



Philosophische Fakultät

Christopher Patrick Auhagen | Melanie Uth

Variation of relative complementizers in Yucatecan Spanish:

A comparison of monolingual and bilingual speakers

Suggested citation referring to the original publication:

Languages 7 (2022) 4, Art. 279

DOI <https://doi.org/10.3390/languages7040279>

ISSN 2226-471X

Secondary publication archived on the publication server of the
University of Potsdam in:

Zweitveröffentlichungen der Universität Potsdam : Philosophische Reihe 181

ISSN: 1866-8380

<https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-584375>

DOI: <https://doi.org/10.25932/publishup-58437>

Article

Variation of Relative Complementizers in Yucatecan Spanish: A Comparison of Monolingual and Bilingual Speakers

Patrick Auhagen * and Melanie Uth *

Department of Romance Linguistics, University of Potsdam, 14469 Potsdam, Germany

* Correspondence: christopher.patrick.auhagen@uni-potsdam.de (P.A.); uth@uni-potsdam.de (M.U.)

Abstract: The starting point of this article is the occurrence of determiner-less and bare *que* relative complementizers like (*en que*, '(in) that', instead of (*en el que*, '(in) which', in Yucatecan Spanish (southeast Mexico). While reference grammars treat complementizers with a determiner as the standard option, previous diachronic research has shown that determiner-less complementizers actually predate relative complementizers with a determiner. Additionally, Yucatecan Spanish has been in long-standing contact with Yucatec Maya. Relative complementation in Yucatec Maya differs from that in Spanish (at least) in that the non-complex complementizer *tu'ux* ('where') is generally the only option for locative complementation. The paper explores monolingual and bilingual data from Yucatecan Spanish to discuss the question whether the determiner-less and bare *que* relative complementizers in our data constitute a historic remnant or a dialectal recast, possibly (but not necessarily) due to language contact. Although our pilot study may not answer these far-reaching questions, it does reveal two separate, but intertwined developments: (i) a generally increased rate of bare *que* relative complementation, across both monolingual speakers of Spanish and Spanish Maya bilinguals, compared to other Spanish varieties, and (ii) a preference for *donde* at the cost of other locative complementizer constructions in the bilingual group. Our analysis thus reveals intriguing differences between the complementizer preferences of monolingual and bilingual speakers, suggesting that different variational patterns caused by different (socio-)linguistic factors can co-develop in parallel in one and the [same] region.

Keywords: relative complementation; variability; language contact; diachrony; Yucatecan Spanish



Citation: Auhagen, Patrick, and Melanie Uth. 2022. Variation of Relative Complementizers in Yucatecan Spanish: A Comparison of Monolingual and Bilingual Speakers. *Languages* 7: 279. <https://doi.org/10.3390/languages7040279>

Academic Editor: Ian MacKenzie

Received: 17 June 2022

Accepted: 18 October 2022

Published: 1 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The present article investigates relative complementizer variation in Yucatecan Spanish, a variety of Spanish that is spoken in southeast Mexico and has been in long-term contact with Yucatec Maya. Relative complementizers are structures that connect a relative clause to a superordinate clause and are found in different sentence configurations, including clefts. A controversial topic in the discussion of complementizer variation in Spanish is the occurrence of bare *que* ('who, that'), where, according to reference grammars, standard varieties require other forms (e.g., *el que* 'who, that', *quien* 'who', *en el que* 'in which', etc.); this controversy has received a fair amount of attention among researchers. This phenomenon has been observed in a number of dialects of Spanish in regions such as in Colombia, Venezuela, the Antilles (Henríquez Ureña 1921), and Argentina (Butt and Benjamin 2000).

The present work has been motivated by an elicitation experiment conducted within a research project on language contact in Quintana Roo, México, where we found a notable degree of inter- and intra-speaker variation of bare *que* and complex forms (i.e., *en el/la que*) in relative clauses in a naturalistic speech corpus (cf. 1a and 1b).¹

- 1 a. *¿Por qué? Por el contexto en el que se encuentra...* (elic03_nmc)
'Why? Because of the context in which [he] finds himself . . . '
- b. *Hay un lugar en Veracruz que hablan con muchas groserías.* (elic03_nmc)
'There is a place in Veracruz that [where] they speak in a vulgar manner.'

Furthermore, in other syntactic configurations involving relative complementation, namely, subject and object clefts, we observed a remarkable number of determiner-less *que* in comparison to complex forms (*el/la que, quien*) in both naturalistic and elicited data (cf. 2a and 2b).

- 2 a. *No, es Cantinflas el que está bailando con su sombrero.* (elic01_lydc)
'No, it is Cantinflas who is dancing with his hat.'
- b. *No, es Cantinflas que está fumando un cigarro.* (elic01_lydc)
'No, it is Cantinflas who is smoking a cigarette.'

In the literature, the discussion of determiner-less *que* in clefts has primarily been restricted to clefts with focalized adverbial or causal preposition constituents, labeled as *que galicado* constructions by Cuervo (1907) (cf., e.g., Dufter 2010 and the references cited therein). Structures in the form of (2b) are surprising, given that such determiner-less complementizers in subject or object clefts are either considered ungrammatical (Di Tullio 1990, 2006) or have, to our knowledge, not been attested in any Spanish context so far. While one example of such a structure is cited in Bentivoglio and Sedano (2017, p. 113), the authors state that this single instance (across two corpora composed of literary works and samples of “habla culta”) comes from a work by the writer, Julio Cortázar. They further argue that, in this particular case, this (single) occurrence of bare *que* may indeed be due to French influence, given that Cortázar spent a fair part of his life in France. Against this background, it is even more striking that the number of sentences containing bare *que* accounts for around 50% of all cases in our data (see Section 3.3 for details).

The reasons for such patterns are controversial in the literature, and various explanations have been put forward. In this context, it is opportune to note that cleft sentences of the type [copula + focal element + *que* + predicate] are rarely included in discussions on determiner-less or bare *que* complementizers in relative clauses, for theoretical reasons (Brucart 2016), a fact that is reflected in the differing explanations given for to the occurrence of bare *que* in cleft sentences, compared to the one in non-clefted relative clauses (see Section 2). Nonetheless, previous researchers, such as Borzi (2018), argue that bare *que* in clefts can be considered to be on par with bare *que* in non-clefted relative clauses. In this paper, we will be following this assumption.

Turning back to the possible causes of the usage of determiner-less or bare *que* relative complementizers in Spanish, we already mentioned the hypothesis that the use of bare *que* in Spanish clefts, with focalized adverbs or causal prepositional constituents, may be due to the historic influence of the French language; hence, we see the label *que galicado* (Cuervo 1907), although such claims have been challenged in other studies (Dufter 2010, see Section 2 for details). Other researchers argue that such bare *que* forms should be understood as processes of semantic bleaching or grammaticalization (Alarcos 1963). In contrast, for non-clefted relative clauses, it has been shown in sociolinguistic diachronic research that relative complementizers with the determiner-less prepositional structure (PREPOSITION + *que*) (e.g., *en que*) predate the complex prepositional complementizers of the type (PREPOSITION (PREP) + DETERMINER (DET) + *que*) (e.g., *en el que*). While the form that has a determiner experiences an increase in usage in the 18th century (Vellón Lahoz 2019), the determiner-less variant still outnumbers the alternative form, at least until the mid-19th century (Blas Arroyo et al. 2019). From a broader historical perspective, then, the complex variant *with* a determiner actually constitutes the innovative form. This, in turn, raises the question of what exactly motivates the usage of determiner-less or bare *que* forms in Spanish relative clauses, synchronically: is it best considered as: (i) a historical remnant, or (ii) a dialectal recast, following the diachronic grammaticalization process

that led to the above-mentioned complex forms? If it transpires that (ii) is more likely, what might be the reasons for the innovative reduction to a determiner-less *que*: general cognitive (economy) principles, the context of massive bilingualism/language contact, or other reasons?

Obviously, these questions are complex and do, actually, require broad cross-dialectal comparisons. This paper aims to contribute one (modest) piece of the puzzle to this complex project by means of a pilot study on relative complementizer usage in Yucatecan Spanish, a Spanish dialect spoken on the Yucatán peninsula in southeast Mexico. Thus, as a first approach to the above-mentioned intricate questions, we will analyze the relative complementizers in a corpus containing both elicited and spontaneous speech data, against the background of the current state of research on complementizer variation in Spanish, focusing on a comparison between monolingual and bilingual speakers. For the purpose of obtaining a more comprehensive picture, we will also include clefts in our analysis, which have often been excluded in similar studies. That is to say, this paper primarily addresses the contrast between the use of: (i) complex prepositional complementizers ([PREP + DET + *que*]) and bare *que*, particularly in the locative domain, on the one hand, and (ii) complex complementizers in cleft sentences ([DET + *que*]) and bare *que*, on the other. When necessary, the determiner-less prepositional complementizers ([PREP + *que*]) will also briefly be touched upon.

The outline of the article is as follows: In Section 2, we will delineate the theoretical background and depict the current state of research regarding relative complementizer variations in Spanish. In Section 3, we will first introduce our research question and hypothesis (Section 3.1), then describe our database and the methodology employed, in order to extract and annotate the relevant data (Section 3.2), and, finally, present the results (Section 3.3). In Section 4, we will discuss our findings in the light of previous studies, in order to provide a first approach concerning the question of whether the complementizer patterns observed in our data of Yucatecan Spanish are best considered either as a historical remnant or as a distinct, innovative change. Finally, Section 5 summarizes our main results and presents our conclusions.

2. Previous Accounts of Distributional Facts

As previously touched upon, relative complementizers are found in both non-clefted relative clauses and in clefts². However, in studies on complementizer variation in relative clauses, clefts are rarely included in this discussion for theoretical reasons (Brucart 2016). Borzi (2018), however, suggests that bare *que* in both clefted and non-clefted relative clauses can be treated on par since the selection of bare *que* in both clause types seems to be governed by similar factors. More precisely, an identified antecedent favors the bare *que* in both clefts and non-clefted relative clauses, as identification seems to render gender- and number-marked complementation obsolete (Borzi 2018, p. 33).

In the discussion of relative complementizer variation, then, researchers have pursued different directions to account for determiner-less variants. One line of argument is the hypothesis that the use of bare *que* can be attributed to an influence from the French language; hence, we see the label *que galicado*, as coined by Cuervo (1907). In French cleft sentences that cleave an adverbial constituent, *que* is the standard complementizer that connects the cleft to the main clause:

3 *Ce fut dans le XV siècle que l'Amérique fut découverte.* (Cuervo 1907, §440)
'It was in the 15th century that America was discovered.'

The fact that Spanish clefts, such as that in (4), follow the structure of the French example in (3) does, at first sight, seem to substantiate the assumption that the use of bare *que* is due to an influence from French in this context.

4 *Fue en el siglo XV que se descubrió la América.* (Dufter 2010, p. 254)
'It was in the 15th century that America was discovered.'

However, the hypothesis that the Spanish bare *que* traces back to a French influence has been challenged, for example, by [Dufter \(2010\)](#), who investigates the use of *que galicado* in cleft sentences with adverbial antecedents. More specifically, [Dufter \(2010\)](#) suggests that such occurrences constitute a pan-Romance phenomenon, rather than a phenomenon in Spanish that was induced by a French influence. In order to test this hypothesis, he conducted a diachronic study, based on the FRANTEXT and CORDE corpora, to assess the development of bare *que* usage from the 16th to the 20th century, in both French and Spanish. Additionally, he pursued a synchronic study, based on the CREA corpus, to map out in which Spanish-speaking regions bare *que* finds the highest usage. [Dufter \(2010, p. 266\)](#) found that bare *que* is used more often in Latin American Spanish than in peninsular Spanish (5.50 vs. 1.34 occurrences per 100,000 words) and that it seems to be most popular among speakers from Venezuela (9.43/100,000 words). Additionally, the results show that in the 16th century, there are more occurrences of bare *que* in Spanish than in French (1.13 vs. 0.29 occurrences per 100,000 words). According to [Dufter \(2010\)](#), this contradicts the notion that bare *que* in Spanish was adopted from the French language, as one would expect French to exhibit a higher degree of occurrences than Spanish in the latter case. In the 17th century, in turn, the data show a turnaround in the frequency rates. In the Spanish corpus, the frequency decreases to a rate of 0.80 per 100,000 words compared to 3.84 in French. According to [Dufter \(2010, pp. 268–69\)](#), this reversal coincides with both the increasingly stereotypical use of adverbial clefts³ in French and the beginning of the refusal of the *que galicado* in Spanish-speaking regions. In French, the use of adverbial clefts increases again in the 18th century (6.33/100,000 words), followed by a drop to 3.14 and 4.55 in the 19th and 20th centuries, respectively. In Spanish, a different developmental pattern can be observed. By the 18th century, there is a slight increase in bare *que* usage (0.94/100,000 words), followed, again, by higher frequencies in the 19th and 20th centuries (2.96 and 1.69/100,000 words, respectively). [Dufter \(2010, p. 267\)](#) attributes the higher number of bare *que* in the 19th and 20th centuries to the higher proportion of Latin American texts in the respective sections of the corpus. Based on these findings, [Dufter \(2010\)](#) concludes that, except for some specific instances, determiner-less or bare *que* usage in adverbial clefts is not caused by a French influence.

Next to this, we note: (i) that determiner-less complementizers in relative complementation, other than cleft sentences, are widespread in medieval Spanish, with the combined forms (DET + *que*) coming into being via reanalysis during the 13th to the 15th centuries only ([Mackenzie 2019, pp. 185–91](#)), and (ii) that determiner-less complementizers in relative sentences other than clefts have also been attested far beyond that date in diachronic research. For example, [Vellón Lahoz \(2019\)](#) investigates the diachronic development of the use of determiner-less prepositional relative complementizers with the preposition *con* (e.g., *con (el) que*) throughout the 18th century. The database is a corpus of ego-documents, i.e., texts of communicative immediacy, such as private letters from Spain to America and vice versa. [Vellón Lahoz \(2019\)](#) takes into account a whole range of possible influencing factors related to the antecedent and the complementizer, as well as to the clause itself, e.g., the presence of a determiner in the antecedent or the meaning or type of the relative clause, respectively. The results show that, at the beginning of the 18th century, the use of complex prepositional complementizers, i.e., *con el que*, was rather low compared to the use of the determiner-less variant, *con que*, with a proportion of 8.5% for cases of *con el que* in the second decade. However, in the final decade, the use of this variant reached a proportion of 21% in the data. This suggests that the determiner-less form actually predates the use of the complex prepositional complementizer with a determiner, suggesting that, diachronically, the latter (*con el que*) constitutes the innovative form, whereas the former (*con que*) seems to be a diachronic remnant of medieval/former stages of Spanish.

Studies on the use of determiner-less complementizers in the subsequent centuries seem to confirm this picture. [Blas Arroyo et al. \(2019\)](#) investigate the frequency of determiner omission in terms of relative complementizers from the 18th to the 20th centuries. They are concerned with the issue of whether the change from determiner-less (*en que*) to

ti', i.e., [REL + PREP]. In contrast, the in situ word order of the corresponding prepositional phrase is *ti' máax*.

Finally, another feature that sets the Yucatec Maya language apart from Spanish is the particular character of the relative pronoun *tu'ux*, 'where'. Consider, e.g., (6):

6	<i>Le tunich</i>	<i>[_{RC} tu'ux</i>	<i>k-u</i>	<i>pak'-a'-l</i>	<i>le graasia]-o'</i>
	the stones	where	AUX-ERG.3	sow-PASS-IND	the stuff
	'The stones between which the stuff (corn) is sowed.' (Adapted from Gutiérrez Bravo 2012, p. 256)				

The complementizer *tu'ux* in (6) expresses local information that in Spanish would translate to *entre las que* 'between which'. However, the complementizer is not modified by any preposition that expresses the semantic content of the preposition *entre*. In fact, the adverbs *tu'un* (*i*) and *tu'ux* (*i*) have various translations—among others, *dónde*, *en qué lugar*, and *por dónde* (Martínez Huchim 2014). In Spanish, relative pronouns need to be preceded by a preposition if they are semantically or grammatically required (Brucart 1999, p. 399), a feature that is lacking in relative complementizers in Yucatec Maya. Thus, it seems that the local relative adverb *tu'ux* has a reduced expression of spatial or directional information, compared to the prepositional complementation in Spanish. Coming back to the issue of relative complementizer variations in Yucatecan Spanish, this phenomenon could, thus, possibly also be influenced by the less structurally complex complementizer system of Yucatec Maya. All in all, the findings considered thus far seem to point in two directions: the determiner-less complementizers in Yucatecan Spanish might either be a historical remnant from Medieval Spanish or are the result of an ongoing process of simplification, possibly driven by other factors, one of which could be the language contact with Yucatec Maya.

In order to approach this question in more detail, we conducted a corpus study that complements previous studies by contributing data from Yucatecan Spanish. Given that the studies described above cover information from the 18th to the 20th centuries, more recent data may contribute to the current body of knowledge on complementizer variation. Additionally, we not only consider relative clauses but also subject and object clefts, to which, according to the literature, bare *que* usage has not yet been attested to any significant degree (see Section 1) in any current variety of Spanish so far.

3. Empirical Analysis

In what follows, we will present the data collected in Quintana Roo, Mexico during several fieldwork visits from 2017 to 2019. Section 3.1 outlines our research question and working hypothesis, to be verified by empirical study; Section 3.2 introduces our database and the methodology of data extraction and annotation; Section 3.3. presents the results of the study.

3.1. Research Question and Working Hypothesis

As stated in Section 2, the overall research question in the background of our pilot study is whether the determiner-less complementizers in a given Spanish variety are best considered as a historical remnant or a dialectal recast. Moreover, in a region of close and long-lasting language contact, such as in Quintana Roo, it is reasonable to consider cross-linguistic influence as being one of the reasons for the observed patterns. That being said, we are obviously fully aware of the fact that there is no direct correspondence between bilingualism (i.e., mastery of the contact language) of the speakers and language contact as a source of linguistic transfer, since contact-induced features may just as easily be characteristic of numerous varieties of the *monolingual* speaker groups. This is what, e.g., Uth and Martínez García (2020) found regarding falling pitch accents in monolingual vs. bilingual Quintana Roo Spanish. On the other hand, it is just as reasonable to suggest, as a working hypothesis for a given feature in a given contact region, that the bilingual speakers might show a stronger influence of the contact language (in our case, Yucatec Maya) in

the language under investigation (in our case, Yucatecan Spanish), as they are generally much more exposed to the contact language than the monolingual speakers. Therefore, although knowing that monolingual speakers of Quintana Roo Spanish exhibit and develop some contact features in a more pronounced way than bilingual speakers, we nevertheless pursue, here, the *working hypothesis* of what, in our view, is a more fundamental assumption: that, if at all, it is the bilingual speakers or data, respectively, who/that result in being more influenced by the contact language as regards the usage of relative complementizers. The above-stated insights with respect to both the diachrony of relative complementizers in (European) Spanish and the morphosyntax of Yucatecan Spanish and Yucatec Maya in the realm of relative complementation thereby lead us to state the following research question (RQ) and working hypothesis (RH) for the purposes of our empirical study.

RQ: Do the speakers of our Quintana Roo Spanish data set exhibit a higher degree of bare *que* usage than what has been found for other varieties? Do the bilingual and monolingual speakers of Yucatecan Spanish in our database differ, with respect to the patterns of relative complementizer usage?

RH: We hypothesize: (i) that we will indeed find a higher degree of bare *que* usage than what has been found for other varieties, and (ii) that the bilingual and the monolingual data do indeed exhibit differences with respect to relative complementation (+/– complex).

With this research question and working hypothesis in mind, we will now turn to the methodology and the results of the empirical study.

3.2. Databases and Methodology

The present study draws on two different databases. In order to verify the research question, we rely upon both a naturalistic and an elicited dataset, which are described one after the other in the following section.

3.2.1. Naturalistic Data

The naturalistic dataset consists of sociolinguistic interviews conducted in Felipe Carrillo Puerto, Mexico, in the context of a research project on language contact in Mexico. In total, the interviews were conducted with 41 language speakers, with an average age of 41 years. Among those speakers, there were 21 speakers who were balanced bilinguals, as well as 20 Spanish-dominant and monolingual speakers. After scanning all interviews, we found that only 24 interviews contained relative clauses, reducing the number of interviews considered from 41 to 24—the remaining interviews have come from 11 bilingual and 13 Spanish-dominant and monolingual speakers. The conversations revolved around cultural and language-related topics, such as food, music, and attitudes toward languages and dialects, and were of an informal character. There was no preconceived interview template in order to achieve a more natural communicative setting. After repeated screenings of the interviews, we ended up with a total number of 202 tokens, which comprise 141 relative clauses and 61 clefts.

For the annotation of the relative clauses, both speaker-related and language-related variables are considered. The former include the variables of *language profile* (with the levels, ‘YS-monolingual/-dominant’ and ‘balanced bilingual’, *gender* (‘male’, ‘female’), and *age*. The language-related variables include the *relative pronoun* (‘que’, ‘cual’, ‘donde’), the *usage domain* of the relative pronoun (‘locative’, ‘temporal’, ‘manner’, ‘partitive’, ‘possessive’, ‘other’), *preposition* (‘none’, ‘a’, ‘con’, ‘de’, ‘en’), and *determiner* (‘none’, ‘el’, ‘la’, ‘los’, ‘las’). As far as the subject and object clefts are concerned, similar criteria were applied. For the speaker-related variables, we annotated *language profile*, *gender*, and *age*. For the language-related variables, we annotated the *relative pronoun* (‘que’, ‘cual’, ‘quien’), *preposition* (‘none’, ‘a’, ‘con’, ‘de’, ‘en’) and *determiner* (‘none’, ‘el’, ‘la’, ‘los’, ‘las’). These variables were selected in order to determine whether they have an impact on the choice of relative complementizers. Examples of the different complementizer configurations that occur in our data, i.e., complex and determiner-less prepositional complementizers, as well as bare *que*, can be found in 7a, 7b, and 7c, respectively.

- 7 a. *Hacen un hueco prenden leña en el que le tiran las piedras.* (elic12_lqr)
They make a hole and set firewood into which they put stones.
- b. *Sí, de hecho no son como en otras ciudades de que puedes llegar.* (elic09_mgv)
Yes, in fact, they are not like other cities that you can reach.
- c. *Son lagunas que puedes llegar en bicicl/, en bici, en bicicleta.* (elic10_cvt)
They are lagoons that you can get to by bike.

3.2.2. Elicitation Data

The second data set comes from a picture-based elicitation experiment on contrastive-corrective focus, which was conducted with 25 speakers, with an average age of 30.5 years. Here, 5 of the participants were balanced bilinguals, and 20 were YS-monolingual or -dominant. In the experiment, the participants were presented with a picture and were subsequently given a statement that the participants were asked to correct, based on the information included in the picture (Figure 1).



Don Gato y su pandilla

¿Aquí, el Chapulín Colorado juega golf, verdad?

Figure 1. Stimulus from the elicitation experiment.

The picture, which is shown to the participants, depicts the character ‘Don Gato’ playing golf. Subsequently, the participants are asked the question: ‘Here, Chapulín Colorado is playing golf, right?’. Given that the content of the picture contrasts with the information given in the question, the participants were expected to give contrastive-corrective statements, such as ‘No, it is Don Gato who is playing golf’. The images used in the experiment were of characters from well-known Mexican TV series, with which the participants were also familiarized prior to the main portion of the experiment. Overall, we obtained a total number of 67 subject and object cleft sentences. For the annotation, we considered the same variables as for the clefts in the naturalistic database.

3.3. Results

An overview of the results for relative complementizers in terms of subject and object clefts in elicitation and interview data is depicted in Figures 2 and 3, respectively.

As far as the elicitation data are concerned, the analysis shows that the complementizer *quien* is not used by bilingual speakers at all, compared to a proportion of 27.5% in the monolingual/YS-dominant data. The complex complementizers, in turn, are the most commonly used complementizer in the bilingual data, with a percentage of 70%, compared to a proportion of 45% in the monolingual/YS-dominant data. While the proportion of complex complementizers is lower in this group than in the bilingual group, it is still the preferred option in the context of the monolingual data. Bare *que*, on the other hand, occurs at a proportion of 30% in the bilingual data and 27.5% in the monolingual data, and, thus, shows a balanced distribution between the two groups (Figure 2).

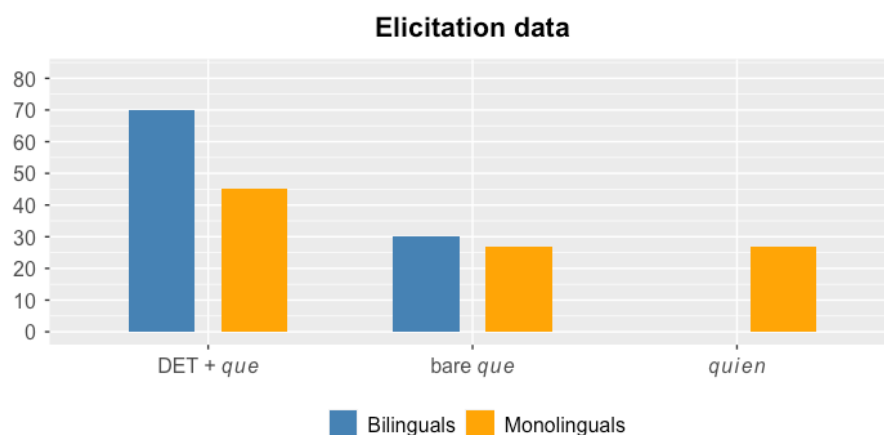


Figure 2. Results for the elicitation cleft data (as a percentage).

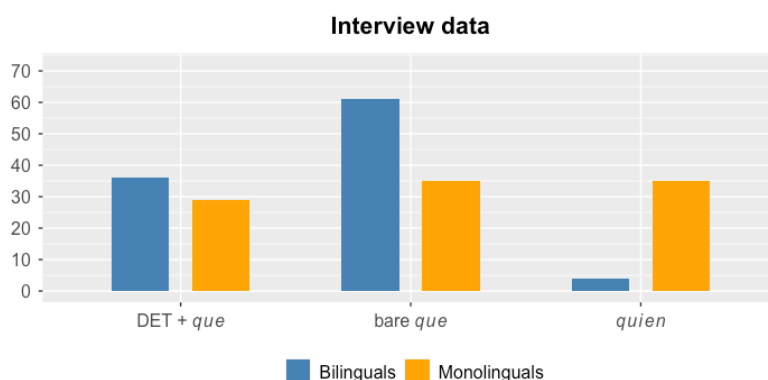


Figure 3. Results for the interview cleft data (as a percentage).

Turning now to the clefts in the interview data, the analysis, again, shows that *quien* is the least preferred option by bilingual speakers (4%) whereas the monolingual/YS-dominant groups exhibit a much higher proportion of *quien* (35.5%). With regard to the complex complementizers ([DET+*que*]), the elicitation data shows a proportion that is higher in bilinguals (36%) than in monolinguals (29%). In comparison to the elicitation data, this proportion is overall lower, but more balanced across the two groups. The proportion of bare *que* is, however, much more increased in the interview data overall, with a proportion of 61% in the bilingual data and 35.5% in the monolingual/YS-dominant data. Overall, there is a much higher proportion of bare *que* in the interview data than in the elicitation data, and it is also less balanced, with the bilinguals showing a much higher proportion than the monolingual/YS-dominant speakers. This may be taken as an effect of cognitive load or economy, respectively, which might have been higher in the interview data, an issue that we will briefly address in the discussion section. In view of the fact that bare *que* usage in clefts is (a) not yet attested in general, and (b) is considered ungrammatical in reference grammars, the number of times it occurred is surprisingly high in both data types and across the groups (Figure 3).

Yet another remarkable effect can be observed in the relative clauses of the naturalistic data set (Figure 3). With regard to the number of determiner-less prepositional complementizers (PREP + *que*), as well as the number of bare *que*, the numbers were quite even between both groups (12 vs. 11 occurrences and 13 vs. 15 occurrences, respectively).

While these results, at first glance, do not reveal any noteworthy patterns, it is interesting to note that, in contrast, the results for two other relative complementizers do show remarkable discrepancies between the two groups. First, with regard to complex prepositional complementizers (PREP + DET + *que*), there were only five occurrences in the bilingual group and eighteen in the monolingual group in total (Figure 4). The production

rate of this complex configuration in the monolingual group is, thus, more than three times as high as it is in the bilingual group.

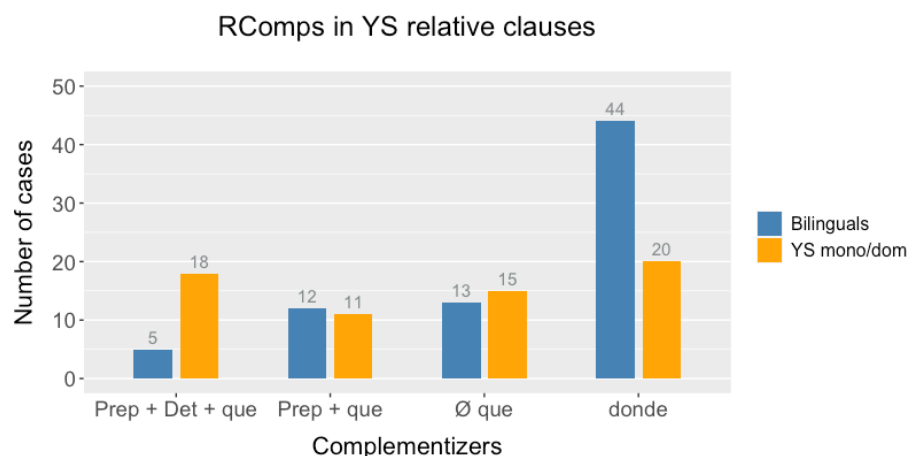


Figure 4. Results for relative complementizers in the relative clause data.

The second difference concerns *donde*, ‘where’, for which the data show that the bilingual group produced 44 cases, whereas the YS-monolingual/dominant speakers only used it 20 times, meaning that the bilinguals produced over twice as many cases of *donde* as the monolinguals (Figure 4). The inverse results of these two relative complementizers, in conjunction with the broad usage of *tu’ux* in Yucatec Maya, suggest that the complementizer *donde* may have a siphoning effect in only the bilingual group; i.e., that the bilingual speakers, but not the monolinguals, gravitate toward *donde* when expressing locative relative clauses, rather than using complex prepositional complementizers.

A subsequent analysis indeed seems to corroborate this notion, as the distributions among the different usage domains show that the bilingual speakers do not seem to use locative complex prepositional complementizers at all (Figure 5).

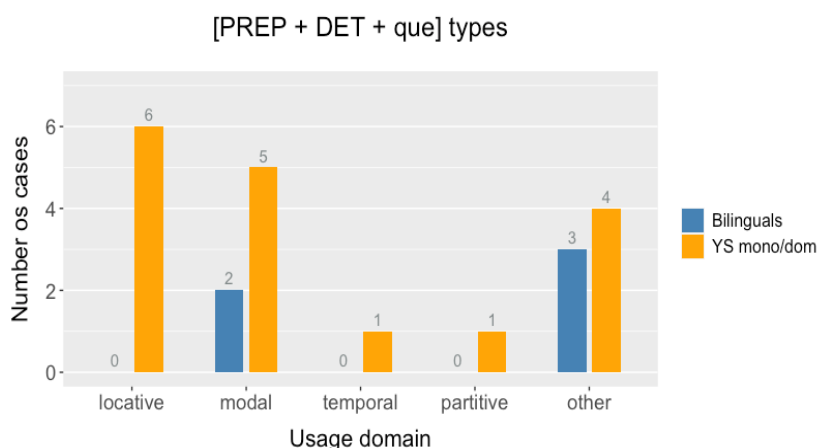


Figure 5. Results for complex complementizers in the relative clause data, according to the usage domains.

It is true that the occurrences of temporal and partitive types are too infrequent to draw any firm conclusions. Nevertheless, note that the bilingual group does produce complex prepositional constructions in other usage domains, which makes the complete absence of such structures in the locative domain rather striking. The varied distribution of this particular relative complementizer prompted us to further analyze the usage domains of bare *que*, despite the similar number of occurrences in both groups (13 vs. 15 in the bilingual and YS-dominant/monolingual group, respectively) and, again, the same pattern

in the locative domain can be observed: the bilingual speakers do not seem to use the locative bare *que* at all, whereas the locative bare *que* is one of the most frequent usages among the YS-dominant/monolingual speakers (Figure 6).

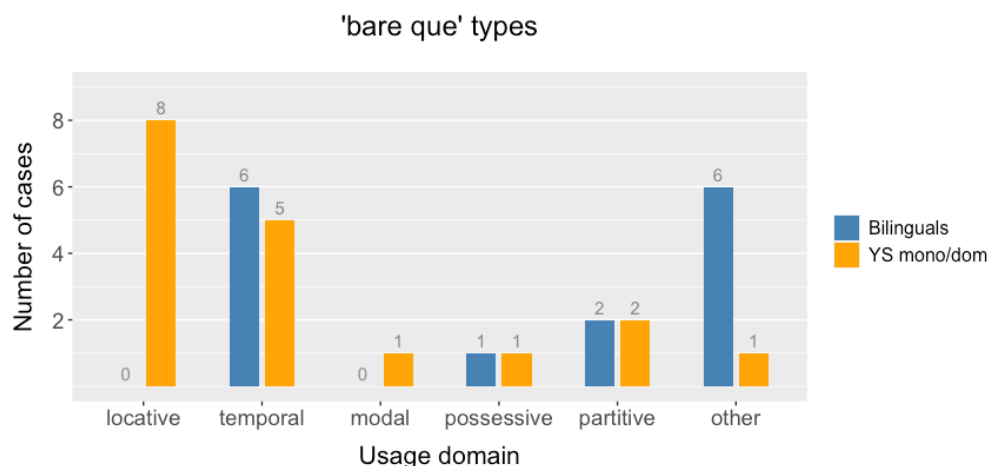


Figure 6. Results for bare *que* in the relative clause data, according to the usage domains.

In more concrete terms, the numbers displayed in Figure 6 reveal two interesting results. First, the bilingual group does not produce any single instance of locative bare *que*, as opposed to the monolingual group, with a total of eight occurrences. Second, in all other usage domains, the groups produce similar numbers, highlighting the particularity of the locative relative complementizers in the bilingual speakers. All in all, the usage domains of bare *que* seem to support the notion of a gravitational effect of *donde* in the bilingual group, suggesting, again, the possible influence of *tu'ux* from Yucatec Maya.

Turning now to the usage domains of determiner-less prepositional complementizers (PREP + *que*), it is, first of all, noteworthy that all occurrences of these constructions included the prototypically locative preposition *en*, 'in', although it appears in different usage domains (locative, manner, temporal, and other). For all four usage domains of determiner-less prepositional complementizers, the two groups show very similar numbers (Figure 7).

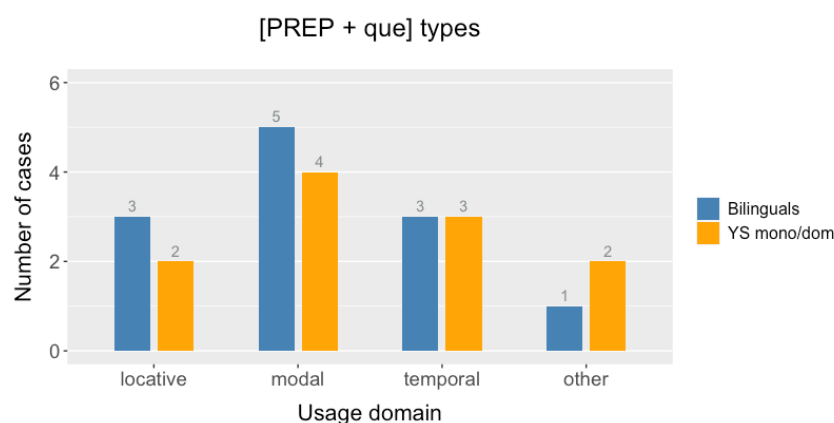


Figure 7. Results for determiner-less complementizers in the relative clause data.

This pattern deviates from that observed in the other two relative complementizer strategies, complex prepositional complementizers and bare *que*, in that the two groups do not differ in the locative usage domain. In the bilingual data, we found three occurrences of determiner-less prepositional complementizers and four in the YS-dominant/monolingual data. Similarly, we found almost even numbers in the remaining usage domains (5 vs. 4 occurrences), temporal (3 vs. 3 occurrences), and other (1 vs. 2 occurrences) usages in the

bilingual and YS-dominant/monolingual data, respectively. That being said, the results for the locative domain should be taken with a pinch of salt, due to the low number of tokens.

To sum up, the YS-dominant/monolingual and the bilingual speaker groups overlap, in some respects, with regard to relative complementizer choice but they also exhibit particular differences. With respect to clefts, both groups show a similar degree of bare *que* (rather than the standard option of DET + *que*). In terms of relative clauses, the results are multi-layered. Looking at the envelope of variation, the bilingual and YS-dominant/monolingual data show an even distribution of determiner-less prepositional complementizers and bare *que*. However, the groups differ in their usage of complex prepositional complementizers and *donde*. The monolingual data show a much higher number of complex prepositional structures compared to the bilingual group, whereas there are many more cases of *donde* in the bilingual data than in the YS-dominant/monolingual data. On the level of usage domains of the relative complementizers, further differences arise (pertaining to locative relative complementizers). Unlike the YS-dominant/monolingual group, the bilingual speakers do not use the locative bare *que* or complex prepositional complementizers at all; however, this is a difference that does not arise in the scattered determiner-less prepositional structures.

4. Discussion

Taken altogether, our analyses of the databases have revealed a high degree of complementizer variation. With regard to the first research question, i.e., if our Quintana Roo Spanish data set exhibits a particularly high degree of bare *que* usage, compared to what has been identified for other varieties, the data supports our affirmative hypothesis. We found a substantial number of bare *que* in subject and object clefts, which have not so far been attested to (at least, to this degree) in other varieties. Additionally, in non-clefted relative clauses, we found that a considerable number of bare *que* co-occurs next to other relative complementizers. As concerns the question of whether the bilingual and monolingual speakers of Yucatecan Spanish in our database differ with respect to the patterns of relative complementizer usage, the data equally support our working hypothesis, in that the bilingual and the monolingual data do indeed exhibit differences with respect to the usage of +/– complex complementizers. As illustrated by Figures 5 and 6 in Section 3.3, the differences between the groups mainly arise in the domain of locative complementation, in that the YS-monolinguals and YS-dominant speakers produce the entire envelope of variation, whereas the bilingual speakers clearly prefer *donde* for locative relative complementation, at the cost of the locative bare *que* and locative complex prepositional complementizers (PREP + DET + *que*). This finding corroborates the assumption of a dialectal recast, the causes of which remain to be investigated. The observation that these patterns are limited to locative complementation raises the possibility of the influence of the Yucatec Maya complementizer *tu'ux* on complementizer variations in bilingual speakers. A notable exception to this pattern was the locative determiner-less prepositional complementizers (PREP + *que*), which were also produced by bilingual speakers, although the number of cases may have been too low to draw firm conclusions.

Nonetheless, it should also be noted that a fair degree of determiner-less *que* and determiner-less prepositional complementizers, outside of the locative usage domain, was shared by both groups. This holds true not only for non-clefted relative clauses but also for clefts. For example, occurrences of bare *que* can be observed in non-clefted relative clause data, with similar rates in both groups. Against the background of previous diachronic studies on relative complementation, the relative clause data are also quite intriguing, in that they, too, show the presence of both the determiner-less prepositional complementizer and diachronically innovative complex prepositional construction. The balanced number may be understood as a continuation of the trend that was observed by Blas Arroyo et al. (2019) and Vellón Lahoz (2019), namely, the increasing use of the innovative complex variant. However, this omits the role of *tu'ux*, the influence of which may well have led to the previously described notable decrease in the use of locative complex prepositional com-

plementizers in the bilingual group. Hence, our results on prepositional complementation with and without a determiner cannot be simply understood as a mere continuation of the development documented in previous studies. Rather, they point to a complex variational phenomenon unifying at least two types of triggers/motives or developments, respectively.

Another pattern shared by both groups, i.e., the usage of bare *que* in subject and object clefts, seems to constitute a region-specific phenomenon as the particular pattern that emerged in our data has, so far, not generally been documented in the literature (see, again, Section 1 for details). This phenomenon is striking in that it occurs in both the bilingual and the YS-monolingual/dominant group. The causes of this pattern are unclear at this point and need to be explored in future research. Another interesting effect that we observed, coincidentally, was that the numbers of bare *que* seem to have been modulated by the task type: In the spontaneous speech data, the proportion of bare *que* usage was higher overall than it was in the experimental data. While this may not be a surprising finding, in itself, we still consider it to be important, as it raises the possibility of cognitive economy in the context of non- or less formal speech as a factor in the variation of relative complementizer usage in Spanish, which should be considered in future research.

Summarizing our central findings, we found a two-way variational development: (i) the variational pattern exclusive to the bilingual group, i.e., the increased use of *donde*, and the accompanying lowered production rate of locative complex prepositional complementizers and bare *que*, which could eventually be attributed to the influence of the Mayan *tu'ux*, and (ii) the patterns of complementizer usage shared by both groups, which can be observed in the presence of bare *que* and determiner-less prepositional structures (PREP + *que*). The causes of the latter are also difficult to pinpoint from our data sample, as the structural variation could be either due to diachronic archaism or could be a case of dialectal recast. We will briefly return to this issue in Section 5.

Overall, the analysis has revealed a complex interplay of different strands of complementizer variation. On the one hand, we encountered complementizer preferences that are exclusive to the bilingual speakers, while on the other hand, we found dialect- or region-specific patterns in the form of bare *que* in both non-clefted relatives and clefts. Additionally, we also detected prepositional complementizers that are both with and without a determiner, the numbers of which are congruent with the pan-Hispanic developments worked out in previous studies, suggesting that the presence of the determiner-less variant may indeed constitute a historic remnant in these particular cases. These layers of variation do not exist in isolation but, instead, most probably interact with each other. For example, the region-specific usage of bare *que* in relative clauses could be influenced by the Mayan *tu'ux* in the bilingual group, to the extent that the bilingual speakers seem to gravitate toward *donde* at the cost of the locative bare *que*. This would suggest that *donde* and the locative bare *que* may be in competition with each other in the bilingual speakers. Similarly, *donde* also seems to have a siphoning effect on the historically innovative form (PREP + DET + *que*) in the bilinguals. As far as the question of grammaticalization is concerned, the presence of both the determiner-less as well as the complex prepositional complementizers suggests that the grammaticalization of (DET + *que*) is still not concluded. Additionally, these particular characteristics, in the form of bare *que* and the increased usage of *donde* in the bilingual group, may ultimately stall or alter the grammaticalization process.

Throughout the analysis, it has become apparent that disentangling the individual developments is a complex task, as both system-internal tendencies and externally induced changes seem to interact with each other. In view of the different tendencies observed between the groups, the question arises as to what exactly causes them. We would like to tentatively address the question of language contact, as this would be one candidate (among others) by which to explain at least some of these patterns, given the long-lasting presence of, and close interchange between, the two languages in one and the same region. Certain patterns in our data, at first glance, do indeed seem to point in this direction. For example, the increased use of *donde* at the cost of other complementizer structures, as well as its usage patterns in the bilingual group, suggest the effect of *tu'ux*. The (supposed) directionality of

influence in this particular example lends itself to analysis along the lines followed by [van Coetsem \(1988, 2000\)](#), [Winford \(2005\)](#), or similar theoretical approaches ([Palacios Alcaine 2007](#); [Palacios 2011](#)). In his framework, [van Coetsem \(1988, 2000\)](#) distinguishes between two types of transfer, which he labels *borrowing* and *imposition*. The difference between these two terms can be explained by referencing: (1) the direction in which the transfer takes place (from the source language to the recipient language) and (2) the agency of the speaker. Thus, borrowing is defined as a transfer where the agent is a speaker of the recipient language, whereas imposition is understood to be a transfer wherein the agent is a speaker of the source language ([van Coetsem 1988](#), p. 3). This approach has been expanded upon by [Winford \(2005\)](#), who crucially suggests that in a bilingual setting, one and the same speaker can be the agent of either type of transfer and can, thus, account for a variety of language-contact phenomena in a wide range of multilingual social configurations/settings. Coming back to the example of the increased use of *donde*, it could be argued that this constitutes a case of imposition, i.e., a case of source language agentivity, given that it is the bilingual speaker group that seemingly transfers the properties of the Yucatec Maya complementizer *tu'ux* onto the Spanish relativizer system. This analysis is far from in-depth, but more importantly, it needs to be ascertained as to whether this is indeed a case of cross-linguistic transfer at all. While our data, despite their limited sample size, reveal some notable tendencies that tend to suggest the possibility of linguistic transfer, they cannot, by themselves, answer the question if the patterns they exhibit are indeed due to language contact. Future research should thus focus on the comparison of complementizer preferences in other varieties of Spanish to confirm or rule out an impact of language contact in the case of our Quintana Roo data (and in the case of other varieties). Additionally, given the limited size of the corpus, the findings of this study should be verified by a larger study with a larger database.

5. Conclusions

This article set out to approach the issue of the high degree of relative complementizer variation that we have found in our elicited and naturalistic speech data analysis of Yucatecan Spanish. Previous diachronic research has shown that determiner-less complementizers actually predate relative complementizers with a determiner, which, thus, constitute an innovative variant ([Blas Arroyo et al. 2019](#); [Mackenzie 2019](#); [Vellón Lahoz 2019](#)). A further important aspect of our research is that the Spanish variety investigated in this article, Yucatecan Spanish, has been in long-standing contact with Yucatec Maya. Relative complementation in Yucatec Maya differs from that in Spanish, in that the complementizer *tu'ux* or 'donde' has a much broader usage than *donde* in standard Spanish. Against this background, we investigated data from elicitation experiments and sociolinguistic interviews to approach the question of whether the determiner-less complementizers in our data should be considered to be historic remnants or rather as a case of dialectal recast, caused by entirely different factors. Our analysis suggests a complex picture: (i) usage of structurally less complex/determiner-less complementizers (PREP + *que*; bare *que*) that can be observed across both monolingual speakers of Spanish and Spanish Maya bilinguals, and (ii) a preference for *donde* at the cost of locative bare *que* and (PREP + DET + *que*) in the bilingual group only. As regards the first finding: (a), the balanced number of complex and non-complex complementizers in our data may be understood as a continuation of the trend that was observed by [Blas Arroyo et al. \(2019\)](#) and [Vellón Lahoz \(2019\)](#), namely, the increasing use of the innovative complex variant, suggesting that we are faced with a historical remnant in this case. The second finding, (b), corroborates the assumption of a dialectal recast, the causes of which remain to be investigated. However, the observation that these patterns are limited to locative complementation raises the possibility of the influence of the Yucatec Maya complementizer *tu'ux* on complementizer variations in bilingual speakers.

Our data not only complement previous findings on relative complementizer variation in Spanish but may also give new motivation for future research, given the broad envelope

of variation in the usage investigated in this article. For example, it is noteworthy that the paradigm of relative complementizers in YS shows a *donde*-effect in the bilingual group, but not among the monolingual speakers, which could open up a window for a more fine-grained investigation of the reasons underlying this particular pattern, one of them possibly being language contact. Additionally, given the low numbers of (PREP + *que*) in our data, future research focusing on this particular configuration may provide new evidence: (i) whether (PREP + *que*) is indeed less susceptible to the influence of *tu'ux*, and, if so, (ii) due to what factors (e.g., processing-related mechanisms). Another complementizer that may benefit from further investigation is *donde*, which, in the bilingual data, seems to compete with both simple and complex locative prepositional complementizers. Finally, due to the limited scope, this article could not address the questions of intra-speaker variation and comparability to other dialects of Spanish. Future research should also focus on providing additional (comparative and experimental) data to complement the tendencies laid bare in this study. Nevertheless, our analysis revealed intriguing differences between the complementizer preferences of YS-monolingual/YS-dominant speakers, on the one hand, and bilingual speakers, on the other, suggesting that different variational patterns caused by different (socio-)linguistic factors can co-develop in parallel in one and the same region. The compelling findings of this early outline of complementizer variation in YS may hopefully serve to inspire more detailed cross-dialectal research along the lines proposed above.

Author Contributions: Conceptualization, P.A. and M.U.; methodology, P.A. and M.U.; investigation, P.A. and M.U.; resources, P.A. and M.U.; data curation, P.A.; writing—original draft preparation, P.A. and M.U.; writing—review and editing, P.A. and M.U.; visualization, P.A.; supervision, M.U.; project administration, M.U.; funding acquisition, M.U. All authors have read and agreed to the published version of the manuscript.

Funding: Funded by the *Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)*—Projekt nummer 491466077. This research is partly funded by the *Deutsche Forschungsgemeinschaft (DFG)* as part of the CRC 1287 “Limits of Variability in Language”, project C09 “Limits of variability in Spanish relative complementation”, directed by Melanie Uth.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors are most grateful to the speakers, without whom this project would not have been possible, for their willingness to participate in this study. We also wish to thank Iñaki Cano (for the support with the visualization) and Daniel Wichmann (for the support with the annotation of the data). Finally, we thank the two anonymous reviewers of the present article for their highly valuable comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

¹ To ensure clarity, we make use of the following terminology: (1) bare *que* (occurrence of *que* where standard grammars would require a complementizer with a determiner or a preposition, i.e., (*en el que*)); (2) determiner-less prepositional complementizers (e.g., *en que*); (3) complex (prepositional) complementizers (e.g., (*con la que*)).

² A cleft sentence is a syntactic configuration in which a constituent is connected to a sentence (from which it was previously detached) by the copula *to be*, following the schema [*to be* + constituent + relative], for example.

³ Adverbial clefts highlight adverbial constituents such as *aquí*—‘here’ in *es aquí donde* (‘it is here, where’), *entonces* in *era entonces que* (‘it was in those days that’), etc. (Dufter 2010, p. 265).

References

- Alarcos, Emilio. 1963. Español “que”. In *Archivum: Revista de la Facultad de Filosofía y Letras*. Oviedo: Universidad de Oviedo, vol. 13, pp. 5–17.
- Andrade, Manuel J. 1955. *A Grammar of Modern Yucatec*. Microfilm Collection of Manuscripts on Middle American Cultural Anthropology Series 7. Chicago: University of Chicago Library, vol. 41. Available online: <https://www.christianlehmann.eu/ling/sprachen/maya/andrade/index.html> (accessed on 27 May 2022).
- Bentivoglio, Paola, and Mercedes Sedano. 2017. El uso del *que galicado* en el español actual. *Boletín de Lingüística* 29: 112–24.
- Blas Arroyo, José Luis, Margarita Porcar Miralles, Mónica Velando Casanova, and Javier Vellón Lahoz, eds. 2019. *Sociolingüística Histórica del Español: Tras las Huellas de la Variación y el Cambio Lingüístico a Través de Textos de Inmediatez Comunicativa*. Lengua y Sociedad en el Mundo Hispánico 41. Madrid and Frankfurt am Main: Iberoamericana, Vervuert.
- Borzi, Claudia Beatriz. 2018. La conceptualización del antecedente como explicación que unifica cuatro fenómenos marginados por los gramáticos. *Cuadernos de la ALFAL* 10: 22–44.
- Brucart, Josep Maria. 1999. La estructura del sintagma nominal: Las oraciones de relativo. In *Gramática Descriptiva de la Lengua Española*. Edited by Ignacio Bosque and Violeta Demonte. Sintaxis Básica de las Clases de Palabras. Madrid: Espasa Calpe, vol. 1, pp. 395–522.
- Brucart, Josep Maria. 2016. Oraciones de Relativo. In *Enciclopedia de Lingüística Hispánica*, 1st ed. Edited by Javier Gutiérrez-Rexach. London: Routledge.
- Butt, John, and Carmen Benjamin. 2000. *A New Reference Grammar of Modern Spanish*, 3rd ed. Chicago: NTC Publishing.
- Cotton, Eleanor Greet, and John M. Sharp. 1988. *Spanish in the Americas*. Washington, DC: Georgetown University Press.
- Cuervo, Rufino José. 1907. *Apuntaciones críticas sobre el Lenguaje Bogotano*, 5th ed. Muy Aumentada y en su Mayor Parte Completamente Refundida. Paris: Ruger & Chernoviz.
- Di Tullio, Angela. 1990. Sobre hendidas y pseudohendidas. *Revista de Lengua y Literatura* 4: 3–16.
- Di Tullio, Ángela. 2006. Clefting in Spoken Discourse. In *Encyclopedia of Language & Linguistics*, 2nd ed. Edited by Keith Brown. Oxford: Elsevier, pp. 483–91.
- Dufter, Andreas. 2010. El *que galicado*: Distribución y descripción gramatical. In *La excepción en la Gramática Española*. Edited by Carsten Sinner and Alfonso Zamorano Aguilar. Madrid: Iberoamericana Vervuert, pp. 255–80.
- Gutiérrez Bravo, Rodrigo. 2012. Relative clauses in Yucatec Maya: Light heads vs. Null domain. In *Relative Clauses in Languages of the Americas: A Typological Overview*. Edited by Bernard Comrie and Zarina Estrada-Fernández. Typological Studies in Language. Amsterdam: John Benjamins, vol. 102, pp. 253–68.
- Henríquez Ureña, Pedro. 1921. Observaciones sobre el español de América. *Revista de Filología Española* 8: 357–90.
- Mackenzie, Ian E. 2019. *Language Structure, Variation and Change: The Case of Old Spanish Syntax*. Cham: Springer International Publishing.
- Martínez Huchim, Ana Patricia. 2014. *Diccionario Maya: Español-Maya, Maya-Español*, 2nd ed. Mérida: Dante.
- Palacios Alcaine, Azucena. 2007. ¿Son compatibles los cambios inducidos por contacto y las tendencias internas al sistema? In *La Romania en Interacción*. Edited by Martina Schrader-Kniffki and Laura Morgenthaler García. Madrid and Frankfurt am Main: Iberoamericana Vervuert, pp. 263–84.
- Palacios, Azucena. 2011. Nuevas perspectivas en el estudio del cambio inducido por contacto: Hacia un modelo dinámico del contacto de lenguas. *Lenguas Modernas* 38: 17–36.
- Uth, Melanie, and Nuria Martínez García. 2020. Circumflex nuclear configurations in Yucatecan Spanish as a supraregional feature: The roles of bilingualism and gender. *Language and Speech* 64: 23830920968130. [CrossRef] [PubMed]
- van Coetsem, Frans. 1988. *Loan Phonology and the Two Transfer Types in Language Contact*. Heidelberg: Winter.
- van Coetsem, Frans. 2000. *A General and Unified Theory of the Transmission Process in Language Contact*. Monographien Zur Sprachwissenschaft Bd. 19. Heidelberg: Winter.
- Vellón Lahoz, Javier. 2019. El artículo en las relativas introducidas por *con* en el siglo XVIII: Contextos y evolución. *Revista de Filología Española* 99: 391–415. [CrossRef]
- Winford, Donald. 2005. Contact-induced changes: Classification and processes. *Diachronica* 22: 373–427. [CrossRef]