

Referentiality in Spanish CPs

Abstract:

In this paper, we discuss the CP domain of embedded clauses in Spanish, specifically in the realm of *que*+embedded question constructions first discussed in Plann 1982. We argue for the existence of (at least) two distinct CP layers (following previous work by Lahiri (2002) Demonte & Fernández-Soriano (2009), and Suñer (1991, 1993)). Following Suñer (1991, 1993), we argue that there are two semantically distinct classes of embedded clauses, although we depart from her by claiming that the relevant distinction should be formulated in terms of referentiality. We claim that her ‘true indirect questions’ are just one case of a non-referential embedded CP (another being a non-referential sentential complement to a non-factive verb). Moreover, we provide evidence that this difference in referentiality corresponds to a structural difference as well: referential CPs have less structure than non-referential embedded CPs. We also offer a classification of embedded clauses based on the presence or absence of an extra CP layer (*cP*) and the presence or absence of a question operator. Finally, we suggest that the overt spell-out of the non-referential head in Spanish embedded clauses is conditioned by the presence of a particular speech-act operator.

Keywords: Spanish, indirect questions, factive and non-factive complements, referentiality, CP layers

1. Introduction

In some Spanish embedded *wh*-questions, the complementizer *que* can appear in a position above the embedded *wh*-word (Brucart 1993, Demonte & Fernández-Soriano 2009, Lahiri 2002, Plann 1982, Suñer 1991, 1993, 1999, Rivero 1980, 1994). This is illustrated in (1a). We refer to this *que* as non-ref(erential)-*que* for reasons we discuss below.¹ In some embedded *wh*-questions, non-ref-*que* cannot appear, as illustrated in (1b). Data in (1) from Suñer (1991: 283).

(1) a. Me preguntaron [**que**] [a quién] invitarás tú al concierto.

To-me they-asked [**that**] [to whom] will-invite you to-the concert

“They asked me whom you will invite to the concert.”

b. Juana no sabía (***que**) [cuándo] visitaría a sus abuelos.

Juana not know (***that**) [when] would-visit to her grandparents

“Juana didn’t know when she would visit her grandparents.”

* We would like to thank the organizers and attendees at the 21st Colloquium on Generative Grammar at Universidad de Sevilla and Universidad Pablo de Olavide in Seville, Spain, where a version of this paper was presented. We would also like to thank two anonymous reviewers for valuable comments and suggestions.

¹ An anonymous reviewer points out that Saito (2010) discusses similar constructions in Japanese, drawing a parallel between Japanese *to* and Spanish *que*. Bhatt and Yoon (1992) had previously discussed some of these properties of Japanese *to* as well as parallel particles in Kashmiri (*ki*) and Korean (*ko*). Bhatt and Yoon (1992) claim that these particles are subordinators, one crucial fact being that they only appear in embedded clauses, never in root clauses. While there are indeed overlaps between Spanish non-ref-*que* and these other particles, if Suñer (1991) is correct in that in root clauses in Spanish the same non-ref-*que* can appear, a complete overlap cannot be maintained, and Spanish non-ref-*que* appears not to be (simply) a subordinator.

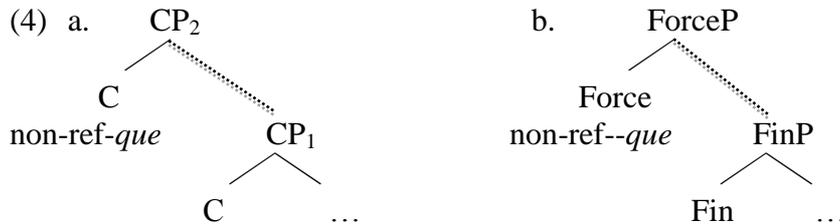
One (widely agreed-upon) generalization, noted first by Plann (1982:303), is that if a verb allows a direct quote of a question as a complement, then it will allow for the presence of non-ref-*que*. If a verb does not allow a direct quote of a question as a complement, then non-ref-*que* is not allowed. This is illustrated in (2) and (3) respectively, data from Suñer (1991:285).²

- (2) a. Juan preguntó/dijo/balbucesó: “¿A quién invitaron?”
 Juan asked/said/babbled: “Whom did they invite?”
- b. Juan preguntó/dijo/balbucesó que a quién habían invitado.
 Juan asked/said/babbled that whom they had invited.
- (3) a. *Pilar confesó/explicó: “¿A quién protegió José?”
 Pilar confessed(= revealed)/explained: “Whom did José protect?”
- b. Pilar confesó/explicó (*que) quién había protegido José.
 Pilar confessed/explained (*that) whom José had protected.

Syntactically, most analyses agree that non-ref-*que* indicates more structure in the CP field, either two CPs (Lahiri 2002, Suñer 1991), as in (4a)

² Note that *decir* “say” has two distinct patterns, one with non-ref-*que*, and one without. Although some have described this pattern simply as optionality of *que*, it is important to note that there are interpretational differences between sentences with and without non-ref-*que*. Lahiri (2002) classifies *decir*-predicates with *que* as Class II (with the complement interpreted as a ‘question utterance’) and *decir*-predicates without *que* as Class IV (taking a propositional complement).

or two positions within a Rizzian left-periphery (Demonte & Fernández-Soriano 2009), as in (4b).³



Independently, de Cuba (2006, 2007) and Haegeman (2006), among others, draw parallel structural conclusions about factive and non-factive sentential complements, examples of which are in (5a) and (5b) respectively.

- (5) a. John regrets that Bill stole the money.
 b. John thinks that Bill stole the money.

They propose that non-factive complements have more syntactic structure than factive complements. Moreover, de Cuba & Ürögdi (2009) argue that this structural difference between sentential complements is related to the referential status of the complement. Specifically, referential CPs have less structure than non-referential CPs. This is illustrated schematically in (6a) and (6b). Non-referentiality arises in the presence of *cP*, which itself embeds referential CP.

³ Note that the dotted lines in the structures in (4) represent the possibility for other positions for movement between the 2 heads represented. Further discussion on this point follows in Section 2.

- (6) a. Referential CP: regrets [CP]
 b. Non-referential cP: thinks [cP [CP]]

In this paper, we make three main claims. The first is that the complements of *lamentar* “regret” and *saber* “know” (and related classes of verbs) are structurally less complex than the complements of *creer* “believe” and *preguntar* “to ask” (and related classes of verbs).⁴ *Saber* “know” and *lamentar* “regret” take a referential CP, and consequently less structure, while *preguntar* “ask” and *creer* “believe” take a non-referential cP, which itself embeds referential CP. This is illustrated schematically in (7).⁵

- (7) a. *saber* “know”/*lamentar* “regret” → Referential CP: [CP]
 b. *preguntar* “ask”/*creer* “believe” → Non-referential cP: [cP [CP]]

Second, a variety of differences in syntactic behavior among verbs that allow for a non-ref-*que* complement will be argued to arise not from the presence of cP itself, but due to the presence of a question operator. This also

⁴ We do not give a typology of the different classes of verbs that take cP and/or CP complements. See, however, Suárez (1991) and Lahiri (2002) for a classification of a wide range of predicates that we feel is generally amenable to our proposal.

⁵ There is considerable overlap between factive complements and an embedded CP structure and non-factive complements and an embedded cP structure. Factivity and referentiality, however, are independent notions, thus there is no one-to-one correspondence between (non-)factivity and (non-)referentiality. See de Cuba & Ürögdi (2009, 2010) for discussion. For ease of exposition, we use the terms *factive complements* and *non-factive complements* to refer to CP and cP respectively in the text. However, see footnotes 6 and 22 for more discussion on the ability of many predicates to select either type of complement, with resulting interpretive differences.

“ask” and *decir* “say”. We suggest that when *que* is spelled out, there is a particular speech-act operator in *cP* relating to a non-initial attempt at introducing a proposition or question into the common ground. Section 5 concludes with a brief recap of our main claim.

2. The left periphery of embedded clauses: structural differences

In this section, we first discuss differences between the ability of the left periphery of referential and non-referential embedded clauses to host moved elements. We see that those embedded clauses that we claim are *cPs* can host displaced constituents, while those embedded clauses that we claim are *CPs* cannot. We take these facts to be consistent with the assumption that non-referential *cPs* contain more structure, while referential *CPs* contain less structure. We then discuss an interesting paradigm of *wh*-extraction facts out of non-referential and referential embedded clauses. We see that the referential status of an embedded clause can explain only part of the extraction paradigm; we must appeal to independent factors to explain the other part: the presence or absence of a question operator.

2.1. Non-referential CPs have more structure

One aspect of our first claim is that there is more structure associated with the sentential complements of verbs like *preguntar* “ask” and *creer* “believe” than with the sentential complements of verbs like *saber* “know” and *lamentar*

“regret”.⁶ That is, there is more structure associated with non-referential sentential complements (i.e. *cP*) than with referential sentential complements (i.e. *CPs*). We take the following set of contrasts with respect to main clause phenomena (MCP) to be consistent with this assumption.⁷ Factive sentential complements do not allow topics, as illustrated in (9), while non-factive complements do, as illustrated in (10).⁸

(9) a. *John regrets that **this book** Mary read

b. *I resent that **this book** he had to examine carefully

(10) a. John thinks that **this book** Mary read.

⁶ Note that semifactives can lose their factivity in questions, if embedded in the antecedent of a conditional, and under certain modals. Semifactives correspond to the Hooper & Thompson 1973 (H&T) class E predicates, and allow main clause phenomena (MCP), unlike true factives (H&T’s class D). Thus, we would expect semifactives like *know* to select either *cP* or *CP* in principle, unless the combination is excluded independently. Bentzen et al. 2007 and Bentzen 2007 show that semifactives can take embedded verb second (a MCP) complements in Mainland Scandinavian for example. In Spanish, *saber* (‘know’) does not allow non-ref-*que*. However, this does not necessarily mean that *saber* never takes a *cP* complement in Spanish, as not all *cP* complements contain non-ref-*que*. See footnote 12.

In addition, Haegeman (2006:1666) discusses examples where a “true factive” (a.k.a. emotive) verb like *regret* behaves more like a non-factive verb, as in (i), her example (24b).

(i) I regret that those details, I cannot reveal to non members.

This instance of *regret* allows the topicalization of *those details* in the embedded clause, raising the possibility that it can take a non-referential *cP* complement. This use of *regret* also allows a modal in the embedded clause, which is not typical of factives (Haegeman 2006:1664:fn. 28). Haegeman argues that such examples are not truly factive: “in this reading *regret* becomes like a speech act verb and, as a result, its complement can be enriched with the ‘speaker deixis’ component which will license a full *CP* complement.” Haegeman (2006:1666). For us, this translates into the proposal that this quite limited use of *regret* (with a meaning akin to “I regret to say...”, or “I regret to inform you...”) does indeed take a non-referential *cP*. For discussion of non-factives selecting *CP* as opposed to *cP*, see footnote 22.

⁷ For more detailed discussion of the truncation analysis for factives, see de Cuba 2006, 2007 and Haegeman 2006.

⁸ The examples in (9) are modified from Maki et al. (1993:3) and the examples in (10) are modified from Hooper & Thompson (1973:479).

- b. I believe/say that **this book** he had to examine carefully.

In a parallel fashion, *preguntar* “ask” complements in Spanish allow clitic left dislocation, illustrated in (11), while *saber* “know” complements do not, illustrated in (12). Data in (11a-b) and (12) from Suñer (1999:2173).

- (11)a. Me preguntaron que **a Juan** qué le había prometido el decano.⁹

Me asked.3pl. that to Juan what him had promised the dean

“They asked me what the dean promised John.”

- b. Le dije que **a su hijo** dónde lo iban a mandar los militares.

Him said.1s that to his son where him go.3pl to send the militaries

“He asked him where the military was going to send his son.”

- c. Juan cree que **ese libro** ya se lo había leído.¹⁰

⁹ It appears that in this sentence, non-ref-*que* must be present. Although at this time we have no explanation for why this might be, observe an interesting contrast with respect verbs like *rogar* “plead” in Spanish, which can omit the complementizer *que*.

- (i)a. Ruego (que) bajen la música. b. Ruego *(que) la música la bajen.
Plead (that) lower the music Plead (that) the music it lower
“I am pleading with them to lower the music.”

Que must appear when there is a left dislocated constituent. Observe that *que* must also appear obligatorily when there is an overt preverbal subject, as illustrated in (ii).

- (ii) Ruego *(que) Jose baje la música.
Plead (that) Jose lower the music
“I am pleading with Jose to lower the music.”

We note, incidentally, that this is in line with Ordóñez and Treviño’s (1999) observations that preverbal subjects in Spanish pattern with other left peripheral elements, and adds support to general approaches which assume that the preverbal subject in Spanish does not occupy Spec,T, but a left peripheral position.

Juan believes.3pl. that **that book** already *se* it had read

“Juan believed that **that book** he had already read.”

(12)a. *Sabía **a Juan** qué le había prometido el decano.¹¹

Knew.1sg to Juan what him had promised the dean

“I knew what the dean had promised John.”

b. *Decidieron **a su hijo** dónde lo iban a mandar los militares.

Decided.3pl to his son where him go.3pl to send the militaries

¹⁰ It may be the case in this example that there is a distinct topic position in Spanish below CP. Observe that this would explain the following data, which illustrate that *saber* “know” can take a left dislocated constituent under *que*.

(i) Juan sabe que ese coche lo han vendido.

Juan know that that car it have sold.

Lit: “Juan knows that that car, they have sold it.”

An anonymous reviewer also points out the following datum, consistent with a potentially lower topic position:

(ii) Sabía qué novelas de Octavio Paz a Juan le iban a gustar.

Knew which novels of Octavio Paz to Juan to.him was to like

“I knew which novels of Octavio Paz’s John was going to like.”

Here *a Juan* is a left-dislocated constituent below a “heavy” *wh*-phrase, following Ordóñez and Treviño (1999). See the discussion in footnote 13 as well.

An alternative possibility here is that this *que* in these instances is, in fact, non-ref-*que*. If so, one consequence for our analysis is that our suggestion in section 4 that one specific type of speech act operator must be present in Spec,cP for non-ref-*que* to be spelled out would have to be broadened to include a wider range of speech act operators. Moreover, a verb like *saber* “know” would be able to take either a referential or a non-referential complement. See footnotes 6 & 12 for related discussion.

¹¹ An anonymous reviewer notes example (i), where a left dislocated constituent appears to the left of the *wh*-word, in apparent conflict with (12a).

(i) Yo no sé, a ti, qué te habrá dicho, pero a mí...

I no know, to you, what to.you will.have said, but to me

“I don’t know, to you, what they said, but to me...”

Other speakers consulted agree with the reviewer’s judgment. Importantly, however, when negation is removed, the same speakers find the sentence ungrammatical, as illustrated in (ii):

(ii) *Yo sé, a ti, qué te dirá, pero a mí...

I know, to you, what to.you will.say, but to me

“I know, to you, what they will say, but to me...”

The presence of negation arguably licenses a non-referential complement of *saber* “know” with the concomitant structural effects as well. See footnote 6 for a related discussion.

“They decided where the military was going to send his son.”

The main point here is that in terms of allowing more material in the embedded left periphery, *preguntar* “ask” and *creer* “believe” complements pattern together on the one hand, while *saber* “know” and *decidir* “decide” complements pattern together on the other.¹² On the not unreasonable assumption that more structure allows for more elements and less structure does not, the facts in (9) to (12) are consistent with our structural claims that there is more structure associated with the sentential complements of verbs

¹² Parallel contrasts can be found in the data from Brucart (1993:95) regarding exclaimatives:

- (i) Cuando la vio llegar a la fiesta, Luis exclamó
when her saw arrive to the party, Luis exclaimed
que qué guapa que estaba María.
that how pretty that was María
“When he saw her arrive at the part, Luis exclaimed how beautiful Mary was.”
- (ii) *Luis sabe que qué bien (que) habla María el inglés.
Luis knows that how well (that) speaks María the English
“Luis knows how well Mary speaks English”

Demonte & Fernández-Soriano (2009), however, provide the following example in Spanish:

- (iii) Lamento que ese coche (*que) no lo compres.
Regret that that car (*that) no it buy
“I regret that this car you are not buying.”

Observe that in (iii), *lamentar* patterns more with *preguntar*, *decir* and *creer* in (11) than with *saber* and *decider* in (12) with respect to the order between the left dislocated *ese coche* and non-ref-*que*. An anonymous reviewer offers a parallel example:

- (iv) Lamento que este libro no lo haya leído María.
Regret that this book no it has read María
“I regret that this book Mary has not read.”

At first glance, this left dislocation pattern is unexpected under our account. We see two possible explanations of these data. First, as noted in footnote 11, it is possible that there is simply a lower topic position available, below CP, which is independent of the referential status of the embedded clause. In fact, as Rivero (1980:367) observes, several classes of verbs allow embedded left dislocated elements in Spanish, including “verbs of saying, volition, and doubt, factives of various types, and implicatives.”

Second, as discussed in footnote 6, *regret* can sometimes pattern with non-factive verbs (in the “regret to inform you” reading) in taking a non-referential *cP* complement. A full treatment of the verbs that can shift between verb classes and the conditions that regulate the shift is clearly warranted. However, we do not carry this task out here.

like *preguntar* “ask” and *creer* “believe” than with the sentential complements of verbs like *saber* “know” and *decidir* “decide”.

2.2. Some extraction facts

It is well documented that factive complements are weak islands for extraction and non-factive complements are not: extraction of adjuncts is generally disallowed from factive complements. This is illustrated in (13) for English and in (14) for Spanish.¹³

(13)a. When_i do you think that John finished the work t_i ?

b. *When_i do you regret that John finished the work t_i ?

(14)a. ¿Cuándo crees que acabó el trabajo Juan t_i ?

when believe.1sg that finished.3sg the work Juan

“When do you believe that John finished the work?”

b. *¿Cuándo lamentas que acabara el trabajo Juan t_i ?

when regret.1sg that finished.3sg the work Juan

“When do you regret that John finished the work?”

In de Cuba & Ürögdi (2009) it is argued that these facts are related to the referentiality of the embedded clause and link this referentiality to differences

¹³ Note that D-linked adjuncts can be extracted from factive complements. We take D-linked adjuncts to be referential, so, following Szabolcsi & Zwarts (1993), they would be predicted to be extractable. See below.

in structure: more structure correlates with non-referentiality, while less structure correlates with referentiality. This is consistent with purely semantic approaches to weak islands such as Szabolcsi & Zwarts (1993), who argue that there can be no variable left unbound within a referential complement. Since non-factive complements are non-referential, extraction can take place.¹⁴ In contrast, sentential factive complements are referential, thus no extraction can take place.

Given that non-referential complements are not islands for extraction ((13a) and (14a)), and since we claim that *preguntar* “ask” complements are non-referential, one might expect complements of *preguntar* not to be islands either. However, non-ref-*que* constructions show classic *wh*-island effects, as illustrated in (15) (from Suñer 1991:301).¹⁵

(15) a. *¿*Quién*_{*i*} preguntaste que cuando llegó *t_i* a este país?

Who_i asked.2sg that when arrived.3sg *t_i* to this country
 “Who did you ask when arrived in this country?”

b. *¿A cuáles de ellos preguntó que quién no les había dado
 to which of them asked.3sg that who not them had.3sg them
 una buena recomendación?¹⁶

¹⁴ Note that under the analysis of de Cuba & Ürögdi (2009), factivity is not the deciding factor for complement size – referentiality is. Therefore, non-factive predicates can sometimes take referential complements. See de Cuba & Ürögdi (2009, 2010) for details.

¹⁵ The glosses in (15) have been modified, and translations have been added, although the data come from Suñer. The same holds for examples (16), (17) and (18).

¹⁶ (15b) is grammatical on the irrelevant matrix reading of *a cuáles* “whom”. Note also that there is a *wh*-island in cases where the extracted *wh*-element is an adjunct.

a good recommendation?

“Which of them did s/he ask who didn’t give a good recommendation to?”

First, note that the island effects are not conditioned by the overt presence of non-ref-*que*, since they remain when non-ref-*que* is absent, as illustrated in (16) from Suñer (1991:303).¹⁷

(16) a. *¿Qué te preguntas dónde compró Luis?

what yourself ask.2sg where bought Luis?

“What did you wonder where Louis bought?”

b. *¿Quién preguntaste si llamó por teléfono?

who ask.2sg if telephoned.3sg by telephone?

“Who did you ask if called on the phone?”

We will claim that the ungrammaticality in (15) and (16) results from the intervention effects of a question operator, given the presence of the embedded *wh*-word. This analysis, however, immediately bears the burden of an explanation for the surprising lack of expected *wh*-island violations in (17) (from Suñer 1991:302).

¹⁷ As we suggest below in Section 4, there is no optionality with respect to the presence vs. absence of overt non-ref-*que*; it is conditioned by certain discourse factors. When there is a specific speech-act operator in *cP*, the head of *cP* is spelled out as non-ref-*que*. If this is the case, then the extraction facts are not related necessarily to the presence/absence of this speech-act operator. There must be an independent factor. As we claim above, this independent factor is the presence or absence of a question operator.

- (17) a. ¿A cuáles de ellos sabes quién no les dio
 to which of them know.2sg who not them give
 una buena recomendación?
 a good recommendation?
 “Which of them do you know who didn’t give a good
 recommendation to?”
- b. ¿Quién_i no recuerdas cuándo llegó *t_i* a este país?
who_i not remember.2sg when arrived.3sg *t_i* to this country
 “Who don’t you remember when s/he arrived to this country?”

In order to explain these differences, we follow the intuition of Suñer (1993, 1999) in claiming that the unavailability of extraction out the complements in (15) and (16) is because they are “true” embedded questions (her *indirect-questions*). In contrast, there are no “true” embedded questions in (17) (her *semi-questions*). This assumption explains the following contrast, also observed by Suñer (1993:57):¹⁸

- (18) a. (Te) digo/repito/recuerdo/se cuáles eran sus actores favoritos:
 You say/repeat/remind/know.1sg which were his/her actors favorite:

¹⁸ Suñer (1991, 1993, 1999) has a series of arguments that show that sentences like (15) and (16) pattern with questions, while sentences like (17) pattern with non-questions. We do not go through all of them here, but see Suñer’s work for ample discussion. Note also that we have modified Suñer’s original examples slightly, by adding the verbs *saber* and *recordar*, and by adding parentheses around indirect object *te* “you”, which is ungrammatical with the verb *saber* “know”.

Nicholson y Depardieu.

Nicholson and Depardieu

“I (will) tell/repeat/remind/know (to you) who his/her favorite actors were: Nicholson and Depardieu.”

b. Te pregunto/repito que cuáles eran sus actores favoritos:

You ask/repeat.1sg. that which were his/her actors favorites:

#Nicholson y Depardieu.

Nicholson and Depardieu.

“I’ll ask you which his/her favorite actors are: Nicholson and Depardieu.”

Whereas an answer is presupposed in (18a), this is not the case in (18b). Stated differently, in (18a) the proposition is resolved, i.e. it is referential; an answer can be provided, and thus its naturalness. In contrast, in (18b), there is no resolved proposition, hence providing an answer which reflects a resolved proposition results in its infelicity.

We argue that there is an intervention effect in (15) and (16) that is not found in (17) due to the presence of a question operator. That is, embedded under *preguntar* “ask” we have the structure in (19a), and under *saber* “know”, the structure in (19b).¹⁹

¹⁹ We are aware that another position is required for the embedded *wh*-word. Perhaps it is in Spec,CP[+REF], or Spec,FinP. We leave this part of the discussion open for now.

- (19) a. *preguntar* “ask” b. *saber* “know”



We also claim that non-factive verbs like *creer* “believe” embed a non-ref cP , just as *preguntar* “ask” does. However, note that unlike the complement of *preguntar* “ask”, the complement of *creer* “believe” does not allow an embedded question, as illustrated in (20).

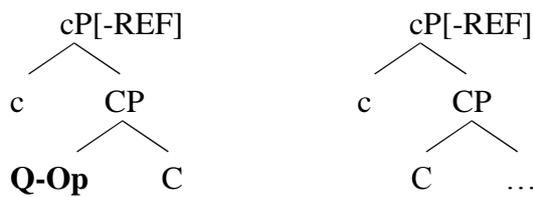
- (20) *¿Crees (que) por qué llegó a este país en balsa?
 Believe.2s.g (that) why arrived.3sg. to this country in raft
 “Do you believe why s/he arrived to this country in a raft?”

We take this to indicate that there is no question operator in the complement of *creer* “believe”. If this is the case, then we would also expect that although *creer* “believe” embeds a cP like *preguntar* “ask”, there should be no *wh*-island effects. This is exactly what we find, illustrated here in (21).

- (21) ¿Quién_i crees que llegó t_i a este país en una balsa?
 who_i believe.2sg that arrived.3sg t_i to this country in a raft
 “Who do you believe that arrived to this country in a raft?”

The difference between *preguntar* “ask” complements and *creer* “believe” complements is summarized in (22).

- (22) a. *preguntar* “ask” b. *creer* “believe”



Note that although there can be a question operator in the left periphery of complements to verbs like *preguntar* “ask”, it is not the case that all verbs that embed non-ref-*que* must have this question operator. That is, not all of these verbs select embedded questions. This has already been pointed out by Lahiri (2002) and Rivero (1994), providing data like those in (23).

- (23)a. Dije que qué bonita estaba el cielo.

I said (that) how nice the sky was. (Lahiri 2002:270)

- b. Dijo que a no molestarle.

said+3s that P not bother+INF-him (Rivero 1994:551)²⁰

“He said not to bother him.”

²⁰ We have altered Rivero’s examples slightly by removing “-“ from between the infinitive *molestar* and the dative *le* for consistency throughout the rest of the text.

In these examples there is no sense in which any question has been introduced by *que*, as Lahiri points out. Lahiri claims that *que* introduces a speech-act. We assume that Lahiri is fundamentally correct, and return to this point in more detail in section 4.

3. Referentiality and proposition resolution

In this section, we offer a working characterization of sentential referentiality and provide examples from discourse contexts and *do so*-replacement in English, as well as sentential referential properties of *it* in English and *lo* “it” in Spanish to support this characterization.

3.1 A working characterization of sentential referentiality

As a working characterization of sentential referentiality, we take a referential CP to be a proposition that refers back to a resolved proposition, where a resolved proposition is a proposition that forms part of the common conversational ground, i.e. the ground shared by the speakers. Consider the following discourse contexts and felicitous and infelicitous uses of referential and non-referential complements in light of this characterization.

In the first discourse context, a parent and a teacher discuss a theft of lunch money that occurred at school. The teacher states (24a) and the parent responds with (24b).

(24)a. Teacher: *Your son stole the lunch money.*

b. Parent: *I regret that my son stole the lunch money.*

The parent's response uses the verb *regret*, which embeds a referential CP. A referential CP refers back to a resolved proposition, one that forms part of the common ground, which in this case is the proposition introduced by the teacher; namely, the sentence in (24a). By contrast, in the same discourse context, and with the same original statement from the teacher from (24b), if the parent responds as in (25), the result is infelicity.²¹

(25) Parent: *#I think that my son stole the lunch money.*

In this case, *think* embeds a non-referential complement. Thus, it cannot refer back to the proposition introduced (under normal intonation) by the teacher. The result is infelicity. Stated differently, *think* introduces an unresolved proposition for acceptance into the common ground, so it is odd to introduce a proposition as unresolved when it is already accepted as part of the common ground shared by the speakers.

²¹ An anonymous reviewer suggests that if *think* is contrastively stressed in (25), then the sentence is fine, with the implication "I can't be sure". While we do not share this judgment in the given context, there may in fact be some influence of contrastive stress on the judgment. As discussed in de Cuba & Ürögdi (2010:47-48), contrastive stress on a typically non-referential complement taking predicate can induce a referential reading of the embedded clause. Indeed, in order for the contrastively stressed version of (25) suggested by the reviewer to be felicitous, the proposition in the embedded clause must already be present in the discourse. In other words, the contrastively stressed version of *think* behaves in just the same way as "regret" in (24b) – it takes a referential CP. The ability of a predicate to select either a referential or a non-referential complement in different circumstances is not unexpected on the view (espoused here, and by de Cuba & Ürögdi (2009, 2010) that referentiality is a property of the embedded clause, not of the selecting predicate. See also the discussion of the informative "regret to inform you" reading of "regret" in footnote 6).

In out-of-the-blue discourse contexts, we find the inverse patterns of felicity with respect to referential and non-referential clauses. Consequently, in a context in which a parent walks up to a teacher and initiates a discourse where there is no previous mention of stealing lunch money nor that the parent's son is involved, the parent's out-of-the blue statement in (26) is infelicitous. (Although, see footnote 6.)

(26) Parent to teacher: #*I regret that my son stole the lunch money.*

Since the proposition *my son stole the lunch money* is not part of the common ground shared between the speakers, no resolved proposition can be referred to, which conflicts with the nature of the embedded sentential complement of *regret*, since it is referential. Now, in the same discourse context, there is no infelicity when the verb embeds a non-referential cP, as is the case for *think*. Thus, the out-of-the blue statement to a teacher by the parent in (27) is perfectly felicitous.

(27) Parent to teacher: *I think that my son stole the lunch money.*

Since the sentential complement is not referring back to any previous proposition, there need not be any previously introduced proposition, so an out-of-the-blue utterance is fine in this case. In fact, the parent introduces this

proposition into the common ground intending for it to be shared by the speakers.²²

3.2 More on referentiality: “do so”, “it” and “todo lo”

In this section, we discuss other elements that appear to interact with the referentiality of the sentential complement. We first discuss syntactic arguments for the differing referential status of some English embedded clauses presented in de Cuba & Ürögdi 2009, and then in Section 3.3 provide a comparable replacement test in Spanish employing *todo lo* (literally “all it”).²³

Observe in (28a) that *do-so* replacement targets the VP, a predicational element, while *it*-replacement targets referential arguments, as illustrated in (28b).

²² An anonymous reviewer wonders whether the following exchange poses a counterexample to our claim that referentiality relates to the common ground shared by the speaker:

- (i) a. De qué te arrepientes en la vida?
Of what you regret en the life
“What do you regret in life?”
b. Me arrepiento de que mis hijos no hayan ido a la universidad.
Me regret of that my children no have gone to the university
“I regret that my children have not gone to University.”

The reviewer rightly points out that in a context where the interlocutors have just met for the first time, the proposition “my children have not gone to University” is not part of the common ground. However, in this context we seem to have a case of what has been called “accommodation” in the semantics and philosophy of language literature (see Beaver & Zeerat 2007 for an overview). In cases of accommodation, a listener accommodates the truth of a proposition into the common ground, despite having not heard the proposition before. It is not uncommon for accommodation, which has been viewed at a repair strategy for a missing presupposition, to occur in cooperative discourse. In the reviewer’s example, it seems clear that the questioner in (ia) is prepared to accommodate the presupposition presented by the responder in (ib). If this is indeed an example of accommodation, then it is not clear that it is a counterexample to our proposal.

²³ For more details, including discussion of referential CPs in Hungarian, see de Cuba & Ürögdi (2009:45-47).

- (28)a. Bill tried the cake, and John did [_{VP} so] too
 b. Bill tried the cake, and John tried [_{DP} it] too

Now consider the contrast exhibited between factives and non-factives in (29), data from (Kiparsky & Kiparsky, 1971:362).

- (29)a. John supposed [that Bill had done it], and Mary supposed [it/so] too.
 b. John regretted [that Bill had done it], and Mary regretted [it/*so] too.

Under a non-factive, as in (29a), the phrase *that Bill had done it* can be replaced with *so* (just like the VP *ate a cake* in (28a)), or with *it*.²⁴ However, only *it* is available under the factive predicate in (29b).

Another piece of evidence for treating CPs as referential expressions comes from the observation (Den Dikken 2008, citing Reeve 2007) that in English *it*-clefts, only specific clefted XPs are compatible with the *wh*-pronoun *which*. Factive complements, interestingly, are also acceptable with *which*.

- (30)a. It's this book which I want to read. (referential)
 b. *It's a doctor which I want to become. (predicative, non-referential)

²⁴ Note that for some speakers, "so" is the only grammatical option in (29a).

- (31) a. It's that John didn't show up which I resent. (referential CP)
 b. *It's that John didn't show up which I believe. (non-referential cP)

Once again, this evidence suggests that the embedded CP here patterns with referential DPs (rather than predicative elements).

3.3 Referentiality and “*todo lo*”

Delbecque & Lamiroy (1999:2030) observe that while in the majority of instances, *lo* “it” can pronominalize an embedded clause in Spanish, there are cases in which it cannot. In these instances, the deictic demonstrative *eso* “that” is used “to refer to what has been said textually in the subordinate clause. This appears to be a mere transposition of direct speech.” They offer the following paradigm to illustrate this.

- (32) a. Eva me {contestó/respondió/objetó} que Juan era un sinvergüenza.
 b. Eva me {contestó/respondió/objetó}: “Juan era un sinvergüenza”.
 c. *Eva me lo{contestó/respondió/objetó}.
 d. Eva me {contestó/respondió/objetó} eso.

In our view, it is not a coincidence that *lo* cannot pronominalize a clause that corresponds to a direct quote. Recall from above that the verbs that allow a direct quote (question) complement also allowed non-ref-*que*. Developing the observation from Delbecque & Lamiroy (1999:2030), we suggest in this section that this pronominal *lo* in combination with the quantifier *todo* “all”

can differentiate referential complements from non-referential complements.²⁵

We take this as support for the conclusion that embedded sentential complements can differ as a function of their referential status.

Brucart (1993 and references therein) discusses the non-referential status of definite DPs that appear with emphatic-*lo*. He observes a contrast which arises in the presence of the quantifier *todo* “all”, taking a discussion of the various interpretations of *lo difícil* (literally “it difficult”) from Bosque & Moreno (1990) as a starting point.

He first discusses two interpretations of *lo difícil* in (32) paraphrased in (33i) and (33ii).²⁶

(33) Nos explicó lo difícil (del artículo).

Us explained.3sg it difficult (of.the article)

i. *Individuating denotation*: He explained to us the difficult part of the article. (i.e. One part of the article is difficult and that part was explained to us.)

ii. *Qualitative denotation*: He explained to us the extreme difficulty of the article. (i.e. It was explained to us that the article is difficult.)

²⁵ An anonymous reviews suggest that the *lo* in the cases we discuss here is the neuter article in Spanish and not a pronominal. We follow Bosque & Moreno (1990) who argue explicitly that these and several other instances of *lo* are in fact pronominal.

²⁶ Bosque & Moreno (1990) observe a third, denotative interpretation, which they refer to as *cuantitativo*. *Individuating* and *qualitative* is our translation of *individuativo* and *cuantitativo* respectively.

The first interpretation is a referential use of *lo difícil*. Observe that it is the only reading available with the quantifier *todo*, illustrated in (34).

- (34) Nos explicó todo lo difícil.
 Us explained.3sg all it difficult
 “He explained every difficult part.”

Now observe in (35) a context in which *lo* is emphatic and non-referential. The *qualitative* use is brought out by the presence of the relative clause. Data from Brucart (1993:78).

- (35) Nos explicó lo difícil que era el problema.
 Us explained.3sg it difficult that was the problem.
 “He explained to us how difficult the problem was.”

Importantly, observe in (36) that *todo* cannot appear on the non-referential *qualitative* use of *lo difícil*.

- (36) *Nos explicó todo lo difícil que era el problema.

What is important for our purposes is that *todo* cannot appear with non-referential *lo*. As noted above, neuter pronominal *lo* can refer back to (most) sentential complements, as illustrated in (37) and (38).

(37) a. María sabía dónde iba su marido todas las noches.

María knew.3sg where go.3sg her husband all the nights.

“Mary knew where her husband went every night.”

b. Juana lo sabía también.

Juana it knew.3sg also.

“Juana knew it too.” → [where Mary’s husband went every night.]

Thus, *lo* in (37b) can refer back to the proposition introduced by the embedded complement of *sabía* in (37a), namely *where Mary’s husband went every night*.

(38) a. María preguntó (que) dónde iba su marido todas las noches.

María asked.3sg (that) where go.3s her husband all the nights

“Mary asked where her husband went every night.”

b. Juana lo preguntó también.

Juana it asked.3sg also

“Juana asked it too.” → [where Mary’s husband went every night.]

In (38b) we see that *lo* can refer back to the complement of *preguntar* in (38a), namely the proposition *where Mary’s husband went every night*.

Now observe that while *lo* and *todo* appear quite naturally with *saber* complements, as in (39), they are degraded with *preguntar* complements, as in (40).

(39) a. María sabía por qué su marido ya no le miraba
 María knew.3sg why her husband already not her looked.3s
 a la cara.
 to the face
 “Mary knew why her husband no longer looked her in the eye.”

b. Lo sabía todo.
 It knew.3sg all.
 “She knew it all.”

(40) a. María preguntó (que) por qué su marido ya
 María asked.3sg (that) why her husband already
 no le miraba a la cara.
 not her looked.3sg to the face
 “Mary asked why her husband no longer looked her in the eye.”

b.??Lo preguntó todo.²⁷

It asked.3sg all

“She asked it all”

The contrast that arises due to the addition of *todo* results from the difference in referentiality of the sentential complements; *cPs* headed by non-ref-*que* are non-referential and *lo todo* is infelicitous with non-referential elements.

4. *Preguntar* and the optionality of *que*

In this final section, we would like to draw attention to the apparent optionality of *que* with *preguntar*, and suggest that, in fact, there is no optionality. We suggest here that the presence of non-ref-*que* is the result of the presence of a speech-act operator in *cP*. Like other speech-act operators, this one introduces a proposition or a question into the common ground. It differs, however, in that it reflects a non-initial attempt on the part of the speaker to have the proposition or question form part of the common ground. That is, from the speaker’s point of view, the proposition or question should

²⁷ This sentence is only good with an interpretation in which there is a previous list of questions and each and every one on that list is asked. In this sense, it appears that *lo* takes on a referential interpretation. One possibility is that since *cP*[-REF] embeds *CP*[+REF], *lo* might be potentially anaphoric on *CP*[+REF]. If this is the case, it is not clear to us why this might be. The contrasts in Spanish also appear to hold in English as well:

- | | |
|---------------------------------|-----------------------------------|
| (i) <i>John regrets it all.</i> | (iii) <i>#John thinks it all.</i> |
| (ii) <i>John knows it all.</i> | (iv) <i>#John asks it all.</i> |

The last example appears to be grammatical, but only on a *request it all* interpretation, as in *John asks it all of his employees, that they give 100%*.

already be shared among the speakers. When this particular speech-act operator appears in *cP*, the head spells-out as *que*. This (non-ref-)*que* reflects the speaker's opinion that it is a non-initial attempt at introducing the following proposition or question into the common ground.

With this in mind, observe the following pair of sentences from Brucart (1993: 97-98):

- (41) a. Se le pidió una candidata y dijo: 'María'.
se him requested.3sg a candidate and said.3sg: 'María'
 "He was asked for a candidate and he said: 'María'."
- b. De repente apareció en el aula y dijo: 'María'.
 of sudden appeared.3sg in the classroom and said.3sg: 'María'
 "All of a sudden he appeared in the room and said: 'María'."

In (41) we see two instances of the verb *decir* taking a name as a direct quote.²⁸ Observe now the contrast in (42), also from Brucart (1993:98).

- (42) a. Se le pidió una candidata y dijo que María.
se him requested.3sg a candidate and said.3sg that María
- b. *De repente apareció en el aula y dijo que María.
 of sudden appeared.3sg in the classroom and said.3sg that María

²⁸ Recall from what was said above that this was a diagnostic for verbs that allow for non-ref-*que*.

Again, we have the same verb *decir* “say” in both cases, yet only in (42a) is non-ref-*que* permitted. This indicates that the ungrammaticality of (42b) is not related to the particular verb itself, but instead to discourse context. In (42a), the second half of the conjunct is an answer to a question, and in the opinion of the speaker, the answer should already be known. In contrast, in (42b) the adverbial *de repente* “all of a sudden” makes it difficult to accommodate a scenario where the speaker would have the opinion that the answer was previously known. Thus, in this case, it is odd to use non-ref-*que*, since the (concealed) proposition or question that is introduced is not understood by the speaker to be a non-initial contribution.²⁹

Consider another example: in a context in which a recently hired man starts a discussion with his new boss and wants to ask a question about his new job, (43) with overt non-ref-*que* is infelicitous.

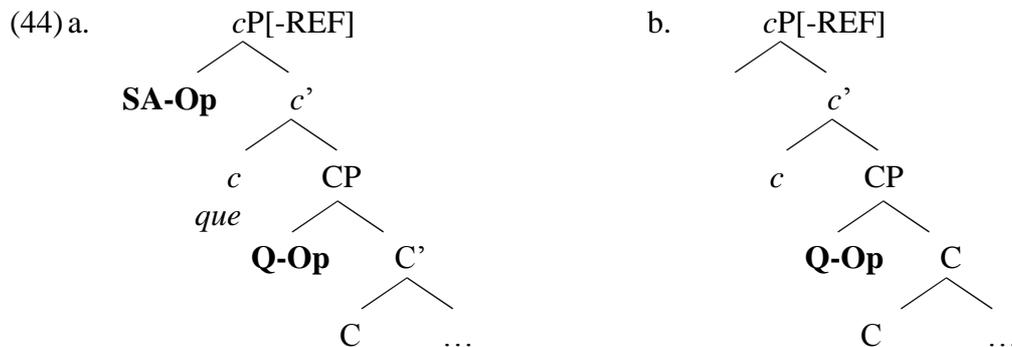
- (43) Quisiera preguntarte (??que) dónde puedo aparcar el coche.
 wanted.1sg ask.you (??that) where can.1sg to.park the car
 “I wanted to ask (*that) where I can park my car.”

The need to indicate that the question *dónde puedo aparcar el coche* “where can I park my car?” should already form part of the common ground is

²⁹ Note that this does not contradict our claim that *cP* containing non-ref-*que* is in fact non-referential. Although the speaker is of the opinion that the proposition should be part of the common ground, the fact that the speaker is using the speech act to reintroduce the proposition shows that the speaker does not yet think it has been completely accepted in the common ground (thus, it is still non-referential).

not clear, precisely because there is no previous discussion of this question in the common ground.

Given that there is a concrete discourse context in which non-ref-*que* is infelicitous, its presence appears to be related to discourse. We have suggested that its presence reflects a non-initial attempt on the part of the speaker to have the proposition or question form part of the common ground. That is, from the speaker's point of view, the proposition or question should already be shared among the speakers. We propose that when this particular speech-act operator (SA-Op) appears in *cP*, the head of *cP* spells-out as *que*, reflecting a non-initial attempt at introducing a proposition or question into the common ground. The trees in (44a) and (44b) illustrate the difference between *preguntar que* and *preguntar* respectively.



Finally, we suggest that the presence of this particular speech-act operator has the potential to explain the following contrasts, observed by Bruccart (1993:92).

- (45) a. Se nos olvidó preguntarle cuándo volvería/volverá.
se us forgot.3sg ask.him when return.cond/return.fut
 “We forgot to ask him when he would/will return.”
- b. Se nos olvidó preguntarle que cuándo volvería/*volverá.³⁰
se us forgot.3sg ask.him que when return.cond/return.fut
 “We forgot to ask him when he would/*will return.”

In (45b), the time of returning is necessarily anchored to the time of the forgetting (embedded predicate), in contrast to (45a), which can also be linked to the time of utterance (Brucart 1993:92). One possibility, the details of which are left to be explored, is that the presence of the speech-act operator in *cP* in (45b) blocks the embedded tense from being anchored to the utterance time.

5. Conclusion: A recap of the main claims

In this paper, we have argued that there are two structurally different types of clausal complements, referential CP and non-referential *cP*. Referential CP, which is structurally less complex, is typically selected by verbs traditionally called *factive*, such as *lamentar* “regret” and *saber* “know”, while non-referential *cP* is structurally more complex, and is typically selected by verbs

³⁰ Note that some speakers may find *volvería* marked here. Although this is the case, *volverá* is still notably worse.

traditionally called *non-factive*, such as *creer* “believe” and *preguntar* “ask” (though there are contexts where factives can select *cP* and non-factives can take CP). The structures are illustrated in (46).

- (46) a. Referential CP: [CP]
 b. Non-referential *cP*: [*cP* [CP]]

We have presented a working characterization of referentiality, claiming that referential complements are accepted as part of the common conversational ground, while non-referential complements are not. In addition, we have suggested that non-referential *que* licenses a particular speech-act operator reflecting a non-initial attempt at introducing a proposition or question into the common ground.

We have presented arguments that the class of complements that Suñer (1991, 1993, 1999) refers to as *semi-questions* are referential CPs, while the class of complements that Suñer refers to as *indirect-questions* are referential *cP*s headed by non-referential *que*. This difference in complement class is exploited to account for syntactic and semantic differences between the classes.

References

- Beaver, David & Henk Zeevat. 2007. In: *Oxford Handbook of Linguistic Interfaces*. Gillian Ramchand & Charles Reiss (eds.). Oxford University Press. 503–538.
- Bentzen, Kristine. 2007. *Order and Structure in Embedded Clauses in Northern Norwegian*. Doctoral Dissertation, University of Tromsø.
- Bentzen, Kristine, Gunnar Hrafn Hrafnbjargarson, Þorbjörg Hróarsdóttir, & Anna-Lena Wiklund. 2007. The Tromsø guide to the Force behind V2. *Working Papers in Scandinavian Syntax* 79:93-118.
- Bhatt, Rakesh & James Yoon. 1992. On the composition of COMP and parameters of V2. In: *Proceedings of WCCFL 10*. Dawn Bates (ed.). Stanford Linguistic Association. 41-52.
- Bosque, Ignacio & Juan Carlos Moreno. 1990. Las construcciones con *lo* y la denotación del neutron. *Lingüística* 2:5-50.
- Brucart, José María. 1993. Sobre la estructura de SCOMP en español. In: *Sintaxi: Teoria y perspectives*. Amadeu Viana (ed.). Lleida: Pagès Editors. 59-102.
- de Cuba, Carlos. 2006. The Adjunction Prohibition and Extraction from Non-Factive CPs. In: *Proceedings of the 25th West Coast Conference on Formal Linguistics*. Donald Baumer, David Montero, and Michael Scanlon (eds.). Somerville, MA: Cascadilla Proceedings Project. 123-131.
- de Cuba, Carlos. 2007. *On (Non)Factivity, Clausal Complementation and the CP-Field*. Doctoral Dissertation. Stony Brook University.

- de Cuba, Carlos & Barbara. Ürögdi. 2009. Eliminating Factivity from Syntax: Sentential Complements in Hungarian. In: *Approaches to Hungarian*: vol. 11. Marcel Den Dikken & Robert Vago (eds.). John Benjamins. 29-63.
- de Cuba, Carlos & Barbara. Ürögdi. 2010). Clearing up the ‘facts’ on complementation. In: *University of Pennsylvania Working Papers in Linguistics* Vol. 16: Iss. 1, Article 6, 41-50.
- den Dikken, Marcel. 2008. A cleft palette: On the landscape of cleft constructions and their syntactic derivations. Talk given at the Cleft Workshop, Berlin: ZAS.
- Demonte, Violeta. & Olga Fernández-Soriano. 2009. Force and finiteness in the Spanish complementizer system. *Probus* 21:23-49.
- Delbecque, Nicole & Béatrice Lamiroy. 1999. In: *Gramática Descriptiva de la lengua española*. Violeta Demonte & Ignacio Bosque (eds.). 1965-2081.
- Haegeman, Liliane. 2006. Conditionals, factives and the left periphery. *Lingua* 116:1651-1669.
- Hooper, Joan, & Sandra Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4:465-497.
- Kiparsky, Paul & Carol Kiparsky. 1971. Fact. In: *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*. Danny D. Steinberg and Leon A. Jakobovits (eds.). Cambridge: Cambridge University Press. 345-369.
- Lahiri, Utpal. 2002. *Questions and answers in embedded contexts*. Oxford: Oxford University Press.

- Ordóñez, Francisco and Esthela Treviño. 1999. Left dislocated subjects and the pro-drop parameter: A case study of Spanish. *Lingua* 107:39-68.
- Maki, Hideki, Lizanne Kaiser & Masao Ochi. 1999. Embedded topicalization in English and Japanese. *Lingua* 109:1–14.
- McCloskey, James. 2006. Questions and Questioning in a Local English. In: *Crosslinguistic Research in Syntax and Semantics: Negation, Tense, and Clausal Architecture*. Raffaella Zanuttini, Héctor Campos, Elena Herburger, Paul H. Portner, (eds.). Georgetown University Press. 87–126.
- Plann, Susan. 1982. Indirect questions in Spanish. *Linguistic Inquiry* 13.2:297-312.
- Rivero, María-Luisa. 1980. On left dislocation and topicalization in Spanish. *Linguistic Inquiry* 11:363-393.
- Rivero, María-Luisa. 1994. On indirect questions, commands and Spanish quotative *que*. *Linguistic Inquiry* 25.3: 547-554.
- Suñer, Margarita. 1991. Indirect questions and the structure of CP: Some consequences. In: *Current Studies in Spanish Linguistics*. Hector Campos & Fernando Martinez-Gil (eds.). Georgetown University Press. 283-312.
- Suñer, Margarita. 1993. About indirect questions and semi-questions. *Linguistics and Philosophy* 16:45-77.
- Suñer, Margarita. 1999. La subordinación sustantiva: la interrogación indirecta. In: *Gramática Descriptiva de la lengua española*. Violeta Demonte & Ignacio Bosque (eds.). 2149-2195

- Szabolcsi, Anna and Frans Zwarts. 1993. Weak islands and an algebraic semantics for scope taking. *Natural Language Semantics* 1:235-284.
(reprinted in *Ways of Scope Taking*, Anna Szabolcsi (ed.). 217-262.
Dordrecht: Kluwer.
- Torrego, Esther.1984. On inversion in Spanish and some of its effects.
Linguistic Inquiry 15.1:103-129.