The propositional nature of non-interrogative *wh*-clauses and their *wh*-words

IVANO CAPONIGRO

University of California San Diego
Linguistics Department
La Jolla, CA 92093-0108
U.S.A.

ivano@ucsd.edu

Draft, 2026/3/31

Abstract

There are at least three major crosslinguistic generalizations concerning interrogative and non-interrogative *wh*-clauses as well as their characterizing restricted class of lexical items—*wh*-words. They are supported by findings that have emerged in the past couple of decades, although further data and analysis are needed, especially for non-interrogative *wh*-clauses. To the best of my knowledge, these generalizations have not yet been fully stated together, let alone accounted for. I argue that a principled explanation of these generalizations requires an analysis of all non-interrogative *wh*-clauses as derived from interrogative *wh*-clauses and therefore sharing a propositional core. This would be the semantic correlate of the shared morphosyntactic core characterizing all *wh*-clauses. In turn, all *wh*-words in *wh*-clauses must be analyzed as logical operators that apply to propositional content. To the best of my knowledge, such a general analysis has not yet been attempted. I formulate and discuss those three generalizations and account for them by developing a "propositional" analysis for each variety of *wh*-clause and its *wh*-words.

Keywords *wh*-clauses – *wh*-words – interrogative clauses – relative clauses – headless relative clauses – free relative clauses – correlative clauses

Table of Contents

1. Introduction ……3
2. Varieties of *wh*-clauses within and across languages ……...4
2.1. Interrogative *wh*-clauses ……4
2.2. Varieties of non-interrogative *wh*-clauses within and across languages ……5
3. Generalizations about *wh*-clauses and their *wh*-words ……6
4. The semantics of wh-clauses and wh-word meaning shifts ……10
4.1. Shared syntactic and semantic core of all *wh*-clauses……10
4.2. Semantic analyses of *wh*-clauses with a single *wh*-phrase……....12
4.2.1 Semantic analysis of embedded interrogative *wh*-clauses (INTs) ……12
4.2.2. Semantic analysis of referential free relative clauses (R-FRs) ……14
4.2.3. Semantic analysis of existential free relative clauses (∃-FRs) ……16
4.2.4. Semantic analysis of correlative clauses (CORs) ……17
4.3. Semantic analyses of *wh*-clauses with multiple *wh*-phrases……19
4.3.1. Semantic analysis of referential free relative clauses with multiple *wh*-phrases 20
4.3.2. Semantic analysis of existential free relative clauses with multiple *wh*-phrases 22
4.3.3. Semantic analysis of interrogative clauses with multiple *wh*-phrases ……23
4.3.4. Semantic analysis of correlative clauses with multiple *wh*-phrases…….....28
5. Varieties of wh-words and their logically related meanings ……31
6. Accounting for the three generalizations ……..34
7. Problems with possible alternative approaches ……35
8. Conclusions……38
Acknowledgements ……..41
References ……41

1. Introduction

There are at least three major crosslinguistic generalizations concerning interrogative and non-interrogative *wh*-clauses as well as their characterizing restricted class of lexical items—*wh*-words.[1] They are supported by findings that have emerged in the past couple of decades, although further data and analysis are needed, especially for non-interrogative *wh*-clauses. To the best of my knowledge, these generalizations have not yet been fully stated together, let alone accounted for. I argue that a principled explanation of these generalizations requires an analysis of all non-interrogative *wh*-clauses as derived from interrogative *wh*-clauses and therefore sharing a propositional core. This would be the semantic correlate of the shared morphosyntactic core characterizing all *wh*-clauses. In turn, all *wh*-words in *wh*-clauses must be analyzed as logical operators that apply to propositional content. To the best of my knowledge, such a general analysis has not yet been attempted. I formulate and discuss those three generalizations and account for them by developing a "propositional" analysis for each variety of *wh*-clause and its *wh*-words.

Whenever attention has been paid to non-interrogative *wh*-clauses (headed relative clauses, free relative clauses, correlative clauses, etc.), the focus has typically been on a single variety, considered independently of other non-interrogative varieties and of interrogative *wh*-clauses. To my knowledge, no unified semantic analysis of all *wh*-clauses has been proposed, let alone one that takes the propositional meaning of *wh*-clauses as primary and as the core building block for the meaning of all other *wh*-clauses.[2] This is the central goal of this paper. I argue that all kinds of *wh*-clauses denote a set of (open) propositions at the same point in their semantic derivation and that differences among varieties of *wh*-clause arise from distinct—but logically related—lexical meanings assigned to their *wh*-words.

The paper proceeds as follows. Section 2 introduces interrogative *wh*-clauses and varieties of non-interrogative *wh*-clauses attested crosslinguistically. Section 3 presents three generalizations concerning the distribution of *wh*-clauses and their *wh*-words within and across languages. Section 4 develops the semantic analysis, beginning with the shared propositional core and then adapting it to each variety of *wh*-clause through systematic logical changes in the meanings of their *wh*-words. Section 5 summarizes, discusses, and compares these logical meaning changes. Section 6 shows how the overall proposal accounts for the empirical facts captured by the three generalizations. Section 7 considers alternative approaches and the challenges they face in explaining these generalizations. Section 8 concludes with broader remarks on the two emerging

[1] Throughout the paper, the term "*wh*-word" refers to the specific lexical item that carries *wh*- marking or an equivalent marker in languages other than English. For instance, *who*, *what*, *which*, and *how much* are *wh*-words. *Wh*-words can (i) form their own phrase, like *who* in *Who left?*, or (ii) be embedded in larger phrases, as in *which books, because of what*, *in the company of who(m)*, *how much expensive food*, and so on. I use the term "*wh*-phrase" to cover both (i) and (ii). Therefore, *who* in English is both a *wh*-word and a *wh*-phrase. By contrast, *which* in *which book* is a *wh*-word but not a *wh*-phrase, while *which book* is a *wh*-phrase but not a *wh*-word. Unless otherwise noted, "*wh*-phrase" will be used as the default term.

[2] Šimík (2026) proposes a unified morphosyntactic and semantic analysis for some of the *wh*-clauses I discuss. I return to this work in § 7.

strategies that languages employ when forming clauses with missing constituents, with or without *wh*-words.

2. Varieties of *wh*-clauses within and across languages

In this section, I introduce the main characters of my investigation—varieties of *wh*-clauses—starting with interrogative *wh*-clauses (§ 2.1) and then turning to non-interrogative *wh*-clauses (§ 2.2).

2.1. Interrogative *wh*-clauses

Wh-clauses and the *wh*-words that characterize them are a well-known morphosyntactic device widely attested across languages for conveying open questions, i.e., requests for a certain kind of information.[3] For instance, the matrix *wh*-clause in (1)a conveys an "open" question that requires either a propositional answer, like the one conveyed by the matrix declarative clause in (1)b, or a short answer, like the one conveyed by the nominal in (1)c.

1) a. What did Andrea cook?
   b. Andrea cooked the lasagna.
   c. The lasagna.

English and many other languages allow for *wh*-clauses like the one in (1)a to be embedded as the complements of a restricted class of predicates, like *know* in (2)a.[4] In this configuration, the entailment from (2)a to (2)b is licensed, with (2)b replacing the *wh*-clause with a declarative clause. In other words, knowing the answer to the question of what Andrea cooked in a world and at time when Andrea cooked lasagna entails knowing that Andrea cooked that specific food in that world at that time. Notice that the knowledge entailed by (2)a is genuinely propositional. This is shown by the unacceptability of (2)c, where the complement of *know* is just the short answer in (1)c.

2) a. Luca knew [what Andrea cooked].
   b. Luca knew [that Andrea cooked the lasagna].
   c. *Luca knew the lasagna.

To sum up, most languages use matrix *wh*-clauses to convey the meaning of an open question, understood as a set of propositions (Hamblin 1973). Many of those languages also allow for the embedding of those propositional *wh*-clauses as the arguments of a restricted class of predicates (*know, tell, remember, ask, wonder*, etc) (Karttunen 1977).

While these facts have been extensively discussed in both the descriptive/typological and the more theoretically-oriented literature, less attention has been paid to the fact that a subset of the

[3] See Abaza (Northwest Caucasian) for a language that doesn't use *wh*-clauses nor *wh*-words to convey question meaning (Arkadiev and Caponigro 2021). Also, I use "question" to refer to a semantic object (i.e., a set of propositions) that is conveyed by an "interrogative clause", a linguistic object.

[4] See Adyghe (Northwest Caucasian) for a language that uses matrix interrogative *wh*-clauses to convey question meaning but doesn't allow for INTs to be embedded nor does it employ *wh*-words to convey the meaning conveyed by embedded interrogative clauses in a language like English (Caponigro and Polinsky 2011).

languages that use *wh*-clauses to convey question meaning also use *wh*-clauses with a very different distribution to convey non-propositional meaning.[5] I adopt the term **'interrogative *wh*-clause' (INT)** to refer to a *wh*-clause that conveys the meaning of a question and **'non-interrogative *wh*-clause'** (**Non-INT**) for a *wh*-clause that conveys non-propositional meaning. These include headed relative clauses, free relative clauses, correlative clauses and others that refer to individuals, or quantify over individuals, or denote sets of individuals. I introduce each variety in § 2.2.

2.2. Varieties of non-interrogative *wh*-clauses within and across languages

In this section, I briefly introduce three varieties of Non-INTs and highlight their morphosyntactic similarity with INTs and with each other, as well as their semantic differences.

The bracketed Non-INT in (3)a is known as a **'free relative clause' (FR)**. It looks identical to the bracketed embedded INT in (2)a, although semantically behaves like the definite DP in (3)b: both bracketed strings in (3) refer to a specific entity in the world—Andrea's food.

3) a. Luca tasted [$_{\text{R-FR}}$ what Andrea cooked].
   b. Luca tasted [$_{\text{Definite DP}}$ the stuff Andrea cooked].

Since the FR in (3)a semantically behaves like a definite, I call it a **'referential free relative clause' (R-FR)** to distinguish it from the FR I introduce next.

Italian can assign three different meanings to what looks like the same *wh*-clause morphosyntactically, occurring in the same syntactic position (complement). The difference depends on the matrix predicates, as shown in (4)a–c. In (4)a, the *wh*-clause occurs in the complement position of existential 'have' and is interpreted as a narrow-scope indefinite nominal involving existential quantification. This is why I label it **'existential free relative clause' (∃-FR)**. In (4)b, the *wh*-clause occurs in the complement position of an ordinary transitive predicate that selects for entity-denoting arguments, and is interpreted as a referential expression; it's therefore an R-FR. In (4)c, the *wh*-clause occurs in the complement position of a predicate selecting for a declarative or interrogative clause as its complement and is interpreted as denoting (the answer to) a question; it's therefore an INT.

4) a. Luca ha [$_{\exists\text{-FR}}$ chi gli cura i bambini].
   Luca has who to_him takes_care_of the children
   'Luca has a person/people who take(s) care of his children.'

[5] *Wh*-clauses can also be used to realize exclamative clauses across languages (e.g., *What delicious food Andrea cooked!* in English). I do not include exclamative *wh*-clauses among Non-INTs and therefore don't discuss them further, since they more closely resemble interrogative *wh*-clauses along at least two dimensions. First, propositional analyses of exclamative *wh*-clauses have been proposed that closely parallel those of INTs (see Zanuttini and Portner 2003). Second, exclamative *wh*-clauses are primarily matrix clauses, although they may be embedded under a restricted class of predicates.

b. Luca conosce bene [R-FR chi gli cura i bambini] .
Luca is_aquainted_with well who to_him takes_care_of the children
'Luca is well acquainted with the person/people who take(s) care of his children.'

c. Luca mi ha detto [INT chi gli cura i bambini].
Luca to_me has said who to_him takes_care_of
'Luca told me who takes care of his children.'

Romanian—another Romance language—is even more productive with Non-INTs. In addition to the varieties already discussed, it also allows for the *wh*-clause in (5)a, which is known as a '**correlative *wh*-clause' (COR)**. It is morphosyntactically identical to the INT in (5)b, but occurs left dislocated with a (boldfaced) preposed *wh*-phrase linked to a (boldfaced) demonstrative in the matrix clause by agreement and interpretation.[6,7]

5) a. [**Ce** scrie azi], **aia** va edita mâine.
what writes today that will edit tomorrow
'What she/he writes today, she/he will edit tomorrow.'

b. Maria se întreabă [ce scrie azi].
Maria REFL wonders what writes today
'Maria wonders what she will write today.'

Other Non-INTs—such as restrictive headed relative clauses, *-ever* free relative clauses (Dayal 1997, von Fintel 2000), free-choice free relative clauses (Caponigro and Fălăuș 2018), and the *wh*-clauses in Rudin constructions (Caponigro and Fălăuș 2022)—also display non-propositional interpretations and can be captured by means of the same formal tools I introduce for the analysis of R-FRs, ∃-FRs, and CORs. I cannot discuss them in this paper solely for reasons of space.

3. Generalizations about *wh*-clauses and their *wh*-words

In this section, I introduce three main crosslinguistic generalizations about the kind of connection (or dependency) Non-INTs and their *wh*-words exhibit with respect to INTs and their *wh*-words within and across languages. To the best of my knowledge, these generalizations have not yet been fully articulated together. They are supported by findings about INTs and Non-INTs across languages that have emerged in the past couple of decades, although further data and analysis are needed, especially about Non-INTs, which have received less attention across descriptive, typological, and theoretical work.

The first generalization is given in (6) and focuses on the distribution of different varieties of *wh*-clauses across languages, highlighting the central role of INTs. Examples from specific languages

[6] Ex. (5)a is from Caponigro and Fălăuş (2023a: ex. 6) with a slightly revised translation; ex. (5)b is from Caponigro and Fălăuş (2023a: ex. 1)

[7] Correlative clauses contain one or more constituents marked either by a *wh*-word or by a demonstrative, depending on the language (Belyaev and Huang 2020). Only correlative clauses with *wh*-constituents are considered here, given the focus on *wh*-clauses, and the abbreviation "COR" refers to them only.

are provided below.

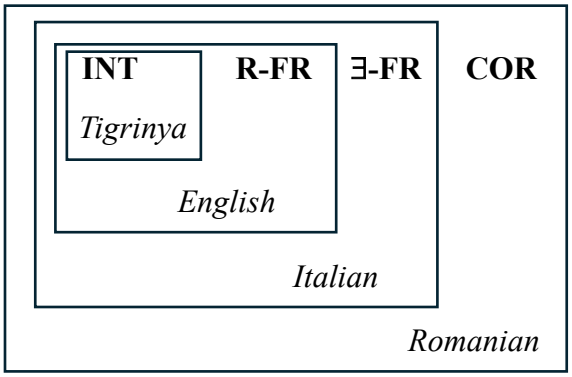
6) **GENERALIZATION 1: *wh*-clauses across languages**[8]

(i) While INTs are attested in almost all languages, only a subset of those languages allows for Non-INTs as well.

(ii) Within the subset of languages that allow for Non-INTs, there is variation in which varieties are attested.

Tigrinya (Ethio-Semitic) only allows for INTs. *Wh*-words cannot occur in any Non-INT (Cacchioli and Caponigro 2026).[9] English allows for some Non-INTs (R-FRs; ex. (3)a), but lacks others (∃-FRs, CORs). Italian allows for the same Non-INTs as English with the addition of ∃-FRs (ex.(4)a,b). Finally, Romanian exhibits the largest inventory of Non-INTs among the four languages mentioned so far (ex. (5)a). This picture is summarized in Table 1.

Table 1. Varieties of *wh*-clauses across four languages



The alignment of *wh*-clauses in Table 1 from left to right is not supposed to convey any implicational hierarchy. While every language with Non-INTs also has INTs, there are languages that use *wh*-clauses to form ∃-FRs but lack R-FRs[10] and languages with CORs but neither variety of FRs.[11]

The second generalization is given in (7) and focuses on the distribution of individual *wh*-words across varieties of *wh*-clauses and languages.

[8] Šimík (2026) argues for a similar but stronger generalization, proposing a precise hierarchy of varieties *wh*-clauses (his varieties only partially overlap with mine). I further discuss Šimík (2026) in § 7.
[9] Tigrinya does use *wh*-clauses to form exclamative clauses (Gioia Cacchioli p.c.), but see fn. 5 for the reason why exclamative clauses are not grouped with non-interrogative *wh*-clauses.
[10] E.g., for Pesh (Chibchan) (Chamoreau 2021).
[11] E.g., Erzya (Finno-Ugric), Maninka (Mande), and Vai (Mande) (De Vries 2002: Appendix II, Table 4, p. 388).

7) **GENERALIZATION 2: Distribution of *wh*-words across *wh*-clauses and languages**[12]

(i) The set of *wh*-words used in Non-INTs in a language tends to be a proper subset of the set of *wh*-words used in INTs in the same language.

(ii) Different subsets of *wh*-words may be licensed within different varieties of Non-INTs in the same language.

(iii) Different subsets of *wh*-words may be licensed within the same variety of Non-INTs across languages.

Table 2 exemplifies the three claims in Generalization 2 by providing the distribution of *wh*-words in different varieties of Non-INTs in three languages: English, Italian, and Romanian.

Table 2. Distribution of *wh*-words across different varieties of Non-INTs in three languages

ENGLISH		ITALIAN			ROMANIAN			
INT	**R-FR**	**INT**	**R-FR**	**∃-FR**	**INT**	**R-FR**	**∃-FR**	**COR**
who	√/*	*chi*	√	√	*cine*	*/√	√	√
what	√	*che cosa* *cosa* *che* *quanto*	* % % √	* √ √ *	*ce*	√	√	√
what + NP	√/*	*che* + NP	*	*	*ce* + NP	√	√	√
which + NP	*	*quale* + NP	*	*	*care* + NP	*	*	√
where	√	*dove*	√	√	*unde*	√	√	√
when	√	*quando*	√	*	*când*	√	√	√
how	√	*come*	√	*	*cum*	√	√	√
how much	√	*quanto*	√	*	*cât*	√	√	√
how much *how many* + NP	√	*quanto/a* *quanti/e* + NP	√	*	*cât/ă* *câti/e* + NP	√	√	√
how + Adj/Adv	√	*quanto* + Adj/Adv	*	*	*cât* + *Adj/Adv*	√	√	√
why	*	*perché*	*	*	*de ce*	√	√	√

Although most English *wh*-words can occur in both INTs and R-FRs, the latter exhibit some restrictions or gaps. For instance, *who* is fully productive in INTs, while R-FRs introduced by *who* are highly restricted (Patterson and Caponigro 2016; Stockwell and Schütze 2022). *What* as a *wh*-phrase is fully productive in R-FRs, while *what* followed by a mass/plural NP complement is productive for at least some speakers, similarly to *how much*/*many* followed by a mass/plural NP

[12] The core observations underlying this generalization already appear in Caponigro (2003, 2004), while a closely related generalization is developed in Chierchia and Caponigro (2013). Part (i) of Generalization 2 is stated in Šimík (2020: ex. 67) as "Generalized Caponigro's generalization".

complement (Caponigro 2025: ex. 5–8). *Why*, instead, can never introduce R-FRs (e.g., ***I left why you left***).

In Italian, the sets of *wh*-words occurring in R-FRs and in ∃-FRs are proper subsets of the set of *wh*-words occurring in INTs and overlap only partially with each other. Finally, Romanian has both the most varieties of Non-INTs and the largest number of *wh*-words used, with few differences from one variety to another. Still, only CORs make use of all *wh*-words attested in INTs.

The third and last generalization is given in (8) and deals with *wh*-clauses with multiple *wh*-phrase, highlighting the close connection between INTs and Non-INTs.

8) **GENERALIZATION 3: Multiple *wh*-phrases across *wh*-clauses**

(i) Languages that allow for INTs with multiple *wh*-phrases[13] also allow for Non-INTs with multiple *wh*-phrases.

(ii) In those languages, distributional and ordering restrictions of *wh*-phrases that hold for INTs also hold for Non-INTs.

For instance, Romanian is well-known for allowing for INTs with multiple *wh*-phrases and requiring all the *wh*-phrases in an INT to be fronted, as shown in (9)a. The same pattern is observed in R-FRs with multiple *wh*-phrases, as shown in (9)b, ∃-FRs with multiple *wh*-phrases, as shown in (9)c, and CORs with multiple *wh*-phrases, as shown in (9)d.[14]

9) a. Bunica se întreabă [**cine** **unde** l-a cunoscut pe Ion].
Grandma REFL asks who where him-has met ACC Ion
'Grandma wonders who met Ion where.'

b. Proprietarul a aranjat [**ce** **unde** a trebuit instalat].
owner-the has arranged what where has needed installed
'The owner arranged the things that needed to be installed in the appropriate place.'

c. Bunica are [**ce** **cui** da de Crăciun].
Grandma has what who.DAT give.INF for Christmas
'Grandma has things to give to people for Christmas.'

d. [**Cine** **ce** aduce], (*acela*) *aia* va mânca.
who what brings that-one that will.3SG eat
'Everybody eats whatever they bring.'

[13] Like English and Romanian and unlike Italian, which only allow for *wh*-clauses with one *wh*-phrase.
[14] (9)a is from Caponigro and Fălăus (2020: ex.44b), (9)b is from Caponigro and Fălăus (2020: ex.4), (9)c is from Caponigro and Fălăus (2023a: ex. 38a), (9)d is from Caponigro and Fălăuş (2023a: ex.49a)

In Romanian, *wh*-phrases in INTs not only are all obligatorily fronted—a feature observed also in Slavic languages[15]—but must follow a strict hierarchical order: 'who' > 'what' > 'where'/'when'/'how'. The very same restriction holds for all Non-INTs with multiple *wh*-phrases. Flipping the order of the *wh*-phrases in each of the *wh*-clauses in (9)a–d would make each sentence unacceptable.

∃-FRs and CORs with multiple *wh*-phrases have already been discussed separately or at least mentioned in the literature in various languages at various degrees, although I am not aware of any unified analysis. R-FRs with multiple *wh*-phrases, instead, are a recent discovery by Caponigro and Fălăuş (2020). In the next section, I develop an analysis for interrogative and Non-INTs that not only accounts for the semantic properties of these *wh*-clauses in a uniform way but also provides an explanation for the three generalizations above.

4. The semantics of *wh*-clauses and *wh*-word meaning shifts

The core idea behind the analysis developed in this section is that all *wh*-clauses denote the same propositional meaning at the same point in their syntactic structure and semantic derivation. In § 4.1, I formulate this core meaning and derive it compositionally. I then sketch how this core meaning can be used to build semantic analyses for all the varieties of *wh*-clauses introduced in § 2, beginning with those with a single *wh*-phrases (§ 4.2) and then those with multiple *wh*-phrases (§ 4.3). None of the analyses are meant to be exhaustive or final. A large body of descriptive, typological, and theoretical literature on INTs has been built over the past fifty years, which I cannot do justice to here. We know much less about R-FRs, ∃-FRs, and CORs, especially crosslinguistically. The proposal that I am about to sketch should be taken as a proof of concept that a propositional analysis can be developed for all *wh*-clauses and their *wh*-words.

4.1. Shared syntactic and semantic core of all *wh*-clauses

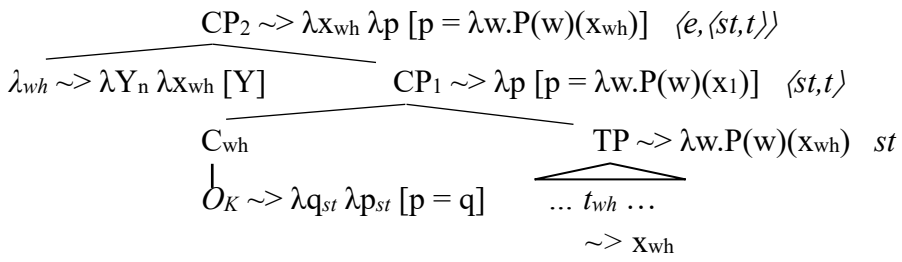
I propose that all *wh*-clauses—INTs and Non-INTs alike—denote a relation between entities and propositions at the same point of their syntactic structure and semantic derivation.[16] This semantic output results from the interplay of three components that I assume all *wh*-clauses share: (i) a *wh*-trace/variable licensed by a *wh*-phrase, (ii) a propositional operator (O_K) in C that lifts a proposition to a set of propositions, and (iii) a set-formation operator (λ) that abstracts over the *wh*-trace/variable in preparation for the clause to combine with its *wh*-phrase. (10) shows in detail the syntactic structure and the semantic properties that I argue are shared by all *wh*-clauses—with

[15] E.g., Bulgarian, Czech, Polish, Russian, and Serbo-Croatian (Rudin 1988).

[16] I owe this core idea entirely to Gennaro Chierchia. In a small workshop organized by Anamaria Fălăuş and me in Nantes in May 2025, Chierchia presented a semantic analysis for INTs and CORs with multiple *wh*-phrases, arguing that they share the same semantic derivation up to the point at which both denote a relation between individuals and propositions (Chierchia 2025b). He also hinted at the possibility of extending this approach to certain other Non-INTs. I am solely responsible for attempting to extend it to additional Non-INTs and for implementing it the way is presented here.

possible morphosyntactic differences within their TPs and above their CP_2 (e.g., tense, mood, word order).

10) *Shared syntactic and semantic features of all* wh-*clauses*:

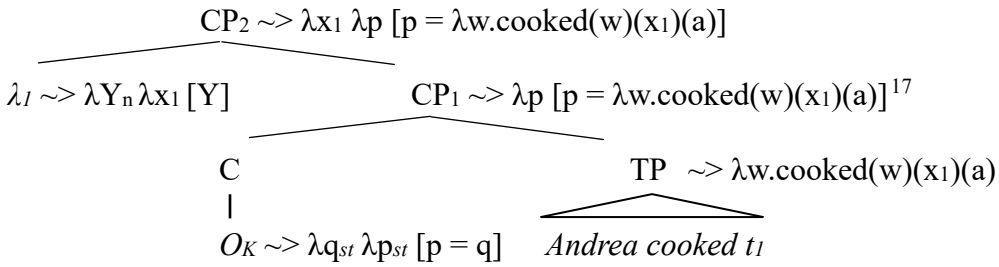


The TP in all *wh*-clauses denotes an open proposition containing (at least) one free variable licensed by (at least) one moved *wh*-phrase and its *wh*-trace t_{wh}. Its sister C_{wh} denotes the operator O_K, which turns the open proposition into the singleton containing it. The definition of the operator O_K is borrowed from Karttunen (1977) (the *K* subscript is a homage to Karttunen and his seminal paper). Karttunen's operator is a crucial component of his fully compositional development of Hamblin's (1973) seminal idea that interrogative clauses denote sets of propositions. O_K is the propositional version of the more general rule within predicate calculus with the λ-operator that allows a given semantic object to be turned into a singleton set containing it. The operator λ_{wh} is also familiar: it's the first step in a two-step procedure by which a moved constituent binds its trace/variable by function application. λ_{wh} abstracts over the *wh*-trace/variable x_1 of its CP_1 sister to return the denotation of their mother node CP_2: a relation between entities *e* and propositions *p* that returns the truth if, given an *e* and a *p*, *p* is identical to the open proposition denoted by TP once it is saturated by *e*.

The proposal is illustrated in (11) with the syntactic structure and the semantic derivation of the relevant portion of the *wh*-clause *what Andrea cooked*.

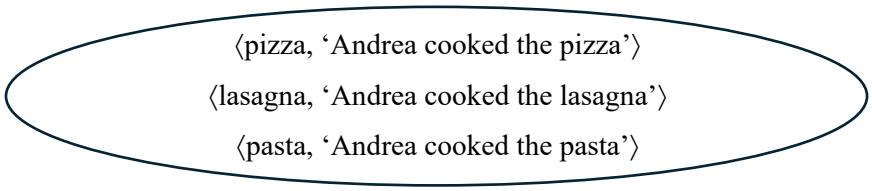
11) a. [$_{CP}$ λ_1 O_K Andrea cooked t_1]

b.



The denotation of CP_2 in (11)b is a relation between entities and propositions or equivalently a set of ordered pairs ⟨entity, proposition⟩ in which the entities *e* range over food and the propositions are of the form 'Andrea cooked *e*', with *e* being the individual occurring as the first element of the pair. This result is represented in a greatly simplified form in (12), assuming a toy model containing only three (food) entities: pizza, lasagna, and pasta.[18]

12)



A clause denoting a relation between entities and propositions is not a syntactic/semantic object that can straightforwardly combine with regular syntactic/semantic objects. It cannot act as a regular transitive predicate because of its syntactic nature as a clause (CP). Nor can it act as the argument of a predicate, because of its semantic type and content (a 2-place relation of type $\langle e,\langle st,t\rangle\rangle$). The role of the *wh*-phrase is to quantify over entities and turn the relation into a set of propositions or to quantify over propositions and turn the relation into a set of entities. I argue that the meaning differences between INTs and Non-INTs, and among different varieties of Non-INTs, are due to differences in lexical meaning between morphosyntactcally identical (or similar) *wh*-words in different *wh*-clauses.

I assume the meaning of *wh*-words in INTs to be the diachronic source for the meanings of *wh*-words in all other *wh*-clauses via logic-driven processes of meaning change. These changes do not apply automatically to all *wh*-words in all kinds of *wh*-clauses, as highlighted by the variations and restrictions in Generalization 1 and Generalization 2. Rather, they need to be licensed *wh*-word by

[17] Henceforth, I provide only the final output of each step of a semantic derivation, omitting all intermediate steps of function application and *λ*-reduction for the sake of simplicity and readability.

[18] In (12) and subsequent examples, nouns like lasagna, pasta, and pizza without any quotation mark stand for semantic objects (entities in the model) rather than linguistic objects (nouns). Similarly, 'Andrea cooked pizza' and other strings in single quotation marks stand for semantic objects (propositions, i.e., functions from worlds to truth values) rather than linguistic objects (sentences).

wh-word, *wh*-clause by *wh*-clause, and language by language. I examine them in detail next, *wh*-clause by *wh*-clause.

4.2. Semantic analyses of *wh*-clauses with a single *wh*-phrase

In this section, I sketch the semantic analyses for all varieties of *wh*-clauses introduced in § 2 when they occur with just one *wh*-phrase. I start with *wh*-clauses conveying propositional content, which I assume are the diachronic source and starting point for all the other varieties: INTs (§ 4.2.1). I then move to *wh*-clauses conveying non-propositional content: R-FRs (§ 4.2.2), ∃-FRs (§ 4.2.3), and CORs (§ 4.2.4).

4.2.1 Semantic analysis of embedded interrogative *wh*-clauses (INTs)

The denotation of a matrix/embedded INT is a set of propositions, following Hamblin (1973) and Karttunen (1977). It results from combining the denotation of the *wh*-phrase with the denotation of its clausal sister. The denotation of the clausal sister is the one already given in (10): a relation between entities and propositions or, equivalently, a set of ordered pairs ⟨entity, proposition⟩. As for the denotation of the *wh*-words in an INT, I propose the one schematized in (13): a function that applies to a 2-place relation *S* between entities and propositions to return the set of propositions *q* that saturate *S* together with entities that satisfy the *WH* restriction (human, thing, location, etc.).

13) *New proposal for semantic contribution of* wh-*words in INTs:*
wh $_{[INT]}$ ~> $\lambda S_{\langle e,\langle st,t\rangle\rangle} \lambda q_{\langle st\rangle} \exists x[WH(w_0)(x) \wedge S(x)(q)]$ [19]

The *WH* restriction and existential quantification over the individual variable *x* in (13) are the same as in Karttunen's denotation of *wh*-words. Karttunen treats a *wh*-word as semantically equivalent to an indefinite pronoun (e.g., *what* is semantically the same as *something*)—both denoting an existential generalized quantifier, as shown in (14).

14) *Semantic contribution of* wh-*words in INTs (Karttunen 1977):*
wh $_{[INT]}$ ~> $\lambda P_{\langle s,et\rangle} \exists x[WH(w_0)(x) \wedge P(w_0)(x)]$

Karttunen then assumes a new semantic rule to combine the generalized quantifier denoted by the *wh*-phrase with the set of propositions denoted by the remainder of the clause. The denotation I am proposing in (13) lexicalizes Karttunen's new combinatorial rule as part of denotation of the *wh*-word by means of the higher-order variable *S* (type ⟨e,⟨st,t⟩⟩) and the *λ*-abstraction over it.

[19] Since lexical meanings are central to my analysis, I usually focus on *wh*-words that form *wh*-phrases on their own without additional lexical material (e.g., *what* in *what Andrea cooked* rather that [*what* NP] in *what food Andrea cooked*), unless otherwise indicated. A general lexical entry for a *wh*-word with extra lexical material in its *wh*-phrase based on (13) would look like (i), where N is the property denoted by the *wh*-phrase minus the *wh*-word (e.g., N would be the property denoted by *food* in *what food*):
(i) *wh-* $_{[INT]}$ ~> $\lambda N \lambda S \lambda q \exists x[N(w_0)(x) \wedge S(x)(q)]$

(15) provides an example of combining the *wh*-phrase *what* with the remainder of its clause. The whole *wh*-clause is unambiguously an interrogative clause since it occurs as the complement of the interrogative predicate *ask*. The semantic contribution of CP_1 in (15) is the semantic core I'm proposing for all *wh*-clauses: a 2-place relation between entities and propositions, as already schematized in (10) and exemplified in (11)b. Once this relation combines with the denotation of *what* [INT] given in (13), it returns the standard set of propositions as the denotation of $CP_{2\ [INT]}$, as exemplified in (15)c.

15) a. Luca asked [what Andrea cooked].

b. $CP_{2\ [INT]} \sim> \lambda q\ \exists x\ [TH(w_0)(x) \wedge q = \lambda w.cooked(w)(x)(a)]$

$what_{1\ [INT]} \sim> \lambda S\ \lambda q\ \exists x[TH^{20}(w_0)(x) \wedge S(x)(q)]$ $\quad CP_1 \sim> \lambda x_1\ \lambda p\ [p = \lambda w.cooked(w)(x_1)(a)]$

λ_1 *O*$_K$ *Andrea cooked* t_1

c.

'Andrea cooked the pizza'

'Andrea cooked the lasagna'

'Andrea cooked the pasta'

4.2.2. Semantic analysis of referential free relative clauses (R-FRs)

R-FRs always refer to either an entity or a kind—exhibiting the same semantic behavior as definite DPs in languages like Italian and Romanian.[21] For instance, the bracketed R-FR in (16)a refers to the maximal (plural) entity resulting from the sum of all the things Andrea cooked. If Andrea cooked the lasagna, the pasta and the pizza, then the R-FR refers to the maximal plural entity resulting from the sum of those three atomic entities: lasagna⊕pasta⊕pizza.[22] The bracketed R-FR in (16)b, instead, refers to the kind of nature that grew on the highlands in the past: ∩highland-nature-in-the-past.[23]

16) a. Luca ate [what Andrea cooked].

b. [What grew on the highlands] is now extinct.

[20] *TH* stands for 'thing', the semantic restriction introduced by the *wh*-word *what*.

[21] See Jacobson (1995) and Dayal (1996) for the first seminal semantic analyses of entity-denoting R-FRs and Caponigro (2003, 2004, 2022) for further developments. See also Caponigro (2003, 2004, 2021) and Šimík (2020) for crosslinguistic data and patterns.

[22] ⊕ is the sum operator in Link (1983) that forms plural individuals from atomic ones.

[23] ∩ is Chierchia's (1982, 1983) nominalization operator that turns sets of entities into kinds.

I derive the denotations of R-FRs starting from the core component shared by all *wh*-clauses: a CP denoting a relation between entities and propositions. This is given in (17)a for the R-FR in (16)a. (17)b illustrates the relation at the world of evaluation with a toy model with just four food items: focaccia, lasagna, pasta, and pizza. In the world of evaluation w_0, Andrea cooked all of them except the focaccia. The proposition 'Andrea cooked the focaccia' forms one of the ordered pairs in (17)b because (17)a imposes no requirement on the included propositions being true in w_0.

17) a. $CP_{WH\text{-}clause} \rightsquigarrow \lambda x_1 \lambda p[p = \lambda w.cooked(w)(x_1)(a)]$

b.

⟨focaccia, 'Andrea cooked the focaccia'⟩
⟨pizza, 'Andrea cooked the pizza'⟩
⟨lasagna, 'Andrea cooked the lasagna'⟩
⟨pasta, 'Andrea cooked the pasta'⟩

The logical contribution of the *wh*-word in an R-FR is schematized in (18). It is a function that applies to a relation *S* between entities and propositions to return a single entity: the largest sum of all entities that satisfy the WH restriction and truthfully saturate the propositional abstract.

18) $wh_{\text{[R-FR]}} \rightsquigarrow \lambda S_{\langle e,\langle st,t\rangle\rangle}\, \iota x \exists q \forall y([WH(w_0)(y) \wedge S(y)(q) \wedge q(w_0)] \rightarrow y \leq x)$

Therefore, the *wh*-word *what* in the R-FR in (16)a applies to the relation in (17)b and returns the maximal entity of the set in (19). This set contains all the entities that occur as the first member of the ordered pairs in the relation in (17)b whose second member is a true proposition in the world of evaluation w_0, i.e., all the entities that were cooked by Andrea in w_0 (the bottom layer). It also contains all their sums, including the largest sum, the maximal one (the boldfaced single entity in the top layer). It is the maximal entity that *what* returns as the final denotation of its R-FR.

19)

lasagna⊕pasta⊕pizza
lasagna⊕pasta pasta⊕pizza lasagna⊕pizza
lasagna pasta pizza

(20)b provides the detailed semantic derivation of the R-FR in (16)a, repeated below as (20)a for convenience.

20) a. Luca ate [*what Andrea cooked*].

b. CP ~> ιx∃q∀y([TH(w_0)(x) ∧ q=λw.cooked(w)(y)(a) ∧ q(w_0)] → y ≤ x)

what$_1$ [R-FR] ~> λSιx∃q∀y([TH(w_0)(y) ∧ S(y)(q) ∧ q(w_0)] → y ≤ x) CP

λ$_1$ *O*$_p$ *Andrea cooked* t$_1$

~> λx$_1$λp[p=λw.cooked(w)(x$_1$)(a)]

R-FRs can also denote kinds, as shown in (16)b. I account for this denotation adopting Chierchia's (1983) *nom* operator ∩ from set of entities *et* to kind entities, as shown in (21), using the denotation schema for *wh*-words in (22), which lexically encodes ∩.

21) [What grew on the highlands] ~> ∩λx∃q[TH(w_0)(x) ∧ q = λw.grew-on-hl(w)(x) ∧ q(w_0)]

22) *wh* [R-FR-∩] ~> λS$_{\langle e,\langle st,t\rangle\rangle}$∩λx∃q[WH($w_0$)(x) ∧ S(x)(q) ∧ q($w_0$)]

The logical operators lexicalized in the denotation of *wh*-words in R-FRs, i.e., ι outputting the maximal entity of a given set (18) and ∩ outputting the kind of a given property in (22), are not new nor stipulative. They are lexicalized in the denotation of the definite determiner in languages like Italian or Romanian, and—with some differences—in English too. For instance, the R-FR denoting a maximal entity in (16)a could be paraphrased with the bracketed definite DP in (23)a, while the R-FR denoting a kind in (16)b could be paraphrased with the bracketed definite DP in (23)b.

23) a. Luca ate [the stuff/food/dishes Andrea cooked].
b. [The nature/plants that grew on the highlands] is/are now extinct.

This parallelism is further strengthened by morphological evidence from languages like Bulgarian, Greek, and Hungarian. In those languages, the *wh*-words occurring in R-FRs carry an extra prefix with respect to those in INTs. Such prefix resembles the definite determiner, at least in Greek.

True propositional content. I conclude with a general remark concerning the analysis of R-FRs and all other Non-INTs. The requirement in (18) for the saturated propositional abstract *q* to be true (*q*(w_0)) characterizes the denotation of all Non-INTs. This requirement captures the intuition that the denotation of a Non-INT results from conditions on propositional content that hold at the world of evaluation w_0. The R-FR *what Andrea cooked* in (16)a denotes the maximal entity in (19) in w_0. This matches the semantic intuitions for that R-FR: the food/stuff/dishes that Andrea cooked in w_0—extensionally, the maximal plural individual lasagna⊕pasta⊕pizza in our toy model. The requirement for the saturated propositional abstract *q* to be true does not apply to INTs. The denotation of an INT should include all the propositions that are possible answers to a question, not just those that are true answers. While there are question-embedding predicates like *know* that

select only for the answers to a question that are true in the world of evaluation, there are also question-embedding predicates like *ask* that require all the possible answers to a question or question-embedding predicates like *agree on* that allow for a subset of those answers, but not necessarily those that are true in the world of evaluation. For instance, *Gioia and Valentina agree on what Andrea cooked* is true if Gioia and Valentina agree that Andrea cooked the focaccia, despite being false in w_0.

4.2.3. Semantic analysis of existential free relative clauses (∃-FRs)

∃-FRs, as the one in (24) from Italian, only occur as complements of a small class of predicates: existential 'be' and/or 'have' introduce ∃-FRs in all languages in which they are attested. In a subset of those languages, ∃-FRs can also occur as complements of intensional predicates like 'seek', 'look for' and others like 'find'.[24]

24) C'è [chi parla sempre].
there's who talks always
'There are people who talk all the time.'

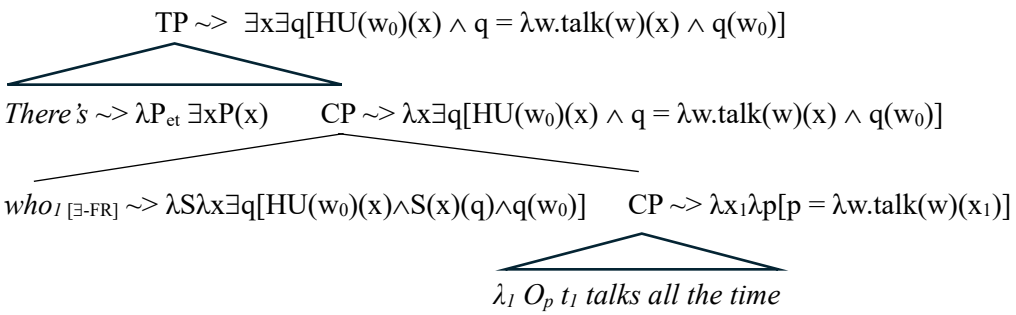
Semantically, ∃-FRs behave like narrow scope indefinites—a semantic behavior resembling that of bare nominals in the complement positions of the same predicates in languages like English, as shown in the translation of (24). Following Caponigro (2001, 2003, 2004), I assume that the denotation of an ∃-FR is a set of entities, but I am revising the denotation of the *wh*-word in an ∃-FR in order to handle the new propositional component. It is now a function that applies to a relation *S* between entities and propositions—the usual core denotation of all *wh*-clauses—to return the set of entities that contains all and only those entities that both satisfy the WH restriction and saturate the propositional abstract so as to make it true at the world of evaluation w_0, as shown in (25).

25) $wh_{[\exists\text{-FR}]} \rightsquigarrow \lambda S_{\langle e,\langle st,t\rangle\rangle} \lambda x\, \exists q\, [WH(w_0)(x) \wedge S(x)(q) \wedge q(w_0)]$

The ∃-FR in (24) occurs as the complement of the existential construction 'there is', which I assume to denote a function $\langle et,t \rangle$ that existentially closes the set it applies to. In (26), I provide the syntactic structure and semantic derivation for the full sentence.[25]

[24] ∃-FRs are not attested in English or in any other Germanic language, with the exception of Yiddish; see Caponigro (2003, 2004, 2021) and Šimík (2011; 2017) for crosslinguistic data and analyses.
[25] Henceforth, I use English translation/glosses and English word order for the syntactic tree and the semantic derivation of examples from languages other than English

26)



4.2.4. Semantic analysis of correlative clauses (CORs)

CORs have received more attention in descriptive and typological literature than R-FRs or ∃-FR. Important formal work has been conducted on their syntax and semantics. Still, further investigation is needed to understand various open issues, including crosslinguistic variation in form and interpretation.[26] My limited goal in this section is to sketch how my propositional analysis of *wh*-clauses can be extended to CORs, based on current findings about CORs across languages. An example of a COR in Romanian is given in (27).

27) [**Ce**$_1$ a gătit Andrea], **aia**$_1$ a mâncat Luca.
what has cooked Andrea that has eaten Luca
'Luca ate what Andrea cooked.'

All CORs share the following features, which also distinguish them from the other *wh*-clauses: (i) they are dislocated to the left of their matrix clause and (ii) each of their *wh*-phrases must be linked to a phrase/constituent in the matrix clause ("anchor") that can be realized as a personal pronoun, a demonstrative pronoun, a demonstrative with its nominal complement, or no overt marker at all.

Following Dayal (1996), I assume that a COR with one *wh*-phrase semantically behaves like an R-FR, although wearing a slightly different logical dress. Like an R-FR, a COR is referential and points at a maximal individual or a kind, but it does so not by denoting directly an entity or a kind (type *e* or *se*), but rather by denoting the corresponding generalized quantifier—the set of all sets containing that entity (⟨et,t⟩) or that kind (⟨⟨s,et⟩,t⟩). The denotations for the *wh*-phrase in a COR in (27) are made to achieve these semantic goals: functions that apply to the usual relation *S* between entities and propositions to return the generalized quantifier of the maximal entity (28)a or the kind (28)b that would be denoted by the corresponding R-FRs.

[26] See Dayal (1996) for a seminal semantic analysis of CORs, with a focus on Hindi. Brasoveanu (2012) provides an overview on the semantics of CORs. Chierchia (2025a) argues for a new analysis of CORs in Hindi and more generally. See Belyaev and Haug (2020) and the references therein for a recent formally informed typological survey of CORs, focusing on their formal features and genesis.

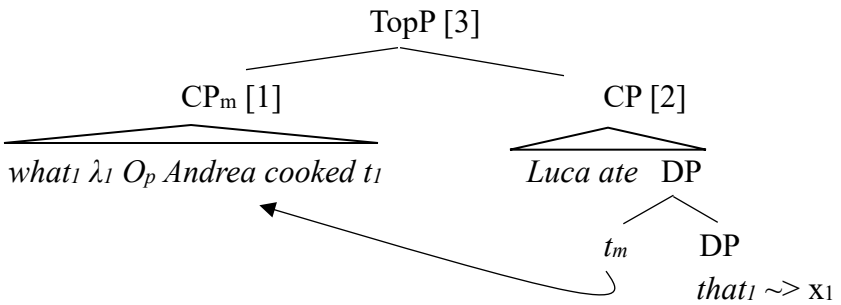
28) a. *wh* $_{[COR\text{-}\iota]}$ ~> $\lambda S_{\langle e,\langle st,t\rangle\rangle}\, \lambda P_{et}.P(\iota x \exists q \forall y([WH(w_0)(y) \wedge S(y)(q) \wedge q(w_0)] \rightarrow y \leq x))$

b. *wh* $_{[COR\text{-}\cap]}$ ~> $\lambda S_{\langle e,\langle st,t\rangle\rangle}\, \lambda P_{\langle se,t\rangle}.P(\cap \lambda x \exists q[WH(w_0)(x) \wedge S(x)(q) \wedge q(w_0)]\ )$

Within a type-shifting view á la Partee (1986), one might not need to assume the COR-specific denotations in (28): individual/kind-denoting expressions can freely shift to generalized quantifiers whenever needed for the semantic computation. Still, two main facts argue for the formal difference in the denotation of a COR and its *wh*-phrase in (28) with respect to an R-FR and its *wh*-phrase in (18) and (22). First, a COR is always left dislocated and linked to an anchor in its matrix clause, while an R-FR occurs inside its matrix clause, as an argument or a clause-internal adjunct. I take these different syntactic positions to mark different modes of semantic combination. Second, Generalization 2 and Table 2 show that the subset of *wh*-words in CORs may not be the same as the one in R-FRs in the very same language. In Romanian, for instance, *cine* 'who' is fully productive in CORs, while it is degraded in R-FRs. Also, *care* 'which' is fully acceptable in CORs, while it is unacceptable in R-FRs. Had the same denotation been assumed for both *wh*-words in both constructions, the differences in distribution would have had to be accounted for by means of an extra mechanism.

In (29), I provide the syntactic structure and semantic derivation of the example of the Romanian COR in (27). Comments follow.

29) a.



b. [1]: CP_m ~> $\lambda P.P(\iota x \forall y \exists q[[TH(w_0)(y) \wedge q = \lambda w.cooked(w)(y)(a) \wedge q(w_0)] \rightarrow y \leq x])$

[2]: CP ~> $\lambda x_1\, [ate(w_0)(x_1)(l)]$

[3]: TopP ~> $ate(w_0)(\iota x \forall y \exists q[[TH(w_0)(y) \wedge q = \lambda w.cooked(w)(y)(a) \wedge q(w_0)] \rightarrow y \leq x)(l)$

Syntactically, I adopt Bhatt's (2003) proposal as revised by Chierchia (2025a): a COR is generated as the adjunct to its demonstrative anchor in the matrix clause and then moved to a Topic position (TopP) on top of the matrix clause. I refer to the cited works for supporting arguments. Semantically, I treat a COR as denoting a generalized quantifier, as discussed above. Its matrix clause denotes the set of entities resulting from abstracting over the trace/variable introduced by the demonstrative anchor. Overall, I treat a COR as a topic, both syntactically and semantically, which combines with its matrix clause as a topic would combine with its comment.

Like R-FRs, CORs can denote kinds as well, as shown in (30). Kind-denoting CORs can be accounted for by the "kind" *wh*-word for CORs in (28)b.

30) [**Ce** crește în zona asta], **aia** e acum pe cale de dispariție.[27]
what grows in area this, that is now on way of disappearance
'What grows in this area is now going extinct.'

4.3. Semantic analyses of *wh*-clauses with multiple *wh*-phrases

In this section, I sketch the semantic analyses for *wh*-clauses with multiple *wh*-phrases. In doing so, I show that my approach can handle *wh*-clauses in all their varieties and forms. At the same time, investigating *wh*-clauses with multiple *wh*-phrases brings further evidence that Non-INTs are closely related to INTs at the semantic level as well and, therefore, supports the main idea of a common propositional core.

One crucial crosslinguistic fact should be kept in mind, at least based on the data currently available. As stated in Generalization 3 (§ 3), Non-INTs allow for multiple *wh*-phrases only in languages that allow for INTs with multiple *wh*-phrases.

Unlike the previous section, I start from R-FRs (§ 4.3.1) and ∃-FRs (§ 4.3.2). Their analysis can be straightforwardly derived from the logical ingredients already introduced for *wh*-clauses with a single *wh*-phrase. In particular, the semantic behavior of both R-FRs and ∃-FRs with multiple *wh*-phrases can be accounted for by assuming that their highest *wh*-phrase has the same denotation as the only *wh*-phrase in R-FRs and ∃-FRs with a single *wh*-phrase, respectively. All the other *wh*-phrases in an R-FR or an ∃-FR are interpreted as the *wh*-phrase in an INT. This conclusion further strengthens the connection between INTs and Non-INTs.

I then deal with INTs with multiple *wh*-phrases (§ 4.3.3). After mentioning their various readings discussed in the literature (single pair, pair list, and functional), I show that existing analyses can be framed within my approach. Finally, I extend this approach to CORs (§ 4.3.4) with multiple *wh*-phrases, since their readings resemble those of INTs. These semantic similarities between INTs and CORs together with the adoption of "functional" *wh*-phrases to account for both further support an approach like the propositional one I am arguing for.

4.3.1. Semantic analysis of referential free relative clauses with multiple *wh*-phrases

R-FRs with multiple *wh*-phrases have only been discovered and analyzed recently, by Caponigro and Fălăuş (2020). Although they mention examples from several languages (German, Czech, Serbian, and American English), including the one in (31), their empirical and theoretical discussion focuses on Romanian, a language—as they show—that productively allows for R-FRs with multiple *wh*-phrases.

[27] Thanks to Anamaria Fălăus for the example.

31) I gave you [what you had to put where]. Caponigro and Fălăuş (2020: ex. 6)

I exemplify my discussion using the new Romanian example in (32), which constitutes a variation on examples in Caponigro and Fălăuş (2020: ex. 1 and 13), by making using of two complex *wh*-phrases, each with a *wh*-word taking a singular count nominal as its complement.

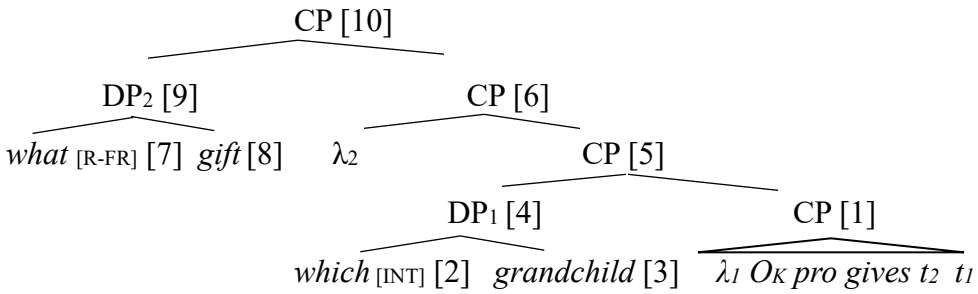
32) Bunica a împachetat [[**ce** **cadou**] [**cărui** **nepot**] dă de Crăciun].[28]
Grandma has wrapped what gift.SG which.SG.DAT grandchild gives for Christmas
'Grandma wrapped the gift she'll give to every grandchild for Christmas.'

The reason for discussing this example is because each of those singular *wh*-phrases would license a singular answer mentioning only one entity (either a gift or a child) in an INT with either *wh*-phrases as its only *wh*-phrase. Also, unlike 'which' + NP_{SG}, 'what' + NP_{SG} can introduce an R-FR with a single *wh*-phrase in Romanian. Such an R-FR would refer to just one atomic gift, as shown in (33).

33) Bunica a împachetat [**ce** **cadou** i-am zis să împacheteze.[29]
Grandma has wrapped what gift her-have.1SG told SUBJ wrap
'Grandma wrapped the gift I told her to wrap.'

Still, an R-FR like the one in (32) ends up referring to more than one atomic gift, because it refers to the largest sum of appropriate atomic gifts, as I show next. The syntactic structure of (32) is provided in (34)a, immediately followed by its semantic derivation in (34)b.

34) a.



b. [1]: CP ~> $\lambda x_1\, \lambda p\, [p = \lambda w.give(w)(x_2)(x_1)(gm)]$
[2]: *which* [INT] ~> $\lambda P\, \lambda S\, \lambda q\, \exists x\, [P(w_0)(x) \wedge S(x)(q)]$
[3]: *grandchild* ~> $gc_{\langle s,et\rangle}$
[4]: DP_1 ~> $\lambda S\, \lambda q\, \exists x_1\, [gc(w_0)(x) \wedge S(x_1)(q)]$
[5]: CP ~> $\lambda q\, \exists x_1\, [gc(w_0)(x) \wedge q = \lambda w.give(w)(x_2)(x_1)(gm)]$
[6]: CP ~> $\lambda x_2\, \lambda q\, \exists x\, [gc(w_0)(x) \wedge q = \lambda w.give(w)(x_2)(x_1)(gm)]$

[28] Thanks to Anamaria Fălăuş for this example. She also noted that, although she finds the example acceptable, she still prefers the variant in Caponigro and Fălăuş (2020: ex. 13) in which the lower *wh*-phrase is realized as *cui* 'who.DAT', rather than *cărui nepot*, as in (32)
[29] Thanks to Anamaria Fălăuş for this example.

[7]: *what* [R-FR] ~> $\lambda P\, \lambda S\, \iota x_2\, \exists q \forall y([P(w_0)(y) \wedge S(y)(q) \wedge q(w_0)] \rightarrow y \leq x_2)$

[8]: *gift* ~> $\text{gift}_{\langle s,et \rangle}$

[9]: DP_2 ~> $\lambda S\, \iota x_2\, \exists q \forall y([\text{gift}(w_0)(y) \wedge S(y)(q) \wedge q(w_0)] \rightarrow y \leq x_2)$

[10]: CP ~> $\iota x_2\, \exists q \forall y([\text{gift}(w_0)(y) \wedge \exists x\, [\text{gc}(w_0)(x) \wedge q = \lambda w.\text{give}(w)(x_2)(x_1)(\text{gm})] \wedge q(w_0)] \rightarrow y \leq x_2)$

The semantic derivation in (34)b is built around the crucial assumption that the lower *wh*-phrase is semantically the same as the morphosyntactically identical one in an INT with one *wh*-phrase: a function that applies to a relation between entities and propositions to return a set of propositions after existentially closing over the set of entities. The higher *wh*-phrase, instead, is semantically the same as the morphosyntactically identical one in an R-FR with a single *wh*-phrase: a function that applies to a relation between entities and propositions to existentially close over the set of propositions and return the largest sum of all the entities that truthfully saturate the propositional abstract. Notice that each *wh*-word has an extra argument (*λP[…P(w0)(x) …]*) which is saturated by the property denoted by the nominal complement of the *wh*-word.

This assumption on the different denotations of the two *wh*-words and their order of combination is further supported by the following empirical contrast. While all other *wh*-words can freely occur in Romanian R-FRs, *care* 'which' + NP can never occur as the only *wh*-phrase or the highest *wh*-phrase (Caponigro and Fălăuş 2020: ex. 13; 2023a: ex. 3 and ex. 5), but only under another *wh*-phrase. This contrast would be expected if *care* 'which' only had one denotation, the one in INTs, unlike other *wh*-words whose denotation have undergone logic-driven changes (see discussion in § 5).[30] My analysis requires a *wh*-phrase c-commanded by another one in a R-FR to be interpreted as in an INT, without any meaning change. On the other hand, the highes*t wh*-phrase must have undergone the meaning change appropriate for R-FRs.

My proposal is significantly simpler than the one in Caponigro and Fălăuş (2020) because it only makes use of *wh*-word denotations that have already been independently introduced for *wh*-clauses with a single *wh*-phrase, without the need for functional *wh*-phrases or functional *wh*-traces. This is possible because R-FRs with multiple *wh*-phrases, like all *wh*-phrases, have a propositional semantic core and all *wh*-words, including those in R-FRs, operate on propositions.[31]

4.3.2. Semantic analysis of existential free relative clauses with multiple *wh*-phrases

Romanian allows for ∃-FRs with multiple *wh*-phrases, unlike English or Italian. Romanian is the only one of the three languages to satisfy both core conditions for ∃-FRs with multiple *wh*-phrases,

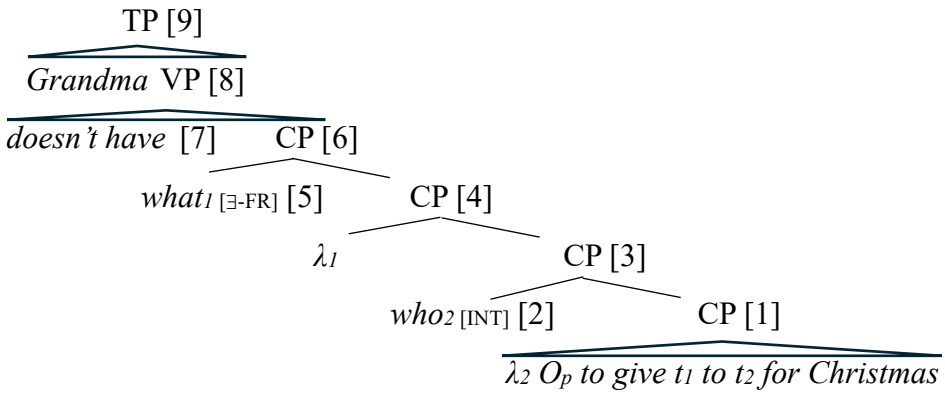
[30] See Caponigro and Fălăuş (2023b) for some speculations on why *care* 'which' has not undergone such meaning change.

[31] Caponigro and Fălăuş (2020) adopt a functional analysis for the lower *wh*-phrase of R-FRs, along the line of functional analyses of INTs with multiple *wh*-phrases (see § 4.3.3 in this paper). Since the variable over functions of type *⟨e,e⟩* and the existential quantifier that binds it—both introduced by the lower *wh*-phrase—are c-commanded by the universal quantifier over individuals introduced by the higher *wh*-phrase, the effect of the function is equivalent to that of plain existential quantification over individuals. I owe this observation to Gennaro Chierchia.

allowing for both (i) INTs with multiple *wh*-phrases (unlike Italian) and (ii) ∃-FRs with a single *wh*-phrase (unlike English). (35) provides an example of a Romanian ∃-FR with two *wh*-phrases, with its syntactic structure and semantic derivation. Comments follow.

35) a. Bunica nu are [ce cui da de Crăciun].
Grandma not has what who.DAT give.INF for Christmas
'Grandma doesn't have anything to give to people for Christmas.'

b.



c. [1]: CP ~> $\lambda x_2\, \lambda p\, [p = \lambda w.\text{give-for-C}(w)(x_1)(x_2)(gm)]$)
[2]: *who2* [INT] ~> $\lambda S\, \lambda q\, \exists x_2[HU(w_0)(x_2) \wedge S(x_2)(q)]$
[3]: CP ~> $\lambda q\, \exists x_2[HU(w_0)(x_2) \wedge q = \lambda w.\text{give-for-C}(w)(x_1)(x_2)(gm)]$
[4]: CP ~> $\lambda x_1\, \lambda q\, \exists x_2[HU(w_0)(x_2) \wedge q = \lambda w.\text{give-for-C}(w)(x_1)(x_2)(gm)]$
[5]: *what1* [∃-FR] ~> $\lambda S\, \lambda x_1\, \exists q[TH(w_0)(x_1) \wedge S(x_1)(q) \wedge q(w_0)]$
[6]: CP ~> $\lambda x_1\, \exists q[TH(w_0)(x_1) \wedge \exists x_2[HU(w_0)(x_2) \wedge q = \lambda w.\text{give-for-C}(w)(x_1)(x_2)(gm)] \wedge q(w_0)]$
[7]: *doesn't have* ~> $\lambda P_{et}\, \lambda y\, \neg\exists x[P(x) \wedge \mathbf{R}_{\langle e,et\rangle}{}^{32}(x)(y)]$
[8]: VP ~ > $\lambda y\, \neg\exists x\, \exists q[TH(w_0)(x) \wedge \exists x_2[HU(w_0)(x_2) \wedge q = \lambda w.\text{give-for-C}(w)(x)(x_2)(gm) \wedge q(w_0)] \wedge \mathbf{R}(x)(y)]$
[9]: TP ~> $\neg\exists x\, \exists q[TH(w_0)(x) \wedge \exists x_2[HU(w_0)(x_2) \wedge q = \lambda w.\text{give-for-C}(w)(x)(x_2)(gm) \wedge q(w_0)]) \wedge \mathbf{R}(x)(gm)]$

The English translation of (35)a conveys Romanian speakers' semantic intuitions that ∃-FRs with multiple *wh*-phrases behave like narrow scope indefinites by replacing them with bare nominals (*things*, *people*). The proposed analysis accounts for these intuitions by adopting a strategy similar to the one for R-FRs with multiple *wh*-phrases in the previous section. Unlike English, Romanian require all *wh*-phrases in INTs to occur at the front, according to a precise order. The same movement and ordering requirements are observed in ∃-FRs. Every *wh*-phrase in an ∃-FR that is c-commanded by (at least) another *wh*-phrase is interpreted as a *wh-[INT]*, i.e., it is assigned the

[32] ***R*** is a contextually salient relation between entities.

same denotation as the corresponding *wh*-phrase in an INT with a single *wh*-phrase ((35)c.[2]). The *wh*-phrase in an ∃-FR that c-commands all the others, instead, is interpreted as a *wh* $_{[\exists\text{-}FR]}$, i.e., it is assigned the same denotation as the *wh*-phrase in an ∃-FR with a single *wh*-phrase ((35)c.[5]).

Like the matrix predicate of the ∃-FR with a single *wh*-phrase in (26), the matrix existential predicate in (35)c.[7] takes a complement denoting a set of entities and existentially quantifies over it. Intuitively, the logical translation of the whole sentence in (35)c.[9] states that there isn't anything that Grandma would give to anybody for Christmas that is related to Grandma by a contextually determined relation ***R*** (e.g., Grandma owns is, Grandma has it on her gift list, etc.).

4.3.3. Semantic analysis of interrogative clauses with multiple *wh*-phrases

In this section, I extend my proposal to INTs with multiple *wh*-phrases, which are well known to license two kinds of answers. For instance, the embedded INT in (36)a can be answered by mentioning just a single pair of ⟨friend, dish⟩, as in (36)b, or with a list of those pairs, as in (36)c. Dayal (1996) argues that, in INTs with two singular *which*-phrases like (36)a, the list is functional in nature, i.e., each friend can be mapped onto only one dish. Therefore, an answer like (36)d is impossible, since it maps Andrea onto two different dishes.

36) a. Luca asked [*which friend* cooked *which dish*].
    b. *Single-pair answer:* Gianni answered [that Andrea cooked the lasagna].
    c. *Pair-list answer:* Gianni answered [that Andrea cooked the lasagna, Paolo cooked the pizza, Leo cooked the risotto].
    d. #Gianni answered [that Andrea cooked the lasagna, Andrea cooked the spaghetti, Paolo cooked the pizza, Leo cooked the risotto].

Notice that an INT with just one singular *wh*-phrase, as in (37)a, only allows for an answer mentioning a single friend, as in (37)b, but not an answer that mentions multiple (or a list of) friends, as in (37)c.

37) a. Luca asked [*which friend* cooked the lasagna].
    b. Gianni answered [that Andrea cooked the lasagna].
    c. #Gianni answered [that Andrea cooked the lasagna and Paolo cooked the lasagna (as well)].

A *which*-phrase with a singular count nominal as its complement always behaves morphosyntactically and semantically as singular, when occurring as the only wh-phrase in an INT. This contrasts with the behavior of a *wh*-phrase like *who*, which is morphosyntactically singular but semantically unspecified for number, and allows for both singular and plural/list answers, as shown in (38).

38) a. Luca asked [*who* cooked the lasagna].
    b. Gianni answered [that Andrea cooked the lasagna].
    c. Gianni answered [that Andrea cooked the lasagna and Paolo cooked the lasagna (as well)].

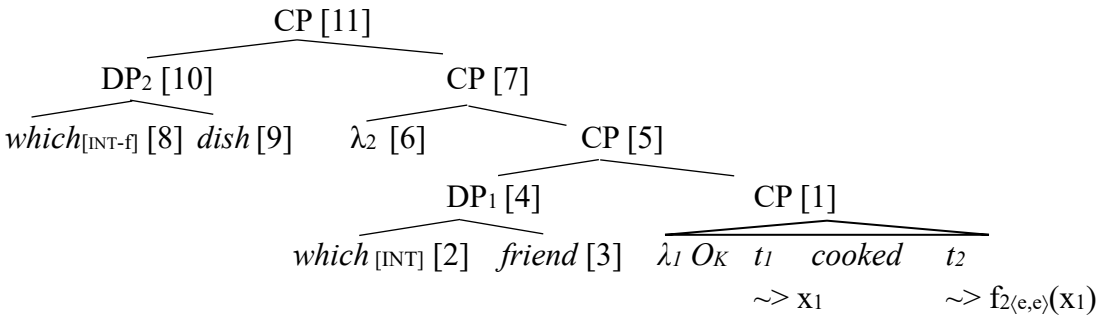
The contrasts in (36)–(38) show that a semantically singular *wh*-phrase changes its semantic behavior when two instances of it occur in the same INT. The reason why this contrast is relevant for my proposal is that CORs exhibit the very same pattern, as I discuss in the next section. This further similarity strengthens the semantic connections between INTs and CORs and provides further support for my proposal, according to which the two *wh*-clauses share the same semantic core and their *wh*-words are logically related.[33]

Karttunen's approach and my own version of it in § 4.2.1 cannot account for the pair-list answer in (36)c, but only for the single-pair answer in (36)b. In the remainder of this section, I develop my proposal to account for the pair-list answer by adopting the so-called "functional" approach.[34] The basic idea is that when two *wh*-phrases occur in the same INT, they can denote different kinds of semantic objects. The *wh*-phrase that is generated higher licenses the usual trace/variable over entities x_1 and is interpreted as in (13) in § 4.2.1. Let's call it an "ordinary" *wh*-phrase (*wh*-$_{[\text{INT}]}$). The *wh*-phrase that is generated lower, instead, licenses a complex variable $f_{2\langle e,e\rangle}(x_1)$ composed of a variable f_2 over Skolem functions (i.e., functions from entities to entities) and the variable x_1 licensed by the ordinary *wh*-phrase occurring as the argument of f_2. Therefore, I call this *wh*-phrase a "functional" *wh*-phrase (*wh*-$_{[\text{INT-f}]}$).[35] The setting of the licensed traces/variables makes it clear that the interpretation of the lower trace/variable depends on the semantic contribution of both *wh*-phrases. Let's go through the example in (39) to examine the details of the semantic derivation of an INT with two *wh*-phrases.

[33] Once again, many thanks to Gennaro Chierchia for enlightening remarks on this issue as well.
[34] This approach was originally championed by Engdahl (1980, 1986) and most extensively developed and defended in Dayal (1996). See Dayal (2016: Ch. 4) for a detailed overview, and Xiang (2023) for a recent take.
[35] As argued by Engdahl (1980, 1986) and Groenendijk and Stokhof (1982, 1983), variables over Skolem functions are independently required in the interpretation of *wh*-phrases by questions which allow for overt functional answers of the following kind:
(i) - Which professor does every student like?
    - His/her/their advisor.

39) a. Luca asked [which friend cooked which dish].

b.



c. [1]: CP ~> $\lambda x_1\ \lambda p\ [p = \lambda w.\text{cooked}(w)(f_2(x_1))(x_1)]$

[2]: *which* [INT] ~> $\lambda P_{\langle s,et\rangle}\ \lambda S\ \lambda q\ \exists x_1\ [P(w_0)(x_1) \wedge S(x_1)(q)]$

[3]: *friend* ~> $\text{friend}_{\langle s,et\rangle}$

[4]: DP_1 ~> $\lambda S\ \lambda q\ \exists x_1\ [\text{friend}(w_0)(x_1) \wedge S(x_1)(q)]$

[5]: CP ~> $\lambda q\ \exists x_1\ [\text{friend}(w_0)(x_1) \wedge q = \lambda w.\text{cooked}(w)(f_2(x_1))(x_1)]$

[6]: λ_2 ~> $\lambda Y_n\ \lambda f_2\ [Y]$

[7]: CP ~> $\lambda f_2\ \lambda q\ \exists x_1\ [\text{friend}(w_0)(x_1) \wedge q = \lambda w.\text{cooked}(w)(f_2(x_1))(x_1)]$

[8]: *which*$_2$ [INT-f] ~> $\lambda P\ \lambda F_{\langle\langle ee\rangle,\langle st,t\rangle\rangle}\ \lambda q_{\langle st\rangle}\ \exists f_2\ [q = \cap\lambda p[F(f_{2[P]})(p)]]$[36]

[9]: *dish* ~> $\text{dish}_{\langle s,et\rangle}$

[10]: DP_2 ~> $\lambda F\ \lambda q\ \exists f_2\ [q = \cap\lambda p[F(f_{2[\text{DISH}]})(p)]]$[37]

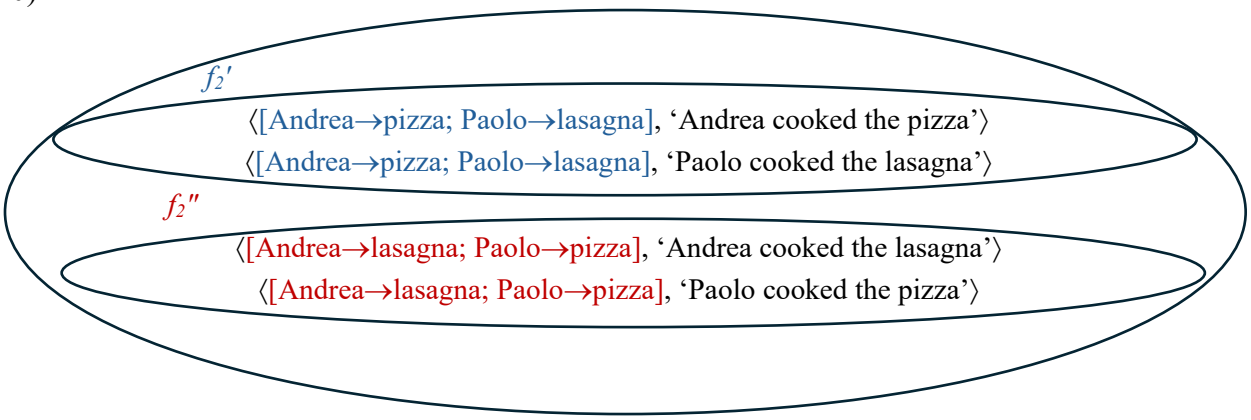
[11]: CP ~> $\lambda q \exists f_2[q=\cap\lambda p \exists x_1[\text{friend}(w_0)(x_1) \wedge p = \lambda w.\text{cooked}(w)(f_{2[\text{DISH}]}(x_1))(x_1)]]$

The *wh*-phrase w*hich friend* in (39)a is generated higher than *which dish*. Let's assume that it *wh*-moves first, while *which dish* moves second, so that *which dish* ends up c-commanding *which friend* and, therefore, entering the semantic derivation after *which friend*. The denotation of *which friend* in (39)c.[4] is the familiar one from INTs with a single *wh*-phrase: a function that applies to a relation between friends and propositions to return the subset of those propositions that are saturated by at least one friend, as shown in (39)c.[5]. The denotation of the actual *wh*-word *which* is a function from a property (the denotation of its nominal complement) to the denotation of its mother node—the full DP *which friend*. Note that all the propositions denoted by CP [5] are open propositions because their logical form contains the free functional variable f_2 over Skolem functions, which is licensed as part of the complex *wh*-trace/variable of the <u>*wh*</u>-phrase *which dish*. The next step in the semantic derivation is the abstraction over f_2 in preparation for the combination with the next *wh*-phrase, *which dish*. The result is a relation between Skolem functions f_2' and

[36] The fully detailed denotation is:
(i) $\lambda P\ \lambda F\ \lambda q\ \exists f_2\ [\forall x[f_2(x) \rightarrow P(w_0)(f_2(x))] \wedge q = \cap\lambda p.F(f_2)(p)]$
The subscript *P* in $f_{2[P]}$ is a shortcut for $\forall x[f_2(x) \rightarrow P(w_0)(f_2(x))$, i.e., the range of f_2.
[37] $f_{2[\text{DISH}]}$ is a shortcut for $\forall x[f_2(x) \rightarrow \text{dish}(w_0)(f_2(x))$,

propositions *q′*, as in (39)c.[7]. This relation holds if the following three conditions are satisfied: (i) all *q′* are derived from the propositional abstract[38] $\lambda w.cooked(w)(f_2(x_1))(x_1)$ once (ii) f_2' is assigned as the value of the functional variable f_2 in the propositional abstract, and (iii) there is at least one entity *e* that is a friend and is assigned as the value of x_1 in the propositional abstract. Let's illustrate this relation between functions and propositions within an extremely simplified toy model in which (i) Andrea and Paolo are the only friends, (ii) lasagna and pizza the only two dishes, and (iii) there are only two Skolem functions from friends to dishes: f_2' and f_2''. Then, the relation between functions and propositions in (39)c.[7] could be represented as in (40).

40)



Both functions f_2' and f_2'' map the set containing Andrea and Paolo onto the set containing lasagna and pizza, but the two mappings differ and therefore different propositions are generated. Each function generates a different set of propositions, resulting in a set of sets of propositions. The next step in the semantic derivation in (39) brings the denotation back to a set of propositions—which is the desired denotation to account for the semantic properties of INTs. Its denotation for the functional *wh*-phrase *which dish* in (39)c.[10] does exactly that, and more. It is a function that applies to a relation *F* between Skolem functions f_2 and sets of propositions *p* to return the set of propositions *q* that satisfy the following conditions. First, each proposition *q* results from the big intersection of all propositions generated by saturating each propositional abstract with one specific function f_2. Second, all f_2 ranges over dishes. Third, there is at least one such function f_2.

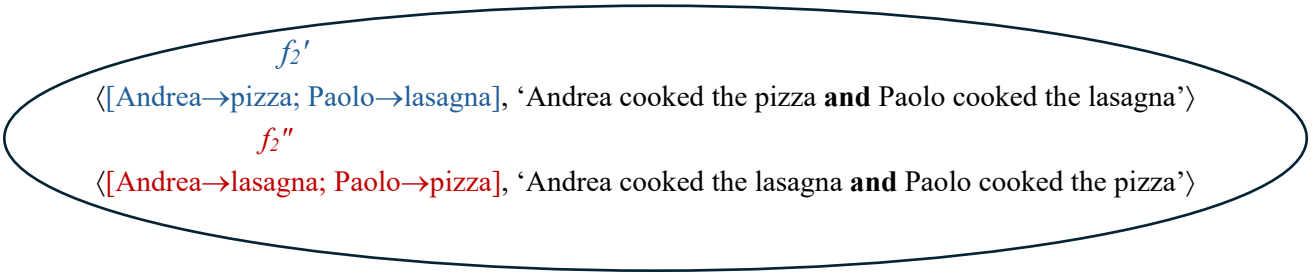
The first condition, the big intersection, can be seen as a type-adjustment that "flattens out" a set of propositions into their conjunction—a single proposition (Dayal 1996: Ch.4, §2.2.3).[39] This step

[38] I use the term "propositional abstract" for a logical expression like $\lambda w.cooked(w)(f_2(x))(x)$ in (39)c.[7] rather than "open proposition" because it has the type of a proposition (*st*) but contains variables that are bound within the larger logical expression of which it is part. By contrast, $\lambda w.cooked(w)(f_2(x))(x)$ in (39)c.[5] is an open proposition because its variables *x* and f_2 are free.

[39] Dayal assumes that a special complementizer is responsible for the "flattening" operation and for the existential quantification over the functional variable. I instead choose to lexicalize both operations in the denotation of functional *wh*-words.

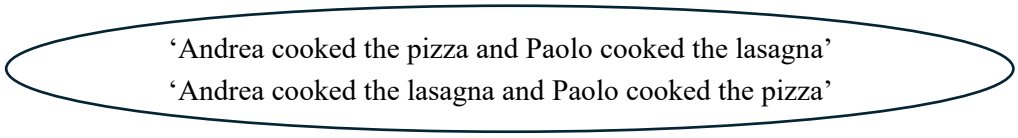
is illustrated in (41). The informational content of (40) and (41) is the same, although the "packaging" (semantic type) is different.

41)



The final component of the logical contribution of the functional *wh*-phrase *which dish* in (39) is to return the set of all and only the propositional members of the ordered pairs in (41), as shown in (42).

42)



The denotation of the functional *wh*-phrase *which dish* in (39)c.[10] is the instantiation of the general schema for the logical translation of functional *wh*-phrases in (43).

43) *wh*-phrase $_{[\text{INT-f}]}$ ~> $\lambda F_{\langle\langle ee\rangle,\langle st,t\rangle\rangle}\, \lambda q_{\langle st\rangle}\, \exists f_{\langle ee\rangle}\, [q = \cap\lambda p[F(f_{WH})(p)]]$[40]

To sum up, *wh*-phrases in INTs have an inherent dual semantic nature—the ordinary one in (13) and the functional one in (43). Either kind of *wh*-phrase can freely enter the semantic derivation at any point, although only one combination results in the correct semantic output—when the functional *wh*-phrase c-commands the ordinary *wh*-phrase after both undergo *wh*-movement. An INT with only functional *wh*-phrases would have the individual variables of its functional *wh*-phrases left unbound at the end of the derivation, resulting in a set of open propositions, rather than the required set of closed propositions. An INT with the ordinary *wh*-phrase c-commanding the functional *wh*-phrase would neutralize the contribution of the functional *wh*-phrase, making it semantically equivalent to the contribution of an ordinary *wh*-phrase (see fn. 31 for related

[40] The fully detailed denotation is:
(i) $\lambda F\, \lambda q\, \exists f_{\langle ee\rangle}\, [\forall \mathbf{x[f(x) \rightarrow WH(x)]} \wedge q = \cap\lambda p.F(f)(p)]$
The subscript *WH* in f_{WH} in (43) is simply shorthand for the boldface string in (i), i.e., the range of *f*. *WH* in *WH(x)* stands for the semantic restriction contributed either by *wh*-words that form their own *wh*-phrase (e.g., *who* humans, *what* things, etc.) or by the lexical material combined with *wh*-words that do not form a *wh*-phrase on their own (e.g.,, *food* in *which food*).

discussion). Two ordinary *wh*-phrases would not contribute to output a pair-list answer, but only a single-pair answer, as discussed at the beginning of this section.

While R-FRs and ∃-FRs with multiple *wh*-phrases do not exhibit clear functional dependencies (and, as shown in § 4.3.1 and § 4.3.2, can be handled without functional *wh*-words), CORs with multiple *wh*-phrases do exhibit readings that more closely resemble those of INTs with multiple *wh*-phrases and, therefore, call for a similar functional analysis. This is what I develop next.

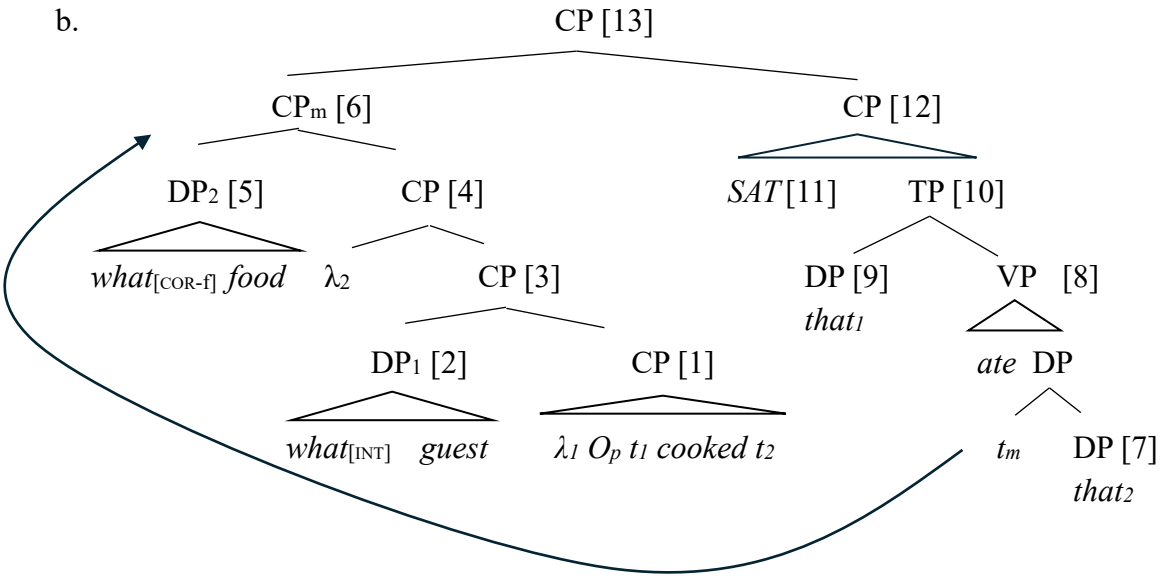
4.3.4. Semantic analysis of correlative clauses with multiple *wh*-phrases

In this section, I touch on the main semantic features of CORs with multiple *wh*-phrases and sketch how my approach can be extended to handle them as well. This section owes even more than the remainder of the paper to Gennaro Chierchia, in particular Chierchia (2025a, 2025b) and several virtual discussions. The syntactic structure that I adopt for CORs is Chierchia's, while my semantic analysis builds on his. It is fair to say that the best of this section is due to Chierchia, while I am the only one responsible for all remaining shortfalls and mistakes.

The example in (44)a is a Romanian COR with two *wh*-phrases with a singular count noun as complement. Its syntactic structure is given in (44)b and its semantic derivation in (44)c. Explanations and remarks follow.

44) a. [**[Ce invitat]$_1$ [ce mancare]$_2$** a gătit], **acela$_1$ aia$_2$** a mâncat.
what guest what food has cooked that that has eaten
'Every guest ate the food they cooked.'

b.



c. [1]: CP ~> $\lambda x_1\, \lambda p\, [p = \lambda w.\text{cooked}(w)(f_{2\langle ee\rangle}(x_1))(x_1)]$

[2]: DP_1 ~> $\lambda S\, \lambda q\, \exists x\, [\text{guest}(w_0)(x) \wedge S(x)(q)]$

[3]: CP ~> $\lambda q\, \exists x\, [\text{guest}(w_0)(x) \wedge q = \lambda w.\text{cooked}(w)(f_2(x))(x)]$

[4]: CP ~> $\lambda f_2\, \lambda q\, \exists x\, [\text{guest}(w_0)(x) \wedge q = \lambda w.\text{cooked}(w)(f_2(x))(x)]$

[5]: DP_2 ~> $\lambda F_{\langle\langle ee\rangle,\langle st,t\rangle\rangle}\, \iota f_{2\langle ee\rangle}\, \exists q_{\langle st\rangle}\, [q = \cap\lambda p[F(f_{2[\text{FOOD}]})(p)] \wedge q(w_0) \wedge \text{MAX}(f_2)]$[41]

[6]: CP ~> $\iota f_2\, \exists q\, \exists x\, [q = \cap\lambda p[\text{guest}(w_0)(x) \wedge p = \lambda w.\text{cooked}(w)(f_{2[\text{FOOD}]}(x))(x)] \wedge q(w_0) \wedge \text{MAX}(f_2)]]$ *in short* $\mathbf{f_{2|COR|}}$

[7]: *that*$_2$ ~> x_2

[8]: VP ~> $\lambda y.\text{ate}(w_0)(x_2)(y)$

[9]: *that*$_1$ ~> x_1

[10]: TP ~> $\text{ate}(w_0)(x_2)(x_1)$ ~> $\lambda x_1 \lambda x_2.\text{ate}(w_0)(x_2)(x_1)$[42]

[11]: SAT ~> $\lambda R\, \lambda f\, \forall x[x \in D(f) \rightarrow R(f(x))(x)]$

[12]: CP ~> $\lambda f\, \forall x[x \in D(f) \rightarrow \text{ate}(w_0)(f(x))(x)]$

[13]: CP ~> $\forall x[x \in D(\mathbf{f_{2|COR|}}) \rightarrow \text{ate}(w_0)(\mathbf{f_{2|COR|}}(x))(x)]$

The overall idea is that a COR with two *wh*-phrases exhibits a different semantic behavior from a COR with just one *wh*-phrase in (i) its denotation and the denotation of its *wh*-phrases, and (ii) the way it semantically combines with its matrix clause. I discuss these differences next, while highlighting features in the semantic derivation in (44)c. I start with the differences in denotation.

Denotation differences. Following Chierchia (2025a), a COR with two *wh*-phrases denotes a specific function *f* of type *⟨e,e⟩*: the unique maximal function *f* that maps entities that saturate the higher free variable (licensed by one of the *wh*-phrases) with entities that saturate the lower free variable (licensed by the other *wh*-phrase). So, the COR in (44) ends up denoting the maximal function from friends to food they cooked, as shown in (44)c.[6]. This denotation results from combining the ordinary *wh*-phrase first, and then the functional *wh*-phrase. This is the same order as in INTs with two *wh*-phrases.[43] The ordinary *wh*-phrase has the same denotation as an ordinary *wh*-phrase in an INT, as shown in (39)c.[4]. The functional *wh*-phrase in a COR (*wh* [f-COR]), instead, is assigned a slightly different denotation than the one assigned to the functional *wh*-phrase in an INT (cf. (43)). The new denotation for functional *wh*-phrases in CORs is schematized in (45).

[41] I have omitted the trivial step in which λ-abstraction over the individual variable *x* takes place. *MAX(f)* is shorthand for a "maximized singleton property" of functions, true of the largest function of a given set of functions—the function with most set of pairs that saturates the open proposition. See the discussion below ex. (45) in the main text. The subscript *FOOD* in $f_{2[FOOD]}$ is shirthand for $\forall x[f_2(x) \rightarrow food(w_0)(f_2(x))$, i.e., the range of f_2.

[42] I'm omitting the trivial step(s) of *λ*-abstraction over the two variables.

[43] The order of combination is the reverse of the order in which the *wh*-phrase obligatorily align in Romanian. I take the surface order to be due to factors other than interpretation alone and assume that reordering can occur before interpretation takes place, yielding the syntactic tree in (44)b.

45) *wh*-phrase $_{[\text{COR-f}]}$ $\rightsquigarrow > \lambda F_{\langle\langle ee\rangle,\langle st,t\rangle\rangle}\, \iota f_{\langle ee\rangle}\, \exists q_{\langle st\rangle}\, [q = \cap\lambda p[F(f_{WH})(p)] \wedge q(w_0) \wedge MAX(f)]$

The denotation in (45) is a function (from relations *F* between Skolem functions and propositions) that returns the unique "maximal function" f_{MAX} out of the set of Skolem functions *f*, rather than existentially quantifying over the set of functions *f*. f_{MAX} is the function with the largest extension of ordered pairs, among the many functions that truthfully saturate the propositional abstract *p* (and therefore *q* as well). Returning to the example in (44), let's again assume a toy model with three friends and three pieces of food as the only relevant entities. Let's also assume that in the world of evaluation w_0 the actual facts are as follows: Andrea cooked and ate the pizza, Paolo cooked and ate the lasagna, and Luca cooked and ate the pasta. Many functions can map friends who cooked food to the food they cooked. A sample is given in (46).

46) *Skolem functions from friends who cooked food to the food they cooked*

f_1	f_2	f_3	f_4
Andrea → pizza	Andrea → **pasta**	Andrea → pizza	Paolo → lasagna
Paolo → lasagna	Paolo → lasagna	Paolo → lasagna	Luca → pasta
Luca → pasta	Luca → **pizza**		

The domains of both f_3 and f_4 lack one human being among the those in the domain of the model: f_3 is missing Luca, while f_4 is missing Andrea. This is why neither f_3 nor f_4 can be the maximal function, i.e., neither of them satisfies the property *MAX* in (44)c.[7]. Functions f_1 and f_2, instead, cover the full domain of human beings in the model. But only f_1 outputs values that truthfully saturate the propositional abstract 'x cooked f(x)' in (44)c.[7] in w_0.[44] The values that f_2 outputs for Andrea and Luca (in boldface in (46)) saturate 'x cooked f(x)' but produce the propositions 'Andrea cooked pasta' and 'Luca cooked pizza', which are false in w_0 (i.e., they do not satisfy the requirement $q(w_0)$ in (44)c.[7]). Therefore, f_1 is the maximal function among those. It can be proved that it exists and is unique.

Combinatory differences. A COR with two *wh*-phrases also differs from a COR with one *wh*-phrase in the way it semantically combines with its matrix clause. The matrix clause of a COR with two *wh*-phrases denotes a 2-place relation *R* over entities. It results from abstracting over the variables x_1 and x_2 that are licensed by the two demonstratives in the matrix clause that are anchored to the two *wh*-phrases in the COR. For instance, the matrix clause in (44) denotes the 2-place relation between friend entities and food entities, as shown in (44)c.[11]. In order to combine the denotation of the COR (a function from entities to entities) with the denotation of its matrix

[44] To be fully precise, it is not the saturated propositional abstract *p* that is required to be true in (44)c.[7], but *q*—the proposition resulting from the big intersection of all *p* (see $q(w_0)$ *in* (44)c.[7]). Given the truth-conditions of big intersection, *q* entails *p*.

clause (a 2-place relation over entities), I adopt (some version of) the semantic rule *Saturation* proposed by Chierchia (2025a: (13)b.iv) with the formulation in (47).

47) *Saturation Rule (based on Chierchia 2025a)*
$\lambda R_{\langle e,et\rangle}\, \lambda f_{\langle ee\rangle}\, \forall x[x \in D(f) \rightarrow R(f(x))(x)]$

Saturation is a higher-order relation between relations *R* between entities and functions *f* from entities to entities. Saturation returns the truth if *R* is true for every entity in the domain of *f* and for every corresponding value of *f*. In other words, *R* ensures that the mapping or functional dependency that is produced by the COR also holds for the matrix clause. What this means for the sentence in (44) is that all the ordered pairs $\langle e_1,e_2\rangle$ with e_1 a friend who prepared food and e_2 the food e_1 prepared also have to be ordered pairs $\langle e_1,e_2\rangle$ of eaters e_1 and the food they ate e_2. In other words, the extension of *f* is included in the domain of ordered pairs of *R*.

The Saturation operator *SAT* is assigned its own functional projection in the left periphery of the matrix clause, SatP. This choice marks a further difference between a COR with two *wh*-phrases and a COR with just one, since the latter is hosted under TopP. Unlike single-*wh* CORs, multiple-*wh* CORs do not function as an ordinary entity-denoting topic.

These differences in the denotations of *wh*-words and rules of combinations (Function Application vs Saturation) in single-*wh* vs multiple-*wh* CORs make the two constructions partially independent and make at least one empirical prediction: it may be possible for a language to have one kind of COR or COR-like construction without having the other. This prediction may be borne out in Romanian, if Rudin constructions are considered together with their requirement for at least two *wh*-phrases (Caponigro and Fălăuș 2022)

5. Varieties of *wh*-words and their logically related meanings

I have proposed a unified account for four different *wh*-clauses beginning from the same syntactic structure and semantic denotation: a CP denoting a relation between entities and propositions. I have shown that the semantic differences among *wh*-clauses can be derived from minimal variations in the logical content of their *wh*-words, starting from the denotations of *wh*- words in INTs. I have taken *wh*-words in INTs as primary because of three main reasons: the available diachronic evidence, the extremely common use of *wh*-clauses to convey question meaning across languages, and the fact that the set of *wh*-words in INTs is larger and often a superset of the sets of *wh*-words in all other *wh*-clauses. In this section, I summarize the logical changes from the denotation of *wh*-words in INTs to those in all other *wh*-clauses investigated to highlight their similarity, regularity, and plausibility.

All non-interrogative *wh*-words impose the requirement that the propositional content of their *wh*-clause be true at the world of evaluation—their saturated propositional abstract *q* to be true at w_0 ($q(w_0)$). Let's label this requirement *True Proposition* in the schematic summaries in (48)–(51) below. I have already discussed why this requirement is natural and necessary for the denotation

of Non-INTs, while too restrictive for INTs, in the paragraph *True propositional content* at the very end of § 4.2.2. The gist of that discussion is that the denotation of non-interrogative *wh*-words is inherently extensional: it is about entities having certain properties that saturate propositional abstracts, making them true in the world of evaluation w_0.

∃-switch and True Proposition are enough to account for the meaning change in *wh*-words from INTs to ∃-FRs, as summarized in (48).[45]

48) *Changes in the meaning of* wh*-words from INTs to ∃-FRs:*

$wh_{[INT]} \rightsquigarrow \lambda S\, \lambda q\, \exists x\, [WH(w_0)(x) \wedge S(x)(q)]$

⇩ + **∃-switch**
⇩ + **True Proposition**

$wh_{[\exists\text{-FR}]} \rightsquigarrow \lambda S\, \boldsymbol{\lambda x\, \exists q}\, [WH(w_0)(x) \wedge S(x)(q) \wedge \mathbf{q(w_0)}]$

The next logical change is needed to account for the logical content of *wh*-words in R-FRs. *Entity* is essentially adding to *wh*-words in R-FRs the logical content of a definite determiner in languages like Italian and Romanian, as summarized in (49). Since ∃-switch makes available the set of entities by switching the existential quantification over the set of propositions, the operators ι or ∩ can apply to the set of entities and return the maximal entity or the kind. See the examples and related discussion in § 4.2.2 for further details.

49) *Changes in the meaning of* wh*-words from INTs to R-FRs:*

$wh_{[INT]} \rightsquigarrow \lambda S\, \lambda q\, \exists x\, [WH(w_0)(x) \wedge S(x)(q)]$

⇩ + **∃-switch**
⇩ + **True Proposition**
⇩ + **Entity**

$wh_{[R\text{-FR}]} \rightsquigarrow \lambda S_{\langle e,\langle st,t\rangle\rangle}\, \boldsymbol{\iota x \exists q} \forall y([WH(w_0)(y) \wedge S(y)(q) \wedge \mathbf{q(w_0)}] \rightarrow \mathbf{y \leq x})$

$wh_{[R\text{-FR-}\cap]} \rightsquigarrow \lambda S\, \boldsymbol{\cap \lambda x\, \exists q}\, [WH(w_0)(x) \wedge S(x)(q) \wedge \mathbf{q(w_0)}]$

The final logical change for ordinary *wh*-words is the one needed for *wh*-words in CORs. Once again, the added logical content is an operation that is independently needed in the grammar, at least if we follow Montague's treatment of referential expressions as generalized quantifiers. *Wh*-words in CORs essentially build generalized quantifiers out of the maximal entities or kinds that are denoted by the corresponding R-FRs, as summarized in (50).

[45] From (48) onward, I use the same color for both the label of a semantic change and the actual change inside the logical translation of the *wh*-phrase.

50) *Changes in the meaning of* wh-*words from INTs to CORs:*

$wh_{[INT]} \sim> \lambda S\, \lambda q\, \exists x\, [WH(w_0)(x) \wedge S(x)(q)]$

+ ∃-switch
+ True Proposition
+ Entity
+ GQ

$wh_{[COR\text{-}\iota]} \sim> \lambda S\, \mathbf{\lambda P_{et}.P(\iota x \exists q} \forall y([WH(w_0)(y) \wedge S(y)(q) \wedge \mathbf{q(w_0)}] \rightarrow \mathbf{y \leq x}))$

$wh_{[COR\text{-}\cap]} \sim> \lambda S\, \mathbf{\lambda P_{\langle se,t\rangle}.P(\cap \lambda x[\exists q}[[WH(w_0)(x) \wedge S(x)(q) \wedge \mathbf{q(w_0)}])$

Let's move to the meaning changes concerning functional *wh*-words from INTs to CORs. In addition to the ∃-switch and True Proposition, *Function* is needed as well, as schematized in (51).

51) *Changes in the meaning of functional* wh-*words in CORs:*

$wh_{[f]} \sim> \lambda F\, \lambda q_{\langle st\rangle}\, \exists f_{\langle ee\rangle}\, [q = \cap \lambda p[F(f_{WH})(p)]]$

+ ∃-switch
+ True Proposition
+ Function

$wh_{[COR\text{-}f]} \sim> \lambda F\, \mathbf{\iota f_{\langle ee\rangle}\, \exists q_{\langle st\rangle}}\, [q = \cap \lambda p[F(f_{WH})(p)] \wedge \mathbf{q(w_0)} \wedge \mathbf{MAX(f)}]$

Since ∃-switch makes available the set of functions by switching the existential quantification over the set of propositions, the operators ι can apply to the set of functions and return the maximal function, as discussed in § 4.3.4. Its logical contribution resembles the one of Entity for ordinary *wh*-words in R-FRs: in both cases, the largest semantic object is selected out of the given set.

The meaning changes for *wh*-words that I have argued for throughout the paper and summarized in this section are built out of shared logical building blocks. Some meaning changes add more logical building blocks than others. Still, this should not be understood as implying that (i) the *wh*-word meaning changes with fewer logical additions are expected to happen more frequently than those with more logical additions nor that (ii) the *wh*-word meaning changes with just logical addition *1* are a necessary condition for those *wh*-word meaning changes requiring logical addition *1* and *2*. There is no conceptual need to assume that logical components pile up progressively or that smaller clusters of logical components should be favored over larger.

There is no factual evidence either (see Table 2). For instance, ∃-FRs require the minimum number of logical changes, as discussed in (48), while R-FRs add an extra logical requirement, as showed in (49). Still, English has R-FRs, but completely lacks ∃-FRs. Italian has both but the *wh*-words occurring in R-FRs form a larger and disjoined set from those occurring in ∃-FRs. Romanian is productive in all its *wh*-clauses. Still, CORs use all the *wh*-words in INTs, while R-FRs only use a

proper subset. This slight difference in productivity of *wh*-words in CORs and R-FRs in Romanian does not correlate with the number of logical changes their *wh*-words have undergone. In fact, the logical changes in the *wh*-words in R-FRs for a proper subset of those in CORs, as apparent by comparing (49) with (50) above.

It should be kept in mind that many factors may participate in the licensing of a certain kind of Non-INT in a language, including morphosyntactic factors like the licensing of *wh*-clauses in argument and/or adjunct and/or dislocated positions or diachronic factors (e.g., the loss of CORs in Modern English in parallel to the rise of headed relative clauses).

6. Accounting for the three generalizations

I have made brief reference throughout the paper to how my analysis accounts for the generalizations in § 3. In this section, I bring those remarks together and summarize the results.

Generalization 1 (§ 3: (6)) concerns the distribution of *wh*-clauses across languages. My analysis assumes that the CP of INTs and its denotation (a relation between entities and propositions) are the core syntactic and semantic building blocks for all Non-INTs. This immediately accounts for Part (i) of Generalization 1: a language can only have Non-INTs if it has INTs in the first place. At the same time, the analysis does not predict any implicational hierarchy among Non-INTs since the meaning changes required in their *wh*-words are independent of one another. Non-semantic factors may also condition the availability of different varieties of non-interrogatives *wh*-clauses in a language, as discussed in § 4. This conclusion aligns with Part (ii) of Generalization 1.

Generalization 2 (§ 3: (7)) concerns the distribution of *wh*-words across varieties of *wh*-clauses and languages. My analysis captures Part (i) of Generalization 2 because *wh*-words in INTs serve as the primary source to which various logically-driven meaning changes apply before *wh*-words can occur in Non-INTs. At the same time, the analysis imposes no dependency among these meaning changes and requires every change to be licensed *wh*-word by *wh*-word in a given language—even within the same variety of *wh*-clause. This predicts, and is fully compatible with, the interlinguistic and crosslinguistic patterns noted in Part (ii) and Part (iii) of Generalization 2.

Generalization 3 (§ 3: (8)) addresses the correlation between the availability of INTs with multiple *wh*-phrases and the availability of Non-INTs with multiple *wh*-phrases, including the distribution and ordering restriction of their *wh*-phrases. Those correlations are expected under my analysis, since all Non-INTs are derived from the same core CP structure as INTs and their *wh*-words are derived from the *wh*-words in INTs via systematic meaning changes.

7. Problems with alternative approaches

I am not aware of any previous attempt to provide a uniform account of all four varieties of *wh*-clauses I have investigated, nor of the three generalizations I have presented. There have been analyses of only one variety of those *wh*-clauses or a proper subset of them, while parts of the generalizations may have been hinted at, but never formulated clearly, let alone all together. In this

section, I briefly discuss some of those previous attempts to highlight difficulties they face, including in handling the three generalizations I have presented.

Although with some differences, Jacobson (1995), Dayal (1996), and Caponigro (2003, 2004) all pursue an analysis of R-FRs that treats *wh*-clauses as set/property denoting and *wh*-phrases as restrictors of those sets. *Iota* type-shifting would then ensure that R-FRs end up as referential by denoting maximal entities. Jacobson sketches how to extend this analysis to INTs, building on Groenendijck and Stokhof's (1982) treatment of ***wh*-clauses as property denoting**. Different type-shifting rules would then apply and make INTs denote (sets of) propositions, while making R-FRs demote entities. Caponigro (2003, 2004) extends this approach to ∃-FRs by arguing the basic set/property denoting *wh*-clause directly combines as the complement of existential predicates. The gist of this line of analysis can be exemplified as in (52) with the *wh*-clause *who came*.

52) Wh-*clauses as denoting sets/properties:*

$\lambda p\, \exists x\, [p = \lambda w[WH(w)(x) \wedge P(w)(x)]]$ $\quad$ $\iota x[WH(w_0)(x) \wedge P(w_0)(x)]$

Question Type-Shifting ⇖ ⇗ **Iota Type-Shifting**

$\lambda x[WH(w_0)(x) \wedge P(w_0)(x)]$

⇓ **∃-predicate**

$\exists x[WH(w_0)(x) \wedge P(w_0)(x)]$

Although it wouldn't be difficult to extend this approach to CORs, it faces major issues with all three generalizations. As schematized in (52), its core syntactic and semantic building block is a kind of *wh*-clause root to which different semantic rules or type-shifters apply to derive the different meanings. There is no reason why there should be an asymmetry between INTs and their *wh*-words on the one hand and Non-INTs on the other—let alone reasons why INTs and their *wh*-words should have a priority status. Therefore, it predicts there should be languages where the set of *wh*-words in INTs is a proper subset of those in Non-INTs. This prediction is not borne out and clashes with the generalizations I have presented.

Another problematic feature of this approach is that it crucially hinges on general semantic rules or type-shifters that uniformly apply to all *wh*-clauses of the same kind, regardless of the specific *wh*-word. Therefore, it cannot straightforwardly explain Generalization 2, which highlights the variations among *wh*-words in different *wh*-clauses within and across languages. An additional ad hoc mechanism would be needed to handle this variation.

The same objection applies to an approach that would treat ***wh*-phrases as existentially quantified expressions**, following Karttunen's (1977) seminal analysis of INTs and the large body of work

that it has inspired. Let's assume that INTs are primary and all Non-INTs are derived from them by means of general semantic rules that are insensitive to specific *wh*-words. Let's also assume that all *wh*-words share the same meaning as existentially quantified expressions. How, then, could we explain the variation in distribution of *wh*-words within a language and across languages that Generalization 2 clearly establishes?

One argument that may be invoked in support of an analysis of *wh*-phrases as indefinite or existentially quantified expressions is the **use of *wh*-words outside *wh*-clauses**. Languages that allow such use treat *wh*-words semantically as indefinites of some sort. Haspelmath (1997, 2013) discusses this fact extensively. The emerging picture is complex: 194 of the 326 languages surveyed use *wh*-words either as indefinites or as roots to form various indefinites (e.g., *some-where*, *any-where*, *where-ever* in English). On the other hand, Haspelmath (1997: 174–176) notes that *wh*-words (i) "prove resistant to etymological analysis" and (ii) "are among the slowest-changing elements in any language". Haspelmath (2013) adds that "[t]he interrogative function is always primary, and the indefinite function is secondary. An interrogative pronoun may lose its indefinite function." So there is no evidence that *wh*-words are derived from indefinites or are "natural indefinites" that can be freely used outside *wh*-clauses. There is, however, evidence that *wh*-words can serve as roots for building words with different semantic properties, when combined with appropriate affixes (e.g., *some-where*, *every-where, no-where* in English).

There are languages that have a rich system of Non-INTs that nevertheless do not allow bare *wh*-roots to occur outside *wh*-clauses—English, Italian, and Romanian among those. At the same time, there are languages like Japanese, Korean, and Mandarin that allow for *wh*-roots in non-*wh*-clauses but lack the Non-INTs I have investigated.

My analysis provides a principled explanation for both the possibility of using *wh*-words as (roots for) indefinites and the fact that this use is orthogonal to their use in Non-INTs, as attested by languages that allow for one use but not the other. I assume the meaning of *wh*-words repeated in (53), as primary, with all other meanings derived via systematic logical changes (see § 5).

53) $wh_{\text{[INT]}} \rightsquigarrow \lambda S\, \lambda q\, \exists x\, [WH(w_0)(x) \wedge S(x)(q)]$

All *wh*-word meanings in both INTs and Non-INTs share propositional content. What varies is how they manipulate that content: they either abstract over a propositional variable to create a set of propositions—as in INTs—or existentially close the proposition variable to abstract over the entity variable and create a set of entities—the starting point to derive the denotation of Non-INTs. Still, one could imagine a third path of logical change in which the denotation in (53) loses its propositional content entirely, turning into (54), a plain existential generalized quantifier:

54) $wh_{\text{[INDEFINITE]}} \rightsquigarrow \lambda P_{et}\, \exists x\, [WH(w_0)(x) \wedge P(x)]$

A further logical step would strip (53) of quantificational force entirely, yielding (55), a simple set-denoting expression—a flexible root for various quantificational affixes.

55) *wh* [ROOT] ~> λx [WH(w_0)(x)]

The logical changes from the denotation of *wh*-word in INTs in (53) to those in (54) and (55) are independent from the changes associated with Non-INTs discussed in § 4. A language may pursue both paths, neither, or one but not the other. Tigrinya shows no use of *wh*-words outside INTs; Japanese pursues only the indefinite/root path; Italian mainly pursues the *wh*-clause path; Romanian may pursue both, as it has a rich inventory of words based on *wh*-roots, in addition to all varieties of *wh*-clauses (Fălăuş and Nicolae 2022).

Šimík (2026) sketches a broad and ambitious **"layered" morphosyntactic and semantic analysis of *wh*-clauses**. I cannot discuss his proposal in detail for reasons of space and because many aspects remain only sketched. I mention a few central features to highlight differences with my approach. According to Šimík (2026), all *wh*-clauses share a core clausal syntactic structure, and all *wh*-words share a core morphosyntactic structure. Functional layers are added the clausal core and the word core in order to account for the different varieties of *wh*-clauses, along the hierarchy in (56).

56) Wh-*hierarchy* (Šimík 2026)
INT < Unconditional < COR < R-FR < Light-Headed Relative < Headed Relative[46]

The *wh*-clauses from INT up to COR in (56) are argued to be propositional, although each combines with its matrix clause via different operators. Their *wh*-words do not trigger λ-abstraction or act as set restrictors or quantifiers. The *wh*-clauses in the right half of the hierarchy in (56), instead, are property-denoting and their *wh*-words license *wh*-traces/variable, λ-abstraction, and set formation. The main motivation for this dichotomy is the generalization that the *wh*-clauses in the left half of the hierarchy never allow for *wh*-in situ. However, counterexamples exist for headed relative clauses and R-FRs, as Šimík acknowledges (Šimík 2026: § 25.5.1). Moreover, while ∃-FRs fit my analysis without difficulty, they cannot be easily placed in Šimík's hierarchy, as he himself notes (Šimík 2026: 851–852). Their morphosyntax makes them resemble INTs, but their semantics is property-based, like the one of R-FRs and headed relative clauses. A similar issue arises for *-ever* free relative clauses. Finally, while my unified approach directly accounts for Generalzation 3 concerning *wh*-clauses with multiple *wh*-phrases, it is unclear how Šimík's dichotomic approach would explain this shared patterning across the two sides of his hierarchy. Still, I find the direction of Simik's investigation extremely stimulating and enriching, especially the impressive crosslinguistic data and the clarity of organization.

8. Conclusions

I have argued that all Non-INTs that are introduced by (a subset of) the *wh*-words that introduce INTs should be analyzed as sharing not only their *wh*-words, but also functional layers in their

[46] I adapted Šimík's formulation by using my own abbreviations for the *wh*-clauses we both discuss and by replacing his abbreviations with the full label for the *wh*-clauses that are not addressed in this paper.

syntactic structures and a propositional component in their denotation. Specifically, I have proposed that all *wh*-clauses share the syntactic structure and the logical translation in (10).

Wh-clauses begin to diverge semantically (and possibly syntactically) when the (first) *wh*-phrase enters the semantic derivation. I have argued that *wh*-words in different *wh*-clauses carry distinct lexical content, which crucially affects the meaning of the clause. Still, the denotational differences among *wh*-words are minimal and fully determined by general logical principles (§ 5). I have also briefly discussed why *wh*-words can often be used on their own to form indefinites or other quantificational expressions across languages, including languages lacking Non-INTs (final part of § 7).

The propositional analysis I have developed offers a principled and unified semantic account for four types of *wh*-clauses attested crosslinguistically—INTs, R-FRs, ∃-FRs, and CORs with one or multiple *wh*-phrase (§ 4)—as well as other Non-INTs that I cannot discuss here in detail for reasons of space. For instance, the *wh*-clauses in Rudin constructions (Caponigro and Fălăuș 2022) can be analyzed in exactly the same ways as CORs with multiple *wh*-phrases in § 4.3.4. Restrictive headed relative clauses would require the logical meaning in (60) for their *wh*-words—a function that applies to the usual relation *S* between entities and open propositions to return a set restrictor that in turn combines with the set-denoting nominal head of the headed relative clause to restrict it.

57) *wh* $_{[\textit{Headed}\ \text{RC}]}$ $\sim>$ $\lambda S_{\langle e,\langle st,t\rangle\rangle}$ $\boldsymbol{\lambda P_{\langle s,et\rangle}}$ $\lambda x\ \exists q\ [WH(w_0)(x) \wedge S(x)(q) \wedge q(w_0) \wedge \mathbf{P(w_0)(x)}]$

Note that (60) is derived from the meaning of the *wh*-words in INTs by the same logical changes employed for ∃-FRs (∃-switch and True Proposition; § 5), with the addition of the boldfaced component required for the *wh*-clause to combine with its nominal head.

I have shown that the propositional analysis can also account for the three generalizations about *wh*-clauses and *wh*-words within and across languages that I have presented (§§ 3, 6), without encountering the difficulties faced by existing approaches (§ 7).

My propositional analysis of Non-INTs is primarily justified by the similarity, and sometimes full identity, between INTs and Non-INTs, beginning with their shared use of *wh*-words. However, some languages display clausal constructions whose distribution and interpretation resemble those of Non-INTs, despite not being *wh*-clauses morphosyntactically. The examples in (57)a and (58)a illustrate what Caponigro et al. (2021) label *Super-Free Relative Clauses*: clauses whose distribution and interpretation resemble those of R-FRs or ∃-FRs, but lack *wh*-words and therefore don't qualify as *wh*-clauses. The bracketed clause in (57)a is introduced by a complementizer (*den*), has a missing object, and exhibits the same distribution and interpretation as the R-FR in Italian in (57)b. Similarly, the bracketed clause in (58)a is introduced by a non-*wh* relativizer (*pV?*), has a missing subject, and shares its distribution and interpretation with the ∃-FR in Italian (58)b.

58) a. y=o-Ø-asi-ko [den i-pampa o-ti-choka-ya][47] *Tlaxcala Náhuatl*
already=PST-S3-arrive-VEN.PST COMP POSS3S-RSN PST-S2SG-cry-IPFV
'The one you cried for has already arrived.'

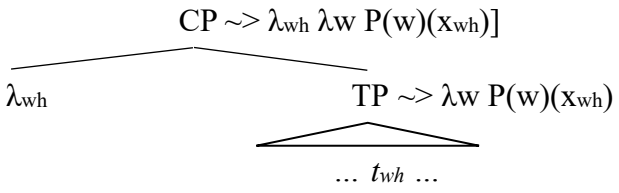
b. È gia arrivato [chi desideravi]. *Italian*
is already arrived who desired-2SG
'The one you desired has already arrived.'

59) a. Ø=ʔit-wɨ [ta=kuʔaʔm-ket-neʔ-wɨ=pVʔ][48] *Sierra Popoluca*
3ABS=be-COMPL 1ABS:INCL=search-descend-PERF-COMPL=REL
'There's someone who looks after us.'

b. C'è [chi si prende cura di noi]. *Italian*
there's who REFL.3 takes care of us
'There's someone who is looking after us.'

Although technically possible, I see no empirical evidence nor principled reasons to analyze the constructions in (57)a and (58)a as "propositional". It is simpler to assume that abstraction applies over the free individual variable licensed by the missing constituent, yielding a set of entities, as argued in Caponigro (2023) and schematized in (59).

60) *Schema of semantic derivation of Super-Free Relative Clauses*



Maximalization through type-shifting applies to the bracketed clause in (57)a, while existential quantification triggered by the matrix predicate would apply to the bracketed clause in (58)a.

Abaza and Adyghe, two Northwest Caucasian languages, provide even stronger evidence for a non-propositional option to be available in the grammar, alongside with a propositional one (Caponigro and Polinski 2011; Arkadiev and Caponigro 2021). Neither language permits embedded interrogative *wh*-clauses; instead, they use non-*wh* relative clauses to convey the meaning that is typically conveyed by embedded interrogative clauses in languages like English. Abaza even lacks true *wh*-words in general and relies on relative constructions to convey the meaning of matrix interrogative clause in English. Consequently, there is no evidence that Abaza and Adyghe ever employ the structure and denotation in (10).

[47] Adapted from Flores-Nájera (2021: ex. 66). Tlaxcala Náhuatl is a Mesoamerican language of the Uto-Aztecan family.
[48] Adapted from López Márquez (2021: ex. 89b). Sierra Popoluca is a Mesoamerican language of the Mixe-Zoquean family.

I expect languages to freely employ the propositional/*wh* option in (10), the non-propositional/non-*wh* option in (59), or both. It is also plausible that a language might initially assign the interrogative structure and propositional content in (10) to a *wh*-clause and then, over time, reanalyze it as the non-interrogative, non-propositional option in (59), driven by morphosyntactic changes in the *wh*-words or by a reduction in the number of *wh*-words permitted in the construction.

The broader conclusion my investigation brings to the linguistic table for further exploration is that the grammar is endowed with two distinct strategies for using clauses with missing (or specially marked) constituents and exploiting their informational richness to refer to individuals, or quantify over individuals, or denotes sets/properties of individuals. One—the non-propositional strategy—applies to clauses without *wh*-words[49] and forms sets of individuals via abstraction over the variable licensed by the missing (or specially marked) constituent, as in (60). At no point in the semantic derivation does abstraction or quantification over propositions take place. This is the strategy proposed by Quine (1960) and Montague (1970) for headed relative clauses in English, and it has since been applied to a range of other clauses. The other strategy that is attested in the grammar is the propositional strategy that I have developed. It starts from INTs—clauses characterized by a specific class of lexical items (*wh*-words) and by the semantic property of always conveying a question (i.e., sets of propositions). Through logic-driven changes in the meaning of their *wh*-words, INTs can shift their denotation and distribution, turning into different kinds of Non-INTs, depending on the language and the specific *wh*-word involved.

Acknowledgements

This paper would never have been conceived without Gennaro Chierchia's crucial contribution, both with his work (Chierchia 2025a, 2025b) and his generous interaction in person and virtually. Among many debts, I owe him the crucial idea of a propositional core for the denotation of correlative clauses and some of the main formal solutions he has developed to implement this idea. He was the one who managed to convince me that it made sense to pursue this idea, despite my initial skepticism. He deserves full credit for all this and more. On the other hand, I'm the only one to blame for any problems or shortfalls in the way I have implemented his idea and extended to many more kinds of *wh*-clauses than he had intended. Many thanks to Anamaria Fălăuş for generously sharing her knowledge of Romanian for being the main force behind the workshop in Nantes where this research started. Thanks to Veneeta Dayal for the discussion at the workshop in Nantes and for encouraging me to write this paper.

[49] Or, at least, without *wh*-words that are identical to—or productively resemble—those found in INTs.

References

Arkadiev, Peter and Ivano Caponigro. 2021. Conveying content questions without *wh*-words: evidence from Abaza. In P. Grosz, L. Martí, H. Pearson, Y. Sudo, and S. Zobel (eds), *Proceedings of Sinn und Bedeutung 25*, 73–94.

Belyaev, Oleg and Dag Haug. 2020. The genesis and typology of correlatives. *Language* 96(4):874–907.

Bhatt, Rajesh. 2003. Locality in correlatives. *Natural Language and Linguistic Theory* 21(3):485–541.

Brasoveanu, Adrian. 2012. Correlatives. *Language and Linguistics Compass* 6: 1–20.

Cacchioli, Gioia and Ivano Caponigro. 2026. Varieties of Headless Relative Clauses in Tigrinya. Ms. University of Geneva and University of California San Diego.

Caponigro, Ivano. 2001. On the semantics of indefinite free relatives. In *Proceedings of the 10th Meeting of the Student Organization of Linguistics in Europe (ConSOLE X)*, eds M. van Koppen, J. Sio, and M. de Vos, 49–62. Leiden: SOLE.

Caponigro, Ivano. 2003. *Free not to ask: on the semantics of free relatives and wh-words cross-linguistically*. PhD diss., University of California, Los Angeles.

Caponigro, Ivano. 2004. The semantic contribution of *wh*-words and type shifts: evidence from free relatives cross-linguistically. In *Proceedings of the 14th Semantics and Linguistic Theory*, ed. R. B. Young, 38–55. CLC Publications.

Caponigro, Ivano. 2021. Introducing Headless Relative Clauses and the findings from Mesoamerican languages. In *Headless relative clauses in Mesoamerican languages*, eds I. Caponigro, H. Torrence & R. Zavala Maldonado, 1–57. New York: Oxford University.

Caponigro, Ivano. 2022. Headless relative clauses and the syntax-semantics mapping: evidence from Mesoamerica. In *SULA 11: Proceedings of the 11th Semantics of Under-represented Languages in the Americas*, eds S. S. Lee & Y. Song, 27-42. GLSA, Linguistics Department, University of Massachusetts at Amherst.

Caponigro, Ivano. 2023. Referring and quantifying without nominals: Headless Relative Clauses across languages. In *Proceedings of the 33rd Semantics and Linguistic Theory Conference (SALT 33)*, eds J. Kim, B. Öney, Y. Zhang, and F. Zhao, 744–774. Linguistic Society of America.

Caponigro, Ivano. 2025. Still free to have a *wh*-phrase: a reply to Donati, Foppolo, Konrad, and Cecchetto (2022). *Linguistic Inquiry*. 56(2):374–400.

Caponigro, Ivano and Anamaria Fălăuş. 2018. Free Choice Free Relative Clauses in Italian and Romanian. *Natural Language and Linguistic Theory* 36(2):323-363.

Caponigro, Ivano and Anamaria Fălăuş. 2020. Unveiling multiple *wh-* free relative clauses and their functional *wh*-words**.** *Semantics and Pragmatics* 13.

Caponigro, Ivano and Anamaria Fălăuş. 2022. The semantics of Rudin constructions in Romanian. In *Proceedings of the 23rd Amsterdam Colloquium*, eds M. Degano, T. Roberts, G. Sbardolini & M. Schouwstra (eds.), 55–61.

Caponigro, Ivano and Anamaria Fălăuş. 2023a. Investigating clausal *wh*-constructions in Romanian. In *Romance languages: recent contributions to linguistic theory*, eds R. Pires de Oliveira and C. Rodrigues. Special issue of *Isogloss. Open Journal of Romance Linguistics* 9(4)/2:1-32.

Caponigro, Ivano and Anamaria Fălăuş. 2023b. 'What' clauses can and 'which' cannot: a Romanian puzzle. *Language* 99(3): e176-e190.

Caponigro, Ivano and Maria Polinsky. 2011. Relative embeddings: A Circassian puzzle for the syntax/semantics interface. *Natural Language and Linguistic Theory*. 29(1):71-122

Chamoreau, Claudine. 2021. Headless Relative Clauses in Pesh. In *Headless Relative Clauses in Mesoamerican Languages*, eds I. Caponigro, H. Torrence, and R. Zavala Maldonado, 509–546. New York: Oxford University Press.

Chierchia, Gennaro. 1982. Nominalization and Montague Grammar. A Semantics without Types for Natural Language. *Linguistics and Philosophy* 5(3):303–354.

Chierchia, Gennaro. 1983. *Topics in the Syntax and Semantics of Infinitives and Gerunds*. PhD diss. University of Massachusetts at Amherst. (Published as G. Chierchia, *Topics in the Syntax and Semantics of Infinitives and Gerunds.* Milton Park: Routledge. 1988.)

Chierchia, Gennaro. 2025a. At the frontier of quantification: Correlatives, movement and scope. In *Proceedings of Sinn und Bedeutung 29*, eds F. Longo and D. Panizza, 319–339. Noto: University of Messina, C.U.M.O.

Chierchia, Gennaro. 2025b. On the intimate connections between questions and relative clauses. Talk given at The Nantes Workshop, May 9–10.

Chierchia, Gennaro and Ivano Caponigro. 2013. Questions on questions and free relatives. Talk given at Sinn und Bedutung 18, September 12.

Dayal, Veneeta. 1996. *Locality in WH Quantification: Questions and Relative Clauses in Hindi*. Dordrecht: Kluwer Academic Publishers.

Dayal, Veneeta. 1997. Free relatives and *ever*: Identity and free choice readings. In *Proceedings from Semantics and Linguistic Theory (SALT) VII*, ed. Aaron Lawson, 99–116. Ithaca, NY: CLC Publications, Cornell University.

Dayal, Veneeta. 2016. *Questions.* Oxford: Oxford University Press.

Engdahl, Elisabet. 1980. *The syntax and semantics of questions in Swedish*. PhD thesis, University of Massachusetts, Amherst.

Engdahl, Elisabet. 1986. *Constituent Questions: The Syntax and Semantics of Questions with Special Reference to Swedish.* Dordrecht: Reidel.

Fălăuş, Anamaria and Andreea C. Nicolae. 2022. Additive free choice items. *Natural Language Semantics* 30:185-214

von Fintel, Kai. 2000. Whatever. In *Proceedings from Semantics and Linguistic Theory (SALT) X*, eds. Brendan Jackson and Tanya Matthews, 27–39. Ithaca, NY: CLC Publications, Cornell University.

Flores-Nájera, Lucero. 2021. Headless relative clauses in Tlaxcala Náhuatl. In *Headless relative clauses in Mesoamerican languages*, eds I. Caponigro, H. Torrence & R. Zavala Maldonado, 79–110. New York: Oxford University

Groenendijk, Jeroen and Martin Stokhof. 1982. Semantic analysis of wh-complements. *Linguistics and Philosophy* 5(2):175–233.

Groenendijk, Jeroen and Martin Stokhof. 1983. Interrogative quantifiers and Skolem-functions. In *Connectedness in Sentence, Discourse and Text: Proceedings of the Tilburg Conference*, eds K. Enlich and H. van Riemsdijk, 71–110. University of Tilburg, Department of Languages and Literature (Reprinted in Groenendijk, J. and M. Stokhof. 1984. *Studies on the Semantics of Questions and the Pragmatics of Answers*, 165–208. PhD diss, Universiteit van Amsterdam.)

Hamblin, Charles L. 1973. Questions in Montague English. *Foundations of Language* 10:41–53.

Haspelmath, Martin. 1997. *Indefinite Pronouns*. Oxford: Oxford University Press.

Haspelmath, Martin. 2013. Indefinite Pronouns. *WALS Online* (v2020.4) http://wals.info/chapter/46 (Accessed on March 14, 2026)

Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In *Quantification in Natural Languages*, eds E. Bach, E. Jelinek, A. Kratzer & B. H. Partee, 451–486. Dordrecht: Kluwer.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1(1): 3–44.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: a lattice-theoretical approach. In *Meaning, Use and the Interpretation of Language,* eds R. Bauerle, C. Schwarze & A. von Stechow, 303–323. Berlin: de Gruyter.

López Márquez, Wendy. 2021. Headless relative clauses in Sierra Popoluca. In *Headless relative clauses in Mesoamerican languages*, eds I. Caponigro, H. Torrence & R. Zavala Maldonado, 475–508. New York: Oxford University

Montague, Richard. 1970. English as a formal language. In B. Visentini (ed.), *Linguaggi nella Società e nella Tecnica*, 189−223. Milan: Edizioni di Comunità. [Reprinted in R. Montague, *Formal Philosophy*, 188−221. New Heaven and London: Yale University Press. 1974.]

Patterson, Gary and Ivano Caponigro. 2016. The puzzling degraded status of who free relative clauses in English. *English Language and Linguistics* 20(2):341−352.

Partee, Barbara H. 1986. Noun phrase interpretation and type-shifting principles. In *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, eds. J. Groenendijk, D. de Jongh and M. Stokhof, 115-143. Dordrecht: Foris.

Prince, Ellen. 1989. Yiddish wh-clauses, subject-postposing, and topicalization. In *Proceedings of the Fifth Eastern States Conference on Linguistics (ESCOL '88)*, eds J. Powers and K. de Jong (eds), 403–415. Columbus, OH: Ohio State University.

Quine, Willard V.O. 1960. *Word and Object*. Cambridge, MA: The MIT Press.

Rudin, Catherine 1988. On multiple wh questions and multiple wh fronting. Natural Language and Linguistics Theory 6:445–501.

Šimík, Radek. 2011. *Modal existential wh-constructions*. PhD diss., University of Groningen.

Šimík, Radek. 2017. Existential *wh*-constructions. In *Oxford Bibliographies in Linguistics*, ed. M. Aronoff. New York: Oxford University Press.

Šimík, Radek. 2020. Free relatives. In D. Gutzmann, L. Matthewson, C. Meier, H. Rullmann, & T. E. Zimmermann (eds.), *The Wiley Blackwell Companion to Semantics*, 1–38. Hoboken, NJ: Wiley.

Šimík, Radek. 2026. From interrogatives to relatives. A comprehensive accounts of *wh*-constructions. In *The Cambridge Handbook of Comparative Syntax*, eds S. Barbiers, N. Corver, and M. Polinsky, 824–865. Cambridge: Cambridge University Press.

Stockwell, Richard and Carson Schütze. 2022. The puzzling nuanced status of *who* free relative clauses in English: A follow-up to Patterson and Caponigro (2015). *English Language and Linguistics* 26(1):185–202.

de Vries, Mark. 2002. *The Syntax of Relativization*. PhD diss., University of Amsterdam.

Xiang, Yimei. Quantifying into *wh*-dependencies: multiple-*wh* questions and questions with a quantifier. *Linguistic and Philosophy* 46:429–482.

Zanuttini, Raffaella and Paul Portner. 2003. Exclamative Clauses: At the Syntax-Semantics Interface. *Language* 79(1):39–81.