ZEITSCHRIFT FÜR SPRACHWISSENSCHAFT
Journal of the Linguistic Society of Germany

Volume 45, Issue 3, Article 6
https://doi.org/10.18148/zs/2026-2015

Max Bonke, Universität zu Köln, Germany, mbonke@uni-koeln.de

Pragmatic constraints on the complementizer *que* 'that' in Spanish embedded gapping

Abstract: Spanish, unlike English, allows the declarative complementizer *que* 'that' in contexts where the finite verb has been elided via gapping. However, the acceptability of *que* in Spanish gapping varies systematically with the type of embedding verb: some verbs permit it readily, while others show a marked degradation in acceptability. Previous work hypothesized that factivity might account for these verb-type effects, but the evidence has been based on inferences from acceptability patterns and lacks support from reliable quantitative data. This paper provides a quantitative and experimental investigation of verb-type constraints on *que* in embedded gapping. A correlation analysis reveals that the degradation of *que* aligns with independently collected measures of verb factivity. However, this correlation is driven by a subset of verbs, i.e., saying and thinking verbs, rather than representing a general pattern. Two experiments further examine the interpretive impact of *que*, testing whether its presence modulates factive inferences and exploring possible connections to additional pragmatic properties (assertion, at-issueness, reportiveness). In both experiments, the presence or absence of *que* has no systematic effect on interpretation. These results reinforce previous rejections of factivity-based accounts; however, the precise source of verb-type effects on *que* in Spanish gapping remains elusive.

Keywords: gapping, clause embedding, complementizer, factivity, quantitative pragmatics

Cite as: Bonke, Max. 2026. Pragmatic constraints on the complementizer *que* 'that' in Spanish embedded gapping. *Zeitschrift für Sprachwissenschaft* 45(3). https://doi.org/10.18148/zs/2026-2015.
Received: 16 September 2025; **accepted:** 27 August 2026; **published:** 16 November 2026
Published by: publikon / dgfs (Linguistic Society of Germany) under a Creative Commons 4.0 BY license

1 Introduction

Gapping[1] is a type of ellipsis that deletes the finite verb in the second of two coordinate clauses (plus, optionally, further identical constituents) (Ross 1968, 1970; Sag 1977; Hankamer 1979; Johnson 2009, 2018, among many others). An example of gapping is in (1), where the underscore indicates the position of the elided verb *lives*.

(1) *Mary lives in England and John _ in Wales.*

As has long been observed, gapping is subject to structural and discourse restrictions. One syntactic domain in which such restrictions are at play is in clausal embedding. When coordinated clauses with gapping serve as the complement of an embedding predicate, only the first conjunct may host a complementizer (e.g., *that*), but not the second, elliptical conjunct. (Fiengo 1974; Wilder 1994, 1997; Hartmann 2000; Taraldsen 2000; Repp 2009; Broekhuis & Corver 2019, among others). This property is illustrated in (2), where the gapping structure in (1) above constitutes the clausal complement of the embedding verb *think*. This structure is only grammatical without the second *that*.

(2) *Anne thinks that Mary lives in England and* (**that*) *John _ in Wales.*

The fact that the complementizer is ungrammatical in the second conjunct is clearly a property of gapping, since in non-elliptical coordinate embedded clauses the complementizer is grammatical; see (3).

(3) *Anne thinks that Mary lives in England and that John lives in Wales.*

In English, there is thus a categorical, ellipsis-specific ban on complementizers from clauses with gapping. Cross-linguistically, however, the possibility of complementizers in gapping is a matter of variation. Other languages with comparable restrictions to English include German (Wilder 1994, 1997; Hartmann 2000; Repp 2009) and Dutch (Broekhuis & Corver 2019), but there are also languages that allow complementizers in gapping (for an overview, see Bonke 2024: Ch. 2).

Among the languages that allow gapping with complementizers is Spanish, as shown in the actual use example in (4). This example contains a coordinate structure with an elliptical second conjunct (*los mexicanos sus lacayos*, lit. 'the Mexicans his lackeys', with elided copula verb *son* 'are'). This coordinate structure functions as the complement of the clause-embedding verb *cree* 'thinks'. Importantly, the elliptical conjunct features the complementizer *que* 'that'.

[1] This paper consists of revised sections of my dissertation (Bonke 2024: Chapters 4/5). The data and analyses for the two experiments presented in this paper are available on OSF: https://osf.io/r4shg.

(4) Spanish
El cree que el mundo es su empresa y ***que*** *los mexicanos _ sus lacayos.*
He thinks that the world is his company and that the Mexicans his lackeys
'He thinks that the world is his company and that the Mexicans are his lackeys.'
(corpus example from Bonke & Repp 2022: 528)

While Spanish allows complementizers in gapping as shown, they are not obligatory. Embedded gapping without the complementizer in the elliptical conjunct (like in English) is also possible; the sentence in (4) above would be equally grammatical without *que* in the second conjunct (*… y los mexicanos sus lacayos*, lit. 'and the Mexicans his lackeys').

This alternation of gapping with vs. without the complementizer in Spanish raises the question as to whether the presence/absence of *que* might be constrained, for example, by grammatical properties or properties of the discourse context. In an acceptability rating study, Bonke & Repp (2022) find constraints on the type of the embedding verb (e.g., *cree* 'thinks' in (4) above). Their results indicate that embedded gapping without *que* is equally acceptable across (types of) embedding verbs, but the variant with *que* is degraded in acceptability with some verb types (e.g., emotives like *lamentar* 'regret', or knowledge verbs like *saber* 'know'). Bonke & Repp (2022) attempt to establish a connection between these acceptability patterns and pragmatic properties of the verbs, focusing on factivity. The hypothesis motivating their experiments is that those verbs that easily allow for gapping with *que* are non-factive, while those that penalize *que* in gapping are factive.

Due to some unexpected patterns in the acceptability data (which are described in detail in Section 2.2), Bonke & Repp (2022) ultimately reject the hypothesis that factivity is crucial in the acceptability of gapping with *que*. However, as I argue in this paper, the empirical case for their conclusion is insufficiently supported. Acceptability data only allows for indirect inferences on the (presumed) interaction of *que* in gapping and clause meaning, which forms a cornerstone of Bonke & Repp's (2022) structural analysis of gapping. Furthermore, Bonke & Repp's (2022) categorization of verbs as factive vs. non-factive relies largely on introspection rather than robust quantitative evidence, which introduces a potential source of misclassification. These empirical issues motivate a re-examination of Bonke & Repp's (2022) gapping data under a more sophisticated, empirically validated treatment of factivity.

In this paper, I return to the factivity hypothesis, putting an empirical focus on the meaning of *que* in gapping. I present a new analysis of existing data on Spanish, which correlates Bonke & Repp's (2022) acceptability data with independently collected quantitative data on the interpretation of factivity for the relevant clause-embedding verbs (Bonke 2024: Ch. 5). I also present the results of two experiments that target the pragmatic contribution of *que* in gapping. The first experiment (Experiment 1) tests whether the presence of *que* in the elliptical conjunct affects

the interpretation of factivity. The results of the correlation analysis and Experiment 1 indicate no meaningful connection between factivity and *que* in gapping, supporting Bonke & Repp's (2022) conclusion with clear and direct empirical evidence. Due to these null results, I conducted the second experiment (Experiment 2), which applied novel diagnostics to gapping clauses with vs. without *que* to explore the hypothesized pragmatic differences more generally. Due to the novel diagnostics and exploratory nature of Experiment 2, the results are to be treated with caution; however, the results parallel the results of Experiment 1 in suggesting no effect of the complementizer on pragmatics.

This paper is structured as follows. Section 2 provides background information on factivity and outlines Bonke & Repp's (2022) study design, findings, and analysis in more detail. Section 3 contains the correlation analysis of the acceptability data and quantitative data on the factivity of the relevant clause-embedding verbs from Bonke (2024: Ch. 5). Section 4 contains Experiment 1 on factivity, Section 5 contains the exploratory Experiment 2, and Section 6 concludes.

2 Gapping, complementizers, and verb-type effects

This section goes into detail on the acceptability patterns of the Spanish complementizer *que* 'that' in embedded gapping. For an example of this structure, see (4) in the introductory section above; further examples follow in Section 2.2 below. Specifically, this section outlines how the acceptability of *que* varies depending on the type of the embedding verb. The basis of this discussion is the acceptability study by Bonke & Repp (2022), who hypothesize that a particular property of embedding verbs, factivity, is crucial for the acceptability of *que* in gapping (but end up rejecting this hypothesis). I start with a brief overview of classical approaches on factivity (Section 2.1), outlining the facts on English, which I assume generalize to Spanish without issue. After that, I turn to Spanish and the verb-type effects on *que* in embedded gapping in Section 2.2.

2.1 Factivity and factive sub-types

Kiparsky & Kiparsky (1970) first described the concept of factivity. They divided clause-embedding predicates into two groups, depending on the discourse status of their complement clause. Factive predicates presuppose the propositional content of their complement clause, while non-factive predicates do not presuppose the propositional content of their complement clause. To illustrate the difference, consider the difference between examples (5) and (6).

(5) *I regret that it is raining.* (Kiparsky & Kiparsky 1970: 147)

(6) *I believe that it is raining.*

When a speaker utters the sentence in (5), they take the proposition *it is raining* for granted and as uncontroversial in discourse. In more technical terms, they assume the propositional content of the complement clause (i.e., that it is raining) to be part of the common ground (Stalnaker 1978, 2002). This presupposed discourse status of the content of the complement clause means that the predicate that embeds it, i.e., the verb *regret*, is factive.

In contrast, someone that utters the sentence in (6) does not presuppose the content of the complement clause, i.e., they do not take the proposition *it is raining* for granted and as uncontroversial in discourse. The content of the complement clause is thus not treated as part of the common ground. Consequently, the embedding predicate *believe* is non-factive.

A core property of factive presuppositions (as of any presupposition) is that they "survive" in entailment-canceling environments. A sentence provides such an environment if it contains an element or structure that manipulates the truth conditions of an equivalent simple, affirmative sentence such that the entailments within that simple sentence disappear. The sentences in (7) to (9) show what this means based on two such elements/structures, negation and question formation. The truth conditions of the simple assertive sentence in (7) are fundamentally different from its negated version in (8) or the polar question in (9). Entailments of the simple sentence (7), e.g., *Someone regrets something*, are not present in (8) and (9). However, like the simple sentence in (7), the sentences in (8) and (9) still contain the presupposition that it is raining, triggered by the factive verb *regret*. This "survival" of the factive presupposition is referred to as "projection": the presupposition projects out of the entailment-canceling environment.

(7) *I regret that it is raining.* [Simple sentence]
Presupposes: *It is raining.*

(8) *I don't regret that it is raining.* [Negation]
Presupposes: *It is raining.*

(9) *Does anybody regret that it is raining?* [Polar question]
Presupposes: *It is raining.*

While the criterion for factivity is relatively clear (presupposition plus projection), it is not always trivial to say whether this criterion applies to an individual clause-embedding predicate. I will demonstrate this by means of an example below. As a result, the classification of a predicate as factive or non-factive may be tricky. Different contributions in the literature on factivity may put the same predicates into different categories. For a general overview of these inconsistencies in classification, I refer to Degen & Tonhauser (2022) and references therein. Here, I will focus on one proposed solution to these difficulties in classification, which splits the class of factive predicates into two sub-classes, semi-factives and true factives (Karttunen 1971),

depending on how consistently they display presupposition projection. This split may be one of several approaches to inconsistent projection patterns; however, it is especially relevant in this paper because it informs Bonke & Repp's (2022) classification of clause-embedding verbs.

The split of factive predicates into true factive and semi-factive sub-types was proposed by Karttunen (1971). Within this split, true factives are assumed to be consistent and invariant in presupposition projection, while semi-factives show variability (see Karttunen 1971; Hooper & Thompson 1973; Hooper 1975 and much subsequent literature). As an example for a semi-factive predicate, Karttunen (1971) gives the verb *discover*; see (10) and (11). Karttunen (1971) argues that *discover* projects a factive presupposition from under negation as in (10), but not from under a polar question as in (11). Thus, *discover* shows variable projection patterns, which differentiates it from other predicates like *regret*, which are assumed to show consistent projection in all entailment-canceling environments. For examples, see (8) and (9) above; cp. also Karttunen (1971) on *regret* specifically. The term "true factive" to refer to consistently presupposing predicates like *regret* is from Hooper & Thompson (1973) and Hooper (1975).

(10) *John didn't discover that he hadn't told the truth.*
Presupposes: *John hadn't told the truth.* (Karttunen 1971: 63)

(11) *Did you discover that you had not told the truth?*
Does not presuppose: *You had not told the truth.* (Karttunen 1971: 63)

This classification of factive predicates raises the question as to whether the split by presupposition behavior correlates with a difference in lexical semantics. Do semi-factives share any common features apart from their variable projectivity? And conversely, do true factives share any common features apart from their (assumed) consistent projectivity? Bonke & Repp (2022) assume that semi-factives are verbs of knowledge and perception (such as *know*, *find out*, *see*), and true factives are emotives (such as *regret*, *hate*, *disapprove*). This assumption has ample precedence in the literature (Hooper & Thompson 1973; Hooper 1975; Beaver 2010; Bîlbîie & de la Fuente 2019, among others).

2.2 Factivity and *que* 'that' in embedded gapping (Bonke & Repp 2022)

Bonke & Repp (2022) derive their hypothesis that factivity and the factive sub-types may be crucial for *que* in gapping from the existing literature on Spanish clausal complementation, where factivity and related properties are sometimes adduced to account for syntactic features in complement clauses (e.g., Demonte & Fernández-Soriano 2009; de Cuba & MacDonald 2013; Bîlbîie & de la Fuente 2019, among others; see also Haegeman 2006; Kastner 2015 on English). As we will see in this section, their experimental results suggest that the type of embedding

verb indeed plays a major role in the acceptability of *que* in gapping. At the same time, the results do not confirm that factivity is the crucial property, leading Bonke & Repp (2022) to reject the factivity hypothesis.

Bonke & Repp's (2022) acceptability study consists of four acceptability judgment experiments. They tested sentences with embedded gapping such as (12) and (13), which varied in the presence vs. absence of *que*, and in the type of embedding verb (e.g., *cree* 'thinks' in (12) vs. *desaprueba* 'disapproves' in (13)). Participants were asked to judge how natural they thought the sentences were within a provided linguistic context (1–2 sentences).

(12) *Santiago* ***cree*** *que el rey fum-a cada día y* { *que* | Ø } *la reina, a veces.*
Santiago thinks that the king smokes-IND every day and that Ø the queen at times
'Santiago thinks that the king smokes every day and that the queen smokes sometimes.'

(13) Santiago ***desaprueba*** *que el rey fum-e cada día y* { *que* | Ø } *la reina, a veces.*
Santiago disapproves that the king smokes-SUBJ every day and that Ø the queen at times
'Santiago disapproves of the fact that the king smokes every day and that the queen smokes sometimes.'

The verbs varied in factivity (non-factive like *cree* vs. factive like *desaprueba*) and mood selection (indicative vs. subjunctive). The difference in mood selection shows in the examples above, where in (12) the embedding verb *cree* 'thinks' selects a complement clause in which the predicate appears in its indicative form (i.e., *fum-a* 'smokes-IND'). In (13), the embedding verb *desaprueba* 'disapproves' selects an embedded predicate in the subjunctive form (i.e., *fum-e* 'smokes-SUBJ').

The study contained the four verb types in Table 1. As discussed in Section 2.1, the factives split into semi-factives and true factives, a distinction that per assumption aligns with different lexical semantic classes: knowledge and perception verbs vs. emotives, respectively. These classes also differ in mood selection preferences, with knowledge and perception verbs favoring indicative complements and emotives favoring subjunctive complements. Non-factives are likewise subdivided by lexical class and mood preference. One set in the sample consists of indicative-preferring non-factives, represented by saying and thinking verbs, which are prototypical non-factives. The other set consists of subjunctive-preferring non-factives, encompassing two lexical semantic classes: desideratives (e.g., *want*, *wish*, *hope*) and verbs with implicit, lexicalized negation (e.g., *deny*, *doubt*, *prohibit*).[2]

[2] Some negation verbs, e.g., *deny* or *doubt*, belong to the class of response stance verbs, which presuppose that the propositional content of their complement is under debate in the discourse (Cattell 1978; Kastner 2015). This presupposition differs from factive presupposition, whereby the content is taken to be true and uncontroversial in discourse.

Table 1: The four verb types in Bonke & Repp (2022)

Factivity	Mood selection	Lexical semantic classes	Examples
Non-factive	Indicative	Verbs of saying and thinking	*decir* 'say' *creer* 'think'
Non-factive	Subjunctive	Desideratives, negation verbs	*querer* 'want' *negar* 'deny'
Factive	Indicative	Verbs of knowledge and perception (= semi-factives)	*saber* 'know' *ver* 'see'
Factive	Subjunctive	Emotives (= true factives)	*lamentar* 'regret' *alabar* 'praise'

Bonke & Repp (2022) hypothesize that non-factive predicates easily allow *que* in gapping, while factives do not allow *que* as easily.[3] This means that with a non-factive embedding verb like *cree* 'thinks' in (12) above, acceptability ratings for the variant with *que* in the gapping conjunct should be comparable to the ratings without *que* (and both should be within the acceptable range). With a factive embedding verb like the emotive *desaprueba* 'disapproves' in (13) above, ratings should be lower with *que* than without *que*. They furthermore hypothesize a difference between the factive sub-types, with semi-factives (knowledge/perception verbs) showing a smaller effect than the true factives (emotives) and thus occupying an intermediate position between non-factives and true factives. The two types of non-factives, i.e., the indicative-taking verbs of saying and thinking on the one hand and the subjunctive-taking desideratives and negation verbs on the other, are not expected to differ.

The experimental results are summarized in Table 2. Gapping without the complementizer received high acceptability ratings across the board and thus serves as the baseline for a by-verb-type comparison to gapping with *que*.[4] The results partially align with the hypotheses: as expected, the indicative-taking non-factives (saying/thinking verbs) overall show no significant difference with vs. without *que* (and ratings indicate high acceptability in general). Also as expected, subjunctive-taking factives (emotives, i.e., true factives) show lower ratings with *que* than without *que*; however, the ratings with *que* are not so low as to indicate outright

[3] The reasons to hypothesize that these verb types show different patterns w.r.t. the acceptability of *que* in gapping stem from observations from the literature on Spanish clause syntax. Generally, complement clauses of factives seem to be more restricted as to the syntactic structures they may contain than complement clauses of non-factives (see e.g., Demonte & Fernández-Soriano 2009: 46 f.; de Cuba & MacDonald 2013: 312, and esp. quantitative evidence from Bîlbîie & de la Fuente 2019). For the sake of conciseness, I will not go into detail on these structural properties of factive complements.

[4] Judgments were given using a slider on a continuous scale (output coding: 0–100). I will keep my discussion of Bonke & Repp's (2022) results on a coarse-grained level and therefore not provide specific numbers. However, in the correlation analysis in Section 3, I use their numeric results.

ungrammaticality. The indicative-taking factives (knowledge/perception verbs, i.e., semi-factives) occupy the expected intermediate position between the subjunctive-taking factives and the (indicative-taking) non-factives. Like the true factives, they show a degradation with *que* compared to without *que*, which, however, is smaller on average than the degradation observed for the true factives and does not show as consistently across experiments. The only verb type that does not pattern as expected is the subjunctive-taking non-factives. Recall that these verbs are predicted to be like the indicative-taking non-factives, i.e., to show no effect of *que*. However, these verbs do show a degradation with *que*, which is comparable in size to that found for the true factives.

Table 2: Summary of Bonke & Repp's (2022) results for the acceptability of embedded gapping with vs. without the complementizer *que* 'that' by the four verb types

Verb type	**Lexical semantic classes**	**Acceptability without *que***	**Acceptability with *que***	**Results as predicted?**
Non-factive + indicative	Saying/thinking	High	High	Yes
Non-factive + subjunctive	Desideratives/ negation verbs	High	**Degraded**	**No**
Factive + indicative	Knowledge/perception (= semi-factives)	High	(Mildly) degraded	Yes
Factive + subjunctive	Emotive (= true factives)	High	Degraded	Yes

It is important to note that these patterns are specific to gapping: in corresponding non-elliptical structures (*…and (that) the queen* smokes *every day*), no degradation of the complementizer was found with subjunctive-taking factives, i.e., the verb type that displays such a degradation most consistently in gapping.

Regarding gapping, Bonke & Repp (2022) furthermore provide a fine-grained examination of the individual verbs within the four types (i.e., 12 verbs per type). This examination shows that their four-way classification of verb types does not ideally capture the variability in the acceptability of gapping with *que*. From the perspective of individual verbs, none of the four verb types exhibits a uniform pattern; instead, distinctions between the types are gradual and emerge only as tendencies. For example, while the subjunctive-taking factives as a type show the largest average degradation with *que*, some of the individual subjunctive-taking factives show considerably smaller degradations, patterning more with the indicative-taking non-factives. Conversely, some of the indicative-taking non-factives show a noticeable degradation, even though taken together as a type they do not show this degradation.

In sum, Bonke & Repp's (2022) results suggest that the acceptability of *que* in gapping is constrained by properties of the embedding verb. When considering the ratings by the four verb types, it is one verb type in particular, indicative-taking non-factives (i.e., saying/thinking verbs), that allows *que* without issue. The other three verb types show different degrees of degradation with *que*.

To account for these findings, Bonke & Repp (2022) propose a syntactic analysis, which assumes a structural difference between gapping with vs. without *que*. According to this proposal, overt *que* indicates a clausal structure of the elliptical conjunct (as opposed to a smaller, sub-clausal structure without *que*). By hypothesis, the clausal structure in this proposal carries specific pragmatic meaning, which is compatible with the meaning of those embedding verbs that show little or no degradation of *que* (i.e., saying/thinking verbs), but clashes with the meaning of the other verbs. This clash in meaning comes at the cost of varying degrees of degraded acceptability, depending on how easily the required interpretation can be constructed with the individual embedding verb. Assuming that the pragmatic property is (non-)factivity, this approach derives the degradation of gapping with *que* under, for example, an emotive embedding verb from the difficulty with which a non-factive sentence meaning could be constructed with such a normally factive verb. Bonke & Repp (2022), however, reject their initial hypothesis that the crucial pragmatic meaning is factivity since the two non-factive sub-types in their experiments do not pattern alike and verb-by-verb analysis reveals heterogeneous effects within types and gradual rather than categorical differences across types.

Nevertheless, it may be premature to reject the factivity-based hypothesis altogether. Such a hypothesis may be coherent with Bonke & Repp's (2022) results if one assumes that their categorization of verbs as factive or non-factive was incorrect in a substantial number of cases. A case in point is the fact that verbs with implicit negation, which formed part of the (subjunctive-taking) non-factives in the experiments, are treated as factive in parts of the literature (e.g., Meinunger 2017). Furthermore, an altogether more sophisticated and nuanced view of factivity may be called for. Recent quantitative research on factivity (Degen & Tonhauser 2022) suggests that factivity is a gradual property rather than a categorical one as often assumed. Coupled with the assumption that some of the verbs were categorized incorrectly, the gradualness of factivity may account for the gradualness in the acceptability data. I pursue this hypothesis in the following section by means of a correlation analysis that connects Bonke & Repp's (2022) acceptability data with quantitative factivity data from Bonke (2024: Ch. 5). Additionally, I experimentally investigate the question of whether the presence of *que* in gapping influences the interpretation of factivity in Section 4.

Alternatively, we may follow Bonke & Repp (2022) in rejecting the factivity-based hypothesis in light of their results. In this case, the question arises as to what determines acceptability instead. Since the type of verb does play a crucial role, an obvious hypothesis would be that instead of

factivity, a different meaning component is the determining factor. This component should relate closely to the indicative-taking non-factives (saying/thinking verbs), since these form the only verb type that showed a negligible effect of *que* in gapping. Bonke & Repp's (2022) analysis focuses on assertion embedding in this regard; Bonke (2024: Ch. 3) mentions at-issueness and reportiveness as further potential meaning components. I explore these options in Section 5.

3 Does the acceptability of *que* 'that' in gapping correlate with the factivity of the clause-embedding verb?

In this section, I connect the acceptability data from Bonke & Repp (2022) on the complementizer *que* 'that' in embedded gapping to independently obtained, quantitative data on factivity from Bonke (2024: Ch. 5). On the basis of the individual embedding verbs, I investigate whether the degree to which *que* is degraded correlates with an experimentally obtained measure of factivity. This means that both measures must receive a numerical value for each embedding verb. For the *que*-degradation, I follow Bonke & Repp (2022) in calculating the difference between the mean acceptability rating in the experimental condition with vs. without *que*. This results in a number between −100 and 100 for every verb,[5] with large positive numbers indicating that the presence of the complementizer leads to a large degradation in acceptability, and small positive numbers indicating that there was a small degradation. Regarding the measure of factivity, I outline Bonke's (2024: Ch. 5) procedure and results in some more detail in the following sub-section, before turning to the correlation analysis of the two measures in Section 3.2.

3.1 Quantitative data on factivity for 44 Spanish verbs (Bonke 2024: Ch. 5)

To obtain a measure of factivity, Bonke (2024: Ch. 5) used (a Spanish translation of) the *certain that*-diagnostic as used by, for example, Tonhauser et al. (2018) and Degen & Tonhauser (2021, 2022). In the experiment, participants judged experimental items such as (14). The items consisted of a sentence containing clausal complementation, which varied as to the embedding verb (sketched in (14) with *afirma* 'claims' and *sabe* 'knows'). The sentences in the items were polar questions, which constituted the necessary entailment-canceling environment to investigate presupposition projection (see Section 2.1).[6] The polar question in turn was the direct speech complement of a speaker (*María*). The experimental materials contained a total of 44 embedding verbs, which also featured in Bonke & Repp's (2022) experiments on embedded gapping.

[5] Bonke & Repp's (2022) collection of acceptability data proceeded on a continuous scale, with the output coded as integers from 0 to 100.

[6] A reviewer points out that the question in (14) shows declarative subject-verb word order, which is possible in Spanish next to verb-subject word order, but may introduce special pragmatic effects. I acknowledge the possibility of such pragmatic effects; however, the crucial property of entailment cancelation is not dependent on word order in interrogatives. Non-inverted word order was used consistently in the experimental items; see Bonke (2024: Ch. 5) for details.

(14) *María: ¿Marco {afirma / sabe / …} que Martina organiza su boda ella sola?*
María: Marco claims knows that Martina organizes her wedding she only
'María: Does Marco claim / know that Martina organizes her wedding all by herself?'

Participants were asked how certain they thought the speaker was that the content of the complement clause was true; see the diagnostic question in (15) for the item in (14) above. They indicated their judgment on a continuous scale with the extremes labeled *no* 'no' and *sí* 'yes' (output coding 1–100). High ratings were thought to indicate high projectivity and therefore high factivity, since participants understood the speaker to take the content of the complement for granted and as uncontroversial in discourse (i.e., presupposed; see Section 2.1 above). Conversely, low ratings were thought to indicate low projectivity, i.e., non-presupposition.

(15) *¿Está María segura de que Martina organiza su boda ella sola?*
is María certain of that Martina organizes her wedding she only
'Is María certain that Martina organizes her wedding all by herself?'

The results show a spread of the verbs along the certainty scale. Some verbs received low certainty ratings (e.g., mean rating of around 30 for *afirma* 'claims' in (14) above) while others received high ratings (e.g., mean rating of around 90 for *sabe* 'knows' in (14) above). This spread along the scale is expected, given the inclusion of (presumed) non-factives (e.g., verbs of saying like *afirma* 'claims') and factives (e.g., knowledge verbs like *sabe* 'knows'). Crucially, however, several verbs did not appear at either end of the scale but instead received intermediate ratings (e.g., *presume* 'presumes', mean rating 57; or *niega* 'denies', mean rating 62). In fact, the results do not overall suggest the existence of well-defined distinct categories (i.e., non-factive vs. factives) at all. Instead, the verbs occupy positions along a spectrum covering the entire scale, which thus blurs the distinction between factive and non-factive.[7]

An interesting finding within these gradual results concerns the verbs with an implicit negation. Recall that Bonke & Repp (2022) classed these as non-factive, while other contributions in the literature treat at least some of them as factives (e.g., Meinunger 2017). In Bonke's (2024: Ch. 5) data, these verbs consistently exhibit intermediate certainty ratings, indicating that they pattern neither with clear non-factives nor with clear factives. The intermediate certainty ratings for this lexical class may explain their inconsistent categorization in the literature. More importantly for the case at hand, the fact that these verbs received unexpectedly high certainty ratings considering Bonke & Repp's (2022) non-factive classification may account for the observed *que*-degradation in their verb type. Recall that the negation verbs, alongside desider-

[7] Degen & Tonhauser (2022) find similar gradual results for English.

atives, formed the subjunctive-taking "non-factives". If enough of these presumed non-factives actually pattern more closely like factives than expected, factivity may still account for the degradation with *que* for their verb type after all.

3.2 Correlating certainty and acceptability ratings

Given that both the certainty ratings from Bonke (2024: Ch. 5) and the effect of gapping-*que* on acceptability as observed in Bonke & Repp (2022) are gradual across the different verbs, a hypothesis regarding their relation may be formulated in gradual terms: the more factive a verb is, the higher its degradation with *que* is expected to be. This hypothesis thus predicts a positive correlation between mean certainty ratings and the average *que*-degradation.

The mean certainty ratings and the average *que*-degradation for the 44 verbs included in both studies are plotted in Figure 1.[8] The vertical axis shows the mean certainty ratings and the horizontal axis shows the average degradation with *que* in embedded gapping (in comparison with the version without *que*). There is one verb, *imagina* 'imagines', with a slightly negative value on the horizontal axis, which means that this verb on average was rated slightly more acceptable with *que* than without.

To test whether the two measures are correlated, I fitted a linear model predicting the average certainty rating per verb from its average degradation with *que* in gapping. The right side of Figure 1 shows the regression line that resulted from this model. The line has a positive slope

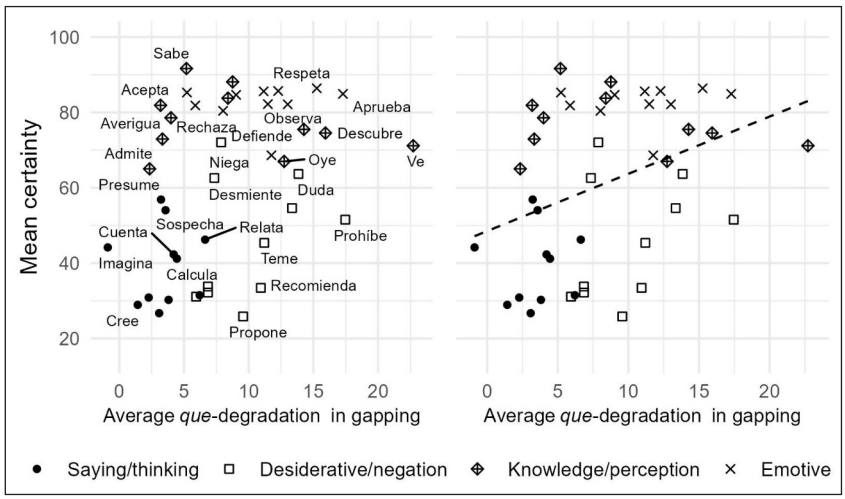


Figure 1: Correlation of the mean degradation of gapping with *que* (Bonke & Repp 2022) and mean certainty rating for the 44 verbs (Bonke 2024: Ch. 5), with labels (left) and regression line (right)

[8] All statistical analyses in this paper were done in R (R Core Team 2026). Figures were generated using ggplot2 (Wickham 2016).

(b = 1.52), which indicates that as predicted, verbs with a higher certainty rating show a higher degradation of *que*. The model indicated that this correlation is statistically significant (p = 0.02). However, the model indicated a low correlation coefficient (adjusted R^2 = 0.11), which means that despite the significant correlation of average certainty with *que*-degradation, certainty can account only for a small part (11%) of the variation within the *que*-degradation data. This lack of explanatory power also shows in the relatively large residuals in Figure 1 (i.e., the distance of the individual data points from the regression line).

The low R^2 is unexpected. If factivity really is the crucial, determining factor for the acceptability of gapping-*que*, this coefficient should be larger and the correlation should be stronger. Despite the statistical significance, the weakness of the correlation casts doubt on the role of factivity for *que*. Another problem results from the role of one particular verb type, the indicative-taking non-factives (= verbs of saying and thinking). Recall that this verb type was the only one that consistently showed only a negligible degradation of *que* (see Section 2.2 above). The low average *que*-effect for this verb type is visible in Figure 1, where the black dots representing the verbs of this type range near the left end of the horizontal scale. This verb type also seems to drive the correlation with the certainty ratings: the black dots are concentrated in an area toward the bottom left in Figure 1, which creates the impression that they are "pulling down" the lower end of the regression line.

To find out whether the indicative-taking non-factives indeed drive the correlation, let us look at the data once more without this verb type. Figure 2 shows the resulting correlation of mean certainty rating and average *que*-degradation in gapping. This figure shows the same data

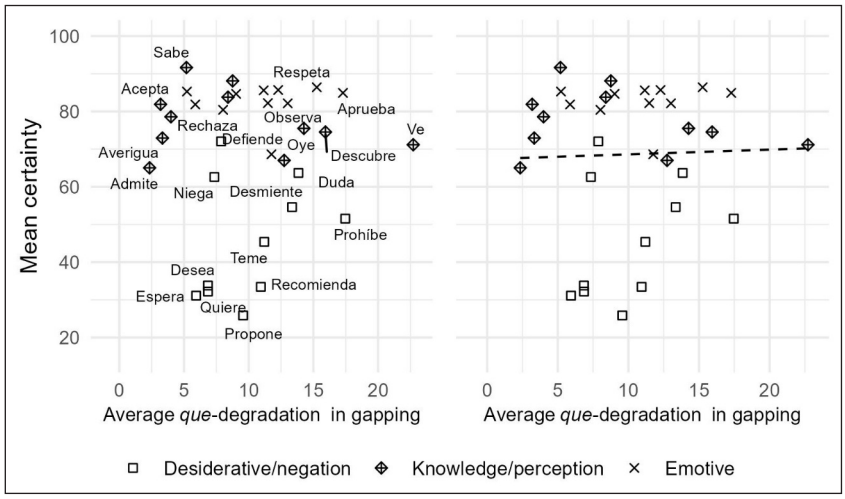


Figure 2: Correlation of the mean degradation of gapping with *que* (Bonke & Repp 2022) and mean certainty rating (Bonke 2024: Ch. 5) without the verbs of saying and thinking, with labels (left) and regression line (right)

as Figure 1 above, the difference being that the black dots representing the indicative-taking non-factives have been taken out. Without these verbs, the figure still shows a large spread of data points in both the vertical and horizontal directions. This means that the data still includes verbs with high and low certainty ratings (vertical axis) and verbs with large and little *que*-degradation (horizontal axis). If factivity determines the acceptability of *que* in gapping, the correlation of certainty ratings and *que*-degradation that we saw above should thus also surface here. However, a linear model fitted to these data indicates a much smaller effect (b = 0.13), which is visualized by the near horizontal regression line on the right side of Figure 2. The linear model furthermore indicated that the effect was not statistically significant (p = 0.87) and that the correlation coefficient was close to zero (adjusted R^2= −0.03). The degradation with *que* thus does not account for certainty ratings in this sub-set of the data, contrary to expectation.

All in all, the correlations of certainty ratings and *que*-degradation presented in this section provide evidence against the hypothesis that factivity is responsible for the *que*-effect in gapping. This conclusion holds even though there is a statistically significant positive correlation of the average certainty rating and *que*-degradation for each verb. The low correlation coefficient, paired with the driving force of the indicative-taking non-factives, shows that certainty, and thus by extension factivity, are unlikely to play a decisive role in the *que*-effect. Instead, the decisive factor is likely an inherent property of the verb type that drives the correlation, i.e., the indicative-taking non-factives (verbs of saying/thinking). This result parallels the conclusions Bonke & Repp's (2022) draw from their acceptability data, which likewise suggests that these verbs are crucial for *que* in gapping, independently of factivity. In the following sections, I try to pin down the special role of the verbs of saying and thinking by presenting two rating experiments on the meaning of *que* in gapping.

4 Experiment 1: Factivity

In the following two sections, I approach the meaning contribution of the complementizer *que* 'that' in Spanish embedded gapping. Specifically, I investigate whether there is a difference in meaning between gapping with *que* and without *que*. The motivation for this procedure stems from Bonke & Repp (2022), who suggest that the verb-type effects they observe may be accounted for under the hypothesis that *que* in gapping is indicative of an aspect of sentence meaning (see Section 2.2). The potential candidates for the relevant sentence meaning they list are (non-) factivity (which is the focus of this section), assertion, at-issueness, and reportiveness (which will be approximated in Section 5).

Bonke & Repp's (2022) reasoning leads to testable hypotheses. If *que* in gapping is associated with non-factive meaning, gapping with *que* is expected to receive a non-factive interpretation across the board. Without *que*, interpretation should be determined by the embedding verb, thus

paralleling the patterns presented in Section 3.1. This predicts that for factive embedding verbs, the effect of gapping-*que* should be strong, lowering certainty ratings to the level of non-factive verbs. Non-factive verbs should receive low ratings, irrespective of the presence of *que*.

The correlation analysis in Section 3.2 suggested that factivity is unlikely to be responsible for the *que*-effect in gapping. However, the comparison of factivity and acceptability was indirect, comparing acceptability data to independent factivity data. From a methodological viewpoint, it is thus justified to return to the issue of factivity once more by means of a more direct approach. Experiment 1 thus applies the *certain that*-diagnostic for factivity (see Section 3.1) to Bonke & Repp's (2022) gapping materials.

4.1 Design, materials, procedure, and participants

Experiment 1 had a 1×2 design, the sole factor being the presence vs. absence of *que* in the elliptical conjunct. For a sample item in both conditions, see (12). This item is identical to an item from Bonke & Repp (2022), with the only difference being that it is presented in a direct speech context. This adjustment of the items was necessary since the *certain that*-diagnostic measures the certainty of the speaker (i.e., *María* in (16), cp. Section 3.1 above).[9] Items were preceded by 1–2 sentences of conversational context.

(16) *María:* *Santiago cree que el rey fuma cada día y* {*que* | Ø} *la reina, a veces.*
María Santiago thinks that the king smokes every day and that Ø the queen at times
'María: Santiago thinks that the king smokes every day and that the queen smokes sometimes.'

The materials contained the 48 embedding verbs of Bonke & Repp's (2022) study and 24 embedded clauses. Each verb combined with two embedding clauses to form a total of 96 sentences like (16), which in the two conditions made a total of 192 test items. The items were distributed across eight lists in a Latin square design, balancing the four verb types. Each list contained 24 test items so that each participant judged each embedded clause once and half of the embedding verbs. In addition to the test items, every list featured ten declarative control items and ten question control items like (17) and (18), which provided baselines for certainty and non-certainty, respectively. Each list thus contained 44 sentences to be judged.

[9] As can be seen in (12), the experimental items did not feature entailment-canceling environments. The use of the *certain that*-diagnostic on these sentences thus raises conceptual difficulties since such environments are crucial for the investigation of (factive) presuppositions (see Section 2.1 above). The certainty ratings for the experimental items may not be straightforwardly equated with factivity, since high certainty ratings may simply reflect the speaker's certainty of their assertion rather than presupposition. However, the ratings for the individual verb types will be consistent with what would be expected for presupposition (see below).

(17) *María: Julio adora el tiempo soleado y fantástico de Gran Canaria.*
María Julio loves the weather sunny and fantastic of Gran Canaria
'María: Julio loves the sunny and fantastic weather of Gran Canaria.'

(18) *María: ¿Elena estudia toda la noche para un examen?*
María Elena studies all the night for an exam
'María: Does Elena spend all night studying for an exam?'

Participants were asked to imagine that they were attending a party hosted by their friend María. María was described as a chatty person who loved talking about other people. Upon entering the kitchen, participants were asked to imagine overhearing a statement (= the experimental items and declarative fillers) or question (= the question fillers) that María made, within the conversational context provided.[10] Participants then indicated how certain they thought María was regarding the content of the elliptical conjunct in case of the target items, or the whole sentence in case of the declarative and question controls. The diagnostic was the *certain that*-diagnostic, e.g., *¿María está segura de que la reina fuma a veces?* 'Is María certain that the queen smokes sometimes?' (see Section 3.1 above for details). To give their ratings, participants used a slider on a continuous scale with the extremes *no* 'no' and *sí* 'yes'. Responses were coded as integers from 1 to 100.

The experiment was an online questionnaire hosted on soscisurvey.de (Leiner 2024). I used prolific.com to recruit monolingual native speakers of European Spanish, who grew up speaking only their native language. In total, 80 participants completed the questionnaire. Only data from participants showing a statistically significant difference between declarative and question controls (assessed via per-participant t-tests) was included in the analysis. All 80 participants met this criterion; therefore, all data entered the analysis.

4.2 Results and discussion

The results are in Figure 3. The ratings for the filler controls show that the *certain that*-diagnostic worked as expected: declarative fillers received high certainty ratings and question fillers received low certainty ratings. The ratings for the two experimental conditions, present

[10] I acknowledge that it may be difficult to imagine being in a situation where you accidentally overhear the utterance someone has made, but where you are simultaneously aware of the conversational context of that utterance. I proceed with the assumption that the participants were not so confused by this slight incoherence in the experimental setup that it affected their ratings tremendously. A further issue with the experimental setup was the varying degrees of degradation of the items in the condition with *que*. These degraded items may have been difficult for participants to interpret regarding speaker certainty. To alleviate this difficulty, the instructions made participants aware that not all of María's utterances may sound very natural, and that they should try to give an intuitive judgment as well as they could.

vs. absent *que*, show hardly any discernible difference, let alone the predicted strong effect.[11] I conclude that *que* in gapping does not determine speaker certainty and, by extension, factivity.

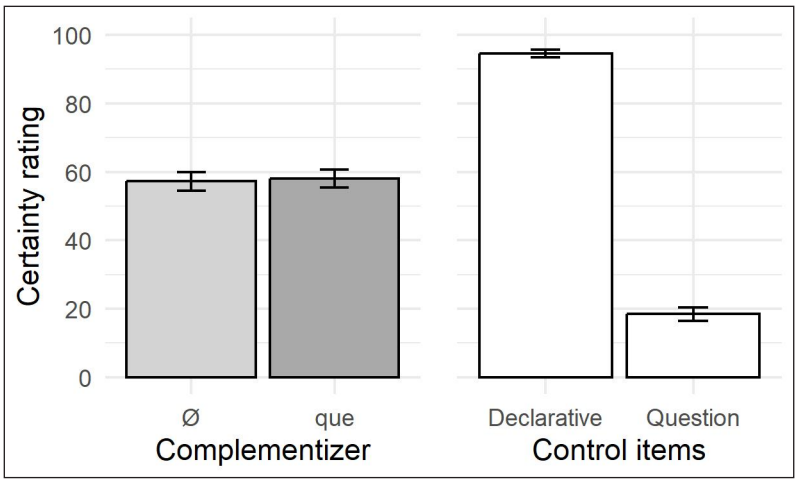


Figure 3: Mean ratings and 95% confidence intervals for experimental conditions and controls (Exp. 1)

A closer look at the four individual verb types reinforces this conclusion. As can be seen in Figure 4, the effect of gapping-*que* on certainty ratings within the verb types is just as negligible as for the overall data. The two factive verb types received higher certainty ratings than the two non-factive types, which suggests that the diagnostic likely picked up on factivity, at least in the items with embedded gapping, where the diagnostic targeted the complement clause.

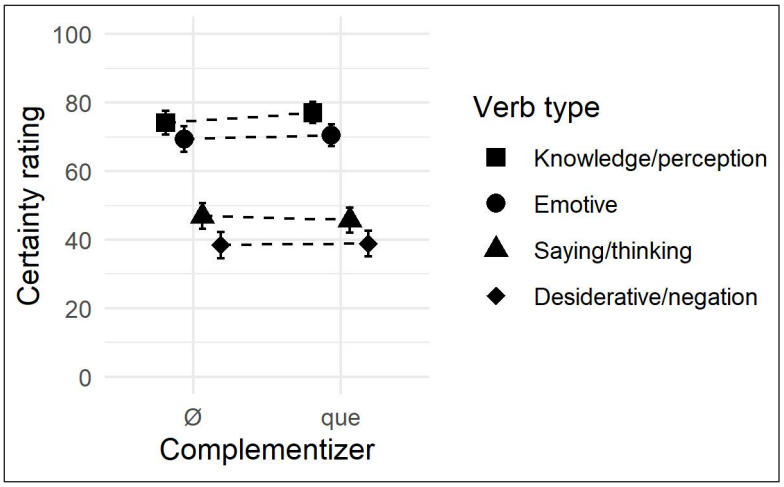


Figure 4: Mean ratings and 95% confidence intervals by verb type and experimental condition (Exp. 1)

[11] Since the size of the effect (or rather lack thereof) sufficiently illuminates the role of *que* in the interpretation of speaker certainty, I did not conduct an inferential analysis on the data. I do not believe the results of such an analysis would have yielded any insights into the questions at hand.

5 Experiment 2

Experiment 1 strongly suggested that the presence of the complementizer *que* 'that' in Spanish embedded gapping has no effect on the interpretation of factivity. This result raises the question of whether *que* affects meaning at all. To explore this issue, Experiment 2 takes a more general approach by applying further pragmatic diagnostics to gapping, which unlike the *certain that*-diagnostic for factivity, have no precedence in the literature. Thus, in Experiment 2, no robust diagnostics target any one specific pragmatic property. Instead, multiple diagnostics are applied in a coarse-grained approach to see whether any effect of *que* on interpretation can be recorded at all, which gives Experiment 2 an exploratory character.

5.1 Experimental materials, diagnostics, and predictions

The 1×2 factorial design and the crucial test items (i.e., embedded gapping with vs. without *que* and different embedding verbs) in Experiment 2 were identical to those of Experiment 1. See Section 4.1 above for a description and see example (16) in that section for a sample item. For each experimental item, there were three statements targeting the interpretation of the elliptical conjunct in the embedded clause. Participants indicated whether they agreed or disagreed with the three statements by clicking a box labeled *sí* 'yes' or another labeled *no* 'no'. For the example item in (16) above, participants gave judgments on the three statements in (19) to (21).

(19) *Possible that not*-diagnostic
Es posible que la reina no fume a veces.
is possible that the queen NEG smokes at times
'It is possible that the queen does not smoke sometimes.'

(20) Importance diagnostic
Lo más importante que dice María es que la reina fuma a veces.
the most important that says María is that the queen smokes at times
'The most important thing that María is saying is that the queen smokes sometimes.'

(21) Reportiveness diagnostic
Santiago le contó algo nuevo a María, concretamente que la reina fumaba a veces.
Santiago her told something new to María specifically that the queen smoked at times
'Santiago told María something new, specifically that the queen smokes at times.'

As mentioned above, the three diagnostics are not standard measurements of any linguistic property. However, they offer an approximation of properties that characterize verbs of saying and thinking, i.e., those verbs that consistently show no degradation of *que* in gapping (Bonke & Repp

2022, cp. Section 2.2 above) and drive the correlation of acceptability and certainty as discussed in Section 3.2. These verbs have been proposed to embed assertive (see Hooper & Thompson 1973; Hooper 1975, and much subsequent work) and at-issue content. These two interconnected properties are targeted by the *possible that not-* and importance diagnostics. The *possible that not*-diagnostic simulates the rejection of the content of the elliptical clause in dialogue, i.e., the potential for an interlocutor to reply with, e.g., *No, that's not true, the queen does not smoke sometimes.*[12] The potential to be targeted for rejection has been connected to asserted content, e.g., by Farkas & Bruce (2009: 82). The importance diagnostic probes whether the content of the elliptical clause represents the main point of the utterance, which captures at-issueness in an intuitive way, see, for example, Simons et al. (2010), who use the term "main point" in a pre-theoretical description of this concept. The reportiveness diagnostic straightforwardly relates to the lexical semantics of saying verbs and extends to many instances of thinking verbs in actual use.

The general reasoning of Experiment 2 follows that of Experiment 1: if gapping with *que* contributes its own meaning component, it should yield a consistent interpretation across verb types (as measured by the proportions of *yes* vs. *no* responses). If this meaning is compatible with verbs of saying and thinking, the presence or absence of *que* should not affect interpretation for these verbs. For the other verb types, gapping with *que* is expected to pattern like saying/thinking verbs, whereas without *que*, interpretations may differ. In other words, saying/thinking verbs should show no effect of *que*, while other verbs should match them with *que* and potentially diverge without it.

An analysis of the filler items included in the experiment revealed that the three diagnostics were likely not ideal to measure specific pragmatic properties like assertion or at-issueness. An in-depth description and critical discussion of the filler items are beyond the scope of this paper, but see Bonke (2024: Ch. 5) for details. Some of the inadequacies of the diagnostics will show in the discussion of the results. Nonetheless, the results are informative regarding the research question of this paper: if *que* in gapping indeed carries meaning, at least one of the diagnostics should suggest an effect for at least some of the verb types, even if it does not perfectly capture the specific pragmatic it targets. I will therefore proceed in describing the experiment using descriptive terms, i.e., without referring to the technical notions at-issueness or assertion.

Experimental items were distributed across eight lists (Latin square). In addition to 24 experimental items, each list contained 16 filler items, so that each participant gave three judgments for 40 sentences. The experiment was an online questionnaire hosted on soscisurvey.de, with participant recruitment via prolific.com. The data of 76 native speakers of European Spanish entered the analysis.

[12] Due to the experimental goal of collecting multiple ratings for each item, it was not feasible to present items in dialogue, i.e., in a format in which the content of the gapping clause was rejected by an interlocutor in discourse.

5.2 Results

The results for the three diagnostics are given below in Figures 5 (*possible that not*), 6 (importance), and 7 (reportiveness). As can be seen in these figures, the effect of *que* on the interpretation of these properties is very small across the board. For all three properties, these small effects obtain for the total data including all verbs (left sides of the figures), as well as for each of the four individual verb types (right sides of the figures). Even for those effects that seem to be stronger than others (e.g., the importance ratings for the saying/thinking verbs in Figure 6 below), the error bars indicating the 95% confidence intervals overlap by a considerable margin. These results do not suggest that *que* in gapping has a strong contribution to pragmatic meaning.[13]

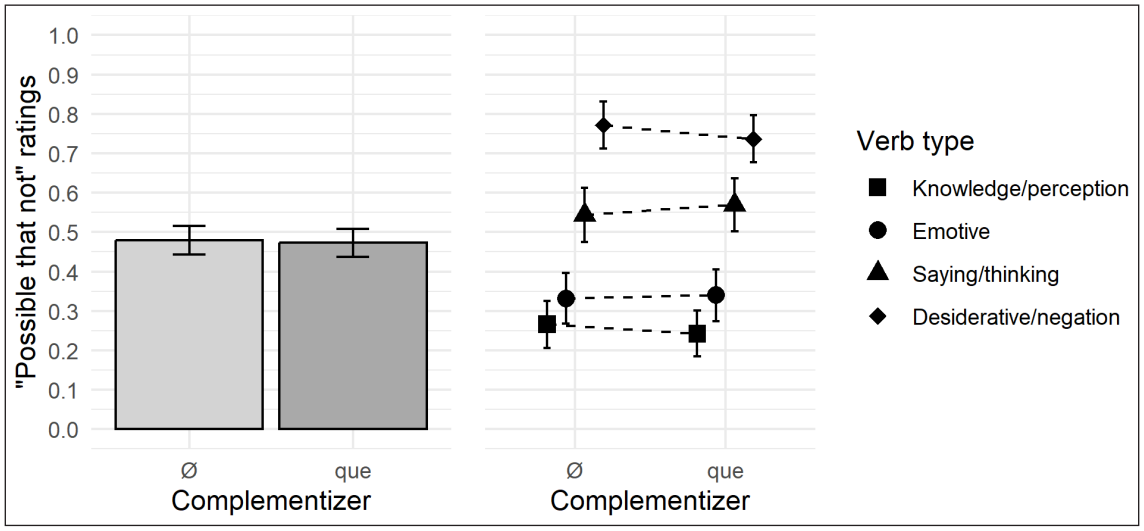


Figure 5: *Possible that not*-diagnostic, Exp. 2 (proportion of *yes*-responses and 95% confidence intervals by experimental condition); all embedding verbs (left) and the four verb types (right)

The results for the *possible that not*-diagnostic (Figure 5) broadly aligned with the initial hypotheses for three of the verb types: both factive types showed relatively low proportions of *yes*-responses with little difference between them, and the verbs of saying and thinking showed a higher proportion, as predicted. However, the desideratives and negation verbs did not match predictions, since instead of patterning with the factives, they yielded the highest proportion of *yes*-responses overall. This finding suggests that the diagnostic did not tap into a property that characterizes verbs of saying and thinking. Possibly, a negative inference associated with the desideratives and negation verbs is responsible for the high proportion of *yes*-responses.

[13] Since a visual inspection of the figures provides sufficient evidence for this conclusion, I will not provide any inferential statistics for the data of Experiment 2. I do not believe that the discussion of *p* values, had they indicated statistical non-significance or (against expectations) statistical significance, would contribute anything to our understanding of the meaning of *que* in gapping.

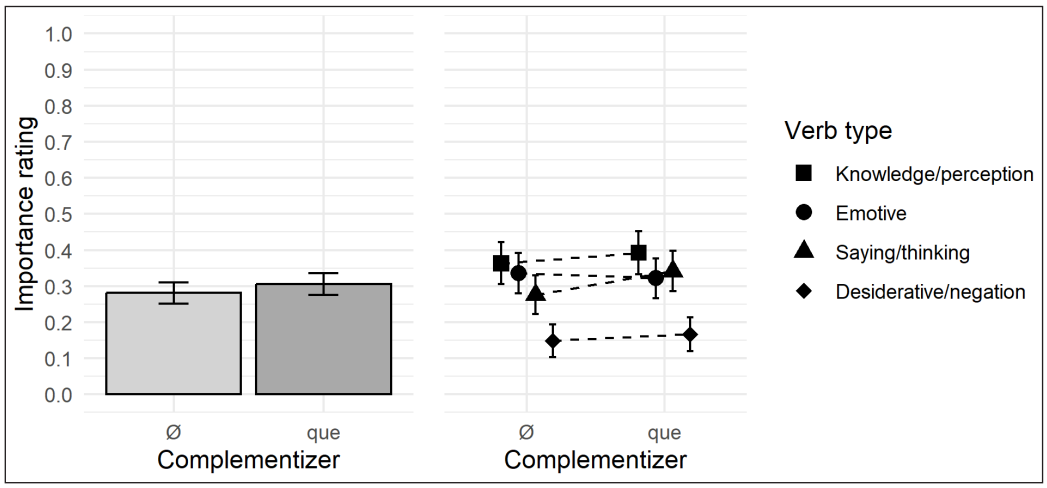


Figure 6: Importance diagnostic, Exp. 2 (proportion of *yes*-responses and 95% confidence intervals by experimental condition); all embedding verbs (left) and the four verb types (right)

The importance ratings in Figure 6 are surprisingly low (below 0.5) across the board. The indicative-taking non-factives should have received much higher ratings, given the assumptions outlined above. A likely explanation for these low ratings is the fact that the experimental items featured coordinate embedded clauses. Test items contained two coordinate clauses with contents of equal value, e.g., [*the king smokes every day*] and [*the queen smokes sometimes*] in (16) above (Section 4.1). The importance diagnostic for this item in (20) above carries a uniqueness inference triggered by the superlative *lo más importante* 'the most important thing'. This inference may have made participants look for one single "most important thing" in the sentence, which may have clashed with the presence of two equally important conjuncts. The analysis of coordinate vs. simple filler items in Bonke (2024: Ch. 5) confirms this suspicion.

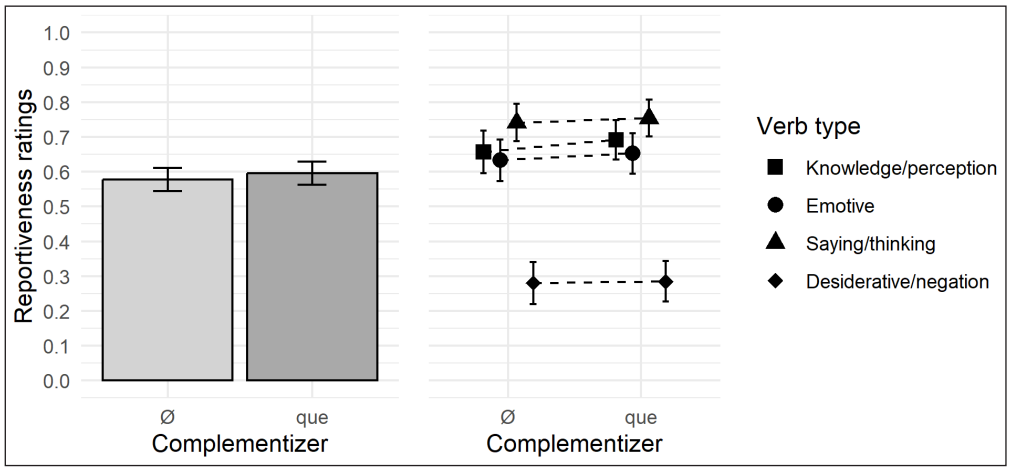


Figure 7: Reportiveness diagnostic, Exp. 2 (proportion of *yes*-responses and 95% confidence intervals by experimental condition); all embedding verbs (left) and the four verb types (right)

Finally, the reportiveness ratings are in Figure 7. Once more, the indicative-taking non-factives should receive a relatively high proportion of *yes*-responses, given that this type contains verbs of saying and thinking. While these verbs do receive the highest *yes*-proportion, the two factive types receive similarly high *yes*-proportions, which do not seem to suggest a categorical difference between these three verb types. Only the subjunctive-taking non-factives receive the expected low proportion of *yes*-responses. This result is surprising because it seems to suggest that emotive verbs (i.e., the subjunctive-taking factives) and verbs of knowledge and perception (i.e., the indicative-taking factives) may quite easily receive reportive ratings.

In sum, the results of Experiment 2 do not provide any evidence for the hypothesized effect of *que* in gapping on interpretation. All three diagnostics deliver null results, both for the pooled data for all verbs as well as the data of the four individual verb types. Additionally, a critical examination of the diagnostics suggests that they do not ideally capture specific pragmatic properties (like assertion or at-issueness). Nevertheless, the results of Experiment 2 are at least somewhat informative: whether or not *que* surfaces in the elliptical conjunct in gapping is most likely not associated with a categorical difference in sentence meaning.

6 Conclusion

In this paper, I set out to investigate the role of the complementizer *que* 'that' in Spanish embedded gapping, with a particular focus on whether its presence is constrained by pragmatic properties. Previous work (Bonke & Repp 2022) investigated the role of such pragmatic properties, especially factivity, as potential explanations for verb-type effects in the acceptability of *que*. The present study presented a new correlation analysis and two experiments to reassess this hypothesis more directly.

The correlation analysis of acceptability data with independently collected quantitative measures of factivity showed only a weak link between the two. While there was a statistically significant correlation, its explanatory power was low, and further inspection revealed that the effect was largely driven by one verb type, namely verbs of saying and thinking. Once these verbs were removed from the analysis, the correlation disappeared, indicating that factivity is unlikely to be the decisive factor in determining the acceptability of *que*.

Experiment 1 tested more directly whether the presence of *que* influences the interpretation of factivity by applying the *certain that*-diagnostic (e.g., Tonhauser et al. 2018) to gapping sentences with and without *que*. The results confirmed the conclusion of the correlation analysis by showing no systematic differences: certainty ratings were unaffected by the complementizer, suggesting that *que* does not alter the interpretation of factivity in embedded gapping.

Experiment 2 extended the investigation beyond factivity to other possible pragmatic contributions of *que* in gapping, using exploratory diagnostics. Across diagnostics and verb types, the results were again null: the presence or absence of *que* did not produce meaningful differences in interpretation. While the diagnostics themselves may not have been ideal tools for measuring specific pragmatic properties, the consistent lack of meaningful effects nevertheless suggests that *que* does not contribute to meaning in the way previously hypothesized.

The origin of the differential sensitivity to gapping-*que* exhibited by different embedding verbs as observed by Bonke & Repp (2022) thus remains unaccounted for. However, based on the present results, we may confidently conclude that the categorization of embedded verbs along pragmatic properties, especially factivity, yields no insight into that origin. Instead, it may be fruitful for future research into *que* in gapping to give up on semantic or pragmatic partitions of embedding verbs altogether and instead turn to the investigation of other potential factors. Unfortunately, the present study does not seem to provide any suggestions as to what these factors might be.

Data Availability Statement: The data and analyses for the two experiments presented in this paper are available on OSF: https://osf.io/r4shg.

Acknowledgements: I would like to thank the audience at *Experimental and Corpus-based Approaches to Ellipsis* 2024 in Poznań and three anonymous reviewers for their helpful comments and discussions. I furthermore thank Sophie Repp and Judith Tonhauser for their guidance and support in the preparation of my PhD dissertation, a part of which forms the basis for this paper.

References

Beaver, David. 2010. Have you noticed that your belly button lint color is related to the color of your clothing? In Rainer Bauerle, Uwe Reyle & Thomas Zimmermann (eds.), *Presuppositions and discourse: Essays offered to Hans Kamp*, 65–100. Leiden: Brill. https://doi.org/10.1163/9789004253162_004.

Bîlbîie, Gabriela & Israel de la Fuente. 2019. Can gapping be embedded? Experimental evidence from Spanish. *Glossa* 4(1). 110. https://doi.org/10.5334/gjgl.782.

Bonke, Max. 2024. *Complementizers in ellipsis*. Köln: Universität zu Köln dissertation.

Bonke, Max & Sophie Repp. 2022. Complementizer deletion in embedded gapping in Spanish. *The Linguistic Review* 39(3). 525–555. https://doi.org/10.1515/tlr-2022-2096.

Broekhuis, Hans & Norbert Corver. 2019. *Syntax of Dutch: Coordination and ellipsis*. Amsterdam: Amsterdam University Press.

Cattell, Ray. 1978. On the source of interrogative adverbs. *Language* 54(1). 61–77. https://doi.org/10.2307/412999.

de Cuba, Carlos & Jonathan E. MacDonald. 2013. On the referential status of embedded polarity answers in Spanish. *Hispanic Linguistics Symposium* 16. 312–323.

Degen, Judith & Judith Tonhauser. 2021. Prior beliefs modulate projection. *Open Mind* 5. 59–70. https://doi.org/10.1162/opmi_a_00042.

Degen, Judith & Judith Tonhauser. 2022. Are there factive predicates? An empirical investigation. *Language* 98(3). 552–591. https://doi.org/10.1353/lan.0.0271.

Demonte, Violeta & Olga Fernández-Soriano. 2009. Force and finiteness in the Spanish complementizer system. *Probus* 21(1). 23–49. https://doi.org/10.1515/prbs.2009.002.

Farkas, Donka F. & Kim B. Bruce. 2009. On reacting to assertions and polar questions. *Journal of Semantics* 27(1). 81–118. https://doi.org/10.1093/jos/ffp010.

Fiengo, Robert W. 1974. *Semantic conditions on surface structures*. Cambridge, MA: MIT dissertation. http://hdl.handle.net/1721.1/13018.

Haegeman, Liliane. 2006. Conditionals, factives and the left periphery. *Lingua* 116(10). 1651–1669. https://doi.org/10.1016/j.lingua.2005.03.014.

Hankamer, Jorge. 1979. *Deletion in coordinate structures*. New York & London: Garland.

Hartmann, Katharina. 2000. *Right node raising and gapping: Interface conditions on prosodic deletion*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.106.

Hooper, Joan B. 1975. On assertive predicates. In John P. Kimball (ed.), *Syntax and Semantics,* vol. 4, 91–124. Leiden: Brill. https://doi.org/10.1163/9789004368828_005.

Hooper, Joan B. & Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4(4). 465–497. http://www.jstor.org/stable/4177789.

Johnson, Kyle. 2009. Gapping is not (VP-) ellipsis. *Linguistic Inquiry* 40(2). 289–328. https://doi.org/10.1162/ling.2009.40.2.289.

Johnson, Kyle. 2018. Gapping and stripping. In Jeroen van Cranenbroek & Tanja Temmerman (eds.), *The Oxford handbook of ellipsis,* 562–604. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198712398.013.24.

Karttunen, Lauri. 1971. Some observations on factivity. *Papers in Linguistics* 4. 55–69. https://doi.org/10.1080/08351817109370248.

Kastner, Itamar. 2015. Factivity mirrors interpretation: The selectional requirements of presuppositional verbs. *Lingua* 164. 156–188. https://doi.org/10.1016/j.lingua.2015.06.004.

Kiparsky, Paul & Carol Kiparsky. 1970. Fact. In Manfred Bierwisch & Karl Erich Heidolph (eds.), *Progress in linguistics*, 143–173. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783111350219.143.

Leiner, Dominik J. 2024. *SoSci Survey (Version 3.5.02)*. https://www.soscisurvey.de, 06 March 2026.

Meinunger, André. 2017. Desires, wishes and hopes – desiderative predicates and presuppositions. *STUF - Language Typology and Universals* 70(4). 611–635. https://doi.org/10.1515/stuf-2017-0026.

R Core Team. 2026. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/, 06 March 2026.

Repp, Sophie. 2009. *Negation in gapping*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199543601.001.0001.

Ross, John R. 1968. *Constraints on variables in syntax*. Bloomington, Indiana: Indiana Univ. Linguistics Club.

Ross, John R. 1970. Gapping and the order of constituents. In Manfred Bierwisch & Karl Erich Heidolph (eds.), *Progress in Linguistics*, 249–259. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783111350219.

Sag, Ivan A. 1977. *Deletion and logical form*. Bloomington, Indiana: Indiana Univ. Linguistics Club.

Simons, Mandy, Judith Tonhauser, David Beaver & Craige Roberts. 2010. What projects and why. In Nan Li & David Lutz (eds.), *Proceedings of SALT 20*, 309–327. Ithaca, NY: CLC Publications.

Stalnaker, Robert. 1978. Assertion. In P. Cole (ed.), *Syntax and semantics*, vol. 9, 315–332. New York: Academic Press.

Stalnaker, Robert. 2002. Assertion. In Paul Portner & Barbara H. Partee (eds.), *Formal semantics: The essential readings*, 147–161. Oxford: Blackwell. [Reprint of Stalnaker (1978).]

Taraldsen, Knut Tarald. 2000. Rethinking gapping and the order of constituents. In Øivin Andersen, Kjersti Fløttum & Torodd Kinn (eds.), *Menneske, språk og fellesskap*, 211–224. Oslo: Novus.

Tonhauser, Judith, David I. Beaver & Judith Degen. 2018. How projective is projective content? Gradience in projectivity and at-issueness. *Journal of Semantics* 35(3). 495–542. https://doi.org/10.1093/jos/ffy007.

Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York: Springer. https://ggplot2.tidyverse.org.

Wilder, Chris. 1994. Coordination, ATB and ellipsis. In Jan-Wouter Zwart (ed.), *Minimalism and Kayne's asymmetry hypothesis*, 291–329. Groningen: Rijksuniversiteit.

Wilder, Chris. 1997. Some properties of ellipsis in coordination. In T. Alan Hall & Artemis Alexiadou (eds.), *Studies on universal grammar and typological variation*, 59–107. Amsterdam: John Benjamins.