuguese than in Spanish. *Claro* is mostly used in AdvC in both LANGUAGES. There is also a considerable number of instances of this MODIFIER in EsAdjC, but only in Portuguese.

Table 4.6: CONSTRUCTION by TEXT TYPE

Text type	academic		fiction		news		oral	
	sp	pt	sp	pt	sp	pt	sp	pt
Adv	57	104	267	160	126	98	389	149
AdvC	2	10	296	152	77	127	394	120
EsAdjC	75	33	79	110	129	162	69	144
	134	147	642	422	332	387	852	413
total	281		1064		719		1265	

Table 4.6 shows the frequency of each CONSTRUCTION per TEXT TYPE per LANGUAGE. The frequency per TEXT TYPE is comparable in the two LANGUAGES,[3] with

[3]While there appears to be a difference in oral texts, there is in reality none to speak of. The tokens per TEXT TYPE in Portuguese are not equally distributed. In particular, the Portuguese oral subcorpus is roughly half the size of the Spanish one (cf. Table 4.1).

the exception that regardless of which CONSTRUCTION the MODIFIERS appear in, they are less frequent in Portuguese fiction texts than in the Spanish equivalents. There is a lower frequency of MODIFIERS in academic and news texts, which can be explained as being due to the fact that these TEXT TYPES often require the authors to maintain or at least simulate an objective perspective. Some assumptions can also be made based on the distribution of the CONSTRUCTIONS in the different TEXT TYPES. AdvC is very rare in academic texts and most frequent in fiction and oral texts. The high number of AdvC in oral texts, in which an interlocutor's perspective can be addressed most directly, is an expected result. Fiction texts have more heterogeneous properties than the other TEXT TYPES. One reason for the similar distribution of AdvC in fiction and oral texts could be that the former also often contain imitations of orality.

4.3.1.4 Results

In this section, I present the results of the tree model that I fitted to the corpus data. Model selection was carried out in a exploratory manner. I fitted a number of models with a varying degree of complexity and settled on the present model based on objective measurements such as the accuracy of the models and also based on hypotheses-driven criteria such as the plausibility and interpretability of the predicted effects. Figure 4.3 visualizes the conditional inference tree model. Three variables are significant in the model: MODIFIER, which shows the greatest effect, TEXT TYPE and LANGUAGE. The dot plot in Figure 4.4 shows the importance of each of these variables, which was determined through a random forest. Finally, the heat map in Figure 4.5 plots the observed vs. predicted values from Table 4.8. It suggests that the model has a high degree of accuracy in its predictions: In fact it has an accuracy of 70%. Accuracy was calculated by dividing the sum of the correctly predicted values, i.e. the values in the diagonal of Table 4.7 ($1137 + 790 + 409 = 2336$), by the sum of all of the predicted values ($2336/3329 = 0.702$).

Table 4.7: Predicted categories

	observed Adv	observed AdvC	observed EsAdjC
predicted Adv	1137	225	232
predicted AdvC	12	790	160
predicted EsAdjC	201	163	409

The first and most significant split in the tree model separates all the data containing *claro* from the data containing the remaining MODIFIERS. This suggests that *claro* displays particular behavior that is unlike the others. On the left branch, containing the *claro*-data (cf. Figure 4.6(a)), Node 2 separates Portuguese and Spanish. Within the Portuguese data, oral and academic texts are paired together and are significantly different from fiction and news texts. In the terminal Nodes (4,5) the distribution of *claro* in the Portuguese TEXT TYPES is plotted. In academic and oral texts AdvC and EsAdjC have a similar frequency (more than 40%). In fiction and news texts, however, AdvC is more frequent. In the Spanish subset (Node 6), there are two subsequent splits on TEXT TYPE. The first one splits academic texts from the rest. In the corresponding terminal Node 10, the plot shows that in this TEXT TYPE *claro* appears frequently in EsAdjC but also in Adv. The second split separates news from fiction and oral texts. In the latter group, AdvC makes up the overwhelming majority (cf. Node 9). In news texts AdvC is also the most frequent CONSTRUCTION in which *claro* appears, but there is also a substantial percentage (more than 20%) of EsAdjC.

In the subset of data containing all MODIFIERS but *claro* (cf. Figure 4.6(b)), the first split separates *natural* and *seguro* from *cierto*, *evidente* and *probable* (Node 11). Within the *natural*/*seguro* subset, the next most important difference is based on LANGUAGE. In the subset that includes Portuguese data with *natural* and *seguro*, the only significant split is on TEXT TYPE (Node 18). In news texts the MODIFIERS appear in all three CONSTRUCTIONS. They are most frequent in Adv but the differences are relatively small (Node 20). In fiction, oral and academic texts, the contrasts are bigger. Adv is by far the most frequent CONSTRUCTION for Portuguese *seguro* and *natural* in these TEXT TYPES (Node 19). In the Spanish subset, there is a split on MODIFIER (Node 13). Text type does not play a role in the distribution of the CONSTRUCTIONS in which Spanish *natural* appears. For Spanish *seguro* there is a significant difference in distribution depending on TEXT TYPES. In news texts it appears in all three CONSTRUCTIONS but is most frequent in Adv (Node 15). In the other three TEXT TYPES, it only appears in Adv and AdvC (Node 16).

The first split in the last set of MODIFIERS splits *cierto* and *evidente* from *probable*. In the subset containing *cierto* and *evidente* the first split is on TEXT TYPE (Node 21, Figure 4.7(a)). There is a significant difference between news texts and the other three TEXT TYPES. In news texts there are no further significant splits. The two MODIFIERS appear in all three CONSTRUCTIONS, but by far most frequently in EsAdjC (Node 40). In the other TEXT TYPES, there is a significant difference between the two MODIFIERS (Node 23). The *cierto* subset is further split on TEXT TYPE (Node 35) and LANGUAGE (Node 37). In academic texts, *cierto*

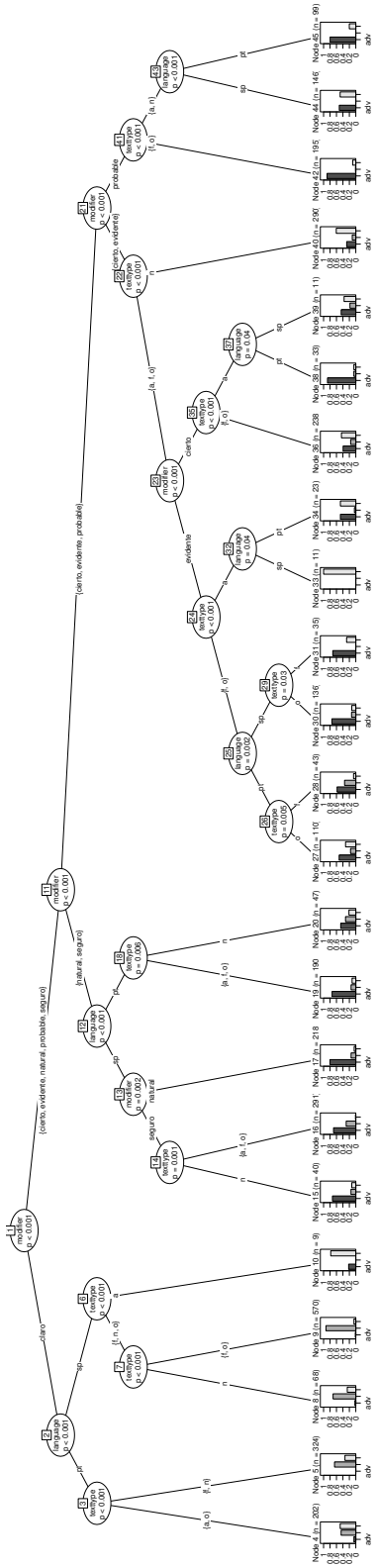


Figure 4.3: Conditional inference tree with splits on MODIFIER, TEXT TYPE and LANGUAGE

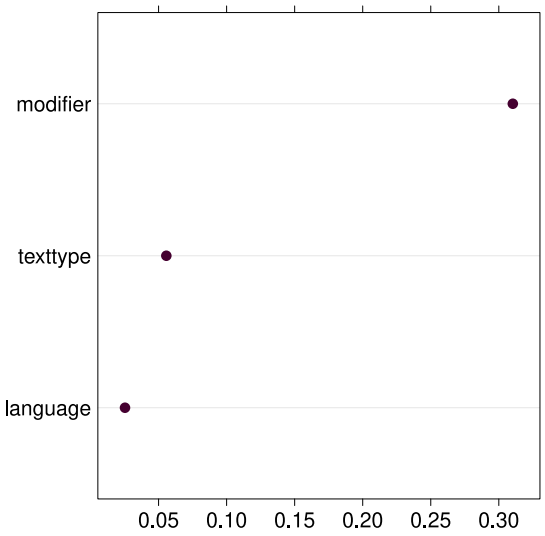


Figure 4.4: Dot plot of variable importance

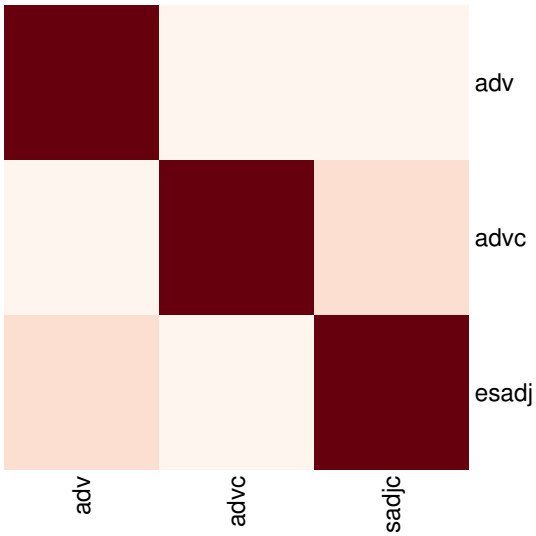
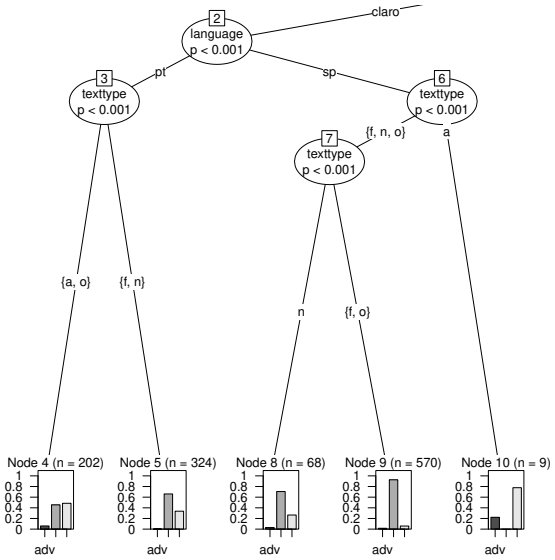


Figure 4.5: Heat map of predicted versus observed values; Darker shades correspond to larger counts.

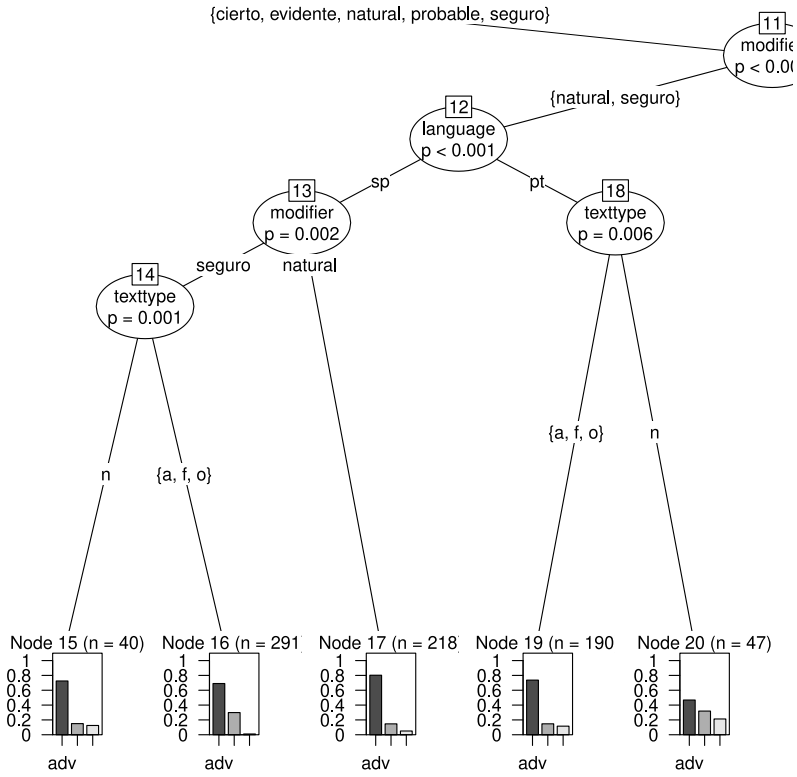
appears most frequently in Adv (Node 44) in both languages. In Portuguese it appears almost exclusively as an adverb, whereas in Spanish, AdvC and EsAdjC are also found. In fiction and oral texts there is no language specific difference. *Cierto* appears in all three CONSTRUCTIONS, but less frequently in AdvC (Node 36). In the *evidente* subset, there is a split separating academic from fiction and oral texts (Node 24). There is again a split on LANGUAGE. In academic texts, *evidente* is only found in EsAdjC in Spanish while in Portuguese it is found in all three constructions but mostly in Adv and EsAdjC. In fiction and oral texts, there is a split on LANGUAGE (Node 25). The TEXT TYPE is significant in both LANGUAGES. Spanish *evidente* appears by far most frequently in Adv in both TEXT TYPES (Node 30, 31). In fiction texts there are no cases of the MODIFIER in AdvC. In Portuguese, the MODIFIER is also most frequent in Adv in both TEXT TYPES, but the contrast between the CONSTRUCTIONS is less pronounced (Node 27, 28).

Within the *probable* subset (Node 21, Figure 4.7(b)), there is a significant difference in fiction and oral texts on the one hand and academic and news texts on the other hand (Node 41). In fiction and oral texts, LANGUAGE does not play a role. The MODIFIER appears by far most frequently in Adv. Lastly, the subset containing the MODIFIER in academic and news texts is split on LANGUAGE. In Spanish, *probable* has roughly the same frequency in Adv and EsAdjC. In Portuguese it has a far higher frequency in Adv than in EsAdjC. There are practically no occurrences of *probable* in AdvC.[4]

[4]There are two occurrences in Spanish, see Table 4.5.

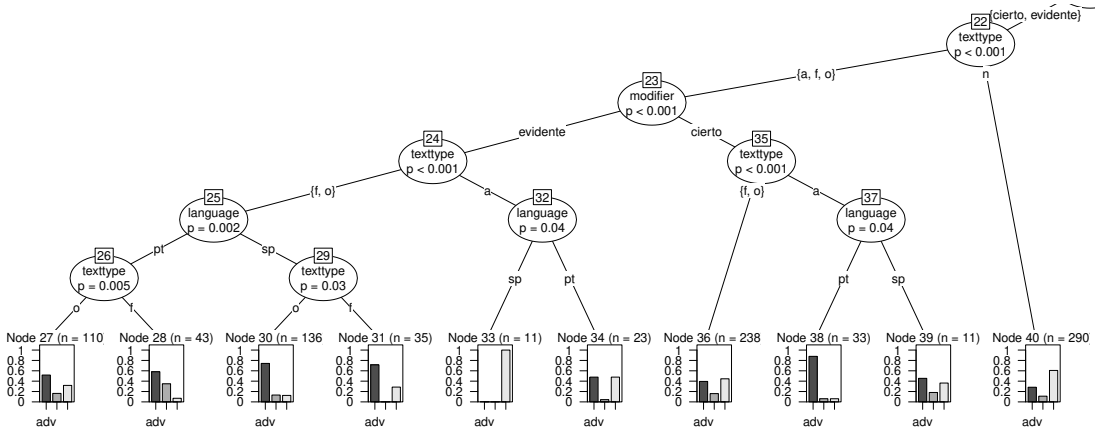


(a) Branch 1.

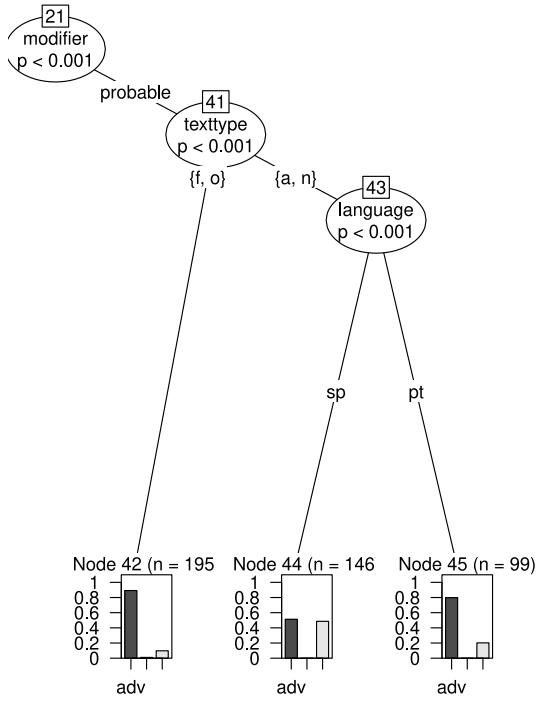


(b) Branch 2.

Figure 4.6: Individual branches of the conditional inference tree in Figure 4.3



(a) Branch 3.



(b) Branch 4.

Figure 4.7: Individual branches of the conditional inference tree in Figure 4.3

4.3.1.5 Discussion

The analysis shows that the MODIFIER *claro* has a very particular distribution that differs from that of the other MODIFIERS. In Spanish, it appears that *claro* is specialized for a use in the AdvC CONSTRUCTION. It is possible that Spanish *claro que* has already grammaticalized or is at least in the process of turning into a fixed pragmatic marker. *Claro* has a high frequency in AdvC in Portuguese too, but also appears in EsAdjC. Given that the two CONSTRUCTIONS have a different meaning, the distribution of *claro* in these CONSTRUCTIONS indicates that the meaning of the MODIFIER may have acquired different nuances in the two LANGUAGES. In Portuguese *claro* seems to be specialized for CONSTRUCTIONS that take perspectives other than the speaker's into account (AdvC and EsAdjC), while in Spanish its meaning is specialized for CONSTRUCTIONS that take the interlocutor's perspective into account (AdvC).

The results in general suggest a greater proximity between AdvC and EsAdjC than between AdvC and Adv, which is consistent with the idea that EsAdjC and AdvC – but not Adv – are attributive (cf. §3.4). In the cases of Adv and EsAdjC, the wrongly predicted data points (EsAdjCs predicted as either Adv or AdvC, and Adv predicted as either EsAdjC or AdvC) are distributed relatively equally across the two categories. In contrast, observed AdvCs are significantly more likely to be predicted to be EsAdjC (160) than Adv (12) (cf. Table 4.7 repeated in Table 4.8).

Table 4.8: Predicted categories

	observed Adv	observed AdvC	observed EsAdjC
predicted Adv	1137	225	232
predicted AdvC	12	790	160
predicted EsAdjC	201	163	409

Aside from *claro*, the other modifiers appear most frequently in Adv. This CONSTRUCTION is unmarked with respect to which speech participant the epistemic and evidential evaluation is centered on. It does not have a preferred reading but permits any type of deictic center that is plausible in a given context. It therefore follows that the other MODIFIERS investigated in this study are less specialized for one reading.

Apart from *probable*, all the MODIFIERS are attested in all three CONSTRUCTIONS. *Probable* is most frequent in Adv and does not appear in AdvC. This suggests that the meaning of this MODIFIER does not easily adapt to a reading that takes the perspective of the interlocutor into account.

The relation between TEXT TYPE and CONSTRUCTION supports the assumption that AdvC refers to an interlocutor's perspective, hence its high frequency in oral and fiction texts. EsAdjC peaks in academic and news texts, in which an authoritarian and apparently objective perspective is often employed and in which referring to the interlocutor's perspective directly is uncommon. Adv is the most common CONSTRUCTION in all texts types. This supports the idea that it is unmarked and therefore adapts easily to all possible deictic centers.

LANGUAGE plays different roles in connection with the different MODIFIERS. An interesting pattern is that the splits on LANGUAGE are at a high level of the tree for *claro*, *natural* and *seguro*. This indicates that the meaning of the MODIFIERS differs between Portuguese and Spanish. In contrast, the splits are at a deeply embedded level for *cierto*, *evidente* and *probable*, suggesting that these three MODIFIERS are less idiosyncratic. For *cierto*, LANGUAGE is not a significant predictor at all suggesting that, with respect to the tested CONSTRUCTIONS, the MODIFIERS have the same meaning in Spanish and Portuguese.

4.3.2 Acceptability judgment experiment on Spanish

In this section I describe the design and the results of an experiment to determine primarily the influence of the variables DEICTIC CENTER and CONSTRUCTION, but also other factors, on the acceptability of Spanish epistemic and evidential MODIFIERS in their respective contexts.

4.3.2.1 Corpus

The stimuli are taken from the 2016 Web/Dialects CdE. The corpus comprises 2 billion tokens and contains data from 21 different Spanish speaking countries. The corpus was built based on data acquired using Google search from 20 million randomly sampled web pages and blogs from all 21 countries.[5] I chose this corpus because of its large size and its informal register that approximates orality, which is essential when eliciting judgments on natural language pragmatics.

4.3.2.2 Data acquisition and selection

I first determined the four most frequent epistemic and evidential modifiers in the corpus which are *cierto, claro, evidente* and *seguro*. For each of these, I extracted all sentence-initial occurrences of the three CONSTRUCTIONS under investigation.

[5] See https://www.corpusdelespanol.org/web-dial/help/textsm.asp [November 4, 2022] for a detailed description of how the data were acquired and processed.

Since sentence position is not annotated in the corpus, I approximated this by extracting all cases of the relevant patterns following a sentence-final punctuation mark.

The final experimental stimuli were identified through stepwise random sampling and controlled selection. The full sample from the corpus was exported to R in order to automatically exclude certain data points and to draw random samples. All the occurrences of Adv followed by a punctuation mark (the adverbs in isolation can function as affirmative particles) and AdvC preceding the affirmative particle *sí* or the negative particle *no* followed by a punctuation mark were excluded from the sample. From the new reduced sample, I drew random samples of 20 items per MODIFIER per CONSTRUCTION. These random samples were then inspected individually. Further items had to be excluded based on my subjective selection. Items were excluded if the CONSTRUCTIONS did not modify declaratives, if the sentence and the contexts were not cohesive or if they contained offensive, sexual or religious content. I then drew another random sample of 3 items for each MODIFIER and CONSTRUCTION, yielding the final count of 36 experimental stimuli.

Table 4.9: Derived vs. underived adverbs in AdvC in the experimental stimuli

	cierto	*claro*	*evidente*	*seguro*
Adv	3	3	3	3
Adv-*mente*C/Adv-0 C	3/1	3/0	1/2	3/0
EsAdjC/EstáAdjC	3/0	0/3	3/0	3/0

In EsAdjC, the MODIFIERS can appear with the copulas *ser* and *estar*. In AdvC the MODIFIERS are sometimes derived and sometimes not. The whole sample extracted from the corpus contained all the cases of each MODIFIER in both varieties for each CONSTRUCTION. The random sampling resulted in the patterns given in Table 4.9. Neither of the two variables had a significant effect in the model.

4.3.2.3 Data modification

The critical target sentences in the scope of the modifiers were shortened so that all of them only constituted simple main sentences. The aim of this shortening was to obtain a more homogeneous set of stimuli and to reduce the number of words, thereby rendering the experiment shorter overall. The size of the preceding context was selected individually for each target sentence. The selection was

carried out based on my subjective judgment. The goal was to maintain the minimum amount of context necessary to make sense of the critical sentence and its modification. I opted for this qualitative criterion because the high heterogeneity of the fragments in the corpus both in length and content made it impossible to apply a quantitative criterion.

Each target sentence was presented in three different conditions that introduce, or more precisely negate, certain readings of the modifiers. These conditions were created by adding a concessive clause before the critical sentence. Each stimulus thus consists of a complex target sentence made up of the concessive clause and the epistemically or evidentially modified sentence extracted from the corpus along with the corresponding preceding context.

4.3.2.4 Experimental design

The experiment was run on Ibexfarm. It started with two simple practice items that explained the task and illustrated how to use the scale. The practice items were followed by 36 experimental stimuli. The experiment was in a Latin square design. Each stimulus was presented to one participant in only one of the three conditions, with the stimuli presented in a random order. The participants were asked to provide a judgment on the acceptability of the concessive, which was underlined, in the relevant context. The judgment was elicited on a five-point Likert scale, with 1 translating to the lowest degree of acceptability and 5 to the highest degree. The participants were instructed to provide judgments that corresponded to their intuitions regarding the naturalness and acceptability of the critical sentences. They were informed that the stimuli originated from a text corpus of an informal and therefore non-standard register. They were asked to disregard anomalies in orthography and other aspects that might not correspond to the norm. The experiment was estimated to take 15 minutes (mean duration 13.8 minutes). No instructions were given as to whether the participants should provide their judgments quickly or slowly.

4.3.2.5 Participants

The participants were recruited through social media by sharing the link with personal contacts and in Facebook groups of Spanish native speakers. In total 61 people participated in the experiment, 35 female and 26 male. The vast majority of them came from Spain (47), followed by Colombia (6) and Argentina (5). One participant each came from Germany, Uruguay and Mexico. The age of the participants ranged from 20 to 66 with a mean of 39.21 years. Contrary to what

one might expect from a recruitment process that relies solely on social media, the age distribution shows that this approach can also reach older participants: 14.75% are 60 or older.

4.3.2.6 Conditions

In the experiment the target sentences are presented in three conditions. Each condition negates certain readings of the MODIFIER.

(2) a. (*yo*-DEICTIC CENTER)
Aunque yo no lo crea, claramente tiene los
although I not it believe.1SG.SBJV.PRS clearly have.3SG.PRS the
ojos de su papá.
eye.PL of his father
'Although I don't believe it, he clearly has his father's eyes.'

b. (*tú*-DEICTIC CENTER)
Aunque tú no lo creas, claramente tiene los
although you not it believe.2SG.SBJV.PRS clearly have.3SG.PRS the
ojos de su papá.
eye.PL of his father
'Although you don't believe it, he clearly has his father's eyes.'

c. (*gente*-DEICTIC CENTER)
Aunque la gente en general no lo crea,
although the people in general not it believe.3SG.SBJV.PRS
claramente tiene los ojos de su papá.
clearly have.3SG.PRS the eye.PL of his father
'Although people in general don't believe it, he clearly has his father's eyes.'

In the first condition (*yo*-DEICTIC CENTER) the reading of the MODIFIER centered on the speaker is negated (2a). In the second condition (*tú*-DEICTIC CENTER), the hearer-centered reading is negated (2b). In the third condition (*gente*-DEICTIC CENTER), a reading where the MODIFIER is centered on a more general group of people is negated (2c).

The aim of using these conditions is to determine whether the three CONSTRUCTIONS make certain readings of the MODIFIERS more prominent. If this turns out to be the case, we expect that the target sentences should be judged low on the acceptability scale in a condition that negates precisely this reading. The *yo*-DEICTIC CENTER condition functions as a control condition. The expectation

is that the speaker is always part of the deictic center present in the MODIFIERS that were selected for this experiment, because all of them imply a strong commitment of the speaker towards the truth of the proposition. The *tú*-DEICTIC CENTER relates to the reading of the MODIFIERS in AdvC. Given the assumption that in this CONSTRUCTION the evaluation of the MODIFIER is centered on the speaker and the hearer, negating that the hearer believes *p* should not be acceptable. Based on this we would expect low acceptability for AdvC in the *tú*-DEICTIC CENTER. The third condition targets EsAdjC. In this CONSTRUCTION the evaluation is centered on a contextually relevant group of people or to an unspecified authority. In the *gente*-DEICTIC CENTER, the DEICTIC CENTER "the people" is negated, hence EsAdjC should be judged lower in this condition.

4.3.2.7 Description of the data and predictors

Table 4.10 summarizes the percentage of judgments for each value on the rating scale. There are larger values at the extremes of the scale, suggesting that in most cases the participants had strong intuitions about the acceptability and unacceptability of the stimuli and they chose intermediate ratings, translating to unclear intuitions, less frequently.

Table 4.10: Percentage of judgments per rating

rating	1	2	3	4	5
percentage of judgments	29.55	11.96	14.05	18.61	25.83

There is always the potential of variation owed to the speaker or to properties of the items that are not captured by the factors one defined. In order to keep this variation in check, the conditional inference tree and random forest model the standardized rather than the raw judgments.

There is considerable variation in the data by design, because the stimuli are taken from a corpus. The length (measured in number of words) is one aspect of this variation. The number of words per context ranges from 62 to 502 (mean 292). The length of the modified sentences ranges from 20 to 145 words (mean: 62). The examples in (3) illustrate two extremes in terms of context length. In spite of these differences, the length of the contexts did not provoke a significant effect. Furthermore, although REACTION TIME is not a concern in this experiment, the correlation between REACTION TIME and ITEM LENGTH (sum of the length of the context plus the length of the targets sentence) is very low (0.07).

(3) a. *Context*: Hablante A – ¿Qué cosas tiene Erin de ti?
'Speaker A – What does Erin have from you?'
Target sentence: Hablante B – Aunque tú no lo creas, claramente tiene los ojos de su papá.
'Speaker B – Although you don't believe it, clearly she has her father's eyes.'

b. *Context*: Hablante A – El objetivo de este seminario es reflexionar sobre aquellos conflictos del mundo (guerras, catástrofes, terrorismo) que en las primeras horas de producirse provocan un intenso seguimiento mediático pero semanas después desaparecen de las páginas de los diarios y de nuestro recuerdo. Decía Kapuscinzky que no entendía por qué los enviados especiales se iban cuando realmente en ese momento había que empezar a contar las historias...
'Speaker A – The goal of this seminar is to reflect on conflicts in the world (wars, catastrophes, terrorism) which, in the first hours after they take place, provoke an intense media following, but weeks after disappear from the pages of the newspapers and from our memories. Kapuscinzky said that he didn't understand why the special envoys left when it was actually at that point that it was time to start telling the stories.'
Target sentence: Hablante B – Aunque yo no lo crea, es evidente que Kapuscinzky no logró convencer a los directores de los medios.'
'Speaker B – Although I don't believe it, it is evident that Kapuscinzky didn't manage to convince the directors of the media.'

The bar charts in Figure 4.8, made using the `lattice` R-package (Sarkar et al. 2021), show the percentage of judgments by DEICTIC CENTER by CONSTRUCTION. The charts suggest that the participants were sensitive to the DEICTIC CENTERS but not to the CONSTRUCTIONS. This is clear from the fact that the plots on the horizontal axes show different patterns, while there is practically no difference on the vertical axes. The judgments for the *yo*-DEICTIC CENTER are far lower than those for the other DEICTIC CENTERS. Roughly 60 percent of the judgments provided for the *yo*-DEICTIC CENTER fall into the lowest category. The judgments for the other two DEICTIC CENTERS have a similar distribution: They are spread across all categories and have the highest percentages at the upper edge of the scale. There is a slight difference between the *tú*- and the *gente*-DEICTIC CENTER. The judgments are more varied for the *gente*-DEICTIC CENTER, which suggests that in some cases the participants were less certain about the acceptability of a reading negating this DEICTIC CENTER but are certain about the acceptability

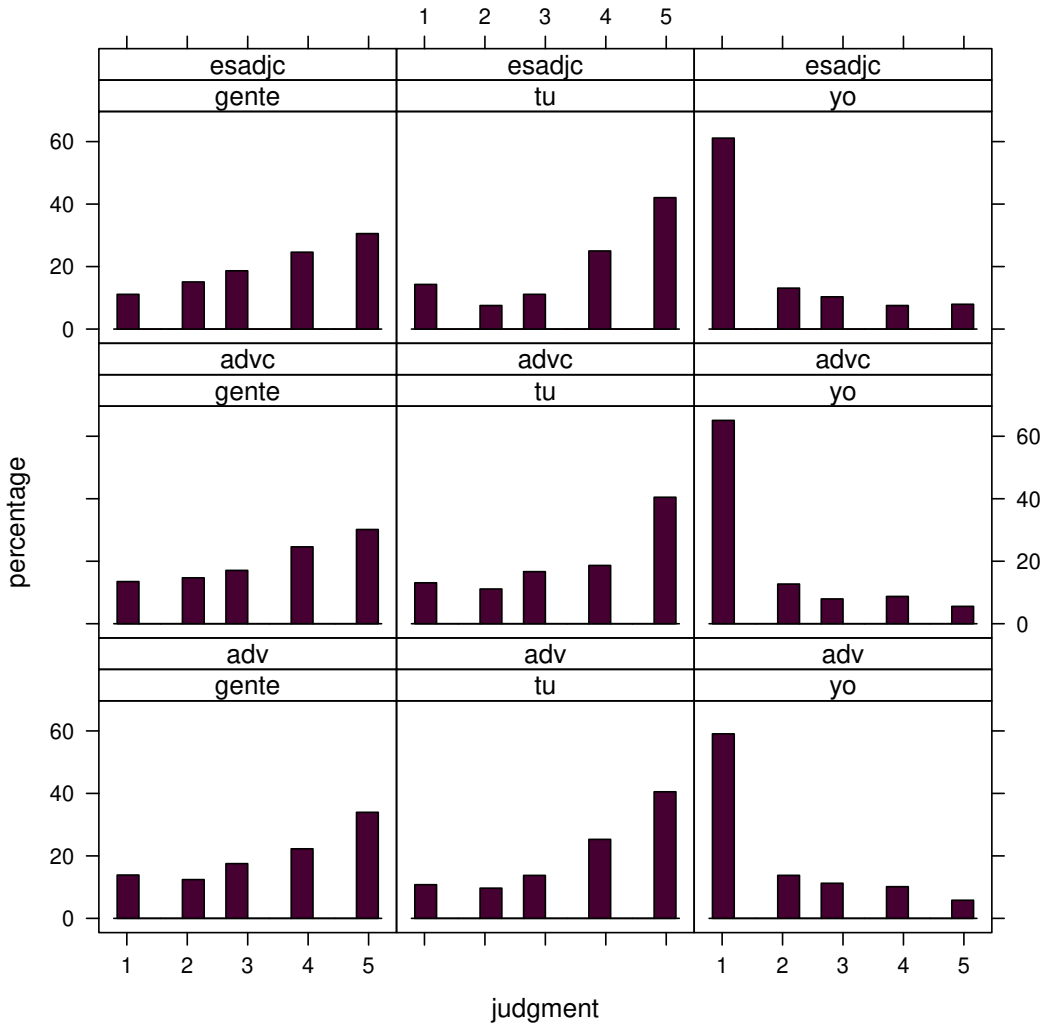


Figure 4.8: Individual bar charts for judgments per DEICTIC CENTER per CONSTRUCTION. The panels in the first row show the judgments for EsAdjC in the different DEICTIC CENTER conditions. The second row shows AdvC and the third row Adv.

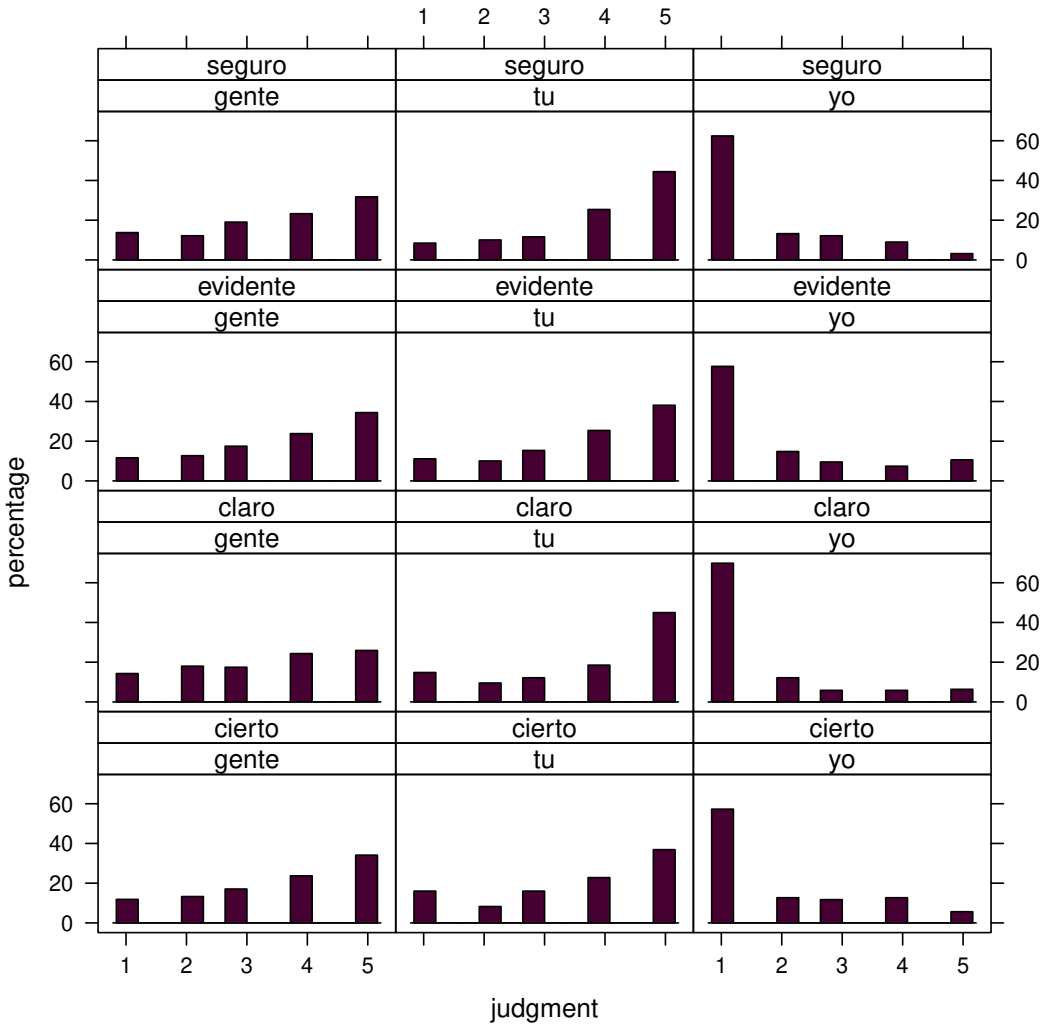


Figure 4.9: Individual bar charts for judgments per DEICTIC CENTER per MODIFIER. The panels in the first row show the judgments for *seguro* in the different DEICTIC CENTER conditions. The second row shows *evidente*, the third row *claro* and the last row *cierto*.

of the *tú*-DEICTIC CENTER. The bar charts in Figure 4.9 show the percentage of judgments by DEICTIC CENTER by MODIFIER. They once again suggest a DEICTIC CENTER effect. The differences between the MODIFIERS are minimal.

The bar charts in Figure 4.10 plot the percentage of judgments by CONSTRUCTION by MODIFIER. The charts support the idea that the participants had preferences for certain MODIFIERS in certain CONSTRUCTIONS. *Cierto* is less acceptable in EsAdjC and most acceptable in Adv. *Claro* has a high percentage of judgments at the lowest end of the scale. Interestingly, the MODIFIER in AdvC has high percentages for the lowest and the highest values on the scale. This suggests that the relation between *claro* and AdvC, which was strongly supported by the results from the corpus analysis, actually plays out differently in an acceptability judgment experiment. *Evidente* is most acceptable in EsAdjC and shows similar patterns as *claro* in AdvC and Adv. *Seguro* has the highest percentage of ratings at the lower extreme of the scale in AdvC. In general it has high percentages for the lowest and the highest values on the scale. This suggests that there is no direct relation between the MODIFIER and the CONSTRUCTION.

4.3.2.8 Results

A conditional inference tree was used to model the data. Model selection was based on the accuracy of the model and on the plausibility and interpretability of the effects. Figure 4.11(a) is a visual representation of the model. Three variables are significant: DEICTIC CENTER, AGE and COUNTRY-PARTICIPANT (nationality of the participants). Other linguistic variables such as CONSTRUCTION and MODIFIER did not show a significant effect. The dot plot in Figure 4.11(b) shows the importance of each variable that entered the calculation of the model. It was determined using a random forest. DEICTIC CENTER is by far the most important variable in this model. The heat map in Figure 4.12 plots the observed vs. the predicted values. The accuracy was calculated by dividing the sum of the correctly-predicted values by the sum of all the predicted values (cf. Figure 4.12). The accuracy is moderate (38%).

The first and most important split in the tree in Figure 4.11(a) shows that there is a significant difference between the *yo*-DEICTIC CENTER and the other two. The tree identifies complex interactions between the variables DEICTIC CENTER and AGE and DEICTIC CENTER and COUNTRY-PARTICIPANT. In the *yo*-DEICTIC CENTER-group, the one speaker from Uruguay gave significantly higher ratings than all the other speakers. In the group of data containing the *tú*- and the *gente*-DEICTIC CENTER there is an interaction with AGE. Younger participants found both readings equally acceptable, while older speakers judged the *tú*-DEICTIC CENTER significantly more acceptable than the *gente*-DEICTIC CENTER.

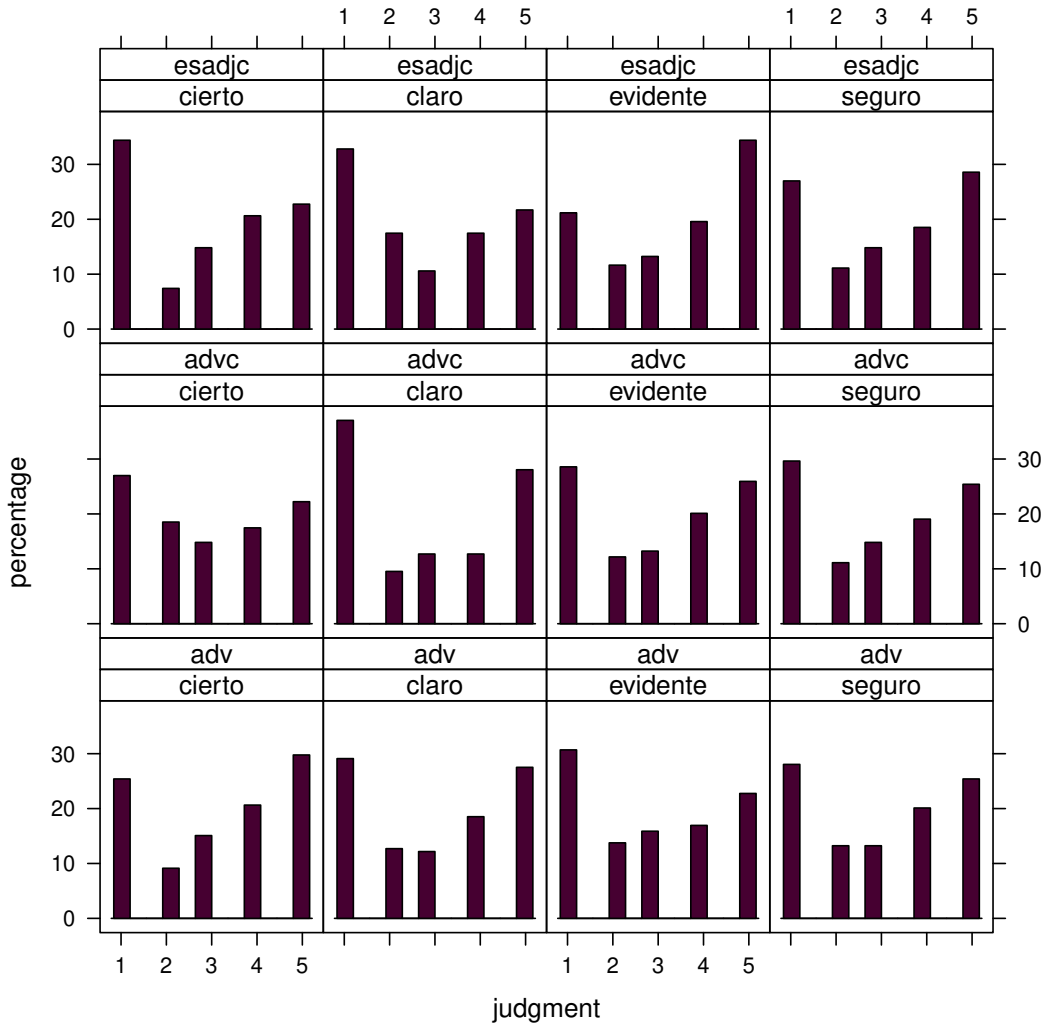


Figure 4.10: Individual bar charts for judgments per CONSTRUCTION per MODIFIER. The panels in the first row show the judgments for EsAdjC in the different DEICTIC CENTER conditions. The second row shows AdvC and the third row Adv.

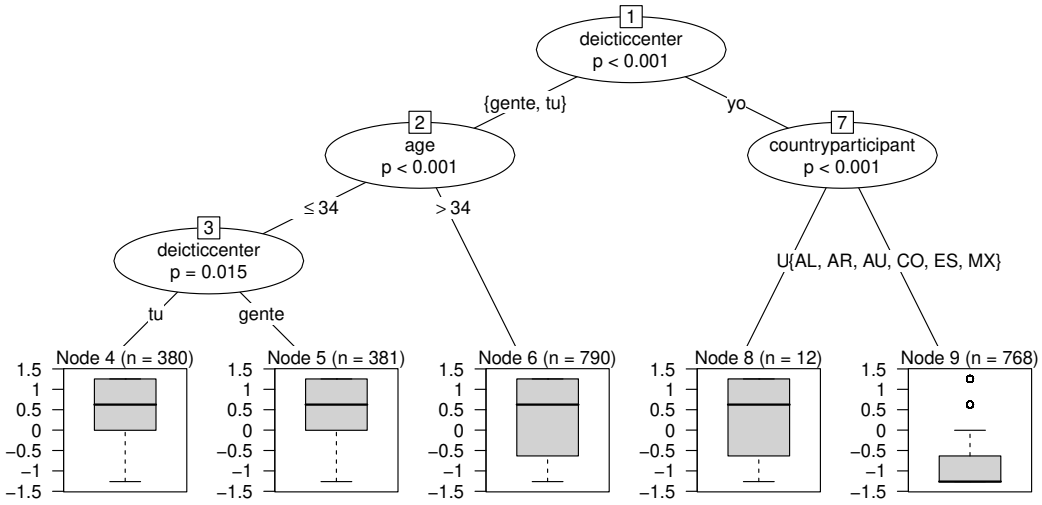
4.3.2.9 Discussion

The results show a solid effect for DEICTIC CENTER. Negating the speaker's perspective is always judged low on the acceptability scale in combination with the MODIFIERS chosen in this experiment. Although no significant interaction with CONSTRUCTION could be detected in the model, Figure 4.8 suggests that negating the speaker's perspective is even worse in AdvC. The model also shows that there is a significant difference between the negation of the perspective of the interlocutor vs. the negation of the perspective of "the people" in the group of older speakers. Negating the interlocutor's perspective is more acceptable in the present experiment. The perspective of the interlocutor is active in all of the items, because all the stimuli are dialogs. My interpretation of these results is that it is easier to address a perspective that is already active, while it is more infelicitous to negate the perspective of "the people", if this perspective was not presented in the context.

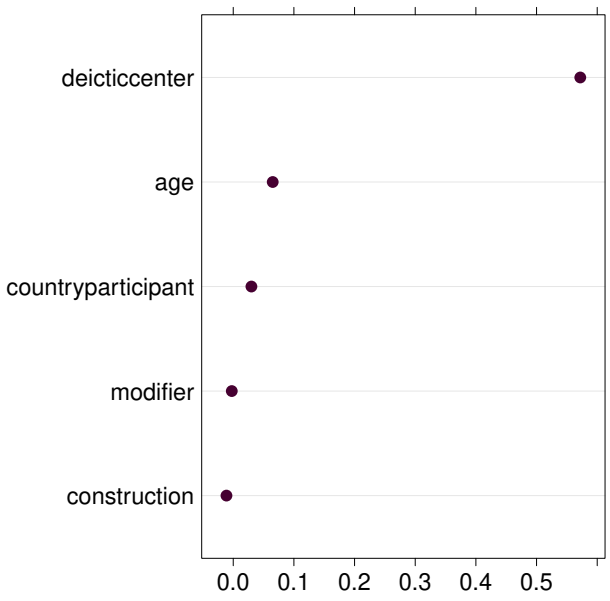
The model does not support an effect of CONSTRUCTION nor an interaction between DEICTIC CENTER and CONSTRUCTION. One possible explanation could be that linguistic expressions of the variable DEICTIC CENTER stood out more than the other variables. The task required the participants to judge the acceptability of concessives, which were underlined, in the contexts they appeared in. This means that the participants were instructed to focus on the DEICTIC CENTER, which might have led them to disregard the other properties. DEICTIC CENTER is also the only strictly linguistic variable that was created by modifying the corpus data. It was introduced in the form of concessive clauses that were not contained in the original data. So another reason why this variable gives rise to a larger effect is potentially the fact that the concessives did not adapt easily to some contexts. In the following experiment on Catalan, in order to counteract this issue, the acquisition of data was carried out differently: Contexts with concessives expressing doubt or disbelief were extracted and the data were manipulated by the addition of the MODIFIER and CONSTRUCTION.

4.3.3 Acceptability judgment experiment on Catalan

The Catalan experiment that will be described in this section again has the aim of determining the factors that influence the acceptability of the MODIFIERS in the three CONSTRUCTIONS.



(a) conditional inference tree



(b) variable importance

Figure 4.11: Conditional inference tree with splits on DEICTIC CENTER, AGE and COUNTRY-PARTICIPANT and a dot plot showing the importance of each variable

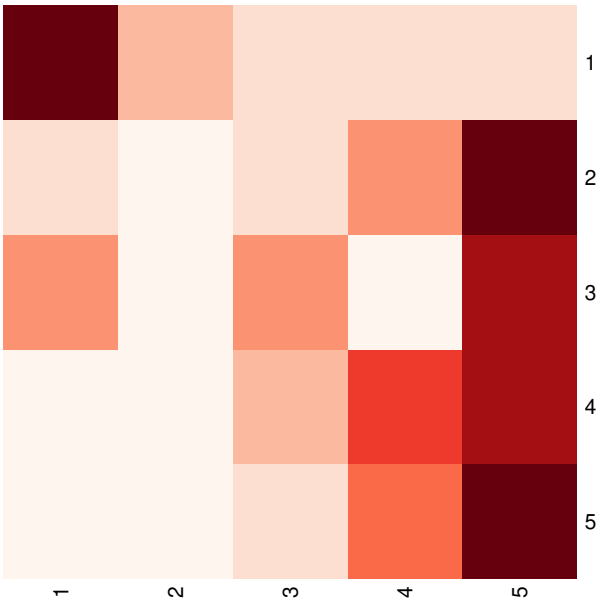


Figure 4.12: Heat map of 5 predicted and 5 observed values; Darker shades correspond to larger counts.

4.3.3.1 Corpus

The stimuli were sampled from the caWac corpus, which is among the largest corpora of contemporary Catalan. It comprises 780 million tokens and was built by a web crawl from the top-level .cat domain in 2013. The fragments can be considered comparable to those used for Spanish in their approximation of an oral register and style, because both are taken from corpora built using web data.

4.3.3.2 Data acquisition and selection

The caWac corpus does not have an online interface, but the data can be downloaded. The file contains xml code, tagging paragraphs and sentences, but no further annotation is provided. The final experimental stimuli were found through stepwise sampling, random sampling and controlled selection.

I first split the files into subfiles that could be handled by a computer with an average RAM. I then wrote Python scripts to select fragments with the critical sequences of words within a context of 500 words on the left and right side. I extracted all the occurrences of sentence-initial *encara que* [0-3 intervening words] $sembli_{subj}$/$sembla_{ind}$ 'although [0-3 intervening words] may seem/seems'. The selected sentences were chosen to approximate the concessives *Encara que*

no t'ho creguis/Encara que a gent no s'ho cregui 'Although you/people in general don't believe it', since the target concessives were absent in the corpus. The data were then processed in R. I kept all the items where the intervening words were a personal pronoun, the negative particle *no*, the neutral pronoun *ho* and combinations thereof, and discarded the rest. From the cases with zero intervening words I kept those in which *sembli/sembla* was followed by one of the following: *mentida* 'lie', *increïble* 'unbelievable', *contradictori* 'contradictory', *estrany* 'strange', *impossible* 'impossible', *erroni* 'wrong', *el contrari* 'the contrary', *una paradoxa* 'a paradox', *paradoxal* 'paradoxical', *difícil de creure* 'hard to believe', *un contrasentit* 'a misunderstanding', *de bojos* 'of crazy people', *rar* 'weird', *un deliri* 'a delirium', *un tòpic* 'a prejudice', *poc probable* 'unlikely', *una contradicció* 'a contradiction', *incomprensible* 'incomprehensible', *il·lògic* 'illogical', *tot el contrari* 'all the contrary', *que no pot ser* 'that it can't be', *subrealista* 'unrealistic', *que no pugui ser* 'that it couldn't be', *una incongruència* 'an incongruence', *una digressió* 'a digression', *absurd* 'absurd'. The rest of the data was discarded.

Out of the clean dataset I drew a random sample of 60 items. I manually excluded all the items that did not have a full sentence following the concessive or contained sensitive sexual, religious or political content. From the remaining data I drew another random sample of 36 items, which constitute the final experimental stimuli.

4.3.3.3 Data modification

The only counterbalanced variable is once again DEICTIC CENTER. MODIFIER and CONSTRUCTION were introduced by randomly assigning each item to one of four equal groups for the four different MODIFIERS (*cert*, *clar*, *evident* and *segur*). Each of these groups was then randomly split into three equal subsets for the CONSTRUCTION (Adv, AdvC and EsAdjC). The result is a set of 12 groups for all the MODIFIER and CONSTRUCTION combinations, each of which contains three data points. The CONSTRUCTION and MODIFIER were added to the data at the beginning of the sentence that follows the concessive.

Just as in Spanish, the Catalan MODIFIERS can appear with *ser* and *estar* in EsAdjC and with or without the derivational morpheme *-ment* in AdvC. In the present experiment, this variation was added to the stimuli respecting the proportion of their distribution in the caWac corpus (cf. Tables 4.11 and 4.12). They were randomly assigned to the experimental stimuli. Neither of the two variables had any effect in the model.

To create the two conditions, the concessives were substituted by *Encara que tu no t'ho creguis* 'Although you don't believe it' for the *tú*-DEICTIC CENTER and *Encara que la gent no s'ho cregui* 'Although the people don't believe it' for the

Table 4.11: Derived vs. underived adverbs in AdvC in the experimental stimuli

	cert	*clar*	*evident*	*segur*
Adv-*ment* C	1	0	3	0
Adv-*0* C	2	3	0	3

Table 4.12: *ser* vs. *estar* in EsAdjC in the experimental stimuli

	cert	*clar*	*evident*	*segur*
ser	3	2	3	2
estar	0	1	0	1

gente-DEICTIC CENTER. The *yo*-DEICTIC CENTER was not included since the experiment on Spanish had already shown that this DEICTIC CENTER is generally unacceptable with all the MODIFIERS tested.

The sentence following the concessive was left at its full length, i.e. the right context ends at the full stop. The length of the left part of the context was selected individually for each stimulus. The selection was carried out with the aim of maintaining the minimum amount of context necessary to make sense of the target sentences. Each stimulus once again consists of a target sentence (concessive+modified sentence) and its context.

4.3.3.4 Experimental design

The experimental design was the same as in the previous experiment. It was run on Ibexfarm. It started with two practice items followed by 36 experimental stimuli. The experiment was in a Latin square design. Each stimulus was presented to each participant in only one of the two conditions, with the stimuli presented in a random order. The acceptability judgments were elicited on a ten-point Likert scale, with 1 translating to the lowest and 10 to the highest degree of acceptability. This more granular scale (with 10 instead of 5 points) was employed to see whether this would lead to more nuanced judgments. The results, however, show that this was not the case (cf. §4.3.3.8).

The participants were instructed to provide judgments on the naturalness of the underlined concessives in the relevant contexts. They were also informed that the data were taken from online corpora and were asked to disregard aspects

that did not correspond to the norm. The experiment was estimated to take 15 minutes (mean duration 16.8 minutes). No instructions were given as to whether the participants should provide the judgments quickly or slowly.

4.3.3.5 Participants

The participants were recruited through social media by sharing the link with personal contacts and in Facebook groups for Catalan native speakers. A total of 24 participants took part in the experiment, 15 female and 9 male. The majority of the participants came from Catalonia (16), 6 from Valencia and 2 from the Balearic Islands. The age of the participants ranges from 19 to 71 with a mean age of 38,9. The percentage of participants over the age of 60 (21%) is even higher than in the previous experiment.

4.3.3.6 Conditions

The stimuli are presented in two conditions. Each condition negates one reading of the MODIFIER. The first condition (4a) negates the reading in which the MODIFIER is centered on the interlocutor (*tú*-DEICTIC CENTER). The second condition (4b) negates the reading in which the MODIFIER is centered on a more general group of people (*gente*-DEICTIC CENTER). The third condition from the previous experiment, in which the DEICTIC CENTER on the speaker is negated, is not included in the present experiment.

(4) a. (*tú*-DEICTIC CENTER)
Encara que no t'ho creguis, és evident que
Although that not you-it believe.2SG.SBJV.PRS be.3SG.PRS evident that
aquest equip també pot perdre.
this team also can.3SG.PRS lose
'Although you don't believe it, it's evident that this team can also lose.'

b. (*gente*-DEICTIC CENTER)
Encara que la gent no s'ho cregui,
Although that not the people them-it believe.3SG.SBJV.PRS
és evident que aquest equip també pot perdre.
be.3SG.PRS evident that this team also can.3SG.PRS lose
'Although the people don't believe it, it's evident that this team can also lose.'

The aim of the study is to determine whether the acceptability of these DEICTIC CENTERS differs depending on the MODIFIER and CONSTRUCTION. The experiment on Spanish provided strong evidence that DEICTIC CENTER has an effect, but no interaction between DEICTIC CENTER and CONSTRUCTION or MODIFIER could be identified.

4.3.3.7 Description of the data and predictors

The judgments of the participants were provided on a ten-point Likert scale. Table 4.13 shows the percentage of the participants' judgments per value of the scale. The judgments fall mostly at the higher end of the scale. 46.14% of the elicited judgments are of a value of 8 or higher. This suggests that the participants found most of the stimuli felicitous.

Table 4.13: Percentage of judgments per rating

rating	1	2	3	4	5
percentage of judgments	12.68	9.66	7.6	6.16	6.16
rating	6	7	8	9	10
percentage of judgments	3.62	7.97	13.29	13.53	19.32

Again, in the statistical analysis I used standardized and not raw judgments.

Apart from the strictly linguistic variables, there are other variables found in the stimuli, that have their origins in the fact that the data were sourced from a corpus. One of them is the length of the text fragments. Length is measured in number of words. Of the variables encoding length, only CONTEXT LENGTH turned out to be a significant predictor in the model. The range of the variable is from 49 (cf. 5a) to 751 (cf. 5b) words per context.

(5) a. *Context:* Per què les crispetes són tan cares en el cinema?
'Why is popcorn so expensive at the cinema?'
Target sentence: Encara que la gent no s'ho creguis, clar que la resolució d'aquesta pregunta és un dels problemes recurrents que es plantejen en economia.
'Although people in general don't believe it, clearly the solution to this question is one of the current issues in economics.'

b. *Context:* Imagineu-vos una empresa de serveis amb diversos centres de producció. L'empresa entra en pèrdues i el consell d'administració

no té més diners per a invertir. Ordenen al director executiu reducció de despeses. Empresa A: es redueixen els costos d'estructura rebaixant personal improductiu. Es fa un estudi de reorganització administrativa que estalvïi processos, i es procura produir dintre tot el que, fins aleshores, es donava a fer a empreses exteriors. Això fa augmentar la productivitat i reduir les despeses. Empresa B: Es rebaixa la producció interna. Es manté el personal improductiu i així es redueix la productivitat del "productiu". Es continua amb la mateixa gestió administrativa, i es manté el donar feina fora a altres tallers.
'Imagine a company in the service sector with various production centers. The company starts to suffer losses and the board of directors doesn't have any money in order to intervene. They order the executive director to reduce expenses. Company A: The structural costs are reduced by letting go of unproductive personnel. They make a study and reorganize the administration in order to economize processes and they try to keep up production, including in the areas that used to be given to external companies. This makes the productivity increase and reduces the expenses. Company B: Reduces in-house production. The unproductive personnel are kept and thus what is reduced is the productivity. They stick to the same administrative management and they keep on giving work to external factories.'
Target sentence: Encara que tu no t'ho creguis, segur que l'Empresa B és la sanitat catalana, aquesta és la realitat dels directors dels hospitals catalans.
'Although you don't believe it, surely Company B is the Catalan Health Department and this is the reality of many Catalan hospital directors.'

The participants were not instructed to move quickly or slowly through the experiment. REACTION TIME did not turn out to be a significant predictor in the model. Again, the correlation between the overall LENGTH OF THE ITEM and the REACTION TIME is small (0.18).

The bar charts in Figure 4.13 plot the judgments for the CONSTRUCTIONS in the two DEICTIC CENTER conditions. They show patterns that differ from those observed in the Spanish results (cf. Figure 4.8). They suggest no DEICTIC CENTER effect. The distribution of the data on the horizontal axes is nearly identical. There is a clear difference between the judgments provided for the lowest and highest acceptability. For EsAdjC, there is a large difference between the percentages for

the highest value and those for the lowest. This suggests that the participants were fairly certain about the acceptability of this CONSTRUCTION. The contrast is even bigger in the *tú*-DEICTIC CENTER. The difference between the high and the low ratings is far smaller in Adv and AdvC.

The plots in Figure 4.14 show the judgments for each MODIFIER in the two DEICTIC CENTER conditions. They indicate again that DEICTIC CENTER does not have a strong effect on the judgments in the Catalan experiment. Most of the bar charts are symmetrical with high percentages at the extremes of the scale and low percentages at the center. The bar charts for *segur* are very similar in the two groups. There are higher percentages at the right extreme of the scale, which means that the participants found *segur* acceptable in both conditions. The pattern is similar for *cert*, but here the percentages of the highest ratings are larger in the *gente*-DEICTIC CENTER than in the *tú*-DEICTIC CENTER. This could suggest that *cert* is commonly interpreted as centered on speech participants, and therefore it is less acceptable to negate the *tú*-DEICTIC CENTER than the *gente*-DEICTIC CENTER. *Evident* shows a very similar distribution for the *gente*-DEICTIC CENTER, so the perspective of "the people" does not appear to be encoded in the meaning of this MODIFIER. The bar chart for the *tú*-DEICTIC CENTER is practically symmetrical, which indicates that the participants are uncertain about the acceptability of *evident* in a reading that negates the interlocutor's perspective. The judgments for *clar* peak at the lower extreme in the *gente*-DEICTIC CENTER, while in the *tú*-DEICTIC CENTER the lowest and the highest ratings have the same percentage.

Finally, the last set of bar charts in Figure 4.15 shows the judgments for each MODIFIER in the different CONSTRUCTIONS. They indicate that some MODIFIER-CONSTRUCTION combinations are preferred (*segur* in Adv, *cert* and *evident* in EsAdjC) while others are dispreferred (*evident* and *clar* in Adv).

4.3.3.8 Results

This section presents the conditional inference tree that was modeled on the basis of the experimental data. Model fitting and selection was exploratory. The model is shown in Figure 4.16(a). There are three variables that show a significant effect: CONSTRUCTION, MODIFIER and NCONTEXT, the length of the context measured in number of words. DEICTIC CENTER is not a significant predictor in this model. The dot plot in Figure 4.16(b), based on a random forest, shows that CONSTRUCTION is the most important variable.

The accuracy (10.14%) is low. This is a result of the way in which accuracy was calculated and the fact that the model predicts a reduced scale. It reduces the 10 values of the observed variables to just a 5 level opposition in the predicted values as can be seen in the heat map in Figure 4.17.

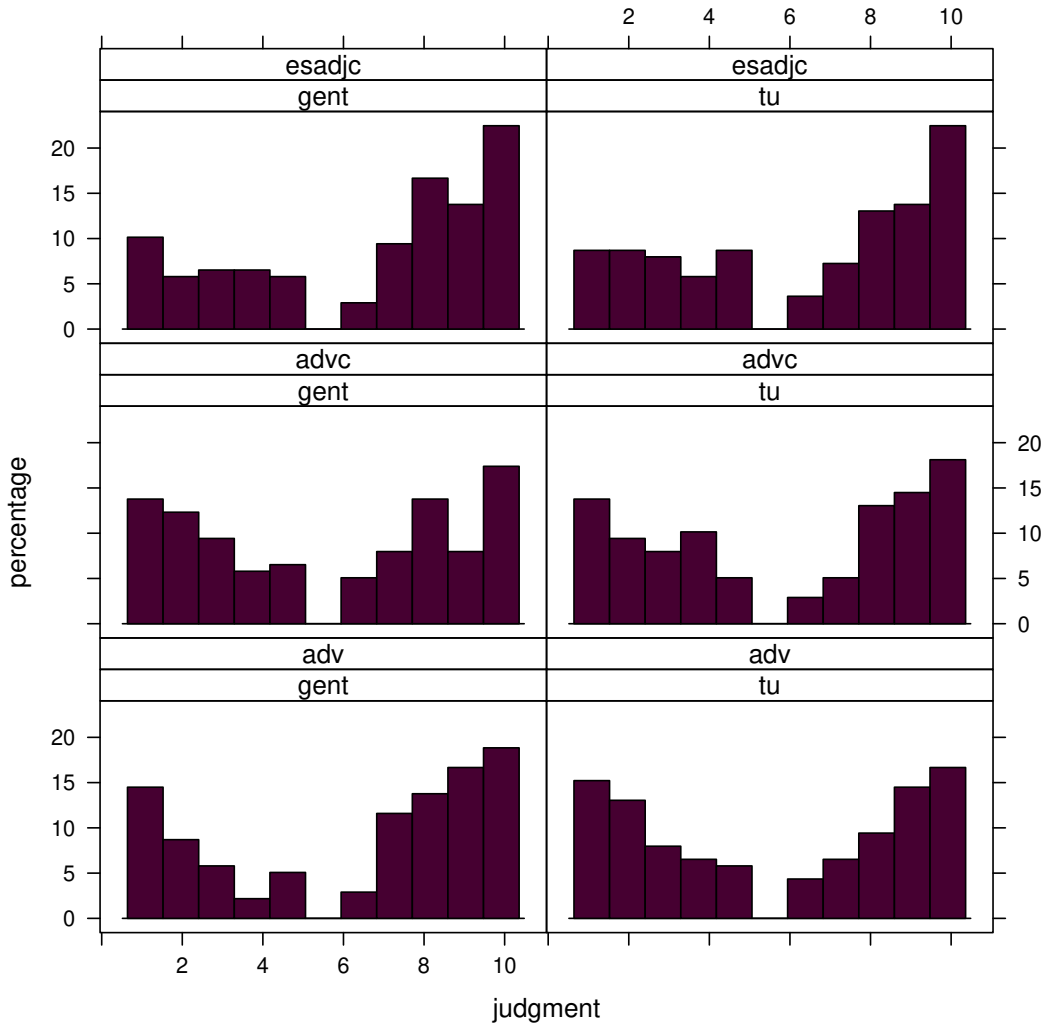


Figure 4.13: Individual bar charts for judgments per DEICTIC CENTER per CONSTRUCTION. The panels in the first row show the judgments for EsAdjC in the different DEICTIC CENTER conditions. The second row shows AdvC and the third row Adv.

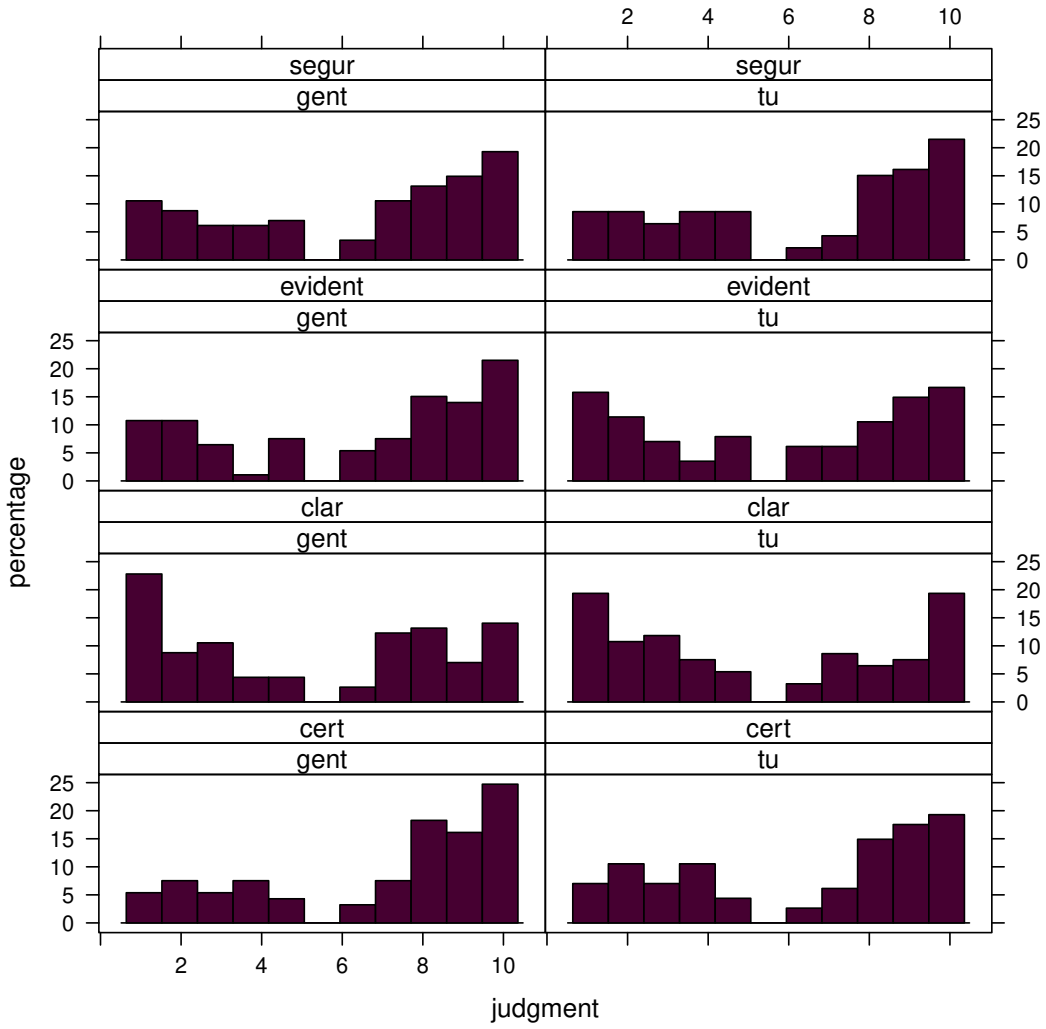


Figure 4.14: Individual bar charts for judgments per DEICTIC CENTER per MODIFIER. The panels in the first row show the judgments for *segur* in the different DEICTIC CENTER conditions. The second row shows *evident*, the third row *clar* and the last row Cert.

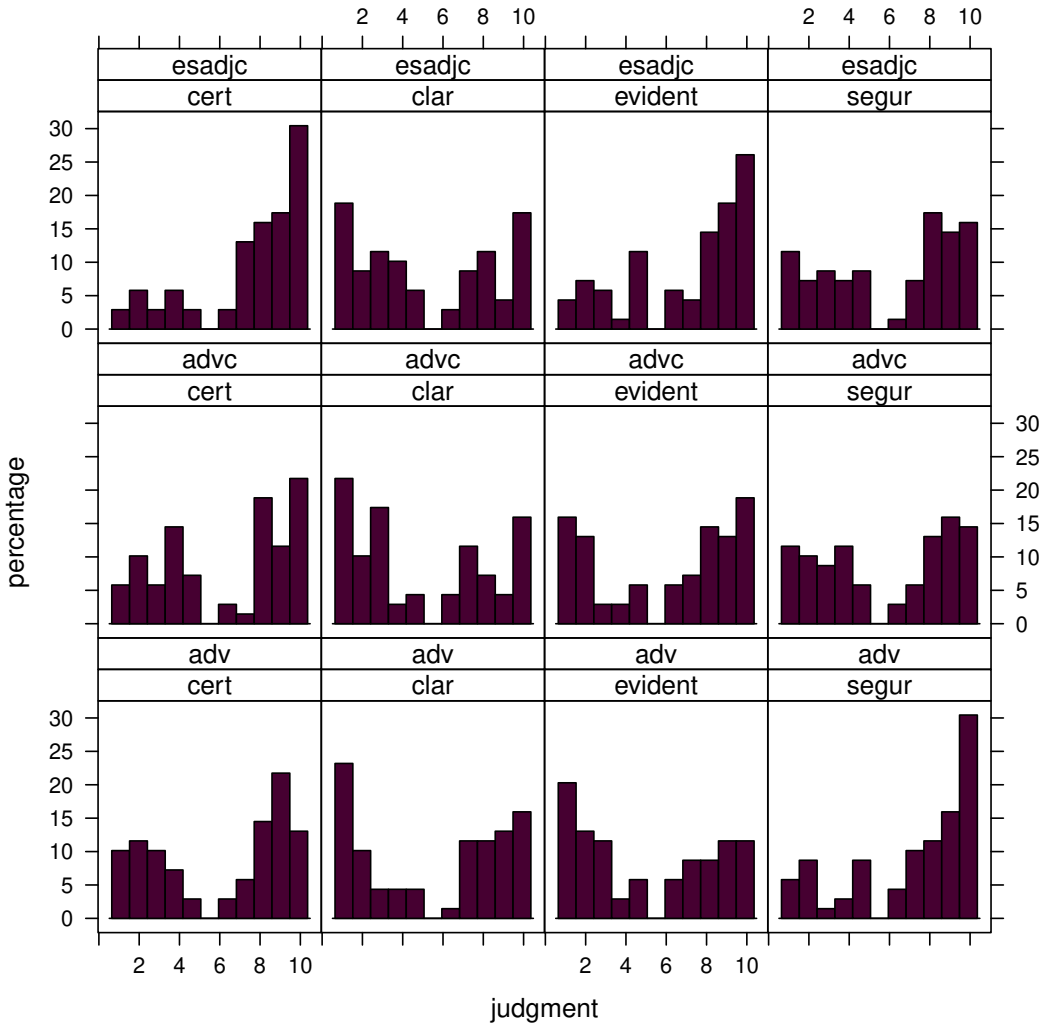


Figure 4.15: Individual bar charts for judgments per CONSTRUCTION per MODIFIER. The panels in the first row show the judgments for EsAdjC in the different DEICTIC CENTER conditions. The second row shows AdvC and the third row Adv.

The first split separates the data containing the modifier *clar* from the rest. MODIFIER *clar* interacts with NCONTEXT: A larger number of words in the context leads to a significantly higher acceptability.

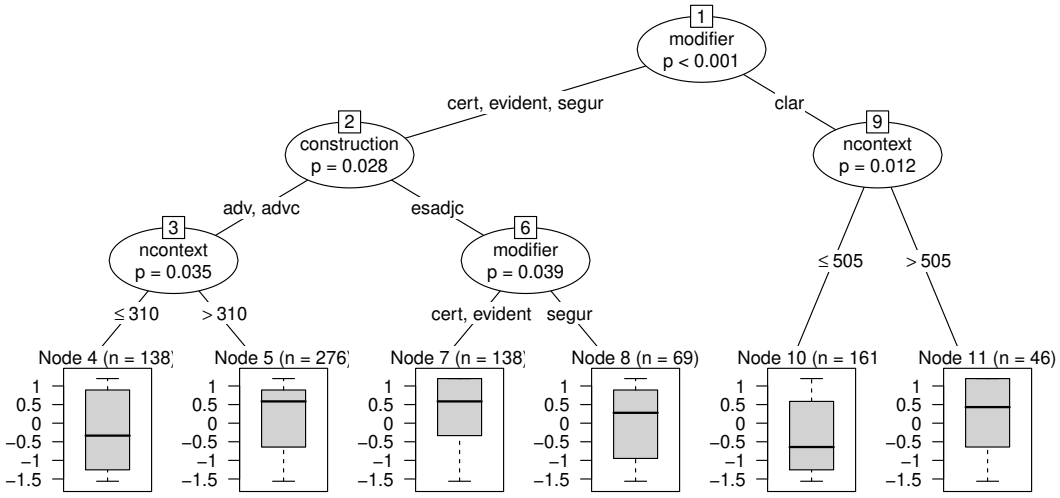
Within the subset of data containing all modifiers but *clar*, the model's strongest variable CONSTRUCTION, splits EsAdjC from Adv and AdvC. The conditional inference tree also reveals a three-way interaction between MODIFIER, CONSTRUCTION and NCONTEXT. For Adv and AdvC there is again an effect of NCONTEXT. In much the same way as before, a larger number of words in the context results in higher acceptability.

4.3.3.9 Discussion

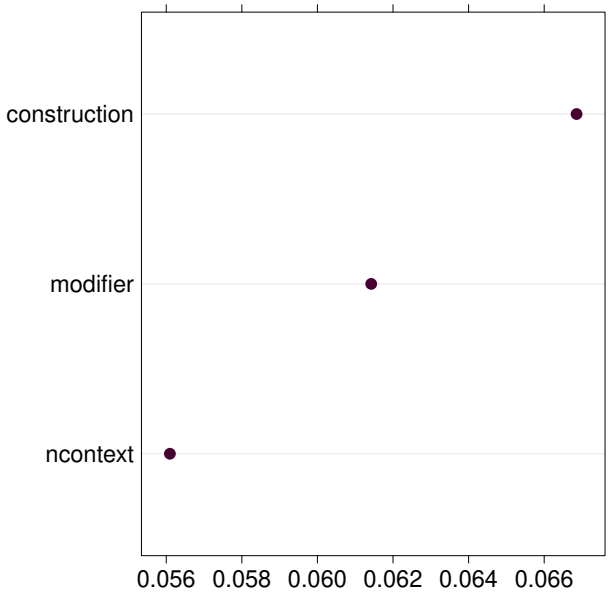
The model suggests that the modifier *clar* is different from the others. The results show that *clar* is less acceptable than the other modifiers in all the contexts tested. Just as in the previous two studies, this suggests that *clar(o)* is unusual in some respect. In the present experiment, the highest degree of acceptability for the MODIFIER is in Adv (mean judgment: 5.55), followed by EsAdjC (mean judgment: 5.25) and the lowest is in AdvC (mean judgment: 4.91). Both conditions tested in the experiment negate the perspective of the interlocutor, which is predicted to be less acceptable with AdvC in the first place but also with the other attributive construction EsAdjC. Finally, Adv should permit all readings of the modifier and therefore a high acceptability in all conditions is expected. This means that the judgments provided for stimuli containing *clar* are in line with the theoretical assumptions. In the corpus data *clar* in AdvC is very frequent in caWac (2303 cases), while in comparison, it occurs as a sentence-initial adverb (Adv) only 387 times in caWac. The highest frequency for *clar* is in EsAdjC (7410, of which 5062 contain the copula *ser*), though it is possible that some of these 5062 are actually AdvC rather then EsAdjC. In recent years *esclar* has emerged as a new evidential modifier resulting from a univerbation of the copula construction *és clar* (see the discussion in §3.3.4). *Esclar* does not yet appear in DIEC (*Diccionarí de la llengua catalana de l'Institut d'Estudis Catalans*), but has been adopted as a popular norm. It appears in the writing of some Catalan authors and is used by various newspapers, which demonstrates the difference in function and interpretation between *és clar* and *esclar*.[6] There is a substantial number of cases (3579) of *esclar* in the corpus, of which 33% are cases of *esclar que*. Some speakers, however, refrain from adopting the spelling *esclar*.[7] One can therefore assume that a number of the occurrences of the evidential adverb *esclar* still appear in the more conservative spelling *és clar* in the caWac corpus. Finally, it is also possible that some of

[6] See for instance https://www.ara.cat/cronica/Aglutinacio-clar-esclar_0_444555554.html.

[7] http://www.elpuntavui.cat/article/7-vista/23-lectorescriu/201549-esclar-o-es-clar.html.



(a) conditional inference tree



(b) variable importance

Figure 4.16: Conditional inference tree with splits on CONSTRUCTION, MODIFIER and NCONTEXT and a dot plot showing the importance of each variable

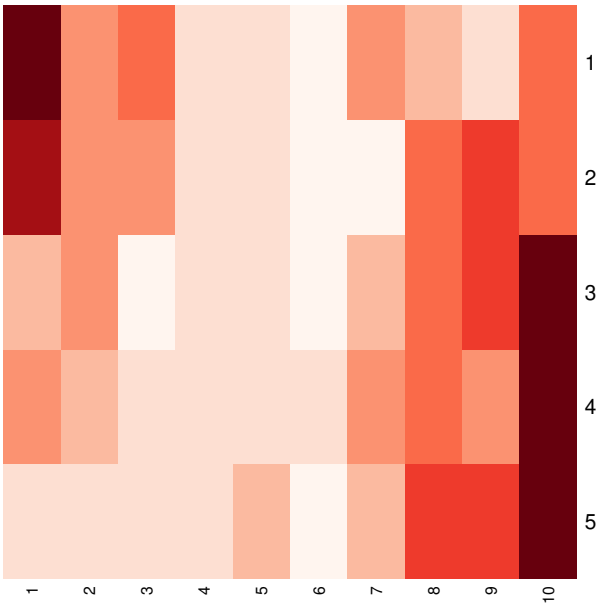


Figure 4.17: Heat map of 5 predicted and 10 observed values; Darker shades correspond to larger counts.

the participants in the experiment interpreted the sequence *és clar*, intended as EsAdjC, as cases of AdvC.

The positive effect of context length on the judgments suggest that more information facilitates the interpretation of the epistemic and evidential MODIFIERS. Or, conversely, it could mean that epistemically or evidentially modified statements are bad in out-of-the-blue contexts.

In the present experiment the model predicts fewer values for the response variable than the original scale. It is not clear whether this means that the participants were certain about the (un)acceptability of the stimuli, making the intermediate levels obsolete, or whether the intermediate levels were not interpretable for the participants.

The moderate fit of the model and the low number of significant variables is also due to the low number of participants. Attempts to reach a larger number of participants in a reasonable amount of time unfortunately failed.

Neither of the two experiments provided results that support direct interaction between CONSTRUCTION and DEICTIC CENTER. The expectation based on the meaning of *que* in AdvC would have been that this CONSTRUCTION should be dispreferred in the *tú*-DEICTIC CENTER, which denies that the interlocutor shares the evaluation. EsAdjC should have been dispreferred in the *gente*-DEICTIC CENTER,

which negates the perspective of a contextually-relevant group of people. A relation between the DEICTIC CENTER and the CONSTRUCTION is however supported by the corpus data. This shows that there is a striking mismatch between production and comprehension. While there might be a preferred DEICTIC CENTER for the MODIFIERS in each CONSTRUCTION, which is suggested by the corpus data, other readings can easily be accommodated, which makes it hard to isolate them in an experimental setting. One of the findings is therefore that other methods that take into account the possibility of accommodation need to be developed in order to test the preferred DEICTIC CENTERS.

4.4 Bias in polar questions

This section focuses on *que*-initial polar questions in Catalan. A prediction of the analysis in §3.4 is that the presence of attributive *que* expresses that the speaker is biased and expects a positive answer from the hearer. The motivation behind the empirical investigation that I will describe here, is to determine how the presence and absence of *que*, along with other linguistic and contextual factors, impact the bias.

Although *que*-initial polar questions are also attested in Spanish (cf. §3.3.2), I have only focused on Catalan here. Corpus data showed that *que*-initial questions are allowed in seemingly the same contexts as in Catalan. However, unlike Catalan, this feature of Spanish grammar does not appear in descriptive grammars of the language, suggesting that there is little awareness of this construction. This could distort the results of an experiment that relies on written stimuli, because, even if the participants use *que*-initial questions in oral registers, they might not accept them in written form.

4.4.1 Corpus

Just as in the experiment described in §4.3.3, the stimuli were sampled from the caWac corpus, which comprises 780 million tokens and was built through a web crawl.

4.4.2 Data acquisition and selection

The caWac corpus if freely available to download. Apart from sentence and paragraph splits, the data contain no annotations. The final experimental stimuli were again found through stepwise sampling, random sampling and controlled selection.

The preparation of the dataset and the sampling was carried out using Python. I extracted all the questions (defined as sentences ending in a question mark) and their preceding 400 words from the corpus. I then split the dataset containing all the questions introduced by *que* from the rest of the data. The remaining data were further reduced by discarding *wh*-questions (defined as questions starting with an interrogative pronoun). I drew a random sample of 30 items from each dataset. The random samples were then manually checked and items were discarded if they constituted negative questions, were not formally complete, contained sensitive content or if they lacked cohesion with the previous context. From the remaining data I drew a random sample of 20 items per dataset, which constitute the experimental stimuli.

4.4.3 Data modification

The experiment contains two counterbalanced variables that were introduced through modification of the corpus data. The first is the presence and absence of sentence-initial *que*. This variable was created by preparing two versions of each stimulus, one with and one without the complementizer. The way this variable is set up makes it possible to investigate the import of *que* isolated from the corresponding question and the context in which it appears. For every question that originally contained *que*, there is a also *que*-less counterpart in the experimental stimuli and vice versa.

The second counterbalanced variable was created by complementing each item with a "yes" or "no" answer to the question. The answer was not just the answer particle on its own, but repeated the prejacent (cf. 6). The motivation was to prevent ambiguity with syntactically complex questions where a simple particle answer could be interpreted as affirming or negating either the main or the embedded content.

(6) *Context:*
A – Estàs pensant en marxar a estudiar a l'estranger però en tens alguns dubtes?
'A – Are you thinking about leaving to study in a foreign country but you have some doubts?'
Target Question:
Version 1: A – Que t'agradaria estar més informat sobre els programes de mobilitat?
Version 2: A – T'agradaria estar més informat sobre els programes de mobilitat?

‘A – Would you like to be more informed about mobility programs?’
Answer:
Version 1: B – Sí, m’agradaria.
‘B – Yes, I would like to be.
Version 2: B – No, no m’agradaria.
‘B – No, I wouldn’t like to be.

Each stimulus is a short dialog that consists of the context, the target question asked by one speaker and an answer provided by another speaker. The context is either the sentences uttered by the first speaker before asking the question or an interaction between two or more speakers. The contexts were again chosen by qualitative rather then quantitative criteria with the aim of maintaining the minimum number of words necessary to make sense of the question.

4.4.4 Experimental design

The experiment was run online on Ibexfarm. It started with three practice items and was followed by 40 experimental items. The experiment is in a Latin square design, which means that each participant was only presented with each stimulus in one of the four conditions. The stimuli were presented in a random order. The participants provided judgments on a five-point Likert scale, with 1 translating to the lowest degree and 5 to the highest degree. The task was to judge the expectedness of a “yes”/“no” answer taking into consideration the form of the question and the context. In a pilot phase I tested two versions of the experiment. In one of them, the participants were asked to judge the degree of expectedness and in the other to judge the degree of surprise. In the pilots, the participants showed more difficulty in judging the degree of surprise than the degree of expectedness, which is why I settled on the second option. The mean duration of the experiment is 18 minutes.

4.4.5 Participant

A total of 46 native Catalan speakers participated in the experiment. Recruitment through social media, which I relied on in the previous experiments, was less successful in this case (it only resulted in the recruitment of 12 participants). The additional 34 participants were hence recruited via Prolific,[8] a platform similar to Amazon Mechanical Turk, which enables researchers to recruit participants online and compensate them for their time. The payment was between £1 and £3.

[8] https://www.prolific.co/

The number of participants is relatively modest given the complexity of the experimental design. Although the experiment was online for several months and I distributed the link widely, a larger number of participants could not be achieved, in part because the number of native Catalan speakers registered on Prolific is very small. There are 21 female and 25 male participants. The vast majority (42) are from Catalonia and 2 each are from the Balearic Islands and Valencia. The age ranges from 18 to 61 years (mean = 31). There is only one participant above the age of 60, and 8.6% are above the age of 50. This means that the participants are younger than in the other experiments. This is probably because the bulk of the recruitment was carried out via Prolific compared to the Facebook-based recruitment in the other two experiments. Prolific attracted a younger group of people, while Facebook has members of all age groups, and the older cohorts proved to be particularly active in the Facebook groups I relied on for distributing the experiments.

4.4.6 Conditions

The stimuli are presented in four conditions resulting from the combinations of the two counterbalanced variables PRESENCE OF QUE and ANSWER (cf. Table 4.14).

Table 4.14: Variables and conditions

	que	no *que*
affirmative answer	C1	C2
negative answer	C3	C4

There are four versions of each item. This is illustrated in Table 4.15 for the stimuli in (6). In condition 1, the question is introduced by *que* and the answer is affirmative. In condition 2, the question is not introduced by *que* and the answer is also affirmative. In condition 3, the question is introduced by *que* and followed by a negative answer. In condition 4, the question is not introduced by *que* and the answer is negative.

The motivation for introducing ANSWER as a variable is to investigate the bias that results from different configurations. The expectation based on the theoretical analysis is that questions introduced by *que* express that the speaker expects a positive answer. This means that C1, where a *que*-question is answered affirmatively, should lead to high acceptability judgments, reflecting the fact that a positive answer is expected. However, in C3, where the answer is negative, we

Table 4.15: Exemplified conditions

	que	no *que*
affirmative answer	C1 A – Que t'agradaria estar més informat sobre els programes de mobilitat? B – Sí, m'agradaria.	C2 A – T'agradaria estar més informat sobre els programes de mobilitat? B – Sí, m'agradaria.
negative answer	C3 A – Que t'agradaria estar més informat sobre els programes de mobilitat? B – No, no m'agradaria.	C4 A – T'agradaria estar més informat sobre els programes de mobilitat? B – No, no m'agradaria.
translation	A – (QUE) Would you like to have more information about mobility programs? B - Yes, I'd like to. / No, I wouldn't like to.	

should expect a low degree of expectedness on the part of speaker A, which should translate to low acceptability. In C2 and C4, the *que*-less questions are neutral and should not per se indicate that the speaker has a bias towards a positive or negative answer. C2 and C4 should not differ because the form of the question does not encode a speaker bias, and so neither a positive nor a negative answer should be expected. I should stress that these are the outcomes that could be expected in a highly controlled hypothesis-driven experiment, but the exploratory design that I have employed here is unlikely to work in the same way. Given that the stimuli are taken from naturally occurring data with inherent variation, the present experiment allows for the possibility that further properties, such as word order, the presence of modifiers or contextual factors, can influence the judgments and give rise to different readings of the questions.

These factors and their interplay with the complementizer can be tested because every target question appears with and without *que*. This means that we can investigate whether the differences between the tested questions go beyond the presence or absence of *que*. In turn, this allows us to study, for instance, whether an original *que*-question remains biased even when *que* is not present.

4.4.7 Description of the data and predictors

The experiment again elicits acceptability judgments on a five-point Likert scale. As in the previous experiments, the largest values are at the extremes of the scale. In this case, they are skewed towards to the top end (cf. Table 4.16).

Table 4.16: Percentage of judgments per rating

rating	1	2	3	4	5
percentage of judgments	23.21	12.45	12.61	13.75	37.99

The response variable in the model discussed below is a standardized version of the raw judgments.

The participants were instructed to choose the value 5 if the answer to the question was completely expected by the speaker asking the question. The extreme degree to which this value was chosen, however, could indicate that the participants did not stick to the instruction. It is possible that they used value 5 to express that the answer was not-unexpected, grouping together cases that are truly expected and cases where there were no prior expectations.

The stimuli are drawn from corpus data and are hence heterogenous by design. There is a large degree of variation in CONTEXT LENGTH (between 19 and 1003, mean = 536), but this did not turn out to be a significant predictor. Moreover, the correlation between REACTION TIME and STIMULI LENGTH (sum of context, target sentence and answer) is low (10.2%). Neither REACTION TIME nor STIMULI LENGTH is a significant predictor in the models.

Different predictors were considered for the statistical model. The counterbalanced variable, PRESENCE OF QUE, is not a significant predictor. The variable ANSWER is significant. The bar charts in Figure 4.18 show the percentage of the raw judgments by PRESENCE OF QUE by ANSWER. The charts suggest that we might find a positive bias, i.e. an expectation of a positive answer, with both types of polar questions. The expectation of a positive answer appears to be stronger in the absence of *que*. The unexpectedness of a negative answer, i.e. the percentage of judgments of value 1 for negative answers, point in a direction predicted by the theory: Although negative answers are judged to be unexpected in more than a quarter of cases for both types of polar questions, the percentage of these judgments is higher when *que* is present (28.7%).

The observed contrasts are even more pronounced in the bar charts in Figure 4.19, which plot the percentages of judgments per ANSWER by ORIGINAL, instead of PRESENCE OF QUE. The variable ORIGINAL has two values: ORIGINAL-QUE,

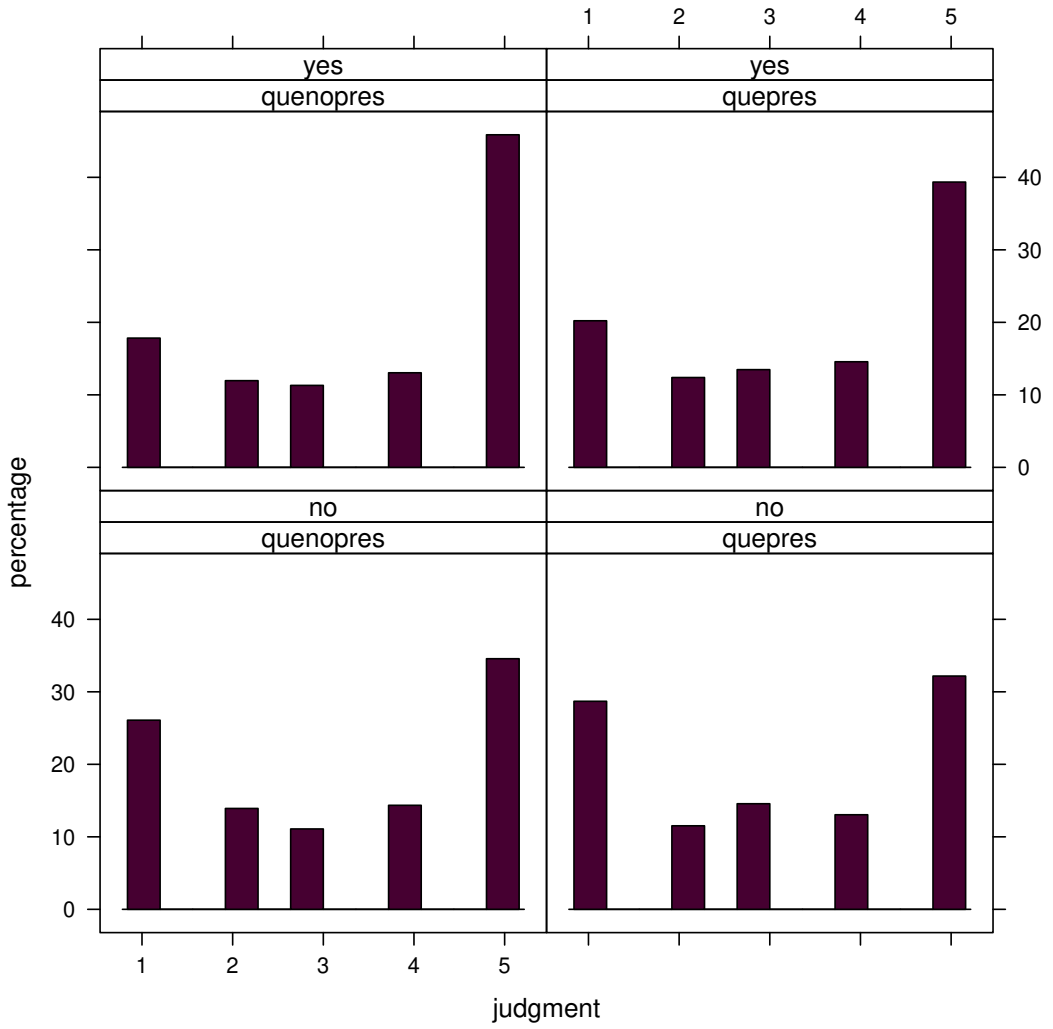


Figure 4.18: Individual bar charts for judgments per PRESENCE OF QUE per ANSWER. The panels in the first row show the judgments when the answer is positive depending on the presence of *que*. The second row shows the judgments when the answer is negative.

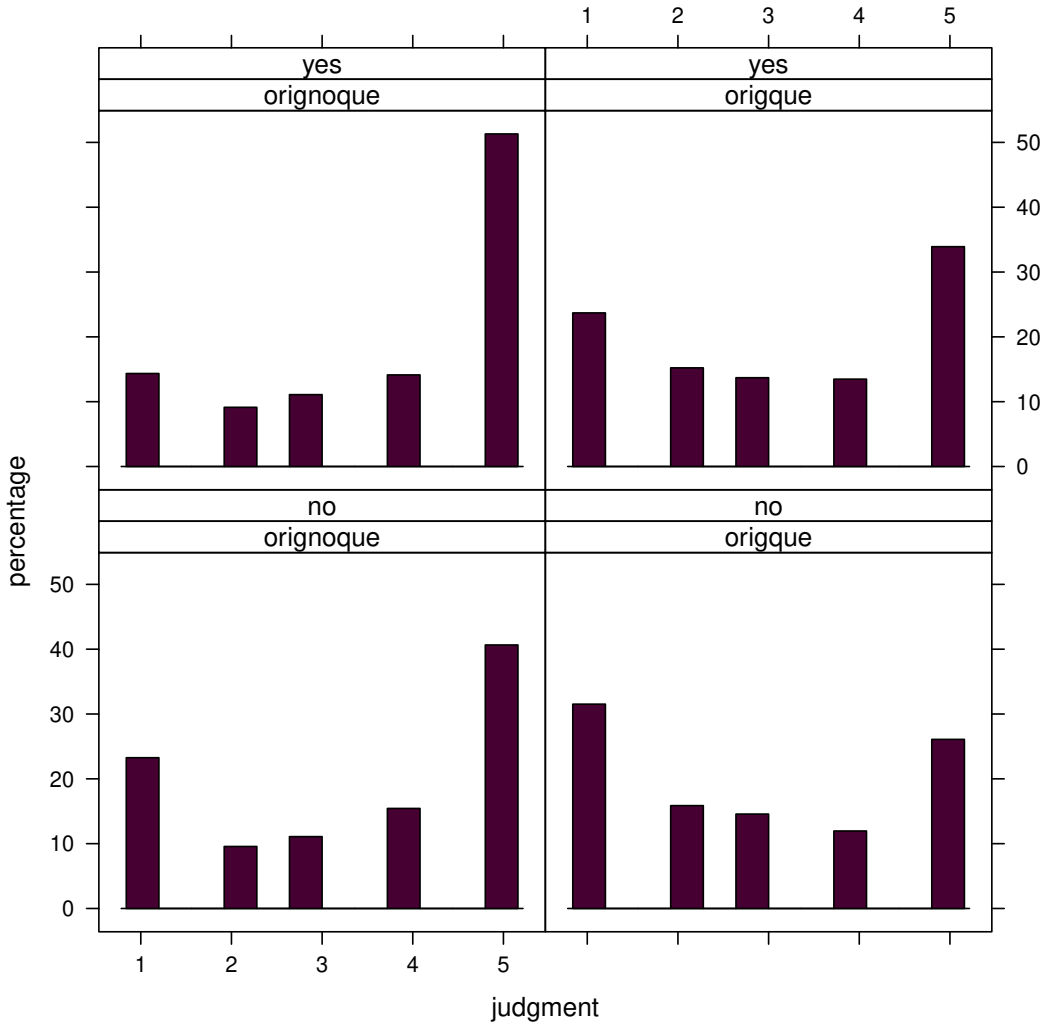


Figure 4.19: Individual bar charts for judgments per ORIGINAL per ANSWER. The panels in the first row show the judgments when the answer is positive depending on whether or not the original version contained *que*. The second row shows the judgments when the answer is negative.

if the original question contained *que* and ORIGINAL-NOQUE, if it did not. This variable is a significant predictor in the model described below. The bar charts indicate that the positive bias is most pronounced for contexts with questions that did not originally contain *que*. However, negative answers are also judged highly in ORIGINAL-NOQUE contexts. This might suggest again that the participants employed value 5 if they considered the answer not-unexpected. The positive answer also appears highly expected in ORIGINAL-QUE contexts. The high degree of unexpectedness of a negative answer in the ORIGINAL-QUE contexts is once again notable.

The bar charts in Figure 4.20 plot the percentage of judgments per category by PRESENCE OF QUE by ORIGINAL. They show once more that ORIGINAL has a stronger impact than PRESENCE OF QUE: The contrast between the two bar charts on the vertical axis is greater than that between the two bar charts on the horizontal axis.

There are three further variables that turned out to be significant predictors in the model. The first is called REALSPEAKER and was created post hoc. It has the value REAL if the interlocutors are introduced by a proper name, and NOT REAL if they are encoded as "speaker A", "speaker B". When creating the stimuli, I used the latter in the cases where no interlocutors were addressed directly, which was the majority of the examples. In the contexts where the interlocutors were introduced with proper names, I adopted those.

The second variable is termed BIAS MARKING. It has three values: NONE, if there is no marking, WO for word order, if the subject-verb inversion typical of unbiased polar questions is not observed, and MOD for modal, if the question contains modal expressions. Both of these properties can give rise to a bias. In (7) a stimulus carrying the value WO is illustrated. The word order is that of a declarative and not that of a polar question. In the pilot phase of the experiment, informants suggested that this word order is used when expressing a positive bias.

(7) (ORIGINAL-NOQUE)
(Que) l'Asia parla igual català que polonès?
QUE the Asia speak.3SG.PRS same Catalan as Polish
'Asia speaks Catalan and Polish equally well?

Similar suggestions were made for modal expressions. In (8), the modal adverb *potser* could itself signal a bias.

(8) (ORIGINAL-QUE)
(Que) potser s' hagués pogut reconduir la cosa?
QUE maybe CL.REFL AUX.3SG.SBJV.PST can.PTCP re-route the issue
'Could the issue maybe have been resolved otherwise?'

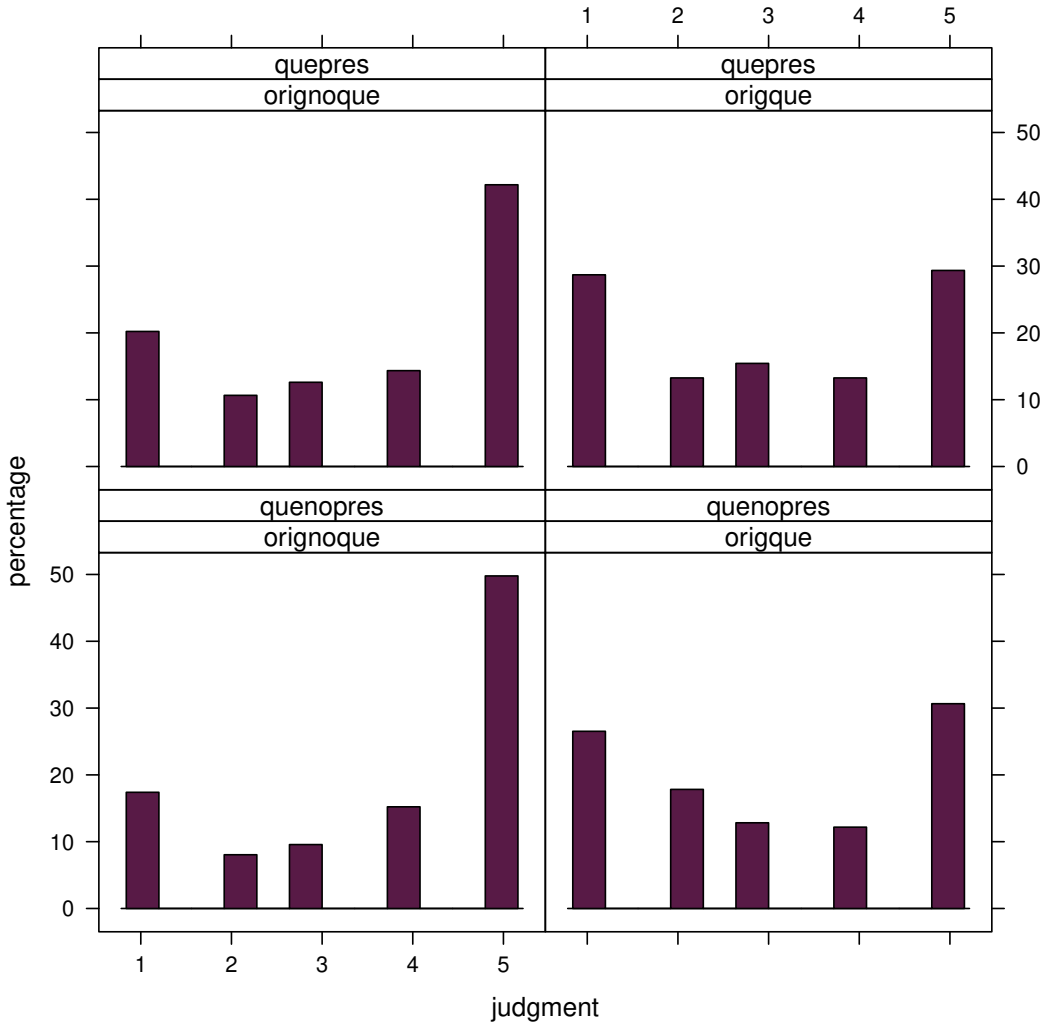


Figure 4.20: Individual bar charts for judgments per PRESENCE OF QUE per ORIGINAL VERSION. The panels in the first row show the judgments when the answer is positive depending on whether or not the original version contained *que*. The second row shows the judgments when the answer is negative.

The distribution of the values in the original data does not suggest a clear preference for these alternative bias markings in polar questions with and without *que*. The majority (9 ORIGINAL-NOQUE, 12 ORIGINAL-QUE) of cases in the original data did not have a special bias marking. Non-neutral word orders were found in 4 examples for each group and modal expressions were found in 7 ORIGINAL-NOQUE and 4 ORIGINAL-QUE questions.

The last variable is DIALOG. Its values encode the number of extra interlocutors present in addition to the addressee. The values are 1, 2 and 3.

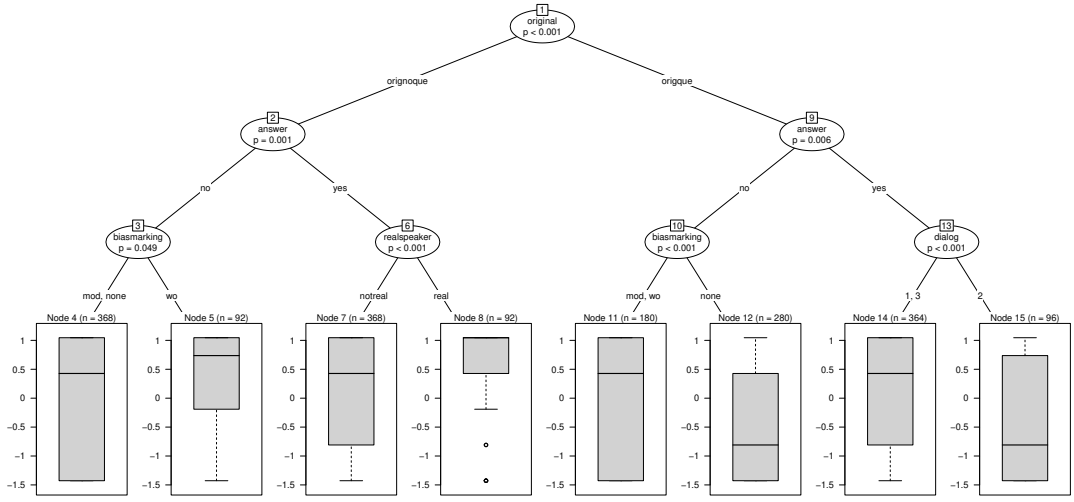
4.4.8 Results

The model described in this section, like the previous models, was established through exploratory selection based on objective measurements and the theoretical plausibility of the model's effects. The conditional inference tree (Figure 4.21(a)) shows significant effects of the variables ORIGINAL, ANSWER, REAL SPEAKER, DIALOG and BIAS MARKING.

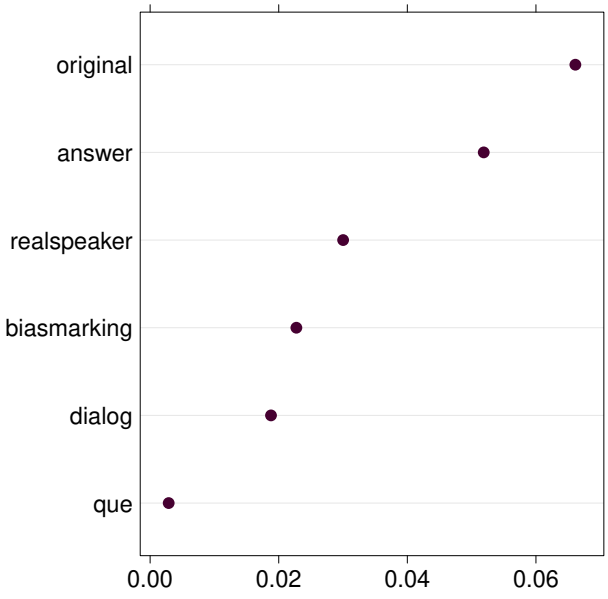
The dot plot in Figure 4.21(b) plots the variable importance calculated with a random forest model. It shows that ORIGINAL and ANSWER are the strongest predictors. The variables REAL SPEAKER, BIAS MARKING and DIALOG play a minor role. The importance of the variable PRESENCE OF *QUE*, which also entered the calculation, is virtually non-existent.

The heatmap in Figure 4.22 plots the observed vs. predicted values. The model predicts more categories than were observed. This again has an effect on the accuracy, which is relatively low (16%) for this model.

The highest split in the model in Figure 4.21(a) on the most important variable ORIGINAL, shows that there is a significant difference between the contexts in which the original version of the target question contained *que* and those in which it did not. The judgments for the ORIGINAL-QUE contexts are significantly lower. The effect is independent of whether *que* is actually present in the stimulus or not and also independent of what the answer is. The model shows complex three-way interactions. In the stimuli that originally contained no *que*, there is a significant difference between positive and negative answers. Both answers are expected, but when the answer is positive, the variable REAL SPEAKER plays an additional role. The positive bias is significantly stronger when the speakers are encoded by proper names. In the stimuli that originally contained *que*-initial polar questions, ANSWER is also the most important predictor. In the contexts where the answer is "yes", the number of interlocutors plays a role. The positive answer is significantly less expected when there are three interlocutors interacting. When the answer is negative the presence of additional bias marking boosts



(a) conditional inference tree



(b) variable importance

Figure 4.21: Conditional inference tree with splits on ORIGINAL, ANSWER, REALSPEAKER, DIALOG and BIAS MARKING and a dot plot showing the importance of each variable.

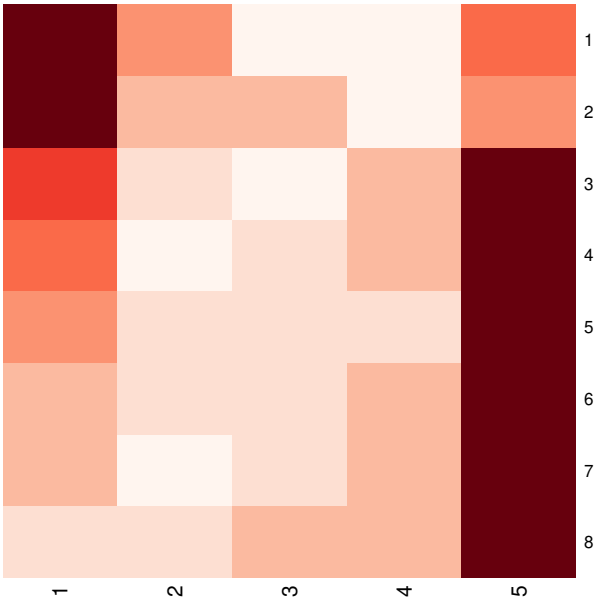


Figure 4.22: Heat map of 8 predicted values versus 5 observed values; Darker shades correspond to larger counts.

the expectation of the negative answer. Node 13 shows that a negative answer to a question without additional bias marking is unexpected in the ORIGINAL-QUE contexts.

4.4.9 Discussion

The model shows a solid effect for ORIGINAL. The main insight that I draw from these results is that the differences between bias in polar questions goes far beyond the mere presence or absence of *que*. The variable ORIGINAL has no obvious direct expression in the stimuli. I therefore conclude that the effect must be a result of multiple factors. The post hoc variables (BIAS MARKING, DIALOG, REAL SPEAKER) showed some effects. However, the low accuracy and the fact that ORIGINAL remains the strongest effect suggest that there are further factors yet to be discovered in the contexts. Another novel insight comes from the importance of the variable ANSWER. The results suggest that there is a general positive bias irrespective of the type and properties of the polar questions.

My interpretation of the importance of the variable REAL SPEAKER is that the use of proper names strengthens existing effects and can therefore be viewed as a tool to create more natural stimuli. In general, the data used as a basis here were not optimal because the caWac corpus is not made up of dialog data. Since

there are no Catalan dialog data freely available, there was no other choice but to work with these imperfect data. However, the use of proper names, appears to be a reasonable means of achieving more authentic stimuli in the future.

The model shows that in ORIGINAL-QUE contexts with no bias marking, the negative answer is unexpected. If a negative answer is unexpected, it conversely suggests that the speakers expected a positive one. Following this reasoning, the results can be taken to suggest that the contexts in which *que*-initial questions normally appear might in fact carry a positive bias. This is in line with the theoretical analysis I presented in §3.4.

4.5 General discussion

To conclude this chapter, I would like to summarize a number of issues that I faced in the empirical investigations described here. I begin with problems that arise from the nature of the languages under investigation and then turn to challenges related to the empirical methods employed. Finally, I describe how my view on the relation between theory and empirical research was informed by the results reported in this chapter.

One challenge in all three experiments was the recruitment of participants. It was extremely time consuming and despite the long process involved, I only managed to achieve a relatively moderate number of participants. This was particularly true of the experiments on Catalan. Prolific or other platforms that pool potential participants and allow them to be compensated could become useful tools. In the present study, however, even relying on Prolific did not allow me to reach enough Catalan speakers. While the numbers of Spanish and Portuguese speakers registered on the platform are definitely higher, attempting to focus on speakers of a particular Spanish or Portuguese variety will likely result in similar problems.

Another issue related to the languages being studied here is the fact that the publicly available resources are somewhat limited. This resulted in a database that was not optimal for the experiments. The results clearly showed that, while corpus data in general are surely a very useful basis for pragmatic experiments, it is not the case that all types of corpus data are adequate. On the contrary, each phenomenon requires specific types of data. The corpora I used for the experiments were web-based. These data represent an informal register and carry oral traits, which is necessary for a study of the pragmatics of phenomena typical of oral speech. The dataset proved somewhat imperfect nonetheless. In particular, the last experiment would have benefited greatly from true dialog data. The

caWac corpus is very large. This allows quantitative analysis to be carried out even for less frequent phenomena. The usefulness of the corpus is limited, however, because it does not contain metadata on the fragments or authors. This means that studying variation linked to extra-grammatical factors is simply not possible on the basis of the available data. The same is true for the web-based parts of the CdE and CdP that I also relied on in this monograph. They do contain information on the country of origin of a fragment. However, these are taken from the domain of the homepage, which is insufficient to identify the country of origin of the author of a posting.

Turning now to the experiments themselves where the tasks appear to have provoked some difficulties for the participants. In the first two experiments, the focus of the participants might have been drawn to the parts of the stimuli that introduced the condition rather than to the sentences containing the target construction. In the last experiment there are indications that the participants used the scale in a different way than was intended and not as it was explained in the practice phase of the experiment. The tasks I designed were complex and went beyond simple acceptability. The problems I faced show the need to test and develop further tasks that allow the investigation of pragmatic properties and are accessible to participants. It also shows that attentiveness and sensitivity in the analysis and interpretation of patterns in the results are an indispensable precondition for drawing conclusions even from imperfect data.

The previous point also touches on the issue of the appropriate means of measuring elicited judgments. I employed a scale, because an ordinal response variable allows for a larger range of statistical modeling than if the elicited judgments were nominal. The experiments carried out here, however, have also shown that there is a need to be aware of certain pitfalls. In all the experiments, the scales were underused, which resulted in models that predicted a different number of categories than were observed. Although this can be a telling result, it also restricts the data modeling possibilities and makes it hard to determine the goodness of fit of models based on the accuracy of the predictions.

Finally, it is important to determine how the empirical method used and described here can inform theory. The approach was exploratory, meaning that I did not set out to confirm or reject a particular theory, but to broaden my understanding of the phenomenon. In my view, the method proved successful in achieving this goal. For instance, one interesting result from the first set of investigations on different deictic centers in AdvC and other constructions is the mismatch between production (which is in accordance with the theoretical predictions) and perception (which does not confirm them). These results suggest that while there are preferred readings, it is always possible to accommodate

other interpretations. A novel insight from the experiment on the Catalan polar questions is that it is not sufficient to consider only the presence or absence of *que* when trying to understand the bias involved. In fact, it seems that it is the context in which each type of question appears that has the greatest predictive power with regard to the bias. These results would not have been possible if the empirical approach had been more traditional and had not relied on corpus data.

5 Conclusion

This monograph set out to investigate the properties of syntactically unselected complementizers in Ibero-Romance. The aim was to show that, despite their atypical behavior, it is possible to maintain the idea that there is only one lexical item *que* in these languages if we adopt a new notion of complementizer. I proposed a definition that treats the complementizer as an underspecified item that adopts its functional interpretation in the syntactic position in which it is merged. Over the course of the book, I have shown that there is a correlation between the syntactic position, the behavior of the complementizer and the interpretation of the proposition in its scope. This has ultimately demonstrated that this new conception of complementizer is indeed adequate. The analysis I proposed to account for the central empirical phenomena was that there are two merge positions in the left periphery that provide different interface features as values for the underspecified complementizer. The position at the left edge of the periphery provides a subordinate feature that results in a subordinate interpretation of the sentence. The lowest position in the left periphery provides an attributive feature that results in the interpretation that a commitment to the proposition is ascribed to the hearer.

The overarching question that I addressed was how the boundary between syntax and pragmatics is organized in Ibero-Romance. In the chapters of this book I argued in favor of a distribution of labor between these two components of grammar and against versions of neo-performative hypotheses that propose that pragmatics should be treated in syntax. The motivation behind this is my conviction that only aspects that are demonstrably impacted by structural factors should also be modeled syntactically. When no syntactic factors appear to be active, pragmatics must kick in. This view grants greater autonomy to each of the grammatical components and was supported empirically in this book.

Chapter 2 was dedicated to *que*-initial reportative sentences. The contribution of syntax and pragmatics was shown to be clearly divided. The information read off the syntactic structure is that the sentence is to be interpreted as subordinate. What it is subordinate to depends on the context and is therefore something contributed by pragmatics. In the chapter, I drew a parallel between *que*-initial reportatives and embedded sentences and showed that the syntactic behavior of

the complementizer in both cases is virtually the same. The empirical distribution is in favor of an analysis of the complementizer as being merged in the highest position in the left periphery, valued with a subordinate feature. The difference between the complementizer in embedded and unembedded contexts is merely that it is syntactically selected in the first case and remains unselected in the latter. I furthermore proposed that the reportative interpretation results from reconstruction. However, since the requirements for syntactic reconstruction are not met, I argued that the type of reconstruction we are dealing with in these cases must be pragmatic. Finally, on the basis of the empirical data I put forward the generalization that *que*-initial reportatives require a salient verb of saying in the context. Portuguese *que*-initial reportatives have a stronger requirement than Spanish or Catalan: It is not sufficient for a verb of saying to be salient, it must be given.

Chapter 3 dealt with a variety of different constructions. In my analysis, all of them involve a low merged complementizer that is valued with an attributive feature. The contributions of syntax and pragmatics were again shown to be well within their domains: Syntax provides an attributive feature but the consequences for the interpretation of the sentences valued in this way, as well as the interaction with other types of meanings, are dealt with in pragmatics. I developed a unified analysis to account for a range of different constructions. This was motivated by the observation that they share a core interpretation which I ascribe to the presence of the attributive feature. I proposed that the feature is linked to the lowest projection of the left periphery. Evidence for this assumption was taken from the fact that in certain embedded contexts a low merged complementizer gives rise to a similar interpretation of the proposition in its scope. I showed that the empirical distribution of the complementizer in the different constructions can be correctly predicted by assuming that the low merged complementizer moves from head to head through the left periphery and only stops if the specifier of the relevant projection is filled by externally merged material. I furthermore showed that while the discourse contribution of attributive *que*, which ascribes a commitment to the proposition to the hearer, is always present, further interpretive effects that arise can be explained on the basis of the sentence type and the interaction with other expressions involved in the particular construction.

Chapter 4 presented corpus-based empirical investigations that aimed to broaden the understanding of the meaning of some of the constructions involving attributive *que*. The approach was exploratory rather than hypotheses-driven. This means that, while the results can tell us about the adequacy of the theoretical analysis, above all they serve to uncover new patterns and relations that

would otherwise have remained undetected. In the case of the empirical investigations focusing on attributive *que* with epistemic and evidential modifiers, one novel insight is that there is a mismatch between production and comprehension. The corpus study suggested that each construction gives rise to a preferred reading of the modifiers. The results of the experiments indicate that speakers are willing to accommodate different readings of the modifiers even if they do not correspond to those made salient by the relevant constructions. With regard to the experiment on Catalan polar questions, the results provide evidence in favor of the analysis proposed. The participants judged negative answers particularly unexpected in the contexts that originally contained *que*-initial polar questions. Consequently, a positive answer was likely expected in these contexts. The exploratory approach made it possible to identify further interesting relations. Two findings were particularly revealing. First, the results suggest that speakers are generally biased toward a positive answer, irrespective of the presence of *que* and other potential bias markers. Second, a bias in polar questions is expressed simultaneously by multiple means, and goes far beyond the mere presence or absence of *que*.

References

Abbott, Barbara. 2003. Some notes on quotation. *Belgian Journal of Linguistics* 17(1). 13–26.

Abney, Steven Paul. 1987. *The English noun phrase in its sentential aspect*. Cambridge, MA: MIT. (Doctoral dissertation).

Aboh, Enoch Oladé. 2004. *The morphosyntax of complement-head sequences: Clause structure and word order patterns in Kwa*. New York: Oxford University Press.

Aboh, Enoch Oladé. 2010. Information structuring begins with the numeration. *Iberia* 2. 12–42.

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford/New York: Oxford University Press.

Alarcos Llorach, Emilio. 1994. *Gramática de la lengua española*. Madrid: Espasa Calpe.

Alcázar, Asier & Mario Saltarelli. 2014. *The syntax of imperatives*. Cambridge: Cambridge University Press.

Alturo, Núria & Mariana Chodorowska. 2009. La gramaticalització de "sisplau". *Els Marges: Revista de Llengua i Literatura* 88. 15–38.

Ambar, Manuela. 2003. Wh-asymmetries. In Anna Maria Di Sciullo (ed.), *Asymmetry in grammar*, 209–249. Amsterdam: John Benjamins.

AnderBois, Scott, Adrian Brasoveanu & Robert Henderson. 2010. Crossing the appositive/at-issue meaning boundary. In Nan Li & David Lutz (eds.), *Proceedings of SALT 20*, vol. 20, 328–346. Ithaca, NY: CLC Publications.

AnderBois, Scott, Adrian Brasoveanu & Robert Henderson. 2015. At-issue proposals and appositive impositions in discourse. *Journal of Semantics* 32(1). 93–138.

Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.

Arregi, Karlos. 1998. Spanish *el que* relative clauses and the doubly filled COMP filter. Manuscript.

Austin, John Langshaw. 1961. *How to do things with words*. Oxford: Oxford University Press.

Baayen, Harald R. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, Harald R. & Elnaz Shafaei-Bajestan. 2019. *LanguageR: Analyzing Linguistic Data: A Practical Introduction to Statistics*. R package version 1.5.0.

Bartol Hernández, José Antonio. 1988. *Las oraciones causales en la Edad Media*. Madrid: Paraninfo.

Batllori, Montserrat & Maria-Lluïsa Hernanz. 2008. Emphatic polarity from Latin to Romance. Poster presented at the 10th Diachronic Generative Syntax Conference, Cornell University, August 7th-9th 2008.

Batllori, Montserrat & Maria-Lluïsa Hernanz. 2013. Emphatic polarity particles in Spanish and Catalan. *Lingua* 128. 9–30.

Batllori, Montserrat, Carlos Eliseo Sánchez & Avelina Suñer. 2000. La evolución del sistema de nexos adverbiales: un proceso cíclico entre pragmática y sintaxis (el caso de 'que'). In José Jesús Bustos de Tovar (ed.), *Lengua, discurso, texto: I Simposio Internacional de Análisis del Discurso*, 227–240. Madrid: Visor.

Batllori, Montserrat & Avellina Suñer. 2005. Valors adverbials associats a la conjunció *que*: Un estudi històric. *Caplletra* 38. 279–301.

Bayer, Josef. 2002. Minimale Annahmen in Syntax und Morphologie. *Linguistische Arbeitsberichte* 79. 277–297.

Bayer, Josef. 2004. Decomposing the left periphery: Dialectal and cross-linguistic evidence. In Horst Lohnstein & Susanne Trissler (eds.), *The syntax and semantics of the left periphery*, 59–96. Berlin: De Gruyter.

Bayer, Josef, Roland Hinterhölzl & Andreas Trotzke. 2015. Issues in discourse-oriented syntax. In Josef Bayer, Roland Hinterhölzl & Andreas Trotzke (eds.), *Discourse-oriented syntax*, 1–12. Amsterdam: John Benjamins.

Bayer, Josef & Hans-Georg Obenauer. 2011. Discourse particles, clause structure, and question types. *The Linguistic Review* 28(4). 449–491.

Beaver, David. 2001. *Presupposition and assertion in dynamic semantics* (Studies in logic, language, and information). Stanford, CA: CSLI Publications.

Beaver, David & Henk Zeevat. 2007. Accommodation. In Gillian Ramchand & Charles Reiss (eds.), *Oxford handbook of linguistic interfaces*, 502–538. Oxford: Oxford University Press.

Belletti, Adriana. 1990. *Generalized verb movement: Aspects of verb syntax*. Turin: Rosenberg & Sellier.

Belletti, Adriana. 2004. Aspects of the low IP area. In Luigi Rizzi (ed.), *The structure of CP and IP* (The Cartography of Syntactic Structures 2), 16–51. Oxford: Oxford University Press.

Belletti, Adriana. 2009. Answering strategies: New information subjects and the nature of clefts. In Adriana Belletti (ed.), *Structures and strategies*, 242–265. New York: Routledge.

Belletti, Adriana. 2013. On Fin: Italian *che*, Japanese *no*, and the selective properties of the copula in clefts. In Yoichi Miyamoto, Daiko Takahashi, Hideki Maki, Masao Ochi, Koji Sugisaki & Asako Uchibori (eds.), *Deep insights, broad perspectives. Essays in honor of Mamoru Saito*, 41–55. Tokyo: Kaitakusha Co., Ltd.

Benincà, Paola. 1996. La struttura della frase esclamativa alla luce del dialetto padovano. In Paola Benincà (ed.), *Saggi di grammatica per Giulio C. Lepschy*, 23–43. Roma: Bulzoni.

Benincà, Paola. 2001. The position of topic and focus in the left periphery. In Guglielmo Cinque & Giampaolo Salvi (eds.), *Current studies in Italian syntax: Essays offered to Lorenzo Renzi*, 39–64. Amsterdam: North Holland.

Benincà, Paola & Cecilia Poletto. 2004. Topic, focus, and V2. In Luigi Rizzi (ed.), *The structure of CP and IP* (The Cartography of Syntactic Structures 2), 52–75. New York/Oxford: Oxford University Press.

Bianchi, Valentina, Giuliano Bocci & Silvio Cruschina. 2016. Focus fronting, unexpectedness, and evaluative implicatures. *Semantics and Pragmatics* 9(3). 1–54.

Bianchi, Valentina & Mara Frascarelli. 2010. Is topic a root phenomenon? *Iberia* 2(1). 43–88.

Bianchi, Valentina & Mara Frascarelli. 2017. Complementizer deletion. In Martin Everaert & Henk van Riemsdijk (eds.), *The Wiley Blackwell companion to syntax*, 1–26. London: Wiley Online Library.

Biezma, María. 2008. An expressive analysis of exclamatives in Spanish. manuscript.

Bocci, Giuliano. 2009. *On syntax and prosody in Italian*. Siena: University of Siena. (Doctoral dissertation).

Bošković, Željko. 2007. On the locality and motivation of move and agree: An even more minimal theory. *Linguistic Inquiry* 38. 589–644.

Bošković, Željko & Howard Lasnik. 2003. On the distribution of null complementizers. *Linguistic Inquiry* 34. 527–546.

Bosque, Ignacio. 1984. Sobre la sintaxis de las oraciones exclamativas. *Hispanic Linguistics* 1(2). 283–304.

Boye, Kasper. 2012. *Epistemic meaning: A crosslinguistic and functional-cognitive study*. Berlin: De Gruyter.

Breiman, Leo. 2001. Random forests. *Machine Learning* 45. 5–32.

Briz, Antonio & Grupo Val.Es.Co. 2002. *Corpus de conversaciones coloquiales*. Special volume of Oralia. Madrid: Arco Libros.

Brody, Michael. 1990. Some remarks on the focus field in Hungarian. *UCL Working Papers in Linguistics* 2. 201–226.

Brody, Michael. 1995. Focus and checking theory. In István Kenesei (ed.), *Levels and structures: Approaches to Hungarian*, 31–43. Szeged: Jate.

Brovetto, Claudia. 2002. Spanish clauses without complementizer. In Teresa Satterfield, Christina Tortora & Diana Cresti (eds.), *Current issues in Romance languages: Selected proceedings from the 29th Linguistic Symposium on Romance Languages (LSRL-29)*, 33–46. Amsterdam: John Benjamins.

Brucart, José María. 1992. Some asymmetries in the functioning of relative pronouns in Spanish. *Catalan Working Papers in Linguistics* 2. 113–143.

Brucart, José María. 1993. Sobre la estructura de SComp en español. In Amadeu Viana (ed.), *Sintaxi. teoria i perspectives*, 59–102. Lleida: Pagès Editors.

Büring, Daniel. 2003. On D-trees, beans, and B-accents. *Linguistics and Philosophy* 26(5). 511–526.

Büring, Daniel & Christine Gunlogson. 2000. Aren't positive and negative polar questions the same? manuscript.

Büring, Daniel & Manuel Kriz. 2013. It's that, and that's it! Exhaustivity and homogeneity presuppositions in clefts (and definites). *Semantics and Pragmatics* 6. 1–29.

Cappelen, Herman & Ernie Lepore. 1997. Varieties of quotation. *Mind* 106(423). 429–450.

Cappelen, Herman & Ernie Lepore. 2003. Varieties of quotation revisited. *Belgian Journal of Linguistics* 17(1). 51–75.

Cappelen, Herman & Ernie Lepore. 2005. Quotation. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosophy*. Stanford, CA: Stanford University.

Carrera de la Red, Fátima. 1982. *Las expresiones causativas en la obras de Gonzalo de Berceo*. Logroño (La Rioja): Instituto de Estudios Riojanos.

Castroviejo, Elena. 2006. *Wh-exclamatives in Catalan*. Barcelona: Universitat de Barcelona. (Doctoral dissertation).

Castroviejo, Elena. 2018. On question tags and confirmation requests. Talk given at the GLOW 41 Sem Workshop: The grammar and pragmatics of interrogatives and their (special) uses, University of Budapest, April 14, 2018.

Celdrán, Eugenio Martínez, Ana María Fernández Planas, Josefina Carrera Sabaté & Janina Espuny Monserrat. 2005. Approche du mappe prosodique dialectal de la langue catalane. *Géolinguistique* 3. 103–151.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles N. Li (ed.), *Subject and topic*, 27–55. New York: Academic Press.

Chiarcos, Christian, Berry Claus & Michael Grabski (eds.). 2011. *Salience: Multidisciplinary Perspectives on its Function in Discourse*. Berlin: De Gruyter.

Chomsky, Noam. 1955. The logical structure of linguistic theory. Manuscript. Published in 1975 by Plenum Press, New York.
Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Ken Hale & Samuel J. Keyser (eds.), *The view from building 20: essays in linguistics in honor of Sylvain Bromberger*, 1–52. Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels, Juan Uriagereka & Samuel Jay Keyser (eds.), *Step by step: essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In Michael J. Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
Cinque, Guglielmo. 1990. *Types of A-bar dependencies*. Cambridge, MA: MIT Press.
Cinque, Guglielmo. 1994. On the evidence for partial N-movement in the Romance DP. In Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi & Rafaella Zanuttini (eds.), *Paths towards universal grammar: Studies in honor of Richard S. Kayne*, 85–110. Washington, DC: Georgetown University Press.
Cinque, Guglielmo. 1999. *Adverbs and functional heads*. Oxford/New York: Oxford University Press.
Colaço, Madalena & Gabriela Matos. 2016. Explicative clauses in Portuguese as a case of parentheses. In Ernestina Carrilho, Alexandra Fiéis, Maria Lobo & Sandra Pereira (eds.), *Romance languages and linguistic theory 10: Selected papers from Going Romance 28, Lisbon*, 43–60. Amsterdam: John Benjamins.
Coniglio, Marco & Iulia Zegrean. 2010. Splitting up Force, evidence from discourse particles. *Venice Working Papers in Linguistics* 20. 7–34.
Cornillie, Bert. 2009. Evidentiality and epistemic modality: On the close relationship between two different categories. *Functions of Language* 16(1). 44–62.
Corr, Alice. 2016. *Ibero-Romance and the syntax of the utterance*. Cambridge: University of Cambridge. (Doctoral dissertation).
Coulmas, Florian. 1986. *Direct and indirect speech*, vol. 31. Berlin: De Gruyter.
Cruschina, Silvio. 2012. *Discourse-related features and functional projections*. Oxford/New York: Oxford University Press.
Cruschina, Silvio. 2013. Beyond the stem and inflectional morphology: An irregular pattern at the level of periphrasis. In Silvio Cruschina, Martin Maiden & John Charles Smith (eds.), *The boundaries of pure morphology: Diachronic and synchronic perspectives*, 262–283. Oxford: Oxford University Press.
Cruschina, Silvio. 2015. The expression of evidentiality and epistemicity: Cases of grammaticalization in Italian and Sicilian. *Probus* 27(1). 1–31.

Cruschina, Silvio, Ion Giurgea & Eva-Maria Remberger. 2015. Focus fronting between declaratives and exclamatives. *Revue Roumaine de Linguistique* 60(2-3). 257–275.

Cruschina, Silvio & Eva-Maria Remberger. 2017a. Before the complementizer: Adverb types and root clause modification. In Martin Hummel & Salvador Valera (eds.), *Adjective–adverb interfaces in Romance*, 81–109. Amsterdam: John Benjamins.

Cruschina, Silvio & Eva-Maria Remberger. 2017b. Focus fronting. In Andreas Dufter & Elisabeth Stark (eds.), *Manual of Romance morphosyntax and syntax*, 502–535. Berlin: Mouton De Gruyter.

Cruschina, Silvio & Eva-Maria Remberger. 2018. Speaker-oriented syntax and root clause complementizers. *Linguistic Variation* 18(2). 336–358.

Cuenca, Maria Josep. 1997. Form-use mappings for tag questions. In Wolf-Andreas Liebert, Gisela Redeker & Linda R. Waugh (eds.), *Discourse and perspective in cognitive linguistics* (Current Issues in Linguistic Theory 151), 3–19. Amsterdam: John Benjamins.

Culicover, Peter W. 1992. Topicalization, inversion and complementizers in English. In Denis Delfitto, Martin Everaert, Arnold Evers & Frits Stuurman (eds.), *Working papers: Going Romance and beyond*, 1–45. Utrecht: University of Utrecht.

Cunha, Celso & Lindley Cintra. 1984. *Nova gramática do português contemporâneo*. Lisbon: João da Sá Costa.

D'Alessandro, Roberta & Claudio Di Felice. 2015. The diachrony of Abruzzese complementation. *Revue Roumaine de Linguistique* 40(2-3). 129–145.

Davidson, Donald. 1968. On saying that. *Synthese* 19(1-2). 130–146.

de Carvalho, Ana Sofia Abreu. 2013. An overview of vocatives in European Portuguese. *Linguistica Atlantica* 32. 52–58.

Déchaine, Rose-Marie, Clare Cook, Jeffrey Muehlbauer & Ryan Waldie. 2017. (De-) constructing evidentiality. *Lingua* 186–187. 21–54.

de Cuba, Carlos & Jonathan E. MacDonald. 2013. Referentiality in Spanish CPs. In Victoria Camacho Taboada, Ángel Jiménez Fernández, Javier Martín Martín & Mariano Reyes Tejedor (eds.), *Information structure and agreement* (Linguistics Today 197), 117–140. Amsterdam: John Benjamins.

Degen, Judith. 2015. Investigating the distribution of *some* (but not *all*) implicatures using corpora and web-based methods. *Semantics and Pragmatics* 8(11). 1–55.

de Haan, Ferdinand. 1999. Evidentiality and epistemic modality: Setting boundaries. *Southwest Journal of Linguistics* 18(1). 83–101.

Del Gobbo, Francesca, Nicola Munaro & Cecilia Poletto. 2015. On sentential particles: A crosslinguistic study. In Sylvie Hancil, Alexander Haselow & Margje Post (eds.), *Final particles*, 359–386. Berlin: De Gruyter.
Delsing, Lars-Olof. 2010. Exclamatives in Scandinavian. *Studia Linguistica* 64(1). 16–36.
Demonte, Violeta & Olga Fernández Soriano. 2009. Force and finiteness in the Spanish complementizer system. *Probus* 21(1). 23–49.
Demonte, Violeta & Olga Fernández Soriano. 2014. Evidentiality and illocutionary force: Spanish matrix *que* at the syntax-pragmatics interface. In Andreas Dufter & Álvaro O. de Toledo (eds.), *Left sentence peripheries in Spanish: diachronic, variationist, and typological perspectives*, 217–252. Amsterdam: John Benjamins.
Diewald, Gabriele & Elena Smirnova. 2010. *Evidentiality in German: Linguistic realization and regularities in grammaticalization*. Berlin/Boston: De Gruyter.
Engdahl, Elisabet. 2006. Information packaging in questions. *Empirical Issues in Syntax and Semantics* 6. 93–111.
Erteschik, Nomi. 1973. *On the nature of island constraints*. Cambridge, MA: MIT. (Doctoral dissertation).
Escandell-Vidal, Victória. 2011. Verum focus y prosodia: Cuando la duración (sí que) importa. *Oralia* 14. 181–201.
Escandell-Vidal, Victória & Manuel Leonetti. 2009a. Fronting and verum focus in Spanish. In Andreas Dufter & Daniel Jakob (eds.), *Focus and background in Romance languages*, 155–204. Amsterdam: John Benjamins.
Escandell-Vidal, Victória & Manuel Leonetti. 2009b. La expresión de verum focus en español. *Español Actual* 92. 11–46.
Espinal, Maria Teresa. 2013. On the structure of vocatives. In Barbara Sonnenhauser & Patrizia Noel Aziz Hanna (eds.), *Vocative! Addressing between system and performance*, 109–132. Berlin: De Gruyter.
Etxepare, Ricardo. 1997. *The grammatical representation of speech events*. College Park, MD: University of Maryland. (Doctoral dissertation).
Etxepare, Ricardo. 2007. Aspects of the quotative construction in Iberian Spanish. *Anuario del Seminario de Filología Vasca Julio de Urquijo: International Journal of Basque Linguistics and Philology* 41(2). 25–58.
Etxepare, Ricardo. 2008. On quotative constructions in Iberian Spanish. In Ritva Laury (ed.), *Crosslinguistic studies of clause combining: The multifunctionality of conjunctions*, 35–77. Amsterdam: John Benjamins.
Etxepare, Ricardo. 2010. From hearsay evidentiality to samesaying relations. *Lingua* 120(3). 604–627.

Etxepare, Ricardo. 2013. Quotative expansions. In S. Baauw, F. Drijkoningen, L. Meroni & M. Pinto (eds.), *Romance languages and linguistic theory*, 93–123. Amsterdam: John Benjamins.

Evans, Nicholas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*, 366–431. Oxford: Oxford University Press.

Faller, Martina. 2006. Evidentiality above and below speech acts. Manuscript.

Farkas, Donka F. & Kim B. Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27(1). 81–118.

Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. *On information structure, meaning and form: Generalizations across languages*. 87–116.

Freites Barros, Francisco. 2006. El marcador de discurso *claro*: Funcionamiento pragmático, metadiscursivo y organizador de la estructura temática. *VERBA* 33. 261–279.

Galves, Charlotte, Maria Aparecida Torres Moraes & Ilza Ribeiro. 2005. Syntax and morphology in the placement of clitics in European and Brazilian Portuguese. *Journal of Portuguese Linguistics* 4(2). 143–177.

Gärtner, Hans-Martin & Markus Steinbach. 2006. A skeptical note on the syntax of speech acts and point of view. In Patrick Brandt & Eric Fuß (eds.), *Form, structure, grammar [= Festschrift für Günther Grewendorf]*, 213–222. Berlin: Akademie-Verlag.

Garzonio, Jacopo. 2004. Interrogative types and left periphery: Some data from the Fiorentino dialect. *Quaderni di lavoro dell'ASIS* 4. 1–19.

Ginzburg, Jonathan. 1996. Interrogatives: Questions, facts and dialogue. In Shalom Lappin (ed.), *The handbook of contemporary semantic theory*, 385–422. Oxford: Blackwell.

Ginzburg, Jonathan. 2012. *The interactive stance*. New York/Oxford: Oxford University Press.

Giorgi, Alessandra. 2010. *About the speaker*. Oxford: Oxford University Press.

González i Planas, Francesc. 2010. *Cartografia de la recomplementació en les llengües romàniques*. Universitat de Girona. (MA thesis).

González i Planas, Francesc. 2014. On quotative recomplementation: Between pragmatics and morphosyntax. *Lingua* 146. 39–74.

González Rodríguez, Raquel. 2008. *La polaridad positiva en español*. Madrid: Universidad Complutense de Madrid. (Doctoral dissertation).

González Rodríguez, Raquel. 2009. *La expresión de la afirmación y la negación*, vol. 107. Madrid: Arco Libros.

González Rodríguez, Raquel. 2016. Polaridad: Afirmación y negación. In Javier Gutiérrez-Rexach (ed.), *Enciclopedia de lingüística hispánica*. London: Routledge.
Gras, Pedro. 2010. *Gramática de construcciones en interacción: Propuesta de un modelo y aplicación al análisis de estructuras independientes con marcas de subordinación en español*. Barcelona: Universitat de Barcelona. (Doctoral dissertation).
Gras, Pedro. 2016. Revisiting the functional typology of insubordination: Insubordinate *que*-constructions in Spanish. In Nicholas Evans & Honoré Watanabe (eds.), *Insubordination*. John Benjamins.
Grewendorf, Günter. 2002. Left dislocation as movement. *Georgetown University Working Papers in Theoretical Linguistics* 2. 31–81.
Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 2. 279–326.
Grosz, Patrick. 2012. *On the grammar of optative constructions*. Amsterdam: John Benjamins.
Gunlogson, Christine. 2003. *True to form: Rising and falling declaratives as questions in English*. New York: Routledge.
Gunlogson, Christine. 2008. A question of commitment. *Belgian Journal of Linguistics* 22. 101–136.
Gutiérrez-Rexach, Javier. 2001. Spanish exclamatives and the interpretation of the left periphery. In Johan Rooryck, Yves D'Hulst & Jan Schroten (eds.), *Romance languages and linguistic theory 1999: Selected papers from Going Romance Leiden 1999*, 167–195. Amsterdam: John Benjamins.
Gutiérrez-Rexach, Javier & Patricia Andueza. 2011. Degree restrictions in Spanish exclamatives. In Luis A. Ortiz-López (ed.), *Selected proceedings of the 13th Hispanic linguistics symposium*, 286–295. Somerville, MA: Cascadilla Proceedings Project.
Gutiérrez-Rexach, Javier & Patricia Andueza. 2016. The pragmatics of embedded exclamatives. In Alessandro Capone & Jacob L. Mey (eds.), *Interdisciplinary studies in pragmatics, culture and society*, 767–790. Dordecht, Heidelberg, London, New York: Springer.
Gutzmann, Daniel & Erik Stei. 2011. How quotation marks what people do with words. *Journal of Pragmatics* 43(10). 2650–2663.
Haegeman, Liliane. 2004. Topicalization, CLLD and the left periphery. *ZAS Papers in Linguistics* 35(1). 157–192.
Haegeman, Liliane. 2006. Conditionals, factives and the left periphery. *Lingua* 116(10). 1651–1669.

Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and composition of the left periphery* (The Cartography of Syntactic Structures 8). Oxford: Oxford University Press.

Haegeman, Liliane. 2014. West Flemish verb-based discourse markers and the articulation of the speech act layer. *Studia Linguistica* 68(1). 116–139.

Hamblin, Charles L. 1971. Mathematical models of dialogue. *Theoria* 37(2). 130–155.

Hankamer, Jorge & Ivan Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7(3). 391–428.

Hernanz, Maria-Lluïsa. 2007. From polarity to modality: Some (a)symmetries between *bien* en *sí* in Spanish. In Luis Eguren & Olga Fernández-Soriano (eds.), *Coreference, modality, and focus: Studies on the syntax–semantics interface* (Linguistik Aktuell/Linguistics Today 111), 105–150. Amsterdam: John Benjamins.

Hernanz, Maria-Lluïsa & Gemma Rigau. 2006. Variación dialectal y periferia izquierda. In Beatriz Fernández & Itziar Laka (eds.), *Andolin gogoan: Essays in honour of Professor Eguzkitza*, 435–452. Bilbao: Servei editorial de la UPV-EHU.

Hill, Virginia. 2006. Stylistic inversion in Romanian. *Studia Linguistica* 60(2). 156–180.

Hill, Virginia. 2007a. Romanian adverbs and the pragmatic field. *The Linguistic Review* (24). 61–86.

Hill, Virginia. 2007b. Vocatives and the pragmatics–syntax interface. *Lingua* 117(12). 2077–2105.

Hill, Virginia. 2013. Features and strategies: The internal syntax of vocative phrases. In Barbara Sonnenhauser & Patrizia Noel Aziz Hanna (eds.), *Vocative! addressing between system and performance*, 133–156. Berlin: De Gruyter.

Hill, Virginia. 2014. *Vocatives: How syntax meets with pragmatics*. Leiden: Brill.

Höhle, Tilman N. 1992. Über Verum-Fokus im Deutschen. *Linguistische Berichte* 4. 112–141.

Hopper, Paul & Elizabeth C. Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.

Hornstein, Norbert. 1995. *Logical form: From GB to minimalism*. Oxford: Blackwell.

Hothorn, Torsten, Kurt Hornik, Carolin Strobl & Achim Zeileis. 2022. *Party: A Laboratory for Recursive Partytioning*. R package version 1.3-10.

Hothorn, Torsten, Kurt Hornik & Achim Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15(3). 651–674.

Hualde, José Ignacio. 1992. *Catalan*. London/New york: Routledge.

Hualde, José Ignacio, Antxon Olarrea, Anna María Escobar & Catherine E. Travis. 2009. *Introducción a la lingüística hispánica*. Cambridge: Cambridge University Press.
Hummel, Martin. 2000. *Adverbale und adverbialisierte Adjektive im Spanischen*. Tübingen: Narr.
Hummel, Martin. 2014. The adjective-adverb interface in Romance and English. In Petra Sleeman, Freek van de Velde & Harry Perridon (eds.), *Adjectives in Germanic and Romance*, 35–71. Amsterdam: John Benjamins.
Hummel, Martin. 2017. Adjectives with adverbial functions in Romance. In Martin Hummel & Salvador Valera (eds.), *Adjective adverb interfaces in Romance*, 12–46. Amsterdam: John Benjamins.
Jacobs, Joachim. 2001. The dimensions of topic-comment. *Linguistics* 39. 641–681.
Jacobson, Pauline. 2012. The myth of silent linguistic material. Presented at the Second East Asian School on Logic, Language, and Computation, Chongqing, China.
Jacobson, Pauline. 2016. The short answer: Implications for direct compositionality (and vice versa). *Language* 92(2). 331–375.
Katz, Jerrold J. & Jerry A. Fodor. 1963. The structure of a semantic theory. *Language* 39(2). 170.
Kaufmann, Magdalena. 2011. *Interpreting imperatives*, vol. 88 (Studies in Linguistics and Philosophy). Dordrecht, Heidelberg, London, New York: Springer.
Kaufmann, Magdalena & Claudia Poschmann. 2013. Embedded imperatives: Empirical evidence from colloquial German. *Linguistic Society of America* 89(3). 619–637.
Kayne, Richard S. 1981. ECP extensions. *Linguistic Inquiry* 12. 93–133.
Kayne, Richard S. 1984. *Connectedness and binary branching*. Dordrecht: Foris.
Kehayov, Petar & Kasper Boye. 2016. Complementizer semantics: An introduction. In Kasper Boye (ed.), *Complementizer semantics in european languages*, 1–12. Berlin: De Gruyter Mouton.
Kocher, Anna. 2014. *Claro que son adverbios: Análisis de un proceso de gramaticalización*. Universität Wien. (MA thesis).
Kocher, Anna. 2017a. From verum to epistemic modality and evidentiality: On the emergence of the Spanish Adv+C construction. *Journal of Historical Linguistics* 7(1/2). 78–111.
Kocher, Anna. 2017b. *Root clause complementation in Catalan*. Universität Wien. (MA thesis).
Kocher, Anna. 2018a. A sentence mood account for spanish verum. Manuscript.

Kocher, Anna. 2018b. Epistemic and evidential modification in Spanish and Portuguese: A quantitative approach. *Languages in Contrast: International Journal for Contrastive Linguistics* 18(1). 99–121.

Kocher, Anna. 2019. *Unselected root complementizers in Ibero-Romance*. Wien: Universität Wien. (Doctoral dissertation).

König, Ekkehard & Peter Siemund. 2007. Speech act distinctions in grammar. In Timothy Shopen (ed.), *Language typology and syntactic description*, 276–324. Cambridge: Cambridge University Press.

Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3-4). 243–276.

Krifka, Manfred. 2013. Response particles as propositional anaphors. *Proceedings of SALT* 23. 1–18.

Laka, I. 1990. *Negation in syntax*. Cambridge, MA: MIT. (Doctoral dissertation).

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19. 335–391.

Larson, Richard. 1990. Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry* 21. 589–632.

Ledgeway, Adam. 2005. Moving through the left periphery: The dual complementiser system in the dialects of southern Italy. *Transactions of the Philological Society* 103(3). 339–396.

Ledgeway, Adam. 2012. *From Latin to Romance: Morphosyntactic typology and change* (Oxford Studies in Diachronic and Historical Linguistics). Oxford: Oxford University Press.

Levinson, Stephen C. 1983. *Pragmatics* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8. 339–359.

Lobo, Maria. 2003. *Aspectos da sintaxe das orações subordinadas adverbiais do português*. Lisbon: Universidade de Lisboa. (Doctoral dissertation).

Lohiniva, Karoliina. 2017. French *bien que* as a scalar subordinator. Talk given at the Linguistic Symposium on Romance Languages 58, University of Delaware, April 21, 2017.

Lohnstein, Horst. 2016. Verum focus. In Caroline Fréy & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 290–313. Oxford & New York: Oxford University Press.

Longobardi, Giuseppe. 1996. *The syntax of N-raising: A minimalist theory*. Utrecht: Utrecht University.

Lopes, Ana Cristina Macário. 2012. Contributos para uma análise semântico-pragmática das causais de enunciação no português europeu contemporâneo. *ALFA: Revista de Linguística* 56(2). 451–468.
Lupşa, Cornelia Daniela. 2011. Deconstructing the adverb plus complementizer construction in Romanian. *Bucharest Working Papers in Linguistics* (2). 91–99.
Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
Malamud, Sophia A. & Tamina Stephenson. 2015. Three ways to avoid commitments: Declarative force modifiers in the conversational scoreboard. *Journal of Semantics* 3(2). 275–311.
Maldonada, Concepción González. 1999. Discurso directo y discurso indirecto. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, vol. 3, 3549–3596. Madrid: Espasa Calpe.
Martín Zorraquino, María Antonia. 1998. Los marcadores del discurso desde el punto de vista gramatical. In María Antonia Martín Zorraquino (ed.), *Los marcadores del discurso: Teoría y análisis*, 19–54. Madrid: Arco Libros.
Martínez Marín, Juan. 1978. *Sintaxis de La Celestina: Contribución al estudio de la sintaxis histórica española*. Granada: Universidad de Granada. (Doctoral dissertation).
Martins, Ana Maria. 2006. Emphatic affirmation and polarity: Contrasting European Portuguese with Brazilian Portuguese, Spanish, Catalan and Galician. In Jenny Doetjes & Paz González (eds.), *Romance languages and linguistic theory 2004: Selected papers from "Going Romance", Leiden, 9 - 11 December 2004* (Current Issues in Linguistic Theory 278), 197–223. Amsterdam: John Benjamins.
Martins, Ana Maria. 2007. Emphatic verb reduplication in European Portuguese. European Dialect Syntax Project. Papers presented at Workshop on Syntactic Doubling, Amsterdam, March 2006.
Martins, Ana Maria. 2013. Emphatic polarity in European Portuguese and beyond. *Lingua* 128. 95–123.
Mascaró i Pons, Ignasi. 1986. Introducció a l'entonació dialectal catalana. *Randa* 22. 5–38.
Matthewson, Lisa. 2015. Evidential restrictions on epistemic modals. In Luis Alonso-Ovalle & Paula Menéndez-Benito (eds.), *Epistemic indefinites: Exploring modality beyond the verbal domain*, 141–160. Oxford: Oxford University Press.
Milner, Jean-Claude. 1978. *De la syntaxe a l'interprétation*. Paris: Seuil.
Mioto, Carlos. 2003. Focalização e quantificação. *Revista Letras* 61. 169–189.
Mioto, Carlos & Maria Lobo. 2016. Wh-movement: Interrogatives, relatives and clefts. In W. Leo Wetzels, João Costa & Sergio Menuzzi (eds.), *The handbook of Portuguese linguistics*, 275–293. Hoboken (NJ): John Wiley & Sons.

Moreira, Juliana. 2013. *O vocativo e a interface sintaxe-pragmática no português brasileiro*. Belo Horizonte: Universidade Federal de Minas Gerais. (Doctoral dissertation).

Moro, Andrea. 2003. Notes on vocative case: A case study in clause structure. In Josep Quer, Jan Schroten, Mauro Scorretti, Petra Sleeman & Els Verheugd (eds.), *Romance languages and linguistic theory 2001*, 247–262. Amsterdam: John Benjamins.

Munaro, Nicola & Cecilia Poletto. 2009. Sentential particles and clausal typing in Venetian dialects. In Benjamin Shaer, Philippa Cook, Werner Frey & Claudia Maienborn (eds.), *Dislocated elements in discourse: Syntactic, semantic, and pragmatic perspectives*, 173–199. New York/London: Routledge.

Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 2, 52–150. Cambridge: Cambridge University Press.

Nuyts, Jan. 2001. Subjectivity as an evidential dimension in epistemic modal expressions. *Journal of Pragmatics* 33. 383–400.

Ocampo, Francisco. 2006. Movement towards discourse is not grammaticalization: The evolution of *claro* from adjective to discourse particle in spoken Spanish. In Nuria Sagarra & Almeida Jacqueline Toribio (eds.), *Selected proceedings of the 9th Hispanic Linguistics Symposium*, 308–319. Somerville, MA: Cascadilla Proceedings Project.

Ouhalla, Jamal. 1991. *Functional categories and parametric variation*. London: Routledge.

Palmer, Frank R. 2001. *Mood and modality*. 2nd edn. Oxford: Oxford University Press.

Payrató, Lluís. 2002. L'enunciació i la modalitat oracional. In Joan Solà, Maria-Rosa Lloret, Manuel Pérez Saldanya & Joan Mascaró (eds.), *Gramàtica del català contemporani*, 1151–1222. Barcelona: Empúries.

Peres, João Andrade & Salvador Mascarenhas. 2006. Notes on sentential connections (predominantly) in Portuguese. *Journal of Portuguese Linguistics* 5. 113–169.

Pesetsky, David. 1995. *Zero syntax*. Cambridge, MA: MIT Press.

Poletto, Cecilia. 2000. *The higher functional field: Evidence from northern Italian dialects*. New York: Oxford University Press.

Poletto, Cecilia & Raffaella Zanuttini. 2003. Making imperatives: Evidence from Central Rhaetoromance. In Christina Tortora (ed.), *The syntax of Italian dialects*, 175–206. Oxford: Oxford University Press.

Poletto, Cecilia & Raffaella Zanuttini. 2010. Sentential particles and remnant movement. In Paola Benincà & Nicola Munaro (eds.), *Mapping the left periphery*, 201–227. Oxford: Oxford University Press.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.

Pons Bordería, Salvador. 2003. "Que" inicial átono como marca de modalidad. *Estudios de Lingüística* 17. 531–545.

Porroche Ballesteros, Margarita. 2000. Algunos aspectos del uso de *que* en el español conversacional (que como introductor de oraciones "independientes"). *Círculo de Lingüística Aplicada* 3. 100–116.

Portner, Paul. 2004. The semantics of imperatives within a theory of clause types. In Robert B. Young & Yuping Zhou (eds.), *Proceedings of SALT XIV*, 235–252. Ithaca, NY: Cornell University.

Portner, Paul. 2007. Imperatives and modals. *Natural Language Semantics* 15(4). 351–383.

Poschmann, Claudia. 2008. All declarative questions are attributive? *Belgian Journal of Linguistics* 22(1). 247–269.

Prieto, Pilar. 1997. Notes sobre l'entonació dialectal del català: Les oracions interrogatives absolutes. In August Bover (ed.), *Actes del Novè Col·loqui d'Estudis Catalans a Nord-Amèrica*, 347–377. Barcelona: Abadia de Montserrat.

Prieto, Pilar. 2002. Entonació. In Joan Solà, Maria-Rosa Lloret, Manuel Pérez Saldanya & Joan Mascaró (eds.), *Gramàtica del català contemporani*, 393–462. Barcelona: Empúries.

Prieto, Pilar & Gemma Rigau. 2007. The syntax-prosody interface: Catalan interrogative sentences headed by *que*. *Journal of Portuguese Linguistics* 6(2). 29–59.

Prince, Ellen F. 1978. A comparison of wh-clefts and it-clefts in discourse. *Language* 54. 883–906.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org/.

Radford, Andrew. 1982. The syntax of verbal wh-exclamatives in Italian. In Nigel Vincent & Martin Harris (eds.), *Studies in the Romance verb*, 185–205. London: Croom Helm.

Raposo, Eduardo Paiva. 2000. Clitic position and verb movement. In João Costa (ed.), *Portuguese syntax: New comparative studies*, 266–297. Oxford: Oxford University Press.

Reinhart, Tanya. 1981. Pragmatics and linguistic: An analysis of sentence topics. *Philosophica* 27(1). 53–94.

Rigau, Gemma. 1984. De com *si* no és conjunció i d'altres elements interrogatius. *Estudis Gramaticals*. 249–278.

Rigau, Gemma & Pilar Prieto. 2005. A typological approach to Catalan interrogative sentences headed by *que*. Report de Recerca CLT.

Rivero, María Luisa. 1982. Las relativas restrictivas con *que*. *Nueva Revista de Filología Hispánica* 31. 195–234.

Rivero, María Luisa. 1994. Negation, imperatives and Wackernagel effects. *Rivista di Linguistica* 6(1). 39–66.

Rivero, María Luisa & Arhonto Terzi. 1995. Imperatives, V-movement and logical mood. *Journal of Linguistics* 31(2). 301–332.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar: Handbook in generative syntax*, 281–337. Dordrecht: Kluwer Academic Publishers.

Rizzi, Luigi. 2001. On the position "Int(errogative)" in the left periphery of the clause. In Guglielmo Cinque & Giampaolo Salvi (eds.), *Current studies in Italian syntax: Essays offered to Lorenzo Renzi* (North-Holland Linguistic Series: Linguistics Variations 59), 267–296. Amsterdam: Elsevier.

Rizzi, Luigi. 2004a. Locality and left periphery. In Adriana Belletti (ed.), *Structures and beyond* (The Cartography of Syntactic Structures 3), 223–251. Oxford/New York: Oxford University Press.

Rizzi, Luigi. 2004b. On the cartography of syntactic structures. In Luigi Rizzi (ed.), *The structure of CP and IP: the cartography of syntactic structures* (The Cartography of Syntactic Structures 2), 3–15. Oxford: Oxford University Press.

Rizzi, Luigi. 2013. Notes on cartography and further explanation. *Probus* 25(1). 197–226.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics-Ohio State University Department of Linguistics* 49. 91–136.

Roberts, Ian. 2001. Head movement. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 113–147. Oxford: Blackwell.

Roberts, Ian. 2004. The C-system in Brythonic Celtic languages, V2 and the EPP. In Luigi Rizzi (ed.), *The structure of CP and IP* (The Cartography of Syntactic Structures 2), 297–328. Oxford: Oxford University Press.

Rochemont, Michael. 2016. Givenness. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 41–63. Oxford: Oxford University Press.

Rodríguez Molina, Javier. 2014. La gramática oculta de la polaridad afirmativa enfática en español antiguo. *RILCE* 30(3). 861–915.

Rodríguez Ramalle, Teresa María. 2003. *La gramática de los adverbios en -mente o cómo expresar maneras, opiniones y actitudes a través de la lengua*. Ediciones de la Universidad Autónoma de Madrid (ed.). Madrid.
Rodríguez Ramalle, Teresa María. 2007. El complementante *que* como marca enfática en el texto periodístico. *Revista Electrónica de Lingüística Aplicada* 6. 41–53.
Rodríguez Ramalle, Teresa María. 2008. El *que* como marca enfática discursiva en adverbios e interjecciones. In Rafael Monroy & Aquilino Sánchez (eds.), *25 años de lingüística aplicada en España: hitos y retos*, 811–816. Murcia: EDITUM.
Rodríguez Ramalle, Teresa María. 2015. A discourse-based approach to some uses of the conjunction *que* in Romance languages. *Languages in Contrast: International Journal for Contrastive Linguistics* 15(1). 125–150.
Roelofsen, Floris & Donka F. Farkas. 2015. Polarity particle responses as a window onto the interpretation of questions and assertions. *Language* 91(2). 359–414.
Rooryck, Johan. 2001. Evidentiality, part I. *Glot International* 5(4). 125–133.
Ross, John Robert. 1970. On declarative sentences. In Roderick Jacobs & Peter Rosenbaum (eds.), *Readings in English transformational grammar*, 222–272. Waltham, MA: Blaisdell.
Roussou, Anna. 2000. On the left periphery: Modal particles and complementisers. *Journal of Greek Linguistics* 1(1). 65–94.
Roussou, Anna. 2010. Selecting complementizers. *Lingua* 120(3). 582–603.
Saab, Andrés. 2015. On long-distance theta-role assignment. *Lingua* 160. 91–126.
Sadock, Jerrold M. 1969. Hypersentences. *Paper in Linguistics* 1(2). 283–370.
Sadock, Jerrold M. 1974. *Toward a linguistic theory of speech acts*. New York: Academic Press.
Sadock, Jerrold M. & Arnold M. Zwicky. 1985. Speech act distinctions in syntax. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 1, 155–196. Cambridge: Cambridge University Press.
Sæbø, Kjell Johan. 2005. The logical basis of exclamatives. Paper presented at 3 eme Journée de Sémantique et Modélisation, Paris.
Sæbø, Kjell Johan. 2010. On the semantics of "embedded exclamatives". *Studia linguistica* 64(1). 116–140.
Saito, Mamoru. 2012. Sentence types and the Japanese right periphery. In Günther Grewendorf & Thomas Ede Zimmermann (eds.), *Discourse and grammar: from sentence types to lexical categories*, 147–175. Berlin: De Gruyter.
Sansiñena, María Sol, Hendrik De Smet & Bert Cornillie. 2015. Between subordinate and insubordinate: Paths toward complementizer-initial main clauses. *Journal of Pragmatics* 77. 3–19.

Sarkar, Deepayan, Felix Andrews, Kevin Wright, Neil Klepeis, Johan Larsson & Paul Murrell. 2021. *Lattice*. R package version 0.20-45.

Schwabe, Kerstin. 2006. German *dass*-clauses. In Valerie Molnár & Susanne Winkler (eds.), *The architecture of focus*, 429–458. Berlin: De Gruyter.

Schwarzschild, Roger. 1999. Givenness, avoid F and other constraints on the placement of accent. *Natural Language Semantics* 7. 144–177.

Sheehan, Michelle. 2012. A new take on partial control: Defective thematic intervention. *Cambridge Occasional Papers in Linguistics* 6. 1–47.

Shlonsky, Ur. 2000. Remarks on the complementizer layer of Standard Arabic. In Jacqueline Lecarme, Jean Lowenstamm & Ur Shlonsky (eds.), *Research in Afroasiatic grammar*, 325–343. Amsterdam: John Benjamins.

Shlonsky, Ur. 2014. Topicalization and focalization: A preliminary exploration of the Hebrew left periphery. In Anna Cardinaletti, Guglielmo Cinque & Yoshio Endo (eds.), *On peripheries*, 327–341. Tokyo: Hitsuzi Synobo.

Shlonsky, Ur & Gabriela Soare. 2011. Where's why. *Linguistic Inquiry* 42. 651–669.

Sigurdsson, Halldór Armann. 2010. On EPP effects. *Studia Linguistica* 64(2). 159–189.

Simons, Mandy. 2003. Presupposition and accommodation: Understanding the Stalnakerian picture. *Philosophical Studies* 112(3). 251–278.

Speas, Margaret. 2004. Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua* 114(3). 255–276.

Speas, Peggy & Carol Tenny. 2003. Configurational properties of point of view roles. In Anna Maria DiSciullo (ed.), *Asymmetry in grammar: Volume 1: Syntax and semantics*, 315–345. Amsterdam: John Benjamins.

Stalnaker, Robert. 1978. Assertion. In Peter Cole (ed.), *Syntax and semantics 9: pragmatics*, 315–332. New York: Academic Press.

Stavrou, Melita. 2013. About the vocative. In Lilia Schürcks, Anastasia Giannakidou & Urtzi Etxeberria (eds.), *The nominal structure in Slavic and beyond*. Berlin: De Gruyter.

Stowell, Timothy. 1981. *Origins of phrase structure*. Cambridge, MA: MIT. (Doctoral dissertation).

Strobl, Carolin, James Malley & Gerhard Tutz. 2009. An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods* 14(4). 323–348.

Sudo, Yasutada. 2013. Biased polar questions in English and Japanese. In Daniel Gutzmann & Hans-Martin Gärtner (eds.), *Beyond expressives: Explorations in use-conditional meaning*, 275–296. Leiden: Brill.

Suñer, Margarita. 1994. V-movement and the licensing of argumental wh-phrases in Spanish. *Natural Language & Linguistic Theory* 12(2). 335–372.
Tagliamonte, Sali A. & Harald R. Baayen. 2012. Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change* 24(2). 135–178.
Tenny, Carol. 2006. Evidentiality, experiencers and the syntax of sentience in Japanese. *Journal of East Asian Linguistics* 15(3). 245–288.
Torrence, Harold. 2013. *The clause structure of Wolof*. Amsterdam: John Benjamins.
Traugott, Elizabeth C. 2010. (inter)subjectivity and (inter)subjectification: a reassessment. *Subjectification, intersubjectification and grammaticalization*. 29–74.
Uriagereka, Juan. 1988. *On government*. Storrs, CT: University of Connecticut. (Doctoral dissertation).
Uriagereka, Juan. 1995. An F position in Western Romance. In Katalin Kiss (ed.), *Discourse configurational languages*, 153–175. Oxford: Oxford University Press.
van der Auwera, Johan & Vladimir A. Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2. 79–124.
van Gelderen, Elly. 2011. *The linguistic cycle*. Oxford: Oxford University Press.
Villa-García, Julio. 2012a. Recomplementation and locality of movement in Spanish. *Probus* (24). 257–314.
Villa-García, Julio. 2012b. *The Spanish complementizer system: Consequences for the syntax of dislocations and subjects, locality of movement, and clausal structure*. Storrs, CT: University of Connecticut. (Doctoral dissertation).
Villa-García, Julio. 2015. *The syntax of multiple* que *sentences in Spanish* (Issues in Hispanic and Lusophone Linguistics 2). Amsterdam: John Benjamins.
Villa-García, Julio & Raquel González Rodríguez. 2020a. Dialectal variation in the expression of *que* in *sí-que* 'yes that' contexts across Spanish: The case of some Latin American Spanish varieties. *Glossa* 5(1). 99.
Villa-García, Julio & Raquel González Rodríguez. 2020b. On the contrasts between *sí* 'yes' and *sí que* 'yes that' in Spanish and the structure of the complementizer phrase domain. *Studies in Hispanic and Lusophone Linguistics* 13(2). 451–484.
Villalba, Xavier. 2003. An exceptional exclamative sentence type in Romance. *Lingua* 113. 713–745.
Villalba, Xavier. 2008. Exclamatives: A thematic guide with many questions and few answers. *Catalan Journal of Linguistics* 7. 9–40.
von Fintel, Kai. 2008. What is presupposition accommodation, again? *Philosophical Perspectives* 22(1). 137–170.

von Fintel, Kai & Anthony S. Gillies. 2010. Must . . . stay . . . strong! *Natural Language Semantics* 18(4). 351–383.

von Fintel, Kai & Anthony S. Gillies. 2011. 'might' made right. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, 108–130. Oxford: Oxford University Press.

Watters, John R. 1979. Focus in Aghem. In Larry Hyman (ed.), *Aghem grammatical structure.* 137–197. Los Angeles, CA: University of Southern California.

Wheeler, Max, Alan Yates & Nicolau Dols. 1999. *Catalan: A comprehensive grammar*. London/New York: Routledge.

Wiltschko, Martina. 2014. *The universal structure of categories: Towards a formal typology* (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press.

Zanuttini, Raffaella. 1997. *Negation and clause structure*. Oxford: Oxford University Press.

Zanuttini, Raffaella. 2008. Encoding the addressee in the syntax: Evidence from English imperative subjects. *Natural Language & Linguistic Theory* 26. 185–218.

Zanuttini, Raffaella, Miok Pak & Paul Portner. 2012. A syntactic analysis of interpretive restrictions on imperative, promissive, and exhortative subjects. *Natural Language & Linguistic Theory* 30. 1231–1274.

Zanuttini, Raffaella & Paul Portner. 2003. Exclamative clauses: At the syntax-semantics interface. *Language* 79(1). 39–81.

Zribi-Hertz, Anne. 1989. Anaphor binding and narrative point of view: English reflexive pronouns in sentence and discourse. *Language* 65(4). 695–727.

Name index

Subject index

Complementizers on edge

This book offers a comparative perspective on the structural and interpretive properties of root-clause complementizers in Ibero-Romance. The driving question the author seeks to answer is where the boundaries between syntax and pragmatics lie in these languages. Contrary to most previous work on these phenomena, the author argues in favor of a relatively strict distribution of labor between the two components of grammar.

The first part of the book is devoted to root complementizers with a reportative interpretation. The second part deals with root complementizers and commitment attribution. Finally, the last part presents the results of empirical studies on the topic.