

Embedded interrogatives as free relatives

Cecilia Poletto and Emanuela Sanfelici

1 Introduction

In this work we intend to show that a typical property of North-Eastern Italian dialects (NEIDs) namely the sequence *wh*-complementizer in embedded interrogatives (which in some dialects also extends to non-standard (under the definition by Obenauer 2006) and even standard main questions originally developed out of a free relative structure where the complementizer is actually not a complementizer but part of the internal structure of the *wh*-item heading the relative clause. The structures we are going to investigate are illustrated in (1), where we see that the *wh*-item is followed by the item *che*, which has always been interpreted as a complementizer.

- (1) No so cossa che el gabia dito.
Not know what that he has said
'I do not know what he said.'

If we consider this phenomenon in traditional terms, it looks like an apparent violation of the traditional "doubly filled comp filter", since the C^0 head is occupied by the complementizer and specifier of the CP by the *wh*-item. Even under a more restricted view, namely that the doubly filled comp filter can only be violated if the head and the specifier agree (see Koopman 1996), there is no morphological hint of an agree operation between the *wh* and *che* that might justify the violation. This structure is problematic even in a more modern framework: for instance the exceptionality of these structures has been noted (see Poletto & Vanelli 1995 and Poletto 2000) within the cartographic framework as cases in which the complementizer does not sit in its usual Force⁰ position as it does in declaratives (see Rizzi 1997), but as a case of a low complementizer probably located in the low Fin/Focus area. Even assuming that the complementizer is first merged in the Fin⁰, since it has [+finiteness] features and then

raises to Force^o, we have a case in which the complementizer does not reach the Force^o position. Alternatively, we have to assume that the Force^o of interrogatives is lower than the Force^o of declaratives, which conceptually empties the notion of Force^o as marking sentence type. Furthermore, the possibility to have at the same time high and low complementizers occurring after left dislocated items and *wh*-items has been identified as a weakness of the cartographic approach (see van ?), where ideally every element should have its own position. Furthermore, *che* is usually the complementizer of declarative clauses, while interrogative clauses (at least yes/no interrogatives use *se* 'if'). Here we intend to explore an alternative which only recently became available since Poletto & Sanfelici (2015) have argued that the "complementizer" *che* found in relative clauses in Italian is actually a determiner-like portion of the internal head of the relative clause. We will argue that the first instances of *wh-che* structures in embedded interrogatives are actually free (or light headed) relative clauses, so that the element *che* found in these cases is not a complementizer at all, but is part of the internal structure of the *wh*-item. In order to show how this might work, we will take various NEIDs into consideration.

2 Venetian

In Venetian, and in the majority of Veneto dialects, the sequence [*wh-che*] is obligatory with all *wh*-items in embedded interrogatives with no exception. The lack of *che* results into ungrammaticality. We provide here some examples to illustrate the point. These structures are so stable that they can occasionally leek into the regional standard Italian of native Veneto speakers:

- (2) No so dove *(che) el sia 'nda.
Not know where that he is gone
- (3) No so chi *(che) le gabia visto.
Not know who that them have seen
- (4) No so come *(che) el lo gabia fato.
Not know how that he it have done
- (5) No so parcossa *(che) el sia ndá via.
Not know why that he be gone away
- (6) No so quanti pomi *(che) el gabia comprá.
Not know how many apples that he have bought

- (7) No so che zente *(che) el gabia visto.
Not know which people that he have seen

Interestingly, Old Venetian does not possess this property in the same way the modern variant does. The following examples are all extracted out of the *Tristano Veneto*, a medieval text dating around the end of the XIII or the beginning of the XIV century. They show that in Old Venetian it is indeed possible not to have *che* after then wh-item in embedded interrogatives:

- (8) çerchando **qui** **novelle** li savaesse dir del chavalier che
looking for which news them could say of.the knight that
persegiva la Bestia Gladisschante
chases the Bestia Gladisschante (370, 32)
- (9) domandé **in que** **logo** die eser questa bataia.
asked in which place should be this battle (69, 45)
- (10) Ma atanto me dixé **in qual** **parte** se va li cavalieri.
But in.as.much to.me said in which PART CL go the knights (145, 23)
- (11) et se algun me domandasse **de qui** era la nave
and if some to.me ask of whom was the ship (168, 8)

This is generally true of complex wh-phrases but also of bare wh-items:

- (12) Lo re domandà Tristan **qui** era lo chavalier che chavalchava
The king asked Tristan who was the knight that was riding
sì solo
alone (89, 47)
- (13) Et ello li domandá **donde** elli era
And he to.them asked where he was (165, 13)
- (14) io non so **qui** elli sia questi tre cavallieri de que le letre
I not know who they are these three knights about whom the letter
parla
talks (22, 14)

Notably, the element meaning *where*, besides the form in (13), also has a different structure, which looks like a free relative clause with a light head, like the pronoun 'there' *lá*. The same is true of the item 'why', which displays the structure of a relative with the light noun 'reason' *chasion*:

- (15) e domandava **lá** o' ch'elo sia lo re
 and asked there where that he was the king (42, 3)
- (16) ela a lor domandà **la chasion perché** eli gera vignudi
 She to them asks the reason why they were come (407)

The same can also occur with the wh-item *qui*, 'who', which has no pronoun or light noun, but can (though it need not, cf. (12)) be followed by *que*:

- (17) voleva saver **qui qu'ello** era
 wanted to know who that he was

Notice that these cases have to be analyzed as embedded interrogatives, since they are selected by the verb *domandar* 'ask' and *saver* 'to know'. However, the case in (18) shows that the light head represented by the deictic pronoun *lá* is not necessary to have the structure *o-que*. This pattern resembles that exhibited by free vs. light-headed relative clauses: a free relative clause can have a null head as in (18) or a light-headed one as in (15) (Poletto & Sanfelici 2015, 2018).

- (18) ma molto elo se maravegiava o' **qu'ella** podeva eser andada
 but much he CL was amazed where that she could have gone
 (368, 28)

The phenomenon *wh-che* in Old Venetian is thus found primarily with 'where', 'who', 'what' and 'why', and in this last case only with a light headed relative clause, while the other wh items can also have a free relative tout court. The sequence is not found with 'how', and 'from where' or with complex wh phrases in the Tristano Veneto. We can conclude that Old Venetian displays the first stage of the phenomenon which then extends to all wh-items, i.e. embedded interrogatives actually display the structure of a free relative clause with a light head noun and the following *che*.

3 Trentino

One further dialectal area that presents this phenomenon, although in a more limited way, is Trentino: here not all wh-items have to be followed by *che*, but only some of them. Garzonio (2007) already notices this fact on the basis of the survey of the ASIt database:

- (19) a. Voi saver ndo che i è nadi i to amizi.
 want know where that CL were gone the your friends [Trento]
- b. No so cosa la mama l abia comprà per zena.
 Not know what the mother cl has bought for dinner
- c. Di-me ci che as vist ieri.
 tell-me who that you have seen yesterday [Val di Non]
- d. Di-me cando torna el Giani.
 Tell-me when comes back the Gianni

The realization of *che* is indeed variable, since the same speaker translates sentences with and without *che* for the same wh-item in the prompt.

The same observation is proved by a further investigation based on her own field work by Polonia (2014) for the Trentino dialects spoken in Val di Sole, Val di Cembra and Val di Non. She notices that the occurrence of wh-*che* is not obligatory with various wh items, but there is a clear tendency to use *che* with the following wh-items:

- (20) Dime kel **che** le te a dit le to sorele.
 Tell-me that that CL to.you have said your sisters [Tassullo]
- (21) No so chi **che** lava giò i piatti.
 Not know who that washes down the dishes [Stedron, Segonzano]

Notice that the structure with the wh-item corresponding to 'what' in (20) is the same used in French embedded interrogatives where there is no wh-items but a light headed relative clause whose head is still clearly the distal demonstrative (*ce que*). Polonia (2014) provides a hierarchy of wh-items which can more frequently be doubled by *che*, although she also notes that the same speaker varies in using *che* after the very same wh-item or not. A simplified summary of the implicational scales she reports for the three valleys she investigates is reported here:

- (22) What/where who how many which when which X how
 many X why/how

The data from Old Venetian and those from Trentino only partially overlap. We can recognize some tendencies, since in all the varieties the elements which display the doubling are where, who, what but Old Venetian also tolerates why (with the light noun chasion) while complex wh-phrases are clearly at the right edge of the spectrum and the element corresponding to 'how' is also generally

not doubled.

4 Preliminary notes for a syntactic analysis

Poletto & Vanelli (1995) propose the following generalization:

- (23) If a dialect displays *wh-che* in main interrogatives it does also in embedded interrogatives.

Here we propose that this generalization can also be read in terms of diachronic development, such that *wh-che* starts out in embedded interrogatives and then expands into main interrogatives. Poletto & Vanelli single out the last step of the development of this structure, namely the point when it extends from embedded to main interrogatives, which we will not deal with here (see Poletto 2000 for an analysis in cartographic terms). At this point two questions arise: (i) how do these structures start out in embedded interrogatives?; (ii) When we take into consideration the data presented in the previous section where we showed that the presence of the *wh-che* in embedded interrogatives depends on the type of *wh* items introducing them, how can the generalization in (23) be implemented.

Addressing the first question, we propose that the trigger for the rise of *wh-che* is the semantic ambiguity which can arise between free and interrogative clauses under certain predicates (see Caponigro 2003). Parry (2003) shows that in Old Piedmontese the first attested cases of *wh-che* are precisely ambiguous between a free relative and an interrogative reading. One might thus simply say that there was a semantic ambiguity between a free relative and an embedded interrogative (along the lines of Cecchetto & Donati 2015) and this is the reason why the structure *wh-che* could be extended from free relative clauses to embedded interrogatives. However, we think that the question is more complex than that, and that the first link towards the intrusion of a relative structure in an interrogative one is not between embedded interrogatives and free relatives altogether but between embedded interrogatives and light headed free relatives. Free relatives can either be expressed by a simple *wh* item, as shown in (24), but Italian varieties (as other Romance languages) have the tendency to realize light heads, especially with the element corresponding to *what*, which, just like in French, is utterly impossible as a free relative (25) and requires a light head (26) (see Munaro 2000).

- (24) Chi ha detto questo, ha sbagliato.
 'Who said this was wrong'
- (25) *Cosa hai detto, é sbagliato.
 'What you said is wrong'
- (26) Quello che hai detto, é sbagliato.
 'The thing that you said is wrong'

As predicted under our proposal, if *wh-che* embedded interrogatives are free/light-headed relative clauses, we expect the embedded interrogative with the element corresponding to 'what' to be impossible at least in certain varieties. This prediction seems to be born out: in certain Trentino dialects the embedded interrogatives on 'what' cannot have the *wh*-item for 'what', but require the light-head *quello che* "that that" as they do in real free relatives (Garzonio 2007). Poletto & Sanfelici (to appear) attribute the spell out of a light head to the fact that, as assumed by Cinque (2013) and originally proposed by Sauerland (1999, 2003), all relative clauses are double headed: they all have an external head located in the DP spine and an internal one located inside the relative clause. Languages differ with respect to the head they lexically spell out, and free relatives can be analyzed as restrictive relatives where neither head is lexically realized. Notice however that there is an "intermediate stage" between real free relatives and headed restrictive relative clauses, namely light headed relatives, which are exactly the type of relatives that first manifest themselves in embedded interrogative contexts. We surmise that the variation found in Old Venetian attests precisely this: the first relative structures to be found in interrogative domains are light headed relatives as the ones found with *lá o que* and *la chasion que*. The fact that the first step of the evolution of the *wh-che* structure is a light headed relative clause is shown by the fact that these are the only possible structures even in languages, like French, where no other *wh*-item presents this possibility. If French represents the first stage of the development of this structure, NEIDs represent the second one: the usage is extended to *wh*-pronouns, where there is no lexical light head. If this were the end of the story, we would expect that all *wh*-items should allow for doubling except for the complex ones already containing a lexical head. However, when we take into consideration the data presented in the previous section where we showed that the presence of the *wh-che* in embedded interrogatives depends on the type of *wh* items introducing them, this expectation is not fully fulfilled. The stage depicted by Polonia (2014) and Garzonio (2007) for Trentino dialects only partially matches this expectation: we

find that the *wh*-items that tolerate the *wh-che* structure are a) those that still display the light head (like the case of *kel che* in (20)); b) those that allow for a null free relative head are not all possible *wh*-items, but there is a second hierarchy even if we exclude light headed relatives. The implicational scale in (23) shows that elements like *why* and *how* are much more resistant, even more than complex *wh*-items. We are forced to assume that there is one more factor at play here. If we leave complex *wh*-items aside, the hierarchy in (23) partially looks like the Keenan and Comrie's case hierarchy, which goes from the less complex to the most complex cases: here evidently *what*, *where* and *who* are less complex than *why* and *how*, which are at the rightmost edge of the scale. It is well known that the Keenan and Comrie's hierarchy is at play in relative clauses in the phenomenon that is known as the case containment condition, i.e. the condition ruling which case is spelled out on the lexical relative head when the case of the main clause and the case of the relative clause do not match. Poletto & Sanfelici (to appear) show that modern Italian generally resolves case mismatches in favor of the external case:

- (27) Partiró con chi hai incontrato ieri.
 'I will leave with whom you met yesterday'
- (28) *A chi hai telefonato é partito.
 'To whom you phoned left'

This is the case also for the NEIDs we are considering, as the translation in modern Venetian shows:

- (29) Ndaró via co chi che te ga visto jeri.
 'I will leave with whom you met yesterday'
- (30) *A chi che ti ghe ga telefoná el ze partito.
 'To whom you phoned left'

Suppose that embedded interrogatives in these dialects have the structure of free relative clauses (as the light headed cases overtly show). This means that they have two heads, a relative internal one, and an external one, which is the (null) DP that is the complement clause of verbs like 'ask'. If we apply the double-head idea to embedded interrogatives, we immediately explain the first step of the evolution, i.e. light headed relatives, which spell out both the external light head and the internal *wh* head: in *la chasion perché*, *la* is the determiner of the whole DP, *chasion* is the external head and *perché* represents the internal head

by means of a relative pronoun. Cases like *lá o che* display the external head (the deictic pronoun *lá*) and the internal one, which is made up of two components [*o che*]. This means that at the successive stage of development, i.e. the one of a real free relative with no lexically spelled out external head, we only have the internal one [*o che*]. One might wonder why an element like *che* intrudes into a wh-item and what portion of its internal structure it spells out. The fact that *che* can be part of a wh-item is actually straightforward if we consider the form corresponding to 'what' in standard Italian, namely *che cosa*. Poletto & Pollock (2009) propose the following internal structure for wh-items:

(31) [DisjP [ExistP [RestrictorP]]]

Suppose that in Italian *cosa* represents the restrictor, since it literally means 'thing' and *che* the existential component as in (32):

(32) [DisjP [ExistP *che* [RestrictorP *cosa*]]]

This structure immediately explains the complex nature of wh-pronouns and the ordering of the elements we observe. An independent argument in favor of the idea that wh-pronouns have the complex structure illustrated in (32) is that there are dialects where two components, i.e. the restrictor and the disjunctive as in (33) and (34) can occur separately giving rise to doubling structures (see Obenauer 2006, Munaro & Poletto 2018):

(33) Cossa inviti-to chi?!
 COSSA invite-you who(m)
 'Who (the hell) are you inviting?!'

(34) Cossa ve-to dove?!
 COSSA go-you where
 'Where (the hell) are you going?!' [Paduan (Central Veneto)]

Notice that our free relative/interrogative case represents the third logical possibility besides the combinations of existential plus restrictor and restrictor plus disjunctive:

(35) [DisjP *o* [ExistP *che* [RestrictorP]]]

Notice that also this combination is found in doubling structures:

- (36) Che 'ncontre-t chi?
Wh meet-you who
'Whom are you meeting?'
- (37) Ch' ö-t qual?
Wh want-you which
'Which one do you want?' [Malonno (Eastern Lombard)]

On this basis we can conclude that the structures [wh *che*] in embedded interrogatives is not a case of low complementizers but the spell out of the ExistentialP, one of the internal projections of the wh-item. In this sense interrogative and relative pronouns are the same, i.e. they contain at least three layers, which can be lexically spelled out or not. The reason why the first embedded interrogatives to manifest the spell out of the existential component *che* are actually ambiguous with free relative clauses has to do with the fact that in general in relative clauses ExistentialP is always spelled out by *che* in Italian varieties.

5 Concluding remarks

In this work we have entertained the hypothesis that the occurrence of *che* after wh-items found in NEIDs is not a case of violation of the doubly filled comp filter, where the complementizer *che* sits in a low C° head. We have proposed that these structures, just like relative clauses, are the spell out of more than one internal projection of the wh-item. This explains why the first attested historically cases are indeed free relative clauses, since relative clauses generally spell out the existential portion of the internal relative head. It can also explain well known cases of wh-doubling in simple wh-questions, which are rather rare even in languages like colloquial German and non-standard Dutch varieties which display a so called scope marker wh- doubling the actual wh-item in long extractions.

References

- Caponigro, I. 2003. *Free not to ask: On the semantics of free relatives and wh-words crosslinguistically*. UCLA dissertation.
- Cecchetto, C. & Donati, C. 2015. *(Re)labeling*. Cambridge, MA: MIT Press.
- Cinque, G. 2013. *Typological studies. Word order and relative clauses*. London: Routledge.

- Garzonio, J. 2007. Complementatori nelle interrogative delle varietà trentine: Variazione diatopica e generazionale. In G. Marcato (ed.), *Dialetto, memoria e fantasia. Atti del convegno, Sappada 28 giugno-2 luglio 2006*. 179–183. Padova: Unipress.
- Koopman, H. 1996. The spec head configuration. *Syntax at Sunset, UCLA Working Papers in Syntax and Semantics* 1. 37–64.
- Munaro, N. 2000. Free relative clauses as defective *wh*-elements: evidence from the northwestern Italian dialects. *University of Venice Working Papers in Linguistics* 10. 89–120.
- Munaro, N. & Poletto, C. 2018. Pattern of *wh*-in-situ and *wh*-doubling in the Northern Italian dialects. Paper presented at the Geneva Workshop on optional *wh*-in-situ, January 2018.
- Obenauer, H.G. 2006. Special interrogatives – left periphery, *wh*-doubling, and (apparently) optional elements. In J. Doetjes & P. Gonzalves (eds.), *Romance Languages and Linguistic Theory 2004 – selected papers from 'Going Romance 2004'*. 247–273. Amsterdam: John Benjamins.
- Parry, M. 2003. Cosa ch'a l'é sta storia? The interaction of pragmatics and syntax in the development of WH-interrogatives with overt complementizer in Piedmontese. In C. Tortora (ed.), *The syntax of Italian dialects*. 152–174. Oxford: Oxford University Press.
- Poletto, C. 2000. *The higher functional field: Evidence from Northern Italian dialects*. Oxford: Oxford University Press.
- Poletto, C. & Pollock, J.Y. 2009. Another look at *wh*-questions in romance: The case of Mendriosiotto and its consequence for the analysis of French in situ and embedded interrogatives. In D. Torck & W.L. Wetzels (eds.), *Romance Languages and Linguistic Theory 2006 – selected papers from 'Going Romance 2006'*. 199–258. Amsterdam: John Benjamins.
- Poletto, C. & Sanfelici, E. 2015. On demonstratives in relative clauses. In M.G. Busà & S. Gesuato (eds.), *Lingue e contesti. Studi in onore di Alberto Mioni*. 561–571. Padua: CLEUP.
- Poletto, C. & Sanfelici, E. 2018. Demonstratives as relative pronouns: New insight from Italian varieties. In M. Coniglio, A. Murphy, E. Schlachter & T. Veenstra (eds.), *Atypical demonstratives. Syntax, semantics and pragmatics*. 95–126. Berlin: Mouton de Gruyter.
- Poletto, C. & Sanfelici, E. to appear. On relative complementizers and relative pronouns. *Linguistic Variation*.
- Poletto, C. & Vanelli, L. 1995. Gli introduttori delle frasi interrogative nei dialetti italiani settentrionali. In E. Banfi, G. Bonfadini, P. Cordin & M. Iliescu (eds.),

- Italia settentrionale: crocevia di idiomi romanzi. Atti del convegno internazionale di studi. Trento, 21-23 ottobre 1993.* 145–158. Berlin: Mouton de Gruyter.
- Polonia, K. 2014. *Wh + che: Die komplexe Struktur einer Interrogativstrategie am Beispiel der Trienter Varietäten*: Goethe University Frankfurt MA thesis.
- Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of grammar: Handbook in generative syntax*. 281–337. Dordrecht: Kluwer Academic Publishers.
- Sauerland, U. 1999. Total reconstruction, PF-movement and derivational order. *Linguistics* 7. 1–38.
- Sauerland, U. 2003. Unpronounced heads in relative clauses. In K. Schwabe & W. Susanne (eds.), *The interfaces*. 205–226. Amsterdam: John Benjamins.