



universität
wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Unselected Root Complementizers in Ibero-Romance“

verfasst von / submitted by

Anna Kocher, MA MA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Doktorin der Philosophie (Dr. phil.)

Wien, 2019 / Vienna 2019

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on the student
record sheet:

A 792 236

Dissertationsgebiet lt. Studienblatt /
field of study as it appears on the student record sheet:

Romanistik

Betreut von / Supervisor:

Univ.-Prof. Dr. Eva-Maria Remberger

Mitbetreut von / Co-Supervisor:

Acknowledgments

Although a thesis is ultimately written by one single individual, the individual who wrote this particular thesis feels that it greatly benefited from the support of others to get the project started, keep it going and, most importantly, finish it. The greatest burden of this effort was surely borne by my supervisor Eva-Maria Remberger, who followed all my adventures with encouragement, guidance and interest. I am thankful for the time she took to discuss my work and especially for the extensive feedback during the final stages that has helped me present my ideas in a more concise way. What made her the most ideal supervisor for me, however, was her way of always seeming to know when it was the right time to push me and when to leave me be.

I'm also grateful to the members of my committee: Daniel Buring, who is a spectacular teacher and has been a source of inspiration since the beginning of my studies, and Josep Quer, whom I got to know during my short research visit to Barcelona and whose engaging questions ended up having a substantial impact on this thesis.

I furthermore want to thank Laia Mayol for being so welcoming and taking so much time to discuss this work with me when I was at UPF in Barcelona in 2017. I remember this stay as an extremely motivating and productive time. Moreover, I'd like to thank Laia for always being available to answer my many questions on experimental pragmatics and Catalan via email.

During my experimental research, I asked advice from pretty much every Catalan and Spanish speaker I happen to know. Luckily, some of them are linguists and didn't tire of helping me out. Apart from Laia, I am thankful to Elena Castroviejo, Isabel Crespí and Maria del Mar Vanrell for their patience and interest in my work and for participating in all my pilot studies. In addition, I need to say thank you to Harald Baayen for getting me excited about statistics and for being such a motivating teacher. I am thankful as well to the other great teachers I was fortunate enough to learn from during my studies at the department of linguistics. I feel they laid the foundation for many of the ideas I developed in this thesis. Most of all, however, I admire them for their passion for the things they teach and the respect and sincere interest they have for their students. Finally, I am also thankful to all the great linguists I've had a chance to meet at conferences and other occasions. Their openness, enthusiasm and questions shaped not only my ideas and but also my conception of what it means to be a scholar. In the last few years I have learned that doing a PhD is so much more than writing a PhD thesis. For everything else, from how to write a readable abstract, to remaining calm during question sessions and all the other valuable lessons, I am greatly in debt to my dear friend Silvio Cruschina. He also read through drafts of pretty much everything I produced during my PhD studies and offered the most constructive criticisms.

The past few years have been a roller coaster ride of glorious heights but also of less glorious lows. In the times when I felt frustrated and overwhelmed, I am immensely

glad that I could count on my peers and friends from the Romance and from the Linguistics department who went through similar experiences. I thank them for being such a supportive community and for helping out with time, advice and coffee and for cheering me on.

Similarly, I am grateful for the support, understanding and love of my friends and my family. Above all, I'm glad they put things into perspective. Lastly, I want to thank Thomas: for all of the above and everything in between.

Contents

1	Introduction	1
1.1	Aims, Scope and Motivation	4
1.2	Theoretical Background	7
1.3	The Analysis in a Nutshell	13
1.4	A Sidenote on <i>Que</i> -Initial Directives	15
1.5	Organization of the Thesis	19
2	On <i>Que</i>-Initial Reportatives	21
2.1	Previous Analyses	23
2.1.1	Etxepare (2013) and Demonte & Fernández Soriano (2014) . . .	23
2.1.2	Corr (2016)	25
2.2	Outline of the Present Analysis	29
2.3	Syntactic Properties	37
2.3.1	Location within the Left Periphery	37
2.3.2	Above the Left Periphery	45
2.4	Contrasting <i>Que</i> -Initial Reports in Portuguese	52
2.5	Pragmatic Requirements	60
2.6	Summary	67
3	On Attributive <i>Que</i>	68
3.1	Previous Analyses	69
3.1.1	Corr (2016)	69
3.1.2	Prieto & Rigau (2007)	76
3.1.3	Ambar (2003), Castroviejo (2006) and Demonte & Fernández So- riano (2009)	80
3.1.4	Cruschina & Remberger (2017a, 2018)	82
3.1.5	Hernanz (2007)	84
3.2	Outline of the Present Analysis	87
3.3	Syntactic Properties	99
3.3.1	<i>Que</i> -initial Declaratives	100
3.3.2	<i>Que</i> -initial Polar Questions	104
3.3.3	<i>Que</i> in Wh-Exclamatives	113
3.3.4	<i>Que</i> with Epistemic and Evidential Modifiers	123
3.3.5	<i>Que</i> in Verum Sentences	131
3.4	Pragmatic Requirements	142
3.5	Summary	161

4	Empirical Support	163
4.1	Corpus-Based Experiments	164
4.2	Statistical Modeling	165
4.2.1	Linear and Generalized Additive Models	165
4.2.2	Conditional Inference Trees	171
4.3	Deictic Centers of Epistemic and Evidential Modifiers	174
4.3.1	Prequel: Comparative Corpus Analysis on Epistemic and Evi- dential Modifiers	175
4.3.2	Deictic Centers of Spanish Epistemic and Evidential Modifiers .	185
4.3.3	Deictic Centers of Catalan Epistemic and Evidential Modifiers .	200
4.4	Bias in Polar Questions	216
4.5	General Discussion	233
5	Conclusion	236
6	References	238

1 Introduction

A complementizer can generally be defined as the element that identifies a sentential complement as such and has the function to subordinate 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. La Joana diu que la Maria és lingüista. (Cat.)
the Joana says that the Maria is linguist
- b. A Joana diz que a Maria é linguista. (Pt.)
the Joana says that the Maria is linguist
- c. Juana dice que María es lingüista. (Sp.)
Juana says that Maria is linguist
‘Jo/uana says that Maria is a linguist.’
- (2) a. Que la Maria es lingüista és fantàstic. (Cat.)
that the Maria is linguist is fantastic
- b. Que a Maria é linguista é fantástico. (Pt.)
that the Maria is linguist is fantastic
- c. Que Maria es lingüista es fantástico. (Sp.)
that Maria is linguist is 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.¹ The examples in (4) show that *que* can co-occur with adverbs when introducing sentential adjuncts.²

- (3) a. La Joana parla d’ un llibre que està llegint. (Cat.)
the Joana talks of a book that is reading

¹There are opposing views regarding whether *que* is 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.

²The presence of the complementizer in sentential adjuncts is subject to adverb-, language- and register-dependent variation.

- b. A Joana fala dum livro que está a ler. (Pt.)
the Joana talks of a book that is PREP read
 - c. Juana habla de un libro que esta leyendo. (Sp.)
Juana talks of a book that is reading
'Jo/uana is talking about a book that she is reading.'
- (4)
- a. Has de netejar el teu dormitori abans que arribi jo. (Cat.)
have.2S to clean the your room before that arrive.SUBJ I
 - b. Tens de limpar o teu quarto antes que chegue eu. (Pt.)
have.2S to clean the your room before that arrive.SUBJ I
 - c. Has de limpiar tu quarto antes que llegue yo. (Sp.)
have.2S to clean your room before that arrive.SUBJ I
'You have to clean your room before I arrive.'

The phenomena that are the central concern of the present thesis pose a problem for a traditional notion of complementizer. I use the term *root-clause complementizers* to refer to the cases relevant for this investigation in which *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 appears that it is not required strictly for syntactic reasons, which is further supported by the fact that in these contexts the presence of the complementizer has an impact on the interpretation of the sentence that it introduces.

- (5)
- a. (Que) la Maria ve a la festa. (Cat.)
QUE the Maria comes to the party
 - b. (Que) María viene a la fiesta. (Pt.)
QUE Maria comes to the party
 - c. (Que) la Maria vem a la festa. (Sp.)
QUE the Maria comes 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. Certament (que) és un arros bó. (Cat.)
 certainly QUE is a rice good
- b. Certamente (que) é um arroz bo. (Pt.)
 certainly QUE is a rice good
- c. Ciertamente (que) es un arroz bueno. (Sp.)
 certainly QUE is a rice good
 ‘Certainly, this is a great rice.’
- (7) a. Que bó (que) és aquest arròs! (Cat.)
 how good QUE is this rice
- b. Que bom (que) é este arroz! (Pt.)
 how good QUE is this rice
- c. ¡Qué bueno (que) es este arroz! (Sp.)
 how good QUE is this rice
 ‘How great this rice is!’
- (8) a. Sí (que) és un arròs bó. (Cat.)
 VERUM QUE is a rice good
- b. Sí (que) es un arroz bueno. (Sp.)
 VERUM QUE is a rice good
 ‘This IS a great rice.’

What all of these examples have in common is that a commitment to the proposition introduced by *que* is attributed to the hearer. In the constructions³ 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

³In this thesis, I use the term *construction* in a theory-neutral way to refer to a pairing of a productive abstract structure with a compositional meaning.

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

the conclusion that we are dealing with three different lexical items: one *que* that is a subordinator and two *ques* that are pragmatic markers, one of which gives rise to a reportative interpretation and 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 thesis will presents arguments for precisely the opposite view. In the following chapters, I will bring in 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

Aims The central aim of this thesis is to determine the structure and interpretation of root-clause complementizer constructions and develop an analysis that can account for their empirical distribution. The results of this thesis will hopefully contribute valuable information on the interaction of *que* with other left-peripheral material and will hence lead to a broader understanding of the left periphery in Ibero-Romance languages.

A second aim is to gain insights into the properties of the complementizer itself. As stated above, the empirical focus of this thesis are those examples in which we encounter an item that looks like a complementizer but does not behave accordingly based on a 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 out as *que*, each encoding a specific meaning. The alternative option I propose, is to assume that there is only one *que* and that we have to adjust our notion of complementizer. In other words, *que* is a complementizer but, crucially, a complementizer is something other than what we thought it was. One of the goals of this thesis is therefore to find empirical evidence in favor of this second option and to show that, with a revised notion of complementizer, apparently atypical behavior like that illustrated in (5)-(8) can also be accounted for. One major empirical point that motivates the assumption that there is only one *que* is that there is no formal difference in the item when it expresses these different functions. This means that there is no formal distinction that would suggest that there are multiple types of *que*. Furthermore, although, it does surface at different points in the functional field, *que* is restricted to appearing at the left periphery of a sentence, the natural habitat of a complementizer.

Scope The examples I discuss here are taken from Catalan and the European varieties of Spanish and Portuguese, although many of the generalizations also hold for non-European varieties of these two languages. Furthermore, where relevant, comparisons with similar phenomena in other Romance and non-Romance languages are pointed out. In this thesis, I have drawn on a variety of different 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 here, which are typically employed in informal registers, because the databases contain oral data (Genre/Historical subcorpus) and web data (Web/Dialects subcorpus) in which informal registers usually prevail. They 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) (Ljubešić & Toral 2014). This corpus does not have an online interface but can be freely downloaded. The Catalan databases were completed by 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.

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 the literature. First, a global goal of this thesis 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. Primarily, I hope to contribute new insights that deepen our understanding of the nature of the complementizer and the interaction

at the left periphery. Additionally, the analysis presented in this thesis is based on slightly different theoretical assumptions than the previous analyses. The most drastic difference is that the present 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 performative structure is a part of the architecture of the clause envisioned as a functional domains above the CP (see for instance the contributions made by Benincà 2001, Garzonio 2004, Hill 2006, Hill 2007a,b, Speas & Tenny 2003, Speas 2004, Tenny 2006, Poletto & Zanuttini 2003, Zanuttini 2008, Zanuttini et al. 2012, Krifka 2013, Haegeman 2014, Wiltschko 2014). Critical reviews of the most prominent account developed by Speas & Tenny (2003) are given in Gärtner & Steinbach (2006) and Alcázar & Saltarelli (2014). The main criticism put forward in Gärtner & Steinbach (2006) is that Speas & Tenny (2003) postulate just one structure as universal, while there are other structures derivable by universal syntactic principles that are disregarded by the authors without a convincing motivation. This 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 motivating factors behind my adoption of this position, will emerge over the course of the following chapters, in particular in sections 2.1 and 3.1, where I review the analyses that have been advocated 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 so dramatic as to warrant rejection out-of-hand. The motivation of this thesis is not to argue that the explanation provided here 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 hold back from presenting my analysis as the more economical alternative. This is 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 be determined here. 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 thesis. I outline my re-conception of what an Ibero-Romance complementizer constitutes and give a brief introduction to my main assumptions and my minimal adaptations of the cartographic approach, which is the theoretical framework in which my analysis is modeled.

The Nature of *Que* In traditional conceptions of complementizers, the function as a subordinator is dominant (cf. the definition I cite at the beginning of this chapter). These conceptions, however, appear inadequate for capturing the behavior of Ibero-Romance *que* exemplified in (5)-(8). This makes it necessary to redefine what an Ibero-Romance complementizer is, in order to comply with one of the goals of the thesis, that is to maintain the idea that there is only one lexical element *que* and that this element is in fact a complementizer. My proposal is inspired by what Bayer (2002, 2004) concludes about German *was* and its Bavarian cognate *wos*. This word is notoriously polyfunctional. Some examples that illustrate this fact are given 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) a. Ich hab da was gesehen. (Ger.)
I have there WAS gesehen
'I have seen something.' (Bayer 2002, 288: ex 23a)
- b. Was auf dem Oberdeck saß, war deutsch und trank Sekt. (Ger.)
WAS on the upper-deck sat was German and drank sparkling wine
'The people that sat on the upper deck, were German and drank sparkling wine.' (Bayer 2002, 288: ex 20b)
- c. Was hast du gegessen? (Ger.)
WAS have you eaten
'What did you eat?' (Bayer 2002, 288: ex 19b)

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) die Frau (die) was am Eck Wiaschtl'n vakauft (Austro-Bav.)
the woman REL.PRON.FS WAS at-the corner sausages sells
‘The woman who sells sausages at the corner’ (Bayer 2002, 290: ex 26a)

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 entail a spreading of the two functions across two items and which is grammatical otherwise (cf. (11c)), is not grammatical here (cf. (11b)).

- (11) a. I woafß, wos-a gern trinkt. (Austro-Bav.)
I know WAS-he preferably drinks
‘I know what he likes to drink.’ (Bayer 2004, 4: 9a)
- b. *I woafß, wos dass-a gern trinkt. (Austro-Bav.)
I know WAS that-he preferably drinks
(Bayer 2004, 4: 9b)
- c. I woafß, wos fiar-a-Bier dass-a gern trinkt. (Austro-Bav.)
I know WAS for-a Bier that-he preferably drinks
‘I know which beer he likes to drink.’ (Bayer 2004, 4: 9c)

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, which in the present thesis is taken 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 is acquired. In order to account for the data that are the center of the present thesis, 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 thesis that these features are not stipulated to account for the unembedded constructions exemplified in (5)-(8), but on the contrary, are the same features that *que* acquires in syntactically embedded contexts like (1)-(4). This again strengthens the claim that there is only one *que*.

Finally, although this will not be elaborated on here, the re-conception of what *que* is, might allow us to (possibly diachronically) account for the interrogative pronouns Cat. *què*, Pt. *quê*, Sp. *qué*. One idea would be that these are again expressions of 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. Rizzi (1997) states that 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) a. Credo [_{ForceP} che] [_{TopP} il tuo libro_i], loro lo_i apprezzerebbero
 believe that the your book they CL.MS appreciate.COND
 molto. (It.)
 much
 ‘I believe that your book, they would appreciate it a lot.’
 (Rizzi 1997, 288: ex 10a)
- b. *Credo, il tuo libro_i, che loro lo_i apprezzerebbero molto. (It.)
 believe the your book that they CL.MS appreciate.COND much
 (Rizzi 1997, 288: ex 10b)

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 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 agreement with the assumption that *di* occupies the lowest left-peripheral position. Additionally, FinP has been taken to be the projection that is targeted by the finite verb in Germanic verb-second configurations (Roberts 2004).⁶ An important observation for the present thesis is furthermore that in certain constructions a finite complementizer can also be merged in this position (see for instance Belletti 2009, 2013, Ledgeway 2005).

- (14) a. *Credo di il tuo libro_i apprezzarlo_i molto. (It.)
 believe to the your book appreciate-CL.MS much
 (Rizzi 1997, 288: ex 11a)
- b. Credo [_{TopP} il tuo libro_i] [_{FinP} di] apprezzarlo_i molto. (It.)
 believe the your book to appreciate-CL.MS much
 ‘I believe to appreciate your book a lot.’ (Rizzi 1997, 288: ex 11b)

⁶For a recent account, see Lohnstein (2015) and Kocher (2018b) who propose that the German finite verb in verb second configurations targets the clause typing head MoodP.

IntP is postulated as the location in which the interrogative complementizer is merged. Evidence for this comes also from Spanish embedded polar questions like (15), which permit the co-occurrence of the finite and the interrogative complementizer and in which, crucially, the interrogative complementizer *si* follows *que*. IntP is furthermore the location where expressions like It. *perché* ‘why’ are analyzed (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 2017) (but see section 3.3.2 for my revised take on complementizer-initial polar questions).

- (15) Me preguntaron [_{ForceP} que] [_{IntP} si] tus amigos ya te
 CL.1S asked.3P that whether your friends already CL.2S
 visitaron a Granada. (Sp.)
 visited in Granada
 ‘They asked me (that) whether your friends already visited you in Granada.’
 (Suñer 1994, 349: ex 30b)

FocP is the projection assumed to host fronted foci and wh-pronouns (cf. (16)). Much work has been dedicated to studying the structure, prosody and interpretation of foci within a cartographic framework (see for instance Belletti 2004, Cruschina et al. 2015, Cruschina & Remberger 2017b, Bianchi et al. 2016).

- (16) Credo [_{ForceP} che] [_{TopP} a Gianni_i] [_{FocP} IL MIO LIBRO] Piero gli_i
 believe that to Gianni the my book Piero
 dovrebbe dare. (It.)
 CL.MS should.COND give
 ‘I believe that to Gianni, Piero should give MY BOOK.’
 (adapted from Rizzi 2013, 203 :ex 5)

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, but Frascarelli & Hinterhölzl (2007) propose that the different positions correlate with different interpretations.

- (17) [_{TopP} Il libro_i], [_{TopP} a Gianni_j], glielo_{i,j} darò senz’altro. (It.)
 the book to Gianni CL.DAT-CL.AKK give.FUT for sure
 ‘The book, to John, I’ll give for sure.’ (adapted from Rizzi 1997, 290 :ex 21)

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) [_{FocP} A PARIGI] [_{ModP} probabilmente] Paolo è già stato (non a
to Paris probably Paolo is already been not to
Londra). (It.)
London
‘To PARIS probably Paolo has already been (not to London).’
(Giorgi 2010, 77: ex 38)

Adaptations In the present thesis, 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]]]]]]]]]

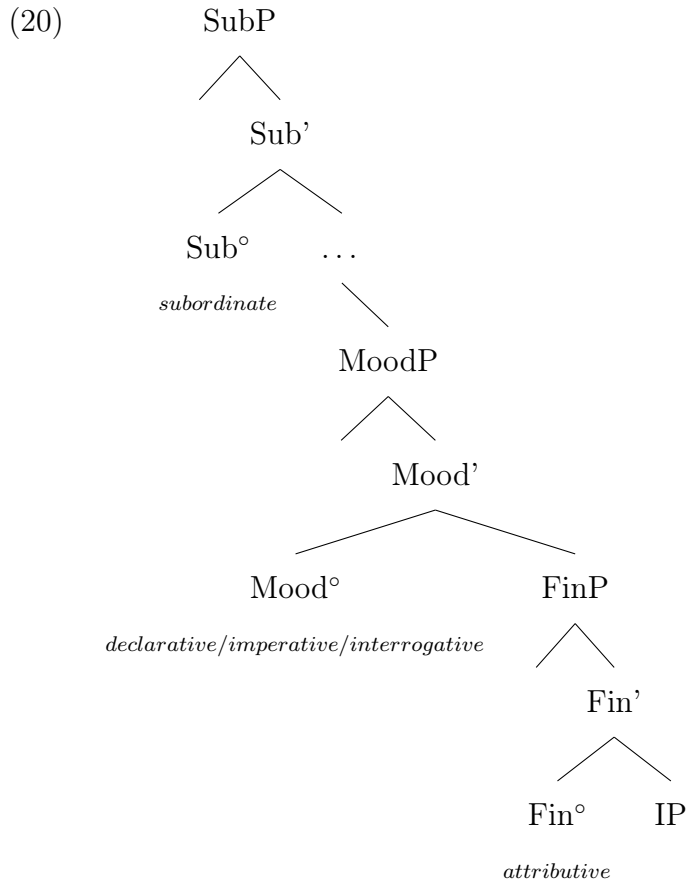
The two adaptations are inspired by Haegeman (2006) (see also Roussou 2010). The first change is the replacement of ForceP by SubP, a functional projection that hosts subordinating conjunctions. According to Haegeman (2006, 1661), Sub is identified as the functional head that subordinates a clause and 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 are separate from each other is evident from data such as (15). In this example, the sentential force is directly encoded through the interrogative complementizer *se*, which is analyzed in IntP. Sentential force is therefore clearly independent of the subordinating function linked to *que*, which is assumed to be merged in the highest projection of the left periphery, SubP.

The second change I adopt is the introduction of a clause typing head, not at the left edge, but in the lower section of the left periphery. In the present thesis, I adopt (sentence) MoodP from Lohnstein (2015) as the functional projection responsible for clause typing. Notably, the position of MoodP is identical to the position in which Haegeman (2006) relocates ForceP. The term Mood is chosen instead of Force in order to prevent a presently unintended conflation of this functional projection with a recent concept 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 that will be motivated empirically 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 section 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 inspired by the fact that there is no formal distinction between the instances of *que* in the different constructions under investigation. 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 but 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. The main aim of the present thesis, as stated above, is to bring in 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 at the center of this thesis, I assume that there are two positions 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) in which the complementizer can be externally merged in. Although complementizers merged in MoodP are not a central concern of this thesis, I briefly return to the core properties of the construction involving a complementizer merged in MoodP in section 1.4. A scheme of the relevant functional projections and their corresponding features is given in (20).



The complementizer in *que*-initial reportative examples like (21a) is analyzed as being merged at the right edge of the left periphery in SubP (cf. (21b)).

- (21) a. Que Juan viene. (Sp.)
 QUE Juan comes
 ‘[reportative] Juan is coming.’
 b. [_{SubP} Que_{subordinate} ... [_{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. (22a) and (22b)).

- (22) a. María ha dicho que Juan viene. (Sp.)
 María has said QUE Juan comes
 ‘María said that Juan is coming.’
 b. [_{SubP} ... [_{IP} María ha dicho [_{SubP} que_{subordinate} ... [_{IP} Juan viene.]]]]

While both of the sentences are introduced by a complementizer carrying a subordinate feature, the difference between the embedded sentence in (22a) and the unembedded sentence in (21a) is that the former is selected by a matrix clause while the latter remains unselected. My claim is that sentences like (21a) receive a reportative interpretation because a verb of saying can be pragmatically reconstructed from the context.

The complementizer in the other root-complementizer constructions examined here is analyzed as being merged in FinP. One example and its corresponding analysis is given in (23a) and (23b).

- (23) a. Ciertamente que Juan viene. (Sp.)
 certainly QUE Juan comes
 ‘Certainly, Juan is coming.’
- b. $[_{\text{SubP}} \dots [_{\text{ModP}} \text{Ciertamente } que_{\text{attributive}} [_{\text{FinP}} \text{t} [_{\text{IP}} \text{Juan viene.}]]]]$

In my account, FinP provides an attributive feature. Once again it has a consequence for the interpretation of the sentence introduced by the complementizer that is valued with this feature. I propose that a commitment to the proposition in the scope of the attributive complementizer is attributed to the hearer. Again, the feature is not stipulated only for the specific phenomenon at hand. On the contrary, it is adopted from de Cuba & MacDonald (2013) who assume a similar feature, *referential* in their terms, 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 here 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 A Sidenote on *Que*-Initial Directives

In this section, I briefly introduce a third type of root-clause complementizer construction attested in all three Ibero-Romance languages. In this construction, *que* appears in directive sentences. The phenomenon differs from the two that are at the center of this thesis in that the complementizer is obligatory and its presence does not give rise

to a special pragmatic interpretation but appears to be required for morphosyntactic reasons.

Que initial directives are jussives, which means that they encode commands directed at a third person who is not currently participating in the conversation.⁷

- (24) a. *Que se vaya.* (Sp.)
 QUE CL.3S go.SUBJ.3S
- b. *Que es vagi.* (Cat.)
 QUE CL.3S go.SUBJ.3S
- c. *Que vá embora.*
 QUE go.SUBJ.3S away
 ‘He/She should go away.’

The comparison between directives addressing a second and a third person is particularly revealing. First of all, the second- and third-person directives differ in their verbal inflection. The verb in the third-person directive in (25a) is inflected for subjunctive while the verb in the second-person directive in (25b) is inflected for imperative mood. Furthermore, the unstressed pronouns are proclitic in (25a) but enclitic in (25b). These structural differences lead to the conclusion that the verb reaches a higher position in imperatives than in third-person directives. In the literature (for instance, Rivero 1994, Rivero & Terzi 1995, Demonte & Fernández Soriano 2009, Alcázar & Saltarelli 2014), the assumption is that the imperative verb targets a left-peripheral position.

⁷A superficially identical construction is employed to express optatives. Optatives 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. (Ia)). 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. (Ib)). Sometimes the second person optatives are employed as polite forms of an imperative (cf. Demonte & Fernández Soriano 2014, who even call them independent imperatives).

- (I) a. *Que tinguis un dia magnífic.* (Cat.)
 QUE have.SUBJ.2S a day magnificent
 ‘May you have a magnificent day!’
- b. *Que tingui un dia magnífic.* (Cat.)
 QUE have.SUBJ.2S a day magnificent
 ‘[honorific:] May you have a magnificent day!’
 ‘May he/she have a magnificent day!’

-
- (25) a. *(Que) es renti les mans. (Cat.)
 QUE CL.3S wash.SUBJ.3S the hands
 ‘He/she should wash her/his hands!’
- b. Renta’t les mans. (Cat.)
 wash.IMP.2S-CL.3S the hands
 ‘Wash your hands!’

The examples in (26) moreover show that third-person directives (cf. (26a)) behave similarly to declaratives (cf. (26b)) in licensing preverbal negation and subjects, both of which are ungrammatical in imperatives (26c). This is further evidence for the idea that the position an imperative verb reaches is higher than the position that a subjunctive or indicative verb reaches.

- (26) a. Que él no lo compre. (Sp.)
 QUE he not it.CL buys.SUBJ.3S
 ‘He/She should not buy it.’
- b. Tu no lo compras. (Sp.)
 you not itCL buysIND.2S
 ‘You don’t buy it.’
- c. (*Tu) (*No) Cómpralo tu. (Sp.)
 you not buy.IMP.2S-it.CL you

In (27) the position of the clitics relative to the imperative and subjunctive verb and the (im)possibility of preverbal subjects in Portuguese confirm that the same pattern holds in this language. Therefore one can conclude that the imperative verb (cf. (27b)) surfaces higher than the subjunctive verb (cf. (27a)) in all three languages.⁸

- (27) a. Que ele o compre. (Pt.)
 QUE it.CL buy.SUBJ.3S
 ‘He/She should buy it.’

⁸As an aside, it must be mentioned that in European Portuguese the clitic placement in declaratives in general requires enclisis (although there are 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 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.

- b. (*Tu) compra-o. (Pt.)
 buy-it.CL
 ‘Buy it!’

For the analysis of third-person directives in Ibero-Romance I adopt an idea outlined in Demonte & Fernández Soriano (2009), which is based on insights from Rivero & Terzi (1995), namely that the complementizer is merged in a left-peripheral position. One difference between Demonte & Fernández Soriano (2009) and the present analysis 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. This is motivated because FinP is not associated with clause typing in the present thesis. There is also an empirical motivation exemplified in (28) 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.

- (28) [_{MoodP} Que [_{TopP} los libros_i [_{FinP} [_{IP} los_i lean los niños y no
 QUE the books CL.3P read.3P.SUBJ the children and not
 las revistas.]]]] (Sp.)
 the magazines
 ‘The children must read the books, not the magazines.’

The basic idea of the analysis assumed here is that MoodP encodes the clause type of each sentence through a feature (see Lohnstein 2015 and Kocher 2018b). For the present purposes, I assume the minimal repertoire of a declarative, an interrogative and an imperative feature.⁹ 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 imperatives directed at a second person, this checking is done by the imperative verb itself. The data discussed in this section have shown that the imperative verb occupies a higher position than either subjunctive verbs in

⁹While 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 three languages or indicative verbs in Catalan and Spanish. The empirical evidence cited in favor of this analysis is the fact that unstressed pronouns are enclitic and the fact that neither subjects nor negative particles can appear before the imperative verb. To explain the merger of *que* in third-person directives, I adopt the argumentation put forward in Rivero & Terzi (1995) and Demonte & Fernández Soriano (2014) who propose a last resort explanation. In imperatives directed to a second person, the verb has the necessary morphological features (=imperative inflection) to check the imperative feature. This triggers the movement of the verb to the projection. In third-person 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°. 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.

To conclude this short sidenote, even though *que*-initial directives are not a central concern in the present context, it is important to stress that the analysis proposed above is very much in accordance with the general assumptions of this thesis. *Que* is again treated as an underspecified element that receives its featural valuation in the position in which it is merged. This again goes to show that, with these simple ideas in place, it is not necessary to assume more than one lexical item *que* to account for the different constructions that the complementizer appears in.

1.5 Organization of the Thesis

This thesis is structured as follows: chapter 2 focuses on *que*-initial reportative sentences such as that 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 section 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 the following section, 2.2, I present my own analysis and compare it to the previous approaches. In section 2.3, I provide the empirical support for my analysis by focusing on the syntactic properties of the construction. In section 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 Spanish and Catalan on the one hand and Portuguese on the other hand. The chapter concludes with section 2.5, which focuses on the pragmatic requirements of *que*-initial reportatives.

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 section 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). The following section, 3.2, details my own proposal for the constructions and compares it to the previous analyses. The syntactic properties that support my analysis are explored in more detail in section 3.3, in which I also discuss some cross-linguistic differences and offers explanations for the contrasts. The last section, 3.4, is dedicated to the pragmatical properties of the constructions.

Chapter 4 presents my empirical studies that further explore the pragmatic contribution of attributive *que* in two constructions. The chapter begins with the methodological background laid out in section 4.1 and the statistical methods described in section 4.2. In section 4.3 I discuss three empirical studies centered on the interpretation of attributive *que* following epistemic and evidential modifiers. In section 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 thesis with a summary of the main points and results.

2 On *Que*-Initial Reportatives

This chapter deals with the phenomenon for which the term insubordination was coined in functional literature (Evans 2007). The term refers to embedded sentences that appear without a matrix clause. The empirical focus of this chapter are *que*-initial sentences that receive a reportative interpretation. In the generative syntactic literature, these sentences have sometimes been called quotatives; however, I use the term reportative instead, because these *que*-initial sentences do not behave like prototypical direct quotations that repeat an expression verbatim (cf. Coulmas 1986, 2).¹⁰ The example in (29) 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.

(29) A: Mentida. (Cat.)
lie

B: Què dius?
what say.2S

A: Perdoni. Que no hi estic d'acord.
excuse.HORT QUE not CL agree

‘A: That’s a lie. B: What did you say? A: Sorry. [reportative:] I don’t agree.’
(ebook-cat)

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 (30a), in which *Juan* is the deictic center, is reported by the original hearer *Maria* in (30b). 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.

¹⁰There 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, Cappelen & Lepore 2005, Maldonado 1999, Abbott 2003, and references therein).

- (30) a. [Juan_j:] T_m 'espero_j a la porta. (Cat.)
 Juan CL.2S wait.1S at the door
 ‘[Juan:] I will wait for you at the door.’
- b. [Maria_m:] Que m_m’ espera_j a la porta. (Cat.)
 Maria QUE CL.1S wait.3S at the door
 ‘[Maria:] [reportative:] He will wait for me at the door.’

This chapter is structured as follows. In section 2.1, I present the previous analyses of this construction put forward by Demonte & Fernández Soriano (2014), Etxepare (2007, 2010, 2013) and Corr (2016). In section 2.2, I introduce my own proposal. The gist of my proposal 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 on the interpretation of the sentences valued with this feature. My principle claim is that the CP of *que*-initial sentences does not differ from its embedded counterparts. In section 2.2, I furthermore 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 section 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 section 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 obligatorily 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 section 2.5, I aim to draw up a unified characterization, by proposing that the reportative interpretation results from pragmatic rather than syntactic reconstruction.

2.1 Previous Analyses

In this section I summarize how *que*-initial reportative sentences have been treated in the generative syntactic literature.¹¹ 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, however, is in how the reportative interpretation arises. 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.¹² The reportative interpretation does not arise through the meaning of an elided verb/noun of saying, but is encoded as a feature in 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 (Etxepare 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. (31)). On his theory, number is restricted to nominals, so he concludes that the silent predicate has to be a noun.

¹¹There 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, Sansiñena et al. 2015).

¹²*Neo-performative hypothesis* is a cover term for different approaches that revive the idea of Ross (1970) and propose that illocutionary force is encoded syntactically. Prominent examples are Speas & Tenny (2003) and Wiltschko (2014).

- (31) Que la lasaña estaba buena y que el vino estaba extraordinario
 QUE the lasagne was good and QUE the wine was great
 resonaron en todo el restaurante
 resounded.PL in all the restaurant
 ‘The saying that the lasagne was good and a saying that the dessert was great
 resounded in the whole restaurant.’ (Etxepare 2010, 620: ex 59)

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.

- (32) [_{LinkerP} [A SAYING] [_{Linker} *que*] [_{ForceP} [...]]]
 (adapted from Etxepare 2013, 98: ex 14, details omitted)

Demonte & Fernández Soriano (2014) revisit Etxepare’s data and claim that there are two distinct instances of *que*. The first instance is not a complementizer but a homophonous reported evidential marker. The second instance is a proper complementizer selected by a silent verb of saying. One of the central distinctions they draw is that the evidential marker does not permit a speech participant to be the source of the reported utterance. For Demonte & Fernández Soriano (2014), (33) is infelicitous because the president cannot report his/her own declaration.

- (33) #Ciudadanos, que se ha / que hemos declarado la guerra. (Sp.)
 citizens that it-has-been that we-have declared the war
 ‘Citizens, someone said that one has/we have declared war.’
 (Demonte & Fernández Soriano, 2014, 154: ex 15)

Corr (2016), however, convincingly shows that this example is actually infelicitous because reported sentences introduced by *que* are inappropriate in the formal register required in a situation where (33) 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 (34).¹³

- (34) Bill, que hemos declarado la guerra. (Sp.)
 Bill that we have declared the war
 ‘Bill, [I said] we’ve declared war.’ (Corr, 2016, 154: ex 16)

¹³See Corr (2016) for further arguments against the assumption of two distinct types of reportative *que*.

In sum, this suggests that the distinction between two different types does not have sufficient empirical support. In line with Corr’s conclusions, there is only one type of *que*-initial reportative constructions. A specialized use as an evidential marker in the sense of Demonte & Fernández Soriano (2014) is not attested. This leaves us with an analysis proposed by Demonte & Fernández Soriano (2014, 39: ex 53) repeated in (35).

(35) (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) are not explicit as to 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 section 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 a number of assumptions that the author makes about the structure of a sentence. She proposes an extension of the highest layers of the clausal structure that is illustrated in (36).

(36) [SA_{high} [SA_{low} [EvalP [EvidP [DeclP ...]]]]]

She assumes a dedicated utterance domain above the CP (similar ideas have been developed in Benincà 2001, Garzonio 2004, Hill 2006, Hill 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 SA_{high} (SpeechAct_{high}) and a low layer termed SA_{low} (SpeechAct_{low}). 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. SA_{low} is furthermore decomposed into projections dedicated to the addressee and the speaker. Corr’s motivation for this structure is the observation that, when CP-external elements

like vocatives and certain discourse markers co-occur, they follow a hierarchical order (see also Moro 2003, Hill 2007b, 2013, 2014, Moreira 2013, Espinal 2013, 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 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 that can be reconstructed from the context (Portuguese) and those without (Spanish, Catalan). Those with a reconstructible matrix clause are not a central concern in Corr (2016). Still, she propose to analyze them as cases of elision. One of Corr’s examples, along with her corresponding analysis, is given in (37a) and (37b).

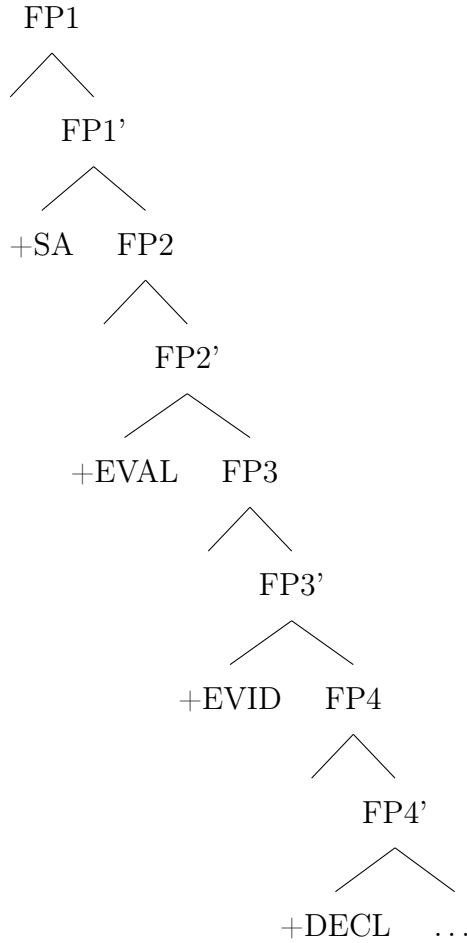
- (37) a. Rotsen, sabes o que me disseram?! Que a época iria começar
 Rotsen know.2S what CL.1S told.3P QUE the season would begin
 a 3 de Dezembro. (Pt.)
 on 3 of December
 ‘Rotsen, do you know what they told me?! That the season was going to
 begin on 3 December.’ (Corr 2016, 149: ex 6)
- b. [CP1 ~~Disseram-me~~ [CP2 que a época iria começar a 3 de Dezembro]]
 (Corr 2016, 149: ex 7)

For Corr (2016), true cases of reportative *que* are only found in Spanish and Catalan and they are namely only the instances that lack a performative verb or clause in the contexts. In these cases, *que* is analyzed as a dedicated evidential complementizer that is merged in the head of her Evidential Phrase. Her analysis of (38a) is given in (38b).

- (38) a. Que quina pallisa que els van clavar. (Cat.)
 QUOT what battering that they go.3PL get.INF
 ‘[I said] what a battering they got.’ (Corr 2016, 161: ex 26)
- b. [_{SA}high [_{SA}low [_{Eval} [_{Evid} QUE_{quot} [_{Decl} [_{Topic} [_{Pol-int} [_{Excl}’ quina pallissa [_{Excl}
 que [_{Wh-int} [_{Focus} [_{Fin} [_{IP} els van clavar]]]]]]]]]]]]]]] (Corr 2016, 188: ex 82)

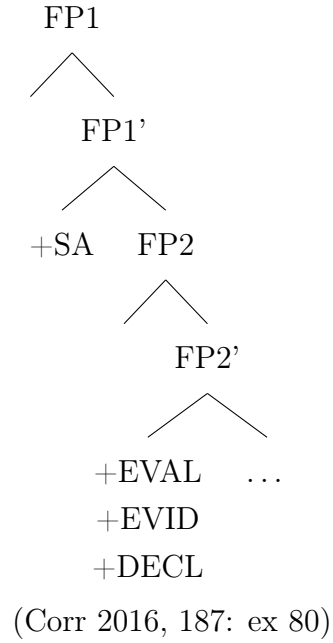
The difference between Spanish and Catalan on the one hand and Portuguese on the other is explained in Corr (2016) as a case of feature scattering versus feature bundling. In Portuguese, there are three features bundled in only one Force head (as in (40)), while in Spanish and Catalan the features are scattered across three heads (EvaluativeP, EvidentialP and DeclarativeP) (as in (39)). Corr (2016) assumes that in true cases of reportative *que* found in Spanish and Catalan the complementizer spells out only the evidential feature. Since the evaluative, evidential and declarative feature are bundled in Portuguese, it is impossible to spell out only one of the features. In Corr's perspective is the reasons 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) study, the analysis might also explain why non-declarative clause types are excluded in Portuguese *que*-initial reportatives (cf. section 2.3.1) because the three features bundled on one head include the declarative feature.

(39) Spanish/Catalan



(Corr 2016, 187: ex 81)

(40) Portuguese



(Corr 2016, 187: ex 80)

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 recoverable *verbum dicendi* in the context have the same syntactic properties and surface at 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 (41), a *verbum dicendi* can be recovered from the context. In the *que*-initial sentence the complementizer precedes a *wh*-phrase just as in (38a), an example with a non-reconstructible matrix clause that according to Corr's analysis should be distinct.

-
- (41) A: Què fa la Tecla? B: Què m' has preguntat? A: Que què fa
 what does the Tecla what CL.1S have asked QUE what does
 la Tecla? (Cat.)
 the Tecla
 'A: What does Tecla do? B: What did you say? A: [reportative:] What does
 Tecla do?' (ebook-cat)

Another possible 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 author considers that *que*-initial reportatives receive a reportative interpretation, how this interpretation arises in her approach is not immediately evident to me. In particular, I am not certain 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 homophonous with but distinct from the default complementizer. In the second option, the reportative interpretation might be encoded syntactically. 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 presentative (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, it is not spelled out which mechanisms ensure that these presentatives with an evidential feature are interpreted as reported sentences.

2.2 Outline of the Present Analysis

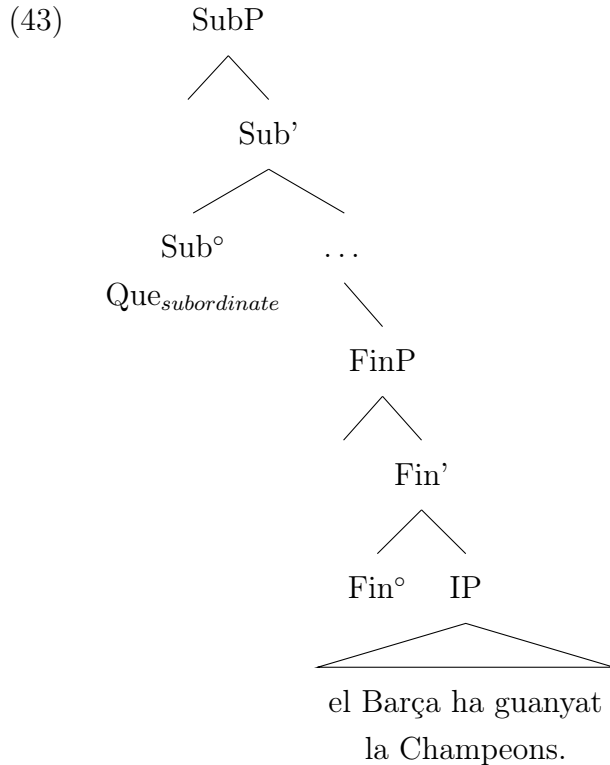
In this section, I present the outline of my own account. The syntactic analysis that I adopt for *que*-initial reportatives builds on the observation that in these apparently unembedded sentences *que* surfaces at the same location as it does in their embedded counterparts. As a consequence, I propose that *que*-initial sentences and their counterparts should be analyzed in the same way. This means that 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 section 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. Moreover, 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. In particular, its use is only felicitous if there is a salient linguistic expression that the sentence can be subordinate to. The pragmatic conditions are described in greater details in section 2.5.

The basic syntactic idea is illustrated in the structure in (43), which corresponds to the analysis I propose for (42). The complementizer is merged directly in the highest head of the split CP SubP where its underspecified feature is valued as *subordinate*.

- (42) Que el Barça ha guanyat la Champeons. (Cat.)
 QUE the Barcelona has won the Championsleague
 ‘[reportative:] Barcelona has won the Champions’ League.’
 (adapted from Demonte & Fernández Soriano 2014, 34)



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 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: As stated above, 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, 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 that I just outlined 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. section 2.3). Consequently, it appears 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 (46). This counts as evidence for the assumption that a complementizer has primarily semantic functions because if its presence were required for purely syntactic reasons it should not be possible to omit it.

For analysis presented in what follows a distinction between indirect speech and direct quotation is essential. Cappelen & Lepore (1997) illustrate the difference with the examples in (44). If Alice uttered (44a), then (44b) is a direct quotation of (44a). It repeats the exact words she uttered. (44c) 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 (44b) is only true if Alice uttered the words that are presented as a quotation whereas (44c) is also true if she didn't.

- (44) 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 (44b) that indicate the beginning and ending of a quote. Indirect speech is not marked orthographically. Furthermore, there are reasons to believe that indirect speech, other than direct quotation, is syntactically embedded. Typically, only indirect speech is introduced by a complementizer (cf. (44c)).

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.¹⁴ This leads to a superficial oral (yet normally not written) ambiguity between a direct quotation and an indirect speech reading exemplified in (45).

- (45) Alice said (that) life is hard.

Similar examples are illustrated for English in (46a) and German in (46b). 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 (46a) and (46b), who said that someone has to leave, while in indirect quotes it is the external speaker, *Mary*, who reports that John said that someone has to leave. The first part of the examples in (46) can in principle be interpreted as either a direct quote 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.¹⁵

¹⁴On complementizer deletion see for instance Erteschik (1973), Kayne (1981, 1984) Stowell (1981), Pesetsky (1995), Bošković & Lasnik (2003), Bianchi & Frascarelli (2017).

¹⁵There is an alternative interpretation in which having Mary as an unwelcome guest leads to John’s

-
- (46) a. [Mary_m:] John_j said I_m have to leave. Apparently I_m 'm not welcome at his_j house.
- b. [Maria_m:] Hans_h hat gesagt, ich_m muss gehen. Scheinbar bin ich_m in Maria Hans has said I must go apparently am I in seinem_h Haus nicht willkommen. (Ger.)
his house not welcome

The parts of the first sentences in (46a), (46b) 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 (47) 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.¹⁶

- (47) a. [Mary_m:] John_j said it again: I_j have to leave. #Apparently I_m am not welcome at his_j house.
- b. [Maria_m:] Hans_h hat es wieder gesagt: Ich_h muss gehen. #Scheinbar bin Maria Hans has it again said I must go apparently am ich_m in seinem_h Haus nicht willkommen. (Ger.)
I in his house not welcome

The Ibero-Romance languages under investigation exhibit a structure that is superficially equivalent to (46a) and (46b) and that also lacks a complementizer. This is exemplified for Spanish in (48a). 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 (46), is not possible. It must be interpreted as a direct quote: 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 inter-

decision to leave his own house. In this scenario, a direct quote reading is possible and the continuations in (47a), (47b) and (48a) are felicitous.

¹⁶For German, an indirect speech reading of (47b) 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.

- (II) [Maria_m:] Hans_h hat die Drohung wiederholt dass ich_m/*_h gehen muss. (Ger.)
Maria Hans has the threat repeated that I have to leave
‘[Maria:] Hans repeated his threat by saying that I have to leave.’

pretation, an overt complementizer is necessary (as in (48b)). Consequently, it appears that in Ibero-Romance complementizer deletion is not an option in these contexts.

- (48) a. [María_m:] Juan_j ha dicho tengo_{j/*m} que irme_{j/*m}. #Aparentemente no
 María Juan has said have.1SG that go-me apparently not
 soy_m bienvenida en su_j casa. (Sp.)
 am welcome.FS in his 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 irme_{*j/m}. Aparentemente
 María Juan has that said have.1SG that go-me apparently
 no soy_m bienvenida en su_j casa. (Sp.)
 not am welcome.FS 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.’

These facts demonstrate 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 (46) 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 (49).

- (49) a. Juan: Tengo que irme. (Sp.)
 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 (49a), the sentence headed by *que* is understood as a report without the need for additional lexical material. In the equivalent version in English in (49b), *That I have to leave* on its own is not sufficient.¹⁷

¹⁷A *that*-initial reported sentence is acceptable in English when a verb of saying is given in the

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 the one illustrated in (42).

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 section 2.1.2) with my own analysis developed here. A summary of the main points is given in table 1.

	nature of <i>que</i>	location	reportative interpretation
Etxepare (2013)	complementizer	LinkerP	silent noun of saying
D&F (2014)	complementizer	ForceP	silent verb of saying
Corr (2016)	evidential complementizer	EvidP	feature in EvidP
present analysis	complementizer	ForceP	pragmatic reconstruction

Table 1: Subordinate *que* in different analyses.

(D&F (2014) = Demonte & Fernández Soriano (2014))

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

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 is not associated with EvidP but with DeclP. The motivation for the splitting of ForceP into subheads is conceptual rather than empirical, and is driven by goal to develop a system in which there is a one-to-one correspondence between a syntactic projection and an

context:

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

abstract pragmatic feature. This, however, makes it difficult to empirically test which of the two adjacent projections *que* 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 the function 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.

Finally, regarding 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) adopts an analysis 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, is based on a reconceptualization of subordination which makes it possible to propose a reduced syntactic analysis without alluding 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 a pragmatic rather than a syntactic type of reconstruction. Unlike the previous accounts, the burden of deriving the interpretation is pushed toward pragmatics rather than syntax.

The analysis proposed in this thesis 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

adopts 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 constitutes a theoretical advantage of my analysis: it works without assuming hidden syntactic layers. Analysis of similar phenomena often rely on a neo-performative approach, for instance the widely adopted structure developed in Speas & Tenny (2003) (cf. for instance Corr’s adaptation). The existence of these types of structures, however, is disputed in the literature (see for instance Gärtner & Steinbach 2006, Alcázar & Saltarelli 2014). The syntactic simplicity of my analysis is achieved by granting a more dominant role to pragmatic mechanisms that are described in section 2.5.

2.3 Syntactic Properties

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

2.3.1 Location within the Left Periphery

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

$$(50) \quad [_{\text{SubP}} [_{\text{TopP}} [_{\text{IntP}} [_{\text{TopP}} [_{\text{ForceP}} [_{\text{TopP}} [_{\text{ModP}} [_{\text{TopP}} [_{\text{MoodP}} [_{\text{TopP}} [_{\text{FinP}}]]]]]]]]]]$$

In (50), I repeat the cartographic structure of the split CP that is adopted in the present thesis and that builds on Rizzi (1997, 2001, 2004a, 2013), Haegeman (2004) and Lohnstein (2015) (see section 1.2). Since, in the proposed analysis, *que* occupies the highest head SubP, it is predicted that it should precede elements that occupy any of the other left-peripheral positions. This expectation is confirmed by the data in (51)-(56).

In (51a) 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.¹⁸ Crucially, all the positions that clitic left dislocated topics can occupy are below SubP. The reverse word order illustrated in (51b), where a clitic left dislocated topic precedes the complementizer, is not grammatical.

- (51) a. [_{SubP} Que] [_{TopP} la tinta_i] la_i hemos de hacer por semanas. (Sp.)
 QUE the ink CL.3S have to make weekly
 ‘(Somebody said) that we have to make the ink weekly.’ (CdE)
 b. *La tinta_i que la_i hemos de hacer por semanas. (Sp.)

As a side note, (51b) 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 section 2.3.2.

In (52a), *que* precedes a polar question introduced by the interrogative complementizer *si*. The relative position and behavior of *si* is 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 supports the proposed analysis. Once again, the reverse word order in which *que* follows *si* is ungrammatical (cf. (52b)).

- (52) a. [_{SubP} Que] [_{IntP} si] eres feliz. (Sp.)
 QUE if are happy
 ‘[reportative:] Are you happy.’
 b. *Si que eres feliz. (Sp.)

In (53a), the complementizer precedes a wh-pronoun. In (54a), it precedes the wh-expression of a wh-exclamative.¹⁹ In accordance with their interpretation and the fact that they cannot co-occur with left dislocated foci, these elements have been analyzed in FocP, which is another position assumed to be located below SubP. It is again ungrammatical to reverse the order of the wh-pronoun and the complementizer (see (53b)). The example in (54b) with the wh-expression followed by *que* in a wh-exclamative is

¹⁸See Frascarelli & Hinterhölzl (2007) who propose that different topic projections result in different topic interpretations.

¹⁹On embedded exclamatives see Zanuttini & Portner (2003), Sæbø (2010) and Gutiérrez-Rexach & Andueza (2016).

- (55) a. ¿Hay mucha corrupción en el gobierno nacional y popular? Los
 is there a lot corruption in the government national and popular the
 kirchneristas dirán que no. [_{SubP} Que] [_{ModP} seguramente] es una
 kirchneristas wloud say that no QUE surely is one
 sola. (Sp.)
 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. (Sp.)
- (56) a. Por ustedes dijo también para luego agregar: [_{SubP} Que] no
 for you told.3S also to afterwards add QUE not
 se enreden: [_{SubP} Que] [_{TopP} Martha] [_{MoodP} sí] va. [_{SubP} Que]
 CL.REFL tangle QUE Martha SÍ goes QUE
 [_{MoodP} sí] va porque ha demostrado que sí puede, con el apoyo
 SÍ goes because has demonstrated that sí can with the help
 de todos. (Sp.)
 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. (Sp.)

Multiple *que* constructions, in which a *que*-initial reported sentences can contain another instance of *que*, are also possible (cf. (57)). The second complementizer is valued with an attributive feature and observes strict adjacency to a wh-expression in the wh-exclamative in (57a), an evidential modifier in AdvC in (57b) and the verum marker *sí* in AffC in (57c). The structure and interpretation of sentences with a low-merged complementizer are dealt with in depth in chapter 3. Regardless of the specific details, what the data in (57) 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 a commitment to the proposition in the scope of the complementizer to the hearer. In the present analysis, this is modeled by assuming that the different projections provide different interface features. Briefly, the idea is that 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.

- (57) a. $[_{\text{SubP}} \text{Que}] [_{\text{FocP}} \text{quina pallissa } \text{que}_i] [_{\text{FinP}} t_i] \text{els van clavar. (Cat.)}$
 QUE what battering QUE they PST get
 ‘(I said) what a battering they got.’ (Corr 2016, 161: ex 26)
- b. $[_{\text{SubP}} \text{Que}] [_{\text{ModP}} \text{claro } \text{que}_i] [_{\text{FinP}} t_i] \text{le viene bien, que qué alegría,}$
 QUE clear QUE CL.3S comes good QUE what joy
 que dónde. (Sp.)
 QUE where
 ‘[reportative:] Of course it was no inconvenience. (She said) what a joy!
 (She said) where?’ (CdE)
- c. $[_{\text{SubP}} \text{Que}] [_{\text{MoodP}} \text{sí } \text{que}_i] [_{\text{FinP}} t_i] \text{heu vingut! (Cat.)}$
 QUE SÍ QUE have.2PL came
 ‘[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. There are two versions of embedded questions in Spanish and Catalan: one in which the complementizer is present and one in which it is absent (see the discussion of the examples (87) in section 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. (58)-(61)). My proposal is that the syntactic structure of the embedded sentence CP in (58) is equivalent to structure of the unembedded sentence CP in (53a). This is supported by the fact that in both cases, *que* precedes the interrogative complementizer *si*.

- (58) Me acerqué a una oficina de turismo, pregunté $[_{\text{SubP}} \text{que}] [_{\text{IntP}} \text{si}]$
 CL.1S approached to a tourist office asked that whether
 estaba abierto el camino hacia Florencia. (Sp.)
 was open the road to Florence
 ‘I went to a tourist office and asked (that) whether the road to Florence was
 open.’ (CdE)

Example (59) parallels (53a). A *wh*-pronoun, analyzed in *FocP*, is preceded by the complementizer *que*.

- (59) Me encontré con un amigo y el amigo me dijo [_{SubP} que] [_{FocP} qué]
 CL.1S met with a friend and the friend CL.1S said that what
 estaba haciendo. (Sp.)
 was.1S doing

‘I ran into a friend and the friend asked (that) what I was doing.’ (CdE)

The examples in (60) and (61) show that syntactically embedded reportatives also permit the occurrence of multiple instances of *que*, just as their unembedded counterparts in (57). The example in (60) demonstrates that the AdvC construction, like in the unembedded sentence in (57b), can be part of an embedded reportative and occupies a position lower than SubP. Example (61) shows that AffC also surfaces below the subordinating complementizer, similar to (57c). The example in (61) 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 predicted by the analysis proposed here. The high *que* is analyzed in SubP and 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 (50)).

- (60) A nivell de relació de parella us haig de dir [_{SubP} que] [_{ModP}
 at level of relationship of couple you have to say that
 evidentment que_i] [_{FinP} t_i] les coses canvien. (Cat.)
 evidentially QUE the things change
 ‘Concerning couples’ relationships I have to tell you that evidently things change.’
 (Kocher 2017, 35: ex 57a from ebook-cat)

- (61) Al museu, mirant un dibuix de Manolo Hugué ha conegut una dona
 at the museum looking a drawing of Manolo Hugué has met a woman
 d’ulls nets, profunds i amb una punta de malícia, i ha pensat [_{SubP}
 of eyes clean deep and with a hint of malice and has thought
 que] [_{ModP} potser] [_{TopP} d’aquesta dona_j] [_{MoodP} sí que_i] [_{FinP} t_i]
 that maybe of this woman SÍ QUE
 se’n_j podria enamorar. (Cat.)
 CL.REFL CL.PART could 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.’ (Kocher 2017, 35: ex 57c from ebook-cat)

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 (52a) and (53a). Another example is shown in (62).

- (62) Pues como te iba diciendo, ella me preguntó por ti. [SubP Que] [FocP
 well as CL.2S was telling she CL.1S asked about you QUE
 qué] tal te va, [SubP que] [IntP si] tienes trabajo y así. (Sp.)
 how are you doing QUE whether have 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 either case, the complementizer obtains in a higher position than the *wh*-pronoun, analyzed in FocP, or the interrogative complementizer *si*, analyzed 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 (63) show that in order to report an imperative, its structure and morphological makeup need to be adapted. Thus (63b), in which the imperative is directly headed by the subordination complementizer, is ungrammatical. In the grammatical version in (63a), the verbal morphology is no longer imperative but subjunctive. The verb furthermore occupies a higher position in the imperative in (63a) than in the reported version in (63c). This is shown by the position of the clitic pronoun, which is enclitic in the true imperative and proclitic in reported imperatives.

- (63) a. ¡Vete! (Sp.)
 go.IMP-CL.2S
 ‘Go away!’
 b. *¡Que vete! (Sp.)
 QUE go.IMP-CL.2S
 intended ‘[reportative:] Go away!’
 c. ¡Que te vayas! (Sp.)
 QUE CL.2S go.SUBJ
 ‘[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 (64b), in which the verb inflected for

imperative mood 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 (64b) than the imperative verb in (63a) (cf. section 1.4 on the structure of *que*-initial directive sentences and imperatives).

- (64) a. Te ordeno que te vayas. (Sp.)
 CL.2S order.1S that CL.2S go.SUBJ
 ‘I order that you leave.’
 b. *Te ordeno que vete. (Sp.)
 CL.2S order.1S that go.IMP-CL.2S
 intended: ‘I order that you leave.’

The data in (63) and (64) are strong evidence in favor of the subordinate nature of *que*-initial reportatives and against the assumption that *que* is a pragmatic marker. One final 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 (65) repeated from (57b). 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 sentences is omitted.

- (65) [SubP Que] [ModP claro que_i] [FinP t_i] le viene bien, [SubP que] [FocP qué
 QUE clear QUE CL.3S comes good QUE what
 alegría], [SubP que] [FocP dónde.] (Sp.)
 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 (66a) the reportative only consists of a wh-pronoun, and in (66b) 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.

- (66) a. Todos volvieron a preguntar que qué. (Sp.)
 all asked.3P again that what
 ‘Everybody asked “what” again.’ (CdE)

-
- b. ¿Cuándo no hay condenación? ¡La Palabra de Dios dice que
 when not there-is damnation the word of god says that
 AHORA! (Sp.)
 now
 ‘When is the moment of no damnation? The word of god says that the
 moment is now.’ (CdE)

Summing up, in the this section I have shown that *que* precedes all other types of phrases that are assumed to be merged in the split CP. Therefore, the complementizer must be analyzed 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 evidence from the literature that shows that this prediction holds true.

2.3.2 Above the Left Periphery

In this section, I discuss 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, 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 (67)).²⁰

- (67) [_{SAhigh} *discourse markers* [_{SAlow} *vocatives* [_{EvalP} [_{EvidP} [_{SpecEvid} *source*] [_{EvidP} [_{Evid⁰} *que_{reportative}*]] [_{DeclP} ...]]]]] (Corr 2016)

Corr (2016) proposes two speech act phrases above the split CP. The higher of these hosts outward-oriented and discourse-activating markers like *oye* ‘listen’. The lower one hosts inward-oriented and discourse-bonding markers like vocatives. As predicted, the complementizer follows both types of expressions in *que*-initial reportatives.

²⁰For a detailed discussion of Corr’s analysis see section 2.1.2.

The data in (68) 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 (68a) shows that the discourse marker *oye*, analyzed in SA_{high} in Corr (2016), can only precede but not follow *que* in reportatives. The same goes for the vocative *Irene* in (68b), analyzed in SA_{low} in Corr (2016). Finally, the example in (68c) shows that a discourse marker and a vocative can co-occur and that, in this case as predicted, they both precede *que*. 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 “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 (68c).

- (68) a. [SA_{high} Oye], que (*oye) si has visto mis llaves. (Sp.)
 DM QUE DM IF have.2S seen my keys
 ‘Hey, (I asked) have you seen my keys?’ (Corr 2016, 182: ex 66)
- b. [SA_{low} Irene], que (*Irene) si has visto mis llaves. (Sp.)
 VOC QUE VOC if have.2S seen my keys
 ‘Irene, (I asked) have you seen my keys?’ (Corr, 2016, 182: ex 67)
- c. [SA_{high} Oiga], [SA_{low} señora Marquesa], (*que) la peinadora que no
 DM VOC QUE the hairdresser QUE not
 puede esperar. (Sp.)
 can wait
 ‘Listen, Lady Marquis, the hairdresser says she cannot wait.’
 (Corr 2016, 182-183: ex 70)

In the example in (68c), the referent of *la peinadora* is the original speaker and at the same time interpreted as the subject of the sentence. Still, in the version illustrated in (68c), *la peinadora* can not 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 (51a)). The only possible analysis for this example is that there is a null subject agreeing with the finite verb *esperar* ‘wait’. However, as discussed in section 2.3.1, *que*-initial reportatives, can be preceded by hanging topics. 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 periphery²¹ while hanging topics are

²¹But cf. Villa-García (2015) for a special type of topic construction where the topic is analyzed as

analyzed as being merged directly in the CP-external position (cf. Villa-García 2015, 165 for a concise comparison).

- (69) La peinadora_i, que esta impaciente_i no puede esperar.
 the hairdresser QUE this impatient not can wait
 ‘[reportative:] The hairdresser, this impatient person cannot wait.’

The example in (69) illustrates the typical properties of hanging topics. *La peinadora* is followed by an intonational break indicated by the comma and it is resumed by the co-referential epithet *esta impaciente*.²²

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 (69). They don’t qualify as hanging topics nevertheless. Importantly, they cannot be resumed because they are not co-referential to an argument or an adjunct of the verb. This can be seen clearly in the example in (70).

- (70) a. La mare, que què t’ has cregut. (Cat.)
 the mother QUE what CL.2S have thought
 ‘[reportative:] The mother: what were you thinking.’ (ebook-cat)
 b. E eu que não, que não. (Pt.)
 and I QUE no NO
 ‘[reportative] and I: no, no.’ (CdP)

In previous accounts, the phrases that precede *que* have been brought in to motivate 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. While it is true that some of the preceding phrases can indeed be interpreted as expressing an original speaker or source, even then the phrases 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 phrase could be analyzed as 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).

being merged directly in a CP-internal position.

²²Some of the native speakers I consulted additionally accepted a version of (69) where *la peinadora* is preceded and followed by *que*. Although further investigation is necessary, this could potentially be a different phenomenon similar to recomplementation analyzed by Villa-García (2015).

The *que*-preceding phrases can be DPs like in (70), which makes the conclusion that they express an original speaker appear plausible. However, *que* can also be preceded by PPs. This is illustrated by the data taken from Etxepare (2013).

- (71) Oye, en Al-Jazeera, que Obama va a atacar Iran. (Sp.)
 Hey in Al-Jazeera QUE Obama goes to attack Iran
 ‘Hey, there’s a saying on Al Jazeera that Obama is going to attack Iran.’
 (Etxepare, 2013, 98: ex 11)

In the example in (71), the phrase that precedes *que* is a locative PP. In Etxepare (2013) *en Al-Jazeera* is analyzed as a adverbial modifier of the elided predicate of saying. 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 (72)-(75), however, could pose a problem for her analysis. In (72), 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*.

- (72) Últimadamente uno no se puede fiar de las noticias: En Al-Jazeera,
 lately one not CL.3S can trust in the news in Al-Jazeera
 que Obama va a atacar Iran. Pero en BBC, que va a retirar sus
 QUE Obama goes to attack Iran but in BBC QUE goes to retreat his
 tropas. (Sp.)
 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 (73), 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 by the analysis in Corr (2016).

-
- (73) Últimadamente uno no se puede fiar de las noticias: Hoy por la mañana, que Obama va a atacar Iran. Pero por la tarde, que va a retirar sus tropas. (Sp.)
 lately one not CL.3S can trust in the news today at the morning QUE Obama goes to attack Iran but at the afternoon QUE goes 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, (74) and (75) show that we can have more than one preceding element. In (74), there are two adjacent PPs. What is crucial is that the element that expresses a potential source is not adjacent to *que*. In (75), we have the DPs *el reportero de Al-Jazeera* and *el reportero de BBC* which qualify as sources again 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). However, the data show that a PP can intervene between a potential source or original speaker and *que*.

- (74) Últimadamente uno no se puede fiar de las noticias: En Al-Jazeera, hoy por la mañana, que Obama va a atacar Iran. Pero, en BBC, por la tarde, que va a retirar sus tropas. (Sp.)
 lately one not CL.3S can trust in the news in Al-Jazeera today at the morning QUE Obama goes to attack Iran but in BBC at the afternoon QUE goes 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.’
- (75) Últimadamente uno no se puede fiar de las noticias: El reportero de Al-Jazeera, hoy por la mañana, que Obama va a atacar Iran. Pero, el reportero de BBC, por la tarde, que va a retirar sus tropas. (Sp.)
 lately one not CL.3S can trust in the news the reporter of Al-Jazeera, today at the morning QUE Obama goes to attack Iran but the reporter of BBC at the afternoon QUE goes 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 (71)-(75) 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 go against 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 gets by without assuming elided material that could be modified, an alternative explanation is called for. My tentative proposal is to characterize the phrases that precede *que* as examples 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 proposal focusing mainly on the interpretation of the phrases that precede *que*. Although a syntactic analysis of frame setters will not be presented here, their high syntactic position could be an additional indication that the proposal is on the right track, since Benincà & Poletto (2004) showed that frame setters are always CP-external.

The first thing Krifka (2008) establishes is that frame setters are not topics. In particular, he underlies 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 (76). In B's answer the topic of the sentences is Daimler-Chrysler and not Germany or America.

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

Although there are apparent similarities, Krifka (2008) argues that frame setters neither constitute contrastive topics. In his theory, contrastive topics are defined as aboutness topics that contain a focus that generates alternative aboutness topic. 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 the expectation that information about a more comprehensive, or distinct entity will be given. Contrastive topics are used to show that the topic of the sentence diverges from this expectation. In (77), 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 uses *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.

(77) A: Does your sister speak Portuguese?

B: [My brother]_{ContrastiveTopic} does. (adapted from Krifka 2008, 268: ex 46)

With frame setters, Krifka (2008) states that the current common ground contains the expectation that more comprehensive or distinct information will be given. The frame setter indicates that the information is less comprehensive and just restricted to a particular dimension specified by the frame. Krifka (2008) illustrates this function with the example in (78). Speaker A's question leads to the expectation of an answer containing comprehensive information on John's 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 among other things. However, the frame restricts his well-being to the dimension of health.

(78) A: How is John?

B: [Healthwise]_{Frame} he is fine. (Krifka 2008, 269: ex 47)

Returning to the issue at hand, in the examples in (71) and (70a) repeated in (79a) and (79b), *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 (79a) gives us the origin of the statement and in (79b) the perspective or speaker that the statement is attributed to.

- (79) a. Oye, [en Al-Jazeera]_{Frame}, que Obama va a atacar Iran. (Sp.)
 Hey in Al-Jazeera QUE Obama goes to attack Iran
 'Hey, there's a saying on Al Jazeera that Obama is going to attack Iran.'
 (Etxepare, 2013, 98: ex 11)
- b. [La mare]_{Frame}, que què t'has cregut. (Cat.)
 the mother QUE what CL.2S have thought
 '[reportative:] The mother: what were you thinking-' (ebook-cat)

Krifka (2008) furthermore points out that explicit frame setters always allude to alternative frames. This can be seen in (76) where the explicit frames *in Germany* and *in America* are contrasted. The examples in (72)-(74) show that the same is true for *que*-preceding adverbials. In (80) the two frames introduce alternative speakers.

- (80) [La mare]_{Frame}, que què t' has cregut. [Jo]_{Frame}, que no volia
the mother QUE what CL.2S have thought I QUE not wanted.1S
tornar-ho a provar. (Cat.)
try-it-again
‘[reportative:] The mother: what were you thinking. [reportative:] I: I didn’t
want to try it again.’ (ebook-cat)

The examples in (79) and (80) suggest that if these phrases are truly frame setters, as I tentatively propose here, their precise function is context dependent. Even material, that is not mentioned explicitly can play a role. In (79a) the frame is interpreted as the origin of a reported statement and in (79b) 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 to syntactically subordinate indirect reportatives and showed that they behave in the same way with respect to CP-internal peripheral material, non-declarative sentences and in permitting reportatives that are not full sentences. This corroborates the idea that unembedded *que*-initial sentences are the same as their embedded counterparts, the only difference being that there is no matrix sentence expressed but the host material must be recovered from the context. 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 tentatively have suggested an analysis of these phrases as frame setters.

2.4 Contrasting *Que*-Initial Reports in Portuguese

The empirical evidence in the previous section was drawn from Spanish and Catalan, but Portuguese also exhibits instances of *que*-initial sentences that are interpreted as a reported speech. While they do differ in some aspects, in this section I make the claim that they are still one and the same syntactic phenomenon as in Spanish and Catalan and I outline my reasons for this assumption. I show that the slightly different properties of Portuguese unembedded reportatives can be explained to result from

independent syntactic properties and additional pragmatic requirements. The example in (81) 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.

- (81) Fiz anestesia geral mas só lá fiquei uma noite, porque
 made.1S anesthetics general but only there stayed one night because
 falei logo com o meu ginec. Que não valia a pena lá ficar
 talked.1S then with the my gynecologist QUE not was worthwhile there stay
 se depois não era vista por mais nenhum medico. (Pt.)
 if afterwards not was.1S seen by more any doctor
 ‘I got a general anesthetics 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)

- (82) a. O Sr. gosta de chafurdar na merda ideológica, filosófica,
 you.HORT like to wallow in-the shit ideological philosophical
 propagandística. Gosta de dizer que é muito bom, sabe muito bem,
 propagandistic like to say that is very good tastes very well
 que possui uma viscosidade suave e um calor aconchegante. [SubP Que]
 QUE posses a viscosity soft and a heat cozy QUE
 [ModP claro que_i] [FinP t_i] podia ser melhor, e será uma vez
 claro QUE could be better and will be once
 estivermos todos afogados em ela, e que um gajo até se
 are.SUBJ.FUT.2P all drowned in her and QUE a guy even CL.3S
 habitua ao cheiro rapidamente. (Pt.)
 habituates 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 all drowned in it. [reportative:] And
 one even gets used to the smell quickly.’ (CdP)
- b. Perguntei se ele lembrava de mim e ele disse [SubP que] [ModP
 asked.1S whether he remembered of me and he said that
 claro que_i] [FinP t_i] lembrava. (Pt.)
 claro QUE remembered
 ‘I asked whether he remembered me and he said that of course he remem-
 bered.’ (CdP)

Concerning 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 (82). The examples illustrate that in embedded (82b) and unembedded (82a) reportatives alike the complementizer surfaces in a high position, preceding for instance AdvC, which is analyzed in ModP. Importantly, this is the same behavior that we saw for Spanish and Catalan reportative sentences.

The examples in (83) 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 (83b) and unembedded reportatives (83a).

- (83) a. Quando cheguei a casa perguntei à minha mãe se agora as
 when arrived.1S at home asked to-the my other if now the
 mulheres bebiam cerveja no café. Que sim. E muito. (Pt.)
 women drank beer at-the coffeehouse QUE 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. 40 alunos o total de 71,4% responderam que o conteúdo associado
 40 students the total of 71,4% answered that the content associated
 à aula prática facilita o aprendizado; 6 (10,7%) responderam
 to-the class practical facilitates the learning 6 (10,7%) answered
 que não e 10 (17,9%) disseram que talvez. (Pt.)
 that no and 10 (17,9%) said 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, regarding their principle syntactic properties, unembedded reportatives are not distinct in Spanish, Catalan and Portuguese. In all three languages the complementizer occupies the same position in embedded and unembedded reportatives, identified with SubP in the present analysis.

There are two aspects in which Portuguese unembedded reportatives differ from the other languages that I will focus on in the following paragraphs. The first has to do with reported questions. While Spanish and Catalan admit *wh*-pronouns and the interrogative complementizer below *que*, Portuguese does not.²³ This is illustrated in

²³There are a handful of cases in CdP that could potentially be analyzed as cases of reported *que*-initial *wh*-questions. However, all of them are from Brazilian Portuguese (see also Corr 2016, 171-172).

(84a) for the *wh*-pronoun *onde* and in (84b) for the interrogative complementizer *se* (see also Corr 2016, 178: ex 61). Notably, the culprit is not the unembedded question introduced by a subordinating element. This can be seen in the example in (84b). A's repetition is different from initial question: it is introduced by the interrogative complementizer *se* marking the sentence as subordinate.

- (84) a. A: Onde és? (Pt.)
 where are.2S
 B: O quê é que disseste?
 what is that said.2S
 A: (*Que) onde és.
 QUE where are.2S
 ‘A: Where are you? B: What did you say? A: [reportative:] where are you.
 b. A: Vens? (Pt.)
 come.2S
 B: O quê é que disseste?
 what is that said.2S
 A: (*Que) se vens.
 QUE if come.2S
 ‘A: Are you coming? B: What did you say? A: [reportative:] are you
 coming.’ (adapted from Corr 2016, 178: ex 61)

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 holds true in syntactically embedded reportatives (see (85) and (86)).

An example is given in (IV).

- (IV) Lá nos anos 60, as pessoas eram mais religiosas, de forma que não comiam carne às sextas. Logo quem não come carne, não come hambúrguer. Então o dono do McDonald's pensou: “Vamos fazer um sanduíche que leva abacaxi em vez de carne.” Óbvio, não deu certo, nem um pouquinho. Que quê tem a ver abacaxi? Isso ninguém sabe explicar... (Pt.)

‘In the 60s people were more religious, in that they didn’t eat meat on Fridays. Consequently, who doesn’t eat meat, doesn’t eat hamburgers. So the owner of McDonald’s thought “Let’s make a sandwich that has pineapple instead of meat.” Obviously this didn’t work out, not even a little bit. [reportative:] What has pineapple to do with this? This nobody can explain...’ (CdP)

- (85) A Joana pergunta (*que) onde és. (Pt.)
 the Joana asks that where are.2S
 ‘Joana asked (that) where you are.’

- (86) A Joana pergunta (*que) se vens. (Pt.)
 the Joana asks that if are.2S
 ‘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. (52a), (53a) and (58), (59)).

As has been briefly mentioned in section 2.3.1, Spanish and Catalan also permit embedded questions without an initial complementizer (87b), paralleling the Portuguese structure in (85). In (87) the two structures are contrasted (cf. 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: (87a) is an embedded wh-question, while (87b) is a declarative in which the wh-pronoun is referential. This is demonstrated by the fact that (87b) can admit a continuation that spells out the referent, while the same continuation is infelicitous following (87a).

- (87) a. Juana repetió que quién vivía en la Rua da Saudade, #era Pereira.
 Juana repeated that who lived in the Rua da Saudade was Pereira
 (Sp.)
 ‘Juana repeated: who lived in Rua da Saudade: #it was Pereira.’
 b. Juana repetió quién vivía en la Rua da Saudade: era Pereira. (Sp.)
 Juana repeated who lived in the Rua da Saudade was Pereira
 ‘Joana repeated who lived in Rua da Saudade: it was Pereira.’

The example in (88) shows that the Portuguese structure is ambiguous. It admits a continuation that spells out the referent but also a continuation that makes the question-reading prominent.

-
- (88) A Joana repetou quem morava na Rua da Saudade:
the Joana repeated who lived in-the Rua da Saudade
- [continuation 1:] era Pereira. (Pt.)
was Pereira
- [continuation 2:] mas ninguem o sabia.
but nobody it knew
- ‘Joana repeated who lived in Rua da Saudade: it was Pereira./but nobody knew it.’

The important empirical facts are that Portuguese does not allow *que* to appear before a wh-pronoun or an interrogative complementizer in unembedded reportatives, but crucially neither in embedded reportatives, while Spanish and Catalan do allow this ordering in both contexts. This once again leads to the conclusion that *que*-initial reportatives are simply syntactically unselected reportatives with essentially the same properties as their embedded counterparts. What is essential to my argumentation is that these facts do not make a uniform syntactic treatment of the Portuguese unembedded reportatives and the equivalent structures in the other two languages implausible. On the contrary, the behavior of Portuguese *que*-initial reportatives is expected based on the proposal developed in this thesis and actually constitutes an argument in its favor. Within the general analyses proposed here, 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. The data are in line with an analysis that assumes that 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 (87b). However, 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 (87a).

I now turn to the second difference that relates to the contextual requirements that render *que*-initial reportatives felicitous. The central contrast is illustrated in (89).

- (89) a. A: No se oye bien. (Sp.)
not CL.3S hear well
- B: ¿Qué? ¿Eh?
what huh

- A: Que no se oye bien.
 QUE not CL.3S hear well
 ‘A: You can’t hear it well. B: Huh? A: (I said) you can’t hear it well.’
- b. A: Não se ouve bem. (Pt.)
 not CL.3S hear well
 B: O quê? Hein?
 what huh
 A: *Que não se ouve bem.
 QUE not CL.3S hear well
 B’: O quê é que disseste?
 what is that said.2S
 A’: Que não se ouve bem.
 QUE not CL.3S hear well
 ‘A: You can’t hear it well. B’: What did you say? A’: That you can’t hear it well.’ (adapted from Corr 2016, 148: ex 5)

While *que*-initial reportatives are felicitous without an explicit *verbum dicendi* in the context in Spanish and Catalan (89a), they are not in Portuguese where *verum dicendi* in the context is required (89b).

Corr (2016) considers this difference substantial enough to claim that the Portuguese-type unembedded reportatives are different to those found in Spanish and Catalan. This leads her to propose two different syntactic analyses (cf. section 2.1.2).

One empirical problem that Corr’s proposal faces is that Spanish and Catalan also show examples of the Portuguese type reportatives, i.e. of *que*-initial sentences where a *verbum dicendi* can be recovered from the context (cf. (90)).

- (90) A: No se oye bien. (Sp.)
 not CL.3S hear well
 B: ¿Qué dices?
 what say.2S
 A: Que no se oye bien.
 QUE not CL.3S hear 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 (89a) and (90) 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 (41) in section 2.1.2). Corr (2016) does not explicitly discuss how these facts are dealt with, nor whether two different analyses for (89a) and (90) are required.

Another aspect that makes the assumption of two different structures less convincing is that Portuguese unembedded reportatives do not actually exhibit any evident syntactic differences. 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. My conclusion is 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 the following section 2.5. In a nutshell the difference is that, while all three languages require salient material that can function as a host to 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 saliency. Extra-linguistic factors, for instance, can also play a role. 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 (89a) in Spanish and Catalan a *verbum dicendi* is salient. This is sufficient to make the use of a *que*-initial reportative sentence felicitous. In Portuguese, on the contrary, a conversational exchange is not enough: a *verbum dicendi* truly must be given to felicitously utter a *que*-initial reportative.

To conclude, 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, and this is indeed the case. Furthermore, I showed 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 manifest in embedded reportatives, which ultimately constitutes an argument in favor of the present analysis. Finally, I defended the claim that the true differences between the construction in the languages are related to pragmatics. In all three languages, there is a requirement for contextually salient material that can function as a host to 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, without the need for it to be explicitly mentioned.

2.5 Pragmatic Requirements

The previous sections focused on the syntactic properties of unembedded reportative sentences. The empirical facts supported the proposed analysis, that the syntactic behavior of *que* in these sentences does not differ from that of their syntactically embedded counterparts. I furthermore proposed that the empirical data can be captured without relying on an underlying performative or evidential syntactic structure responsible for the reportative interpretation. The claim is that the only information that the syntactic structure provides us with is that the sentence is subordinate.²⁴ The aim of the present section is to describe how the reportative interpretation arises. I suggest that this is not encoded in a hidden syntactic structure or a dedicated syntactic features but instead results from the contexts in which the sentences appear. In a nutshell, 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 to the subordinate proposition.

Corr (2016) observes that *que*-initial reportatives are not asserted and that the speaker is not committed to their truth. She illustrates this by the fact that with an unmarked declarative in (91a), the speaker cannot negate that she believes that the proposition is true, but in (91b), with a *que*-initial declarative, she can.

- (91) a. Juana y Aique están casados, #pero no es verdad. (Sp.)
 Juana and Aique are married but not is truth

²⁴I 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.

‘Juana and Aique are married, but it’s not actually true.’

(adapted from Corr 2016, 158: ex 23)

- b. Que Juana y Aique están casados, pero no es verdad. (Sp.)

QUE Juana and Aique are married but not is truth

‘Someone said Juana and Aique are married, but it’s not actually true.’

(adapted from Corr 2016, 158: ex 24)

This fact allows to draw a parallel with indirect embedded reportatives. In (92) the embedded indirect reportative is felicitously even though it is followed by a continuation that clarifies that the speaker made a statement he does not consider to be true.

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

Que-initial sentences have furthermore been shown to 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 (93) 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.

- (93) María: Antes de salir he hablado con mi madre. Le he dicho
Maria before of leave have.1S talked with my mother CL.3S have.1S told
que estoy a punto de encontrarme contigo. Y ella, que ¿qué tal Jorge?
that am about to meet-CL.1S with-YOU and she QUE how is Jorge
(Sp.)

Jorge: #Estoy bien.

Jorge am 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, that 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 (94).

- (94) [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 esta hecho. (Sp.)

DM QUE already is done

‘Hey, it’s already done.’ (adapted from Etxepare 2010, 613: ex 34)

The reportative sentence introduced by *que* is traceable in the contexts 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 (95) and (96) 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).

- (95) Avisa el comissari. Que ja pot venir. (Cat.)

notify the inspector QUE already can.3S come

‘Notify the inspector. [reportative:] He can already come.’ (ebook-cat)

In (95) *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 (96) the *que*-initial sentence is hypothetical. The speaker thereby introduces a question that she anticipates before the hearer has even asked it.

- (96) És una cosa que no t’ havia explicat mai, estimada. No vaig

is a thing that not CL.2S had told ever my love not PST

goser dir- t’ ho. Que per què ho vaig fer? Perquè jo

have courage tell CL.2S CL.NEUT QUE why CL.NEUT PST do because I

no sóc el meu pare. (Cat.)

not am 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 the claim of Etxepare (2010) and Corr (2016), for a *que*-initial reportative to be felicitous a previous speech event in which the original sentence was 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 (94) but also new as in (95) and (96).

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 section 2.4, that 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 (Chiaros 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. However, givenness is not a necessary requirement. 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 *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 (97) to (100), I exemplify cases of *que*-initial reportatives in different contexts to illustrate my claim. In (97) 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. As predicted, the host material contains the *verbum dicendi* *dir* ‘say’.

- (97) Tenia ganes de començar o amb el rus o amb l'arameu, però
 wanted to start either with the Russian or with the Aramenaic but
 la mare va entrar a l'habitació i va dir ni
 the mother PST enter to the room and PST say not even
 parlar-ne. Que ja estava bé amb aquelles llengües que
 talk about-CL.PART QUE already was well with these languages that
 sabia. (Cat.)
 knew
 '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 (98) repeated from (89b), B's question introduces the host material. This example shows that there is no identity required. 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.

- (98) A: Não se ouve bem. B: O quê é que disseste? A: Que não se ouve
 not CL.3S hear well what is that said.2S QUE not CL.3S hear
 bem. (Pt.)
 well
 'A: You can't hear it well. B': What did you say? A': That you can't hear it
 well.' (adapted from Corr 2016, 148: ex 5)

The example in (99) is different because no *verbum dicendi* is given. The relevant *que*-initial reportative is uttered within a conversational exchange. This seems to be sufficient in Spanish and Catalan to make unembedded reportatives felicitous. The conclusion within the present explanation must be that a *verbum dicendi* is salient in a conversational exchange.

- (99) Inf.a - Hoy es el día de salida. Inf.b - ¿Eh? Inf.a. - Que hoy es el
 today is the day of departure Hm? QUE today is the
 día de salida. (Sp.)
 day of departure
 'A: Today is the day of departure. B: Hm? A: [reportative:] Today is the day
 of departure.' (CdE)

Finally, in the example in (100) repeated from (96), the *que*-initial reportative appears in a context without a given *verbum dicendi* and without an active interlocutor. Just

as (99), an unembedded reportative in such a context is only possible in Spanish and Catalan. In the fragment, there is only one speaker present. However, a hearer is addressed with *estimada*. This could suggest that activating a hearer perspective is enough for a *verbum dicendi* to become salient.

- (100) És una cosa que no t' havia explicat mai, estimada. [...] Que per què
is a thing that not CL.2S had told ever my love QUE why CL.NEUT
ho vaig fer? (Cat.)
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 that illustrate 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 described 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) is sufficient. A case of pragmatic control is illustrated in (101).

- (101) [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 (102a), illustrating the same issue. The pronoun *he* is pragmatically inferred to be *Tony*, a referent who is salient but not given. In (102b), *it* is understood to mean *diving off the high diving board*, once again a salient yet not given expression.

- (102) 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 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 (99) (100)) or it can be made salient through an explicit mention (cf. examples (97), (98)).

To conclude, in sections 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* parallels 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. The present 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. This section has only offered a description of the pragmatic requirements. In order to formalize the generalizations in the future, a number of issues need to be resolved. The most pressing of these is the concept of pragmatic reconstruction: it requires further research in order to determine what it is that is reconstructed (salient expression, entire matrix clause, etc.). Furthermore, it needs to

be established what being a host for a subordinate sentence actually means and what properties the salient host has to have to be identified as such. Finally, it should be noted that a more developed version of this proposal could also be reconciled with the analysis in Etxepare (2013) (and possibly Demonte & Fernández Soriano 2014) since it could present a solution to the question of how the elided noun of saying enters the common ground.

2.6 Summary

This chapter has dealt with *que*-initial reported sentences. I examined their properties and developed an analysis to account for them. A new theoretical concept introduced in this chapter is the 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 analyze *que* as a normal complementizer without having to stipulate a new linguistic category. 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 of unembedded *que*-initial sentences surfaces in the same position as the complementizer in their embedded counterparts, suggesting that they can be analyzed in the same way as heads of the highest left-peripheral position SubP valued with the interface feature *subordinate*. The syntactic analysis is simple and, contrary to the 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.

3 On Attributive *Que*

This chapter explores the syntactic behavior and discourse contribution of the complementizer in the constructions exemplified in (103).

- (103) a. Tranquilo, tío. Que no muerdo. (Sp.)
 calm man QUE not bite.1S
 ‘Chill, man. I don’t bite.’
- b. Que no muerdes? (Sp.)
 QUE not bite.2S
 ‘You don’t bite?’
- c. Que chato que é. (Pt.)
 how annoying QUE are.3S
 ‘How annoying this is.’
- d. Segur que son amics. (Cat.)
 sure QUE are.3P friends
 ‘Surely, they are friends.’
- e. Sí que son amics. (Cat.)
 VERUM QUE are.3P friends
 ‘They ARE friends.’

In (103a), the complementizer appears sentence-initially in a declarative and in (103b), in a polar question. In (103c), the complementizer surfaces below the *wh*-expression of a *wh*-exclamative. In (103d), for which I use the term AdvC, it follows an epistemic modifier and in (103e), I will call AffC, the complementizer follows the particle *sí* in a *verum* construction.

In the various sections of this chapter, I will argue that in all of these constructions the complementizer is merged at the right-most 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 choice to use the label *attributive que* (see section 3.4 for a more detailed description of the meaning involved).

This chapter is organized as follows: In section 3.1, I discuss the previous analyses that have been proposed for the relevant constructions in the literature. In section 3.2, I outline my own account and compare it to previous approaches. In section 3.3, I report the empirical evidence behind the syntactic analysis and discuss further syntactic properties of the individual constructions. Finally, section 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 that has given rise to the arguments laid out in this chapter is that the different constructions exemplified in (103) all involve an instance of the complementizer merged in the same position and valued with the same feature. A connection between all these constructions has not been made explicitly in the literature to date; consequently, the analyses that I discuss in the following sections, do not account for the whole phenomenon at once but deal separately with the individual constructions. Also, 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 account defended in this thesis. Section 3.1.1 discusses the accounts for *que*-initial declaratives with a non-reportative interpretation. Section 3.1.2 reviews the treatment in the literature of *que* in polar questions. In section 3.1.3, I present the previous analyses for *que* in wh-exclamatives. In section 3.1.4 I summarize the accounts for *que* when following epistemic and evidential modifiers and in section 3.1.5, the accounts for *que* in verum marked sentences.

3.1.1 Corr (2016)

The most extensive study of *que*-initial sentences with a non-reportative reading is found in Corr (2016). The author distinguishes between two types, exclamative QUE and conjunctive QUE, on an interpretive and structural level. Exclamative QUE is illustrated in (104).

- (104) a. (Ai) que t' atrapo! (Cat.)
 DM EXCL CL.2S catch.1S
 ‘I’m coming to get you!’ (Corr 2016, 88: ex 9)

- b. Ai, quo o gato se me foi ao peixe! (Pt.)
 DM EXCL the cat CL.REFL CL.1S went to the fish
 ‘The cat went off after the fish!’ (Corr 2016, 88: ex 11)
- c. ¡Que hemos salido en la radio, oiga! (Sp.)
 EXCL AUX.1P go in the radio DM
 ‘We’re on the radio, look!’ (Corr 2016, 92: ex 31)

Similar data have been addressed by Biezma (2008) on Spanish and Ledgeway (2012) on Romance in general. Otherwise the use of a complementizer in these types of exclamatives has gone mostly unnoticed in the formal literature on Ibero-Romance. Similar phenomena found in other languages have been mentioned by some authors (cf. Sæbø 2005 on French, Kerstin 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, i.e. 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 (105) (repeated from (36) see section 2.1.2 for a characterization of the UP).

- (105) [SA_{high} [SA_{low} [EvalP [EvidP [DeclP ...]]]]]

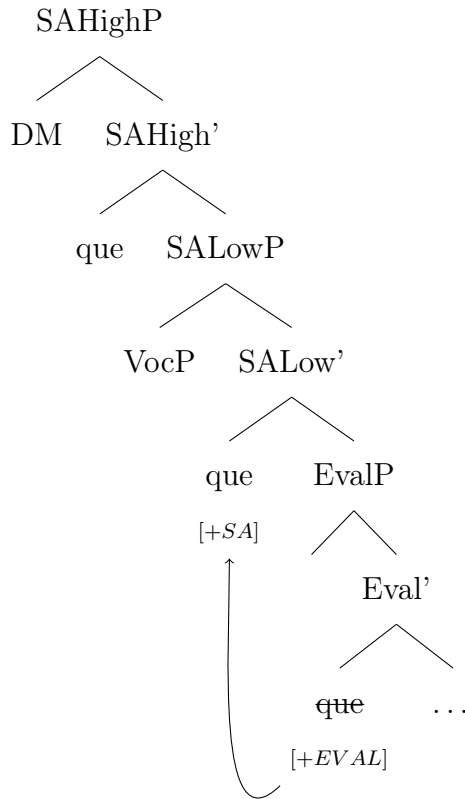
The data in (106) show that exclamative *QUE* follows discourse markers and vocatives.

- (106) a. (Ai/ apa) que (*ai/ apa) em poso vermella. (Cat.)
 DM DM EXCL DM DM CL.1S put red
 ‘Ohh/gosh, I’ve gone red!’ (Corr 2016, 132: ex 118)
- b. (Amor) que (*amor) em poso vermella. (Cat.)
 love EXCL love CL.1S put red
 ‘Darling, you make me blush!’ (Corr 2016, 134: ex 122)

To capture the syntactic properties, Corr (2016) proposes the analysis in (107). Exclamative *QUE* is assumed to be merged in Eval^o 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*, that 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. (108)).²⁵

- (107) Licensing of exclamative QUE in the Ibero-Romance left periphery according to Corr (2016)



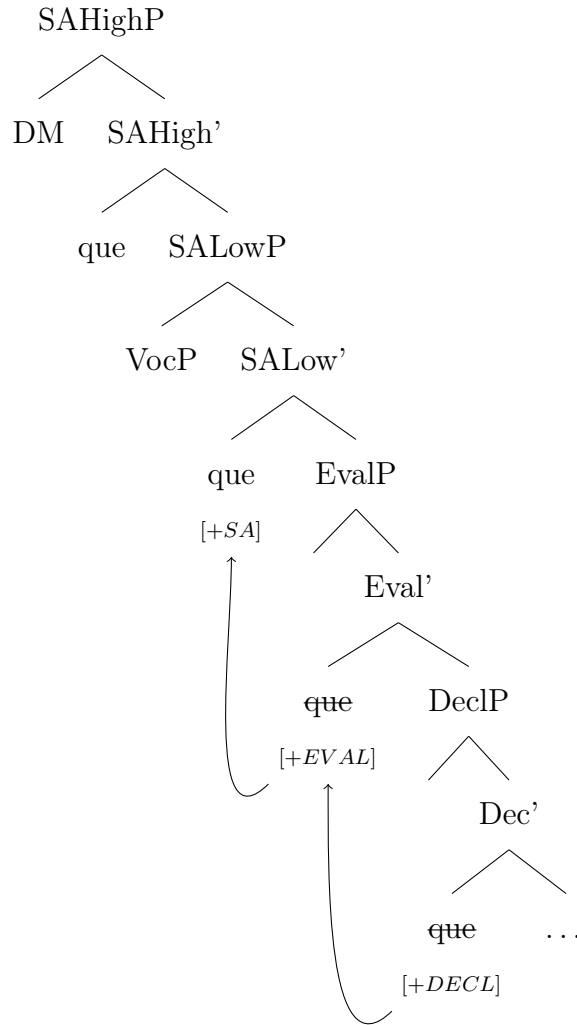
(Corr 2016, 137: ex 134)

- (108) Ai, fofinha (*que) olha que estas com sorte! (Pt.)
 DM cute.DM EXCL DM QUE are.2S with luck
 ‘Oh, darling, gosh aren’t you lucky.’ (Corr, 2016, 137: ex 132)

²⁵In my analysis, one possible explanation for the ungrammaticality of *que* above certain discourse markers in examples such as (108), is that we would be 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.

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).

- (109) Licensing of exclamative QUE in the European Portuguese left periphery according to Corr (2016)



(Corr 2016, 236: ex 112)

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 language exclamative QUE is merged lower, in Decl° where it picks up the additional declarative feature (as

in the structure in (109)).²⁶

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).

An example from contemporary Spanish is given in (110). Conjunctive QUE sentences often, but not always (see the example in (111)), follow imperatives.

- (110) No me pises, que llevo chanclas. (Sp.)
not CL.1S step.SUBJ2S CONJ wear.1S flip flops
'Don't step on me, I'm wearing flipflops.' (adapted from Corr 2016, 229: ex 90)

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.²⁷ She illustrates this with the example in (111).

- (111) 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. (Cat.)
QUE CL.1S CL.PART go.1S of holidays
'I'm going on holiday.'
b. *Perquè/ car me' n vaig de vacances. (Cat.)
because for CL.1S CL.PART go.1S of holidays
(Corr 2016, 226: ex 84)

Conjunctive QUE, but not a true causal connective like *perquè/car* (cf. (111b)), is felicitous without a previous utterance and can serve to explain a non-linguistic situation

²⁶The 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 section 2.1.2). It is thus surprising that the author presents an analysis for exclamative QUE in European Portuguese that does not involve EvidP.

²⁷For more arguments again a primary causal or explicative nature of these complementizer, see section 3.4.

(cf. (111a)). 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.

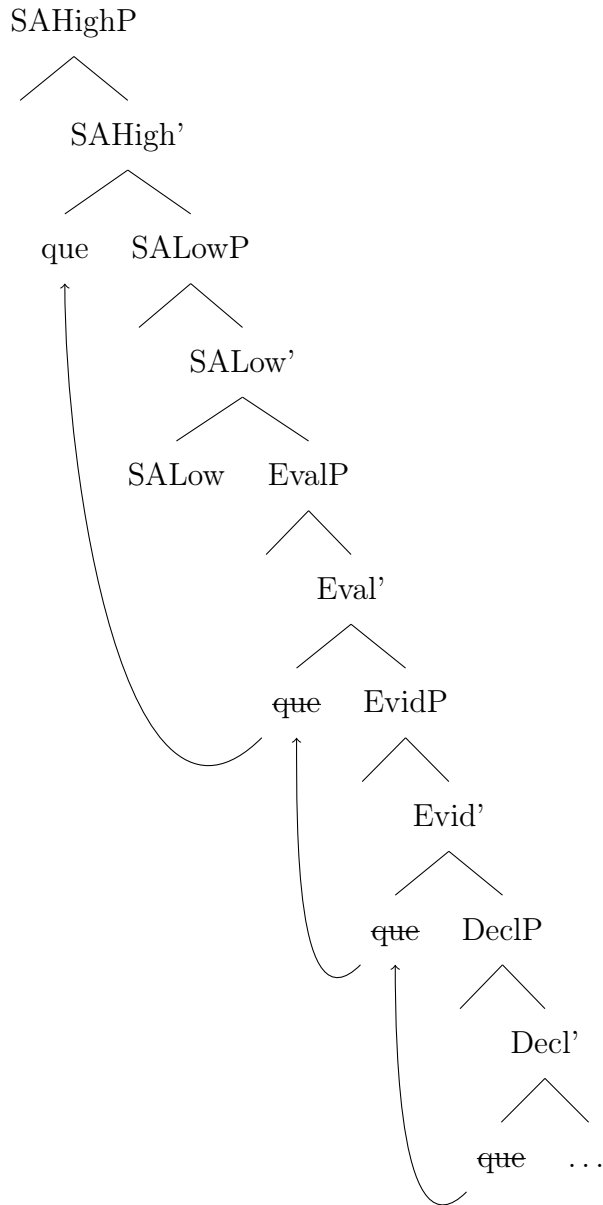
Corr (2016) maps out the syntactic position against the backdrop of her UP. With the data in (112), she shows that conjunctive QUE cannot co-occur with discourse makers or vocatives.

- (112) a. ¡Escúchame, (*oye) que (*oye) vamos a llegar tarde! (Sp.)
 listen.IMP=CL.1S DM CONJ DM go.1P to arrive late
 ('Listen, *hey we're going to arrive late!') (Corr 2016, 235: ex 109)
- b. ¡Escúchame, (*María) que (*María) vamos a llegar tarde! (Sp.)
 listen.IMP=CL.1S María CONJ María go.1P to arrive late
 ('Listen, *María we're going to arrive late!') (Corr 2016, 235: ex 110)

Corr (2016) proposes that conjunctive QUE reaches SAHigh°, the highest head in the UP. The complementizer is assumed to be merged in Decl°, the lowest head of her split ForceP. On its way to SAHigh°, it passes through Evid° and Eval°. Conjunctive QUE is valued with the features hosted by the individual heads.²⁸ Corr (2016) assumes that the merger occurs in DeclP because conjunctive QUE is restricted to declaratives. The evidential feature picked up in Evid° is explained on the grounds that conjunctive QUE shows a 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.

²⁸As an alternative to the movement analysis, Corr (2016) proposes a structure in which the features are valued simultaneously by a syncretic head (cf. Corr 2016, 236: ex 111).

-
- (113) Head movement analysis of Ibero-Romance conjunctive QUE according to Corr (2016)



(Corr 2016, 236: ex 111)

To conclude this section, I will now point out a potential problem with the analysis. The main criticism 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. In fact, the explanation presented in Corr (2016) is only founded on the theoretical

assumptions made by the author when she relates some interpretative properties of the constructions to the effects of the features in her assumed UP. This is a potential issue because, as in her analysis of reportative *que* discussed in section 2.1.2, the question of how the abstract features result in the specific meaning is not answered satisfactorily. In principle, the empirical data are also compatible with an analysis assuming that *que* in both of the two constructions 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 (106) 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 the one 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 (112), show that conjunctive QUE is incompatible with vocatives and discourse markers, therefore the assumption that the complementizer here reaches a high UP projection lacks compelling empirical support. Again, the conclusion that the complementizer simply occupies the highest CP head is equally plausible. For now, then, pending further empirical evidence of movement to a CP-external projection, the position I adopt in the present thesis 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 in the three languages under discussion here, has 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, as I will show in section 3.3.2, 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 sums up 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 or absence of *que* in different pragmatic contexts. They argue that *que* is virtually unrestricted in its presence in Minorcan Catalan,

whereas the other dialects are more selective. All the 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 (114).

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

The anti-expectational nature of the context becomes evident from the statement that follows the *que*-initial polar question where the speaker explicitly states 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. (115)). According to the authors, the choice of the particle depends on the dialect.

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

Castroviejo (2018) suggests that, at least in sentence final position like in (116), a subset of these particles are not synonymous but encode different meanings (see also section 3.4).

- (116) T' has tallat els cabells, oi?/ eh? (Cat.)
 CL.2S have.2S cut the hair OI EH
 'You had your hair cut, right?/ huh?' (Castroviejo 2018, ex 19)

Other non-neutral polar questions that permit *que* are rhetorical questions as in (117).²⁹

²⁹Although this is not central to the present discussion, in Kocher (2017), I propose that questions like (117) 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 and constitute an exaggerated version of the question for which they stand in.

- (117) Que et penses que tinc quatre mans, jo?
 QUE you think that have.1S four hands I
 ‘Do you think I have four hands?’ (Prieto & Rigau, 2007, 18: 36b)

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.

- (118) Que em deixes el teu apartament de la platja, aquest cap de setmana?
 QUE me leave the your apartment of the beach this weekend
 ‘Would you let me use your apartment by the beach this weekend?’
 (Prieto & Rigau 2007, 4: ex 8a)

The use of *que* in (118) is felicitous only if the hearer has offered the apartment to the speaker previously. In a context where this is not the case, they are not felicitous according to Prieto & Rigau (2007).

- (119) Que puc fumar?
 QUE can smoke
 ‘Can I smoke?’
 (Prieto & Rigau 2007, 4: ex 9a)

Similarly, (119) 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 to the structure of these constructions, and assume that *que* is merged at the lowest edge of the left periphery in FinP. They propose the structures in (120).

- (120) a. $[_{\text{ForceP}} [_{\text{Operator}} \text{Oi}] [_{\text{Force}} +\textit{confirmative interrogative} [_{\text{FinP}} \text{que} [_{\text{IP}} \text{en Pere no} \\ \text{OI} \text{QUE the Pere not} \\ \text{va a Barcelona?}]]]]]$
 goes to Barcelona
 ‘Pere isn’t going to Barcelona, right?’ (Prieto & Rigau 2007, 25: ex 56a)
- b. $[_{\text{ForceP}} [_{\text{Operator}}] [_{\text{Force}} +\textit{anti-expect./neutral interrogative} [_{\text{FinP}} \text{que} [_{\text{IP}} \text{no} \\ \text{QUE not} \\ \text{volies un collaret?}]]]]]$
 wanted.2S a necklace
 ‘Didn’t you want a necklace?’ (Prieto & Rigau 2007, 25: ex 56b)

In their approach, 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.

To conclude this section, I will evaluate 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.

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

Que les taronges_j les_j vas comprar tu? (Cat.)

QUE the oranges CL.FP PST buy you

‘The oranges, you bought them, didn’t you?’ (Kocher 2017, 49: ex 98a)

The data in (121) pose a problem for the analysis proposed in (120b), because in the example, a clitic left dislocated topic follows rather than precedes *que*. Furthermore, based on their analysis for particle questions in (120a), we would expect that certain phrases should be able to intervene between the particle and the complementizer. However, the data in (122) show that the particle and the complementizer must be adjacent (more data are discussed in section 3.3.2).

(122) *Oi en Jordi que l’has convidat tu?

OI the Jordi QUE CL.MS have.2S invited tu

‘intended: You are the one that invited Jordi, right?’

To reconcile these data I will propose a revised analysis in section 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 position of the left periphery.

Prieto & Rigau (2007) offer a detailed characterization of different pragmatic contexts that license *que* in polar questions. Based on these, I propose a generalization that, as will be seen in section 3.4, has the potential to ultimately enable the assumption 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 (114), while the speaker’s belief was the opposite, contextual evidence (see Kocher 2017 and section 3.4 for a revised definition of the notion of contextual evidence presented in Buring & Gunlogson 2000), suggest that the answer is going to be positive. In confirmatory contexts like (115), the speaker’s belief itself makes her expect a positive

site for dislocated elements and can be considered parallel to a Rizgian 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.

- (124) [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 section 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.

- (125) [EvaluativeP *que* *livro_i* [Evaluative' [AssertiveP [Assertive' *que/t_i* [XP *o* *João_j* [WhP *t_i*
what book QUE the João
[Wh' [FocusP *t_i* [Focus' [XP [IP *t_j* *leu* *t_i*]]]]]]]] (Pt.)
read

‘What a book (that) John read!’ (Ambar 2003, 238-239: ex 88-89)

Demonte & Fernández Soriano (2009) propose an account of Spanish wh-exclamatives that relies on a cartography of the left periphery 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. (126)).

- (126) [FocP *¡Qué rico*] ... [FinP (que)] [IP *está!*] (Sp.)
how good QUE is

‘How good this is!’

(Demonte & Fernández Soriano 2009, 33 : ex 19a, analysis added by the author)

All these analyses are based on different theoretical assumptions, making it difficult to compare them. Demonte & Fernández Soriano (2009) assume a Rizgian 1997 style structure of the left periphery that is similar to my proposal. One potential issue with their analysis, however, is that it cannot account for the data in (127).

- (127) a. **Qué raro* a *Juana_i* *que* *la_i* *has* *invitado* *pero* *no* a
how strange DOM Juana QUE CL.FS have.2S invited but not DOM
María. (Sp.)
María

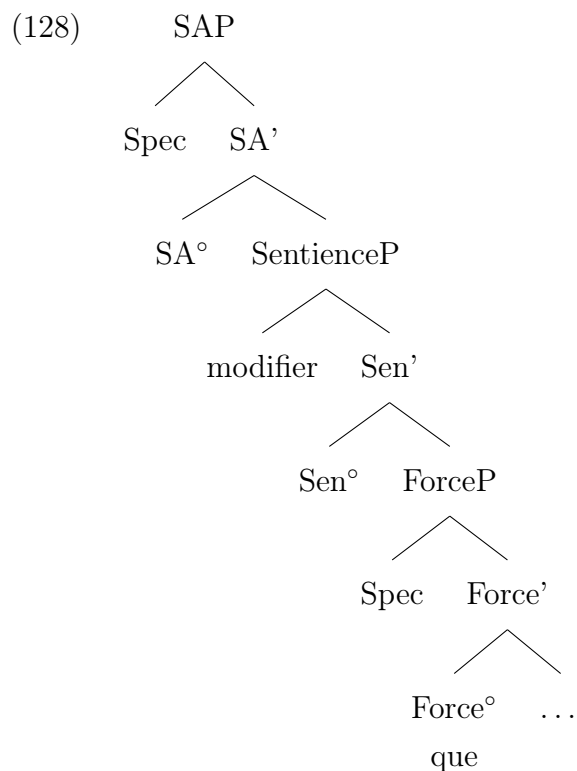
- b. A Juana_i qué raro que la_i has invitado pero no a
 DOM Juana how strange QUE CL.FS have.2S invited but not DOM
 María.
 María
- c. Qué raro que a Juana_i la_i has invitado pero no a
 how strange QUE DOM Juana CL.FS have.2S invited but not DOM
 María.
 María
- ‘How strange that you invited Juana but not Maria.’

The example in (127a) shows that a dislocated topic cannot intervene between the *wh*-expression and the complementizer, even though their account predicts that it should be able to. (127b) shows that the *wh*-expression can be preceded by a dislocated topic, which is in line with the analysis by Demonte & Fernández Soriano (2009). 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, Freitas Barros 2006, Hernanz & Rigau 2006, Ocampo 2006, Rodríguez Ramalle 2007; Rodríguez-Ramalle 2008, 2015, Gras 2010, Sansiñena et al. 2015). Apart from my own analysis (Kocher (2014, 2017), the previous more extensive accounts of the construction focus primarily on other Romance languages and not 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. 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 analyze the modifier in a CP-external projection. In the structure in (128), as an example, I illustrate the analysis proposed in Cruschina & Remberger (2018) based on Speas & Tenny (2003).



(adapted from Cruschina & Remberger 2018, 350: ex 23)

One potential issue for an analysis along these lines are data such as those in (129).

- (129) a. Disse que certamente que iria ver logo resultados. (Pt.)
 said that certainly QUE go.COND see soon results
 ‘S/he said that certainly I would see results soon.’ (CdP)
- b. Otra canción que claro que escuchamos todos y que podría parecer
 other song that claro QUE listen-to all and that could seem
 muy buena, es “Realmente no estoy tan solo”. (Sp.)
 very good is Realmente no 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. I per això us hem preparat un article que segur que us
 and therefore CL.2P have.1P prepared an article that sure QUE CL.2P
 serà útil un moment o altre. (Cat.)
 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)

These examples show that the construction is not restricted to root contexts but also appear in embedded contexts, and in particular also in relative sentences (cf. an appositive relative in (129b) and a restrictive relative in (129c)). 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 (129a) 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 (129).

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 a number of studies, notably in Martins (2006, 2013), González Rodríguez (2008, 2009, 2016) Escandell-Vidal & Leonetti (2009b), Escandell-Vidal & Leonetti (2009a), Escandell-Vidal (2011) and Rodríguez Molina (2014). The most widely adopted analysis of the verum construction in Spanish is developed in Hernanz (2007) (see also Batllori & Hernanz 2008 for an application of the analysis to diachronic data). The analysis has also been extended to Catalan in Batllori & Hernanz (2013).



(adapted from Hernanz 2007, 144: ex 87)

Hernanz (2007) proposes the structure in (130). The complementizer is merged as the head of ForceP and *sí* is assumed in its specifier.

Verum sentences do not require the presence of *que* in either language. Hernanz (2007) compares *sí que*-sentences (as in (131a)) with its *que*-less equivalents (as in (131b)).

- (131) a. *Sí que ha llovido hoy.* (Sp.)
 yes that has rained today
 ‘It HAS indeed rained today.’ (Hernanz 2007, 134: ex 3a)
- b. *Sí ha llovido hoy.* (Sp.)
 yes has rained today
 ‘It HAS rained today.’ (Hernanz 2007, 134: ex 1a)

She notes that they differ in that the version with the complementizer stresses an already mentioned proposition. She adopts an idea from Etxepare (1997) and states that a proposition introduced by the verum marker requires a linguistic antecedent. Hernanz (2007) proposes to attribute this function to ForceP. Since this aspect of meaning, according to the author, is not present in the *que*-less verum sentences, she adopts a different syntactic analysis for these, which is exemplified in (132).

- (132) [ForceP [TopicP [FocusP *sí*_i [PolP *t_i* [IP ...]]]]]
 (adapted from Hernanz 2007, 129: ex 48)

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 that sentence negation *no* is located in (see Laka 1990). Hernanz (2007) argues that the *sí* in these contexts has focal properties, which is the reasons 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 section 3.3.5 and Kocher 2017 for a more detailed discussion). Hernanz’s analysis for *sí que*-sentences fails to account for data such as given in (133) which show that *sí que* is also attested in embedded sentences (see (133a), (133b)) and can be preceded by a clitic left dislocated topic (see (133c)).

- (133) a. En el bar de la Confederació General del Treball (CGT) confirmen
in the bar of the Confederació General del Treball (CGT) confirm
que sí que hay huelga. (Sp.)
that VERUM QUE there is strike
‘In the bar of the CGT they confirm that there is a strike going on.’
(Kocher 2017, 94: ex 32b from caWac)
- b. A banda d’aquest dissentiment inicial, hi ha dos partits de
at side of this disagreement initial there are two parties of
l’oposició que sí que la votaran afirmativament. (Cat.)
the opposition that VERUM QUE it vote affirmative
‘Concerning this initial disagreement there are two parties of the opposition
that WILL vote in favour of it.’
(Kocher 2017, 45: ex 88b from caWac)
- c. A López_i sí que le_i he visto agredir a dos de mis
DOM López VERUM QUE CL.MS have seen attack at two of my
jugadores.
players
‘I HAVE seen López attack two of my players.’
(Kocher 2017, 94: ex 32a from caWac)

Finally, there is also a piece of data that is problematic Hernanz’s analysis for the *que*-less *sí*-sentences. Given that they are assumed to start out in PolP, examples like (134) are unexpected.³¹

- (134) Eso sí no podía faltar en ninguna casa. (Sp.)
this VERUM no could miss in any house
‘This could NOT be missing in any house.’ (Kocher 2017, 94: 31a from CdE)

The example shows that *sí* can co-occur with sentence negation *no*, which is usually analyzed in PolP, therefore the same position that *sí* is assumed to start out in Hernanz (2007). The alternative analysis I adopt that will be presented in section 3.3.5 can account for these data. It is compositional. This means that *sí* always occupies the same left-peripheral position and the only difference between the two alternative means of expressing verum lies in the presence of *que*.

³¹These data contradict Hernanz (2007, 139: fn 8), where it is claimed that in these configurations *sí* cannot co-occur with negative particles.

3.2 Outline of the Present Analysis

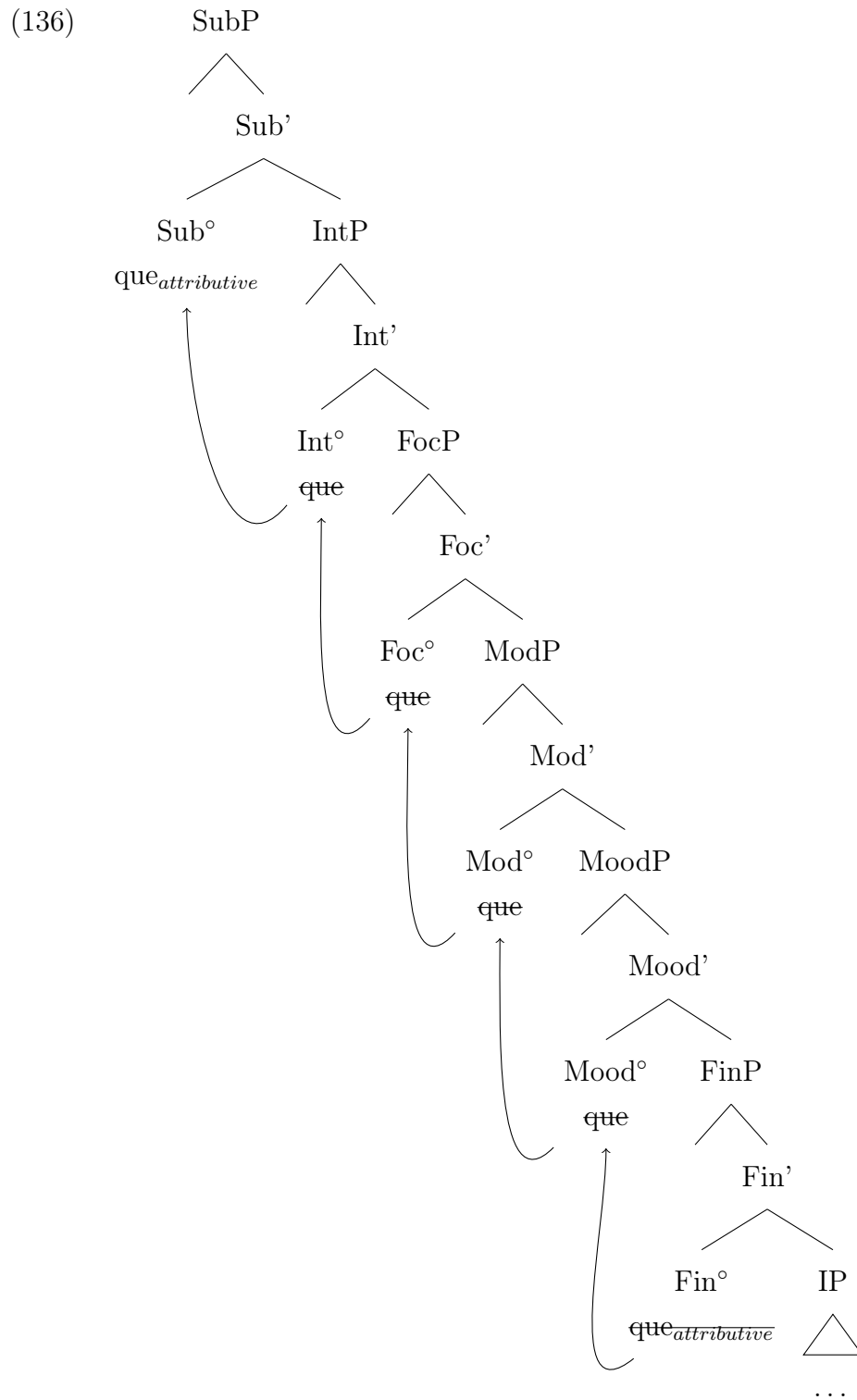
One central aim of this thesis is to develop an analysis that is compositional on a structural and interpretive level. In very simple terms, what I mean by this is that my goal is to develop an account in which, unless there is empirical evidence to the contrary, each element involved in the construction 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 (135) (repeated from (19)) that I assume for the present thesis. To repeat, the main differences between the split CP used here and the one developed by Rizzi is that I adopt SubP instead of ForceP and assume a lower projection termed MoodP which takes over the functions originally associated with ForceP (cf. section 1.2).

- (135) [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. section 3.4). However, the word order suggests that the complementizer does not occupy the same position in the different constructions (see sections 3.3.1 to 3.3.5). One possible way of accounting for these facts, which will not be outlined in detail here, would be to assume that a lexical item with a dedicated function is inserted directly into the different positions. This option could then account for the shared meaning and explain why the complementizer surfaces in different positions. This approach, however, 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 thesis that there is only one lexical item *que* whose meaning is determined in the syntactic position in which it is externally merged, and allows accounting in a unified way for the constructions involving attributive *que* and the *que*-initial reportatives discussed in chapter 2.



The position hosting the attributive feature that results in the merger of the complementizer is at the right edge of the left periphery. The complementizer reaches its final

surface position through head-to-head movement. As will be shown over the course of the following sections, 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 (136) illustrates these basic ideas. In the following paragraphs, I briefly present some theoretical support for each of the basic assumption in the analysis. The detailed empirical basis is given in section 3.3 and 3.4.

The assumption that the complementizer starts out in the lowest projection of the left periphery, FinP, goes back to 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 on Abruzzese). The data from the Corsican variety reported in Ledgeway (2012) are particularly interesting for the present investigation, because the different morphological forms coincide with different readings of the proposition. *Chì*, the complementizer merged in ForceP, introduces a declarative. An example is given in (137a) where it heads a sentence embedded under a verb of saying. Contrastingly, *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. The interesting parallel, however, is that they are also able to introduce main clauses as is illustrated in (137b).

- (137) a. *Dì a Caccara chì, à ott' ore sì Diu vole, saremu in*
 tell.IMP.2S to Grandma that at eight hours if God wants we.shall.be in
casa (Cor.)
 house
 'Tell Grandma that, God willing, we shall be home by eight o'clock'
 (Ledgeway 2012, 175: ex i.a)
- b. *Chè vo un caschete! (Cor.)*
 that you= not fall
 'Watch you don't fall.' (Ledgeway 2012, 175: ex i.b)

The empirical distribution of the Ibero-Romance complementizer in the relevant constructions (cf. section 3.3) shows that it always occupies positions above FinP and can also occupy the position immediately above it. These facts can be accounted for by assuming that the complementizer is base-generated below the lowest head in the

relevant constructions 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. The feature I propose is termed *attributive*. My approach is based on insights presented in de Cuba & MacDonald (2013). These authors show that complement clauses of factive and non-factive verbs have a different interpretation and also a different structure. An 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 (138) and (139).

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

- (139) 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 (138) are integrated at the IP-level and have a more reduced structure than adverbial clauses like (139), which she considers to be adjoined to the host clause at a later stage in the derivation. The latter clauses start with a SubP and contain essentially the full set of projections that are also found in root clauses. Instead, she claims that adverbial complements integrated at the IP-level like (138) 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.

- (140) a. John thinks that this book Mary read. (*non-referential*)
(de Cuba & MacDonald 2013, 8: ex 10a)
b. *John regrets that this book Mary read. (*referential*)
(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) (140b) *John doesn't regret that Mary read this book.*, it is still presupposed that Mary read this book. In the negated version of (140a), *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 (140), shows that while non-factive sentences admit a left dislocated topicalization of the object *this book* (cf. (140a), the same dislocation is ungrammatical in factive complements (cf. (140b)). The authors draw the conclusion that the structure 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. (141a)), it is ungrammatical in a factive complement clause (cf. (141b)).³²

- (141) a. Juan cree que ese libro_i ya se lo_i había leído. (Sp.)
 Juan believes that that book already CL it had read
 (*non-referential*)
 ‘Juan believed that that book he had already read.’
 (de Cuba & MacDonald 2013, 9-10: ex 11c)
- b. *Sabía a Juan_i qué le_i había prometido el decano. (Sp.) (*referential*)
 knew to Juan what him had promised the dean
 ‘I knew what the dean had promised John.’
 (de Cuba & MacDonald 2013, 10: ex 12a)

In de Cuba & MacDonald (2013) the terms introduced to distinguish between the two types of complement sentences are *referential* vs. *non-referential*. The logic behind these terms is that a sentence introduced by a factive verb is considered by the authors as referential in the sense that there is a sentence in the common ground that it refers

³²The critical example of a complement sentence under the factive verb *saber* ‘to know’ in (141b) 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 (V), in which case, the complementizer follows *qué*, which is an indication that the complementizer must be located lower than FocP.

- (V) ¿Sabes qué que le dice una madre a su hijo informático? (Peruvian Sp.)
 know what QUE him tells a mother to her son computer scientist
 ‘You know what (that) a mother tells her computer scientist son?’ (CdE)

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 section 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.

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

Table 2: Properties of referential and non-referential complements in de Cuba & MacDonald (2013) and Villa-García (2015).

The aim of de Cuba & MacDonald (2013) is to capture the contrast in embedded sentences. The proposal I make is that their insights can be adapted to account for unembedded sentences, which are the main focus of this thesis. 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 section 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 assumption is mainly based on the empirical facts that are given in section 3.3, which show that the complementizer surfaces at different positions in the different constructions. The assumption that the complementizer moves, however, is not unprecedented. Independently Rizzi (1997), Polletto (2000), Roberts (2001), Ledgeway (2005), Belletti (2009, 2013) assume the merger of a complementizer in a lower CP position and movement to a higher one. The precise reasons behind the movement of the complementizer is not a central concern for the current line of argumentation. For the present purposes, 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 aside 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 also 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, Villa-García 2015, Corr 2016).

My complementizer-movement analysis predicts that *que* can in principle move all the way up to the left edge of the functional field, SubP.³³ 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 section 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

³³In Corr (2016) it is suggested that the complementizer can even reach a position in her performative field above the CP.

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 phrase if the specifier of the projection of which the complementizer is currently the head 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 this idea is given in section 3.3. The following examples show the analyses I assume for the different constructions. For attributive *que* in declaratives (shown in (142a)) and polar questions (shown in (142b)), since there is no material hindering it, the movement of complementizer is unrestricted and consequently *que* does in fact reach the highest projection SubP.

- (142) a. Tranquilo, tío. [SubP Que_{attributive}] ... [FinP ~~que~~_{attributive}] [IP no muerdo.]
 calm man QUE QUE not bite.1S
 (Sp.)

‘Chill, man. I don’t bite.’

- b. [SubP Que_{attributive}] ... [FinP ~~que~~_{attributive}] [IP no muerdes?] (Sp.)
 QUE QUE not bite.2S
 ‘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 in the head of this phrase (cf. (143)).

- (143) [FocP Que chato que_{attributive}] ... [FinP ~~que~~_{attributive}] [IP é.] (Pt.)
 how annoying QUE QUE are.3S
 ‘How annoying this is.’

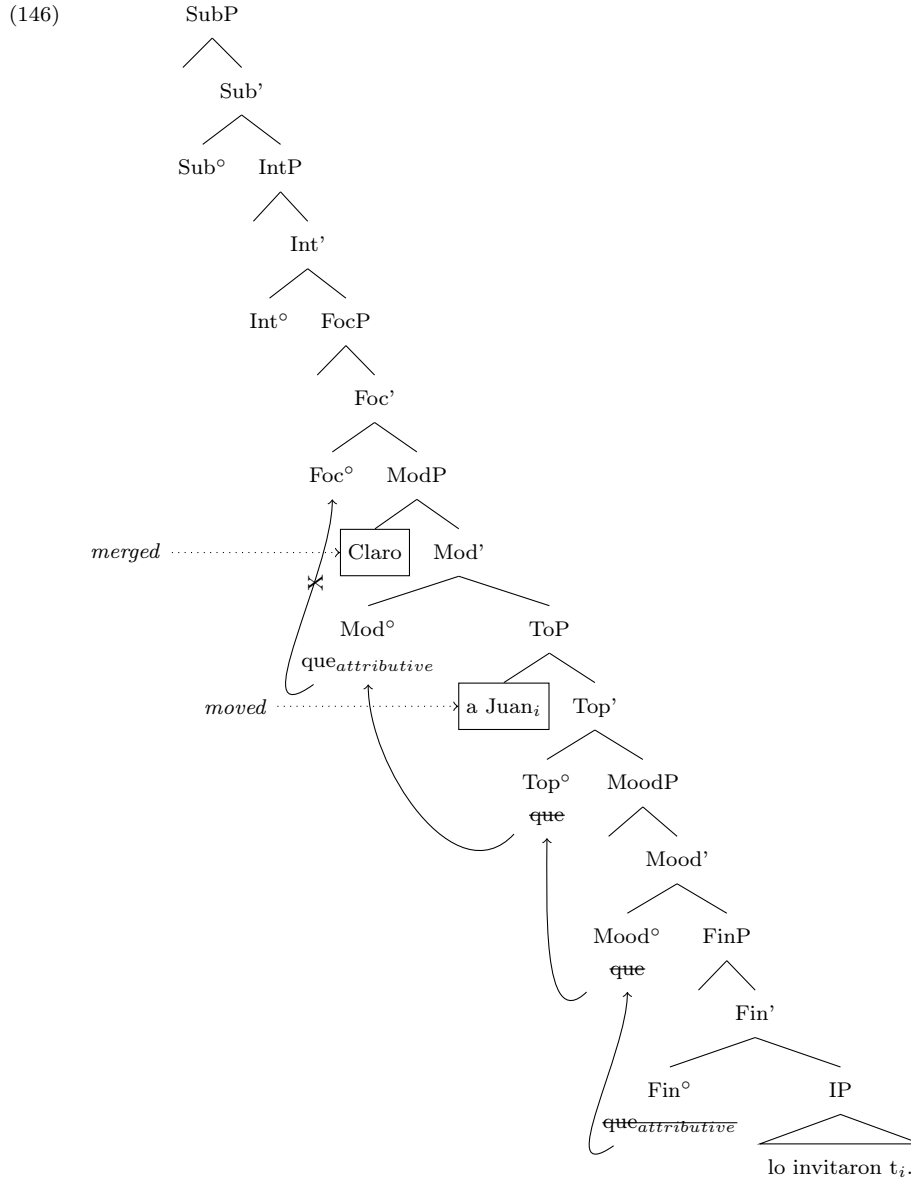
Similarly, epistemic and evidential modifiers in AdvC as well as the verum particle *sí* in AffC are assumed to be merged directly in the left-peripheral position in which they appear in the surface structure. The analyses I assume for these constructions are given in (144a) and (144b).

- (144) a. [ModP Segur que_{attributive}] ... [FinP ~~que~~_{attributive}] [IP son amics.] (Cat.)
 sure QUE QUE are.3P friends
 ‘Surely, they are friends.’
 b. [MoodP Sí que_{attributive}] ... [FinP ~~que~~_{attributive}] [IP son amics.] (Cat.)
 VERUM QUE QUE are.3P friends
 ‘They ARE friends.’

To repeat the main idea, I assume that in the examples in (143)-(144), intervening material blocks the movement of the complementizer to the next phrase, which results in the surface word order we observe in each of the constructions. The example in (145) is more complex because in addition to the evidential modifier preceding *que*, there is also a clitic left-dislocated topic below the attributive complementizer.

- (145) Claro que a Juan_i lo_i invitaron. (Sp.)
 clear QUE to Juan CL.MS invited.3P
 ‘Clearly, they invited Juan.’

The analysis for examples like (145) is illustrated in the structure in (146). The complementizer starts out in Fin° 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 section 3.1. In table 3, I summarize the main aspects of my own account and in table 4 the main aspects of the central accounts from the literature.

In my analysis, *que* is treated as a complementizer. In the present thesis, an Ibero-Romance complementizer is considered an underspecified element of the form *que* that occupies a left-peripheral head position and acquires its functional interpretation from a feature in the location in which it is merged (see section 1.2). This allows us to maintain that in the current constructions we are dealing with the same lexical item as in *que*-initial reportatives.

	<i>que</i> -initial declaratives	polar question	wh- exclamatives	AdvC	AffC
nature of <i>que</i>	complementizer				
location	FinP... >SubP	FinP... >SubP	FinP... >FocP	FinP>... >ModP	FinP>... >MoodP
interpretation	attributive feature in FinP				

Table 3: Attributive *que* in the present analysis.

	<i>que</i> -initial declaratives	polar questions
author	Corr (2016)	Prieto & Rigau (2007)
nature of <i>que</i>	illocutionary complementizer	complementizer
location	exclamative: EvalP>SA _{low} P	conjunctive: DeclP>EvidP> EvalP>SA _{high} P
interpretation	features in UP	operator in Force
	wh-exclamatives	AdvC
author	Demonte & Fernández Soriano (2014)	Cruschina & Remberger (2018)
nature of <i>que</i>	complementizer	complementizer
location	FinP	ForceP
interpretation	none	attributed to active performative structure

Table 4: Attributive *que* in the previous analyses.

The different functions and syntactic behaviors follow from the assumption that the complementizer is merged in different syntactic projections where it receives different 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*, as it is already merged in the highest position of the functional field. Most 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 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 distinctness is expressed in terms of different merge positions. Notably,

this is very close to the theory I put forward in the present study.

In my analysis, the location of the complementizer in the different constructions is derived straightforwardly by head movement that is conditioned 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 sentence initial *que* 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 U(terrance)P. As pointed out in section 3.1.1, the motivation for this movement is mainly theoretical; while my analysis in which *que* reaches SubP also captures the empirical data. Prieto & Rigau (2007), along the same lines as in my own analysis, propose that *que* is merged in FinP, but they do not assume further movement. I showed in section 3.1.2, that the analysis provided in Prieto & Rigau (2007) fails to account for data, that my analysis does predict by assuming that the complementizer reaches SubP in this case as well. In the account proposed for wh-exclamatives in Demonte & Fernández Soriano (2009), *que* is also analyzed in FinP. The empirical facts again pose some problems for this account, but support my own assumption that the complementizer movement stops in the head of FocP, adjacent to the wh-expression (see section 3.1.3). The analysis 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 sections 3.1.4 and 3.1.5). Again, my own analysis makes the correct predictions with respect to these data.

Finally, the backbone of my analysis is the assumption that in all the constructions under investigation here, 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 detailed characterization of the interpretation of the constructions is given in section 3.4. The basic idea is that the attributive complementizer attributes a commitment to the proposition to the hearer. While 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, as different features in the syntactic structure. These 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 construc-

tions, 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 *que* is merged in guarantees a factive interpretation of the content of the wh-exclamatives. However, the feature responsible for this interpretation is also present and checked in her derivation of *que*-less wh-exclamatives. As a result, I conclude that there is also no particular interpretive function attributed 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 give rise to the interpretation. In Prieto & Rigau (2007), while *que* itself is not equipped with interpretive features and is deemed optional, an operator in ForceP expressed through prosodic means is assumed to be responsible for 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. I conclude that they mean that the interpretive effect is more strongly connected to the modifier, merged directly in the CP-external structure, than to the complementizer itself analyzed in ForceP.

In the following section 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 the analysis are confirmed by the word order we observe.

3.3 Syntactic Properties

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

The first two assumptions are based on theoretical and empirical considerations. A central observation 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. section 3.4). Since the complementizer surfaces in different positions in the constructions, we are left with 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 thesis, I argue in favor of the second option, which is a better fit in light of the general ideas proposed here.

The following sections illustrate 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 displayed by the constructions and show how these can be accounted for in the present analysis. In section 3.3.1, I focus on the syntactic properties of attributive *que*-initial declaratives and in section 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 section 3.3.3, I focus on *wh*-exclamatives, in section 3.3.4 on *que* following epistemic and evidential modifiers (AdvC) and in section 3.3.5 on *que* in verum sentences (AffC).

3.3.1 *Que*-initial Declaratives

Since the basic idea is that the movement of attributive *que* is only blocked by externally merged material in the specifier of a projection, one prediction that follows is that if nothing intervenes the complementizer can reach the highest projection of the functional field, i.e. SubP. My proposal is that this is precisely where it surfaces in attributive *que*-initial declaratives. As a consequence, on a superficial level, these *que*-initial sentences mirror *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 section 3.4. Examples (147) and (148) illustrate the difference between *que*-initial reportatives and attributive *que*-declarative.

- (147) Pare: És dolent demanar a un fill que llegeixi un llibre? (Cat.)
 father is bad demand of a son that read.SUBJ a book

Mare: Que té nou anys.
 mother QUE has nine years

‘Father: Is it a bad thing to ask your son to read a book? Mother: He’s only nine!’ (ebook-cat)

(148) Mare: Té nou anys. (Cat.)
mother has nine years

Pare: Eh?
father huh

Mare: Que té nou anys.
mother QUE has nine years

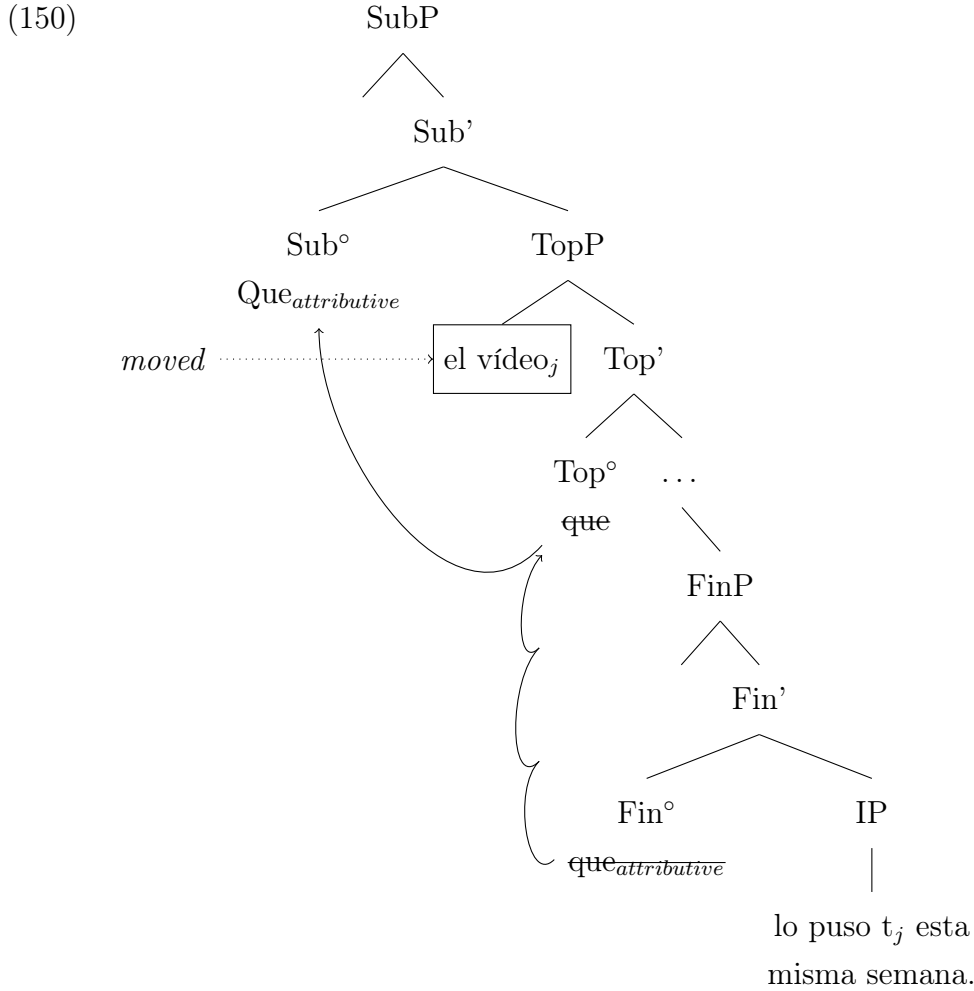
‘Mother: S/he is nine years old. Father: Huh? Mother: [reportative:] S/he is nine years old.’

In (147) 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 possibly communicates that she does not consider the books that the father gave to the son to be age appropriate. In (148), 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 the sentence as a reported version of the previous statement.

The proposed analysis of attributive *que* in SubP is not only based on theoretical considerations. There are also some empirical facts that lead to the same conclusion. First, attributive *que* precedes left-dislocated topics (cf. (149a)). This is expected because moved phrases, as these topics are, do not constitute an obstacle for the head movement of the attributive complementizer.

- (149) a. ¡Pues has sido muy rápida! [SubP Que] [TopP el vídeo_i] lo_i puso
well have.2S been very fast QUE the video CL.MS put
esta misma semana, jajaja. (Sp.)
this same week hahaha
‘Well, you’ve been very fast! The video, I just put it up this same week, haha.’ (CdE)
- b. Chicas, si son fashionistas, abran la mente! [SubP Que] [TopP
girls if are.3P fashionistas open.3P the mind QUE the fashion
la moda_i] la_i hacen todos! (Sp.)
CL.FS make.3P everyone
‘Girls, if you consider yourselves fashionistas, open your mind! Fashion is made by everyone.’ (CdE)

A schematic derivation of the example in (149a) is given in (150). For ease of exposition, the intermediate projections through which the complementizer moves are not displayed.



A further piece of evidence in favor of the assumption that attributive *que* reaches the highest position in the left periphery is that it appears to be impossible to embed attributive *que*-initial declaratives (cf. also Corr (2016)), hence the ungrammaticality of (151).³⁴

³⁴There are different theoretical alternatives to account for the fact that a sequence of two *ques* is disallowed in these contexts. One option is to assume that one of the complementizers, including its value, is deleted at PF. Most likely, the deleted complementizer is the one with the attributive feature since the subordinate feature needs to be visible in syntactically selected sentences (cf. section 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.

-
- (151) La mare va dir que (*que) té nou anys. (Cat.)
 the mother PST say that QUE has nine years

Notably, it is not because attributive *que* is disallowed in embedded contexts, as can be seen in (152) where AdvC is embedded (see also sections 3.3.3, 3.3.4 and 3.3.5 where further examples illustrating embedded attributive *que* constructions are discussed).

- (152) Só queria dizer que obviamente que, por falta de informação, podemos,
 just wanted.1S say that obviously QUE for lack of information can.1P
 por vezes fazer dietas mais tontas. (Pt.)
 at times do diets more stupid
 ‘I just wanted to say that obviously, for lack of information, at times we can
 end up doing stupider diets.’ (CdP)

Attributive *que* can furthermore introduce the answer particles Cat./Sp. *sí*, Pt. *sim* and Cat./Sp. *no*, Pt. *não* (the examples in (153)).

- (153) a. – Solfieri, não é um conto, isso tudo? – Pelo inferno, que não! (Pt.)
 Solfieri not is a swindle this all for-the hell QUE no
 ‘– Solfieri, isn’t all this a swindle? — Hell no!’ (CdP)
- b. La multitud responde: Que sí, que sí. (Sp.)
 the crowd answers QUE yes QUE yes
 ‘The crowd answers: Yes. Yes.’ (CdE)
- c. CARLES: Home no, tampoc no cal. PEP: Que sí, home. Que
 Carles man no either not is necessary Pep QUE yes man QUE
 sí. (Cat.)
 yes (Cat.)
 ‘Carles: Man, don’t, it’s not worth it either. Pep: Yes it is, man. Yes it is.’
 (caWac)

The interpretive effect is that the affirmation or negation is emphatic, with particular stress on the fact that the hearer shares the commitment, which is in line with the general contribution of attributive *que* (cf. also section 3.4). Concerning their 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.

- (154) [_{SubP} Que_i] ...[_{FinP} t_i] [_{PolP} sí/sim/no/não.]

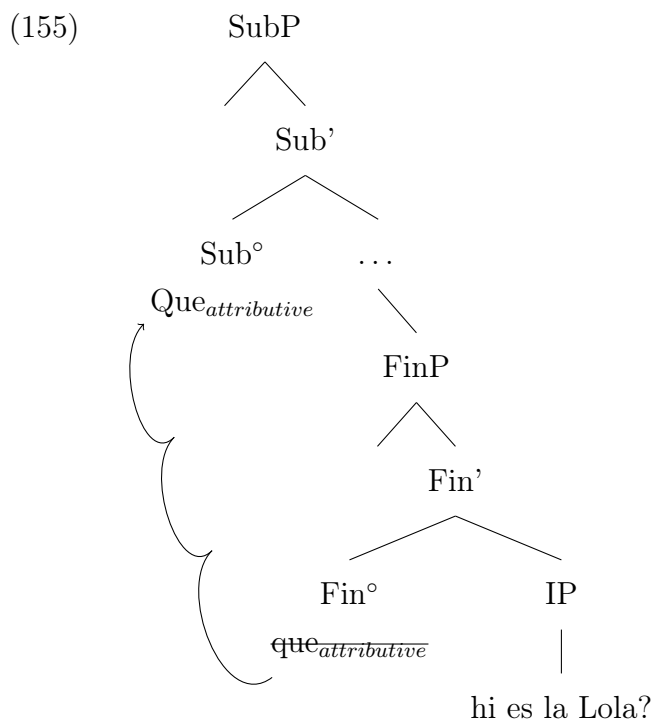
There are no strong theoretical consideration behind this terminological choice and I am open to adopting other notations such as Laka's ΣP used by Martins (2013), if they prove more adequate.

In my analysis given in (154), I assume that also in these cases attributive *que* moves to SubP. This should be taken as a tentative 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 (154) is that answer particles appear in a structure that projects a left periphery. One prediction that follows is that we expect the answer particles to be preceded by left-peripheral material. Additionally, if attributive *que* is present, we expect it to be subject to the same restrictions as we observe in full sentences. That this is indeed the case will be demonstrated in sections 3.3.4 and 3.3.5, where I will show that the answer particles are also compatible with AdvC and AffC.

Initial attributive *que*-declaratives 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 contrasts arises with respect to the occurrence of attributive *que* in polar questions, which is covered in the following section 3.3.2. This is a very productive construction in Catalan, while it is potentially restricted in Spanish and virtually unattested in Portuguese.

3.3.2 *Que*-initial Polar Questions

For the structure I assume for *que* in polar questions, a key point in my analysis is once again that the complementizer moves from head to head 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 it is argued to occupy in the attributive *que*-declaratives discussed in the previous section 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 section 3.2 without requiring further stipulations. The relevant points are shown in (155), where again the intermediate positions through which *que* passes are omitted in the structure.



I need to stress 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 for the current 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.

The syntactic properties of polar questions with and without *que* show no apparent discrepancies. This makes it plausible to conclude that *que* is not present for syntactic reasons. My proposal, on the contrary, is that attributive *que* has an impact on the interpretation of the sentence it introduces just as in declaratives. The effect of the presence of the complementizer in a polar question is that it is not perceived as neutral; instead, the speaker is biased toward a positive answer (cf. also Kocher 2017). *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”.

- (156) La Caterina va entrar i va córrer cap al lavabo amb el
the Caterina PST enter and PST run in direction to the bathroom with the
paraigua que regalimava. – Que plou? (Cat.)
umbrella that dripped QUE rains
‘Catarina entered and ran to bathroom with a dripping umbrella. – Is it rain-
ing?’ (ebook-cat)

This can be seen in example (156). The idea is that in this case 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 section 3.4, where I offer an explanation that allows a unified vision of the discourse contribution across all constructions involving attributive *que*. I propose that the seemingly different effect in polar question results 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 Mito 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.

- (157) B: Son cinco bloques, con pista de tenis colectiva ¿no? Entonces, ahí
are five blocks of flats with tennis court collective no then there
jugamos. A: ¿Que tenéis apartamento no? B: No, es un piso en un
play QUE have apartment no no is a flat in a
bloque. (Madrileño Sp.)
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 (157) can be characterized in a similar way to the Catalan example in (156). 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 towards 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-final tag *no*.

Examples such as (157) therefore suggest that attributive *que* is also licensed in Spanish polar questions. In the following, however, I will rely mostly on Catalan examples to illustrate the core syntactic properties and will then 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 (158a) when *que* is interpreted as attributive.

- (158) a. A: M' he d' anar, tinc un camí llarg a casa. (Cat.)
 CL.1S have.1S to leave have a way long to home
 B: (*Que_{attributive}) on vius?
 QUE where live..2S
 '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 reparative interpretation (cf. (159)).

- (159) A: On vius? (Cat.)
 where live.2S
 B: Què has dit?
 what have.2S said
 A: Que_{subordinate} on vius.
 QUE where live.2S
 'A: Where do you live? B: What did you say? A: [reparative:] 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. (160)). 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.

- (160) B: M' he trobat amb la teva directora de tesi l' altre dia i
 CL.1S have.1S met with the your thesis supervisor the other day and
 m' ha preguntat per tu. (Cat.)
 CL.1S has asked about you

A: Què t' ha preguntat?
what CL.2S has asked

B: Que si has acabat la tesi.
QUE if have finished 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 (161). 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.

(161) A: Fem un cafe la setmana que ve? Ara tinc temps. (Cat.)
make.1P a coffee the week that comes now have.1S time

B: Que has acabat la tesi?
QUE have.2S finished 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 regards to the word order, the analysis in (155) predicts that the complementizer should be followed rather than preceded by left-peripheral material. Evidence for the high position of *que* in polar questions is provided by data like (162) (repeated from (121)), which show that the complementizer precedes a clitic left-dislocated topic.

(162) [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 comprar tu?]
QUE the oranges CL.FP PST buy you
(Cat.)

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

As shown in section 3.1, a topic can precede *que* in polar questions. However, these topics exhibit properties that are 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 in the example illustrated in (163).

- (163) [Context: Marta and Maria are at a party. Marta sees that Maria's college 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:]

[_{αP} En Jordi_i,] [_{SubP} que] [_{TopP} aquest idiota_i ... [_{IP} l_i' has convidat
the Jordi QUE that idiot CL.MS have.2S invited
tu?] (Cat.)
you
'Did you invite that idiot Jordi?'

Additionally, according to my informants, it is not sufficient to have a co-referential epithet or DP but the example is 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 (163) do not constitute counter-evidence for my current analysis of *que* in polar questions. My tentative analysis sketched in (163) is that *en Jordi* is directly merged in a CP-external position and that *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) coincide with a confirmatory reading of the question. I briefly return to their pragmatic function in section 3.4.

- (164) Oi que ens entenem? (Cat.)
OI QUE CL.2P understand.2P
'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 the present thesis. 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 (165), 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 suggesting that they are located in the specifier and head of the same projection.

- (165) *Oi en Jordi que l’has convidat tu?
 OI the Jordi QUE CL.MS have.2S invited tu
 ‘intended: You are the one who invited Jordi, right?’

The example in (165) is consistent with the analysis that follows from the general assumption presented in this chapter. Just as in *que*-initial polar questions without an additional pragmatic marker, 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 (166), which shows that it precedes a clitic left-dislocated topic.

- (166) [SubP Oi que_i] [TopP aquesta pregunta_j t_i] ... [FinP t_i] [IP no se l_j’
 OI QUE this question not CL.REFL CL.FS
 havien fet mai?] (Cat.)
 have.3P made never
 ‘They never asked themselves this question, right?’ (caWac)

An additional piece of evidence for *que* reaching a high position in polar questions is that, just as in the case of *que*-initial declaratives discussed in section 3.3.1, *que*-initial polar questions are not found in embedded contexts (cf. (167a), (167b)). This is predicted by the present analysis because the two instances of *que* would compete for the same projection, SubP (but see footnote 34 for some suggestions of how to account for this theoretically).

- (167) a. La Maria va preguntar que (si) (*que) l’ he convidat jo.
 the Maria PST ask that whether QUE CL have.1S invited I
 b. La Maria va preguntar que (si) (*oi que) hi es la Lola.
 the Maria PST ask that whether OI QUE CL.LOC is the Lola

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 mostly disregarded. The review of corpus data, however, shows that Spanish does have *que*-initial polar questions. Moreover, they do appear to have the same function, that of expressing that the speaker is biased towards a positive answer (cf. (157)). Another example that illustrates this is given in (168).

-
- (168) – Creo que no se le pagan 15 millones mensuales para que
 believe that not CL.REFL CL.DAT pay.3P 15 millions monthly so that
 ande haciendo proselitismo político junto a Horst Golborne. – ¿Qué? –
 goes doing proselytism politic together to Horst Golborne what
 ¿Que no se llama Horst? Ah perdón, me confundí... (Chilean Sp.)
 QUE not CL.REFL calls Horst ah sorry CL.1S confused
 ‘– I think that he doesn’t get 15 millions 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. (157) from Madrileño and (168) from Chilean Spanish).

The example in (169), 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 also places the complementizer in SubP in the Spanish examples.

- (169) [SubP ¿Que_i] [TopP la respuesta_j t_i] [FinP t_i [IP la_i publicamos en
 QUE the answer CL.FS publish.1P in
 periódicos de provincias? (US Sp.)
 newspapers of provinces
 ‘Are we publishing the answer in provincial newspapers?’ (CdE)

These data suggest that there is no evidence for a systematic syntactic difference between Spanish and Catalan with respect to *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 5). 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 ori-

gin. 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 the 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 of the subcorpus: 20.4 million) of *que*-initial polar question in the subcorpus. In comparison, the Catalan caWac contains 3124 occurrences (total tokens of 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, 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 at this point. What the data however once again confirm is that *que*-initial polar questions are unlikely to result from contact with Catalan, since the attestations also stem from various non-European varieties.

	fiction	press	oral	total
Argentina	21	1	7	29
Bolivia	0	0	4	4
Chile	19	0	5	24
Columbia	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
NA	20	0	0	20
total	126	15	124	180

Table 5: Absolute numbers of *que*-initial polar questions in the 2001 contemporary subcorpus of CdE. (NA=not available)

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.

- (170) a. A Joana pergunta (*que) se vens. (Pt.)
 the Joana asks that if come.2S
- b. Juana pregunta que si vienes. (Sp.)
 Juana asks that if come.2S
- c. La Joana pregunta que si vens. (Cat.)
 the Joana asks that if come.2S
 ‘Jo/uana asks if you are coming.’

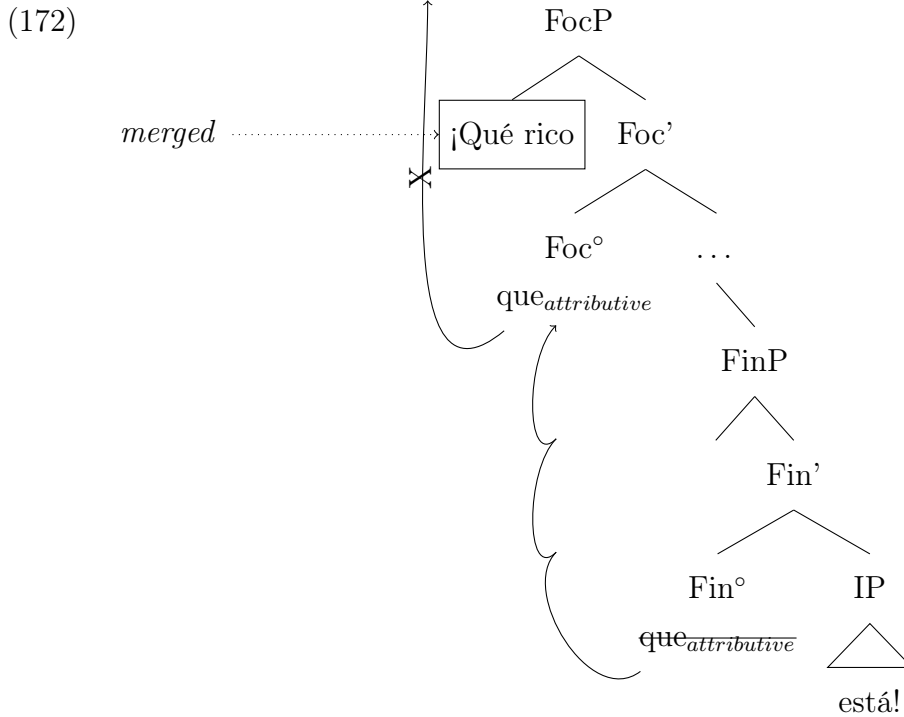
Corr (2016) offers an explanation for this discrepancy. The author hypothesizes that Portuguese *que* is more restricted than the cognate item in Spanish and Catalan in that it is specialized for declaratives and cannot appear in other clause types. One of the reasons behind this hypothesis is the ungrammaticality of *que* in embedded and reported questions above the interrogative complementizer *se* in Portuguese (see (170a)) that contrasts with the other two languages (see (170b), (170c); cf. section 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, which is 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 (171) (repeated from (126)) is given in (172).

- (171) ¡Qué rico que está! (Sp.)
 how good QUE is
 ‘How good this is!’ (Demonte & Fernández Soriano, 2009, 33 : ex 19a)

Again, the intermediate positions through which the complementizer passes, are not shown 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.



One main argument for this analysis, which assumes that the *wh*-expression is merged in rather than moved to the left periphery, is that it is in keeping with the idea that the movement of the 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. One crucial aspect to be explained given this analysis is how the *wh*-expression ends up being interpreted as an argument or adjunct dependent on the IP-internal verb. Non-local theta- and even case-assignment is not an unprecedented assumption 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 (173), 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.

- (173) Susi dice que a los alumnos, que les vana dar regalos. (Sp.)
 Susi says that DAT the students that CL.3P go to give presents
 ‘Susi says that they are going to give the students presents.’
 (Villa-García 2015[18: ex 23])

In order to explain how these topics are assigned their case, Villa-García (2015) adopts the agreement mechanism of 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.³⁵ 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 (174b), the left-peripheral wh-expressions cannot bear case marking when followed by attributive *que*.

- (174) a. Han dado un premio a una estudiante inteligente.
 have.3P given the prize DAT a student intelligent
 ‘They gave the prize to an intelligent student.’
 b. *¡A qué estudiante más inteligente que han dado el premio!
 DAT which student more intelligent QUE have.3P given the prize
 c. *¡Qué estudiante más inteligente que han dado el premio!
 which student more intelligent QUE have.3P given the prize
 intended: ‘What an intelligent student they gave the prize to!’

In the preajacent, i.e. the declarative equivalent to the wh-exclamative in (174a), the object of the verb is introduced by the dative marker *a*. The corresponding wh-exclamative in (174b), in which the dative object is a left-peripheral wh-expression, is judged ungrammatical by my informants. Furthermore, example (174c) shows that an unmarked wh-expression is not grammatical either.³⁶

These data contrast with the example in (175), which shows that in the absence of attributive *que* a dative marked object can be a left-peripheral wh-expression.

³⁵In 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.

³⁶There appears to be a small degree of variation involved, as a minority of my informants judged examples like (174b) marginally acceptable. Additionally, some of the informants found (175) only grammatical with a dative clitic co-referent to the wh-expression.

- (175) ¡A qué estudiante más inteligente (le) han dado el premio!
 DAT which student more intelligent CL.DAT have.3P given the prize
 ‘What an intelligent student they gave the prize to!’

This example constitutes a crucial motivation for the idea, I will return to below, which is 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 case, 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.

- (176) a. Que coisa mais idiota que fazem aos animais. (Pt.)
 what thing more stupid QUE do.3P to-the animals
 ‘What stupid things people do to animals.’ (CdE)
 b. Que malament que anem. (Cat.)
 how badly QUE go.2P
 ‘How badly it is going for us.’ (caWac)

The potentially more complex are cases are the ones in which the wh-expression appears to be an argument of the verb, as in (176a). In line with Villa-García (2015), I adopt the idea that for theta-role assignment, clausemate hood 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 current argumentation 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 (176a). 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 (i.e. c-commanded by *que coisa mais idiota*) to the thematic head.

An interpretation of wh-exclamatives such as (176b), in which the wh-expression appears to be a modifier of a verb can also be provided under this account. 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 correspondent 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 could be accounted for, I will now evaluate the predictions of the analysis outlined in (172). 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 (177). In the analysis, this is captured by postulating that they appear in the specifier and head of the same projection.

- (177) *Que estrany al Jordi_i que li_i hagin trucat. (Cat.)
 how strange DAT-the Jordi QUE CL.DAT have.SUBJ.3P called
 (adapted from Kocher 2017, 48: ex 95a)

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 examples (178a) and (178b).

- (178) a. [_{FocP} Que estrany que_i] [_{TopP} al Jordi_j t_i] [_{FinP} t_j] [_{IP} li_j
 how strange QUE DAT-the Jordi CL.DAT
 hagin trucat. (Cat.)
 have.SUBJ.3P called
 ‘How strange that they called Jordi.’
 (adapted from Kocher 2017, 48: ex 95b)

- b. $[\text{TopP Al} \quad \text{Jordi}_i]$ $[\text{FocP que estrany que}_j]$ $[\text{FinP } t_j]$ $[\text{IP li}_i]$
 DAT-the Jordi how strange QUE CL.DAT
 hagin trucat. (Cat.)
 have.SUBJ.3P called
 ‘How strange that they called Jordi.’
 (adapted from Kocher 2017, 48: ex 95c)

Moreover, since the complementizer only moves up to FocP but not higher, we expect that wh-exclamatives with attributive *que* can appear in embedded contexts as well as in *que*-initial reportatives, both of which have the high complementizer merged in SubP and are valued with a subordinate feature, in my analysis. Examples (179a) and (179b) confirm that this is indeed the case.

- (179) a. $[\text{SubP María dijo } [\text{SubP que}_{\text{subordinate}}] [\text{FocP qué suerte que}_i] [\text{FinP } t_i] [\text{IP no}]]$
 María said that what luck QUE not
 salió de fin de semana.] (Sp.)
 went-out.3S for weekend
 ‘María said how lucky that she didn’t go out on the weekend.’ (CdE)
- b. $[\text{SubP Que}_{\text{subordinate}}] [\text{FocP que bom que}_i] [\text{FinP } t_i] [\text{IP você vai estar}]]$
 QUE how good QUE you will be-there
 no dia 1 de setembro.] (Pt.)
 on the day 1 of September
 ‘[reportative:] How good that you will be there on September 1st.’ (CdP)

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 the present chapter in which 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 that disallows movement across material that is merged in the left periphery. I will now show that this assumption is not just a theoretical necessity, but 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.³⁷ The existence of parallel examples like (180a) and (180b) lead

³⁷There might be some cross-linguistic variation with respect to the possibility of omitting the

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).

- (180) a. Quin nas més gros que tens. (Cat.)
 what nose more big QUE have.2S
- b. Quin nas més gros tens. (Cat.)
 what nose more big have.2S
 ‘What a big nose you have.’

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 the facts support the assumption 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. My idea is that the two examples in (180a) and (180b) are not structurally equivalent and furthermore I believe that *que* only appears to be optional, when in fact this is not the case. Evidence in favor of this assumption is the contrast between (181a) and (181b) repeated from (174b) and (175). The data show 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.

- (181) a. *¡A qué estudiante más inteligente que han dado el premio!
 DAT which student more intelligent QUE have.3P given the prize
 intended: ‘What an intelligent student they gave the prize to!’
- b. ¡A qué estudiante más inteligente (le) han dado el premio!
 DAT which student more intelligent CL.DAT have.3P given 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 (182), where crucially, the complementizer cannot be omitted (cf. also González i Planas (2010)).

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.

- (182) a. Quina sort *(que) neva. (Cat.)
 what luck QUE snows
 ‘What luck that it is snowing!’
 b. Que mal *(que) isso passou. (Pt.)
 how bad QUE this happened
 ‘How bad that this happened.’
 c. Qué lástima *(que) no vienes. (Sp.)
 what pity QUE not come.2S
 ‘What a pity that you’re not coming.’

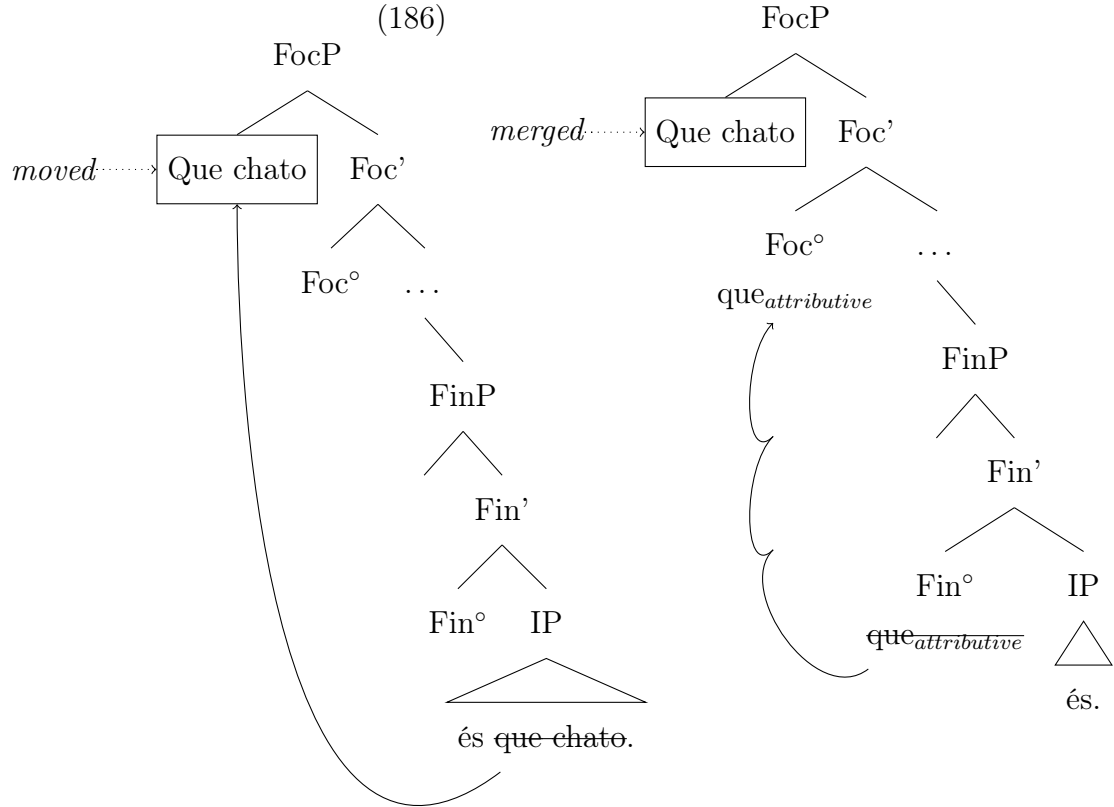
The difference between the examples in (180) and in (182) is that in the former the *wh*-expression can be reconstructed in the core sentence, while in the latter it cannot. More examples in which the *wh*-expression appears to be syntactically dependent on the verb are given below. It is a predicate depending on a copula verb in (183), an argument in (180) and an adjunct in (184). On a side note, a contrast seems to be also present in English between *wh*-exclamatives where the *wh*-expressions 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 appear to have a function in the IP (for instance (182) vs. (183a)), allows the presence of a complementizer below the *wh*-expression.

- (183) a. Que chato que és. (Pt.)
 how annoying QUE are.2S
 b. Que chato és.
 how annoying are.2S
 ‘How annoying you are.’
 (184) a. Qué bien que nos ha ido. (Sp.)
 how well QUE CL.1P has gone
 b. Qué bien nos ha ido. (Sp.)
 how well CL.1P has gone
 ‘How well it went for us.’

In sum, while parallel examples like those illustrated in (180), (183), (184) could lead to the conclusion that a derivation by movement is a tenable analysis, examples such as those in (182), 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. My proposal is 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 those in which the wh-expression is followed by *que*, the wh-expression is merged in the left periphery (cf. (186)), while in those without *que*, the wh-expression is promoted to the left periphery through movement from an IP-internal position (cf. (185)).

(185)



The idea is therefore that examples like (180a), (183a), (184a) and (182) have the same underlying structure in the CP, while examples like (180b), (183b) and (184b) differ structurally. Further motivation behind the assumption that there are distinct derivations comes from the position of subjects. The complementizer-less wh-exclamatives do not allow a preverbal subject: In all three languages the subject must necessarily follow the verb (see (187)). Castroviejo (2006) takes this as evidence 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.

- (187) a. Que blanca i bonica es presenta des d'allí la ciutat. (Cat.)
 how white and pretty CL.REFL presents from there the city
 'How white and pretty the city looks from there.' (caWac)
- b. Qué suerte tenemos los lectores de tener quien nos escriba.
 what luck have.1P the readers to have who CL.1P write.SUBJ.3S
 (Sp.)
 'What luck we readers have that someone writes to us.' (CdE)
- c. Que sorte tive eu do ter conhecido. (Pt.)
 what luck had.1S I to-CL.MS have known
 'What luck I had to have known him.' (CdP)

In contrast, Catalan, Portuguese and Spanish all allow preverbal subjects in wh-exclamatives with an attributive complementizer (as in (188)). Crucially, this is the case irrespective of whether the wh-expression appears to have a function in the IP as in (188b) or when it does not as in (188a), (188c).³⁸

- (188) a. Que trist que tu no sàpigues ni escriure ni expressar
 how sad QUE you not know.SUBJ.PST.2S neither write nor express
 -te correctament. (Cat.)
 CL.2S correctly
 'How sad that you don't know neither how to write nor how to express
 yourself correctly.' (caWac)
- b. Que parva que eu sou! (Pt.)
 what fool QUE I am
 'What a fool I am.' (CdP)

³⁸In the examples in (188a) and (188c), 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 great deal. For instance, example (188c) contrasts with (182c), which contains the same evaluative term, *lástima* 'pity', a verb inflected for indicative. As subjunctive is most commonly encountered in complement clauses, one might be lead to the conclusion that we are dealing with elliptical structures ([_{CP1} Qué lástima ~~(es)~~ [_{CP1} que la vida sea tan corta.]]). However, as can be seen in the example in (VI), subjunctive is not restricted to complement clauses but also appears in main clauses. This why, for now, I stick to the idea that, despite the subjunctive mood, also these wh-exclamatives can be analyzed as main clauses.

- (VI) Ojalá nos dejen. (Sp.)
 hopefully CL.1P leave.SUBJ.3P
 'Hopefully they will leave us alone.' (CdE)

-
- c. Qué lástima que la vida sea tan corta y las dictaduras tan largas. (Sp.)
 What pity QUE the life is.SUBJ so short and the dictatorships so long
 ‘What a pity that life is so short and dictatorships are so long.’ (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 (189)).

- (189) Hala, que que ben (*que) fala a irlandesa! (Galician)
 wow QUE what well QUE speaks the Irish
 ‘Wow, the Irish girl speaks so well!’ (Corr 2016, 90: ex 24)

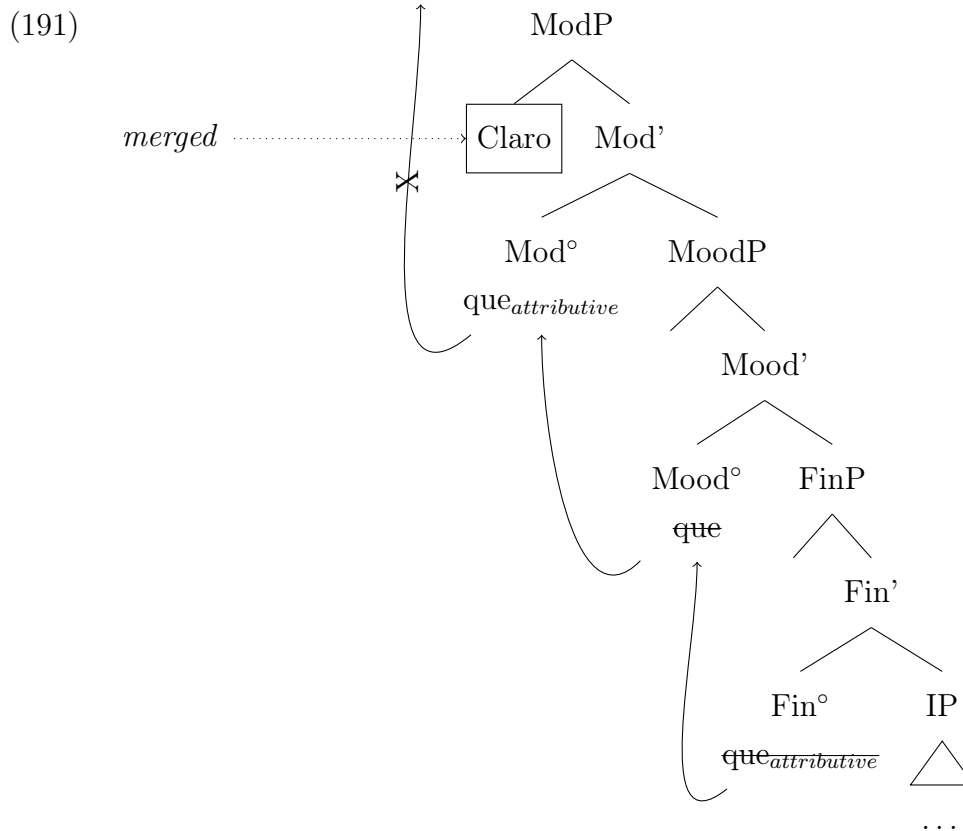
- (190) [_{SubP} Que_{j,attributive}] ... [_{FocP} que ben_i t_j] ... [_{FinP} t_j]_{[IP} fala t_i a irlandesa!]

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 tentative analysis is given in (190). 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, the data discussed here support my assumption 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 thesis, 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

The present 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 adopt 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.



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.³⁹ Based on the universal word order restrictions of co-occurring modifiers described in Cinque (1999), Giorgi (2010) and van Gelderen (2011) actually propose that ModP consists of three sub-projections dedicated to evaluative, evidential and epistemic meaning. 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. (192), (193)). This is in line with what the proposal in Giorgi (2010) and van Gelderen (2011) would suggest.

³⁹The proposal that sentence adverbs are merged in the CP is different from the analysis in 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.

-
- (192) [_{ModP} Felizmente que] na Suíça se faz imensas coisas
 fortunately QUE in the Switzerland CL.REFL make immense things
 sem referendos. (Pt.)
 without referendums
 ‘Fortunately there are immense things happening in Switzerland without a referendum.’ (CdP)
- (193) [_{ModP} Lamentavelmente que] no va a ocorrer. (Sp.)
 lamentably QUE not go to occur
 ‘Lamentably, it won’t occur.’ (CdE)

The idea introduced above allows the adoption of a strictly compositional account, in which the modifiers are merged in the same place irrespective of whether *que* is present or not.

The assumption that epistemic, evidential (and evaluative) modifiers are merged in rather than moved to the left periphery is able to explain why attributive *que* follows precisely these types of modifiers but precedes other types. This can be seen in (194) in which the complementizer follows the evidential modifiers *obviamente* and *claro* but precedes *imediatamente* and *rapidamente*. This word order is expected in the present proposal, because the lower adverbs are preposed from an IP internal position. As moved elements, they do not constitute an obstacle to complementizer movement.

- (194) a. [_{ModP} Obviamente que] imediatamente le surgirá la duda. (Sp.)
 obviously QUE immediately him arise.FUT the doubt
 ‘Obviously, doubt will arise immediately.’ (CdE)
- b. [_{ModP} Claro que] rapidamente deixa de ser confortável fazer chamadas
 clear QUE rapidly stop be comfortable make calls
 com o tablet ao ouvido. (Pt.)
 with the table at-the ear
 ‘Clearly, it quickly stops being comfortable to make calls with a tablet at the ear.’ (CdP)

As mentioned before in section 3.3.1, epistemic and evidential modifiers can also appear in short emphatic affirmations and negations introduced by attributive *que* (cf. (153)). As expected, the modifiers precede the complementizer. My tentative analysis relies once again on a polarity projection PolP sandwiched between FinP and IP.

- (195) a. M' havia venjat jo mai de ningú? [_{ModP} Segur que_i] [_{FinP} t_i] [_{PolP} sí.]
 CL.1S had avenge I never of no one sure QUE yes
 (Cat.)
 ‘Had I never sought revenge from anyone? Sure I did.’ (ebook-cat)
- b. Me preguntas si veo todo eso en tus ojos. [_{ModP} Naturalmente
 CL.1S ask.2S if see.1S all this in your eyes naturally
 que_i] [_{FinP} t_i] [_{PolP} sí], amada mía. (Sp.)
 QUE yes beloved mine
 ‘You ask whether I see all of this in your eyes. Naturally I do, my beloved.’
 (CdE)
- c. Mas isso não significa que o veganismo é uma política a se
 but this not mean that the veganism is a politics to CL.REFL
 abandonar. [_{ModP} Claro que_i] [_{FinP} t_i] [_{PolP} não.] (Pt.)
 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 they should not allow any left-peripheral material to intervene between them. The ungrammaticality of the intervening topic in (196a) shows that this prediction holds true. A further prediction is that AdvC can be followed and preceded by a topic since in the present analysis it is merged in neither the lowest nor highest position of the left periphery. The grammaticality of (196b) with a preceding topic and (196c) with a topic following the complementizer show that this is indeed the case.

- (196) a. [_{ModP} *Segur] [_{TopP} aquest llibre_i] que l'_i ha llegit. (Cat.)
 sure this book QUE CL.MS has read
 ‘Surely, he indeed has read this book.’
- b. [_{TopP} Aquest llibre_i] [_{ModP} segur que] l'_i ha llegit. (Cat.)
 this book sure QUE CL.MS has read
 ‘Surely, he indeed has read this book.’
- c. [_{ModP} Segur que] [_{TopP} aquest llibre_i] l'_i ha llegit. (Cat.)
 sure QUE this book CL.MS has read
 ‘Surely, he indeed has read this book.’

Furthermore, AdvC appears in embedded contexts. In (197a), it appears in the com-

plement of a verb of saying, in (197b), it appears in an appositive relative clause and in (197c) in a restrictive relative clause. This behavior is also predicted by the analysis proposed for AdvC, and moreover constitutes strong evidence in favor of the assumption that there is more than one merge position accessible to the complementizer in Ibero-Romance.

- (197) a. Disse [_{SubP} que] [_{ModP} certamente que] iria ver logo resultados. (Pt.)
 said that certainly QUE go.COND see soon results
 ‘S/he said that certainly I would see results soon.’ (CdP)
- b. Otra canción [_{SubP} que] [_{ModP} claro que] escuchamos todos y que
 other song that claro QUE listen-to all and that
 podría parecer muy buena, es “Realmente no estoy tan solo”. (Sp.)
 could seem very good is Realmente no 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. És gent [_{SubP} que] [_{ModP} segur que] has vist pero mai has
 are people that sure QUE have.2S seen but never have.2S
 passat un cap de setmana amb ells. (Cat.)
 spent a weekend with them
 ‘These are people that you surely met but with whom you never spent a
 whole weekend.’ (caWac)

Moreover, AdvC can also appear in *que*-initial reportatives as illustrated in example (198) repeated from (57b) in section 2.3.1. This again supports the idea that the two instances of *que* have different functions and are merged in distinct positions.

- (198) [_{SubP} Que] [_{ModP} claro que_i] [_{FinP} t_i] le viene bien, que qué alegría, que
 QUE clear QUE CL.3S comes good QUE what joy QUE
 dónde. (Sp.)
 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, 2017, 2018a). One significant feature of AdvC is that it admits both underived modifiers (cf. (199a)), which formally coincide with adjectives, and derived modifiers (cf. (199b)), which take the adverbial derivational suffix *mente* (cf. also Cruschina & Remberger (2017a), Cruschina & Remberger (2018)).

- (199) a. Cierto que la culpa no es suya. (Sp.)
 certain QUE the fault not is his
 ‘Certainly, it’s not his fault.’ (CdE)
- b. Ciertamente que la culpa no es suya. (Sp.)
 certainly QUE the fault not is his

In Kocher (2014, 2018a), I provided empirical motivation for the claim that despite these distinct forms, we are dealing with one and the same underlying structure. 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 section 4.3). I propose that they are all sentential adverbs. This is not a controversial assumption for examples like (199b), but requires further justification for examples like (199a). In Kocher (2014, 2017), I showed that, contra Martín Zorraquino (1998), Freitas Barros (2006) and, Ocampo (2006), it is not tenable to assume that cases like (199a) 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, Hummel 2017). What is more, when the construction emerged in the 16th century (Kocher 2017), 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. As will be shown in section 3.4 and 4.3, the copula construction has a different interpretation than the AdvC construction, which is consistent regardless 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, we expect not only to find derived and underived adverbs, but also epistemic and evidential adverbials. The examples in (200a), (200b) and (200c) show that this prediction holds true.

- (200) a. Con certeza que crearemos productos muy interesantes en esta
 with certainty QUE create.FUT.2P products very interesting in this
 nueva etapa. (Sp.)
 new phase
 ‘Certainly, we will create very interesting products in this new phase.’ (CdE)

-
- b. Sem dúvida que nos Açores há espaço e público para este
without doubt QUE in the Azores there is space and audience for this
tipo de eventos! (Pt.)
type of events
'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. (Pt.)
of-certain QUE the conductor fell-asleep
'Certainly the conductor 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 types of modifiers would be expected, such as *fundamental* in (201a), which is found frequently in copula constructions. These are impossible, however, in AdvC (see (201b)).

- (201) a. É fundamental que a criança explore o seu mundo. (Pt.)
is fundamental that a child explores the its world
'It is fundamental that a child explores its world.' (CdP)
- b. *Fundamentalmente/ Fundamental que a criança explore o seu
fundamentally fundamental that a child explores the its
mundo. (Pt.)
world
'It is fundamental that a child explores its world.'

One additional argument is that the modifier in AdvC cannot be modified (compare (202a) and (202b); see also Cruschina & Remberger 2017a, 2018). This would once again not be expected if it were just a version of a copula construction.

- (202) a. É bem certo que as rivalidades se semeiam. (Pt.)
is good certain that the rivalries CL.REFL spread
'It is pretty certain that the rivalries are spreading.' (CdP)
- b. *Bem certamente/ certo que as rivalidades se semeiam. (Pt.)
good certainly certain that the rivalries CL.REFL spread

Furthermore, AdvC is incompatible with morphological negation (cf. (203a) and (203b)), 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).

- (203) a. És impossible que es belluguin tant sense la meva
 is impossible that CL.REFL move.SUBJ.3P so without the my
 voluntat. (Cat.)
 will
 ‘It is impossible that they move this much against my will.’ (ebook-cat)
- b. *Impossiblement que es belluguin tant sense la meva
 impossibly that CL.REFL fight.SUBJ.3P so without the my
 voluntat.
 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, 2017 and section 4.3).

To close this section, I provide a discussion of a development in Catalan that illustrates the productivity of AdvC. Recently, *esclar* (< *és clar*) seems to have emerged as a new evidential adverb (see section 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’. 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 2003.⁴⁰ Similarly, the Catalan evidential adverb *esclar* resulted from univerbation, originating from EsAdjC (cf. (204)).

- (204) [_{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 (205)).

- (205) a. Jo, és clar, en això no sé què aconsellar-te. (Cat.)
 I is clear in that not know.1S 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. (Cat.)
 I not know.1S ever the my clients nor they to me is clear
 ‘I never know my clients, nor they know me - that is clear.’ (caWac)

⁴⁰See also Alturo & Chodorowska (2009) on the grammaticalization of *sisplau*.

Typical features of grammaticalization can also be observed in the case of *esclar* such as phonological reduction ([eskla] > [əskla]) and syntactic reanalysis (clause > adverb).

- (206) a. Sola no hi he estat mai, esclar, però sé què vull dir.
 alone not there have.1S been ever clearly but know.1S what want.1S say
 (Cat.)
 ‘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
 not CL treat.3S clearly to deny the existence of a organization
 social. (Cat.)
 social
 ‘Evidently, it’s not about denying the existence of a social organization.’
 (caWac)

The examples in (206) illustrates the syntactic mobility of *esclar*, which parallels the syntactic mobility of epistemic and evidential adverbs.

- (207) a. ¿Com era possible que una obra mestra [...] tingués un final feliç
 how was possible that a masterpiece had a happy ending
 com aquell? -Esclar que per a Espert el suïcidi de la protagonista
 like that clearly QUE for E. the suicide of the protagonist
 també es pot entendre com un final feliç. (Cat.)
 also CL can.3S understand 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)

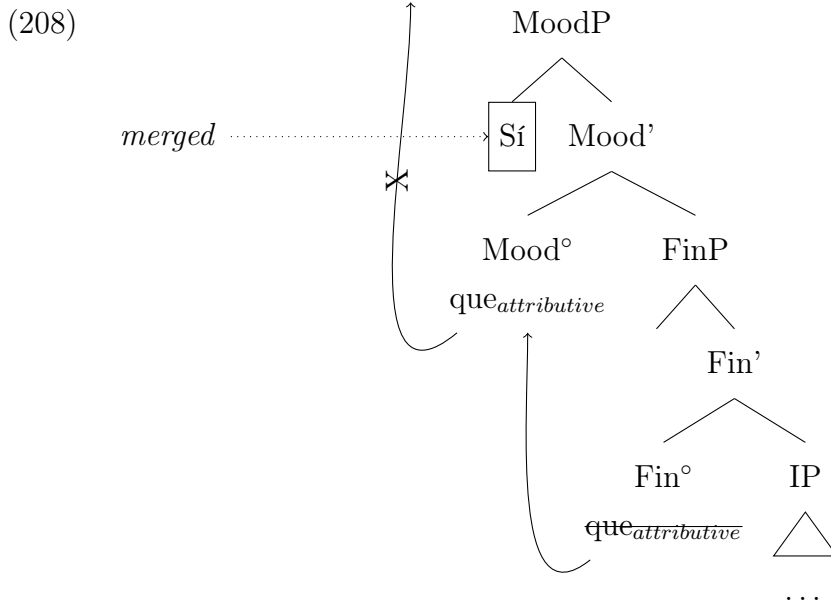
One final piece of evidence demonstrating the adverb-hood of *esclar* is the fact that it can appear in AdvC (207). This development consequently also constitutes compelling evidence for the compositional structure of the constructions, that I argue for in this chapter. The reasoning is that, if these were fixed grammaticalized expressions, we would not expect a productive extension of the construction to novel modifiers.

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 to 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. It is illustrated in (208). 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 as mentioned in section 1.4. The verum marker is directly merged in the specifier of this head. This idea is inspired by Lohnstein (2015), 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 2016). 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.

Lohnstein (2015) uses German data to support his proposal. German is a V2 language, which means that the finite verb necessarily 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 (2015) proposes that the position that is 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.⁴¹ The main empirical evidence for his idea is given in (209). In German declaratives *verum* is marked through stress on the finite verb (as in (209a)). In embedded declaratives, in which the finite verb is sentence final, the stress is not on the verb, but on the complementizer (209b). Lohnstein (2015) takes this as evidence that the complementizer occupies MoodP, hence why the focus feature is expressed as stress on this element.

- (209) a. Karl HAT den Hund gefüttert. (Ger)
 Karl has the dog fed
 ‘Karl DID feed the dog.’ (Lohnstein 2015, 1: ex 1a)
- b. Aber Maria glaubt DASS Karl in Urlaub gefahren ist. (Ger.)
 but Maria believes that Karl in holiday driven is
 ‘But Maria believes that Karl DID go on holiday.’
 (Lohnstein 2015, 2: ex 1e)

In Kocher (2017, 2018b), I show that the analysis proposed by Lohnstein (2015) 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 on, and *sí* is merged directly in the specifier of the projection to satisfy this requirement. This analysis moreover offers a straightforward explanation for why the *verum* marker *sí* does not obtain in imperatives (cf. (210a) vs. (210b)). As shown in section 1.4, this clause type requires local checking of the sentence mood feature and the imperative verb moves to Mood°. Consequently, there is no need for the merger of *sí*, because the focus feature can be expressed on the verb itself.

⁴¹A relation between sentence mood and *verum* has already been observed by Höhle (1992) who offers the first extensive study of *verum* focus.

- (210) a. *¡Sí que cógete una silla de una vez! (Sp.)
 VERUM QUE grab-CL.2S a chair for once
 (Kocher 2018b, 27: ex 49)
- b. ¡CÓGETE una silla de una vez! (Sp.)
 grab-CL.2S a chair for once
 ‘GRAB a chair, at once!’ (Kocher 2018b, 27: ex 49)

The presence of *que* is not required for Spanish and Catalan verum sentences (cf. (211)).

- (211) Sí (que) ve a la festa. (Cat.)
 VERUM QUE comes 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 section 3.4. With regard to the syntactic structure, the idea of the present analysis is again that *sí* is merged in the same place irrespective of whether *que* is present or not. This allows me to maintain 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 2016). The data in (212) constitute the principal evidence against this assumption. They show that *sí* cannot occupy this position, because it can co-occur with sentential negation.

- (212) a. [_{TopP} Eso] [_{MoodP} sí] [_{PolP} no podía faltar en ninguna casa.] (Sp.)
 this VERUM not could miss in any house
 ‘This could NOT be missed in any house.’
 (Kocher 2017, 94: ex 31a from CdE)
- b. [_{TopP} Aixó] [_{MoodP} sí] [_{PolP} no sé fins quan durarà amb l’]
 this VERUM not know.1S until when take.FUT with the
 ajuntament que tenim.]
 city hall that have.2S
 ‘I do NOT know how long this will take with the city hall we have.’ (caWac)

One prediction that follows from the analysis presented here is again that it should not be possible for *sí* and *que* to be separated by other left-peripheral material. The example in (213) shows that this prediction holds true. A clitic left-dislocated topic cannot intervene between *sí* and *que*.

- (213) *Sí aquesta noia_i que la_i conec. (Cat.)
 VERUM this girl QUE CL.FS know.1S

Because of its relative position within the left periphery, we expect that it should be possible for the sequence of *sí* and *que* to be both followed and preceded by topics. The examples in (214a) and (214b) show that this is indeed the case.

- (214) a. [_{TopP} Aquesta noia_i] [_{MoodP} sí que] la_i conec.
 this girl VERUM QUE CL.FS know.1S
 ‘This girl I DO know.’
- b. [_{MoodP} Sí que] [_{TopP} aquesta noia_i] la_i conec.
 VERUM QUE this girl CL.FS know.1S
 ‘This girl I DO know.’

The analysis moreover predicts that *sí que* should follow modifiers merged in ModP, which is also what we find in the data.⁴²

- (215) [_{ModP} Segurament] [_{MoodP} sí que] és el millor dissenyat a nivell
 surely VERUM QUE is the best design at level
 funcional (o això crec jo). (Cat.)
 functional or this believe.1S 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 (216a) and (216b), it follows a verb of saying. In (216c), it appears in an appositive relative clause. In (216b), a topic and an epistemic modifier are merged above AffC. This demonstrates the availability of a large number of structural positions between AffC and the subordinating complementizer analyzed in SubP.

- (216) a. Ella le dijo [_{SubP} que] [_{MoodP} sí que] le extrañaba. (Sp.)
 she CL.MS told that VERUM QUE CL.MS missed
 ‘She told him that she DID miss him.’ (CdE)

⁴²There are some examples in which the reverse order obtains. Although a definite analyses of these cases is pending further investigation, it is notable that examples of *sí* following AdvC are restricted to cases involving *claro que*. However, there are indications that *claro que* is undergoing a process of grammaticalization and turning into a fixed expression (Kocher (2014, 2018a), see also section 4.3). It might therefore no longer be analyzed compositionally by some speakers, which could be a potential way of accounting for these examples.

- b. [...] i ha pensat [SubP que] [ModP potser] [TopP d' aquesta dona_i]
 and has thought that maybe of this woman
 [MoodP sí que] se 'n_i podria enamorar. (Cat.)
 VERUM QUE CL.REFL CL.PART could 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 technologies no són dolentes, el [SubP que] [MoodP sí que] pot
 the technologies are not bad the that VERUM QUE can
 ser perjudicial és l'ús que en fem. (Cat.)
 be prejudicial is the use that CL.PART make.2P
 'Technologies themselves are not bad, what CAN be prejudicial is our use of
 them.' (caWac)

(217) 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.

- (217) A: No vas a venir, ¿pues no?
 not go.2S to come right
 B: Sí que vengo.
 VERUM QUE come.1S
 A: ¿Eh?
 huh
 B: [SubP Que] ... [MoodP sí que] vengo.
 QUE VERUM QUE come.1S
 'A: You're not coming, right? B: I AM coming. A: Huh? B: [reportative:] I AM coming.'

Finally, the data shown in (218), 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.

- (218) a. ¡Eso [MoodP sí que_i] [FinP t_i] [PolP no!] (Sp.)
 this VERUM QUE no
 'This: no way!' (CdE)
- b. Ara [MoodP sí que_i] [FinP t_i] [PolP sí!] (Cat.)
 now VERUM QUE yes
 'Now: absolutely yes!' (caWac)

I will now turn to the question of cross-linguistic variation. As stated at the top of this section, AffC is not available in European Portuguese. Contrary to the claim put forward in Martins (2013), I did however find attestations of the Portuguese equivalent *sim que* in Brazilian Portuguese in the CdP. Moreover, it shows the same behavior as in Spanish and Catalan. Therefore I propose that the analysis proposed for AffC can also be extended to the cognate construction in Brazilian Portuguese. The data in (219) show that there are versions with and without *que* and that *sim* can be preceded by a topic.

- (219) a. [TopP Isso] [MoodP sim que] é uma idéia completamente errada.
 that VERUM QUE is an idea completely wrong
 (B.Pt.)

‘That IS a completely wrong idea.’ (CdP)

- b. [TopP Isto] [MoodP sim] é ser feliz. (B.Pt.)
 this VERUM is be happy

‘This IS being happy.’ (CdP)

The data in (220) show that the versions of the *verum* construction both with and without *que* also appear in embedded sentences. Therefore, as predicted by the analysis, they appear to be merged lower than SubP.

- (220) a. Acho [SubP que] [MoodP sim que] a melhor compra é aquela
 believe.1S that VERUM QUE the best purchase is that
 que conseguimos combinar bem com o nosso «closet». (B.Pt.)
 that manage.1P combine well with the our closet
 ‘I believe the best purchase IS the one that we manage to combine well with
 our existing closet.’ (CdP)
- b. Eu procuro mostrar [SubP que] [MoodP sim] é difícil. (B.Pt.)
 I try show that VERUM is hard
 ‘I am trying to show that it IS hard.’ (CdP)

Finally, the example in (221) shows that both *verum* constructions are also compatible with sentential negation, suggesting that they are not merged in a polarity projection immediately above the IP but in a left-peripheral position.

- (221) a. Aí sim que não vai sair. (B.Pt.)
 there VERUM QUE not go.3S leave
 ‘You will NOT get out there.’ (CdP)

- b. Isso *sim* não tem. (B.Pt.)
 that VERUM not has
 ‘S/he does NOT have that.’ (CdP)

In European Portuguese, *sim que* is not allowed. According to Martins (2013) there are two alternative strategies to express verum in this variety: verb reduplication and final or post-verbal *sim*. The verb reduplication strategy, in which the finite verb is doubled, is illustrated in (222).

- (222) A: O João não comprou o carro. (Pt.)
 the João not bought the car
 B: O João comprou o carro comprou.
 the João bought the car bought
 ‘A: João didn’t buy the car. B: João DID buy the car.’ (Martins 2013, 97: ex 1)

The analysis proposed in Martins (2013) is given in (223).

- (223) $[_{Top} [_{ele} comprou_i o carro]_k [_{Top'} [_{CP} [_{C'} [_{C} comprou_i] [_{\Sigma P} ele_j [_{\Sigma'} comprou_i] [_{IP} I'] comprou_i [_{VP} ele_j comprou_i o carro]]]]]]_k]]]$ (Martins 2013, 101: 8c)

Her proposal is that verb reduplication results from the phonetic realization of two copies of the verbal movement chain (see also Nunes 2001, 2004). The finite verb moves from the IP to ΣP , a polarity position that is 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 last step of 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° and the (nonspecified) 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 reconciled fairly straightforwardly with the analysis developed in this chapter. My tentative proposal is that the non-specified C-position in Martins (2013) could be MoodP. If this can be shown to be correct, these data could constitute further support for the idea that the underlying structure of verum is universal.

The second alternative strategy that European Portuguese employs according to Martins (2013) is one in which *sim* appears post-verbally. This is exemplified in (224).

- (224) Comprou *sim*. (Pt.)
 bought AFF
 ‘Yes, he DID buy it.’ (Martins 2013, 117: ex 57b)

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.

I believe there is evidence that suggests that post-verbal *sim* is not completely equivalent to *sí/sim (que)*. In fact, speakers use *verum* for different purposes. One of these purposes is to emphasize the truth value of a proposition. But 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 the present thesis. 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. From the empirical evidence, we can conclude that Sp., Cat. *sí (que)* and B.Pt. *sim que* can express both functions. This is illustrated with Spanish data in (225). In (225a), speaker A states that the interlocutors disagreed in one aspect and speaker B emphatically stresses the truth of this statement. In (225b) the sentence preceding the one headed by AffC, states that what follows contradicts the interlocutor’s belief.

- (225) a. A: Recuerdo perfectamente que discrepamos en este aspecto. B:
 remember.1S perfectly that disagree.1P in this aspect
 Sí que discrepamos.
 VERUM QUE disagree.1P
 ‘A: I remember perfectly that we disagreed in this aspect. B: We DID in fact disagree.’ (CdE)
- b. No estoy de acuerdo con tí. Sí que hay sistemas de pensiones
 not am in agreement with you VERUM QUE are systems of pensions
 que son sostenibles.
 that are sustainable
 ‘I don’t agree with you. There ARE systems of pensions that are sustainable.’
 (CdE)

The post-verbal *sim* strategy in Portuguese (see (224)) appears to be specialized for expressing counter-assertive focus. Support for this hypothesis comes from a contrast noted in Martins (2013) between initial and final *sim* answering strategies. While initial *sim* is used as an answer to a neutral, unbiased question (226a), in (226b) the final *sim* is used to rebut the negative bias of the question.

- (226) a. A: Ele comprou o carro vermelho? (Pt.)
 he bought the car red
 B: Sim, comprou.
 AFF bought
 ‘A: Did he buy the red car? B: Yes, he did.’ (Martins 2013, 117: ex 56a-b)
- b. A: Ele não comprou o carro vermelho, (pois não)? (Pt.)
 he not bought the car red POIS not
 B: Comprou sim.
 bought AFF
 ‘A: He didn’t buy the car, did he? B: Yes, he DID.’
 (Martins 2013, 117: ex 57a-b)

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° , 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 (227) 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).

- (227) a. E ainda esqueceu-se de consultar um dicionário para vê que a
 and yet forgot-CL to consult a dictionary to see that the
 palavra presidenta existe sim na língua portuguesa. (B. Pt.)
 word presidente exists SIM in the language portuguese
 ‘And yet it was forgotten to consult a dictionary to see that the word ‘pres-
 identa’ DOES exists in the Portuguese language.’ (CdP)
- b. Escribir novelas tiene sí algo de riesgo. (Columbian Sp.)
 write novels has SI something of risk
 ‘Writing novels DOES carry some risks.’ (CdP)

Final *sí/sim* is furthermore compatible with attributive *que*. In (228) this is illustrated by an example where it co-occurs with AdvC.

- (228) a. Claro que existe sim uma relação. (B.Pt.)
 clear QUE exists SIM a relation
 ‘Clearly, there DOES exists a relation.’ (CdP)
- b. Claro que la creatividad debe sí girar entorno a la música.
 clear QUE the creativity should SI revolve around of the music
 (E.Sp.)

‘Clearly, the creativity SHOULD revolve around music.’ (CdE)

Although a consistent analysis is pending further investigation, the data reviewed so far suggest that post-verbal *sim* is merged in a low position. Consequently, there appears to be more than one merge positions for *sím/sí*. Two positions were already explained in the previous discussion: I propose that in the AffC construction it is merged directly in the left periphery and in a short affirmative answer and their emphatic variants in a polarity position between FinP and IP, as exemplified in (229).

- (229) a. [_{PolP} Sí.] (Sp.)
 yes
- b. [_{SubP} Que_i] [_{FinP} t_i] [_{PolP} sí.] (Sp.)
 QUE yes
- c. [_{MoodP} Sí que_i] [_{FinP} t_i] [_{PolP} sí.] (Sp.)
 VERUM QUE yes
- d. [_{ModP} Claro que_i] [_{FinP} t_i] [_{PolP} sí.] (Sp.)
 clear QUE yes

The post-verbal *sí/sim* constructions, in my view, constitutes evidence for a third, lower merge position accessible to this particle. Finally, the various functions of *sí/sim* additionally suggests that this particle, like *que*, acquires its functional meaning in the position it is merged in. This suggests that *que* is not an exception but that functional items more generally might share this property.

To conclude, I briefly summarize the major points. In the previous 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 sections 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. Moreover, in the previous sections 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 that is merged in the left periphery but can cross moved phrases. In the following section 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 core meaning present in all the different constructions which gives rise to different effects depending on the clause type and the meaning of the elements with which attributive *que* co-occurs.

3.4 Pragmatic Requirements

In the present section I focus on the meaning contributed by attributive *que*. My proposal is that 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. This is illustrated by the contrast between the examples in (230) and (231).

- (230) A: Cual es la falsa idea que tienen los doctorandos al inicio de
 what is the false idea that have the PhD students at-the beginning of
 sus estudios?
 their studies
- B: Que seguramente acabarán su tesis a tiempo.
 That surely finish.FUT.3S their thesis on time
- B': #Que seguro que acabarán su tesis a tiempo.
 That sure QUE finish.FUT.3S 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 (230), 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 in itself is not attributive.

(231) A: Qué dicen los doctorandos al inicio de sus estudios?
what say.3S the PhD students at-the beginning of their studies

B: Que seguramente acabarán su tesis a tiempo.
That surely finish.FUT.3S their thesis on time

B': Que seguro que acabarán su tesis a tiempo.
That sure QUE finish.FUT.3S 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 (231), 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 the following, 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 details 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 coin the term *informative presupposition* to capture this fact. 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 (232b) as a reaction to (232a), because it directly denies the content of the appositive clause.

- (232) 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 Stanford 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.

- (233) A: Seguro que o Pedro chegou à hora. (Pt.)
 sure QUE the Pedro arrived on-the time
 B: Não, chegou atrasado.
 no arrived late
 ‘A: Surely Peter arrived on time. B: No, he arrived late.’

At present, this crucial contrast between 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 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 of context is given in (234). Speaker E. asserts the same content that is then taken up again by speaker M. in the proposition introduced by AdvC. This can be used as a tool to emphasize the fact that the commitment to the proposition is shared by both interlocutors.

- (234) E. - Isso é uma aposta e o governo pode sair perdendo. M. -
 this is a gamble and the government can leave losing
 Claro que é uma aposta. (Pt.)
 clearly QUE is 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. This way, attributive *que* can 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 (235), where AffC introduces a proposition that contains a potentially controversial opinion the speaker holds about the hearer.

- (235) No em valores prou. No em fas cas. A ell sí que
 not CL.1S appreciate.2S enough not CL.1S make.2S case DAT he VERUM QUE
 n' hi faries. (Cat.)
 CL.PART CL.DAT make.COND.2S
 '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.

- (236) "Oye, María. Yo creo que aquí hay seis pies." "¿Pero qué dices?" "Yo
 listen María I believe that here are six feet but what say.2S I
 estoy convencido de que aquí hay seis pies." "¿Que no, hombre, que sólo hay
 am convinced of that here are six feet QUE no man QUE only are
 cuatro!" (Sp.)
 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)

A context in which the attributive *que*-initial sentence rebuts an opposing hearer-commitment is exemplified in (236). 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.

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). The 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) to mean 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 as *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 to this approach that is developed in Malamud & Stephenson (2015) is that there are projected versions of other parts of the scoreboard. Particularly relevant for the phenomenon under discussion here is the assumption that there are 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 6. 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.

current		projected	
CG	$\{\}$	CG*	$\{\{p\}\}$
DC _{Speaker}	$\{p\}$	DC* _{Speaker}	$\{\{\}\}$
DC _{Hearer}	$\{\}$	DC* _{Hearer}	$\{\{\}\}$
Table	$\langle p \rangle$	Table*	$\langle \langle \rangle \rangle$

Table 6: Conversational scoreboard when asserting p .

With the 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. My proposal is that in all the different constructions containing attributive *que* the proposition is added to the hearer’s DC. Moreover, the contribution of *que* in declaratives and polar questions only differs in that the proposition is added to the current DC of the hearer in the former and to the projected DC of the hearer in the latter.

Table 7 illustrates the basic case of attributive *que*-initial declaratives. I assume that they have all the 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.

current		projected	
CG	$\{\}$	CG*	$\{\{p\}\}$
DC _{Speaker}	$\{p\}$	DC* _{Speaker}	$\{\{\}\}$
DC _{Hearer}	$\{p\}$	DC* _{Hearer}	$\{\{\}\}$
Table	$\langle p \rangle$	Table*	$\langle \langle \rangle \rangle$

Table 7: Conversational scoreboard when asserting attributive *Que p*.

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 for the assumption made by some authors in the literature that the function of *que* in these contexts is to mark 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 am hesitant to relate this causal meaning to the complementizer itself, because a causality interpretation is available even in the absence of *que*. This means that even without a special marker, the interpretation that two juxtaposed sentences are causally related, is available.

- (237) a. No menges pipas. (Que) fa castellà. (Cat.)
 not eat sunflower seeds QUE makes Castillian
 ‘Don’t eat sunflower seeds. That makes you (look) Spanish.’ (ebook-cat)
- b. Keep calm. (Que) hoje é sexta-feira. (Pt.)
 keep calm QUE today is Friday
 ‘Keep calm. Today is Friday.’

In (237a) 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. In (237b), the *que*-initial sentence could be interpreted as a reason why you should keep calm. My idea is that, if the sentence is introduced by *que*, in the present cases, it is additionally marked as uncontroversial information that according to the speaker both speaker and hearer are committed to.

Additionally, there are cases, where the *que*-initial declarative follows a directive, in which a causal interpretation is not really consistent. One such example is given in (238).

- (238) *Salid.* Que no vos mataré. (Sp.)
 leave.IMP.2P QUE not CL.2P kill.FUT.1S
 ‘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 reasons for saying *Salid*.⁴³ The motivation for the second sentence is to reassure the hearers and make them believe that there is no imminent threat to their life. The idea is that by adding attributive *que* at the beginning of the sentence, the speaker anticipates the hearers’ agreement with it, which makes it sound even more persuasive and less threatening.

⁴³An explicative interpretation, however, whereby not killing the hearers could be the explanation for why they should come out, is possible.

My proposal for AdvC-declaratives is that they involve the same conversational states as attributive *que*-initial declaratives, illustrated in table 8. 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 evidential modifiers like *clearly*, imply that the speaker is committed to the truth of the proposition while others like *allegedly* do not.

	current		projected
CG	{}	CG*	{{ <i>p</i> }}
DC _{Speaker}	{ <i>claro</i> (<i>p</i>)}	DC* _{Speaker}	{{}}
DC _{Hearer}	{ <i>p</i> }	DC* _{Hearer}	{{}}
Table	< <i>p</i> >	Table*	<<>>

Table 8: Conversational scoreboard when uttering *Claro que p*.

To understand the effect of attributive *que* in combination with evidential and epistemic modifiers, it is useful to contrast AdvC with similar constructions that involve the same epistemic and evidential modifiers but no attributive complementizer (see Kocher 2018a and section 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 (239) illustrate the three constructions for comparison here.

- (239) a. Obviamente/ Segurament que Pedro viene a la fiesta. (Sp.)
 obviously/ surely QUE Pedro comes to the party
 ‘Obviously/Surely, Pedro will in fact come to the party.’
 b. Obviamente/ Segurament en Pere ve a la festa. (Cat.)
 obviously/surely the Pere comes to the party
 ‘Pere will obviously/surely come to the party.’
 c. É óbvio/ seguro que o Pedro vem à festa. (Pt.)
 is obvious/sure that the Pedro comes to-the party
 ‘It is obvious/sure that the Pedro will come to the party.’

In (239a), the modifier appears in AdvC, containing attributive *que*. In (239b) it functions as a simple adverb and in (239c), 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 base are built on the information states of different speech participants and are based on 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 (240). A statement like *Peter will probably come* can be interpreted from the perspective of different speech participants and thus anchored to different deictic centers.⁴⁴

- (240) a. (In the view of what I know) Peter will probably come.
 (*deictic center=speaker*)
- b. (In the view of what you know) Peter will probably come.
 (*deictic center=hearer*)
- c. (In the view of what you and I know) Peter will probably come.
 (*deictic center=speaker and hearer*)
- d. (In the view of what is generally known) Peter will probably come.
 (*deictic center=a (contextually relevant) group or authority*)

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 (241), 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 (241a), Billy takes up the reading of the modal that is anchored to the speaker Alex, while in the reaction in (241b), he takes up and rejects a reading anchored to a perspective that involves Billy, the hearer, i.e. himself.

⁴⁴A 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 (2018a) for an application of this concept to Spanish and Portuguese modifiers.

-
- (241) 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.⁴⁵ I therefore also propose that when the speaker utters *obvious(p)* all the contextually relevant readings of *obvious(p)* are put into play (cf. (242)).

- (242) 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*)
- 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 (239b), 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*. The idea is that, 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 (240c) and (242c). 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

⁴⁵I make no claims here about the relation between epistemic modality and evidentiality, but see 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.

same implicature does not arise automatically. The adverbs do not have a prominent reading and 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 (240d) and (242d). 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 section 4.3, which indicate that the prominent reading of the modifiers can be overridden if contextual factors make other readings available.

To conclude the discussion on AdvC, I will briefly discuss the tentative conversational scoreboards that I assume for pure adverbial modification and for EsAdjC and will observe how they contrast with the conversational scoreboard assumed for AdvC. Table 9 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.

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 9: Conversational scoreboard when uttering *Claramente 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*	$\langle \langle \rangle \rangle$

Table 10: Conversational scoreboard when uttering *Está claro p*.

Table 10 represents the conversational scoreboard I assume for EsAdjC, the equivalent biclausal copula construction in which a proposition is modified by a predicative ad-

jective. 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 8, 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.

I will now present my characterization of 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 for these should be treated as preliminary for the time being. 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. (243a)) but they can be confirmed (cf. (243b)).

- (243) a. A: Com és d'alt en Pau? B1: Molt alt. B2: Fa 1.90m. B3: #Que
 how is of tall the Pau very tall makes 1.90m how
 alt que és! (Cat.)
 tall QUE is
 '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 is 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 tentative scoreboard I propose for a *que*-less wh-exclamative like (244b) repeated from (183b) is given in table 11.

- (244) a. Que chato que és. (Pt.)
 how annoying QUE are.2S

- b. Que chato és.
 how annoying are.2S
 ‘How annoying you are.’

current		projected	
CG	$\{p\}$	CG*	$\{\{\}\}$
DC _{Speaker}	$\{Excl(p)\}$	DC* _{Speaker}	$\{\{\}\}$
DC _{Hearer}	$\{\}$	DC* _{Hearer}	$\{\{\}\}$
Table	$\langle \rangle$	Table*	$\langle \langle \rangle \rangle$

Table 11: Conversational scoreboard when uttering *Qué x p*.

The proposition and the exclamative import are added to the speaker’s DC.⁴⁶ As stated above, exclamatives do not assert the prejacent, i.e. the underlying proposition. In the current model, I represent this by not putting the proposition on the table nor into the projected CG. Exclamatives are factive, therefore they presuppose the truth of the prejacent. I propose to model this by imposing the proposition on the current CG.

The main argument of the present section is that the discourse contribution of the attributive complementizer is the same in all the constructions that it is involved in. Therefore, in wh-exclamatives with *que* like (244a) repeated from (183a), 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 12, is the same as the one I assume for *que*-less exclamatives.

current		projected	
CG	$\{p\}$	CG*	$\{\{\}\}$
DC _{Speaker}	$\{Excl(p)\}$	DC* _{Speaker}	$\{\{\}\}$
DC _{Hearer}	$\{p\}$	DC* _{Hearer}	$\{\{\}\}$
Table	$\langle \rangle$	Table*	$\langle \langle \rangle \rangle$

Table 12: Conversational scoreboard when uttering *Qué x que p*.

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 could explain why many authors consider *que* to be optional and to have no effect on the interpretation (see for instance

⁴⁶The 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.

Villalba 2003, Castroviejo 2006). In the *wh*-exclamatives that permit versions without the complementizer, we expect 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. (245)).

- (245) A: És alt en Pau? B: Sí (que) ho és! (Cat.)
 is tall the Pau VERUM CL.NEUT is
 ‘A: Is Pau tall? B: He IS (indeed)!’

In the current model, this means that the speaker does put the proposition on the table. However, like exclamatives, they are factive and thus presuppose the truth of the prejacent. In my current proposal this is modeled as an imposition of the proposition on the CG.

- (246) a. Pablo sí que es alto. (Sp.)
 Pablo VERUM QUE is tall
 b. Pablo sí es alto. (Sp.)
 Pablo VERUM is tall
 ‘Pablo IS tall.’

The tentative scoreboard for a *que*-less verum sentence like (246b) is given in table 13. The verum marked proposition is part of the speaker’s DC. The prejacent is put on the table and imposed on the current CG.

current		projected	
CG	{ <i>p</i> }	CG*	{{}}
DC _{Speaker}	{ <i>verum</i> (<i>p</i>)}	DC* _{Speaker}	{{}}
DC _{Hearer}	{}	DC* _{Hearer}	{{}}
Table	< <i>p</i> >	Table*	<<>>

Table 13: Conversational scoreboard when uttering *Sí p*.

Table 14 illustrates the conversational scoreboard I assume for *sí que*-sentences like (246a). The presence of the attributive complementizer again marks 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.

current		projected	
CG	$\{p\}$	CG*	$\{\{\}\}$
DC _{Speaker}	$\{verum(p)\}$	DC* _{Speaker}	$\{\{\}\}$
DC _{Hearer}	$\{p\}$	DC* _{Hearer}	$\{\{\}\}$
Table	$\langle p \rangle$	Table*	$\langle \langle \rangle \rangle$

 Table 14: Conversational scoreboard when uttering *Sí que p*.

The final part of this section is dedicated to the discourse contribution of attributive *que* in polar questions like (247).

- (247) Que tens pressa? (Cat.)
 QUE have.2S stress
 ‘Are you stressed out?’ (ebook-cat)

My proposal is that here too the function of attributive *que* is to attribute the 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 15 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.⁴⁷ The positive and the negative version 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.

current		projected	
CG	$\{\}$	CG*	$\{\{p\}, \{\neg p\}\}$
DC _{Speaker}	$\{\}$	DC* _{Speaker}	$\{\{\}\}$
DC _{Hearer}	$\{\}$	DC* _{Hearer}	$\{\{\}\}$
Table	$\langle p \rangle$	Table*	$\langle \langle \rangle \rangle$

 Table 15: Conversational scoreboard when asking p ?

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 16). This means that the prejacent is put on the table and the positive and negative version are added to the projected CG. In section 3.1.2, I proposed the generalization that *que* in polar questions signals that the speaker expects a positive answer from the hearer. I model this

⁴⁷For 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).

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, 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.

current		projected	
CG	{}	CG*	{{p}, {¬p}}
DC _{Speaker}	{}	DC* _{Speaker}	{{}}
DC _{Hearer}	{}	DC* _{Hearer}	{{p}}
Table	< p >	Table*	<<>>

Table 16: Conversational scoreboard when asking *Que p?*

In section 3.1.2, 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 17), but contextual evidence contradicts her belief. This is illustrated in the example in (248) repeated from (114).

- (248) Que vindràs a Barcelona? No em pensava pas que ens acompanyessis.
 QUE will come to B. not me thought NEG that us accompany
 ‘Are you coming to Barcelona? I didn’t think you were coming with us.’
 (Prieto & Rigau 2007, 15)

I define contextual evidence as evidence accessible to the speaker in the current discourse situation or in a previous situation (Kocher 2017).⁴⁸ Importantly, given the contextual

⁴⁸This 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 the speaker and the hearer. In Kocher (2017), 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.

evidence and despite the fact that it is in disagreement with 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.

previous		current		projected	
CG	{}	CG	{}	CG*	{{p}}, {¬p}}
DC _{Speaker}	{¬p}	DC _{Speaker}	{}	DC* _{Speaker}	{{}}
DC _{Hearer}	{}	DC _{Hearer}	{}	DC* _{Hearer}	{{p}}
Table	<>	Table	< p >	Table*	<<>>

Table 17: Conversational scoreboard when asking an anti-expectational *Que p?*

Confirmatory polar questions such as (249) present a different picture.

(249) [Context: Anna meets her friend Carles. He is tanned and seems relaxed. Anna asks:]

Que has estat de vacances?

QUE have.2S been on vacation

‘Have you been on vacation?’ (Kocher 2017, 19: ex 40a)

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 what it is founded on – in this case of (249), indirect contextual evidence – she expects that the answer of the hearer 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.

current		projected	
CG	{}	CG*	{{p}}, {¬p}}
DC _{Speaker}	{}	DC* _{Speaker}	{{p}}
DC _{Hearer}	{}	DC* _{Hearer}	{{p}}
Table	< p >	Table*	<<>>

Table 18: Conversational scoreboard when asking a confirmatory *Que p?*

To conclude the discussion on *que* in polar questions, I will now provide a tentative characterization of the scoreboard of these questions when they contain the question particles *oi* or *en*. 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.(250)).

- (250) a. Sense fer res es fa més llarga l' estona. Eh que
 without do nothing CL.IMP makes more long the time period EH QUE
 sí? (Cat.)
 yes
 ‘When doing nothing time grows longer. Right?’ (caWac)
- b. No tots els partits estan acusats d' haver -se finançat il·legalment,
 not all the parties are accused of have CL.REFL financed illegally
 oi que no? (Cat.)
 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 a request 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 (251a): these sentences are infelicitous with *eh?* but felicitous with *oi?*. On the other hand, in requests for confirmation of opinions like in (251b) repeated from (116), both particles are allowed.

- (251) a. La Terra és rodona, oi?/ #eh? (Cat.)
 the earth is round OI EH
 ‘The Earth is round, right/ #huh?’ (Castroviejo 2018, ex 18)
- b. T’ has tallat els cabells, oi?/ eh? (Cat.)
 CL.2S have.2S cut the hair OI EH
 ‘You had your hair cut, right?/ huh?’ (Castroviejo 2018, ex 19)

Castroviejo (2018) proposes the following conversational scoreboards to model the discourse contribution of the sentence final particles. Table 19 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. Castroviejo (2018) proposes a different contribution for *eh*. The corresponding scoreboard is given in table 20.

current		projected	
CG	{}	CG*	{{p}}
DC _{Speaker}	{}	DC* _{Speaker}	{{p}}
DC _{Hearer}	{}	DC* _{Hearer}	{{}}
Table	< p >	Table*	<<>>

Table 19: Conversational scoreboard when uttering *p*, *oi?* adopted from Castroviejo (2018)

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. (251a)). 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.

current		projected	
CG	{}	CG*	{{p}}
DC _{Speaker}	{}	DC* _{Speaker}	{{}}
DC _{Hearer}	{}	DC* _{Hearer}	{{p}}
Table	< p >	Table*	<<>>

Table 20: Conversational scoreboard when uttering *p*, *eh?* adopted from Castroviejo (2018)

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.

- (252) Oi/ Eh que acabaràs la feina? (Cat.)
 OI EH finish.FUT.2S the work
 ‘You’ll finish your work, right?’ (Castroviejo 2018, ex 15)

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 tentative scoreboard in table 21.

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 question. 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

current		projected	
CG	$\{\}$	CG*	$\{\{p\}\}, \{\{\neg p\}\}$
DC _{Speaker}	$\{\}$	DC* _{Speaker}	$\{\{p\}\}$
DC _{Hearer}	$\{\}$	DC* _{Hearer}	$\{\{p\}\}$
Table	$\langle p \rangle$	Table*	$\langle \langle \rangle \rangle$

Table 21: Conversational scoreboard when uttering *Oi/Eh p?*

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 the conversational scoreboards 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 some 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.

4 Empirical Support

In this chapter I present empirical studies that focus on the pragmatic effects of attributive *que*. The first motivation behind this more extensive examination of these constructions is that only a small number of existing investigations focus on interpretation of them. This is why more detailed research has proven necessary. Since there are no previous analyses that can be used as a basis for predictions, the empirical studies that will be described in the following sections, approach the phenomena in an exploratory manner. The second motivation for the the extensive examination that I have chosen to carry out has to do with the phenomena in and of itself. The pragmatic effect of attributive *que* is strongly context-dependent. Furthermore, 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 way of acquiring an initial impression of the possible pragmatic effects in the constructions. Thirdly, given these properties, it seems necessary to take native speaker's judgments on the acceptability and felicitousness of the constructions into account. Finally, this approach 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. Section 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 sampled from corpus data. In the following section 4.2, I provide some background information on the statistical tools that I used to model the data. Section 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 section 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 section 4.5, which provides a general discussion on how exploratory empirical research can inform

our theoretic choices.

4.1 Corpus-Based Experiments

For the experiments I use a method that is very much inspired by the work described in Degen (2015). It differs from more traditional experimental approaches in that the stimuli are not constructed by the experimenters but are sampled from corpora. This means that the stimuli are heterogeneous by design. This design proves particularly useful in pragmatics research, partly 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 analyses can be employed to understand the data, hence 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. More traditional experiments concentrate on a few of these factors. While this leads to very clear results with respect to the effects of these few factors in the precise experimental contexts, it is 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 established methods, which does carry some risks, but offers the significant advantage of finding new approaches to the data and possibly 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 speakers' perspectives. In the best case scenario, it provides us with fine grained information on aspects of a construction that cannot be extracted from the corpora alone. This then leads to a more complete picture of the issues under investigation.

In Degen (2015), all occurrences containing the relevant pattern 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 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 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 fashioned 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 to a certain degree in many areas of linguistic research, for instance when a normative bias could influence the judgments. However, in the area that I investigate 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, enables the participants to be 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 I try to answer.

4.2 Statistical Modeling

I follow Baayen (2008) in ascribing 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.

4.2.1 Linear and Generalized Additive Models

Generalized additive models (hence GAMs) are statistical models that are generalized from linear models. Modeling data with a linear regression means that we expect that

one or more predictor variables will have a linear effect on the response variable. This means that we assume that when plotting our data points we can approximate them by a straight line. This can be seen in figure 1.

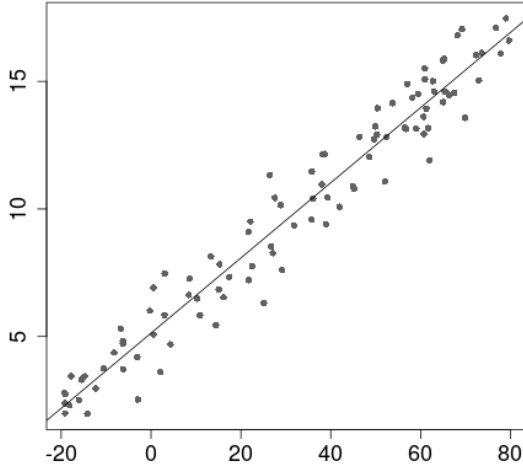


Figure 1: Random data points and their linear regression with the function $y = 5.33 + 0.15 * x$.

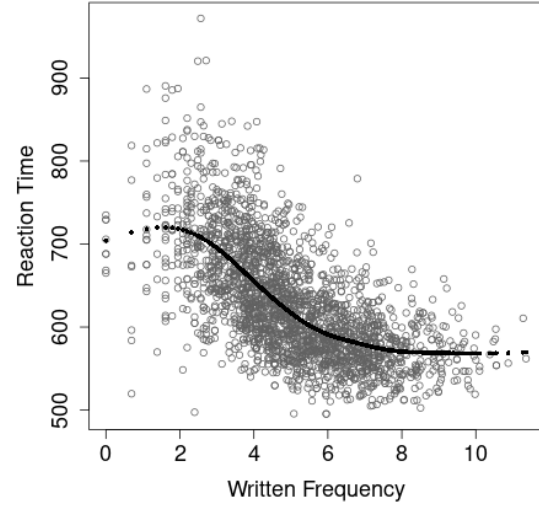


Figure 2: Example of a non-linear effect. The plot shows the effect of **written frequency** on the REACTION TIME in a lexical decision task in the subset of young participants using the built-in dataset **English** from the **languageR** package. The black dotted line give the predicted values of a GAM illustrating the non-linear effect of written frequency on the REACTION TIME.

A simple linear model with one predictor has the function given in (253). A response variable y is predicted as a linear function of β_0 , the intercept (the point where the line meets the y-axis, 5.33 in our toy example), and the linear effect or coefficient β_1 (the slope of the line, 0.15 in our toy example) of a predictor variable x_1 .

$$(253) \quad y = \beta_0 + \beta_1 * x_1$$

In GAMs the assumption that the effect of a predictor is linear is relaxed in order to allow the modeling of data such as that illustrated in figure 2, where the effects is non-linear. Non-linear effects are very frequent in linguistics. GAMs are therefore a useful tool that can help to approximate and understand the complexity of linguistic data in a more principled way.

GAMs are additive in the sense that the functions are built cumulatively on the basis of functional addition. For instance, the simple function given below figure 1 can be calculated as an addition of the horizontal line for the intercept function $y = 5.33$ and the slope function $y = 0.15x$. More complex non-linear functions can be calculated in a very similar way. In the figure in 3 this is shown for a 4th order polynomial function. Each basis function is multiplied by a weight. Varying these weights means that, out of a limited derivable set of basis functions, we are able to generate all 4th order or lower polynomial functions.

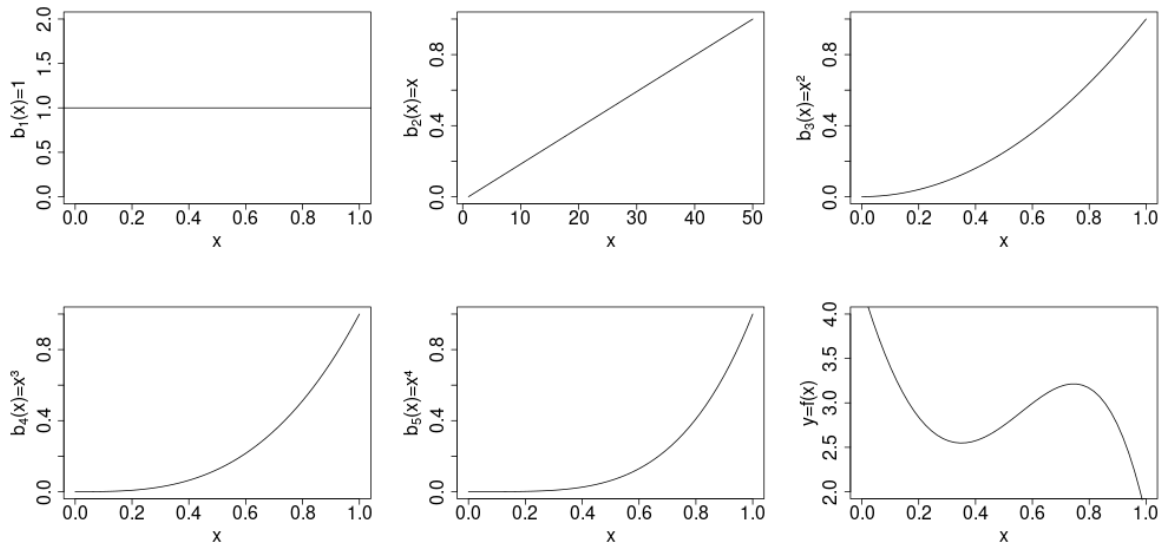


Figure 3: Basis functions and the resulting additive function $f(x)$. The first five panels illustrate the basis functions for a 4th order polynomial function. The basis functions are multiplied by a weight b_i and added up to give the final function $f(x)$ in the sixth panel. Varying the weights by which the functions are multiplied we can produce any 4th order or lower polynomial function. (adapted from Wood 2017, 163)

The weighted basis functions used in GAMs are not polynomial but spline functions. The drawback of high order polynomials is that they can yield oscillations, i.e. interpolations, that have maxima and minima far greater than the original data. Spline interpolations counteract the interpolation error and yields a smoother fit by splitting the data into smaller intervals and interpolating them with a low order polynomial (on polynomial and spline interpolation see Wood 2017, 162, 215-220 and Baayen forthcoming).

Interactions in regression models yield multidimensional effects that can once again be

non-linear. These effects can also be modeled with a GAM in a similar way as two-dimensional effects are modeled. For instance, when two variables interact, the basis functions are three-dimensional and are once again multiplied by weights. In this case, functional addition does not give rise to a wiggly curve but to a wiggly surface that approximates the effect of the interaction.

When fitting any model to actual data, one goal is always to find a way of remaining faithful to the data while keeping the model simple enough so that it can be generalized to a population or to new data. In basic linear models this balance is obtained by choosing the coefficients such that the sum of the squared differences between predicted and observed values is as small as possible (cf. Baayen forthcoming). In GAMs overfitting, i.e. being overly faithful to the data, results in extreme wigglyness of the curve. In order to counteract overfitting in these models, the functions of GAMs are calculated in a way that penalizes wigglyness. This is achieved by means of a smoothing parameter λ that keeps the wigglyness in check. When dealing with three dimensional basis functions, the smooth is achieved through a mathematical operation called tensor product.

To illustrate a GAM, I use the built-in dataset `English` from the `languageR` package in R that was also used in the illustration in figure 2. This GAM, like all the other GAMs presented in this thesis, are modeled using the `mgcv` package in R. The dataset comes from a lexical decision experiment and contains data on the participants and on the lexical items used. The present model is fitted to a subset of the data that contains only young speakers. The model is set out to predict the REACTION TIME of the participants, which encodes the time it took the participants to judge whether a stimulus is a word or a non-word. The only predictor I take into account for this simple illustration is the WRITTEN FREQUENCY of the lexical item.

A. parametric coefficients	Estimate	St. Error	t-value	p-value
(Intercept)	629.547	1.071	587.7	<0.0001
B. smooth terms	edf	Ref.df	F-value	p-value
s(WrittenFrequency)	6.565	7.649	245.9	<0.0001

Table 22: Model summary of the GAM fitted to the REACTION TIMES with one non-linear effect `s(WrittenFrequency)`.

Table 22 summarizes the model. It reports the estimated coefficient (Estimate), the standard error (St. Error), the t-value and the p-value for the linear, parametric effects. In the this case, the only linear term is the intercept. The estimate of the intercept

gives us the point where the regression line meets the y-axis. In a mathematical sense, it is the mean value of the response variable; in our case, the REACTION TIME, when all predictors - here only the written frequency - are set to zero. This means that the intercept can be interpreted as the mean time it takes the participants to react to a visual, written stimulus. In other words, in the experiment it took 629.457 milliseconds until the visual stimulus was recognized as such.

If we had an additional linear predictor in the model, say the NUMBER OF LETTERS in a word, the estimate coefficient would report the average amount of time added to the REACTION TIME for each additional letter in a word. Assume this coefficient is 50 milliseconds. So in order to calculate the predicted REACTION TIME of a three letter word like *fog* in a simple model that considers only the effect of the NUMBER OF LETTERS, we substitute the variables in the linear equation in (253): predicted reaction time of *fog*: $629.547 + 50 \cdot 3$. We take $50 \cdot 3$ for the three letters in the word and add it to the 629.547 estimated for the intercept. This means that our hypothetical model would predict that it takes 779.547 milliseconds to judge whether *fog* is a word or a non-word. The standard error of the estimate is calculated by taking the sum of the root of the squared differences between the observed and the predicted values divided by the number of pairs of observed and predicted values.⁴⁹ This is a measure of the accuracy of the prediction. The smaller the standard error of the estimate, the closer the predicted values are to the observed values. Consequently, the smaller the standard error the more accurate the prediction. The t-value is the estimate coefficient divided by its standard error. The p-value is calculated by checking the t-value against the t-distribution. The p-value tells us whether the effect of the predictor is strong enough to consider it statistically significant. The level of statistical significance is normally fixed at a level of 0.05. With a p-value of 0.05 there is only a 5% percent chance of finding the same distribution in a random sample. The p-value therefore gives the probability of randomly finding a result as extreme as those we observe in a dataset in which the predictor has no effect.

Our model summarized in table 22 contains, in addition to the intercept, a non-linear effect of WRITTEN FREQUENCY. This effect is characterized by the effective degrees of freedom (edf), the reference degrees of freedom (ref.df), the f-value and the t-value. The reference degrees of freedom give the total number of basis functions used. The effective degrees of freedom give us the basis functions estimated to approximate the effect of

⁴⁹ $\sigma_{est} = \sqrt{\frac{\sum (observed - predicted)^2}{N}}$

the predictor. This is achieved by multiplying the total number of basis functions (the reference degrees of freedom) minus 1 (the straight line function for the intercept, because the intercept is reported separately) by its corresponding weights. The higher the effective degrees of freedom, i.e. the more basis functions are used, the wigglier the curve described by the estimated function. The f-value is the result of a test that compares the current model to the intercept-only model, taking into account the squared residual (observed-predicted values) sums and the degrees of freedom of the models. The test checks whether the predictor improves the model. The p-value checks the f-value against the f-distribution and once again tells us whether the results obtained are extreme enough to consider the effect of the predictor statistically significant.

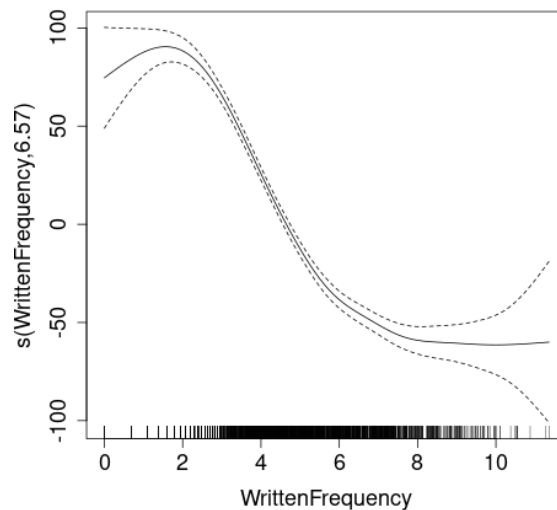


Figure 4: Non linear effect of written frequency.

Figure 4 plots the function estimated for the effect of WRITTEN FREQUENCY. The area between the dotted lines corresponds to the confidence interval. This is once again a measure of accuracy. If the area is small, as in the present figure, we can be confident that the predictions are accurate. The confidence interval is strongly related to the number of observations we have for a particular value. Information about the distribution of the data is contained in this figure in the form of the small horizontal lines directly above the x-axis, the so-called *rug*. The closer together the lines are, the larger the amount of data for that particular frequency and thus the smaller the confidence interval.

The graph shows that when we go from a WRITTEN FREQUENCY of 0 to 2, the REACTION TIME increases. This can probably be attributed to the fact that non-words all have a WRITTEN FREQUENCY of 0 and the participants were a little faster to recognize those than the real words with an extremely low frequency of 1 or 2. After a WRITTEN FREQUENCY of about 2, the effect is reversed: a higher written frequency results in lower REACTION TIME. The steepness of the slope of the curve in the interval of a WRITTEN FREQUENCY of 2 to about 7 shows that an increased WRITTEN FREQUENCY drastically shortens the REACTION TIME in this interval. For the extremely frequent words with frequencies larger than 7, the curve flattens. This means that increasing the WRITTEN FREQUENCY for instance from 7 to 8 hardly has any impact on further reducing the REACTION TIME.

To calculate the REACTION TIME for a word with the WRITTEN FREQUENCY of 0, we read the y-value of the graph at the x-value of 0, which is at 74.6157, and add it to the intercept. This means that our model predicts that it should take 704.1627 milliseconds to decide whether a word with a WRITTEN FREQUENCY of 0 is a word or not.

In this toy example, the response variable was numeric, namely the reaction time. In the analyses of the experimental results, I make use of ordinal GAMs because I have elicited the judgments in my experiments on Likert scales. The ratings provided by the participants could be taken straightforwardly as the numeric response variable; however, this would be based on the assumption that the distance between each step of the scale is equally spaced. Often, however, this is not how the participants use the scale. In ordinal GAMs, the assumption is that the steps may not be equal between each rating step. Factors with ordered levels which may or may not be equally spaced are termed *ordinal factors*. In ordinal GAMs, the actual integer values of the response variable are predicted by the other variables in the model.⁵⁰

4.2.2 Conditional Inference Trees

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

⁵⁰See Baayen & Divjak (2017) for a comparison between a model that treats the rating variable as numeric and one that treats it as ordinal that proves the superiority of the ordinal model.

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. However, they do this based not on the mathematical equation in (253) but on binary recursive splits. Conditional inference trees test whether a predictor is associated with a response variable and choose the 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).

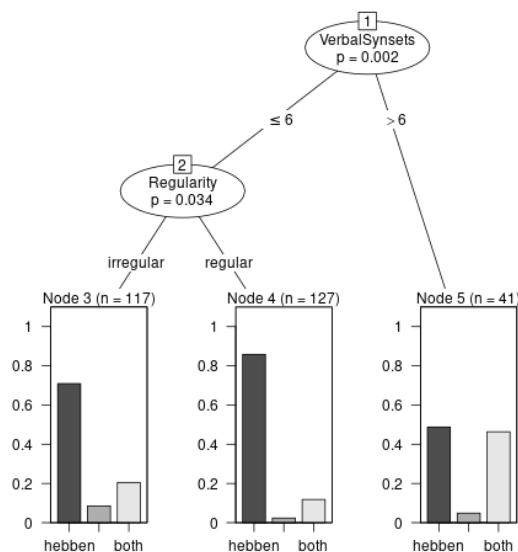


Figure 5: Conditional inference tree.

The tree was fit to the built-in dataset `auxiliaries` from the `languageR` package. 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.

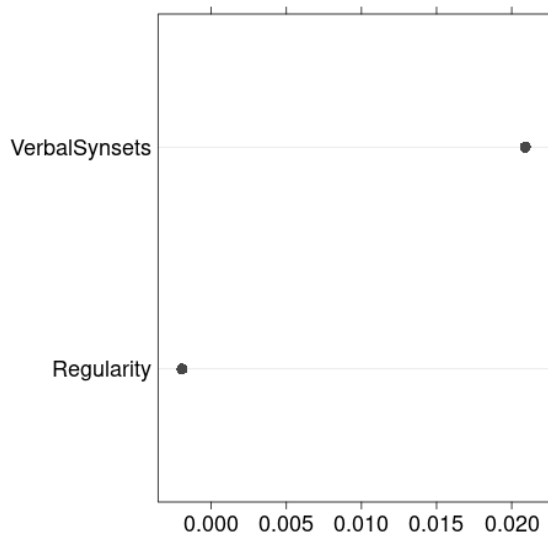


Figure 6: Dot plot of variable importance.

The example tree in figure 5 as well as all the other conditional inference trees in this thesis are modeled using the **party** package in R. In 5 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 with regard to 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). The figure in 6 is based on a random forest calculated for the same model as the conditional inference tree in 5. The figure shows the relative importance of the two predictors in the model. The number of verbal synonyms is detected as the most im-

portant predictor. This is also evident from the conditional inference tree in 5. The dot plot in 6 additionally shows that its importance is far greater than that of the second predictor.

4.3 Deictic Centers of Epistemic and Evidential Modifiers

In the present section, I report three empirical studies, all of which focus on the interpretation of epistemic and evidential modifiers in different constructions. The goal is to relate linguistic and extra-linguistic factors to the use of the modifiers in the 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 (254).

- (254) a. Seguro que Juan viene. (AdvC) (Sp.)
 b. Seguramente Juan viene. (Adv) (Sp.)
 c. Es seguro que Juan viene. (EsAdjC) (Sp.)
 ‘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 sections 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 shared perspective between multiple speech participants. The function I propose for attributive *que* is that it attributes the 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 section 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, whatever perspective is most plausible in the context is adopted.

The first approach to the phenomenon is a corpus analysis that compares the distri-

bution of the constructions in Spanish and Portuguese in different text types. The assumption is that different text types have different likelihoods to invite certain 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 oral than in academic 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 acceptability. 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. This approach offers the potential to learn from the data and draw novel conclusions that could not have been formulated a priori.

4.3.1 Prequel: Comparative Corpus Analysis on Epistemic and Evidential Modifiers

In the present 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 corpora publicly available, 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-crawl corpora, they have the advantage of consisting of different text types. TEXT TYPE was introduced as a variable in the study and proved to be a significant predictor. A subset of the data has been used in a previous study reported in Kocher (2018a). The results presented here are still novel, because both the database and the statistical approach is 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.

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 23). Sample size was considered as a potential predictor, but showed no significant effect.

Text type	academic	press	oral	fiction	total
Corpus del Español	5.1 m	5.1 m	5.1 m	5.1 m	20.4 m
Corpus do Português	5.9 m	6.6 m	2.2 m	6 m	20.7 m

Table 23: Number of tokens in million words per TEXT TYPE in CdE and CdP.

Sampling I collected the sample through the online interface and then exported it to R for adding annotations and performing 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*⁵¹, *claro*, *evidente*, *natural*⁵², *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 24). For EsAdjC, I included versions with copulas *ser* and *estar* (cf. table 25). Neither of these variables showed a significant effect in the modeling.

	cierto		claro		evidente		natural		probable		seguro	
	sp	pt	sp	pt	sp	pt	sp	pt	sp	pt	sp	pt
<i>-mente</i>	14	8	0	0	18	32	30	36	2	0	11	3
<i>-0</i>	34	9	576	306	0	11	2	4	0	0	82	0

Table 24: 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
<i>estar</i>	0	8	34	17	0	0	0	0	0	0	0	0
<i>ser</i>	126	81	24	190	64	96	11	31	85	25	8	1

Table 25: Distribution of *ser* and *estar* per modifier in EsAdjC.

⁵¹In the following I will use the Spanish cognate to refer to all the modifiers.

⁵²For a discussion of the evidential nature of *natural* in Spanish and Portuguese see Kocher (2014).

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 the present section I describe the distribution of these variables in the sample.

	Adv	AdvC	EsAdjC	total
pt	511	409	449	1369
sp	839	769	352	1960
total	1350	1178	801	3329

Table 26: CONSTRUCTION by LANGUAGE.

Table 26 shows the occurrence of each CONSTRUCTION per LANGUAGE. It shows that in general there are more epistemic and evidentially marked propositions 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, while they 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. 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) in addition to European Spanish (Andean, Canarian, Caribbean, Central American, Chilean, Mexican and Rioplatense). 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 only the variable LANGUAGE with the two values Spanish and Portuguese.

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 27: CONSTRUCTION by MODIFIER.

Table 27 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.

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 28: CONSTRUCTION by TEXT TYPE.

Table 28 shows the frequency of each CONSTRUCTION per TEXT TYPE per LANGUAGE. The frequency per TEXT TYPE is comparable in the two LANGUAGES,⁵³ with 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. This is compatible with the idea that AdvC refers to the perspective of an interlocutor. The high number of AdvC in oral texts, in which an interlocutor’s perspective can be addressed most directly, is therefore 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.

Results In the present section, I present the results of the tree model that I fitted to the corpus data. Model selection was carried out in an 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

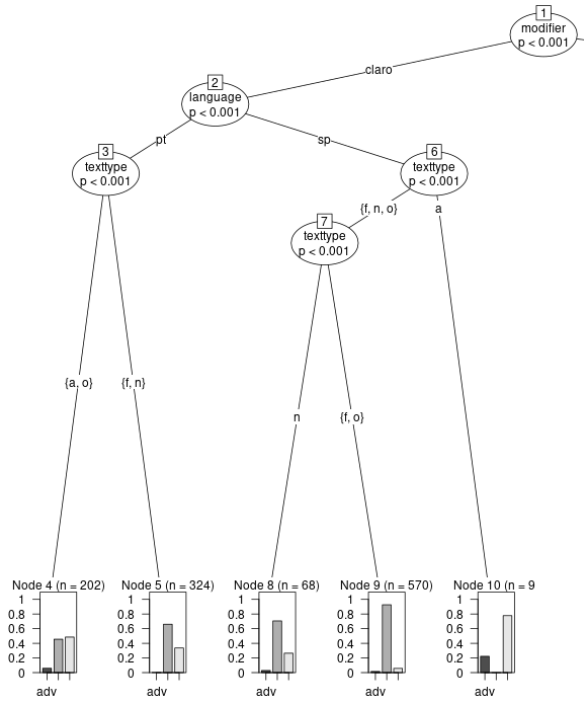
⁵³While there appears to be a difference in oral texts, there is in reality none to speak of. The tokens per TEXT TYPE in Portuguese are not equally distributed. In particular, the Portuguese oral subcorpus is roughly half the size of the Spanish one (cf. table 23).

hypotheses-driven criteria such as the plausibility and interpretability of the predicted effects. Figure 7 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 8 shows the importance of each of these variables, which was determined through a random forest. Finally, the heat map in 9 plots the observed vs. predicted values from table 30. 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 the table in 29 ($1137+790+409=2336$), by the sum of all of the predicted values ($2336/3329=0.701$).

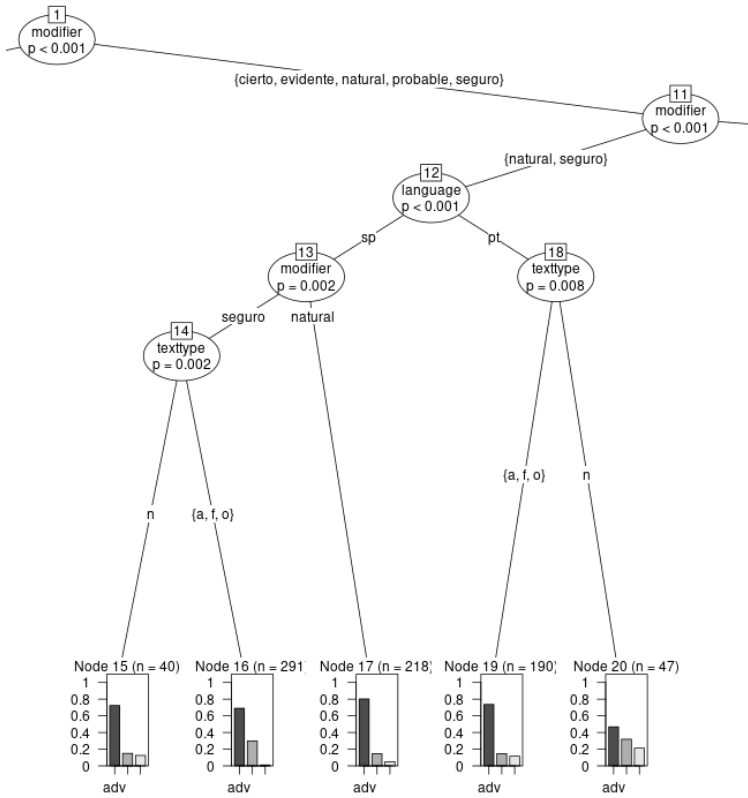
	observed Adv	observed AdvC	observed EsAdjC
predicted Adv	1137	225	232
predicted AdvC	12	790	160
predicted EsAdjC	201	163	409

Table 29: Predicted categories.

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 10(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 make up 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.



(a) Branch 1.



(b) Branch 2.

Figure 10: Individual Branches of the conditional inference tree in 7.

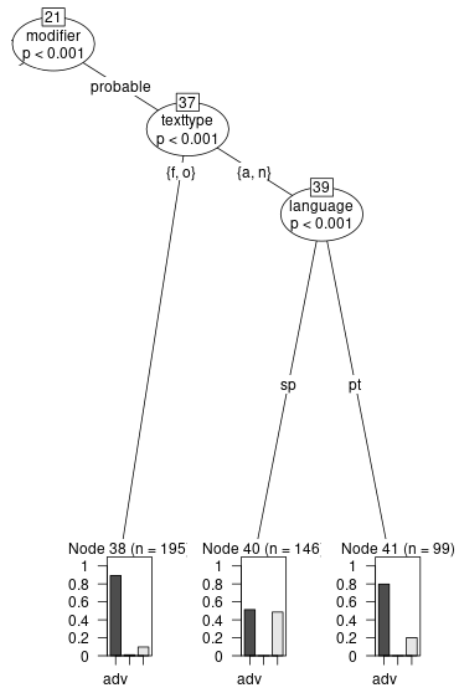
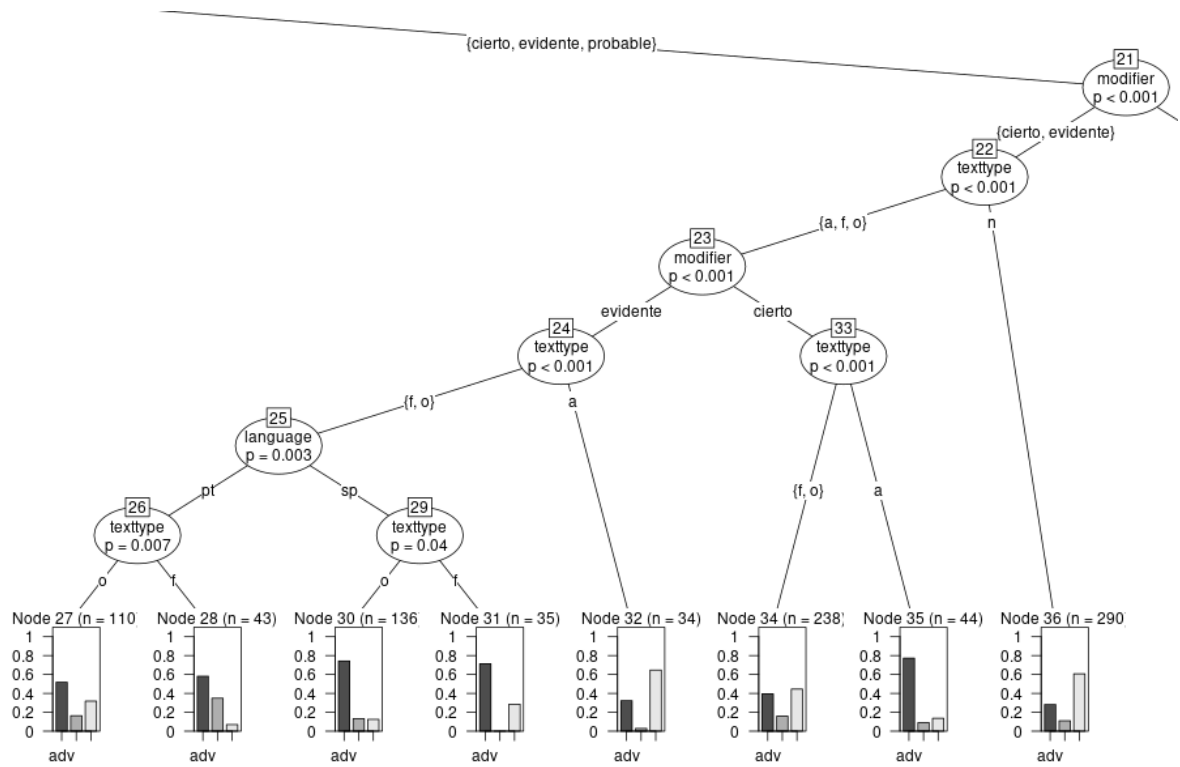


Figure 11: Individual Branches of the conditional inference tree in 7.

In the subset of data containing all MODIFIERS but *claro* (cf. figure 10(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, 11(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 36). 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 33) but not on LANGUAGE. In academic texts, *cierto* appears most frequently in Adv (Node 44), while in fiction and oral texts, it appears in all three CONSTRUCTIONS, but less frequently in AdvC (Node 34). In the *evidente* subset, there is a split separating academic from fiction and oral texts (Node 24). There is once again no split on LANGUAGE. In academic texts, EsAdjC is the most frequent CONSTRUCTION. There are also cases of Adv, but AdvC is absent from the data (Node 32). 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, 29).

Within the *probable* subset (Node 21, figure 11(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 37). 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.⁵⁴

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 the use in the AdvC CONSTRUCTION. It is possible that Spanish *claro que* grammaticalized to a 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. section 3.4). In the cases of Adv and EsAdjC, the wrongly predicted data points (EsAdjCs predicted as Adv or AdvC and Adv predicted as EsAdjC and 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 29 repeated in 30).

	observed Adv	observed AdvC	observed EsAdjC
predicted Adv	1137	225	232
predicted AdvC	12	790	160
predicted EsAdjC	201	163	409

Table 30: Predicted categories.

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 follows that the other MODIFIERS investigated in the present study are less specialized for one reading.

⁵⁴There are two occurrences in Spanish. See table 27

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 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*. This suggests that the latter three MODIFIERS are not as idiosyncratic as the other three. For *cierto*, LANGUAGE is not a significant predictor at all suggesting that, with respect to the tested CONSTRUCTIONS, the MODIFIERS has the same meaning in Spanish and Portuguese.

4.3.2 Deictic Centers of Spanish Epistemic and Evidential Modifiers

In the present section I describe the design and the results of an experiment set out to determine primarily the influence of the variables DEICTIC CENTER and CONSTRUCTION but also of other linguistic and extra-linguistic factors, on the acceptability of Spanish epistemic and evidential MODIFIERS in their respective contexts.

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.⁵⁵ I chose this corpus because of its large size and its informal register that approximates orality, which I consider necessary in eliciting judgments on natural language pragmatics.

⁵⁵See <https://www.corpusdelespanol.org/web-dial/help/textsm.asp> [February 28, 2019] for a detailed description of how the data were acquired and processed.

Data Acquisition and Selection I first determined the four most frequent epistemic and evidential modifiers in the corpus (*cierto*, *claro*, *evidente*, *seguro*). For each of these, I extracted all sentence-initial occurrences of the three CONSTRUCTIONS under investigation. Since sentence position is not annotated in the corpus, I again approximated this by extracting all cases of the relevant patterns following a sentence-final punctuation mark.

The final experimental stimuli were found through step wise 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.

	cierto	claro	evidente	seguro
Adv	3	3	3	3
Adv-menteC/Adv-0 C	3/1	3/0	1/2	3/0
EsAdjC/EstáAdjC	3/0	0/3	3/0	3/0

Table 31: Derived vs. underived adverbs in AdvC in the experimental stimuli.

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 31. Neither of the two variables had a significant effect in the model.

Data Modification The critical target sentences in the scope of the modifiers, were shortened so that all 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 context was selected

individually for each target sentence, with the selection done based on my own 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.

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. The stimuli were 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 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.

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 were 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.

Conditions In the experiment the target sentences are presented in three conditions. Each condition negates certain readings of the MODIFIER.

(255) a. (*yo*-DEICTIC CENTER)

Aunque yo no lo crea, claramente tiene los ojos de su papá.
although I not it believe.1S clearly has the eyes 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 ojos de su papá.
although you not it believe.2S clearly has the eyes 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, claramente tiene los ojos
although the people in general not it believe.3S clearly has the eyes
de su papá.
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 (255a). In the second condition (*tú*-DEICTIC CENTER), the hearer-centered reading is negated (255b). In the third condition (*gente*-DEICTIC CENTER), a reading where the MODIFIER is centered on a more general group of people is negated (255c).

The motivation behind 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 to 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 judgments for AdvC in the *tú*-DEICTIC CENTER. The third condition targets EsAdjC. The idea is that 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 low in this condition.

Description of the Data and Predictors Table 32 summarizes the percentage of judgments for each value on the rating scale. There are larger values at the extremes of the scale. This suggests that in most cases the participants had strong intuitions on the acceptability and unacceptability of the stimuli and they chose intermediate ratings, translating to unclear intuitions, less frequently.

rating	1	2	3	4	5
percentage of judgments	29.55	11.96	14.05	18.61	25.83

Table 32: Percentage of judgments per rating.

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 range from 20 to 145 words (mean: 62). The examples in (256) illustrate two extremes in terms of context length measured in words. In spite of these differences, the length of the contexts in words did not seem to create any 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).

- (256) 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...

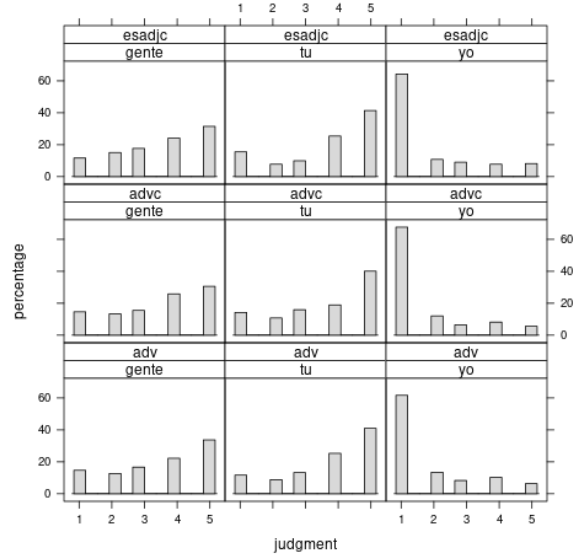
‘Speaker A – The goal of this seminar is to reflect on conflicts in the world (wars, catastrophes, terrorism) which, in the first hours after they take place, provoke an intense media following, but weeks after disappear from the pages of the newspapers and from our memories. Kapuscinzky said that he 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.’

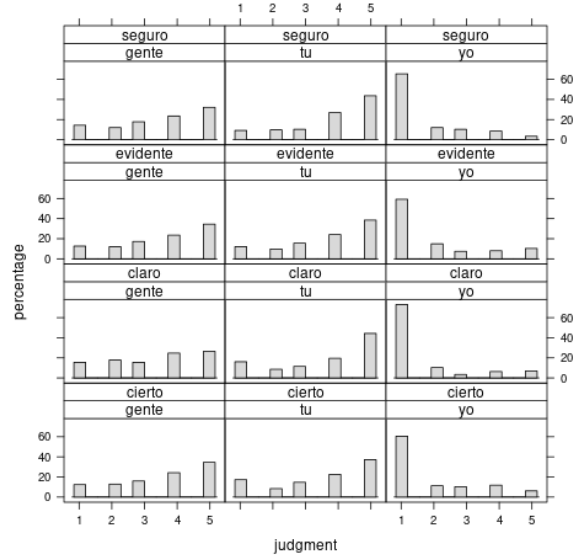
‘Speaker B – Although I don’t believe it, it is evident that Kapuscinzky didn’t manage to convince the directors of the media.’

The bar charts in 12(a), made using the `lattice` R-package, 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 comparable frequencies. They are distributed 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 this DEICTIC CENTER but are certain about the acceptability of the *tú*-DEICTIC CENTER. The bar charts in 12(b) 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 13 plot the percentage of judgments by CONSTRUCTION by MODIFIER.

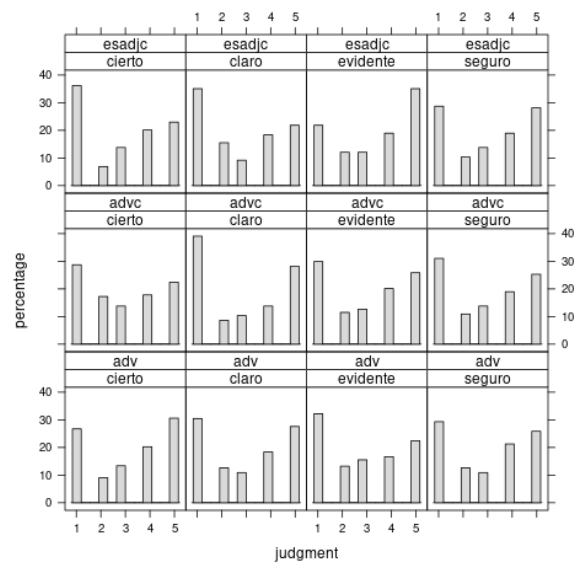


(a) Individual bar charts for judgments per DEICTIC CENTER per CONSTRUCTION. The panels in the first row show the judgments for EsAdjC in the different DEICTIC CENTER conditions. The second row shows AdvC and the third row Adv.



(b) 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*.

Figure 12: Bar charts for the distribution of the judgments across different variables.



(a) 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.

Figure 13: Bar charts for the distribution of the judgments across different variables.

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 edge 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 in the low 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.

Results As stated at the beginning, the model described in the present section was selected in an exploratory manner. I fitted a number of different models with varying degrees of complexity. Model selection was carried out based on objective measurements such as the comparison of the model's AIC and BIC⁵⁶ and the accuracy of the model's predictions. The fitting and the selection, however, were also strongly determined by a priori hypotheses and the theoretical plausibility and interpretability of the model's effects. I settled on the present, rather simple, model. A more complex model, containing a non-linear factorial interaction, displayed comparable accuracy, AIC and BIC as the present model, but the additional effects supported by the model did not contribute to the understanding of the questions I set out to answer.

⁵⁶The *Akaike Information Criterion* and the *Bayesian Information Criterion* are measures that are used to compare the goodness of fit of nested models. The values of the AIC and BIC tell us whether the addition of a new predictor results in a better fit of the model compared to a simpler model. AIC and BIC are based on the maximum likelihood estimates and penalize for the inclusion of more predictors. This is the case because a model with a larger number of predictors almost always has a better fit than a simpler model. The penalization of the AIC is independent of the sample size. This results in an advantage for models with a high number of predictors fitted to large samples. With the BIC the penalization of additional factors is larger when the sample size grows.

A. parametric coefficients	Estimate	St. Error	t-value	p-value
(Intercept)	0.9998	0.1710	5.847	<0.0001
Deictic Center-tu	0.3676	0.1503	2.445	0.0146
Deictic Center-yo	-2.4813	0.1550	-16.013	<0.0001
B. smooth terms	edf	Ref.df	F-value	p-value
s(Item)	61.43	108	1.452	<0.0001
s(Participant)	54.68	60	8.304	<0.0001

Table 33: Model summary of the GAM fitted to the ordered categorical acceptability rating with a main effect of DEICTIC CENTER. The terms s(Item) and s(participant) denote the by-ITEM and by-PARTICIPANT random intercepts.

The model contains a linear effect of the variable DEICTIC CENTER and two random effects for ITEM and PARTICIPANT. The main effect shows that all three DEICTIC CENTERS have a significant effect on the judgments. There is a significant difference between the *gente*-DEICTIC CENTER, which negates the DEICTIC CENTER on a general group of people and the *tú*-DEICTIC CENTER, which negates the DEICTIC CENTER on the interlocutor. The *tú*-DEICTIC CENTER gives rise to a significantly higher judgment, but the effect is relatively small. On the contrary, the effect of the *yo*-DEICTIC CENTER, the negation of the speaker’s belief, is strongly negative.

The two random effects are also significant. A random effect for ITEMS is included due to the heterogeneity of the corpus-sampled stimuli. Other ITEM-dependent variables, other than the DEICTIC CENTER, did not show any conclusive effects. The figures in 14 indicate a lack of structure in the by-ITEM intercepts.

The random effect for PARTICIPANTS is included because we expect individual differences in judgments on the stimuli. Other PARTICIPANT-dependent variables such as AGE, GENDER or COUNTRY OF ORIGIN did not show any conclusive effects on the judgments. The figures in 15 indicate no further structure in the by-PARTICIPANT intercepts.

The plots in 16 show that the residuals (observed-predicted values) do not have a normal distribution. Since I modeled an ordered categorical variable, the distribution does not belong to the exponential family, and there are no conditions on the distribution of the residuals. The plots show that there are residuals on the extremes but none in the center, which suggests that the predicted values probably fall into two categories and not five.

The plots in 17 suggest that the accuracy of the prediction is not very high in this model. For the intermediate observed values, there is hardly any difference between

the corresponding fitted values, and they span a large interval. The fitted values for the observed values 1 and 5 also cover large intervals, but the differences between the two are clearly visible. The plot in 17(b) shows again that the model predicts well at the extremes but not for the intermediate values. The residuals approach 0 at the largest and smallest fitted values, which shows that the accuracy of prediction at the extremes is high. At the intermediate parts of the fitted values, there are high positive and negative values for the residuals. For the lowest extreme, there are a few outliers that have very high positive residuals, which represent the data points that have higher observed values but are wrongly fitted to a low value. For the highest extreme, the reverse pattern is observed.

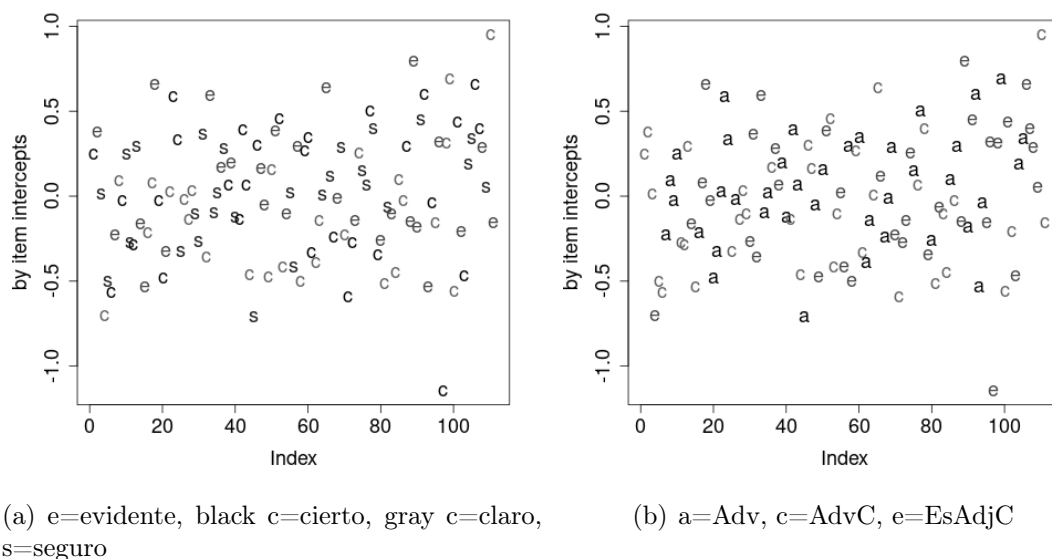
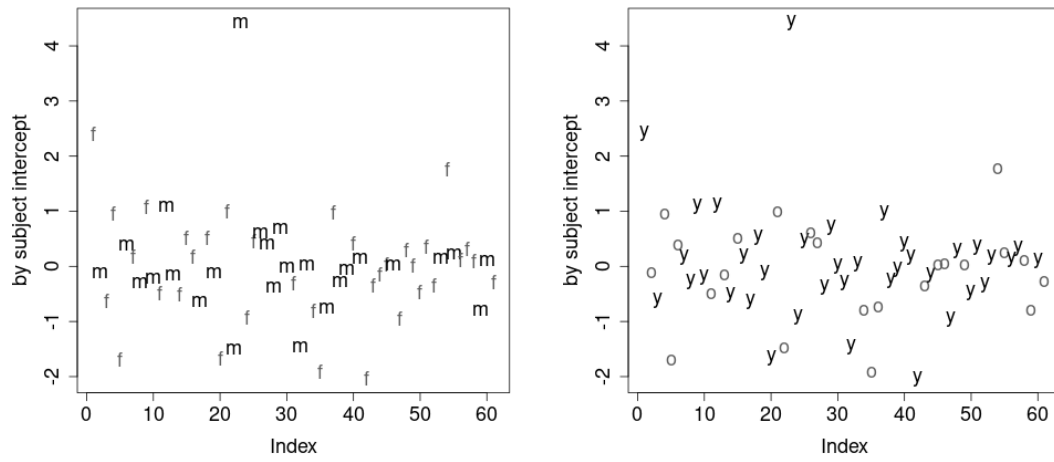
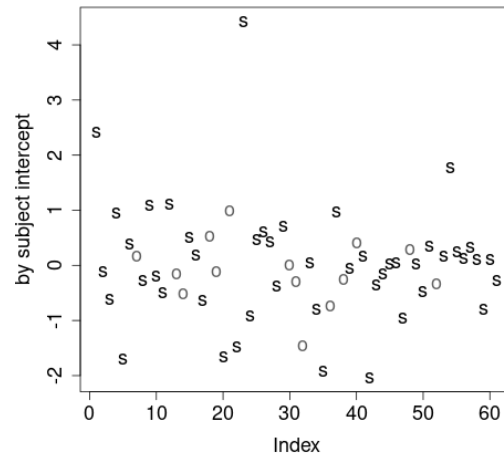


Figure 14: By-ITEM intercepts with ITEM-dependent variables.



(a) f=female, m=male

(b) o=old, y=young



(c) s=Spain, o=other

Figure 15: By-PARTICIPANT intercepts with PARTICIPANT-dependent variables.

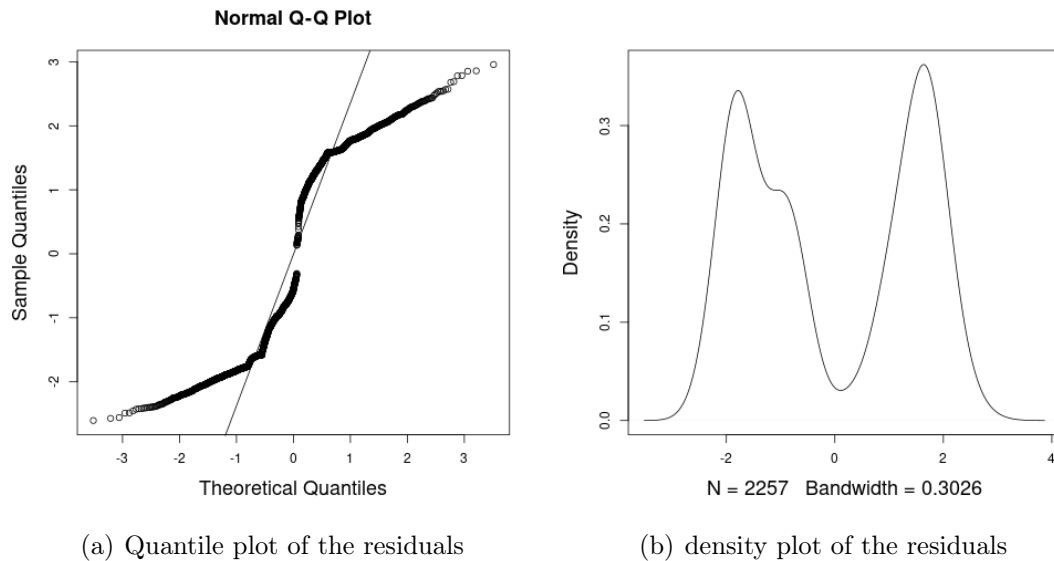


Figure 16: Quantile and density plot of the residuals of the model.

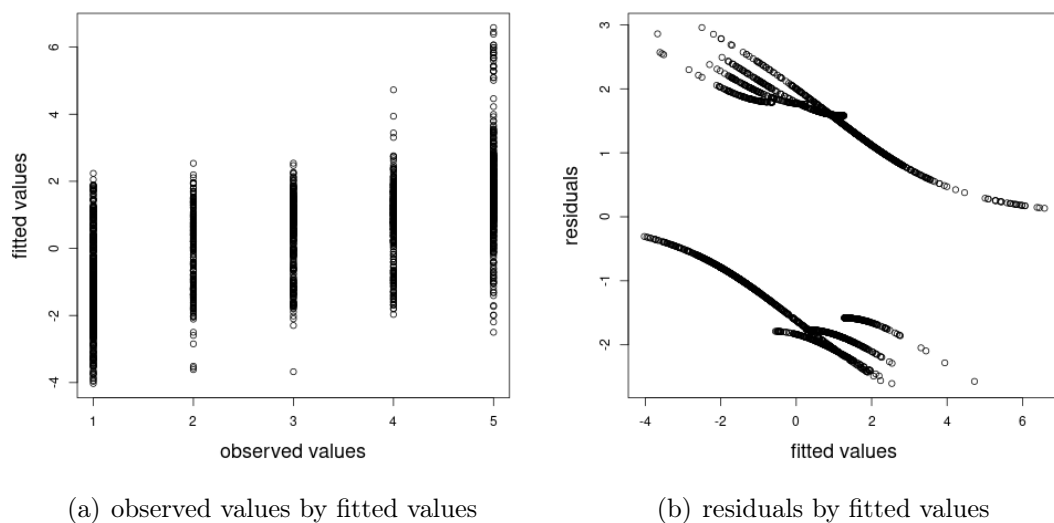


Figure 17: Correspondence of fitted values vs. observed values and fitted values vs. residuals.

Translating the numeric predicted values back to categorical ratings, it emerges that the predictors in the model only support a three-point scale and not a five-point scale. The table of raw judgments in 32 also shows that most judgments are found at the extremes of the scale. The heat map in 18 shows that the two lowest and highest observed values

are mostly fitted to the values 1 and 5, while the observed value 3 is fitted to either of the two.

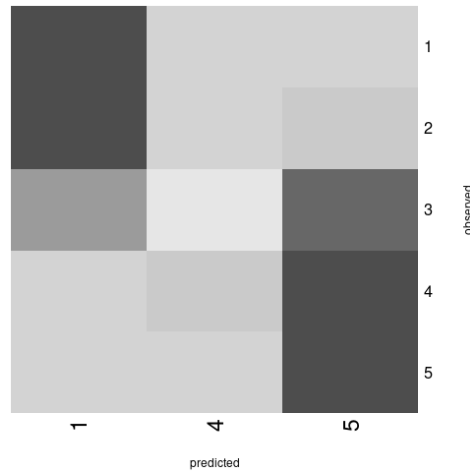


Figure 18: Heat map of 3 predicted versus 5 observed values. Darker shades of gray correspond to larger counts.

The deviance explained (50.1%) and the accuracy (33%) of the model are moderate. The accuracy was calculated by dividing the sum of the correctly-predicted values by the sum of all the predicted values. The low accuracy therefore results from the fact that the model predicts only 3 (cf. figure 18) categories while the observed values fall into 5 categories. This gives rise to a high number of wrongly-predicted values.

A conditional inference tree (cf. figure 19(a)) with splits on DEICTIC CENTER, AGE and COUNTRY-PARTICIPANT (nationality of the participants) achieves a slightly higher accuracy (38%). The other linguistic variables CONSTRUCTION and MODIFIER did not show any significant effect here either. The tree shows complex interactions between DEICTIC CENTER and AGE and DEICTIC CENTER and COUNTRY-PARTICIPANT, which were not detectable in the GAM.

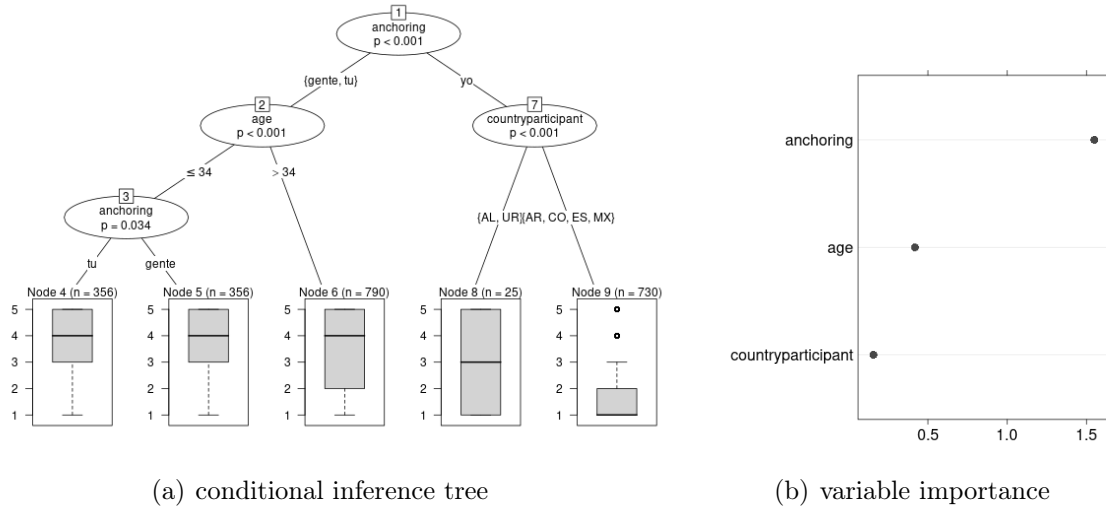


Figure 19: Conditional inference tree with splits on DEICTIC CENTER, AGE and COUNTRY-PARTICIPANT and a dot plot showing the importance of each variable.

The dot plot in 19(b) shows the variable importance that was determined using a random forest. Just as in the GAM, DEICTIC CENTER is by far the most important variable in this model.

Discussion The results show a solid effect for DEICTIC CENTER. Negating the speaker’s perspective is always judged low in combination with the MODIFIERS chosen in this experiment. Although this is a significant effect in the model, the figure in 12(a) suggests that negating the speaker’s perspective might be even worse in AdvC. This could indicate that in this CONSTRUCTION the speaker’s perspective is necessarily present. 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”. Negating the interlocutor’s perspective is more acceptable in the present experiment. The perspective of the interlocutor is built into the data, 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.

Contrary to the expectations formulated based on theoretical analysis, the model does not support an effect of CONSTRUCTION, let alone 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 unnaturalness of the sentences in 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.

The model supports two random effects. The random effect of ITEM is expected due to the heterogeneity of the corpus-based stimuli. The random effect of PARTICIPANT is also expected, because, with human ratings, PARTICIPANT-based variation is very common.

The distribution of the judgments is strongest at the extremes of the scale. This results in a model, in which the five-point scale is reduced to a three point scale. There are several possible reasons behind this particular response behavior. One explanation could be that the participants had very clear judgments on what is acceptable and what is not. The intermediate steps on the scale were not necessary to report their judgments. From a more pessimistic perspective, it could be the case that the intermediate spaces on the scale were not interpretable for the participants. It appears to have been clear what constituted an odd-sounding example and a natural-sounding one, but an intermediate example might have been less clear. Additionally, the scale only had labels at the left and right extremes (*extraño* and *natural*), which might have made this effect even larger.

4.3.3 Deictic Centers of Catalan Epistemic and Evidential Modifiers

The Catalan experiment that will be described in this section is again carried out with aim of determining the factors that influence the acceptability of the MODIFIERS in the three CONSTRUCTIONS.

Corpus The stimuli were sampled from the caWac corpus (Ljubešić & Toral 2014), 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 of web data.

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 step wise 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’, *errori* ‘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 puguí 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.

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. table 34 and 35). They were randomly assigned to the experimental stimuli. Neither of the two variables had any effect in the model.

	cert	clar	evident	segur
Adv- <i>ment</i> C	1	0	3	0
Adv- <i>0</i> C	2	3	0	3

Table 34: Derived vs. underived adverbs in AdvC in the experimental stimuli.

	cert	clar	evident	segur
<i>ser</i>	3	2	3	2
<i>estar</i>	0	1	0	1

Table 35: *ser* vs. *estar* in EsAdjC in the experimental stimuli.

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 *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.

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 is in a Latin square design. Each stimulus was presented to each participant in only one of the two conditions. The stimuli were 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).

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.

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 of whom were 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.

Conditions The stimuli are presented in two conditions. Each condition negates one reading of the MODIFIER. The first condition (257a) negates the reading in which the MODIFIER is centered on the interlocutor (*tú*-DEICTIC CENTER). The second condition (257b) 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.

(257) a. (*tú*-DEICTIC CENTER)

Encara que no t’ho creguis, és evident que aquest equip també
 Although that not you-it believe.2S is evident that this team also
 pot perdre.
 can.3S 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, és evident que aquest
 Although that not the people them-it believe.3S is evident that this
 equip també pot
 team also can.3S lose

‘Although the people don’t believe it, it’s evident that this team can also lose.’

The motivation behind these conditions 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. The aim of the current experiment is to see whether the changes in the set up of this experiment makes it possible to find the expected effects.

Description of the Data and Predictors The judgments of the participants were provided on a ten-point Likert scale. Table 36 shows the percentage of the participants’ judgments per value of the scale. They fall mostly at the higher edge 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.

rating	1	2	3	4	5	6	7	8	9	10
percentage of judgments	12.68	9.66	7.6	6.16	6.16	3.62	7.97	13.29	13.53	19.32

Table 36: Percentage of judgments per rating.

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 again 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. (258a)) to 751 (cf. (258b)) words per context.

- (258) a. **Context:** Per què les crispetes són tan cares en el cinema?

‘Why are 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 estalviï 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 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 is 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.

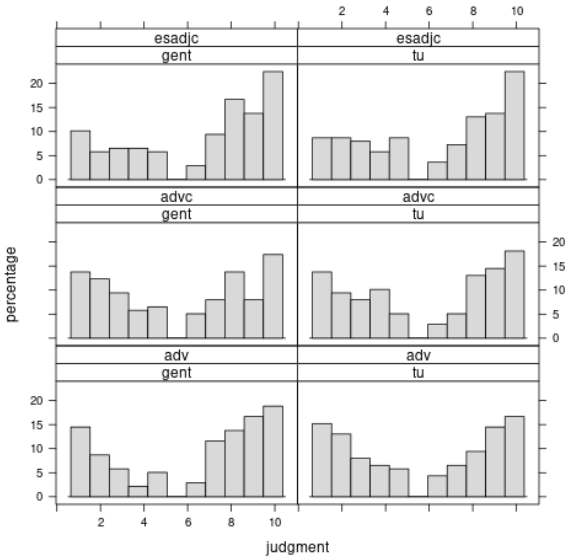
‘Although you don’t believe it, surely Company B is the Catalan Health Department and this is the reality of many Catalan hospital directors.’

The participants were not instructed to move fast or slow through the experiment. Consequently, 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 20 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 12(a)). 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. We should not therefore expect to find a linear relation between the acceptability and these two CONSTRUCTIONS.

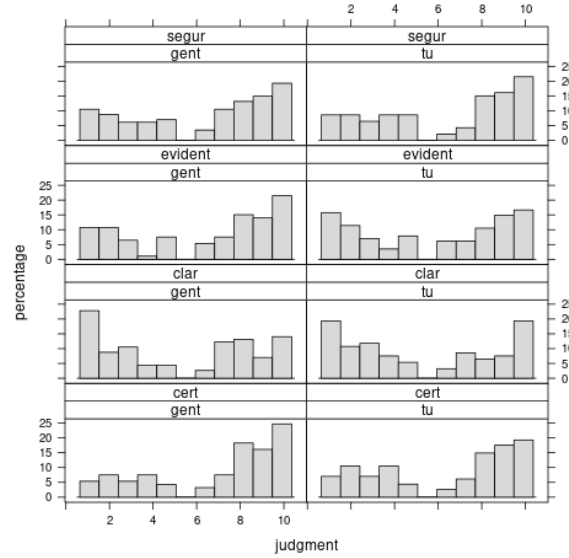
The plots in 21(a) show the judgments for each MODIFIER in the two DEICTIC CENTER conditions. They indicate again that DEICTIC CENTER does not have a strong effect on the judgments in the Catalan experiment. Most of the bar charts are symmetrical with high percentages at the extremes of the scale and 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 bigger 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.

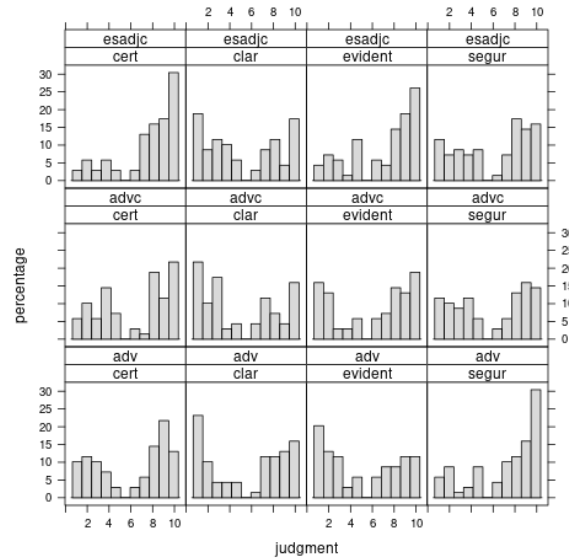


(a) Individual bar charts for judgments per DEICTIC CENTER per CONSTRUCTION. The panels in the first row show the judgments for EsAdjC in the different DEICTIC CENTER conditions. The second row shows AdvC and the third row Adv.

Figure 20: Bar charts for the distribution of the judgments across the different variables.



(a) Individual bar charts for judgments per DEICTIC CENTER per MODIFIER. The panels in the first row show the judgments for *secur* in the different DEICTIC CENTER conditions. The second row shows *evident*, the third row *clar* and the last row *Cert*.



(b) 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*.

Figure 21: Bar charts for the distribution of the judgments across the different variables.

Finally, the last set of bar charts in 21(b) 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 disliked (*evident* and *clar* in Adv).

Results Model fitting and selection was again exploratory. I fitted a number of different models relying on hypotheses derived from the theory and the inspection of the data. I settled on the present model based on objective measurements (AIC, BIC and accuracy of the model) and its theoretical plausibility.

The model supports two main effects, of MODIFIER and CONSTRUCTION, and one random effect of ITEM. It furthermore contains an interaction effect with a smooth of NCONTEXT (context length measured in number of words) by MODIFIER.

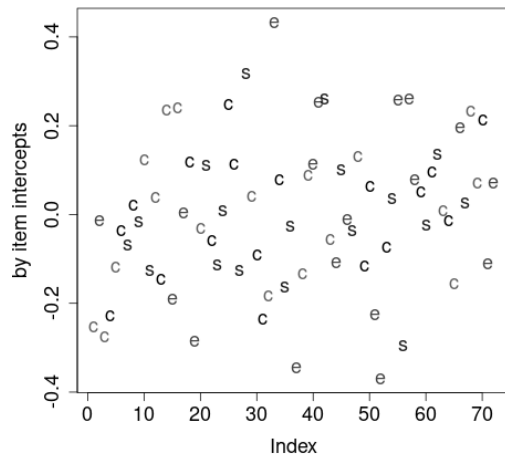
A. parametric coefficients	Estimate	St. Error	t-value	p-value
(Intercept)	1.1306	0.2067	5.471	<0.0001
Modifier-clar	-0.6807	0.2250	-3.026	0.00256
Modifier-evident	-0.2842	0.2403	-1.183	0.23732
Modifier-segur	-0.1796	0.2259	-0.795	0.42682
Construction-AdvC	0.1030	0.1953	0.527	0.59826
Construction-EsAdjC	0.5141	0.2012	2.556	0.01078
B. smooth terms	edf	Ref.df	F-value	p-value
s(Item)	17.809	62.000	0.403	0.0246
s(Ncontext):Modifier-cert	1	1	0.004	0.9497
s(Ncontext):Modifier-clar	1	1	6.168	0.0132
s(Ncontext):Modifier-evident	1	1	0.094	0.7590
s(Ncontext):Modifier-segur	2.001	2.326	3.139	0.0284

Table 37: Model summary of the GAM fitted to the ordered categorical acceptability rating with a main effect of MODIFIER and CONSTRUCTION. The term s(Item) denotes the by-ITEM random intercepts. The terms s(NCONTEXT):MODIFIER denotes smooth functions of the interaction of the individual MODIFIERS with the length of the context measured in number of words.

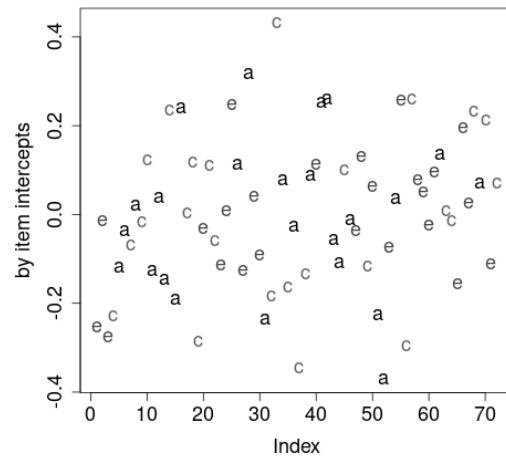
Clar is the only MODIFIER that has a significant effect compared to the reference level *cert*. The effect is negative. Stimuli containing *clar* were judged significantly lower than those with *cert*. Among the CONSTRUCTIONS, EsAdjC is a significant positive predictor when compared to Adv. There is no significant difference between Adv and AdvC. The model contains a significant random effect for ITEM. The plots in 22 show

that there is no further structure in the by-ITEM random intercepts. Random intercepts for PARTICIPANTS were included in an alternative model, but they were not significant.

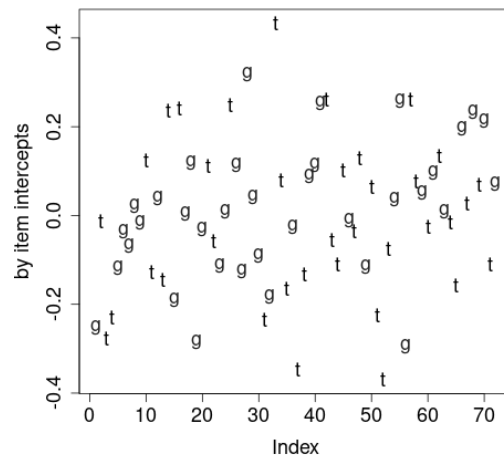
The model also contains interactions in the form of smooths for NCONTEXT (context length measured in number of words) by MODIFIER. While context length has an overall positive effect with all the MODIFIERS, the strength and significance of the effect differs depending on the MODIFIER. The effect is only significant with the two epistemic MODIFIERS *clar* and *segur*. It is nowhere near the level of statistical significance for *cert* and *evident*, which can also be seen in the very mild slope of the effects in 23. The effect is linear with *cert*, *clar* and *evident* and non-linear with *segur*. The effect curve is very steep and positive until around 500 words, at which points it reverses the direction and becomes negative.



(a) e=evidente, black c=cierto, gray c=claro, s=seguro



(b) a=Adv, c=AdvC, e=EsAdjc



(c) t=tú, g=gente

Figure 22: By-ITEM intercepts with ITEM-dependent variables.

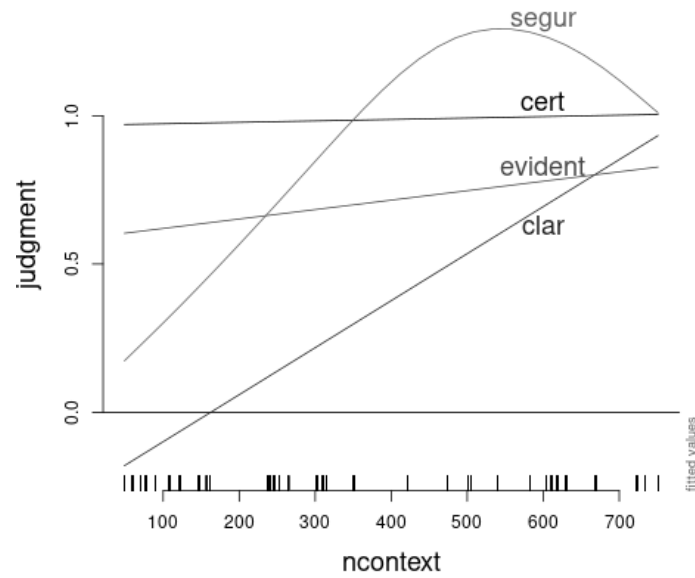


Figure 23: Contrasting linear and non-linear effects of context length per MODIFIER on the judgments.

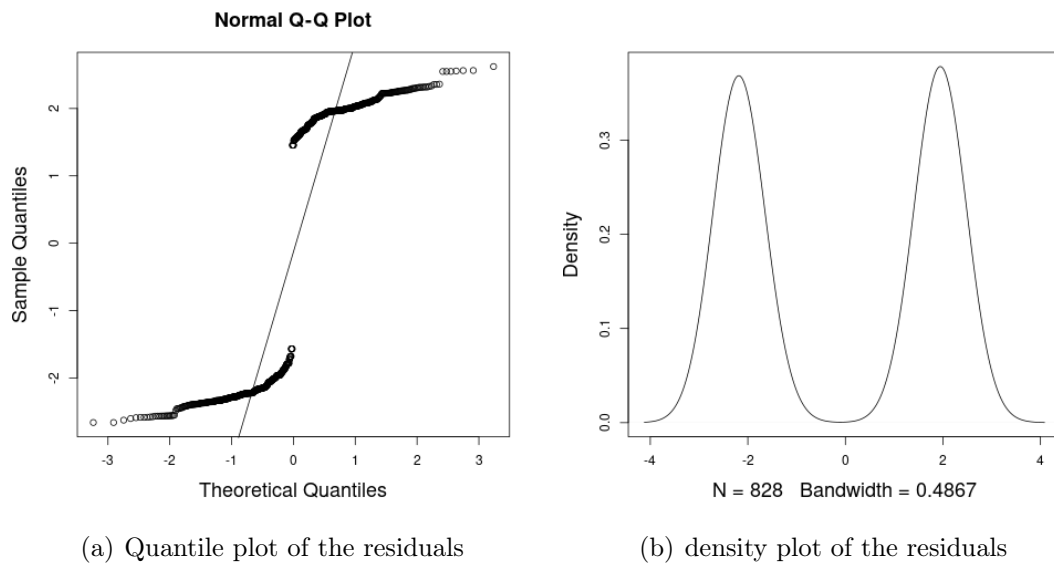


Figure 24: Quantile and density plot of the residuals of the model.

The residual plots for the present model are even more extreme than those corresponding to the Spanish study (cf. the figures in 16). The residuals form two peaks at the

extremes 24(a), 24(b) in the density plot. This indicates once again that the scale on which the judgments were elicited was reduced drastically in the model. The residuals do not have a normal distribution, which is also not unexpected given the ordinal nature of the response variable.

The plots in 25 suggest that the model moderately fits the data. The plot in 25(a) shows that there is no close correspondence between fitted and observed values. In particular it shows that the fitted values for each observed value span practically across the entire range of the fitted values. The second plot in 25(b) shows substantial positive and negative residuals for all the fitted values.

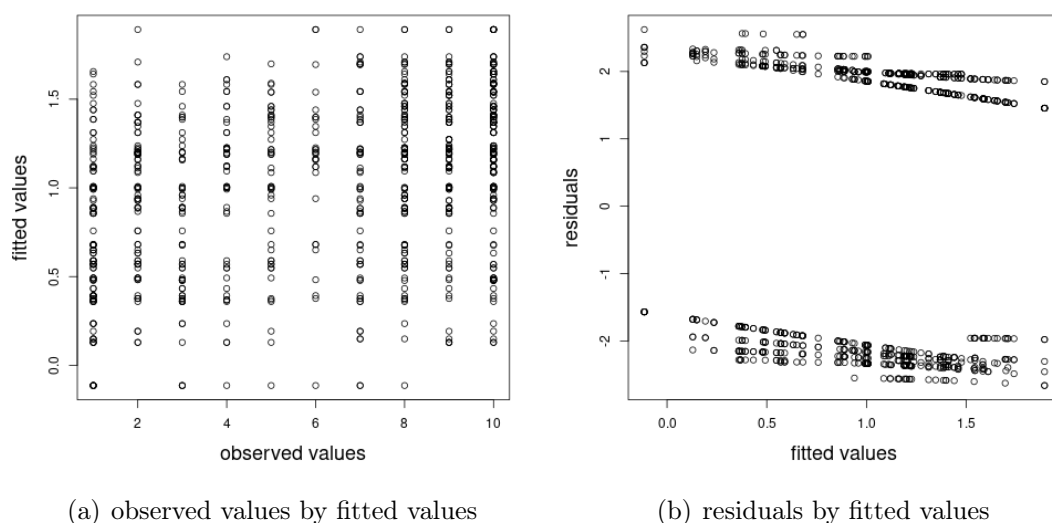


Figure 25: Correspondence of fitted vs. observed values and fitted values vs. residuals.

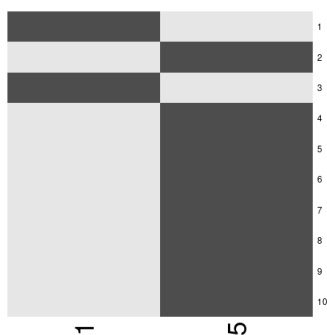


Figure 26: Heat map of 2 predicted values versus 10 observed values. Darker shades of grey correspond to larger counts.

The deviance explained (56.9%) is moderate but larger than that observed for Spanish. However, the accuracy (13,64%) is very low. This is once again related to the way in which accuracy was calculated and the fact that the model predicts a reduced scale. In the present case, the outcome is even more drastic than in the previous experiment, because the model reduces the 10 values of the observed variables to just a two level opposition in the predicted values (cf. 26). The observed values 1 and 3 are fitted to the value 1, whereas all the others, strikingly also the value 2 is fitted to the value 5.

The conditional inference tree (accuracy: 10.14%) and the dot plot in 27 support the same variable effects as those in the GAM. It shows that the strongest effect is that of CONSTRUCTION, which splits EsAdjC from Adv and AdvC. The effect of CONSTRUCTION, however, only obtains within the group of MODIFIERS other than *clar*. The conditional inference tree also reveals a three-way interaction between MODIFIER, CONSTRUCTION and NCONTEXT within the same group, which was impossible to detect in the GAM.

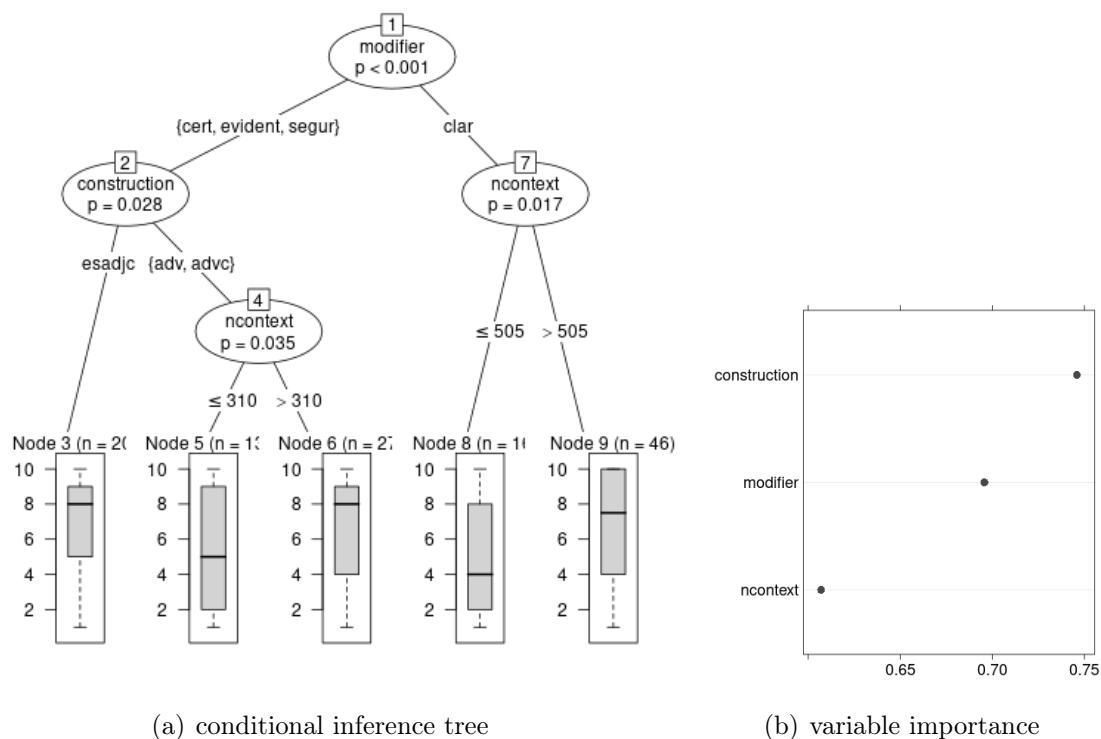


Figure 27: Conditional inference tree with splits on CONSTRUCTION, MODIFIER and NCONTEXT and a dot plot showing the importance of each variable.

Discussion The GAM and the ctree suggest 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 particular. Interestingly, there seems to be a mismatch in the comprehension and production of Catalan *clar* in the different CONSTRUCTIONS. 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). Conversely, AdvC is very frequent in caWac (2303 cases), while in comparison, the highest-rated Adv only occurs 387 times in caWac. The highest frequency is for *clar* in EsAdjC (7410, of which 5062 contain the copula *ser*). It is possible that some of these 5062 cases could actually be analyzed as AdvC rather than EsAdjC. This analysis is plausible given the fact that 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 section 3.3.4). *Esclar* does not yet appear in DIEC (*Diccionari 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 shows the fact that *és clar* and *esclar* have different functions and interpretations is acknowledged.⁵⁷ 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*.⁵⁸ 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. Conversely, it is also possible that some of 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, this could mean that epistemically or evidentially modified statements are bad in out-of-the-blue contexts. In the present experiment the model yet again predicts fewer values for the response variable than the original scale. An open question remains regarding 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

⁵⁷See for instance https://www.ara.cat/cronica/Aglutinacio-clar-esclar_0_444555554.html [February 28, 2019]

⁵⁸<http://www.elpuntavui.cat/article/7-vista/23-lectorescriu/201549-esclar-o-es-clar.html> [February 28, 2019]

were not interpretable for the participants. The moderate fit of the model and the lack of significant effects is probably the result of 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 an 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. On the one hand, this shows that there is a striking mismatch between production and comprehension. On the other hand, the lack of significant results in the experiment suggest that this relation cannot be tested in the manner attempted in the experiments. While there might be a preferred DEICTIC CENTER for the MODIFIERS in each CONSTRUCTION, other readings can easily be accommodated. 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 section 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 shortly, 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. section 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. This suggests that there is little awareness of this construction. It is very possible that 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.

Corpus Just as in the experiment described in section 4.3.3, the stimuli were sampled from the caWac corpus (Ljubešić & Toral 2014), which comprises 780 million tokens and was built through a web crawl.

Data Acquisition and Selection The caWac corpus is freely available to download. Apart from sentence and paragraph splits, the data contain no other annotations. The final experimental stimuli were again found through step wise 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.

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. This is possible because for every question that was originally contained *que*, there is also a *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. (259)). 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.

(259) **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 than quantitative criteria with the aim of maintaining the minimum number of words necessary to make sense of the question.

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. The degree of expectedness addresses directly the aims of the study. 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, this is why I settled on the second option. The mean

duration of the experiment is 18 minutes.

Participant A total of 46 native Catalan speaker participated in the experiment. Recruitment through social media, which I relied on in the previous experiments, was less successful in this instance (it only resulted in the recruitment of 12 participants). This is why the additional 34 participants were recruited via Prolific,⁵⁹ 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. 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. One of the reasons for this is that the number of native Catalan speakers registered on Prolific is very small. In fact, there are currently no more than the 34 participants I managed to recruit. 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. The first platform 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.

Conditions The stimuli are presented in four conditions resulting from the combinations of the two counterbalanced variables PRESENCE OF QUE and ANSWER (cf. table 38).

	<i>que</i>	<i>no que</i>
affirmative answer	C1	C2
negative answer	C3	C4

Table 38: Variables and conditions.

There are four versions of each item. This is illustrated in 39 for the stimuli in (259). In condition 1, the question is introduced by *que* and the answer is affirmative. In

⁵⁹<https://www.prolific.ac/>

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.

	<i>que</i>	no <i>que</i>
context	A – Estàs pensant en marxar a estudiar a l'estranger però en tens alguns dubtes?	
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.

Table 39: Exemplified conditions.

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* are a tool for the speaker to express that he/she expects a positive answer. This means that C1, where a *que*-question is answered affirmatively, should lead to high judgments reflecting the fact that a positive answer is expected. However, in C3, where the answer is negative, we should expect a low degree of expectedness on the part of speaker A, which should translate to low judgments. 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. Therefore the expectation in a hypothetical laboratory example would be that there should not be any difference between the judgments for C2 and C4 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 one could expect in a highly controlled hypothesis-driven experiment. However, this is not how things work with the exploratory design that I have employed here. Given the fact 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.

Testing these factors and their interplay with the complementizer is possible 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.

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. 40).

rating	1	2	3	4	5
percentage of judgments	23.21	12.45	12.61	13.75	37.99

Table 40: Percentage of judgments per rating.

The participants were instructed to choose 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. Possibly, 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. This interpretation of the raw data might be supported by the fact that the statistical model, which is described in what follows, only predicts two values. Crucially, all the observed values except the value 1 are fitted to the higher value in the model. This could suggest that the participants did not make a gradual distinction between different degrees of expectedness but merely between unexpected and the rest. Given the fact that the stimuli are drawn from corpus data, they are 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.

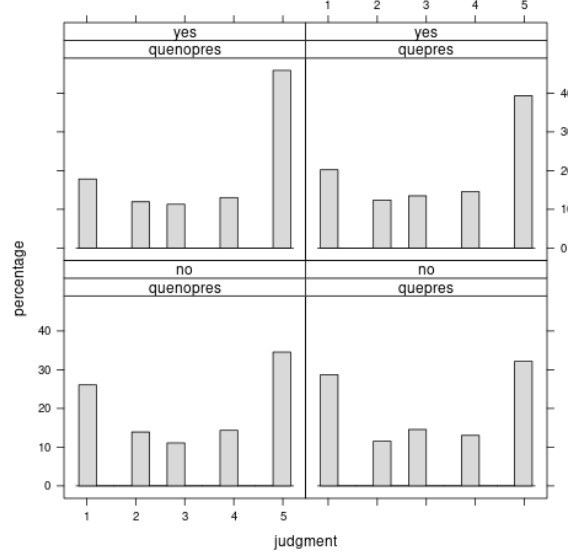
Different predictors were considered for the statistical model. The counterbalanced variable, PRESENCE OF QUE, is not a significant predictor. On the contrary, the variable ANSWER is significant. The bar charts in 28(a) show the percentage of the raw

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 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%).

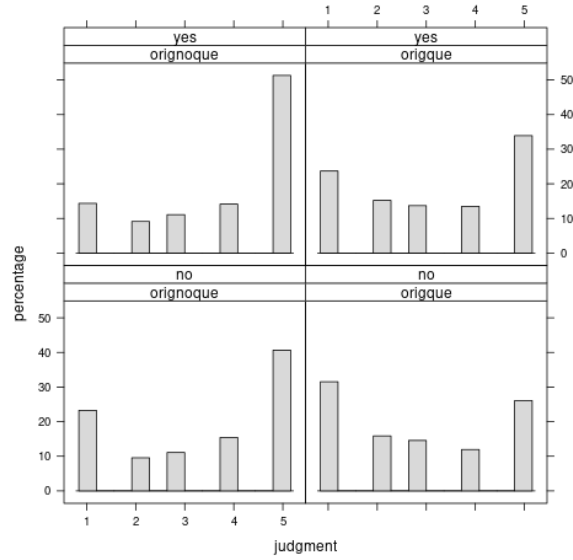
The observed contrasts are even more pronounced in the bar charts in 28(b), which plot the percentages of judgments per ANSWER by ORIGINAL, instead of PRESENCE OF QUE, by ANSWER. The variable ORIGINAL has two values: ORIGINAL-QUE, if the original question contained *que* and ORIGINAL-NOQUE, if it did 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 expected 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. Notable again is the high degree of unexpectedness of a negative answer in the ORIGINAL-QUE contexts.

The bar charts in 29(a) 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 one is called REALSPEAKER. This variable 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.

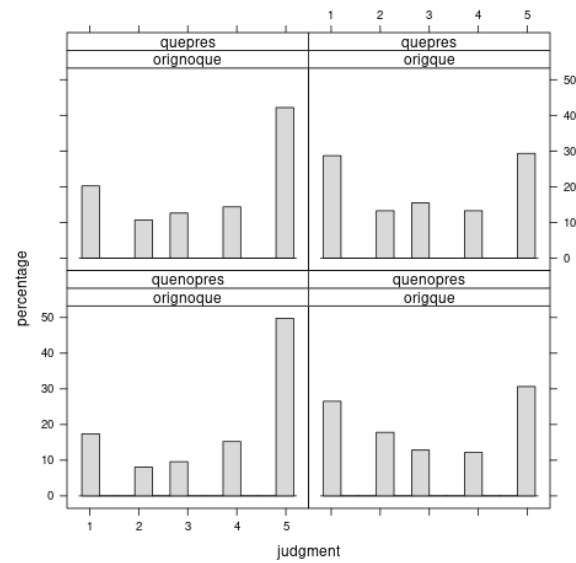


(a) Individual bar charts for judgments per PRES-ENCE OF QUE per ANSWER. The panels in the first row show the judgments when the answer is positive depending on the presence of *que*. The second row shows the judgments when the answer is negative.



(b) 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.

Figure 28: Bar charts for the distribution of the judgments across the different variables.



(a) 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.

Figure 29: Bar charts for the distribution of the judgments across the different variables.

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. The motivation behind this variable is that word order and modal expressions also result in a biased reading. In (260) a stimuli 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.

- (260) (Que) l'Àsia parla igual català que polonès?
 QUE the Asia speaks same Catalan as Polish
 'Asia speaks Catalan and Polish equally well? (ORIGINAL-NOQUE)

Similar suggestions were made for modal expressions. In (261), the modal adverb *potser*, could itself signal a bias.

- (261) (Que) potser s' hagués pogut reconduir la cosa?
 QUE maybe CL.REFL had could rerout the issue
 'Could the issue maybe have been resolved otherwise?' (ORIGINAL-QUE)

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.

Results The model described in this section, like the previous models, was established through exploratory selection based on objective measurements (AIC, BIC, accuracy) and the theoretical plausibility of the model's effects.

A. parametric coefficients	Estimate	St. Error	t-value	p-value
(Intercept)	0.5758	0.1784	3.228	0.001
Original-que	-0.6791	0.1717	-3.955	<0.0001
Answer-yes	0.3928	0.1706	2.303	0.02142
Realspeaker-real	-0.3444	0.3995	-0.862	0.38885
Answer-yes:Realspeaker-real	1.6008	0.5679	2.819	0.00488
B. smooth terms	edf	Ref.df	F-value	p-value
s(Item)	106.90	155	2.294	<0.0001
s(Participant)	35.49	45	3.955	<0.0001

Table 41: Model summary of the GAM fitted to the ordered categorical acceptability rating with a main effect of ORIGINAL, ANSWER and REALSPEAKER and the interaction effect of REALSPEAKER:ANSWER. The terms s(Item) and s(Participant) denote the by-ITEM and by-PARTICIPANT random intercepts.

The selected model is summarized in table 41. It supports two significant linear main effects of ORIGINAL and ANSWER. REALSPEAKER is not a significant predictor on its own but in interaction with ANSWER it is a significant predictor. There are also two random effects for ITEM and PARTICIPANT. The variable ORIGINAL turned out to be the strongest linear predictor in the model. What this means is 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. This means in these contexts that any type of answer was significantly less expected than in the ORIGINAL-NOQUE contexts. Crucially, in this model, the effect is independent of whether *que* actually is present in the stimulus or not and also independent of what the answer is.

The second significant linear predictor is ANSWER: there are significantly higher judgments for stimuli with a positive answer. This means that, across the board, there is a significantly higher expectation of a positive than a negative answer. This suggests that there is a general positive bias, irrespective of the type and properties of the polar question. The interaction between ANSWER and REALSPEAKER is also a predictor in the model. In particular, there is a significant difference between contexts where the answer is positive and the speakers are encoded through proper names on the one hand and contexts where the answer is negative and the speakers are encoded merely through the letters A and B on the other hand.

There are two significant random effects supported by the model. The random effect for PARTICIPANT is expected because with human judgments, individual differences are very common. Other PARTICIPANT-dependent variables like AGE, GENDER and REGION

showed no significant effects in the model. The figures in 30 show that there is no further structure in the by-PARTICIPANT intercepts. The random effect for ITEMS is motivated by the heterogeneity of the corpus-sampled stimuli. Further ITEM-dependent variables that are not already contained in the model, such as PRESENCE OF QUE and DIALOG encoding the number of interlocutors did not show any significant effects in the model. The figures in 31 show the lack of structure in the by-*item* intercepts.

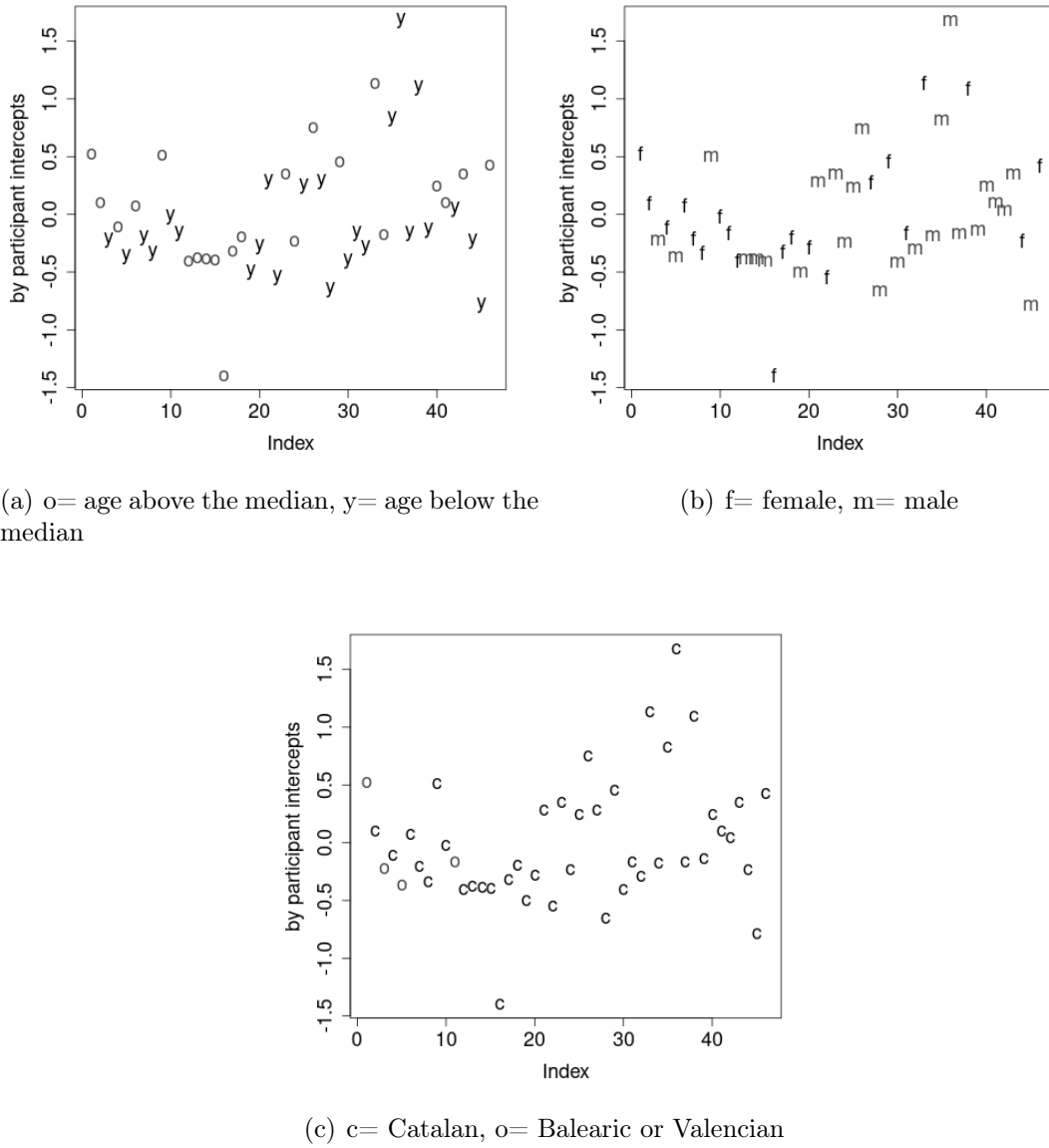
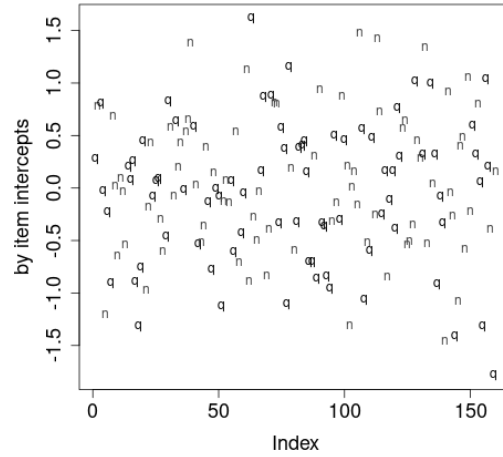
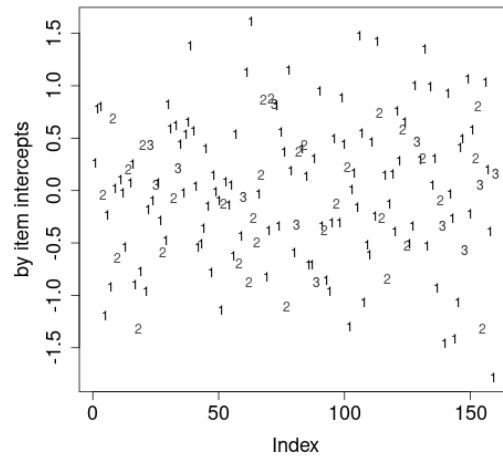


Figure 30: By-PARTICIPANT intercepts with PARTICIPANT-dependent variables.



(a) q= *que* present, n= no *que* present



(b) 1= dialog with 1 interlocutor, 2= dialog with 2 interlocutors, 3= dialog with 3 interlocutors

Figure 31: By-ITEM intercepts with ITEM-dependent variables.

The residual plots in 32 show that once again the residuals do not have a normal distribution. I repeat that this is also not necessarily an expected distribution given the present modeling. Here too, the residuals are grouped at the extremes with a gap in the center, suggesting that once more the model predicts fewer values than were observed.

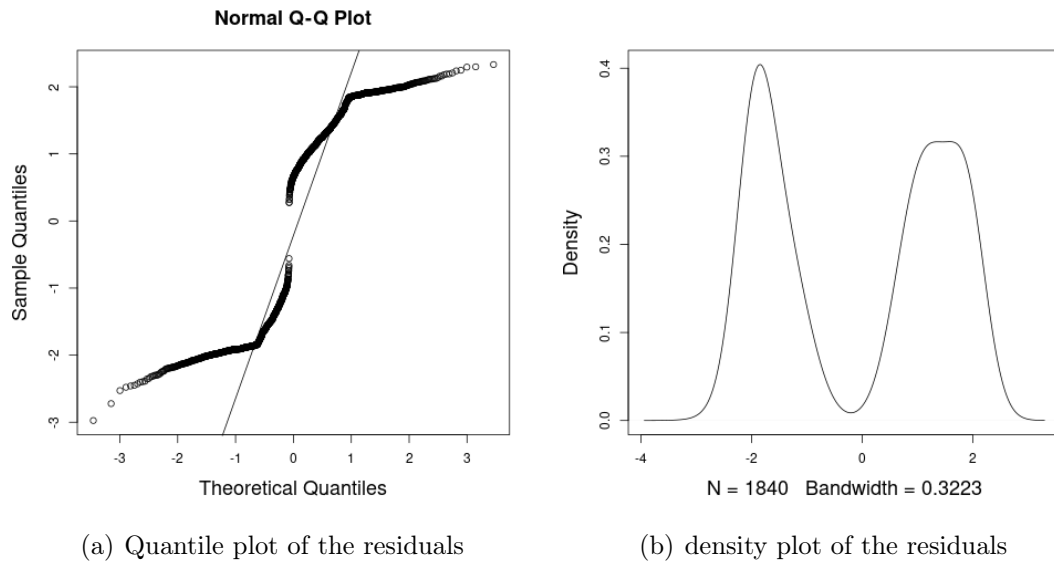


Figure 32: Quantile and density plot of the residuals of the model.

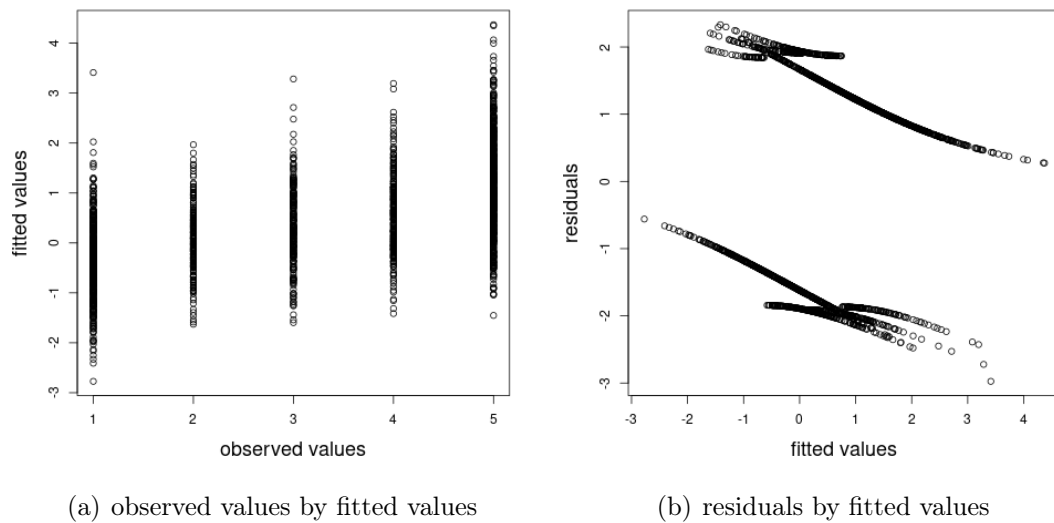


Figure 33: Correspondence of fitted vs. observed values and fitted values vs. residuals.

As in the previous results, the plots in 33 indicate a moderate accuracy in the predictions of the model. In 33(a), the intermediate observed values show hardly any difference in the values they are fitted to and they span large intervals. At the extremes, i.e. for the observed value 1 and 5, the predictions appear more accurate and there is a clear

difference between the two. The plot in 33(b) again shows that the model predicts well for the extreme observed values but not for the intermediate observed values. For the intermediate fitted values there are large positive and negative residuals.

The heatmap in 34 shows that the suspicion formulated above is true: the model does predict fewer categories than were observed, and in fact only predicts two. The two observed values 1 and 2 are fitted to the lower category, while all the others are fitted to the higher category.

The deviance explained (14.2%) and the accuracy of the model (22.8%) are low. The low accuracy is again due to the fact that the model predicts fewer categories than are observed. The low deviance explained suggests that there is more variation in the data than has been captured through the predictors in the model.

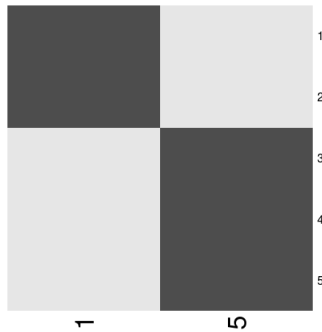
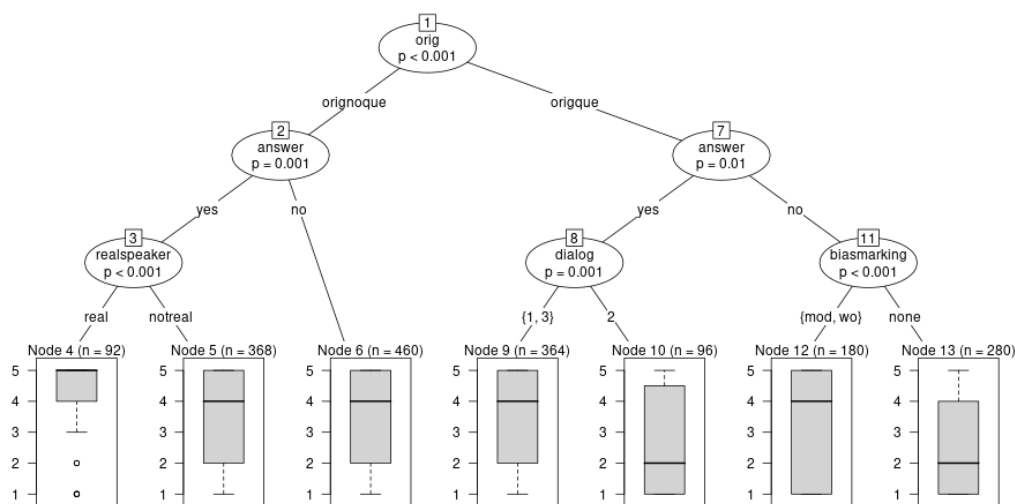


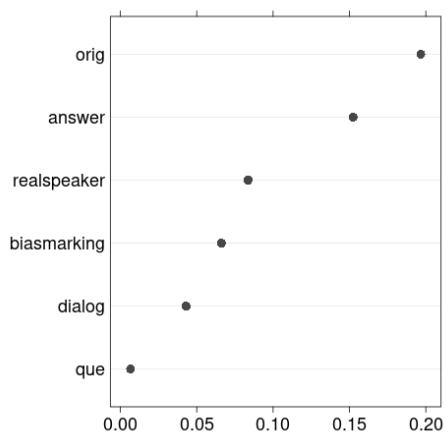
Figure 34: Heat map of 2 predicted values versus 10 observed values. Darker shades of grey correspond to larger counts.

A conditional inference tree (figure 35(a)) with splits on ORIGINAL, ANSWER, REAL SPEAKER, DIALOG and BIAS MARKING reaches a lower accuracy than the GAM model (19%). However, it shows complex three-way interactions that could not be detected in the previous model. In particular, the tree model shows that within 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 more unexpected when there are two interlocutors interacting with the the speaker who provides the answer. When the answer is negative the presence of additional bias marking boosts 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. The dot plot in 35(b) plots the variable importance calculated with a random forest model. Just like the GAM model, this 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 showed no significant effect in the GAM model, is virtually inexistent.



(a) conditional inference tree



(b) variable importance

Figure 35: Conditional inference tree with splits on ORIGINAL, ANSWER, REALSPEAKER, DIALOG and BIAS MARKING and a dot plot showing the importance of each variable.

Discussion The models show a solid effect for ORIGINAL. The main insight I draw from these results is that the differences between bias in polar questions goes far beyond a 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) targeted some potential aspects and showed some modest effects. However, the low deviance explained by the GAM model 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 GAM model only predicts two categories. As mentioned in the discussion of the raw judgments, I suspect that these two categories are the result of a division between unexpected answers and all the other cases. Importantly, the high-value category therefore comprises expected answers but also answers where there are no prior expectations. If this is truly the case, there is evidence in the results that the theoretical assumption that polar questions with *que* carry a positive bias is on the right track. In particular, the tree model shows that in ORIGINAL-QUE contexts with no bias marking, the negative answer is unexpected. If a negative answer is unexpected, conversely a positive one is expected. 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.

One last point that I would like to address again is the imperfect nature of the data. For the present experiment, one problematic aspect already mentioned above is the fact that the data are not dialogs. As a consequence, more data modification than intended had to be performed. In particular, an answer had to be added, since not all of the questions in the corpus were originally followed by one. Some of these may have been intended as rhetorical questions that do not require an answer but in fact assert the

prejacent of the question. This could be an additional explanation of why the positive answer was expected to such a high degree.

4.5 General Discussion

To conclude this chapter, I would like to summarize a number of issues that I faced in the empirical investigations reported 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 allows for compensation could become useful tools. In the current study, however, even relying on Prolific did not allow me to reach enough Catalan speakers. While the number 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, the phenomena require 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 however restricted 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 thesis. They do contain information on the country of origin of a fragment. However, these are taken from the domain of the

homepage. To identify the country of origin of the author of a posting, the domain of the homepage is simply insufficient.

Turning now to the experiments themselves, 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 the 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 fewer categories than were observed. Although this can be a telling result, as it was in the last experiment, it also restricts the possibilities of data modeling and makes it hard to determine the goodness of fit of models based on the accuracy of the predictions.

A final point is to determine how the empirical method used and described here can inform the theory. The approach was exploratory. This means that in adopting it 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 compared 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. The conclusion I draw for my own research is that a broad methodological repertoire is the key. This is the only approach that allows us to do justice to the complexity of linguistic phenomena. However, one question remains: what conclusions should be drawn for generative linguistic research and theory building? The data suggest that the phenomena under investigation are influenced by multiple factors. The major challenge that researchers face in light of the growing importance of big data is how results like these can be useful for a generative theory that aims for clear generalizations.

5 Conclusion

This thesis 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 element that adopts its functional interpretation in the syntactic position in which it is merged. Over the course of the thesis, I have shown that there is a correlation between the syntactic position and behavior of the complementizer and the interpretation of the proposition in its scope, which has ultimately demonstrated that this new conception of complementizer is indeed adequate. The analysis I put forward 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.

Chapter 2 was dedicated to *que*-initial reportative sentences. I drew a parallel between these and embedded sentences and showed that in these contrasted contexts, the syntactic behavior of the complementizer is virtually the same. The empirical distribution supported my proposal to analyze the complementizer as being merged in the highest position in the left periphery valued and with a subordinate feature. The difference between the complementizer in embedded and unembedded contexts is merely that they are syntactically selected in the first case and remain 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 is found in 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 which according to my analysis all involve a low merged complementizer that is valued with an attributive feature. The unified account is justified because there is a common core interpretation associated

with the attributive feature, that can be observed in all the different constructions. I proposed that the attributive feature should be linked to the lowest projection of the left periphery. Further 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 material that was externally merged in this same position. 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.

The final chapter 4 presented corpus-based empirical investigations that aimed to broaden the understanding of some of the constructions involving attributive *que*. The approach adopted was exploratory rather than strictly hypotheses-driven. This means that, while the results can reveal 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, only some of the theory-based expectations could be confirmed. However, one central novel insight about these constructions 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, however the results of the experiments indicate that speakers are willing to accommodate different readings of the modifiers even though they do not correspond to those made salient by the relevant constructions. With regard to the experiment regarding Catalan polar questions, the results did provide (indirect) evidence in favor of the analysis proposed. The participants judged negative answers to *que*-initial polar questions as unexpected. One may conclude, then, that a positive answer was 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*.

6 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 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 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. *Proceedings of SALT* 20. 328–346.
- 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 (MIT).
- Auwers, Johan van der & Vladimir A. Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2. 79–124.
- Baayen, R Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.
- Baayen, R Harald. forthcoming. *Analyzing linguistic data: A practical introduction to statistics using R. second edition*. Cambridge University Press.
- Baayen, R Harald & Dagmar Divjak. 2017. Ordinal GAMMs: a new window on hu-

-
- man ratings. In Anastasia Makarova, Stephen M. Dickey & Dagmar Divjak (eds.), *Thoughts on language: Studies in cognitive linguistics in honor of Laura A. Janda*, 39–56. Bloomington, IN: Slavica Publishers.
- Bartol Hernández, José Antonio. 1988. *Las oraciones causales en la Edad Media*. Madrid: Paraninfo.
- Batllo, Montserrat & Avellina Suñer. 2005. Valors adverbials associats a la conjunció *que*: un estudi històric. *Caplletra* 38. 279–301.
- Batllo, Montserrat & M. Lluïsa Hernanz. 2013. Emphatic polarity particles in Spanish and Catalan. *Lingua* 128. 9–30.
- Batllo, 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. <http://filcat.uab.cat/clt/publicacions/reports/pdf/GGT-09-08.pdf>.
- Batllo, 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 de Bustos Tovar (ed.), *Lengua, discurso, texto : I simposio internacional de análisis del discurso*, 227–240. Madrid: Visor.
- 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. 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.
- 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*, vol. 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. Kaitakusha Co., Ltd.
- Benincà, Paola & Cecilia Poletto. 2004. Topic, focus, and V2. In Luigi Rizzi (ed.), *The structure of CP and IP: The cartography of syntactic structures volume 2*, 52–75. New York, Oxford: Oxford University Press.
- 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.
- Bianchi, Valentina, Giuliano Bocci & Silvio Cruschina. 2016. Focus fronting, unexpectedness, and evaluative implicatures. *Semantics and Pragmatics* 9(3). 1–54. doi: 10.3765/sp.9.3. <http://dx.doi.org/10.3765/sp.9.3>.
- 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 2*, 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 dissertation.
- 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*. de Gruyter Mouton.
- Bošković, Željko. 2007. On the locality and motivation of move and agree: An even more minimal theory. *Linguistic Inquiry* 38. 589–644.
- Breiman, Leo. 2001. Random forests. *Machine Learning* 45. 5–32.
- 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, Hungary: 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. John Benjamins.
- Brucart, Josep M. 1992. Some asymmetries in the functioning of relative pronouns in Spanish. *Catalan Working Papers in Linguistics* 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). 511–526.
- Büring, Daniel & Christine Gunlogson. 2000. Aren't positive and negative polar questions the same? Manuscript.
- 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. <http://plato.stanford.edu/archives/win2008/entries/quotation/>.
- Carvalho, Ana Sofia Abreu de. 2013. An overview of vocatives in European Portuguese. *Linguistica Atlantica* 32. 52–58.
- Castroviejo, Elena. 2006. *Wh-exclamatives in catalan*: Universitat de Barcelona 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 Ma 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: Multidisci-*

-
- plinary perspectives on its function in discourse*. Berlin: De Gruyter Mouton.
- 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: The 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. 1994. On the evidence for partial N-movement in the Romance DP. In Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi & Raff aella Zanuttini (eds.), *Paths towards universal grammar. Studies in honor of Richard S. Kayne*, 85–110. Washington, DC: Georgetown University Press.
- Cinque, Guglielmo. 1990. *Types of A-bar dependencies*. Cambridge, MA: MIT 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. John Benjamins.
- Coniglio, Marco & Iulia Zegrean. 2010. Splitting up Force, evidence from discourse particles. *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*: University of Cambridge dissertation.
- Coulmas, Florian. 1986. *Direct and indirect speech*, vol. 31. Berlin: Walter 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 per-*

-
- spectives*, 262–283. Oxford University Press.
- Cruschina, Silvio. 2015. The expression of evidentiality and epistemicity: Cases of grammaticalization in Italian and Sicilian. *Probus* 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.
- de Cuba, Carlos & Jonathan E. MacDonald. 2013. Referentiality in Spanish CPs. *Information Structure and Agreement* 117–140. doi:10.1075/la.197.04cub. <http://dx.doi.org/10.1075/la.197.04cub>.
- Cuenca, Maria Josep. 1997. Form-use mappings for tag questions. *Current Issues in Linguistic Theory* 3–19. doi:10.1075/cilt.151.04cue. <http://dx.doi.org/10.1075/cilt.151.04cue>.
- 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.
- Déchaine, Rose-Marie, Clare Cook, Jeffrey Muehlbauer & Ryan Waldie. 2017. (De-)constructing evidentiality. *Lingua* 186–187. 21–54.
- 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. doi:10.3765/sp.8.11.
- 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 Mouton.

-
- 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. doi:10.1515/prbs.2009.002. <http://dx.doi.org/10.1515/prbs.2009.002>.
- Demonte, Violeta & Olga Fernández Soriano. 2014. Evidentiality and illocutionary force: Spanish matrix *que* at the syntax-pragmatics interface. In Andreas Dufter & Alvero O. de Toledo (eds.), *Left sentence peripheries in Spanish: Diachronic, variationist, and typological perspectives*, 217–252. 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*: MIT 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. Mouton de Gruyter.
- Etxepare, Rica. 1997. *The grammatical representation of speech events*: University of Maryland 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. John Benjamins.
- Etxepare, Ricardo. 2010. From hearsay evidentiality to samesaying relations. *Lingua* 120(3). 604–627. doi:10.1016/j.lingua.2008.07.009. <http://dx.doi.org/10.1016/j.lingua.2008.07.009>.

-
- 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 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. doi:10.1093/jos/ffp010.
- von Fintel, Kai & Anthony S. Gillies. 2010. Must . . . stay . . . strong! *Natural Language Semantics* 18(4). 351–383. doi:10.1007/s11050-010-9058-2. <http://dx.doi.org/10.1007/s11050-010-9058-2>.
- von Fintel, Kai & Anthony S. Gillies. 2011. ‘might’ made right. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, 108–130. Oxford: Oxford University Press.
- Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. *Linguistik Aktuell/Linguistics Today* 87–116. doi:10.1075/la.100.07fra. <http://dx.doi.org/10.1075/la.100.07fra>.
- 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* 143–177.
- Garzonio, Jacopo. 2004. Interrogative types and left periphery: Some data from the Fiorentino dialect. *Quaderni di lavoro dell’ASIS* 4. 1–19.
- van Gelderen, Elly. 2011. *The linguistic cycle*. Oxford: Oxford University Press.
- Ginzburg, Johnathan. 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 (OUP). doi:10.1093/acprof:oso/9780199571895.001.0001. <http://dx.doi.org/10.1093/acprof:oso/9780199571895.001.0001>.
- González Rodríguez, Raquel. 2008. *La polaridad positiva en español*: Universidad Complutense de Madrid, Servicio de Publicaciones 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*: Universitat de Barcelona dissertation.
- 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 & 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. Dordrecht, Heidelberg, London, New York: Springer.
- Gutiérrez-Rexach, Javier. 2001. Spanish exclamatives and the interpretation of the left periphery. In Yves D’Hust ad Johan Rooryck & Jan Schroten (eds.), *Romance languages and linguistic theory 1999: Selected papers from Going Romance leiden 1999*, 167–195. John Benjamins.
- Gutzmann, Daniel & Erik Stei. 2011. How quotation marks what people do with words. *Journal of Pragmatics* 43(10). 2650–2663.
- 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. Akademie-Verlag.
- de Haan, Ferdinand. 1999. Evidentiality and epistemic modality: Setting boundaries. *Southwest Journal of Linguistics* 18(1). 83–101.
- 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. doi:10.1016/j.lingua.2005.03.014. <http://dx.doi.org/10.1016/j.lingua.2005.03.014>.
- Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and composition of the left periphery: The cartography of syntactic structures*. 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. doi:10.1111/stul.12023. <http://dx.doi.org/10.1111/stul.12023>.
- 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, M Lluïsa. 2007. From polarity to modality: some (a)symmetries between *bien en sí* in Spanish. In Luis Eguren & Olga Fernández-Soriano (eds.), *Studies in Spanish syntax*, 105–150. Venezia: Editrice Cafoscarina.
- Hernanz, M 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. Servei editorial de la UPV-EHU.
- Hill, Virginia. 2006. Stylistic inversion in Romanian. *Studia Linguistica* 60(2). 156–180. doi:10.1111/j.1467-9582.2006.00136.x. <http://dx.doi.org/10.1111/j.1467-9582.2006.00136.x>.
- 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. doi:10.1016/j.lingua.2007.01.002. <http://dx.doi.org/10.1016/j.lingua.2007.01.002>.
- 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. Mouton de Gruyter.
- Hill, Virginia. 2014. *Vocatives: How syntax meets with pragmatics*. Leiden: Brill.
- Hopper, Paul & Elisabeth Closs Traugott. 2003. *Gramaticalization*. Cambridge University Press.
- Hornstein, Norbert. 1995. *Logical form: From GB to minimalism*. Oxford: Blackwell.
- 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 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. John Benjamins.
- Höhle, Tilman N. 1992. Über Verum-Fokus im Deutschen. In *Informationsstruktur und Grammatik*, 112–141. Berlin: Springer. http://link.springer.com/content/pdf/10.1007/978-3-663-12176-3_5.pdf.
- 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.
- 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.
- Kerstin Schwabe, K. 2006. 2006. German *dass*-clauses. In Valerie Molnár & Susanne Winkler (eds.), *The architecture of focus*, 429–458. Berlin: Mouton de Gruyter.
- Kocher, Anna. 2014. *Claro que son adverbios: Análisis de un proceso de gramaticalización*. University of Vienna MA thesis.
- Kocher, Anna. 2017. 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. 2017. *Root clause complementation in Catalan*: University of Vienna MA thesis.
- Kocher, Anna. 2018a. 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. 2018b. Sentence mood and verum focus: A unified analysis for German, English and Spanish. Manuscript.
- Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3-4). 243–276.
- Krifka, Manfred. 2013. Response particles as propositional anaphors. In *Proceedings of salt*, vol. 23, 1–18.
- Laka, I. 1990. *Negation in syntax*: MIT 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. doi:10.1111/j.1467-968x.2005.00157.x. <http://dx.doi.org/10.1111/j.1467-968x.2005.00157.x>.
- Ledgeway, Adam. 2012. *From Latin to Romance: Morphosyntactic typology and change* Oxford Studies in Diachronic and Historical Linguistics. Oxford University Press.
- Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. John Benjamins Publishing Company.
- Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8. 339–359.
- Ljubešić, Nikola & Antonio Toral. 2014. caWaC - a web corpus of catalan and its application to language modeling and machine translation. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the ninth international conference on language resources and evaluation (LREC'14)*, Reykjavik, Iceland: European Language Resources Association (ELRA).
- Lobo, Maria. 2003. *Aspectos da sintaxe das orações subordinadas adverbiais do português*: Universidade de Lisboa 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. 2015. Verum focus. In Caroline Fréy & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 2015. 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* 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: Ed. Espasa Calpe.
- Martins, Ana Maria. 2006. Emphatic affirmation and polarity. *Current Issues in Linguistic Theory* 197–223. doi:10.1075/cilt.278.10mar. <http://dx.doi.org/10.1075/cilt.278.10mar>.
- 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.
- 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*: Universidad de Granada dissertation.
- 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 University Press (OUP). doi:10.

-
- 1093/acprof:oso/9780199665297.003.0007. <http://dx.doi.org/10.1093/acprof:oso/9780199665297.003.0007>.
- Milner, Jean-Claude. 1978. *De la syntaxe a l'interprétation*. Paris: Seuil.
- Mioto, Carlos. 2003. Focalização e quantificação. *Revista Letras* 61. 169–189.
- Moreira, Juliana. 2013. *O vocativo e a interface sintaxe-pragmática no português brasileiro*: Universidade Federal de Minas Gerais, Belo Horizonte 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: 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 and 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 Pragmatic* 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*. Oxford: Oxford University Press 2nd edn.
- 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. 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.
- 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(146). 39–74. doi:10.1016/j.lingua.2014.04.007. <http://dx.doi.org/10.1016/j.lingua.2014.04.007>.

-
- 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.
- Mascaró i Pons, Ignasi. 1986. Introducció a l'entonació dialectal catalana. *Randa* 22. 5–38.
- Pons Borderia, 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. 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 *Actes del novè col·loqui d'estudis catalans a nord-amèrica*, 347–377. Publicacions de l'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. Empúries.
- Prieto, Pilar & Gemma Rigau. 2007. The syntax-prosody interface: Catalan interrogative sentences headed by *que*. *Journal of Portuguese Linguistics* 6(2). 29–59.
- 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.
- Carrera de la Red, Fátima. 1982. *Las expresiones causativas en la obras de Gonzalo de Berceo*. Logroño (La Rioja): Instituto de Estudios Riojanos.
- 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. 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.
- Rivero, María Luisa. 1982. Las relativas restrictivas con *que*. *Nueva Revista de Filología Hispánica* 31. 195–234.
- 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*, 281–337. Kluwen 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 offered to Lorenzo Renzi*, vol. 14, 267–296. Elsevier.
- Rizzi, Luigi. 2004a. Locality and left periphery. In Adriana Belletti (ed.), *Structures and beyond. The cartography of syntactic structures*, vol. 3, 223–251. 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*, vol. 2, 3–15. Oxford: Oxford University Press.
- Rizzi, Luigi. 2013. Notes on cartography and further explanation. *International Journal of Latin and Romance Linguistics* 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. doi:10.3765/sp.5.6.
- Roberts, Ian. 2001. Head movement. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 113–147. Oxford: Blackwell Publishers

-
- Ltd.
- Roberts, Ian. 2004. The C-system in Brythonic Celtic languages, V2 and the EPP. In Luigi Rizzi (ed.), *The structure of CP and IP*, 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. 2003. *La gramática de los adverbios en -mente o cómo expresar maneras, opiniones y actitudes a través de la lengua*. Madrid.
- Rodríguez-Ramalle, Teresa. 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 – 811–816. Murcia: EDITUM.
- Rodríguez-Ramalle, Teresa M. 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. doi:10.1075/lic.15.1.07ram. <http://dx.doi.org/10.1075/lic.15.1.07ram>.
- Rodríguez Ramalle, Teresa María. 2007. El complementante "que" como marca enfática en el texto periodístico. *revista electrónica de lingüística aplicada* 6. 41–53.
- 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. *Glott 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, A. 2000. On the left periphery: Modal particles and complementisers. *Journal of Greek Linguistics* 1(1). 65–94. doi:10.1075/jgl.1.05rou. <http://dx.doi.org/10.1075/jgl.1.05rou>.
- Roussou, Anna. 2010. Selecting complementizers. *Lingua* 120(3). 582–603. doi:10.1016/j.lingua.2008.08.006. <http://dx.doi.org/10.1016/j.lingua.2008.08.006>.
- Saab, Andrés. 2015. On long-distance theta-role assignment. *Lingua* 160. 91–126.
- 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. Walter 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. doi:10.1016/j.pragma.2014.12.004. <http://dx.doi.org/10.1016/j.pragma.2014.12.004>.
- 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, Philadelphia: 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.
- Speas, Margaret. 2004. Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua* 114(3). 255–276. doi:10.1016/S0024-3841(03)00030-5. [http://dx.doi.org/10.1016/S0024-3841\(03\)00030-5](http://dx.doi.org/10.1016/S0024-3841(03)00030-5).
- Speas, Peggy & Carol Tenny. 2003. Configurational properties of point of view roles. In Anna Maria DiSciullo (ed.), *Asymmetry in grammar*, vol. 1, 315–345. 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: Mouton de Gruyter.
- Stowell, Timothy. 1981. *Origins of phrase structure*: MIT 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.
- 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.

-
- 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* 116–140.
- Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change* 24(02). 135–178. doi:10.1017/s0954394512000129. <http://dx.doi.org/10.1017/S0954394512000129>.
- Tenny, Carol. 2006. Evidentiality, experiencers and the syntax of sentience in Japanese. *Journal of East Asian Linguistics* (15). 245–288.
- Torrence, Harold. 2013. *The clause structure of Wolof*. Amsterdam: Benjamins.
- Traugott, Elizabeth. 2010. (inter)subjectivity and (inter)subjectification: A reassessment. *Topics in English Linguistics* 29–74. doi:10.1515/9783110226102.1.29. <http://dx.doi.org/10.1515/9783110226102.1.29>.
- Uriagereka, Juan. 1988. *On government*. Storrs, CT: University of Connecticut dissertation.
- Uriagereka, Juan. 1995a. An F position in Western Romance. In Katalin Kiss (ed.), *Discourse configurational languages*, 153–175. Oxford University Press.
- Uriagereka, Juan. 1995b. An F position in western Romance, vol. 26, 79–123. 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*: University of Connecticut dissertation.
- Villa-García, Julio. 2015. *The syntax of multiple que sentences in Spanish* (Issues in Hispanic and Lusophone Linguistics 2). Amsterdam: John Benjamins Publishing Company. doi:10.1075/ihll.2. <http://dx.doi.org/10.1075/ihll.2>.
- Villa-García, Julio & Raquel González Rodríguez. 2016. On emphatic affirmative polarity markers and predicates of truth: *sí* and *sí que* in Spanish. Ms.
- 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.
- 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.
- Wood, Simon N. 2017. *Generalized additive models: an introduction with R*. CRC 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. *Nat Language & Linguistic Theory* 30. 1231–1274.
- Zanuttini, Raffaella & Paul Portner. 2003. Exclamative clauses: At the syntax-semantics interface. *Language* 79(1). 39–81.

Abstracts

Abstract (English)

This thesis deals with the structural and interpretive properties of Ibero-Romance root clauses containing a syntactically unselected complementizer (cf. I, II).

- (I) Que en Jordi va a la festa. (Cat.)
QUE the Jordi goes to the party
‘Reportedly, Jordi is going to the party.’
- (II) Claro que en Jordi va a la festa. (Cat.)
clear QUE the Jordi goes to the party
‘Clearly, Jordi is indeed going to the party.’

Usually, a complementizer is defined as an element that identifies a complement clause that is syntactically selected. In the data that are the center of the present investigation, the complementizer does not express its prototypical function and appears in apparently unembedded contexts. The two main arguments I defend in my thesis are that the sentences in (I) and (II) are syntactically unselected and that the presence of *que* has an impact on the interpretation of the sentences in its scope. In (I), *que* gives rise to a reportative reading. In (II) the effect of *que* is that a commitment to the unmodified proposition is attributed to the hearer. One aim of this thesis is to show that, in spite of this untypical behavior, it is possible to maintain the idea that there is only one lexical element *que* in Ibero-Romance. This is achieved by refining the conception of what constitutes a complementizer. My proposal is that the Ibero-Romance complementizer is an underspecified element that acquires its functional meaning in the syntactic position in which it is merged, where it is valued with an interface feature. The analysis is developed within a cartographic framework (Rizzi 1997 ff.). Through the interplay with other left-peripheral material, I show that in fact the different interpretations correlate with different merge positions of the complementizer. It is merged at the highest projection of the split CP in examples like (I) and at the lowest projection in examples like (II).

Abstract (deutsch)

Diese Dissertation beschäftigt sich mit den strukturellen und interpretativen Eigenschaften von iberoromanischen Hauptsätzen, die durch einen syntaktisch unselegierten Komplementierer eingeleitet werden (vgl. I, II).

- (I) Que en Jordi va a la festa. (Kat.)
QUE der Jordi geht zu der Feier
‘Angeblich geht Jordi zu der Feier.’
- (II) Claro que en Jordi va a la festa. (Kat.)
klar QUE der Jordi geht zu der Feier
‘Klar, Jordi geht in der Tat zu der Feier geht.’

Normalerweise wird ein Komplementierer definiert als ein Element, das syntaktisch eingebettete Komplementsätze identifiziert. In den Daten, die im Zentrum dieser Untersuchung stehen, drückt der Komplementierer also nicht seine prototypische Funktion aus. Ein zentrales Argument der Dissertation ist, dass die Sätze in (I) und (II) nicht eingebettet sind und dass die Anwesenheit von *que* Konsequenzen für die Interpretation der Sätze hat. Dies führt zu einer repositiven Lesart des Satzes in (I) und schreibt dem Hörer ein Commitment zur unmodifizierten Proposition in (II) zu. Ein Ziel dieser Dissertation ist es zu motivieren, dass es, trotz dieses untypischen Verhaltens, möglich ist, die Annahme zu vertreten, dass es in den untersuchten Sprachen nur ein lexikalisches Item *que* gibt. Diese Annahme basiert auf einer neuen Konzeption der Kategorie “Komplementierer”. Ich schlage vor, dass ein iberoromanischer Komplementierer ein unterspezifiziertes Element ist, das seine funktionelle Bedeutung in der syntaktischen Position erwirbt, in der es basisgeneriert wird. Ich entwickle eine Analyse im theoretischen Rahmen der syntaktischen Kartographie (Rizzi 1997 ff.). Durch die Interaktion mit anderem links-peripherem Material wird gezeigt, dass die unterschiedlichen Interpretationen tatsächlich mit unterschiedlichen Basispositionen im funktionalen Feld korrelieren. In Beispielen wie (I) wird der Komplementierer in der höchsten Position der gesplitteten CP basisgeneriert, während er in Beispielen wie (II) in der niedrigsten Position basisgeneriert wird.