

# ‘What’ clauses can and ‘which’ cannot: a Romanian puzzle

IVANO CAPONIGRO

University of California San Diego

ANAMARIA FĂLĂUȘ

CNRS, Laboratoire de Linguistique  
de Nantes

Draft, 2021/1/31

## 1. The puzzle

Romanian exhibits a puzzling contrast that, to the best of our knowledge, hasn’t yet been noticed in the language nor has it been attested in any other Romance or Indo-European language. Like in English, Romanian interrogative *wh*-clauses can be introduced by the complex *wh*-phrases ‘what’ + NP (*ce* in Romanian) or ‘which’ + NP (*care* in Romanian), as shown in (1).

(1) *Interrogative clauses introduced by ‘what’/ ‘which’ + NP*

Ana se întrebă [**ce/ care**] {**carte/cărți/ mâncare**} i-ar plăcea].  
Ana REFL wonders what/ which book/ books/ food CL.DAT-would like  
‘Ana wonders {what/which} {book/books/food} she would like.’

This parallelism breaks down in non-interrogative *wh*-constructions whose distribution and interpretation resemble those of NPs/DPs or PPs: Maximal Free Relative Clauses (henceforth, Max-FRs), which have been argued to semantically behave like definite DPs<sup>1</sup>, and Existential Free Relative Clauses (henceforth, Ex-FRs), which only occur as complements of a restricted class of predicates and have been argued to semantically behave like narrow scope indefinites.<sup>2</sup> Both *wh*-clauses are extremely productive in Romanian: they can be introduced by all the simple *wh*-words and complex *wh*-phrases

---

<sup>1</sup> See Jacobson (1995), Dayal (1996), Caponigro (2003, 2004). They all argue that Max-FRs are *wh*-clauses that denote the unique maximal individual of a set like definite DPs, borrowing from the analysis of definite DPs in Sharvy (1980) and Link (1983). Their proposals differ in the semantics they assume for *wh*-words in Max-FRs. See Šimík (2020) for a detailed and complete overview of generalizations and proposal about Max-FRs and free relative clauses in general, and Caponigro *et al.* (2021) for further detailed crosslinguistic evidence from fifteen Mesoamerican languages.

<sup>2</sup> See Izvorski (1998), Caponigro (2003, 2004), Grosu (2004), and Šimík (2011, 2013). Their analyses are significantly different: Izvorski treats them as interrogative *wh*-clauses syntactically; Caponigro treats them as a subgroup of free relative clauses; Grosu and Šimík, instead, argue that they are not necessarily clausal and, therefore, prefer to treat them as a separate construction (“Modal Existential Construction”). See Šimík (2017) for an overview of the work on this construction and Caponigro *et al.* (2021) for further crosslinguistic evidence from fifteen Mesoamerican languages.

that can introduce interrogative *wh*-clauses (see Grosu 2013, Caponigro & Fălăuș 2018, a.o.). This includes *ce* 'what' + NP, as shown in (2) and (3) with a singular count, a plural count, and a mass NP complement.

(2) *Max-FRs introduced by 'what' + NP*

Ana a luat [**ce** {**carte/ cărți/ mâncare**}a luat și Maria].  
 Ana has taken what book books food has taken also Maria  
 'Ana took the book/books/food Maria took as well.'

(3) *Ex-FRs introduced by 'what' + NP*

Maria nu are [**ce** {**carte/ cărți/ mâncare**} să ia ].  
 Maria not has what book books food SUBJ take.3SG  
 'Maria doesn't have any book/books/food to take.'

There's only one exception: *care* 'which' + NP, which cannot introduce either non-interrogative construction regardless of the kind of NP complement occurring with the *wh*-determiner, as shown in (4) and (5).<sup>3</sup> We are not aware of any variation about these judgments across Romanian varieties or speakers.

(4) *Max-FRs introduced by 'which' + NP*

a. \*Ana a luat [**care** **carte/ mâncare** a luat-o și Maria].  
 Ana has taken which book/ food has taken-CL.ACC.3SG also Maria  
 b. \*Ana a luat [**care** **cărți** le-a luat și Maria].  
 Ana has taken which books CL.ACC.3PL-has taken also Maria

(5) *Ex-FRs introduced by 'which' + NP*

\*Maria nu are [**care** **carte/ cărți/ mâncare** să (o/le) ia ].  
 Maria not has which book books food SUBJ CL.3SG/CL.3PL take.3SG

Unlike the pattern in the interrogative *wh*-clause in (1), the pattern in Max-FRs and Ex-FRs in (2)–(5) doesn't generalize to other Indo-European languages. A systematic investigation of these two options in either construction across Indo-European languages is still lacking. There is at least one Romance language, namely Italian, where neither option is allowed in either construction. The incorrect generalization of this language-specific restriction to all languages with Max-FRs has led to theories of labelling that erroneously predict the impossibility of 'what' + NP in Max-FRs in any

---

<sup>3</sup> As pointed out by a Reviewer, this contrast is not explicitly mentioned in previous work, but it can be inferred from the description of Romanian FRs in Grosu (2013: 645) and the fact that *care* is missing from this list: "The relative elements that can head a DP (i.e. relative determiners) are *ce* 'what', *orice* 'whatever' and *oricare* 'whichever'."

language.<sup>4</sup> Germanic languages do not seem to have Ex-FRs at all. On the other hand, English has a productive system of Max-FRs, including those introduced by *what* + NP. The latter exhibit a behavior that is partially different from and less clear-cut than the behavior of their Romanian counterpart, though. Andrews (1975) reports them as acceptable with a plural or mass NP complement, but unacceptable with a singular count NP complement, as shown in (6)a,b.<sup>5</sup>

- (6) a. I drank [**what beer**]/[\***what glass of milk**] was provided].  
 b. Fred hid [**\*what weapon** was]/[**what weapons** were] on the table].
- (7) % Ana took [**what {book/books/food}** Maria took].

The English native speakers we asked confirmed Andrews' judgments, but found (7)—the English equivalents of the fully acceptable Romanian examples in (2)—either all degraded or all acceptable, regardless of the NP complement. This variation in acceptability of Max-FRs introduced by *what* + NP depending on the examples and the speakers needs a study on its own. On the other hand, FRs introduced by *which* + NP, which Andrews (1975) doesn't mention, are all less acceptable than their *what* + NP counterpart without any kind of variation, at least according to our consultants.

In this squib, we focus on the puzzle in (2)–(5) showing that, in Romanian, 'what' + NP can introduce Max-FRs without restrictions, while 'which' + NP cannot. We suggest a line of explanation that links this puzzle to a meaning difference in D-linking or discourse anaphoricity between 'what' + NP and 'which' + NP in interrogative clauses—a meaning difference that we show is even sharper in Romanian than in English and makes 'which' + NP in Romanian resemble the partitive construction 'which of the' + NP. The fact that 'what' + NP can introduce Max-FRs and Ex-FRs follows as the default option for a language like Romanian that has a productive system of *wh*-clauses, where all *wh*-words and *wh*-phrases can be used, unless some extra factor intervenes and blocks this extension. We suggest that this is exactly the case for 'which' + NP.<sup>6</sup> The strong presuppositional anaphoric requirement for its NP to refer to a set that has already been explicitly introduced in the discourse clashes with the non-anaphoric nature of Max-FRs and Ex-FRs.

<sup>4</sup> E.g., Donati & Cecchetto (2011); see also critical remarks in Caponigro (2019).

<sup>5</sup> (6)a is adapted from Andrews (1975: p. 76, ex. 76b,c) and (6)b from Andrews (1975: p. 76, ex. 77b,c). Andrews (1975) calls the bracketed clauses in (6)a,b "internally headed relative clauses." See de Vries (2002: Ch. 2), in particular Sec. 6.3.2, for a thorough discussion and morpho-syntactic and semantic arguments supporting our claim that the bracketed clauses in (6)a,b and in the Romanian examples in (2) and (3) are free relative clauses, rather than internally headed relative clauses or correlative clauses.

<sup>6</sup> Both *ce* and *care* can be used without NP complements in Romanian (much like English *what*). In these uses, they exhibit different properties from those characterizing their use as *wh*-determiners with an overt NP complement. In this paper, we focus exclusively on the latter use.

Before focusing on the puzzle, it's worth mentioning that Romanian does allow *care* 'which' + NP in certain non-interrogative *wh*-clauses, but, crucially, none of these constructions is a Max-FR nor an Ex-FR.<sup>7</sup> For instance, the bracketed clause in (8), taken from Brașoveanu (2008: 49), is a correlative clause.

- (8) [**Pe care om** l-a interogat Securitatea], în *acela*  
 ACC which person CL.ACC.3SG-has interrogated Security-the in DEM.3MSG  
 nu mai am încredere  
 not anymore have.1SG trust  
 'I don't trust any person interrogated by the secret police anymore'

It occurs left dislocated, while Max-FRs and Ex-FRs occur clause-internally, as arguments or adjuncts. Also, it licenses the italicized demonstrative pronoun in the matrix clause, which can be optional in subject position (Romanian is subject pro-drop). Last, the bracketed clause in (8) cannot be replaced by nor interpreted as a definite nor an indefinite DP. Its semantic contribution adds a universal or free choice component to the interpretation of the whole sentence (see Brașoveanu 2008). This interpretative option is never available for Max-FRs nor Ex-FRs.

The bracketed clause in (9), instead, is a kind of *wh*-clause originally discussed in Rudin (1986) and recently described in Caponigro & Fălăuș (2020).

- (9) A mâncat [**care copil** ce a vrut.]  
 has eaten which child what has wanted  
 'Every child ate whatever they wanted.'

It obligatorily involves a *wh*-clause with more than one *wh*-phrase, each of which is related to an argument of *both* the matrix and the embedded predicate. For instance, in (9) the children eating are the ones that want to eat and the things they eat are the things they want to eat. Crucially, *care* + NP can only be replaced by a universally quantified DP or the free choice *wh*-DP *oricare/orice copil* 'any child' in order to obtain a sentence that is roughly truth-conditional equivalent to the original one.<sup>8</sup>

---

<sup>7</sup> We thank our Reviewers for encouraging us to mention these other constructions that we had initially omitted for lack of space. We also thank our Editor for granting us extra space.

<sup>8</sup> The free choice determiners *oricare* and *orice* morphologically result from the combination of the free choice prefix *ori* with the *wh*-words *care* 'which' and *ce* 'what', respectively. They may therefore resemble *wh*-determiners *whichever* and *whatever* in English. Semantically, though, *oricare* and *orice* behave more like the free choice determiner *any* in English (Caponigro & Fălăuș 2018).

Our Reviewers find examples like those in (10)a,b acceptable, while our consultants do not.<sup>9</sup> Crucially, even our Reviewers seem to agree that the bracketed strings in (10)a,b cannot be interpreted like definite nor indefinite DPs, but necessarily trigger some kind of free choice inference, based on the translations they provided.

(10) a. [ **Care student** rezolvă primul problema] câștigă premiul.  
 which student solves first.the problem.the wins prize.the  
 'Whoever student solves the problem first wins the prize.'

b. Întoarce-te cu [ **care tren** vrei].  
 return-REFL with which train want.2SG  
 'Take any train you want (to go back).'

Taken together, these examples and the properties they exhibit reinforce the puzzle concerning the unacceptability of the examples in (4)–(5): why is *care* 'which' + NP the only *wh*-phrase to be ruled out from Max-FRs (non-interrogative *wh*-clauses interpreted like definite DPs) and Ex-FRs (non-interrogative *wh*-clauses interpreted as indefinite DPs), while it is productive in interrogative clauses and in other non-interrogative *wh*-clauses?<sup>10</sup>

The remainder of the squib is structured as follows. In Section 2, we show the discourse-anaphoric contrast between 'what' + NP and 'which' + NP in interrogative clauses in Romanian building on insights from the literature on D-linking and Dayal (2017, *in preparation*). In Section 3, we use the findings in Section 2 to suggest an account for the contrast in Max-FRs and Ex-FRs in (2)–(5). Section 4 concludes with some general remarks and outstanding issues.

<sup>9</sup> More precisely, seven out of our eight consultants rejected sentences (10)a,b and spontaneously suggested to replace *care* 'which' with *oricare* or *orice* 'any'. The eighth consultant judged them marginally acceptable.

<sup>10</sup> The only Max-FRs and Ex-FRs where *care* 'which' NP is possible (though still marginal for most of our consultants) are those where the *wh*-phrase needs to be in the oblique (dative/genitive) form, as the bold-faced forms in (i) and (ii).

(i) I-am scris [ **cărui profesor** i-ai scris și tu].  
 CL.DAT-have.1SG written which.DAT professor CL.DAT-have.2SG written also you  
 'I wrote to the (same) professor you wrote to.'

(ii) Nu avem [ **cărui coleg** să îi cerem ajutorul].  
 not have.1PL which colleague SUBJ CL.DAT.3SG ask.1PL help  
 'We don't have any colleague that we can ask for help.'

The *wh*-word *ce* 'what' lacks an oblique morphological form, while *care* 'which' inflects for case, person, number, and gender. Therefore, the forms in (i) and (ii) could be coerced into acting as suppletive forms of *ce*, rather than true instances of *care*, explaining why our consultants find them less degraded.

## 2. 'What' + NP versus 'which' + NP in interrogative clauses

We are not aware of any systematic comparative investigation of *ce* 'what' + NP vs. *care* 'which' + NP in interrogative clauses in Romanian. The literature on Romanian seems to imply a difference between the two *wh*-determiners that is usually cast in terms of D-linking (e.g. Pesetsky 1987, Dayal 2016). Specifically, it is assumed that *care*, but not *ce*, comes with a contextual salience/familiarity presupposition (see Comorovski 1996, Rațiu 2011, Giurgea 2013, Giurgea & Grosu 2019, a.o.).<sup>11</sup> Accordingly, *care* can only be used "when the domain of entities from which values are to be selected is contextually available" (Farkas 2013: 220).

In this section, we probe the D-linking/discourse-anaphoric properties of *care* + NP and *ce* + NP by applying some of the tests that Dayal (2017, *in preparation*) develops to distinguish between *which* + NP and the phrasal *what* (without an NP complement) in English. We conclude that *care* + NP carries a strong presuppositional requirement: the set denoted by its NP must have been explicitly introduced in the discourse in a way that makes it akin to the partitive construction 'which of the' + NP. On the other hand, *ce* + NP carries no presuppositional restriction whatsoever.

Let's start by considering a context in which the set denoted by the NP complement of *care* or *ce* is part of common knowledge, but it is not introduced explicitly, like the one in (11). Both Ana and Maria know that every year there is a list of nominees for the Nobel prize. Still this shared knowledge is not enough to license the use of *care* in the absence of some explicit overt introduction of the set of relevant authors in the discourse. The partitive construction *care dintre* 'which of' + NP patterns alike in being infelicitous in this context; *ce* + NP, instead, is fully felicitous.

- (11) CONTEXT: *Ana just said that the Nobel prize for literature was finally announced. The list of nominees has not been introduced in the discourse. Maria asks:*

{ #Care autor } / { Ce autor } / { #Care dintre autori } a câștigat  
 which author / what author which of authors has won  
 premiul Nobel anul acesta?  
 prize-the Nobel year-the this  
 '{Which author}' / '{What author}' / '{Which of the authors}' won the Nobel prize this  
 year?'

<sup>11</sup> In the case of direct objects, D-linking in Romanian is also reflected by clitic doubling: while *ce* + NP never gets clitic doubled, *care* + NP is obligatorily clitic doubled (Comorovski 1996, Giurgea & Grosu 2019). Since this difference is orthogonal to our discussion, we set it aside in the rest of the paper.

Let's now consider Context 1 in (12), in which the set denoted by the NP complement of *care* or *ce* is not mentioned explicitly but it can be inferred from other pieces of information that are provided. Ana's mentioning her trip to the nearby bookstore to buy a book for Lia makes it clear that book is a member of the set of books in that bookstore. Still this is not enough to license the use of *care* + NP or the closely related partitive construction, whereas *ce* + NP is fully acceptable, as shown in (12)a–c. If we slightly change the scenario as in Context 2 in (12), and have Ana explicitly mention a list of books Maria is familiar with, then *care* + NP and the partitive construction join *ce* + NP in being felicitous, as shown in (12)a–c.<sup>12</sup>

(12) CONTEXT 1: *Ana says: "I went to the nearby bookstore and got a book for Lia".  
Maria asks:*

CONTEXT 2: *Ana says: "I went to the nearby bookstore and got a book for Lia  
out of the list you recommended". Maria asks:*

- |                                       |                             |                  |
|---------------------------------------|-----------------------------|------------------|
| a. <b>Care</b> <b>carte</b> ai        | luat-o?                     | # (in Context 1) |
| which book have.2SG                   | taken-CL.3SG                | √ (in Context 2) |
| 'Which book did you get?'             |                             |                  |
| b. <b>Ce</b> <b>carte</b> ai          | luat?                       | √ (in Context 1) |
| what book have.2SG                    | taken                       | √ (in Context 2) |
| 'What book did you get?'              |                             |                  |
| c. <b>Care</b> <b>dintre cărți</b> ai | luat-o?                     | # (in Context 1) |
| which of                              | books have.2SG taken-CL.3SG | √ (in Context 2) |
| 'Which of the books did you get?'     |                             |                  |

The same pragmatic constraint holds for embedded interrogative *wh*-clauses that are introduced by *care* + NP. In the example in (1), *care* + NP can be used felicitously only if the discourse makes it clear that Ana has a certain set of books in mind and the speaker is aware of that. On the other hand, *ce* + NP doesn't exhibit this restriction and can be used regardless of the availability of a contextually salient/familiar set.

To sum up, we have shown that *care* + NP in interrogative clauses in Romanian is subject to very strict discourse conditions: the set denoted by its NP is presupposed to have been previously introduced in the discourse explicitly.<sup>13</sup>

<sup>12</sup> Neither (12)a nor (12)c would be acceptable if Ana explicitly mentioned a list of books that Maria is not familiar with.

<sup>13</sup> A possible source for this presuppositional component may be the fact that in old Romanian *care* is the only *wh*-word that incorporates the enclitic definite determiner—*carele* (M.SG), *carii* (M.PL), *carea* (F.SG), *carele* (F.PL) (Pană Dindelegan 2016).

If we assume Karttunen's (1977) semantics for interrogative *wh*-clauses, we can extend his treatment of *wh*-words as existential generalized quantifiers to *ce* + NP and *care* + NP as in (13) and (14), respectively.

$$(13) \textit{ce}_{\text{INTERR}} \textit{'what'} \rightsquigarrow \lambda P \lambda Q \exists y [P(y) \wedge Q(y)]$$

$$(14) \textit{care}_{\text{INTERR}} \textit{'which'} \rightsquigarrow \lambda P \lambda Q: \mathbf{P} \in \mathbf{D}_{\text{REF}}. \exists y [P(y) \wedge Q(y)]$$

While *ce* semantically behaves like a plain indefinite determiner, *care* introduces the presuppositional requirement that the set P—the denotation of its NP complement—be already a member of the set of discourse referents.<sup>14</sup>

The reader may have noticed that (some of) the contexts above that do not license *care* + NP in Romanian may, instead, be appropriate for *which* + NP in English (e.g., Context 1 in (12)). Similarly, the contexts licensing *ce* + NP may not be exactly the same as those licensing *what* + NP (e.g., Context 2 in (9)). These possible crosslinguistic differences deserve further investigation. Here we just want to raise the general issue that, even if the repertoire of complex *wh*-phrases in a language includes some form of 'what' + NP and 'which' + NP, it should not be taken for granted that the way these two complex *wh*-phrases are sensitive to discourse conditions and the way they differ from each other in this respect are necessarily identical across languages. Consequently, they should be carefully investigated in each language.

### 3. Back to the puzzle

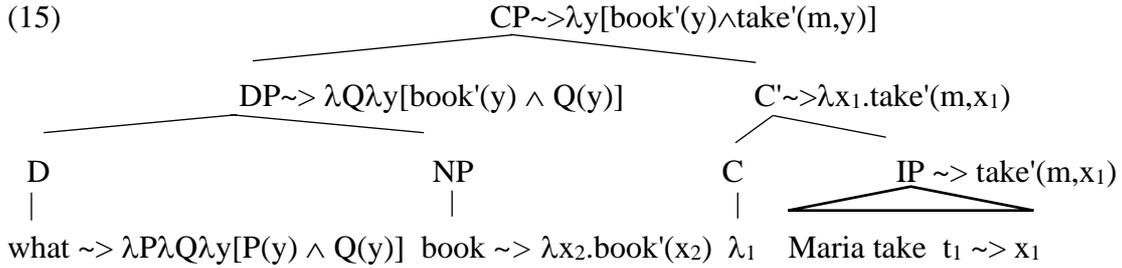
Let's now return to the puzzle in (2)–(5) we started with. In this section, we sketch a solution in light of the observations and conclusion in Sections 1 and 2. We first provide an explicit semantics for 'what' + NP in Max-FRs and Ex-FRs in Romanian and then speculate on why a similar move does not extend to 'which' + NP.

As mentioned in Section 1, Max-FRs behave like definite DPs semantically, while Ex-FRs resemble narrow scope indefinites (see fn. 1 and fn. 2). Still, we argue they share most of their semantic derivation, including the meaning of their *wh*-phrases. Let us look at the concrete example in (15), showing the semantic derivation of the CP of the Max-FR in (1) and the Ex-FR in (2) with lexical entries in English and the verbal form 'take' standing for both the indicative form in the Max-FR and the subjunctive form in the Ex-FR in the Romanian original clauses.

---

<sup>14</sup> The logical translation in (14) assumes that a set and its intensional counterpart (a property) can act as referents in the discourse (see Nouwen 2003). It is also possible to reformulate the presuppositional content in (14) without such an assumption, as shown in (i), with the referent being the maximal individual of the set P rather than P itself.

(i)  $\textit{care}_{\text{INTERR}} \rightsquigarrow \lambda P \lambda Q: \iota x \mathbf{P}(x) \in \mathbf{D}_{\text{REF}}. \exists y [P(y) \wedge Q(y)]$



The logical translation of C' in (15) denotes the set of objects that Maria took or can take. This is the standard semantic contribution of a *wh*-clause in general right before combining with its *wh*-phrase. The logical translation of the NP is standard as well, denoting a set of individuals. The crucial new assumption is about the meaning of the *wh*-determiner *ce* 'what' in a free relative clause (henceforth, FR), regardless whether it is a Max-FR or an Ex-FR. The logical translation of the determiner *ce* in an FR (henceforth,  $ce_{FR}$ ) is repeated in (16), while the resulting translation for the whole *wh*-phrase is given in its general form in (17), with NP' standing for the logical translation of the NP complement of *ce*.

$$(16) \quad ce_{FR} \text{ 'what' } \sim \lambda P \lambda Q \lambda y [P(y) \wedge Q(y)]$$

$$(17) \quad [DP \ ce_{FR} + NP] \sim \lambda Q \lambda y [NP'(y) \wedge Q(y)]$$

$ce_{FR}$  semantically behaves like a purely functional/logical element: it applies to two sets of individuals to return their intersection. Once it combines with its NP complement, it turns into a set restrictor: it applies to a set of individuals Q to return the subset of Q containing those members that are in the denotation of the NP complement as well. In (15),  $ce_{FR}$  combines with 'book' to return the function that restricts whatever set it applies to next to its subset containing books. When the *wh*-DP combines with the C' denoting the set of individuals that Maria took or can take, it returns the subset of those individuals that are books, as shown by the logical translation in CP in (15). This set is existentially closed by the matrix existential predicate in an Ex-FRs, while it is type-shifted into its maximal individual in a Max-FR.

If we compare the semantic contribution of  $ce_{FR}$  in FRs in (16) to the one of  $ce_{INTERR}$  in interrogative *wh*-clauses in (13), we can see that both *wh*-words apply to two sets and intersect them.  $ce_{FR}$  doesn't do anything else semantically, while  $ce_{INTERR}$  existentially quantifies over the intersection. In other words,  $ce_{INTERR}$  semantically behaves like a plain existential determiner (e.g., *a* in English or *un(M)/o(F)* in Romanian). This meaning change from existential quantification to  $\lambda$ -abstraction/set-formation applies not just to  $ce_{FR}$ , but to all the *wh*-words that can occur in FRs. It is possible to provide a semantics for interrogative *wh*-clauses that assumes the same set-restrictor meaning for *wh*-words as the one in FRs, as sketched in Caponigro (2003: Ch.

6) and discussed extensively in George (2011). This approach would account even more easily for the fact that all *wh*-words in interrogative clauses can occur in FRs as well in Romanian, with the exception of *care*, to which we now turn.

If it is so natural for *wh*-words in Romanian to occur in FRs, what prevents *care* from joining all its *wh*-colleagues in FRs? In other words, why doesn't Romanian allow for a *care*<sub>FR</sub> whose logical translation would look like the one in (18)?

(18) \**care*<sub>FR</sub> 'which'  $\sim \lambda P \lambda Q \lambda y: \mathbf{P} \in \mathbf{D}_{\text{REF}}. [P(y) \wedge Q(y)]$

(18) contains exactly the same presuppositional discourse-requirement as *care*<sub>INTERR</sub> in (14),<sup>15</sup> minus the existential quantifier, i.e., the meaning change we have argued to take place with all other *wh*-words in Romanian when they occur in an FR. But *care*<sub>FR</sub> is not an actual lexical item of Romanian—this is why it is starred in (18).

We do not have an account for this puzzle—this is why we are writing a squib about it rather than a full paper. Still, we can offer some observations and speculations that we hope will help in advancing towards a solution. The lack of *care*<sub>FR</sub> may be due to a clash between its presuppositional requirement and the semantic function of an FR. Specifically, an FR is a device that natural language can make use of to define a set by means of the informational richness of a clause without relying on a nominal head, as in headed relative clauses. Similarly to the FR in (15), the FR *what Maria took* ends up denoting the set of all and only the things that Maria took by a purely compositional semantic mechanism, rather than relying on the speaker's and listener's familiarity with it. On the other hand, the core property of a discourse-anaphoric element is to refer back to a referent that has already been introduced in the discourse: the personal pronoun *he* refers back to a male individual, while the demonstrative pronoun *that* refers back to a non-human object. Both pronouns can introduce a light-headed relative clause in English. In this use, they lose their anaphoric properties. Neither *he* in (19) nor *that* in (20) triggers any reference to any previously mentioned individual.

(19) **He** who laughs last, laughs best.

---

<sup>15</sup> A Reviewer suggested to extend the presuppositional requirement in (18) to the set the variable Q refers to as well, with (18) turning into (i):

(i) \**care*<sub>FR</sub>  $\sim \lambda P \lambda Q \lambda y: \mathbf{P}, \mathbf{Q} \in \mathbf{D}_{\text{REF}}. [P(y) \wedge Q(y)]$

Our formulation of (18) aims at keeping the meaning of *care*<sub>FR</sub> as minimally different from *care*<sub>INTERR</sub> as possible, in the same way as we did with *ce*<sub>FR</sub> and *ce*<sub>INTERR</sub>. The logical translation in (i), instead, increases the difference in meaning between *care*<sub>FR</sub> and *care*<sub>INTERR</sub> by requiring the set introduced by the remainder of the FR to be presupposed as well. Since *care*<sub>FR</sub> is not an actual lexical item in Romanian, it is not possible to test the presuppositional differences between its possible meaning in (18) and the one in (i). Therefore, we maintain our formulation in (18) for simplicity, since we believe it does not significantly affect the solution we are hinting at.

(20) A very small part of acting is **that** which takes place on the stage!<sup>16</sup>

Notice that the option of losing anaphoric features depends on the individual lexical item. It is permitted with *he*, although sounding archaic or proverb like, but it is completely banned with a personal pronoun like *it* (cf. *\*it that* or *\*it which*). It is fully productive with *that which* (or *those who*), while it is not possible with *this* or *these*.<sup>17</sup> What this pattern suggests is that *care* in Romanian would need to lose its anaphoric features to be able to occur in a non-interrogative clause. This is indeed the case for the use of *care* in non-interrogative *wh*-clauses other than Max-FRs and Ex-FRs that we touch on towards the end of Section 1: none of them exhibit any dependencies on existing discourse referents. If the *wh*-determiner *care* had undergone a similar meaning change in order to be licensed in Max-FRs and FRs, it would have become fully identical to *ce<sub>FR</sub>* 'what' + NP. Some form of blocking may disfavor a lexical item from assuming exactly the same meaning as another lexical item with exactly the same distribution.<sup>18</sup> Also, *care* would have to change its meaning along two dimensions (loss of existential quantification and loss of anaphoricity) in order to be able to occur in Max-FRs and Ex-FRs, while *ce* loses its existential strength only. Although neither ban is absolute, the combination of both may be harder to overcome and explains the lack of *care<sub>FR</sub>*.

Ex-FRs may add an extra reason to ban *care<sub>FR</sub>* on top of the two we just discussed. Remember that Ex-FRs very naturally occur as the complement of the existential predicate *a avea* 'to have'. It looks like *care<sub>INTERR</sub>* cannot be based generated in this syntactic position either, as shown in (21)a. On the other hand, *ce<sub>INTERR</sub>* exhibits no such restriction while triggering a 'kind' interpretation, as shown by the acceptability of the interrogative clause in (21)b and the 'kind' answer it licenses.

---

<sup>16</sup> T.S. Eliot, *The Sacred Wood*.

<sup>17</sup> Cf. *Those who laugh last, laugh best* vs. *\*These who laugh last, laugh best*; also, cf. *\*A very small part of acting is this which takes place on the stage!*

<sup>18</sup> See Embick & Marantz (2008) and references therein for the general notion of blocking and Chierchia (1998) for blocking applied to semantic rules and type shifting.

- (21) a. \*[ **Care prieteni**] îi ai \_\_?<sup>19</sup>  
 which friends CL.3PL have.2SG  
 ('Which friends do you have?')
- b. [ **Ce prieteni**] ai \_\_?  
 what friends have.2SG  
 'What kind of friends do you have?'  
 (ANSWER 1: × 'The friends you met yesterday.'; ANSWER 2: ✓ 'Friends that I can fully trust.')

Whatever reason may be responsible for this ban of *care*<sub>INTERR</sub>, it may further disfavor the meaning switch to *care*<sub>FR</sub>.

#### 4. Conclusions

We have presented a previously unnoticed and puzzling contrast between the unconstrained productivity of 'what' + NP in Max-FRs and Ex-FRs and the complete ban on 'which' + NP from the same two constructions in Romanian—a language in which both constructions are highly productive and can be introduced by all the other *wh*-words and *wh*-phrases that are attested in interrogative clauses. We have also shown that 'which' + NP in interrogative clauses in Romanian clearly exhibit strong discourse-anaphoric requirements, unlike 'what' + NP. We have connected the two findings by suggesting that discourse anaphoricity and set formation out of a complex expression may not be compatible. Lexical blocking and further restrictions on the complement of existential predicates may play a role as well.

The constraints we have invoked (and their interplay) are neither absolute nor exhaustive. Some languages may apply further constraints and ban both 'what' + NP and 'which' + NP from introducing FRs. This would be the case for Italian, as mentioned earlier. Others may ban 'which' + NP from FRs and allow 'what' + NP in FRs, but with restrictions, rather than as freely as in Romanian. This would be the case for English, if Andrews' (1975) generalizations hold under further scrutiny: 'what' +

---

<sup>19</sup> Our consultants find (i) acceptable, with *care* obligatorily preceded by the preposition *de* 'of'. As made clear by the translation of the matrix interrogative *wh*-clause and by its only possible answer, *care* in this use only allows for a 'kind' interpretation without any discourse anaphoricity. This is the same semantic behavior as the interrogative clause with *ce* in (21)b.

(i) **De care prieteni** ai ?  
 of which friends have.2SG  
 'What kind of friends do you have?'  
 (ANSWER 1: × 'The friends you met yesterday.'; ANSWER 2: ✓ 'Friends that I can fully trust.')

NP in Max-FRs in English would apply to two sets to return not their intersection, but rather a kind—the kind that is instantiated by the members of the intersection of those two sets.<sup>20</sup> Because these constraints are “soft”, there may be languages that overcome the ban on complex anaphoric descriptions and/or semantic lexical blocking and allow for ‘which’ + NP in FRs. We have no clear evidence of any language of this kind yet.<sup>21</sup>

We hope that our squib will encourage further language-specific and crosslinguistic investigation on FRs introduced by ‘what’ + NP and ‘which’ + NP, with special attention to the semantic/pragmatic properties that these *wh*-phrases exhibit in interrogative clauses and the way these properties transfer across *wh*-constructions.

### Acknowledgements

Thanks to our Romanian consultants and to Radek Šimík, two anonymous reviewers, Gennaro Chierchia, Donka Farkas, Veneeta Dayal, Daniel B. Kane, and Carson Schütze. We the authors are solely responsible for all the remaining mistakes.

### References

- Andrews, Avery D. 1975. Studies in the syntax of relative and comparative clauses. Doctoral dissertation, MIT.
- Brașoveanu, Adrian. 2008. Uniqueness effects in correlatives. In *SuB 12: Proceedings of the 12<sup>th</sup> Sinn und Bedeutung*, ed. by Alte Grøn, 47–65.  
<https://ojs.ub.uni-konstanz.de/sub/index.php/sub/issue/view/17>
- Caponigro, Ivano. 2003. Free not to ask: on the semantics of free relatives and *wh*-words cross-linguistically. Doctoral dissertation, University of California, Los Angeles.
- Caponigro, Ivano. 2004. The semantic contribution of *wh*-words and type shifts: evidence from free relatives cross-linguistically. In *SALT 14: Proceedings of the 14<sup>th</sup> Semantics and Linguistic Theory*, ed. by Robert B. Young, 38–55. CLC Publications. <https://doi.org/10.3765/salt.v14i0.2906>
- Caponigro, Ivano. 2019. In defense of what(ever) free relative clauses they dismiss: A reply to Donati and Cecchetto (2011). *Linguistic Inquiry* 50(2):356–371.

---

<sup>20</sup> Heim (1987) suggests that *what* + NP in interrogative clauses in English licenses a variable ranging over subkinds of the kind denoted by its NP complement, while *which* + NP is inherently partitive and licenses a variable ranging over individuals.

<sup>21</sup> Royer (2021) reports that in Chuj (Mayan) Max-FRs can be introduced by both ‘what’ + NP and ‘which’ + NP (Royer 2021: ex. 49 and 50); the only example with ‘which’ seems to be heavily discourse dependent, based on its translation, but it lacks an overt NP complement.

- Caponigro, Ivano, and Anamaria Fălăuș. 2018. Free choice free relatives in Italian and Romanian. *Natural Language and Linguistic Theory* 36(2):323–363.  
<https://doi.org/10.1007/s11049-017-9375-y>
- Caponigro, Ivano, and Anamaria Fălăuș. 2020. Unveiling Multiple *wh*- Free Relative Clauses and their functional *wh*-words. *Semantics and Pragmatics* 13.  
<http://dx.doi.org/10.3765/sp.13.2>
- Caponigro, Ivano, Harold Torrence, and Roberto Zavala Maldonado, eds. 2021. *Headless relative clauses in Mesoamerican languages*. New York: Oxford University Press.
- Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6:339–405.
- Comorovski, Ileana. 1996. *Interrogative phrases and the syntax–semantics interface*. Dordrecht: Kluwer. <https://doi.org/10.1007/978-94-015-8688-7>
- Dayal, Veneeta. 1996. *Locality in WH quantification: questions and relative clauses in Hindi*. Dordrecht: Kluwer. <https://doi.org/10.1007/978-94-011-4808-5>
- Dayal, Veneeta. 2016. *Questions*. Oxford: Oxford University Press.  
<https://doi.org/10.1093/acprof:oso/9780199281268.001.0001>
- Dayal, Veneeta. 2017. Deconstructing the D in D-linking. Paper presented at *David Pesetsky@60, Panel on questions*, February 11, MIT.
- Dayal, Veneeta. in preparation. *WH-scope taking in questions*, LI Monographs, Cambridge, MA: MIT Press.
- Donati, Caterina, and Carlo Cecchetto. 2011. Relabeling heads: A unified account of relativization structures. *Linguistic Inquiry* 42:519–560.
- Embick, David, and Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39:1–53
- Farkas, Donka. 2013. The semantics of determiners. In *A reference grammar of Romanian*, ed. by Carmen Dobrovie-Sorin and Ion Giurgea, 175–230. Amsterdam: John Benjamins.
- George, B. R. 2011. Question embedding and the semantics of answers. Doctoral dissertation, University of California, Los Angeles.
- Giurgea, Ion. 2013. The syntax of determiners and other functional categories. In *A reference grammar of Romanian*, ed. by Carmen Dobrovie-Sorin and Ion Giurgea, 97–174. Amsterdam: John Benjamins.
- Giurgea, Ion, and Alexander Grosu. 2019. Interrogatives and *wh*-movement. Ms. University of Bucharest and Tel Aviv University.
- Grosu, Alexander. 2004. The syntax-semantics of Modal Existential *wh* Constructions, in *Balkan syntax and semantics*, ed. by Olga Mišeska Tomić, 405–438. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.67.21gro>

- Grosu, Alexander. 2013. Relative clause constructions and unbounded dependencies. In *A reference grammar of Romanian*, ed. by Carmen Dobrovie-Sorin and Ion Giurgea, 597–662. Amsterdam: John Benjamins.
- Heim, Irene. 1987. Where does the definiteness restriction apply? Evidence from the definiteness of variables. In *The representation of (in)definiteness*, ed. by Eric Reuland and Alice ter Meulen, 21–42. Cambridge, MA: MIT Press.
- Izvorski, Roumyana. 1998. Non-indicative wh-complements of possessive and existential predicates. In *NELS 28: Proceedings of the 28th Annual Meeting of the North East Linguistic Society*, ed. by Pius N. Tamanji and Kiyomi Kusumoto, 159–173, Amherst, MA: GLSA Publications.
- Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In *Quantification in natural languages*, ed. by Emmon Bach, Eloise Jelinek, Angelika Kratzer and Barbara Partee, 451–486. Dordrecht: Kluwer.  
[https://doi.org/10.1007/978-94-017-2817-1\\_15](https://doi.org/10.1007/978-94-017-2817-1_15)
- Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1(1):3–44. <https://doi.org/10.1007/bf00351935>
- Link, Godehard. 1983. The logical analysis of plurals and mass terms: a lattice-theoretical approach. In *Meaning, use and the interpretation of language*, ed. by Rainer Bauerle, Christoph Schwarze and Arnim von Stechow, 303–323. Berlin: de Gruyter. <https://doi.org/10.1515/9783110852820.302>
- Nouwen, Rick. 2003. Plural pronominal anaphora in context: dynamic aspects of quantification. Utrecht: LOT.  
[https://www.lotpublications.nl/Documents/084\\_fulltext.pdf](https://www.lotpublications.nl/Documents/084_fulltext.pdf)
- Pană Dindelegan, Gabriela, ed. 2016. *The syntax of Old Romanian*, 463–561. Oxford: Oxford University Press.  
<https://doi.org/10.1093/acprof:oso/9780198712350.001.0001>
- Pesetsky, David. 1987. Wh-in-situ: movement and unselective binding. In *The representation of (in)definiteness*, ed. by Eric Reuland and Alice ter Meulen, 98–129. Cambridge, MA: MIT Press.
- Rațiu, Dafina. 2011. De la syntaxe à la sémantique des propositions interrogatives. Étude comparative des questions multiples en roumain. Doctoral dissertation, University of Nantes.
- Royer, Justin. 2021. Headless relative clauses in Chuj. In *Headless relative clauses in Mesoamerican languages*, ed. by Ivano Caponigro, Harold Torrence, and Roberto Zavala Maldonado, 327–361. New York: Oxford University Press.
- Sharvy, Richard. 1980. A more general theory of definite descriptions. *The Philosophical Review* 89(4):607–624. <https://doi.org/10.2307/2184738>

- Šimík, Radek. 2011. Modal existential wh-constructions Doctoral dissertation, University of Groningen.
- Šimík, Radek. 2013. Modal existential *wh*-constructions as affordance descriptions. In *SuB 17: Proceedings of the 17<sup>th</sup> Sinn und Bedeutung*, ed. by Emmanuel Chemla, Vincent Homer, and Grégoire Winterstein, 563–580.  
<http://semanticsarchive.net/Archive/Dk3NGEwY/Simik.pdf>
- Šimík, Radek. 2017. Existential wh-constructions. In *Oxford bibliographies in linguistics*, ed. by Mark Aronoff. New York: Oxford University Press.  
<https://doi.org/10.1093/obo/9780199772810-0162>
- Šimík, Radek. 2020. Free relatives. In *The Wiley Blackwell companion to semantics*, ed. by Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann, and Thomas Ede Zimmermann, 1–38. Hoboken, NJ: Wiley.  
<https://doi.org/10.1002/9781118788516.sem093>
- de Vries, Mark. 2002. *The syntax of relativization*. Utrecht: LOT.  
[https://www.lotpublications.nl/Documents/053\\_fulltext.pdf](https://www.lotpublications.nl/Documents/053_fulltext.pdf)