A slightly different version appeared in: Ivano Caponigro, Harold Torrence, and Roberto Zavala Maldonado (eds), *Headless Relative Clauses in Mesoamerican languages*, pp. 1–57. New York: Oxford University Press, 2021.

CHAPTER 1

# Introducing Headless Relative Clauses and the findings from Mesoamerican Languages

**Ivano Caponigro**

October 23, 2020

*Abstract*

This introductory chapter pursues several goals. First, it introduces the characters at the center of the volume: both the main characters, i.e., varieties of headless relative clauses, and the equally important supporting characters, i.e., headed relative clauses and *wh*- interrogative clauses. The next chapters can, therefore, assume that the reader is already familiar with these core constructions and just focus on their manifestations in the specific languages under investigation. Second, this chapter presents the definitions, methodologies, and tests that all the authors of the volume have adopted, so that each subsequent, language-specific chapter can make use of them without further introduction or justification. In doing so, this chapter also fulfills a third goal: to provide a concise guide to scholars who are interested in pursuing further investigation of headless relative clauses in Mesoamerican and other languages. Fourth, the current chapter aims to highlight commonalities and differences in the findings from the other chapters and discusses how those findings contribute to the current understanding of headless relative clauses typologically and theoretically and of human language in general.

*Keywords:* Mesoamerican language, Chibchan language, headless relative clause, free relative clause, light-headed relative clause, headed relative clause, *wh*-expression, interrogative clause, definite, existential construction

The Front Matter including the index and the preface of the volume is available here: http://idiom.ucsd.edu/~ivano/Papers/2020_CaponigroTorrenceZavala_FrontMatter_2020-10-23.pdf

**Table of Contents**

# 1. Introduction

## 1.1 Headless relative clauses

Languages can use nouns as lexical devices to refer to sets of objects in the world that share a similar property or are subsumed under a shared concept, or—depending on the theory of meaning that is endorsed—to refer to that property or concept itself. For instance, the noun *immigrant* refers to the set of all and only the people who share the property of having moved to a country from elsewhere, or—under a different view—it refers to the related concept we have in our mind.[1] Nominal modifiers like adjectives can help restrict the set a noun refers to. For instance, *Italian immigrant(s)* refers to the subset of people who are originally from Italy but have moved to another country.

Modified or unmodified nominals can then be used to form referential expressions by adding specific markers like definite determiners, demonstrative determiners, etc., or no marking at all—depending on the language. For instance, *that immigrant* refers to a contextually salient person who is an immigrant and is currently pointed at or was previously mentioned in the conversation.

A different group of markers—quantifiers—is used to quantify over the set that a noun refers to. For instance, *every immigrant*—with the universal quantifier *every* as its determiner—doesn't refer to any particular person or set of people. What it does, instead, is to impose some form of relationship between the set of immigrants and the set that is referred to by the remainder of the sentence. Thus, *Every immigrant is Italian* is true if all the members of the set of immigrants are also members of the set of Italians. On the other hand, *One immigrant is Italian*—with the numeral quantifier *one*—is true if at least one member of the set of immigrants is in the set of Italians as well. Unlike reference, quantification has to be triggered by an overt marker—an overt quantifier.[2]

Another very common feature across languages is the use of clauses to refer to states of affairs—facts that hold in the actual world or some other possible world. The declarative clause *That immigrant is Italian* refers to a state of affairs in which the immigrant that the speaker is pointing at is from Italy in the actual world. Clauses can also be used to ask for information on states of affairs or request that a state of affairs holds. For instance, the *wh*- interrogative clause *Who is Italian?* asks for the identity of the Italian person in the state of affairs that is described by the preceding declarative clause, while the imperative clause *Help that immigrant!* conveys the speaker's request that the state of affairs in which the hearer helps the immigrant be actualized.

[1] See Chierchia & McConnell-Ginet (2000, Ch. 1) for a brief introduction to these two families of approaches to language meaning.

[2] We leave aside whether bare nominals that are interpreted as indefinites should be treated as quantificational expressions.

However, there is a well-known family of clauses across languages that exhibit a different behavior. Syntactically they combine with a noun (their 'head'), and semantically they restrict that noun in the same way as adjectives do. These clauses are known as *headed relative clauses*. For instance, the definite nominal *the immigrant Pablo is helping* refers to a specific person who is an immigrant and who is helped by Pablo. The string *Pablo is helping* is a clause with a subject and an inflected verb. It is missing its object, though, and it does not refer to a state of affairs, but rather to the set of people that Pablo is helping. Notice that this modification can become quite rich and complex, as in the phrase *the immigrant Pablo is helping together with other volunteers after the police arrested and deported her family*. It looks as if languages are taking advantage of the lexical richness and structural complexity of clauses to restrict the meaning of a noun in ways that no single lexical item like an adjective could ever do.

Languages can even go one step further and use clauses not to combine them with and modify nouns, but to actually act like referential or quantificational nominals themselves, both distributionally and semantically. Such clauses can refer to individual objects or quantify over sets of objects. Some of them can even combine with the same determiners as nouns. For instance, the *wh-* clause *what Pablo made* in *Maria tasted what Pablo made* is missing its object and refers to the food that Pablo made, even if no nominal head introduces it. We label clauses like this one *headless relative clauses* (henceforth, *[−H]RCs* throughout this chapter). They are at the center of our investigation across Mesoamerican languages.

### 1.2. Mesoamerican languages

The term *Mesoamerica* (or *Meso-America*) refers to both a geographical region and a cultural and linguistic area. The region extends "from the Pánuco River in northern Mexico to the Lempa River in El Salvador, but also includes the Pacific coast of Nicaragua and Costa Rica" (Campbell 1997: 156). The indigenous populations from this region share a large number of cultural and linguistic traits (Campbell et al. 1986). Mesoamerican languages form a "group of more than 125 languages classified into some 10 language families (including language isolates) that are native to Mesoamerica," including Uto-Aztecan, Oto-Manguean, Mayan, and Mixe-Zoquean. (Campbell 2014). Word order restrictions are among their common features: no Mesoamerican language exhibits verb-final basic word order, "although Mesoamerica is bounded by languages to both the north and the south that have SOV basic word order" (Campbell 2014).

[−H]RCs have received less attention in the linguistic literature than two related constructions, namely headed relative clauses and *wh-* interrogative clauses, despite the many morpho-syntactic and semantic puzzles they raise within and across languages and for our understanding of human language in general. As most Mesoamerican languages are understudied, it is not surprising that even less is known about [−H]RCs in these languages. Nonetheless, preliminary investigations have shown that they are present and productive (see §6).

### 1.3. This volume

This volume begins to fill in the gaps in the study of [−H]RCs in Mesoamerican languages—languages which are all threatened or endangered. It investigates fifteen languages from five language families—all of which are Mesoamerican but one. It provides the most extensive study of [−H]RCs in each of the languages investigated, as well as the broadest and most articulated crosslinguistic study of [−H]RCs that has been undertaken to date. It also contributes to the broader enterprise of understanding and accounting for the syntactic and semantic behavior of [−H]RCs crosslinguistically. Finally, it offers the methodology we have developed as a model for future language-specific or comparative investigations of [−H]RCs, both in Mesoamerica and beyond.

All of the fifteen chapters in this volume are language specific, except for this introductory chapter, and constitute original contributions to typological and theoretical linguistics. They cover morpho-syntactic issues such as determining what elements can introduce or characterize [−H]RCs (*wh*-words, specialized relative markers, articles, deictics, etc.), the conditions under which those elements are enriched by further morpho-syntactic marking (e.g., *what+ever* in English), the circumstances under which a clause can have the distribution of a nominal or prepositional phrase, and the syntactic structure of [−H]RCs. All of the chapters also touch on semantic issues, in particular, the interpretation of different kinds of [−H]RCs as definite, indefinite, free choice, etc.

This volume presents the results of a genuinely collaborative project. The writing of the individual chapters was preceded by two preparatory workshops which aimed to provide a common understanding of the various constructions involved and their crosslinguistic variation; by collective brainstorming on definitions, methodologies, and tests; and by group and one-on-one discussions. Each chapter (except this one) focuses on a single specific language (or at most two), to do justice to the breadth of the empirical investigation within each language and the need to provide the language-specific background.

### 1.4. This chapter

This first chapter pursues several goals. First, it introduces the characters at the center of the volume—both the main characters, i.e., the documented varieties of [−H]RCs, and the equally important supporting characters, i.e., headed relative clauses and *wh*- interrogative clauses. The next chapters can, therefore, assume that the reader is already familiar with these constructions and can focus on their manifestations in the specific languages under investigation. Second, this chapter presents the definitions, methodologies, and syntactic/semantics tests that our team has adopted, so that each language-specific chapter can make use of them without having to introduce and justify them. In doing so, this chapter fulfills a third goal: to provide a concise guide to scholars who are interested in pursuing further investigation of [−H]RCs in Mesoamerican and other

languages. Fourth, the current chapter aims to highlight commonalities and differences in the findings from the other chapters and discusses how those findings contribute to our current understanding of [−H]RCs typologically and theoretically. Although comparative data and generalizations across languages come up occasionally elsewhere in the volume, it is only in this introductory chapter that the findings are systematically summarized and compared across all languages under study.

Examples in this chapter are mainly from English and Spanish. We intentionally avoided citing random linguistic examples from Mesoamerican languages because they may not be fully understood—or may even be misunderstood—without the language-specific background and explanations that are provided in their respective chapter. We would like our volume to speak to the broadest linguistic audience possible, regardless of language expertise or theoretical commitments. The Mesoamerican experts among our readers will find ample data from Mesoamerican languages in the next chapters.

The remainder of this chapter unfolds as follows. First, we define [−H]RCs by introducing the properties that they all share (§2). Then, we characterize different varieties of [−H]RCs and indicate how to distinguish among them (§3). Next, we compare [−H]RCs to two related and better-known constructions: *wh-* interrogative clauses (§4) and headed relative clauses (§5). We discuss the significance of [−H]RCs and the reasons to study them across languages in general, and in Mesoamerican languages in particular (§6). We introduce the format that each language-specific chapter in the volume follows and the reasons behind this choice of uniformity (§7). Finally, we present the main findings across all the language-specific chapters, highlight similarities and differences among [−H]RCs in the languages under investigation, and discuss how our findings support and further develop the current understanding of [-H]RCs, relative constructions, and clausal embedding in general (§8). We conclude with some general remarks (§9).

## 2. Definition of headless relative clauses

In this section, we introduce our main characters. The label "headless relative clauses" (which we shorten to *([−H]RCs)*, as noted earlier in this chapter) is used throughout the volume to refer to a family of constructions that share the properties in (1):

(1) *Properties characterizing [−H]RCs:*

I. They are embedded clauses.[3]

II. They lack a constituent—an argument or an adjunct (which can sometimes be marked by a resumptive pronominal form).

[3] Or, equivalently, "dependent clauses" or "subordinated clauses."

III. They lack an "external nominal head"—a nominal head that precedes or follows them and is linked[4] to the missing constituent.

IV. They exhibit the same distribution and interpretation as nominal phrases (DPs[5]) or prepositional/postpositional phrases (PPs).

Examples of [−H]RCs are given in brackets in (2)–(4).

(2) Frida was very friendly with [$_{[-H]RC}$ *those* Susana invited __ from Mexico].

(3) Paloma buys [$_{[-H]RC}$ *whatever painting* Frida makes __ ].

(4) Mercedes lives [$_{[-H]RC}$ *where* she was born __ ].

The [−H]RCs in (2)–(4) share the four properties in (1). They are all embedded clauses inside a matrix clause (Property I). They all lack a missing, marked by the underscore (Property II): in (2) and (3), what is missing is the object of the transitive predicates *invited* and *makes*, while in (4) it is the locative adjunct to the intransitive predicate *was born*. None of the [−H]RCs is preceded or followed by a nominal head (Property III). The [−H]RC in (2) is introduced by the italicized demonstrative *those*, which crucially occurs without a nominal complement. The [−H]RC in (3) is introduced by the italicized complex *wh*-phrase *whatever painting*, which does contain a noun. The noun *painting*, though, is not external to the [−H]RC, but is part of it.[6] The [−H]RC in (4) is introduced by the italicized *wh*-word, which sits in the left peripheral position of the [−H]RC. Last, the [−H]RCs in (2), (3), and (4) can all be replaced and paraphrased by a DP or a PP (Property IV). The [−H]RCs in (2) and (3) have the same distribution and interpretation as the bracketed DPs in (5) and (6), respectively. The bracketed string in (4), by contrast, shows an example of a [−H]RC with the same distribution and interpretation as the bracketed PP in (7).

(5) Frida was very friendly with [$_{DP}$ the people that Susana invited from Mexico].
(6) Paloma buys [$_{DP}$ any painting that Frida makes].
(7) Mercedes lives [$_{PP}$ in the place where she was born].

Despite the similarities highlighted in (1), [−H]RCs exhibit differences in interpretation, distribution, and/or morpho-syntactic properties that are discussed in the next section (§3). Overall,

[4] "Linked" is used here as a general label for any kind of morpho-syntactic (e.g., gender or number agreement) or semantic connection.
[5] In this introductory chapter, we use the syntactic label *Determiner Phrase* (*DP*) to refer to a full nominal constituent including a determiner, like *the old cat* or *almost every book on the shelves*, while we reserve the label *Noun Phrase* (*NP*) for just the noun and its modifiers, without the determiner (and its modifiers), like *old cat* and *book on the shelves*. See Abney (1987) and, more recently, Salzmann (2018) for arguments in favor of this distinction. Some of the chapters in the volume follow the same distinction, while others make use of the label *NP* only.
[6] See Caponigro (2019) for data and arguments from English and crosslinguistically.

the label "headless relative clauses ([−H]RCs)" should be taken as a way to identify not just one single construction, but a cluster of related constructions, and to distinguish them from another cluster of related but different constructions—headed relative clauses.

## 3. Varieties of headless relative clauses

This section introduces different varieties of [−H]RCs and highlights similarities and differences we have relied on in our investigation. We have identified three main varieties of [−H]RCs: *free relative clauses*, *light-headed relative clauses*, and *super-free relative clauses*. They all share the four properties in (1) in the preceding, which we summarize and further specify with the syntactic schema in (8)a and the feature bundle in (8)b.

(8) *Summary of the properties characterizing [−H]RCs:*

a. [(DET) [$_{CP}$ (*wh*-/REL/COMP) …. __ …]]$_{DP/PP}$[7]

b. [±D, –N, ±WH]

[−H]RCs have no nominal head (–N). Some have a "light head," which we have assigned the label Determiner (±D) for convenience and whose nature we shall discuss later. [−H]RCs can feature a *wh*-expression[8] (*wh*-) from the set of those introducing *wh*- interrogative clauses with possible extra morphological marking (±WH), a non-*wh* relativizer (REL) of the same kind as those introducing headed relative clauses (including non-*wh* relative pronouns), a general complementizer (COMP) of the same kind as those introducing complement clauses (and maybe headed relative clauses too), a combination of those, or no marking at all. All [−H]RCs have a missing constituent or gap ( __ ).

We discuss each kind of [−H]RC in turn, highlighting the main criteria that we have made use of to identify the [−H]RC in the languages under study. We also introduce further distinctions within each kind. Free relative clauses are discussed next (§ 3.1), then light-headed relative clauses (§ 3.2.1), and finally super-free relative clauses (§3.2.2).

[7] The subscript *DP/PP* occurring at the far right of the syntactic schema in (8)a does not indicate the actual syntactic categories of [−H]RCs, but indicates the distributional similarities of [−H]RCs with actual DPs and PPs. The subscript *CP* stands for *Complementizer Phrase*—the phrasal category of a full, tensed clause, including *wh*- clauses of any kind.

[8] We use the term *wh-expression* to refer both to a single *wh*-word occurring on its own (e.g., *who, where, when,* etc.) as well as to a *wh*-phrase that consists of a *wh*-word and other material, like a complement (e.g., *what book, how much food*) or a preposition (e.g., *by means of what*, *with which friend*) or both (e.g., *together with how many other people*).

### 3.1. Free relative clauses

*Free relative clauses* (*FRs*) are [−H]RCs without a D head that are introduced by *wh*-expressions that either are identical to those in *wh*- interrogative clauses or carry extra morphology. The *wh*-expression in a [−H]RC may co-occur with a relativizer or a complementizer. The characterizing features of [−H]RCs are summarized by the syntactic schema in (9)a and the feature bundle in (9)b.

(9) *Properties characterizing FRs*
   a. [$_{CP}$ *wh*- (REL/COMP) …. __ …]$_{DP/PP}$
   b. [–D, –N, +WH]

The bracketed strings in (3) and (4) in the preceding are examples of FRs in English. The FR in (4) is introduced by the *wh*-word *where*, while the FR in (3) is introduced by the *wh*-phrase *whatever painting*. All three main varieties of FRs that are attested crosslinguistically occur in the languages in our study as well: maximal free relative clauses, existential free relative clauses, and free-choice free relative clauses. The crucial distinction is semantic, as revealed by these labels. This distinction often correlates with morpho-syntactic differences as well.[9] We discuss and exemplify each variety of FR in turn in §§3.1.1–3.1.3.

#### *3.1.1. Maximal free relative clauses*

*Maximal free relative clauses* (*Max-FRs*) are those FRs that satisfy the properties in (10):

(10) *Properties characterizing Max-FRs:*

   a. DEFINITENESS. A Max-FR can be replaced and paraphrased by a definite DP—a DP introduced by a definite marker or determiner in a language that has them, like *the* in English or *el/la/los/las* in Spanish—or by a PP with a definite DP as its complement.

   b. REFERENTIALITY. A Max-FR is interpreted as referential: it refers to an individual.[10] In this respect, Max-FRs are like proper names, definite DPs, and DPs introduced by demonstratives.

[9] See Šimík (forthcoming) for a thorough overview of the semantic properties of FRs and the analyses that have been suggested, and van Riemsdijk (2017) for a detailed overview of their syntactic properties and related syntactic proposals.

[10] We follow a common use in semantics and adopt the term ‘individual’ to characterize any object in the ontology of the domain of discourse—any human or non-human, animate or inanimate, abstract or concrete object we can talk about, including places, time points or intervals, manners, and reasons.

c. MAXIMALITY: A Max-FR is interpreted as maximal: it refers to the largest ('maximal') individual of a set of (atomic and plural) individuals. This is the same semantic behavior as seen with definite DPs.

Let's explore the properties in (10) by looking at concrete examples of Max-FRs, like those from English and Spanish in brackets in (11)a–(14)a.

(11) a. I bought [$_{\text{Max-FR}}$ **what** __ is on the table].
b. I bought [$_{\text{Definite DP}}$ the things on the table].

(12) a. Admiro a [$_{\text{Max-FR}}$ **quien** __ trabaja duro]. *Spanish*
admire.1SG ANIM.OBJ[11] who works hard
'I admire those who work hard.'
b. Admiro a [$_{\text{Definite DP}}$ la gente que trabaja duro].
admire.1SG ANIM.OBJ the people.SG that works hard
'I admire the people who work hard.'

(13) a. I went [$_{\text{Max-FR}}$ **where** my friends are vacationing __ ].
b. I went [$_{\text{PP}}$ to [$_{\text{Definite DP}}$ the place(s) my friends are vacationing]].

(14) a. Lo hice [$_{\text{Max-FR}}$ **como** tú lo hiciste __ ]. *Spanish*
it did.1SG how you it did.2SG
'I did it {how}/{in the same way(s)} you did it.'
b. Lo hice [$_{\text{PP}}$ de [$_{\text{Definite DP}}$ la misma manera en que tú lo hiciste]].
it did.1SG of the same way in which you it did.2SG
'I did it in the same way you did it.'

All the Max-FRs in (11)a–(14)a satisfy the "Definiteness" property in (10)a: they can all be replaced and paraphrased with definite DPs, as shown in (11)b–(14)b. They also satisfy the "Referentiality" property in (10)b: the Max-FR in (11)a refers to the object(s) on the table, the Max-FR in (12)a refers to the people who work hard, the Max-FRs in (13)a to the place(s) where the speaker's friends are on vacation, and the Max-FR in (14)a refers to the way(s) in which the addressee did what they were doing. Referentiality is also a semantic property of the definite DPs replacing the Max-FRs in (11)b–(14)b, and of definite DPs in general.

Last, all the Max-FRs in (11)a–(14)a exhibit the "Maximality" property in (10)c. If there is an apple, an orange, and a peach on the table, the Max-FR in (11)a refers to the plural individual

[11] ANIM: animate.

that is formed by combining those three atomic individuals. Crucially, it cannot refer to anything smaller than that—such as the atomic individual consisting of the apple or the non-maximal plural individual made up of only the apple and the orange. This is the same semantic behavior as found with the plural definite DP *the things on the table*. Notice that the Max-FR in (11)a cannot be interpreted as triggering quantification over a set of individuals, unlike the indefinite DP *some of the things on the table*. Similarly, the Max-FR in (12)a has to refer to the maximal individual made up of all the hard workers, like the plural definite DP *the people who work hard* and unlike the indefinite DP *some of the people who work hard*. The same semantic behavior is observed when a set of places (13)a or a set of manners (14)a is involved: the Max-FR has to refer to the maximal individual of those sets. The Max-FR in (13)a means the same thing as the PP *to the places where my friends are vacationing* with a definite DP as its complement. It cannot mean the same as the indefinite DP *to some of the places where my friends are vacationing*. The Max-FR in (14)a receives the same interpretation as the PP *in the same way(s) you did it* with a definite DP as its complement, while it can never be interpreted as a PP with an indefinite DP complement like *in some of the ways you did it*.[12]

The precise syntactic structure of Max-FRs is an open issue—in particular, whether Max-FRs are headed by a silent head (a D head and/or an N head).[13] Unless otherwise stated, the chapters in this volume assume Max-FRs—and FRs at large—to be just clauses (CPs), as in the general syntactic schema in (9)a.

Notice that not all *wh*-expressions which occur in *wh*- interrogative clauses necessarily occur in Max-FRs. Both English and Spanish only allow for a proper subset of *wh*-expressions to occur, and the particular subset is different in each language. For instance, the *wh*-word *quien*(*es*) can easily occur in Max-FRs in Spanish, whereas *who*, its equivalent in English, is highly restricted (see Patterson & Caponigro 2016). On the other hand, the *wh*-word *what* can productively introduce Max-FRs in English, whereas the equivalent in Spanish, *que*, is completely banned from Max-FRs. In our investigation, we have aimed to collect all the *wh*-expressions occurring in *wh*-interrogative clauses and to check if each of them can occur in Max-FRs (or in any other kind of FR).

The properties in (10) are partially redundant on purpose. Although definiteness is equivalent to the combination of referentiality and maximality within the semantic framework that we are adopting, the replacement and paraphrase test in (10)a provides a quick preliminary step to assess whether a *wh*- clause is a Max-FR.

[12] The semantic view of Max-FRs we just sketched is directly based on Jacobson (1995), Dayal (1996), and Caponigro (2003, 2004), who, in turn, are inspired by the analysis of definite DPs in Sharvy (1980) and Link (1983). See Šimik (forthcoming) for a thorough overview of the main semantic properties of Max-FRs and related proposals. As for the semantic properties of different kinds of DPs, see Farkas (2018) for a general overview, Heim (2011) for an overview of the semantics of definite DPs, Wolter (2009) for an overview of the semantics of demonstrative DPs, and Szabolcsi (2010) for an overview of the semantics of quantificational DPs.

[13] See van Riemsdijk (2017) for a review of this issue.

### *3.1.2. Existential free relative clauses*

*Existential free relative clauses* (*Ex-FRs*) are those FRs satisfying the properties in (15)a–b and exhibit the syntactic structure schematized in (15)c.[14]

(15) *Properties characterizing Ex-FRs:*

a. EXISTENTIAL MEANING. Ex-FRs can be replaced and paraphrased by existentially quantified nominal expressions—indefinite DPs, which are introduced by indefinite markers (e.g., determiners like *a* in English or *un/una/unos/unas* in Spanish), or bare nominals (e.g., *I love books*), in languages that allow for either.

b. EXISTENTIAL PREDICATE. When attested in a language, Ex-FRs can always occur as the complement of existential 'be' and existential 'have' in that language.

c. [$V_{[existential]}$ [$_{CP}$ *wh-* (REL/COMP) … __ …]$_{Indefinite\ DP/Bare\ Nominal}$ ]

Ex-FRs are not attested in English or other Germanic languages,[15] but are common in Romance, Balto-Slavic, and Semitic languages, as well as in the Mesoamerican languages we have studied. Examples of Ex-FRs from Spanish and Italian are given in brackets in (16)a–(21)a. Comments follow.

(16) a. No tengo [$_{Ex\text{-}FR}$ **qué** decirles]. *Spanish*
not have.1SG what say.INF_to.them
'I don't have anything to tell them.'
b. No tengo [$_{Indefinite\ DP}$ nada que decirles].
not have.1SG nothing that say.INF_ to.them
'I don't have anything to tell them.'

(17) a. Tengo [$_{Ex\text{-}FR}$ **con quién** hablar]. *Spanish*
have.1SG with who speak.INF
'I have somebody/people I can talk to.'
b. Tengo [$_{Indefinite\ DP}$ un amigo con quien hablar].
have.1SG a friend with who speak.INF

[14] This discussion of EX-FRs owes much to Izvorski (1998), Grosu (2004, 2013), Caponigro (2003, 2004), Šimík (2011), and work cited there (Šimík 2017 provides a thorough review of the relevant literature on Ex-FRs and related constructions). It also benefitted from Judith Aissen's lecture on Ex-FRs at the first workshop that we organized as part of our project on "Headless Relative Clauses in Mesoamerican Languages".
[15] With the possible exception of Yiddish and Yiddish-influenced varieties of English (Caponigro 2003: §3.2.4).

'I have a friend I can talk to.'

(18) a. No hay [$_{\text{Ex-FR}}$ **dónde** ir]. *Spanish*
not there's where go.INF
'There's nowhere to go.'
b. No hay [$_{\text{Indefinite DP}}$ un lugar donde ir].
not there's a place where go.INF
'There's no place to go to.'

(19) a. C'è [$_{\text{Ex-FR}}$ **chi** dice sempre sì]. *Italian*
there's who say.IND.PRS.3SG always yes
'There are people who say yes all the time.'
b. Ci sono [$_{\text{Indefinite DP}}$ (delle) persone che dicono sempre sì].
there are (some) people that say. IND.PRS.3PL always yes
'There are people who say yes all the time.'

(20) a. Non aveva [$_{\text{Ex-FR}}$ **chi** le curasse i bambini]. *Italian*
not have.IND.IPFV.3SG who to.her look-after.SBJV.IPFV.3SG the children
'She didn't have anybody who looked after her children.'
b. Non aveva [$_{\text{Indefinite DP}}$ nessuno che le curasse i bambini].
not have.IND.IPFV.3SG nobody that to.her look-after.SBJV.IPFV.3SG the children
'She didn't have anybody who looked after her children.'

(21) a. Ho [$_{\text{Ex-FR}}$ **di che** lamentarmi]. *Italian*
have.IND.PRS.1SG of what complain.INF
'I have something/things to complain about.'
b. Ho [$_{\text{Indefinite DP}}$ (delle) cose di cui lamentarmi].
have.IND.PRS.1SG (some) things of which complain.INF
'I have things to complain about.'

All the Ex-FRs in (16)a–(21)a satisfy the properties in (15)a–b. Examples (16)b–(21)b shows that the Ex-FRs can be replaced and paraphrased with existentially quantified DPs, as required by the "Existential Meaning" property in (15)a. For instance, (17)a asserts the existence of at least one person the speaker can talk to. In other words, the set of people the speaker can talk to is not empty. The Ex-FR is, therefore, semantically close to the indefinite DP in (17)b. Example (20)a with negation on the matrix existential predicate asserts that it is not true that there is at least one individual in the set of people who could take care of the children of the contextually salient woman that the speaker is referring to. In other words, that woman didn't have anybody who could

take care of her children. Therefore, the Ex-FR in (20)a is semantically close to the negative indefinite in (20)b. As for the "Existential Predicate" property in (15)b, all the Ex-FRs in (16)a–(21) occur as the complements of existential 'be' or existential 'have' in Spanish and Italian. The two properties are, of course, connected. Existential predicates select for indefinites as their complements (exclusively so in many languages). It is not surprising, therefore, that if *wh*- clauses are allowed as their complements, they will be interpreted as indefinite DPs.

The two properties in (15)a–b are attested in Ex-FRs across languages, including all the languages in this volume. Ex-FRs also exhibit parameters of variation when compared across languages. We touch on three of them next.

PREDICATES. A first parameter of crosslinguistic variation concerns the predicates whose complement position can introduce Ex-FRs (no other argument positions can productively introduce Ex-FRs across languages). This is always a highly restricted set, which includes existential 'be' and existential 'have.'[16] There are languages that allow for more than just those two existential predicates. Šimík (2011) offers a rich survey of Ex-FRs in Indo-European and Finno-Ugric languages and predicates that can introduce them. Although our investigation of Ex-FRs will mainly focus on existential 'be' and 'have', we may provide examples of Ex-FRs in the complement position of other predicates as well. These are predicates like 'find', 'look for/seek', 'choose', 'give', 'get', 'take', 'send', 'bring', 'buy', and 'build' (Šimík 2011: §2.2)—verbs of "coming into being, view, or availability, or causation of one of these" (Grosu 2004) whose lexical meaning supports existential quantification over their indefinite internal argument (Szabolcsi 1986).

MOOD/MODALITY. Another parameter of variation is the mood/modal properties of Ex-FRs. In many languages, including some in this volume, Ex-FRs require a non-indicative/irrealis verb form. Depending on the language and its morpho-syntactic resources, this form could be infinitive, subjunctive, or have some other irrealis marking. Semantically, this marking is interpreted as conveying some form of modality with existential force (possibility) and circumstantial flavor (Šimík 2011: Ch. 2, §2.2.6). This requirement doesn't hold for all languages, though, or even for all examples of Ex-FRs in the same language. For instance, the examples from Italian in (19)a–(21)a above show three different options: the Ex-FR in (19)a is in the indicative, the one in (20)a is in the subjunctive, and the one in (21)a is infinitival. Some of the languages we have investigated don't impose any restrictions on the mood of their Ex-FRs, while some others do, as we discuss in §8.3. They offer an intriguing challenge to the generalization put forward most forcefully by Šimík (2011) that Ex-FRs in all languages can only occur in infinitival form or, if the language doesn't make use of an infinitive, in the subjunctive or some other non-indicative/irrealis mood.

[16] These two existential predicates can be realized by just one form, with an extra argument added in the case of existential 'have'. This is the case in Hebrew and the languages we have investigated.

*WH*-EXPRESSIONS. The third and last dimension of variation that will be relevant for our investigation has to do with *wh*-expressions. As with Max-FRs, the *wh*-expressions that can introduce Ex-FRs are a subset of those introducing *wh*- interrogative clauses – but not necessarily the same subset that permits Max-FRs. For instance, the *wh*-word *qué* 'what' in Spanish can introduce Ex-FRs (and *wh*- interrogative clauses, of course), while it is unacceptable in Max-FRs. Crosslinguistic differences are possible too, even between closely related languages. In Italian, another Romance language closely related to Spanish, the *wh*-words *quando* 'when' and *come* 'how' cannot introduce Ex-FRs at all, unlike in Spanish, while they can easily introduce Max-FRs in Italian. It follows that it is important to investigate which specific *wh*-expressions can introduce Ex-FRs, in the same way that it mattered for the study of Max-FRs. This is how we proceeded in the present investigation.

The syntactic structure of Ex-FRs is an open issue. We assume that they are clauses (CPs). Grosu (2004, 2013) and Šimík (2011) argue that Ex-FRs should not even be called FRs because they are not clauses. They propose the label *Modal Existential Construction* (*MEC*) instead. They argue that peculiar features of Ex-FRs/MECs in the languages they investigate (e.g., infinitival or subjunctive mood, modal interpretation, lack of overt non-*wh* subjects, etc.) can be accounted for by assuming a syntactic structure smaller than a clause. Details aside, we think that this hypothesis faces problems with handling some of the crosslinguistic facts. For instance, the Italian Ex-FR in (19)a is in the indicative mood and doesn't receive any modal interpretation of the kind Grosu and Šimík argue for. Most of the Mesoamerican languages we have investigated don't impose any mood or modal restrictions on their Ex-FRs either, as we discuss in detail in §8.3. Throughout the volume, we use the label "Ex-FRs" for all the constructions satisfying (15)a–b and assume them to be CPs in the complement position of an existential predicate, as schematized in (15)c.

### *3.1.3. Free-choice free relative clauses*

*Free-choice free relative clauses* (*FC-FRs*) are those FRs that satisfy the properties in (22)a–b.[17] Examples are given in (23) and (24).

(22) *Properties characterizing FC-FRs:*

a. FREE-CHOICE INFERENCE. A sentence containing an FC-FR obligatorily triggers an inference of ignorance or indifference.

b. FREE-CHOICE MARKER. An FC-FR always contain a free-choice (FC) marker.

[17] This discussion of FC-FRs clauses owes much to Dayal (1997) and von Fintel (2000). It also benefitted from Scott AnderBois' lecture on FC-FRs at the first workshop we organized as part of the "Headless Relative Clauses in Mesoamerican Languages" project. See Šimík (forthcoming) for an overview of the relevant issues and literature. Free-choice free relative clauses in English (and in other languages as well) are often labeled "*-ever* free relative clauses".

(23) a. [$_{\text{FC-FR}}$ **Whatever** Paloma is cooking right now] uses onions.[18]
b. *Asserted content:* [$_{\text{Max-FR/DP}}${What}/{The stuff that} Paloma is cooking right now] uses onions.
c. *Ignorance FC inference:* The speaker doesn't know what Paloma is cooking right now.

(24) a. Pablo (simply) voted for [$_{\text{FC-FR}}$ **whoever** was at the top of the ballot].[19]
b. *Asserted content:* Pablo voted for [$_{\text{DP}}$ the person who was at the top of the ballot].
c. *Indifference FC inference:* Pablo didn't care about who was at the top of the ballot.

Example (23)a shows a bracketed FC-FR in the subject position of its matrix clause. It is introduced by the bolded *wh*-word *whatever*, which results from the morphological enrichment of the *wh*-root *what* with the FC suffix *-ever*. The asserted content of (23)a—the state of affairs that (23)a depicts—is the same as the one of (23)b, in which the FC-FR has been replaced with a Max-FR and a definite DP. The FC-FR in (23)a also obligatorily triggers the ignorance inference that the speaker doesn't know the identity of what Paloma is cooking, as stated in (23)c. Notice that both the Max-FR and the definite DP in (23)b are compatible with a situation in which the speaker doesn't know the identity of what Paloma is cooking, but, crucially, they are also compatible with a situation in which the speaker does know what Paloma is cooking. In other words, they do not obligatorily trigger an ignorance inference.

Example (24)a shows a bracketed FC-FR in the complement position of the preposition *for* in the matrix clause. It is introduced by the morphologically enriched *wh*-word *whoever* in bold. Example (24)a asserts the same as (24)b, in which the FC-FR has been replaced with a definite DP. Unlike (24)b, though, (24)a with an FC-FR necessarily triggers the indifference inference that Pablo—the individual the matrix subject refers to—doesn't care about the actual identity of the candidate who was at the top of the ballot, as stated in (24)c.

We group these two inferences under the same "free choice" label to highlight the fact that they both trigger a form of variation on the reference of an FC-FR. It does refer to a maximal individual, like a Max-FR, but which maximal individual it refers to can vary depending on the relevant modality and the related modal agent: epistemic modality and the speaker in (23), or counterfactual modality and Pablo—the individual the matrix subject refers to—in (24).

FC INFERENCES. It is a parameter of crosslinguistic variation whether both inferences are triggered in FC-FRs or only one—and if only one, which one. English FC-FRs allow for either, depending on various factors (tense, aspect, discourse conditions). Italian and Romanian FC-FRs allow only for ignorance inferences (Caponigro & Fălăuş 2018). It has been an open issue whether

[18] Adapted from Dayal (1997: ex. (27a)).
[19] Adapted from von Fintel (2000: ex. (18)).

there are languages whose FC-FRs only allow for indifference inferences. Our investigation may provide a positive answer to this question, as we highlight in the remarks about Ch'ol in §8.4.

MORPHO-SYNTAX OF FC MARKER. Another parameter of variation across languages is the morpho-syntactic nature of the FC marker whose presence obligatorily triggers FC inferences. It can be realized as a suffix on the *wh*-word, like *-ever* in English FC-FRs. It can be realized as a prefix, as in (25) from Romanian, or a clitic, as in Ilokano in (26). Finally, it can be an independent word, as in (27) in Appalachian English. All of these morpho-syntactic options for FC markers are attested across the languages we have studied, as we discuss in §8.4.

(25) Elena detestă [FC-FR pe **ori**-cine o critică].[20] *Romanian*
Elena hates ACC FC-who her criticizes
'Elena hates anybody that criticizes her.'

(26) napintas ti boses [FC-FR ti sino=**man** nga agkankanta][21] *Ilokano*
STAT.beautiful DET voice DET who=FC COMP AP.PROG.sing
'Whoever is singing has a beautiful voice.'

(27) You should return [FC-FR **ever** what you have finished reading] *Appalachian English*
to the library.[22]
'You should return whatever you finished reading to the library.'

ORIGIN OF FC MARKER. A third parameter of crosslinguistic variation for FC-FRs has to do with the (diachronic or synchronic) source of FC markers. Whenever such source is reconstructable, it is limited to a restricted number of options. It can be derived from disjunction, as in Romanian, from a scalar/additive particle, as in English and Ilokano, from an exhaustivity operator like 'only', as in Tlaxcala Nahuatl (see Chapter 3), or from an irrealis/desiderative particle, as in Spanish with the FC suffix *-quiera*.

WH-EXPRESSIONS. The fourth and last parameter of crosslinguistic and intra-linguistic variation we examine deals with the range of *wh*-expressions which can introduce FC-FRs. In English, for instance, all *wh*-words can combine with the FC suffix *-ever*, with the exception of *why*, due to the unclear status of *whyever*. The same is observed in Romanian. In Italian, however, only the *wh*-words for 'who' (*chi-unque*), 'where' (*dov-unque*), and 'which' (*qual-unque*) can combine with the FC suffix. In some cases, the forms resulting from *wh*-root+FC have shifted

[20] From Caponigro & Fălăuş (2018: ex. (1)).
[21] From Collins (2015: ex. (10a)). Abbreviations not in the Leipzig Glossing Rules: AP: actor pivot; STAT: stativity marker. We glossed =*man* as 'FC' for consistency. Ilokano is an Austronesian language spoken in the Philippines.
[22] From Johnson (2015: ex. (2a)).

meaning and use: *com-unque* (‘how-FC’) means ‘in any way’ or ‘regardless’. In some other cases, the morphological combination is not even attested: **quand-unque* (‘when-FC’).

Another aspect of this same parameter of variation is the fact that the set of *wh*-expressions introducing FC-FRs may not be the same as the one introducing Max-FRs or Ex-FRs. In English, for instance, Max-FRs introduced by complex *wh*-phrases like *what*/*which* + NP are judged significantly degraded, while the corresponding FC-FRs are fully acceptable, as shown in (28)a vs. (28)b.

(28) a. * She met with [$_{\text{Max-FR}}$ **what**/**which** student stopped by her office].
b. She met with [$_{\text{FC-FR}}$ **whatever**/**whichever** student stopped by her office].

Therefore, in our investigation of FC-FRs and FRs in general, we have checked whether or not each *wh*-expression can occur in each kind of FRs in a given language. We discuss this research strategy further and present the findings that it has generated in §8, with the help of several comparative tables.

FC inferences and FC markers are the characterizing properties that all FC-FRs share, including those in this volume. There are at least a couple of other features that FC-FRs or their *wh*-expressions may exhibit, which we have explored in our investigation, although not as systematically: the use of subjunctive/irrealis mood inside the FC-FR and the use of an FC *wh*-expression which does not, however, introduce an FC-FR. Some languages can or even prefer to use subjunctive/irrealis in FC-FRs. For instance, in my variety of Italian (from the area around Milan), the sentence in (29)a is fully acceptable with the FC-FRs in the subjunctive, but highly degraded if in the indicative. The contrast disappears in the corresponding Max-FRs (29)b.

(29) a. Parlava con [$_{\text{FC-FR}}$ chiunque gli **desse** / **??dava** *Italian*
speak.IND.IPFV.3SG with who-FC to.him give.SUBJ.IPFV.3SG / give.IND.IPFV.3SG
retta].
attention
‘He spoke with anybody who paid attention to him.’
b. Parlava con [$_{\text{FC-FR}}$ chi gli **desse** / **dava**
speak. IND.IPFV.3SG with who to.him give.SUBJ.IPFV.3SG/ give.IND.IPFV.3SG
retta].
attention
‘He spoke with those who paid attention to him.’

A common crosslinguistic feature of *wh*-expressions with an FC marker is their absolute use: they can be used on their own without introducing FC-FRs, even when the corresponding bare

*wh*-expressions without an FC marker cannot. (30)a–b shows two examples from English, while (31) shows an example from Italian.

(30) a. I can sleep **wherever/*where**.
b. Eat **whatever/*what**!

(31) Parlava con **chiunque / *chi.**
speak.IND.IPFV.3SG with who-FC / who
'He spoke to everybody.'

This concludes our introduction to the three main varieties of FRs that are attested across languages and those of their features that have been particularly relevant for our investigation. Next we move to different kinds of [−H]RCs.

### 3.2. Other headless relative clauses

In this section, we introduce two more kinds of [−H]RCs. Like FRs, they lack a nominal head. Unlike FRs, they don't require the presence of a *wh*-expression. Light-headed relative clauses are discussed in §3.2.1, while super-free relative clauses are discussed in §3.2.2.

#### *3.2.1. Light-headed relative clauses*

*Light-headed relative clauses* (*LHRs*) are [−H]RCs with an overt D head that can be followed by a *wh*-expression, a relative marker, a complementizer, or no marker at all.[23] Their characterizing features are summarized by the syntactic schema in (32)a (inspired by the analysis in Citko 2004) and the feature bundle in (32)b.

(32) *Properties characterizing LHRs:*
a. [ D [$_{CP}$ (*wh*-/REL/COMP) …. __ …] ]$_{DP/PP}$
b. [+D, –N, ±WH]

D heads in LHRs can be of three main kinds, though they are not necessarily all instantiated in all languages with LHRs: articles, demonstratives, or quantifiers. Some D heads behave like pronominals and never take an NP complement; some others only allow an adnominal use with an NP complement when they are not introducing an LHR; and some exhibit both patterns. What all D heads share is that they occur—with or without other material—where DPs (or PPs) can occur and none of them has N features (e.g., none of them can occur as the complement of a D or can

[23] We borrow the label "light-headed relative clauses" from Citko (2004), although we expand its coverage by also including constructions that lack *wh*-expressions.

refer to a set of individuals). Examples of LHRs are given in (33)–(38), with D heads in bold and the whole LHRs in brackets. Comments follow.

(33) *Article D + COMP*
[LHR **Lo** que tú crees] no es cierto.[24] *Spanish*
the.N COMP you.SG believe.2SG not is true
'{What}/{That which} you believe is not true.'

(34) *Demonstrative D + REL*
Mi ha comprato [LHR **ciò** di cui avevo bisogno]. *Italian*
to.me has bought PRO.DEM.3SG. of REL had.1SG need
'He bought me what I needed.' (Lit. 'He bought that, for which I had need.')

(35) *Demonstrative D + WH/COMP/∅*
I like [LHR **those** who/that/∅ you like].

(36) *Demonstrative D + WH*
Jan czyta [LHR **to**, co Maria czyta].[25] *Polish*
J. reads this what M. reads
'Jan reads what Maria reads.'

(37) *Existential D + WH*
Zjawi się [LHR **ktoś,** kto ci może potóc].[26] *Polish*
appear self somebody who you can help
'Somebody who can help you will appear.'

(38) *Universal D + WH*
Pojadę [LHR **wszędzie,** gdzie rosną magnolia].[27] *Polish*
go.1SG everywhere where grow magnolias
'I will go everywhere magnolias grow.'

Example (33) shows an LHR from Spanish that is introduced by a definite article[28] followed by a declarative complementizer. The Italian example in (34) presents an LHR that is introduced by a

[24] From Plann (1980: ex. (I.1.a)). Glosses and translation have been adapted.
[25] From Citko (2004: ex. (3)).
[26] From Citko (2004: ex. (7a)).
[27] From Citko (2004: ex. (9b)).
[28] See Plann (1980: Ch. 1) for arguments that *lo* is a definite article rather than a pronoun. The article *lo* is never used with an NP complement or anywhere else in the language.

pronominal (PRO) demonstrative that can never take an NP complement, followed by an oblique relative pronoun as the complement of a preposition. Example (35) shows an example of an LHR in English that is introduced by a demonstrative D head and can be followed by a *wh*-word (and/or relative pronoun), a declarative relative complementizer, or no marker at all. Polish introduces its LHRs by means of demonstrative or quantificational D heads, as shown in (36)–(38). The demonstrative D head can take an NP complement when it is not introducing an LHR, while the quantificational D head acts like a pronominal and never allows for an NP complement. Both D heads have to be followed by a *wh*-word (rather than a *wh*-phrase). No *wh*-word can introduce a headed relative clause in Polish (Citko 2004).[29]

Overall, the emerging picture depicts LHRs as a family of constructions, rather than just a single construction. Across languages and even within the same language, there may be LHRs whose morpho-syntax is closer to that of [−H]RCs or closer to headed relative clauses or to neither. We have kept this observation in mind in our investigation and have discovered all those options for LHRs in the languages we have studied.

We briefly highlight the main differences between LHRs and FRs that we have made use of in our study in (39). We do the same for LHRs and headed relative clauses after introducing the latter in §4.

(39) *Main differences between LHRs and FRs:*

a. *D head.* This is the core difference: LHRs always have an overt D head by definition, while FRs never have one.

b. *Markers.* LHRs can be introduced by a relativizer (including a relative pronoun), a general complementizer, a *wh*-expression, a combination of those markers, or by no marker at all; FRs must be introduced by a *wh*-expression (with or without other markers) by definition.

c. *Subset of wh-expressions.* Even when LHRs are introduced by *wh*-expressions, as in Polish, the subset of admissible *wh*-expressions in LHRs doesn't have to be the same as the one in FRs (in Polish, *why* can introduce LHRs but not FRs; Citko 2004: fn. 7).

d. *Meaning.* The semantic behavior of LHRs is determined by their D head: if their D head is definite, they are referential, like definite DPs; if their D head is indefinite, they denote whatever indefinite DPs are assumed to denote; if their D head is a quantifier, they denote

[29] Citko (2004) specifically shows that the *wh*-words for 'who', 'what', 'where', 'when', 'how', and 'why' can all introduce LHRs in Polish, while the *wh*-word for 'which' cannot. On the other hand, 'which' is the only *wh*-word that can occur in headed relative clauses in Polish.

what quantificational DPs denote.[30] By contrast, FRs exhibit a fixed semantic behavior, as discussed in §3.1: they behave like definite DPs (Max-FRs), unless they are in the complement position of existential predicates (Ex-FRs) or their *wh*-expressions are morpho-syntactically enriched by an FC marker (FC-FRs).

e. *FC marking*. A clear difference between LHRs and FC-FRs is that the presence of a D head is incompatible with an FC marker, as shown in (i) vs. (ii):
   i. I chose [$_{LHR}$ [$_{D}$ that] [**which** is on the table]].
   ii. * I chose [$_{LHR}$ [$_{D}$ that] [**whichever** is on the table]].

### *3.2.2. Super-free relative clauses*

*Super-free relative clauses* (*SFRs*) are the last variety of [−H]RCs that we investigate. Like all [−H]RCs, they lack an N head. Unlike LHRs, they lack a D head as well, resembling FRs. But they are even "freer" than FRs in lacking a *wh*-expression as well. This is why we label them "super-free" in this chapter, although a slightly different terminology is adopted in other chapters, depending on language-specific features and authors' choices. The syntactic structure we have assumed for SFRs is summarized in (40)a and their main features are bundled in (40)b.

(40) *Properties characterizing SFRs:*
   a. [$_{CP}$ (REL/COMP) …. __ …]$_{DP/PP}$
   b. [–D, –N, –WH]

SFRs can be introduced by a non-*wh* relativizer, as in (41), by a general complementizer, as in (42), or by no marker at all, as in (43). Crucially, the bracketed in SFRs in (41)–(43) are all full, tensed clauses, rather than nominalized tenseless clauses or subclausal constituents.

(41) Bhí [ a raibh __ san Oileán ] ag féachaint ar na naomhóga.[31] *Irish*
   was REL was in.the Island look [PROG] on the currachs
   'Everyone who was in the Island was watching the currachs.'

[30] Indefinites have been argued to denote existential generalized quantifiers, properties, and/or individuals. Quantificational DPs are usually assumed to denote various kinds of generalized quantifiers. See Szabolcsi (2010) for a thorough overview of the semantics of indefinites and quantificational DPs, including the definition of 'generalized quantifier'. A terminological clarification may be useful. Within the referential theory of meaning we are adopting, every linguistic expression *refers to* or, equivalently, *denotes* something in the world, including abstract mathematical objects like sets, functions, and generalized quantifiers. However, in this chapter we restrict the use of the verb *refer*, the adjective *referential*, and the noun *reference* to linguistic expressions that refer to/denote individuals.

[31] In the source of this example, McCloskey 2002 (ex. (55a)), *a* is glossed *aN*. In McCloskey's analysis, *aN* is not a *wh*-word, but instead a complementizer associated with a *wh*-dependency created by binding rather than movement. It can introduce headed and headless relative clauses and *wh*- interrogative clauses, but not declarative clauses. We chose to gloss it as "relativizer" for consistency with what we have been assuming so far.

(42) Għamil-t [ li għid-t-l-i.].[32] *Maltese*
do.PFV-1SG COMP say.PFV-2SG-DAT-1SG
'I did what you told me.'

(43) [ máki-and'əh nĭh=yɨʔ=ʔĭh ni-ĕ-ew-ĭt]=yɨʔ-íʔ ?[33] *Hup*
Mark-ASSOC.PL POSS=FOC=M be-PVF-FLR-OBL=FOC-INT
'It was with [that one who used to be associated with Mark's group] (that you went)?'

This concludes the introduction to the main characters in this volume—the varieties of [−H]RCs that we have discovered in the languages we have investigated. Next, we briefly introduce the two supporting characters that are closely related to [−H]RCs—headed relative clauses (§4) and *wh*- interrogative clauses (§5). Both constructions are wide-spread across languages and have received much more attention in the literature than [−H]RCs. Our main goal is just to touch on some of their core properties in order to highlight the similarities and differences with [−H]RCs that we have kept in mind in our investigation.

## 4. Headless relative clauses vs. headed relative clauses

In this section, we briefly compare the varieties of [−H]RCs we have investigated with a better-known family of relative clauses: headed relative clauses. We focus on the subset of headed relative clauses known as *restrictive externally-headed relative clauses* (*[+H]RCs*, for short).[34] [+H]RCs can be characterized for our purposes by the properties in (44).

(44) *Properties characterizing [+H]RCs:*

I. They are embedded clauses.

II. They lack a constituent—an argument or an adjunct (which can sometimes be marked by a resumptive pronominal form).

[32] From Sadler & Camilleri (2018: ex. (31)). Maltese is a Semitic language spoken in the Mediterranean island of Malta.

[33] From Epps (2012: ex. (7)). Glosses have been minimally adapted for consistency. Abbreviations not in the Leipzig Glossing Rules: ASSOC.PL: associative plural; FLR: filler; INT: interrogative; TAG: interactive tag. Hup is one of the four Nadahup languages, spoken by the Hupda and Yohup, indigenous Amazonian peoples who live in Brazil and Colombia.

[34] Internally headed relative clauses lack a gap, a *wh*-expression, and a relative pronoun, which makes them crucially different in their morpho-syntax from [+H]Rs (and [−H]Rs) and thus easy to distinguish. Appositive headed relative clauses exhibit a different semantic behavior from [+H]Rs: they do not restrict the meaning of their nominal head, but add an extra qualification to it, usually at a level that doesn't affect the truth conditions of the sentence.

III. They have an "external nominal head"—a nominal head that precedes or follows them and is linked to the missing constituent.

IV. Their distribution and interpretation resemble those of nominal modifiers like adjective phrases (AdjP) and PPs acting as adjuncts within DPs.

Properties I and II in (44) are shared with [−H]RCs (cf. (1)I–II): both families of constructions involve embedded clauses with a missing constituent (or a resumptive pronoun). Properties III and IV, by contrast, distinguish [+H]RCs from [−H]RCs (cf. (1)III–IV). Unlike [−H]RCs, [+H]RCs are preceded or followed by an overt nominal head, which they restrict semantically.[35] For instance, (45) shows the bracketed [+H]RC restricting the meaning of the nominal head *music*: Frida doesn't love music in general; what she loves is the subset of music that comes from Mexico. This semantic function is similar to that of the bracketed AdjP in (46) or the bracketed PP in (47), except that the richer clausal structure of [+H]RCs allows for the creation of semantically richer and more articulated nominal modifiers. For instance, it wouldn't be possible to semantically restrict *music* in the way the [+H]RC in (48) does by way of an AdjP or a PP.

(45) Frida loves [$_{N}$ <u>music</u> ] [$_{[+H]RC}$ that comes from Mexico].

(46) Frida loves [$_{AdjP}$ truly Mexican] [$_{N}$ <u>music</u> ].

(47) Frida loves [$_{N}$ <u>music</u> ] [$_{PP}$ from Mexico].

(48) Frida loves [$_{N}$ <u>music</u> ] [$_{[+H]RC}$ that her father played for her mother and her when Frida was a little girl].

[+H]RCs are widely attested. Some languages make an even more extensive use of them to convey the meaning that is conveyed by embedded declarative or interrogative clauses in most other languages.[36] There is a very large descriptive, typological, and theoretical literature devoted to [+H]RCs.[37] Our characterization of [+H]RCs in (44) doesn't aim to summarize these findings, but only highlights the properties of [+H]RCs that are most relevant for our comparison with [−H]RCs, as we shall discuss next.

[+H]RCs as characterized by the properties in (44) can easily be distinguished from any of the varieties of [−H]RC which we presented in §3: all [+H]RCs have a nominal head, while no [−H]RC does (Property III). This implies that [+H]RCs never occur as arguments on their own, but only together with their N head (Property IV).

[35] The original idea of treating [+H]RCs as set restrictors is first sketched in Quine (1960) and then fully implemented in Montague (1970). See Chierchia & McConnell-Ginet (2000) for an introductory presentation.
[36] Cf. Adyghe, a Northwest Caucasian language discussed in Caponigro & Polinsky (2011).
[37] See Andrews (2007) and de Vries (2018) for two overviews focusing on the typology and the syntax of [+H]RCs, and the literature therein.

[+H]RCs and [−H]RCs can make use of similar morpho-syntactic devices, though. Depending on the language, [+H]RCs can be introduced by *wh*-expressions, like FRs and LHRs. Still, it cannot be taken for granted that the *wh*-expressions introducing [+H]RCs are necessarily the same subset of *wh*-expressions which introduce [−H]RCs. For instance, the *wh*-word *which* in English can introduce [+H]RCs but not FRs, while the *wh*-word *what* can productively introduce FRs in all varieties of English we are aware of, but is banned in [+H]RCs (at least in most varieties of English).

[+H]RCs can also be introduced by a special set of relativizers (including relative pronouns) or general complementizers or involve just a gap with no markers, as is possible with LHRs or SFRs as well. In a language in which LHRs and SFRs are introduced by the same devices as [+H]RCs, a question arises as to whether this language has true LHRs and SFRs or just some [+H]RCs with a "silent" nominal head. We have conducted our investigation under the assumption that silent N heads need to be licensed by syntactic and semantic/pragmatic conditions, as is the case for all silent elements. Whenever we did not observe constraints of this kind, we assumed that LHRs and SFRs were constructions different from [+H]RCs.

We conclude this section by quickly mentioning a construction that is often subsumed under the large umbrella concept of (headed) relative clauses: correlative clauses.[38] An example from Romanian is given in brackets in (49).

(49) [**Cine** __ m-a rănit odată], în *acela* nu mai am încredere.[39] *Romanian*
who me-has hurt once in DEM.3SG.M not anymore have.1SG trust
'I don't trust anyone who hurt me.'

Correlative clauses are embedded clauses that, like [+H]RCs and [−H]RCs, have a gap (the subject in (49)) and a marker in clause-initial position associated with the gap (the bolded *wh*-word *cine* 'who' in (49)). Unlike [−H]RCs, though, correlative clauses are always dislocated—sentence initially or sentence finally (the correlative clause is sentence initial in (49)). Also unlike [−H]RCs, their matrix clause contains a demonstrative pronoun which is linked to the correlative clause (the italicized demonstrative pronoun *acela* in (49)). Last, the set of elements introducing correlative clauses is not necessarily identical to the one that introduces [−H]RCs or [+H]RCs. In the specific case of (49), the correlative clause is introduced by the *wh*-word *cine* 'who' and the whole sentence is fully acceptable. On the other hand, sentences with Max-FRs introduced by *cine* have a degraded level of acceptability in Romanian, while those with [+H]RCs introduced by *cine* are completely unacceptable.

[38] See Liptak (2009) for an overview of correlative clauses crosslinguistically and issues related to them.
[39] Thanks to Anamaria Fălăuș for the example.

## 5. Headless relative clauses vs. *wh-* interrogative clauses

In this section, we compare the varieties of [−H]RCs of §3 with another well-known construction, which may look identical to FRs in some languages—embedded *wh-* interrogative clauses.[40] They can be characterized by the properties in (50).

(50) *Properties characterizing embedded wh- interrogative clauses:*

I. They are embedded clauses.

II. They lack a constituent—an argument or an adjunct (which can sometimes be marked by a resumptive pronominal form).

III. They are introduced by a *wh*-expression (or more than one).

IV. They occur as the argument of interrogative predicates, i.e., predicates selecting for an argument denoting a question or a proposition.

Properties I and II are shared with all [−H]RCs (cf. (1)I–II). Property III is shared with FRs (see (9)). Property IV crucially differentiates *wh-* interrogative clauses[41] from all [−H]RCs and [+H]RCs. Let's discuss them further by looking at some examples of *wh-* interrogative clauses and comparing them with FRs.

(51) a. Frida asked [$_{\text{Interrogative}}$ what Chavela sang __ ].
b. Paloma told Frida [$_{\text{Interrogative}}$ what Chavela sang __ ].
c. Frida played [$_{\text{Max-FR}}$ what Chavela sang __ ].

The bracketed clauses in (51)a–b are identical *wh-* interrogative clauses that are embedded (Property I), lack their direct objects (Property II), and are introduced by the *wh*-word *what* (Property III). These properties are all shared with the (superficially) identical Max-FR in (51)c. It is Property IV, i.e., the selectional properties of the matrix predicate, that makes the difference. The matrix verb *asked* in (51)a selects for a complement denoting a question; the matrix verb *told* in (51)b selects for a complement denoting a proposition; the matrix verb *played* in (51)c selects for a complement denoting an individual, in particular an object that can be played. It follows that the three identical *wh-* clauses in (51)a–c receive three different interpretations: those of a question, a proposition, and an individual, respectively. Current semantic analyses of interrogative clauses have established a close connection between the semantic object labeled "question" and

[40] This discussion of *wh-* interrogative clauses has benefited from Judith Aissen's lecture on interrogative clauses at the first workshop that we organized as part of the "Headless Relative Clauses in Mesoamerican Languages" project.
[41] Henceforth, "*wh-* interrogative clause" will be used as short for "embedded *wh-* interrogative clause."

the semantic object labeled "proposition": a question can be defined as a set of propositions.[42] Therefore, we can assume that the *wh-* clauses in (51)a–b are syntactically the same and still derive the appropriate semantic differences. The *wh-* clause in (51)c, by contrast, denotes an individual. This is a semantic object that is quite different from a proposition or a question. It is also the same semantic object that is denoted by Max-FRs, as we discussed in §3.1.1. Therefore, the *wh-* clause in (51)c is a Max-FR. But how can one distinguish between *wh-* interrogative clauses and Max-FRs in languages in which they look morpho-syntactically identical without having to rely on intuitions about meaning that may be hard to grasp or share?

A similar issue can arise for the distinction between *wh-* interrogative clauses and Ex-FRs. (52)a shows an example of an embedded *wh-* interrogative clause in Spanish (in brackets) occurring as the complement of 'know', while (52)b shows an example of an Ex-FR (in brackets) in the complement position of existential 'have'.

(52) a. Frida no sabe [$_{\text{Interrogative}}$ adónde ir]. *Spanish*
Frida not knows to.where go.INF
'Frida doesn't know where to go.'
b. Frida no tiene [$_{\text{Ex-FR}}$ adónde ir].
Frida not has to.where go.INF
'Frida doesn't have a place where she can go.'

Since the two *wh-* clauses in (52)a–b look identical, how can one (i) conclude that they are different and (ii) determine which one is which, without relying on intuitions about their meaning? In our investigation, we made use of two main tests to distinguish between *wh-* interrogative clauses and FRs in languages where they superficially look the same: the Embedding Predicate Test and the Substitution Test. We discuss each in turn.

(53) *The Embedding Predicate Test*
A *wh-* clause is an FR if it occurs as the argument Y of a predicate that only selects for a Y denoting an individual (e.g., both arguments of *love, build, buy, make*, etc.) or a predicate that only selects an indefinite or a quantificational Y (e.g., the complement of existential predicates 'be' and 'have'). A *wh-* clause is an interrogative clause if it occurs as the argument X of a predicate that only selects for an X denoting a question (e.g., the complement of *ask, wonder*).

The Embedding Predicate Test not only offers guidance in examining naturally produced data or corpora, but is also useful in constructing examples whose acceptability can be tested with

[42] See Dayal (2016) for a recent and detailed survey on the semantics and pragmatic of interrogative clauses.

consultants. It is important to choose embedding predicates that unambiguously select for one type of construction. For instance, in a language like English, predicates like *ask* and *love* and their complement positions are excellent to distinguish a *wh*- interrogative clause from an FR, as we have just discussed. On the other hand, a predicate like *know* is lexically ambiguous between the meaning of 'be able to answer the question' and 'be acquainted with', as shown in example (54). The *wh*- clause in (54)a has to be a *wh*- interrogative clause if the sentence is interpreted as in (54)b. It has to be an FR if the sentence receives the interpretation in (54)c.

(54) a. Frida knows [$_{\text{Interrogative/FR}}$ what Chavela sang __ ].
b. "INTERROGATIVE" MEANING: Frida can answer the question: "What did Chavela sing?"
c. "FR" MEANING: Frida is acquainted/familiar with the song that Chavela sang.

Therefore, we preferred not to use predicates like *know*—or not to rely exclusively on them—to draw any solid conclusions. Luckily, the equivalent of *know* in many languages is realized by two different forms: one that only takes a complement denoting a question/proposition (e.g., *saber* in Spanish or *sapere* in Italian) and another one that only takes a complement denoting an individual (e.g., *conocer* in Spanish or *conoscere* in Italian).
The other test we adopted is the Substitution Test, given in (55).

(55) *The Substitution Test*
A *wh*- clause is an FR if it can be replaced by a nominal expression referring to an object (a DP like *the beautiful songs*) or a set of objects (an NP like *beautiful songs*). A *wh*- clause is a *wh*- interrogative clause if it can be replaced by a polar (i.e., *yes/no*) interrogative clause.

For instance, we can conclude that the bracketed *wh*- clause in (56)a is a *wh*- interrogative clause because it can be replaced by the polar interrogative clause in (56)b, resulting in a fully acceptable and interpretable sentence. On the other hand, replacing the *wh*- clause with a definite DP results in an unacceptable sentence, as shown in (56)c.

(56) a. Frida wonders [$_{\textit{Wh}\text{- Interrogative}}$ what Chavela sang __ ].
b. Frida wonders [$_{\text{Polar Interrogative}}$ if Chavela sang *La Llorona*].
c. * Frida wonders [$_{\text{Definite DP}}$ that song— *La Llorona*].

The opposite pattern holds with FRs. We can concluded that the *wh*- clause in (57)a is a Max-FR because it can be replaced by the definite DP in (57)b, but not by the polar interrogative in (57)c.

(57) a. Frida loves [$_{\text{Max-FR}}$ what Chavela sang __ ].
b. Frida loves [$_{\text{Definite DP}}$ that song— *La Llorona*].
c. *Frida loves [$_{\text{Polar Interrogative}}$ if Chavela sang *La Llorona*].

The Spanish examples in (52)a and (52)b, repeated in (58)a and (59)a for convenience, illustrate a *wh-* interrogative clause, (58)a, which is superficially identical to an Ex-FR, (59)a. The Substitution Test helps distinguish between the two in this case as well.

(58) a. Frida no sabe [$_{\text{Interrogative}}$ adónde ir]. *Spanish*
Frida not knows to.where go.INF
'Frida doesn't know where to go.'
b. Frida no sabe [$_{\text{Polar Interrogative}}$ si tiene que irse a casa].
Frida not knows if must.3SG COMP go.INF to home
'Frida doesn't know whether she should go home.'
c. *Frida no sabe [$_{\text{DP}}$ lugares seguros para ir].
Frida not knows places safe for go.INF
['Frida can't answer the question: "Safe places where to go."']

(59) a. Frida no tiene [$_{\text{Ex-FR}}$ adónde ir]. *Spanish*
Frida not has to.where go.INF
'Frida doesn't have a place where she can go.'
b. *Frida no tiene [$_{\text{Polar Interrogative}}$ si tiene que irse a casa].
Frida not has if must.3SG go.INF to home
['Frida doesn't have whether she should go home.']
c. Frida no tiene [$_{\text{DP}}$ un lugar seguro para ir].
Frida not has a place safe for go.INF
'Frida doesn't have a safe place where she can go.'

A polar interrogative clause can replace the *wh-* clause in (58)a, as shown in (58)b, while a DP cannot, as shown in (58)c. Therefore, the *wh-* clause in (58)a must be a *wh-* interrogative clause. On the other hand, a DP can replace the *wh-* clause in (59)a, as shown in (59)c, while a polar interrogative clause cannot, as shown in (59)b. It follows that the *wh-* clause in (59)a is an FR.

A language-specific variant of the Substitution Test can be constructed by using those *wh*-expressions that introduce only *wh-* interrogative clauses. Most languages only allow a proper subset of *wh*-expressions to introduce both *wh-* interrogative clauses and FRs. In English, for instance, complex *wh*-phrases like *which*/*what* + NP can introduce only *wh-* interrogative clauses, not FRs, as shown in (60)a–b, respectively.

(60) a. Frida wonders [$_{wh\text{- Interrogative}}$ **what/which song** Chavela sang __ ].
b. *Frida loves [$_{wh\text{- Interrogative}}$ **what/which song** Chavela sang __ ].

For English, then, the Substitution Test can also make use of *wh*- interrogative clauses introduced by *which*/*what* + NP, rather than just polar interrogative clauses.

## 6. Why study headless relative clauses crosslinguistically in Mesoamerican languages

After defining [−H]RCs, the main subject of our investigation, introducing its various manifestations, and showing that they all differ from [+H]RCs and interrogative clauses, we can now return to some broader questions behind our project to further elaborate on them: Why should we study [−H]RCs to begin with, why do so crosslinguistically, and why look at Mesoamerican languages in particular?

In the previous sections, we have defined [−H]RCs as embedded clauses that may resemble more common and better-studied kinds of embedded clauses morpho-syntactically, but differ from them distributionally and semantically. Unlike other embedded clauses, [−H]RCs occur in positions where we would otherwise find non-clausal constituents, such as DPs, NPs, and PPs. Unlike embedded declarative clauses and various kinds of clausal adjuncts, [−H]RCs do not convey propositional content—they are not about facts or states of affairs. Unlike embedded interrogative clauses, they do not convey a question (a set of propositions) or an answer to question (a proposition). [−H]RCs refer to individuals like definite DPs, or sets of individuals like NPs, or trigger quantification over individuals like quantificational DPs, or impose locative, temporal, manner, and other kinds of restrictions on events as PPs can do. This is non-trivial semantic behavior, which sheds new light on the complex mapping between syntax and semantics in natural languages and the extreme richness and flexibility of clauses in natural language—not just on the morpho-syntactic side but also semantically. We return to this issue in §9.

Those [−H]RCs that are introduced by *wh*-expressions contribute significantly to our understanding of the meaning of *wh*-expressions across constructions and across languages. Most proposals about the semantic contributions of *wh*-expressions are primarily (if not exclusively) grounded in their behavior in *wh*- interrogative clauses. Recent work[43] has highlighted the fact that the systematic use of *wh*-expressions in [−H]RCs calls for an account which explains how their semantic contribution in [−H]RCs relates to their semantic contribution in *wh*- interrogative clauses: if the two are not the same, the relation has to be systematic and principled, given the regularity with which *wh*-expressions introduce both clause types across languages, which we discuss further in §9. More data and more analytical work are needed from a variety of languages, especially from those outside the better-known Indo-European family. Our volume contributes to this need.

[43] Caponigro (2003), Caponigro & Fălăuș (2018), Caponigro & Fălăuș (2020).

A crosslinguistic perspective is crucial to reaching a better understanding of what is general and what is language-specific in the morpho-syntactic and semantic behavior of [−H]RCs and their associated *wh*-expressions. Our project and volume contribute to this endeavor by providing a systematic and rich description of [−H]RCs in fifteen languages—all Mesoamerican but one. We also offer a model for how this kind of research can be conducted by providing definitions, tools, and tests. Although relatively little is known about the details of [−H]RCs in most Mesoamerican languages, the available work indicates that they are found in most (if not all) of the major language phyla of the region, including those in (61).[44]

(61) *Mesoamerican languages where [−H]RCs have been attested so far:*[45]

UTO-AZTECAN:

various languages (Langacker 1979), Southeastern Tepehuan (Willett 1991), Nahuatl (Beller 1979; Brockway 1979; Sischo 1979; Tuggy 1979; Hill & Hill 1981; Flores-Nájera 2019)

OTO-MANGUEAN:

Chinantec (Westley 1991), Mixtec (Caponigro et al. 2013), Otomi (Voigtlander & Echegoyen 1985), Tlapanec (Carrasco Zúñiga 2006), Zapotec (Ramos 2015)

MAYAN:

Kaqchikel (García Matzar and Rodríguez Guaján 1997; Torrence et al. 2016; Guarcax González 2016), Tseltal (Polian 2013), Yucatec (Gutiérrez Bravo 2012, 2015)

MIXE-ZOQUEAN:

Sierra Popoluca (Boudreault 2009; Jiménez Jiménez 2019)

TOTONACAN:

Highland Totonac (Aschmann 1984), Upper Necaxa Totonac (Beck 2016); Coahuitlán Totonac (Moore 2016)

ISOLATE:

Purepecha (Castro 2013)

Even when [−H]RCs are explicitly discussed, the focus is on their morpho-syntax, while their semantic properties are usually not mentioned.

[44] See also Comrie & Estrada-Fernández (2012), which investigates patterns of relative clause formation across languages from the Americas and dedicates two chapters to some varieties of [−H]RCs in Hup (Epps 2012; see ex.(43) and fn. 33) and Yucatec Maya (Gutiérrez Bravo 2012); Gutiérrez Bravo (2015), which contains a chapter on LHRs and one on FRs; and Palancar et al. (forthcoming), which focuses on relative clauses in Mesoamerican languages, including some [−H]RCs.

[45] Thanks to Roberto Zavala and Harold Torrence for (61) and the bibliographical references therein.

This volume presents the collective work of a team of twenty-one scholars who have investigated [−H]RCs in fifteen languages from five language families. The languages that we have studied and the scholars who have studied them are listed in (62), in the same order and grouping as the remaining chapters in the volume (with the chapter numbers in parentheses).

(62) *Languages investigated in this volume and scholars investigating them:*

**Mesoamerican**

UTO-AZTECAN

*Southeastern Tepehuan* (*O'dam*), by Gabriela García Salido (Chapter 2)

*Tlaxcala Nahuatl*, by Lucero Flores-Nájera (Chapter 3)

OTO-MANGUEAN

*Acazulco Otomi*, by Néstor Hernández-Green (Chapter 4)

*Matlatzinca*, by Enrique Palancar and Leonardo Carranza Martínez (Chapter 5)

*Iliatenco Me'phaa*, by Philip T. Duncan and Harold Torrence (Chapter 6)

*San Pedro Mixtepec Zapotec*, by Pafnuncio Antonio Ramos (Chapter 7)

MAYAN

*K'iche'*, by Telma Angelina Can Pixabaj (Chapter 8)

*Q'anjob'al*, by Eladio Mateo Toledo (B'alam) (Chapter 9)

*Chuj*, by Justin Royer (Chapter 10)

*Ch'ol*, by Juan Jesús Vázquez Álvarez and Jessica Coon (Chapter 11)

*Tsotsil* and *Tseltal* (*Tseltalan*), by Gilles Polian and Judith Aissen (Chapter 12)

*Yucatec Maya*, by Scott AnderBois and Miguel Oscar Chan Dzul (Chapter 13)

MIXE-ZOQUEAN

*Sierra Popoluca* , by Wendy López Márquez (Chapter 14)

**Non-Mesoamerican**

CHIBCHAN

*Pesh*, by Claudine Chamoreau (Chapter 15)

## 7. Language-specific chapters and their structure

All of the language-specific chapters follow the same structure. We made this choice in order to ensure that exactly the same ground would be covered in each language and to facilitate comparison across languages and chapters. This strategy also allows us to use this introductory chapter as the place in the volume where all the relevant constructions are defined, differences are discussed, methodologies and tests are introduced, and broader comparative remarks are presented. In this way, each language-specific chapter can focus just on the investigation of how [−H]RCs and related constructions manifest themselves in that language without taking space away from new language data and generalization by repeating the same introductory remarks and definitions

over and over. It follows that each language-specific chapter crucially presupposes this introductory chapter, which should be kept in mind while reading the chapters that follow.

The uniformity across chapters and the advantages we have just mentioned were made possible by two main factors. First, even if they had not studied [−H]RCs before, all the participants came to the project with a deep knowledge of the structure of their language. This allowed us to discuss ahead of time definitions and strategies that would be compatible with all the languages and potential issues. Second, we organized two extended preparatory workshops as core components of our project. In the first workshop, we introduced the main characters and related supporting characters of our investigation, discussed methodological and definitional issues, and agreed on the overall structure of the chapters. In the second workshop, six months later, we presented the results, discussed them together, and revised them based on collective and individual feedback.

The remainder of this section consists of a sketch of the template and the terminology that all chapters in the volume follow. Each chapter may introduce further subsections, with more or less emphasis on specific issues, and some terminological particularities.

*Section 1. Introduction and Basic Features of Language X*
This section presents basic information about the language such as where it is spoken, the number of speakers, how the data were collected (elicitation, texts), and relevant phonological, morphological, and syntactic features.

*Section 2. Wh- Interrogative Clauses in Language X*
This section lists the *wh*-expressions in the language and discusses the morpho-syntax of *wh*- interrogative clauses. Specifically, the section discusses the *wh*-strategies employed by the language (e.g., *wh*-movement vs. *wh* in-situ), matrix vs. embedded interrogative clauses, multiple-*wh* interrogative clauses, the (un)availability of a clefting strategy, and so-called pied piping with inversion.

*Section 3. Headed Relative Clauses in Language X*
This section lays out the strategies that the language uses to form [+H]RCs. The focus is on situating the language's [+H]RCs in the typological and crosslinguistic landscape. The section addresses topics such as relative pronouns vs. other relative markers, the use of *wh*-expressions as relative pronouns, and the use of gaps vs. resumptive pronouns to mark the missing constituent within [+H]RCs.

*Section 4. Headless Relative Clauses in Language X*
This section presents the varieties of [−H]RCs that are attested in the language. It assumes the definitions and distinctions that are provided in the introductory chapter. If there is a type of

[−H]RC that is listed in the introductory chapter but is not found in the language, this is explicitly noted. The discussion specifically steps the reader through the following:

*Section 4.1. Free Relative Clauses in Language X*
This section presents all the varieties of FRs that are attested in the language, in particular:

*Section 4.1.1. Maximal Free Relative Clauses*

*Section 4.1.2. Existential Free Relative Clauses*

*Section 4.1.3. Free-Choice Free Relative Clauses*

*Section 4.2. Light-Headed Relative Clauses in Language X*
This section presents all the varieties of LHRs that are attested in the language and the kinds of light heads introducing them.

*Section 4.3. Super-Free Relative Clauses in Language X*
This section presents all varieties of SFRs that are attested in the language, if any.

*Section 5. Conclusions*
This section summarizes the main findings with special emphasis on commonalities and differences among the different [−H]RCs and between [−H]RCs and related constructions. Specific findings about the language are highlighted as well.

## 8. Main comparative findings

We shall not attempt to provide a summary here of the content of each of the subsequent chapters. It would not be possible to do justice to the languages and the respective discussions, since each chapter provides rich and dense language-specific information, with data that may look unclear or be misunderstood if presented out of context. Instead, this section provides the main *comparative* findings emerging from the fifteen languages we have investigated—a perspective that is present only to a limited extent in the individual chapters, given their language-specific focus. Nevertheless, the reader should keep in mind that the crosslinguistic, typological findings that we are about to present can be fully understood and appreciated only by going through each chapter. Similarly, the tables in this section only offer a rough summary of the distributional data, unavoidably involving approximations or simplifications. The reader should refer back to the specific chapters for more detailed tables and supporting examples, descriptions, and generalizations.

We proceed as follows in this section. First, we touch on the distribution of the different varieties of [−H]RCs across the languages we have investigated (§8.1). Then, we highlight our crosslinguistic findings and summarize our main language-specific findings on Max-FRs (§8.2), Ex-FRs (§8.3), FC-FRs (§8.4), the *wh*-expressions that introduce them (§8.5), and the other [−H]RCs (§8.6).

### 8.1. Productivity of headless relative clauses

Headless relative clauses are extremely productive in the fifteen languages we have investigated, as summarized in Table 1.1 All of these languages have Max-FRs and Ex-FRs. Southeastern Tepehuan, Tsotsil, and Tseltal lack FC-FRs, while the situation is uncertain in Ch'ol (Chapter 11: §4.1.3). All the languages have LHRs. SFRs as well are attested in all but two of them: Southeastern Tepehuan and Acazulco Otomi. The other three Oto-Manguean languages we have investigated do allow for SFRs, and several other Otomi languages allow for SFRs as well (Chapter 4: §5). Finally, Pesh uses SFRs with a maximal interpretation and frequency that resemble those of Max-FRs in other languages. The pattern that is summarized in Table 1.1 shows both intra-linguistic and crosslinguistic variation.

**Table 1.1. Varieties of headless relative clauses in the languages investigated in this volume**

| | | UTO-AZT | | OTO-MANGUEAN | | | | MAYAN | | | | | | MI-ZO | CHI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Te** | **Na** | **AO** | **Ma** | **IM** | **SZ** | **K'** | **Q'** | **Cj** | **Cl** | **TT** | **YM** | **SP** | **Pe** |
| | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| **FR** | **Max** | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| | **Ex** | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| | **FC** | * | √ | √ | √ | √ | √ | √ | √ | √ | ? | * | √ | √ | √ |
| **LHR** | | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| **SFR** | | * | √ | * | √ | √ | √ | √ | ? | √ | ? | ? | ? | √ | √ |

NOTE: **Marks:** √: attested; *: not attested; ?: unclear if attested. **Language family name abbreviations:** UTO-AZT: Uto-Aztecan; MI-ZO: Mixe-Zoquean; CHI: Chibchan. **Language name abbreviations:** Te: Southeastern Tepehuan; Na: Tlaxcala Nahuatl; AO: Acazulco Otomi; Ma: Matlatzinca; IM: Iliatenco Me'phaa; SZ: San Pedro Mixtepec Zapotec; K': K'iche'; Q': Q'anjob'al; Cj: Chuj; Cl: Ch'ol; TT: Tsotsil and Tseltal; YM: Yucatec Maya; SP: Sierra Popoluca; Pe: Pesh. **Numbers:** the number under the language name abbreviation indicates the chapter in the volume where the language is discussed.

Clearly, just because a language has one kind of [−H]RC, it doesn't follow that it will have all the others as well. Meanwhile, the fact that a language has all kinds of [−H]RCs does not ensure that a related language from the same language family—or even the same language group—will also have all types of [−H]RCs. Therefore, the study of [−H]RCs in a language requires a detailed investigation of each variety – as much so in this comparative analysis as in the language-specific chapters that follow. Thus, we now look more closely at each type of [−H]RC and the specific patterns that emerge.

### 8.2. Findings on maximal free relative clauses

All fifteen languages we have investigated have Max-FRs, but with different levels of productivity, if measured by the relative size of the subset of *wh*-expressions found in *wh*-interrogative clauses that can occur in Max-FRs. Our findings are summarized in Table 1.2.

**Table 1.2. Use of *wh*-expressions in maximal free relative clauses across the languages investigated in this volume**

| | UTO-AZT | | OTO-MANGUEAN | | | | MAYAN | | | | | | M-Z | CHI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Te** 2 | **Na** 3 | **AO** 4 | **Ma** 5 | **IM** 6 | **SZ** 7 | **K'** 8 | **Q'** 9 | **Cj** 10 | **Cl** 11 | **TT** 12 | **YM** 13 | **SP** 14 | **Pe** 15 |
| 'who' | √ | √ | √ | √ | % | √ | √ | √ | √ | √ | √ | √ | √ | * |
| 'what' | √ | √ | √ | √ | * | √ | √ | √ | √ | √ | √ | √ | √ | * |
| 'where' | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| 'when' | √ | * | * | * | √ | -- | * | √ | * | √ | √ | -- | √ | * |
| 'why' | √ | -- | * | √ | √ | √ | * | -- | % | * | * | * | * | * |
| 'how' | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | * |
| 'what/which' (+NP) | * | √/* | * | * | √ | √ | √ | √ | √ | √ | (√) | * | √ | * |
| 'how much/many' (+NP) | √ | * | * | * | √ | √ | √ | √ | √ | √ | √ | ? | √ | * |

NOTE: **Marks:** √: attested; *: not attested; ?: unclear if attested; --: no simple *wh*-word conveying the relevant meaning; %: speaker variation. **Shading:** languages in which *wh*-expressions in Max-FRs are morpho-syntactically different from those in interrogative clauses. **Language family name abbreviations:** UTO-AZT: Uto-Aztecan; MI-ZO: Mixe-Zoquean; CHI: Chibchan. **Language name abbreviations:** Te: Southeastern Tepehuan; Na: Tlaxcala Nahuatl; AO: Acazulco Otomi; Ma: Matlatzinca; IM: Iliatenco Me'phaa; SZ: San Pedro Mixtepec Zapotec; K': K'iche'; Q': Q'anjob'al; Cj: Chuj; Cl: Ch'ol; TT: Tsotsil and Tseltal; YM: Yucatec Maya; SP: Sierra Popoluca; Pe: Pesh. **Numbers:** the number under the language name abbreviation indicates the chapter in the volume where the language is discussed.

Two languages from different language families allow all *wh*-expressions to introduce Max-FRs: San Pedro Mixtepec Zapotec and Q'anjob'al. Three languages from different language families—Southeastern Tepehuan, Iliatenco Me'phaa, and Ch'ol— allow all *wh*-expressions but one. All the languages allow for most of their *wh*-expressions to occur in Max-FRs, except for Pesh, which only allows the *wh*-expression for 'where'.

The most common *wh*-expressions that can introduce Max-FRs are 'where', which is found in all languages under study, and 'who' and 'how', which are banned only in Pesh. By contrast, 'why' is the least common *wh*-expression—a pattern that is also attested in Indo-European, Finno-Ugric, and Semitic (e.g., Caponigro 2003: §1.4).[46] Two of the Mesoamerican languages covered

[46] To the best of our knowledge, Romanian (Caponigro & Fălăuș 2018) and Teramano (Mantenuto & Caponigro 2020) are the only Indo-European languages for which it has been reported that 'why' can introduce Max-FRs.

do not even have a simple *wh*-word corresponding to 'why', not even in interrogative clauses. Still, four of the languages from two different language families allow for 'why' to introduce Max-FRs, and one more language from a third language family exhibits speaker variation with respect to 'why' in Max-FRs. Overall, our investigation shows that Max-FRs introduced by 'why' are possible in Mesoamerican languages. Therefore, an account of their more restricted distribution across languages cannot be grounded on absolute principles, but needs to allow for language-specific variation.

Max-FRs introduced by complex *wh*-phrases without a free-choice morpheme are rare in the languages that had been surveyed before our investigation, though not unheard of. This has led to incorrect conclusions and theories.[47] The Mesoamerican languages that we have studied here provide strong evidence toward more reliable generalizations. Half of them allow for Max-FRs with 'what'/'which' + NP, and a couple more also allow for Max-FRs that are introduced by 'how much/many' + NP.

Two of the languages, which are from different language families and whose columns have been shaded in Table 1.2, exhibit another less common pattern—less common at least among the languages whose Max-FRs have been studied so far. Usually, the *wh*-expressions in Max-FRs are morpho-syntactically the same as those in interrogative clauses. Still, there are languages like Modern Greek and Hungarian in which the *wh*-words in Max-FRs are built out of the *wh*-roots from interrogative clauses with an extra morpheme resembling the definite determiner. Bare *wh*-words are not allowed in Max-FRs in those languages. San Pedro Mixtepec Zapotec obligatorily enriches its *wh*-words with the prefix *tel-* of uncertain origin (Chapter 7: §4.1.1). The very same enriched forms are used as the obligatory stems for *wh*-words in FC-FRs, while *wh*-words in Ex-FRs occur bare. This is the same pattern observed in Modern Greek. In K'iche', on the other hand, the *wh*-expressions in Max-FRs obligatorily require what looks like a determiner at their right edge. The morpho-syntactic status of this element is not fully clear, although it has to occur after complex *wh*-expressions with a postposition, rather than right after the *wh*-word (Chapter 8: ex. 44). The *wh*-words in FC-FRs and Ex-FRs in K'iche' never combine with determiners. Lastly, Ch'ol exhibits an intriguing difference between *wh*-expressions in matrix vs. embedded *wh*- interrogative clauses: the matrix ones carry an extra suffix, which is dispreferred or impossible on the embedded ones. It may be a marker of clefting, but the issue needs further investigation. *Wh*-expressions in Max-FRs (and any other FRs) make use of the simpler *wh*-forms occurring in embedded clauses (Chapter 11: §2).

[47] See Caponigro (2019) for data and arguments against proposals that are built on the assumption that complex *wh*-phrases can never introduce FRs across languages.

### 8.3. Findings on existential free relative clauses

Ex-FRs are attested in all fifteen languages with different degrees of productivity, as shown in Table 1.3.

**Table 1.3. Use of *wh*-expressions in existential free relative clauses across the languages investigated in this volume**

| | UTO-AZT | | OTO-MANGUEAN | | | | MAYAN | | | | | | MI-ZO | CHI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Te** 2 | **Na** 3 | **AO** 4 | **Ma** 5 | **IM** 6 | **SZ** 7 | **K'** 8 | **Q'** 9 | **Cj** 10 | **Cl** 11 | **TT** 12 | **YM** 13 | **SP** 14 | **Pe** 15 |
| 'who' | √ | √ | √ | √ | % | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| 'what' | √ | √ | √ | √ | * | √ | % | √ | √ | √ | √ | √ | √ | √ |
| 'where' | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| 'when' | * | √ | * | * | * | -- | √ | √ | * | √ | √ | -- | √ | √ |
| 'why' | * | -- | * | ? | * | √ | √ | -- | % | √ | √ | √ | √ | * |
| 'how' | √ | √ | √ | ? | * | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| 'what/which' (+NP) | * | * | √ | * | √ | * | √ | √ | * | √ | √ | * | √ | √ |
| 'how much/ many' (+NP) | √ | * | √ | * | √ | * | % | √ | √ | √ | √ | ? | √ | √ |

NOTE: **Marks:** √: attested; *: not attested; ?: unclear if attested; --: no simple *wh*-word conveying the relevant meaning; %: speaker variation. **Shading:** languages in which Ex-FRs are necessarily marked as non-finite/modalized/irrealis. **Language family name abbreviations:** UTO-AZT: Uto-Aztecan; MI-ZO: Mixe-Zoquean; CHI: Chibchan. **Language name abbreviations:** Te: Southeastern Tepehuan; Na: Tlaxcala Nahuatl; AO: Acazulco Otomi; Ma: Matlatzinca; IM: Iliatenco Me'phaa; SZ: San Pedro Mixtepec Zapotec; K': K'iche'; Q': Q'anjob'al; Cj: Chuj; Cl: Ch'ol; TT: Tsotsil and Tseltal; YM: Yucatec Maya; SP: Sierra Popoluca; Pe: Pesh. **Numbers:** the number under the language name abbreviation indicates the chapter in the volume where the language is discussed.

All the languages allow for most of their *wh*-expressions to introduce Ex-FRs. Four Mayan languages (Ch'ol, Q'anjob'al, Tsotsil, and Tseltal) and Sierra Popoluca allow all their *wh*-expressions to introduce Ex-FRs, while Pesh allows this for all its *wh*-expressions but one. The *wh*-words for 'who' and 'where' can introduce Ex-FRs in all languages.

A common crosslinguistic pattern that has been observed regarding the TAM (tense/aspect/mood) properties of Ex-FRs previously is that they tend to be marked as non-finite, or subjunctive, or incompletive, or irrealis, depending on the morpho-syntactic resources of the language. Semantically, they trigger a modal interpretation. Grosu (2004, 2013) and Šimík (2011) have argued that this TAM and modal restriction is a constitutive property of Ex-FRs, as we

discussed in §3.1.2. Still, this is not the only attested pattern, and the fifteen languages we have investigated shed further light on this issue. The two Uto-Aztecan languages (Southeastern Tepehuan and Tlaxcala Nahuatl) and two of the Oto-Manguean languages (Acazulco Otomi and Matlazinca) exhibit the same restriction, requiring obligatory irrealis marking in their Ex-FRs. These are the first four languages from the left in Table 1.3, shaded to highlight this commonality. All the other languages, though, freely allow for Ex-FRs with finite/completive/realis/past TAM marking, and Ex-FRs with this TAM marking do not receive a modal interpretation. This is the same kind of morpho-syntactic and semantic behavior that is observed in Max-FRs in those languages. Since TAM and modal restrictions have been at the center of the debate about Ex-FRs, in the following we summarize our findings about the languages whose Ex-FRs do not exhibit any of the expected finiteness, mood, or aspectual restrictions:

- *Iliatenco Me'phaa* allows for Ex-FRs with realis and imperfective aspect marking (Chapter 6: ex. (45b)) or realis and perfective aspect marking (Chapter 6: ex. (45c)).
- *San Pedro Mixtepec Zapotec* easily allows for completive Ex-FRs (Chapter 7: §4.1.2, most examples).
- *K'iche'* allows for completive Ex-FRs (Chapter 8: exs. (48), (50), (54), (55)), as well as incompletive Ex-FRs.
- *Q'anjob'al* allows for completive Ex-FRs (Chapter 9: exs. (43c), (45b)), besides Ex-FR with incompletive or potential markers.
- *Chuj* allows for Ex-FRs marked as perfective (Chapter 10: exs. (68), (71), (80)), as well as ones marked as imperfective or showing no aspectual marking at all.[48]
- *Ch'ol* exhibits the same pattern as Chuj, allowing Ex-FRs with perfective marking (Chapter 11: exs. (85), (87), (90), (91)), or imperfective marking, or no aspectual marking at all.
- *Tsotsil* and *Tseltal* allow for Ex-FRs with completive marking (Chapter 12: exs. (61), (62), (70), (71), (72), (74)), as well as Ex-FRs with incompletive marking or no marking at all.
- *Yucatec Maya* behaves like the other Mayan languages: Ex-FRs can be marked as perfective. Although only one example is provided (Chapter 13: ex. (23c)), the issue is explicitly addressed and the conclusion is drawn that there are no morpho-syntactic differences between Max-FRs and Ex-FRs in the language (Chapter 13: §2.1.2).
- *Sierra Popoluca* allows for Ex-FRs with completive marking (Chapter 14: exs. (60), (61)), besides ones with incompletive marking.

[48] This observation about Ex-FRs in Chuj is already stated in Kotek & Erlewine (2016), although supported by a much smaller set of data.

- *Pesh* allows its Ex-FRs to exhibit the certainty marker (Chapter 15: exs. (45), (46), (57), (58), (59), (60)), besides the dubitative or uncertainty markers.

In conclusion, most of the languages we have investigated do not impose any TAM or semantic restrictions on their Ex-FRs. Ex-FRs and Max-FRs are morpho-syntactically identical in those languages, except for the morpho-syntactic enrichment of the *wh*-expressions in Max-FRs in San Pedro Mixtepec Zapotec and K'iche' that we discussed in §8.2. The four languages that do exhibit TAM restrictions in their Ex-FRs are not concentrated in just one language family. A restriction is attested in Ex-FRs in both the Uto-Aztecan languages, two of the four Oto-Manguean languages, but none of the seven Mayan languages, the Mixe-Zoquean language, or the Chibchan language. Whatever factor affects the presence or absence of TAM and semantic restrictions in Ex-FRs, it must be independent enough to be able to cut across language families.

Lastly, Tlaxcala Nahuatl, one of the languages with TAM and semantic restrictions on their Ex-FRs, allows an optional clausal subordinator to precede *wh*-expressions in Ex-FRs—an option that is available for Max-FRs and FC-FRs as well. This indicates that Ex-FRs are full clauses in Tlaxcala Nahuatl. Therefore, the presence of obligatory TAM and semantic restrictions cannot be taken as a sufficient evidence that Ex-FRs are non-clausal, contrary to what Grosu (2004, 2013) and Šimík (2011) suggest.

### 8.4. Findings on free-choice free relative clauses

FC-FRs are attested in twelve of the languages we have investigated, with various degrees of productivity, as shown in Table 1.4.

Mayan languages exhibit the full array of variation: Yucatec Maya and Q'anjob'al allow for all their *wh*-expressions to introduce FC-FRs, Ch'ol and Chuj allow for all their *wh*-expressions but one, K'iche' excludes three of eight *wh*-expressions, while Tsotsil and Tseltal do not have FC-FRs at all.[49] Among the Oto-Manguean languages, Iliatenco Me'phaa and San Pedro Mixtepec Zapotec allow for (almost all) their *wh*-expressions to introduce FC-FRs, while Acazulco Otomi and Matlatzinca are much more restrictive. Our two Uto-Aztecan languages show opposite patterns with respect to each other: Southeastern Tepehuan does not have FC-FRs at all, while Tlaxcala Nahuatl allows all its *wh*-expressions but two to introduce FC-FRs. Finally, FC-FRs are highly productive in both Sierra Popoluca and Pesh. Among the languages allowing for FC-FRs, the most productive *wh*-expressions introducing them are those roughly corresponding to ‘what’ and ‘where’, closely followed by ‘who’.

[49] Tsotsil and Tseltal form FC items out of *wh*-expressions; they just don't use them to introduce FC-FRs. Max-FRs can be used to convey a free-choice interpretation (see Chapter 12: §4.1.3).

**Table 1.4. Use of *wh*-expressions in free-choice free relative clauses across the languages investigated in this volume**

| | UTO-AZT | | OTO-MANGUEAN | | | | MAYAN | | | | | | M-Z | CHI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Te** | **Na** | **AO** | **Ma** | **IM** | **SZ** | **K'** | **Q'** | **Cj** | **Cl** | **TT** | **YM** | **SP** | **Pe** |
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| | | +F | +F | F+ | F+ | +F | F+ | +F | F+ | +F | | F+ | +F | FF |
| 'who' | * | √ | √ | * | √ | √ | √ | √ | √ | ? | * | √ | √ | √ |
| 'what' | * | √ | √ | √ | √ | √ | √ | √ | √ | ? | * | √ | √ | √ |
| 'where' | * | √ | √ | √ | √ | √ | √ | √ | √ | ? | * | √ | √ | √ |
| 'when' | * | √ | * | * | √ | -- | * | √ | √ | ? | * | -- | √ | √ |
| 'why' | * | -- | * | * | √ | √ | * | -- | * | * | * | √ | * | * |
| 'how' | * | √ | * | * | √ | √ | * | √ | √ | ? | * | √ | √ | n/a |
| 'which/what' (+NP) | * | * | * | * | √ | √ | √ | √ | √ | ? | * | √ | √ | √ |
| 'how much/ many' (+NP) | * | * | * | ? | √ | * | √ | √ | √ | ? | * | √ | √ | √ |

NOTE: **Marks:** √: attested; *: not attested; ?: unclear if attested; --: no simple *wh*-word conveying the relevant meaning; n/a: data not available; F+: free choice markers precedes wh-words; +F: free choice markers follow wh-words; FF: free choice marking results from reduplication of *wh*-expressions. **Language family name abbreviations:** UTO-AZT: Uto-Aztecan; M-Z: Mixe-Zoquean; CHI: Chibchan. **Language name abbreviations:** Te: Southeastern Tepehuan; Na: Tlaxcala Nahuatl; AO: Acazulco Otomi; Ma: Matlatzinca; IM: Iliatenco Me'phaa; SZ: San Pedro Mixtepec Zapotec; K': K'iche'; Q': Q'anjob'al; Cj: Chuj; Cl: Ch'ol; TT: Tsotsil and Tseltal; YM: Yucatec Maya; SP: Sierra Popoluca; Pe: Pesh. **Numbers:** The number under the language name abbreviation indicates the chapter in the volume where the language is discussed.

As mentioned in § 3.1.3, *wh*-expressions in FC-FRs are enriched by means of varying morpho-syntactic devices, depending on the language. Our languages exhibit three main patterns, also highlighted in Table 1.4: in six of them, FC marking follows the *wh*-word or the whole *wh*-phrase (F+); in five, FC marking precedes the *wh*-expression (+F); and in one of them, FC marking is realized by means of the full reduplication of the whole *wh*-expression (FF). Brief characterizations of FC marking in each language are provided in the following. The reader should consult the individual chapters for a more complete picture.

- *Tlaxcala Nahuatl.* The FC marker is a suffix on the *wh*-word, rather than the whole *wh*-phrase. A subordinator can optionally precede the *wh*-expression (just as in Max-FRs and Ex-FRs). The FC suffix looks like a combination of the exhaustivity operator 'only' with the uncertainty marker 'maybe' in Colonial Nahuatl. FC *wh*-expressions can also occur by themselves, without introducing FC-FRs.

- *Acazulco Otomi.* The FC marker follows the *wh*-expression and can be separated from it by optional pro-forms that can occur in other FRs and interrogative clauses as well. The FC marker is homophonous with the matrix conjunction indicating consequence (‘so’).
- *Matlazinca.* The FC marker is a prefix to the *wh*-word that doesn’t resemble any other element in the language.
- *Iliatenco Me'phaa.* The FC marker immediately precedes the *wh*-expression and is homophonous with the temporal marker meaning ‘until’. FC *wh*-expressions can occur by themselves, without introducing FC-FRs.
- *San Pedro Mixtepec Zapotec.* The FC marker acts as a phrasal suffix combining with the whole *wh*-expression. It is homophonous with the temporal marker meaning ‘until’. FC *wh*-expressions can occur by themselves without introducing FC-FRs. *Wh*-words in FC-FRs are formed out of the morphologically enriched *wh*-words that are used in Max-FRs, which come with an extra prefix (§8.2).
- *K'iche'.* FC marking is realized by an extra prefix on the *wh*-word, together with another marker that looks like the exhaustivity operator (‘only’) and immediately precedes the *wh*-expression. FC *wh*-expressions can occur by themselves, without introducing FC-FRs.
- *Q'anjob'al.* FC marking is realized by means of a clitic suffix and is homophonous with a temporal element receiving a multiplicity of interpretations (‘always’, ‘suddenly’, ‘as soon as’).
- *Chuj.* FC marking is realized by a complex marker that precedes the *wh*-expression and looks like a combination of the possibility modal (‘can’) with the exhaustivity operator (‘only’). FC *wh*-expressions can occur by themselves, without introducing FC-FRs.
- *Ch'ol.* The FC marker is a second-position clitic to the right of the *wh*-expression and resembles the exhaustivity operator ‘only’. FC-FRs trigger an indifference inference but no ignorance inference—a pattern that has not yet been observed in FC-FRs crosslinguistically (§3.1.3). Further investigation is needed to fully assess whether this construction is a true FC-FR (Chapter 11: §4.1.3). This is why we marked the *wh*-expressions that can occur in this construction with “?” in Table 1.4 (and Table 1.5).
- *Yucatec Maya.* FC marking can be realized by either one of two markers preceding the *wh*-expression. The two markers may trigger slightly different FC inferences (partial variation vs. total variation). They both trigger ignorance inferences; it is unclear if they can trigger indifference inferences as well. FC *wh*-expressions can occur by themselves, without introducing FC-FRs.

- *Sierra Popoluca.* The FC marker is a clitic to the right of the *wh*-word, rather than of the whole *wh*-phrase. It is homophonous with the adverb 'already'. FC-FRs trigger both ignorance and indifference inferences. FC *wh*-expressions can only be used to introduce FC-FRs and can never occur by themselves.
- *Pesh.* FC marking is realized by means of reduplication: the whole *wh*-expression is fully reduplicated. Although this is the only language among those that we have investigated that exhibits this pattern, reduplication is an attested way to realize FC marking elsewhere. Latin is a known example (e.g., *quisquis* 'whoever' from *quis* 'who').

Overall, our twelve languages with FC-FRs show that the use of FC markers preceding or following *wh*-words or *wh*-expressions in FC-FRs can vary within the same language family. We see the full array of option on display: prefixes, suffixes, clitics, and independent words.
They also show that FC markers in different languages can be derived from a similar element without necessarily exhibiting the same morpho-syntactic behavior. For instance, the FC marker in Iliatenco Me'phaa and the one in San Pedro Mixtepec Zapotec are both homophonous with the temporal preposition/conjunction 'until'. Still, one behaves like a prefix, the other like a suffix. Of the Mayan languages, K'iche', Chuj, and Ch'ol all make use of the exhaustivity operator ('only') as the FC marker. In Ch'ol, it behaves as a second-position clitic to the right of the *wh*-expression, whereas in the other two languages it acts as a marker preceding the *wh*-expression.

Another crosslinguistic fact that is confirmed by some of our languages is that *wh*-expressions with FC marking can be used on their own, without introducing FC-FRs, a well-known fact in English and other better-studied languages (e.g., *I can eat <u>whatever (food)</u>* or *Go <u>wherever!</u>*).

Last, some of our languages bring further evidence in support of a more nuanced view of the inferential properties of FC-FRs. The assumption that FC-FRs may trigger both ignorance inferences and indifference inferences is due to the dual inferential behavior of English FC-FRs, which has been taken as the paradigm for FC-FRs in general. Recent crosslinguistic investigation has shown that languages may vary in this respect as well. For instance, the Italian and Romanian morpho-syntactic equivalents of FC-FRs in English don't trigger an indifference inference and impose different conditions on the ignorance inference (Caponigro & Fălăuş 2018). Giannakidou & Cheng (2006) argue that Greek FC-FRs trigger a different kind of ignorance inference that is crucially linked to a modal/intensional component. Our investigation barely touched on the subtle inferential properties of FC-FRs; much more detailed investigation is needed. Still, Ch'ol and Yucatec Maya provide preliminary evidence that only one inference may be triggered – and not necessarily the same inference in each language: Ch'ol FC-FRs only trigger an indifference inference, while Yucatec Maya FC-FRs trigger an ignorance inference but may not trigger an indifference inference.

### 8.5. The distribution of *wh*-expressions across constructions

Having reached the end of our survey, we would like to draw attention to the overall distribution of *wh*-expressions across constructions that our investigation has revealed. The distributional patterns we found provide further evidence that the three varieties of FRs are different constructions with respect to one another and with respect to [+H]RCs and *wh*- interrogative clauses.

The crucial assumption is that, all else being equal, a necessary (but not sufficient) condition for two *wh*- constructions to be the same morpho-syntactically is that they be introduced by the same *wh*-expressions, unless independently motivated restrictions intervene.

Table 1.5 summarizes the distributions of the main *wh*-expressions across all the languages we have studied and, for each language, across the three varieties of FRs and [+H]RCs. We do not indicate the distribution of *wh*-expressions in *wh*- interrogative clauses since, by definition, all *wh*-expressions occur in *wh*- interrogative clauses. The picture that emerges is complex but extremely telling—we believe—and worth some elucidation.

First, in none of our languages is the distribution of *wh*-expressions in all FRs the same as in [+H]RCs. This immediately excludes the possibility that all FRs can be reduced to [+H]RCs in specific languages, let alone in all languages. Even more strongly, in none of our languages does the distribution of *wh*-expressions in [+H]RCs match their distribution in any variety of FR. Pesh may look like the single exception: the only *wh*-word that can introduce a [+H]RC is also the only *wh*-word that can introduce a Max-FR. But the very limited inventory of *wh*-expressions in both constructions prevents us from drawing any strong conclusions.

On the other hand, there are clear cases in which most *wh*-expressions can occur in some if not all FRs, while none is allowed in [+H]RCs. Southeastern Tepehuan and San Pedro Mixtepec Zapotec ban *wh*-expressions in [+H]RCs, but allow many in Max-FRs and Ex-FRs. San Pedro Mixtepec Zapotec allows them in FC-FRs as well. Q'anjob'al allows only one *wh*-expression in [+H]RCs, while it allows for all its *wh*-expressions to occur in all FRs. For all the languages, the number of *wh*-expressions in FRs is the same or greater than the number in [+H]RCs. All these asymmetries provide further evidence for the conclusion that FRs cannot be reduced to [+H]RCs.

A similar line of argument supports the conclusion that the three varieties of FRs are conceptually different from one another. None of the languages allows for exactly the same *wh*-expressions in all varieties of FRs, except for Q'anjob'al. There are extreme cases in which no *wh*-expression is allowed in FC-FRs (i.e., the construction is not attested at all), while most *wh*-expressions can occur in Max-FRs and Ex-FRs, as in Southeastern Tepehuan, Tsotsil, and

Table 5. U**se of *wh*-expressions in [+]HRs, Max-FRs, Ex-FRs, and FC-FRs across the languages investigated in this volume**

| | UTO-AZTECAN | | | | | | | | OTO-MANGUEAN | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **SE Tepe** 2 | | | | **Tlax Nahua** 3 | | | | **Acazulco Oto** 4 | | | | **Matlatzinca** 5 | | | | **Iliatenco Me'** 6 | | | | **SPMZ** 7 | | | |
| | **H** | **M** | **E** | **F** | **H** | **M** | **E** | **F** | **H** | **M** | **E** | **F** | **H** | **M** | **E** | **F** | **H** | **M** | **E** | **F** | **H** | **M** | **E** | **F** |
| 'who' | * | √ | √ | * | √ | √ | √ | √ | √ | √ | √ | √ | * | √ | √ | * | % | % | % | √ | * | √ | √ | √ |
| 'what' | * | √ | √ | * | √ | √ | √ | √ | * | √ | √ | √ | * | √ | √ | √ | * | * | * | √ | * | √ | √ | √ |
| 'where' | * | √ | √ | * | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | * | √ | √ | √ |
| 'when' | * | √ | * | * | * | * | √ | √ | * | * | * | * | * | * | * | * | * | √ | * | √ | -- | -- | -- | -- |
| 'why' | * | √ | * | * | -- | -- | -- | -- | * | * | * | * | * | √ | ? | * | * | √ | * | √ | * | √ | √ | √ |
| 'how' | * | √ | √ | * | * | √ | √ | √ | * | √ | √ | * | * | √ | ? | * | n/a | √ | * | √ | * | √ | √ | √ |
| 'which/what' (+NP) | * | * | * | * | * | √/* | * | * | * | * | √ | * | * | * | * | * | * | √ | √ | √ | * | √ | * | √ |
| 'how much/many' (+NP) | * | √ | √ | * | * | * | * | * | √ | * | √ | * | * | * | * | ? | * | √ | √ | √ | * | √ | * | * |

| | MAYAN | | | | | | | | | | | | | | | | | | | | | | | | MIXE-ZO | | | | CHIBCHAN | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **K'iche'** 8 | | | | **Q'anjob'al** 9 | | | | **Chuj** 10 | | | | **Ch'ol** 11 | | | | **Tseltalan** 12 | | | | **Yucatec Maya** 13 | | | | **Sierra Po** 14 | | | | **Pesh** 15 | | | |
| | **H** | **M** | **E** | **F** | **H** | **M** | **E** | **F** | **H** | **M** | **E** | **F** | **H** | **M** | **E** | **F** | **H** | **M** | **E** | **F** | **H** | **M** | **E** | **F** | **H** | **M** | **E** | **F** | **H** | **M** | **E** | **F** |
| 'who' | (√) | √ | √ | √ | * | √ | √ | √ | * | √ | √ | √ | * | √ | √ | ? | √ | √ | √ | * | (√) | √ | √ | √ | * | √ | √ | √ | * | * | √ | √ |
| 'what' | * | √ | % | √ | * | √ | √ | √ | * | √ | √ | √ | * | √ | √ | ? | * | √ | √ | * | (√) | √ | √ | √ | * | √ | √ | √ | * | * | √ | √ |
| 'where' | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ? | √ | √ | √ | * | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| 'when' | * | * | √ | * | * | √ | √ | √ | * | * | * | √ | √ | √ | √ | ? | √ | √ | √ | * | -- | -- | -- | -- | √ | √ | √ | √ | * | * | √ | √ |
| 'why' | * | * | √ | * | -- | -- | -- | -- | √ | % | % | * | * | * | √ | * | * | * | √ | * | √ | * | √ | √ | √ | * | √ | * | * | * | * | * |
| 'how' | * | √ | √ | * | * | √ | √ | √ | √ | √ | √ | √ | * | √ | √ | ? | * | √ | √ | * | √ | √ | √ | √ | √ | √ | √ | √ | * | * | √ | n/a |
| 'which/what' (+NP) | * | √ | √ | √ | * | √ | √ | √ | √ | √ | * | √ | √ | √ | √ | ? | * | (√) | √ | * | * | * | * | √ | * | √ | √ | √ | * | * | √ | √ |
| 'how much/many' (+NP) | * | √ | % | √ | * | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ? | √ | √ | √ | * | * | ? | ? | √ | √ | √ | √ | √ | * | * | √ | √ |

Note. **Marks:** √: attested; ; (√): attested with restrictions; *: not attested; ?: unclear if attested; --: no simple *wh*-word conveying the relevant meaning; ; %: speaker variation; ; n/a: data not available; **Relative clause name abbreviations**: H: headed relative clauses; M: Max-FRs; E: Ex-FRs; F: FC-FRs. **Language family name abbreviations:** MIXE-ZO: Mixe-Zoquean. **Language name abbreviations:** SE Tepe: Southeastern Tepehuan; Tlax Nahua: Tlaxcala Nahuatl; Acazulco Oto: Acazulco Otomi; Iliatenco Me': Iliatenco Me'phaa; SPMZ: San Pedro Mixtepec Zapotec; Tseltalan: Tsotsil and Tseltal; Sierra Po: Sierra Popoluca. **Numbers:** the number under the language name abbreviation indicates the chapter in the volume where the language is discussed.

Tseltal.[50] On the other hand, in Q'anjob'al, the three varieties of FRs all seem to be instantiations of the same basic construction. They all allow for all *wh*-expressions to occur and do not exhibit any differences beyond the expected ones: restricted distribution for Ex-FRs, extra morpho-syntax for FC-FRs, and different meanings.

Q'anjob'al is also the only language for which the evidence from *wh*-expressions supports the hypothesis that FRs are morpho-syntactically *wh*- interrogative clauses. For the other languages, this hypothesis would be supported for some varieties of FRs—those allowing for all *wh*-expressions: Max-FRs in San Pedro Mixtepec Zapotec (although with an extra prefix on each *wh*-expression), Ex-FRs in Ch'ol, Tsotsil, Tseltal, and Sierra Popoluca, FC-FRs in Iliatenco Me'phaa and Yucatec Maya (although with extra FC marking on each *wh*-expression).

In conclusion, the distribution of *wh*-expressions across the varieties of FRs, in [+H]RCs, and in *wh*- interrogative clauses strongly supports the default assumption behind our investigation: these five constructions should be considered related but different, unless there is strong evidence to the contrary.

### 8.6. Findings on other headless relative clauses

All the Mesoamerican languages we investigated show the other forms of [−H]RCs (besides FRs) that we introduced in §3.2. We have already commented on the productivity of LHRs and SFRs across these languages (cf. §8.1). In this section, we first highlight some common features characterizing these constructions across the languages studied and then provide language-specific details, some of which support and exemplify the general remarks.

With respect to LHRs, our languages exhibit all the crosslinguistic patterns we have discussed in §3.2.1. Their D head can be a deictic or a pronominal that never takes an NP complement, or it can be a plain determiner that requires an NP complement when not introducing an LHR, or, finally, it can be a demonstrative, a quantifier, or a numeral that can occur with or without an NP complement.

The nature of the D head determines the overall semantic behavior of LHRs in almost every language we looked at. If the D head is a definite determiner or a pronoun, then the whole LHR refers to a maximal individual; if D is a quantifier, then the LHR behaves like a quantificational DP, etc. Acazulco Otomi and San Pedro Mixtepec Zapotec are the only clear exceptions. Their LHRs are introduced by pronominal forms that seem to be compatible with more than one interpretation. Still, the default interpretation is the referential and maximal one (we speculate on why this is the case in our concluding remarks in §9). The interpretation as an indefinite has to be triggered by an existential predicate. In Acazulco Otomi, LHRs can trigger FC inferences as well, but only if the FC marker occurs. Notice that, alongside languages like Acazulco Otomi, there are languages like Yucatec Maya that ban FC markers from any LHR or SFR.

[50] See the qualification in fn. 49 for Tsotsil and Tseltal.

Five of the Mayan languages we have investigated (Q'anjob'al, Ch'ol, Tsotsil, Tseltal, and Yucatec Maya) seem to have two kinds of LHRs: one which is a genuine [−H]RC lacking a syntactic N head and one which might be better analyzed as a [+H]RC with a silent N head. The most common piece of evidence in favor of the silent N head analysis of some LHRs is a restriction on their interpretation. These LHRs build their meaning out of a set of individuals that has to be contextually provided. If they refer to a maximal individual, then it has to be the maximal individual of a set that has been made salient in the discourse. Similarly, if they behave like quantificational expressions, then they must quantify over a set that has been introduced in the previous discourse. An overt N denotes a set lexically. A silent N still denotes a set, but via the discourse, behaving as some sort of pronominal whose denotation has to be a set, rather than an individual. We label this semantic behavior "discourse anaphoric" and discuss it further in the language-specific remarks that follow.

As for SFRs, morpho-syntactically they appear in three shapes: either they are introduced by no marker at all, or by a complementizer/relative subordinator, or a relative marker (if the language has a special set of relative markers that are independent from *wh*-expressions). The latter pattern is less common, but it is the one found in San Pedro Mixtepec Zapotec.

Semantically, SFRs across our languages are interpreted as referential and maximal by default. If they are allowed to occur as the complement of existential predicates, then they are interpreted as indefinites. There are languages that only allow SFRs as the complement of existential predicates. All our Mayan languages behave this way, with the exception of Ch'ol. This consistent semantic behavior of the same construction across languages and language families cannot be by chance. We speculate further on this concurrence in our concluding remarks in §9.

Details about LHRs and SFRs in each language under study are provided next. As in previous sections, the reader should consult the language-specific chapters for a more complete picture.

- *Southeastern Tepehuan.* LHRs can only be introduced by demonstrative pronouns as their D heads (these demonstratives can never take an NP complement). *Wh*-words can optionally occur as well, while they are never allowed to introduce [+H]RCs. There is no evidence for SFRs in the language.

- *Tlaxcala Nahuatl.* The language allows for three kinds of LHRs. One kind is composed of LHRs that are introduced by a demonstrative or quantificational D followed by the relative subordinator that also occurs in [+H]RCs and other [−H]RCs. These LHRs are the only kind to be obligatorily interpreted as discourse anaphoric. This may suggest that they are [+H]RCs with a silent N complement of their D head. A second kind of LHR is introduced by the same D head as the previous one, but a *wh*-word replaces the relative subordinator. The third kind of LHRs is introduced by the definite determiner followed by the *wh*-word

for ‘who’ or ‘what’, or by the relative subordinator. SFRs are introduced by the relative subordinator and can only be interpreted as referential and maximal.

- *Acazulco Otomi.* LHRs can be introduced by one of the pronominal forms as their D head. The language also has determiners taking NP complements, but they cannot introduce LHRs. One kind of LHR has its D head immediately followed by a *wh*-expression, the other has no other marking. Both kinds allow for all three semantic patterns that are attested in FRs. By default, LHRs refer to a maximal individual, like Max-FRs. They behave like indefinites when they occur as complements of existential predicates, like Ex-FRs. Finally, like FC-FRs, they trigger FC inferences when the FC marker occurs right after the D head. There does not seem to be evidence for SFRs.
- *Matlazinca.* LHRs are only introduced by a distal demonstrative pronoun, do not contain any *wh*-word, and only occur as the subject or the object of their matrix clause. SFRs can be introduced by the relative subordinator or nothing at all. They are always interpreted as referential and maximal. The SFR without a subordinator can occur as the complement of an existential predicate and be interpreted existentially, as in Ex-FRs.
- *Iliatenco Me'phaa.* LHRs can only be introduced by the indefinite marker as their D head, followed by the relative subordinator. The language has no definite marker, while demonstratives are not allowed as the D heads of LHRs. No *wh*-word is attested in LHRs. SFRs are introduced by the relative subordinator. They are interpreted as referential and maximal by default, and as existential when occurring in the complement position of existential predicates.
- *San Pedro Mixtepec Zapotec.* LHRs are introduced by pronominal D heads followed by the relative subordinator. The subordinator is required, unlike in [+H]RCs. LHRs are interpreted as referential and maximal by default, unless they occur as the complement of an existential predicate. SFRs, by contrast, are introduced by the same set of non-*wh* relative pronouns that introduce [+H]RCs. Like LHRs, they are interpreted as referential and maximal, unless in the complement position of an existential predicate.
- *K'iche'.* LHRs can be introduced by the distal, medial, or proximal D heads, but not by the indefinite D head. LHRs have no subordinator or *wh*-word. They are interpreted as referential and maximal only. SFRs have no subordinator (nor any other marking) and can only occur as the complement of an existential predicate.
- *Q'anjob'al.* There are two main kinds of LHRs. One kind is introduced by pronominal D heads and has no subordinator or *wh*-word (except for locative LHRs). Their morpho-syntax resembles that of [+H]RCs. The second kind of LHR is introduced by an indefinite determiner or a demonstrative as its D head, optionally followed by a *wh*-word. SFRs have no subordinator (nor any other special marking) and can only occur as the

complement of a plain existential predicate, unlike Ex-FRs, which occur with a broader set of predicates. Both kinds of LHRs and SFRs in the language are obligatorily interpreted as "discourse anaphoric"—their interpretation must rely on a domain of individuals that has already been introduced in the discourse.

- *Chuj*. LHRs allow for a noun classifier, a demonstrative, a plural marker, or a quantifier as their D head, optionally followed by a *wh*-word but without any subordinator. SFRs have no subordinator either and can only occur as the complement of a plain existential predicate, unlike Ex-FRs, which occur with a broader set of predicates.
- *Ch'ol.* All LHRs are introduced by an element from the same set of determiners and demonstratives, but exhibit two different patterns. The LHRs whose D head is followed by the relative subordinator are obligatorily interpreted as discourse anaphoric. This may suggest that they could be [+H]RCs with a silent N head. LHRs with a *wh*-word, by contrast, do not exhibit this restriction and fully qualify as [−H]RCs. SFRs are marked by the relative complementizer, and their interpretation is obligatorily discourse anaphoric. This constraint may suggest that SFRs too are [+H]RCs with a silent N (a silent D may be present or not, since the language allows for bare nominals without a D).[51] They are the only SFRs in the seven Mayan languages that we have investigated that are not restricted to the complement position of existential predicates.
- *Tsotsil and Tseltal.* Both languages have two kinds of LHRs. One kind takes a definite article or a demonstrative as its D head, followed by no other marker. These LHRs are obligatorily interpreted as discourse anaphoric and may be [+H]RCs with a silent N head. The other kind of LHR is only introduced by the definite article, followed by any *wh*-word. They are not obligatorily discourse anaphoric and evidence shows they are not [+H]RCs. SFRs exhibit no subordinator or relativizer, mainly occur as the complement of existential predicates, and enforce a discourse-anaphoric interpretation.
- *Yucatec Maya.* LHRs come in two flavors, resembling a pattern that we have already encountered in other Mayan languages. Both are introduced by a determiner, a numeral, or a quantifier as D head. The way the two kinds of LHRs differ is in what comes after the D head and how they are interpreted. One kind has its D head followed by no other marker—no *wh*-word nor subordinator/relativizer. These LHRs are necessarily interpreted as discourse anaphoric. This indicates, as with other Mayan languages, that these clauses could be analyzed as [+H]RCs with a silent N head. The second kind of LHR has its D head immediately followed by a *wh*-word and has no interpretative restriction. SFRs have

[51] See AnderBois et al. (2019) for a different take, according to which these constructions are true SFRs, at least in Ch'ol and Yucatec Maya.

no subordinator, only occur as the complement of an existential predicates, and enforce a discourse-anaphoric interpretation as well. None of these [−H]RCs allows for FC marking.

- *Sierra Popoluca.* LHRs come in two varieties. Both are introduced by demonstrative pronouns, are interpreted as referential and maximal, and can be clearly distinguished from [+H]RCs. The two kinds of LHRs differ in what follows their D head, which can be either one of the relativizers or one of the *wh*-words. SFRs are introduced by one of the relativizers. They are interpreted as referential and maximal by default, unless in the complement position of an existential predicate.
- *Pesh.* LHRs can take demonstratives, quantifiers, or numerals as their D heads. Their interpretation depends on the D head. SFRs are very common. They look like LHRs without a D head and are always interpreted as referential and maximal.

## 9. Some general remarks and conclusions

The crosslinguistic investigation in this volume reveals a world of [−H]RCs that is rich and productive, although it has been largely ignored so far. An extensive typological database like Dryer & Haspelmath (2013) mentions [−H]RCs only to exclude them from their inventory of constructions.[52] The study of contact–induced grammaticalization in Heine & Kuteva (2006: 204–205) starts from Haspelmath's (1998: 279–280) observation that relative clauses that are introduced by *wh*-expressions are common in "Standard Average European" (SAE) and goes on to claim that they are much less common in other languages families and that, whenever they are attested outside SAE, this is due to language contact with SAE. Our findings clearly show that [−H]RCs that are introduced by *wh*-expressions are common, varied, and productive in Mesoamerican languages, to a degree that is often much higher than what is found in Spanish or English or any other language within the SAE group.

Our findings from the fifteen languages we have studied provide strong reasons to believe that there is more to discover in other Mesoamerican and Chibchan languages. We hope that our work will correct some misunderstandings by broadening the empirical landscape and will inspire further research on [−H]RCs within and beyond Mesoamerican and Chibchan, by providing a framework, definitions, a methodology, and ample evidence and examples.

Before letting each language speak for itself in the following chapters, we would like to conclude with some more general remarks about the significance and the role of [−H]RCs within the study of human language, including the logic that underlies it.

[52] "Headless relative clauses (like English *what I bought at the store*) are not relevant to this map" Dryer (2013a: §1). Later in the same section, ex. (9a) provides an example of [−]HRC from Awa Pit. Dryer (2013b: ex. (34d)) contains an example of a [−]HRC from Nadëb. I have not been able to find [−]HRCs mentioned anywhere else in the database.

[−H]RCs are full clauses, exhibiting the rich morpho-syntax and the semantic complexity that clauses do. Still, they behave like nominals, both distributionally and semantically. They do so according to very precise constraints —a picture that our findings fully support. If a [−H]RC is introduced by a D or is morpho-syntactically enriched with an FC morpheme, then its specific kind of DP-like semantic behavior is determined by its D (definite, indefinite, demonstrative, quantifier, etc.) or by its FC marker (different kinds of free choice). If neither element occurs, then [−H]RCs behave like definite DPs by default. If they can occur as the complement of an existential predicate, then they obligatorily behave like existentially quantified DPs.

This semantic behavior is far from trivial. In principle, we could expect a DP-like expression like a [−H]RCs without an overt marker of definiteness or quantification to exhibit all the possible semantic behaviors that DPs exhibit, as long as they are consistent with the rest of the sentence and the discourse. For instance, the bracketed Max-FR in (63)a should be able to be interpreted as any of the quantificational DPs in (63)b, given the right context. But this is never the case, in language after language.

(63) a. [What is on the table] is expensive.
b. [Some/several/many/most/all of things on the table] are expensive.

Alternatively, we might expect unmarked [−H]RCs to behave like unmarked DPs (or NPs), also known as *bare nominals*— nominals with no overt D or FC marker. But this is not the case either. Bare nominals refer to individuals (including kinds) or convey existential quantification, depending on the aspectual/modal property of the sentence in which they occur.[53] For instance, the English bare nominal *bagels* is interpreted as referring to the maximal plural individual resulting from the sum of all the bagels in the generic modal sentence in (64)a, while it is interpreted as existentially quantified in the episodic sentence in the progressive form in (64)b.[54] The latter behavior is roughly equivalent to that of the DP with the overt existential quantifier *some* in (64)c. Crucially, the bracketed [−H]RC in (64)d occurs in the same position in the same episodic sentence as the bare nominal in (64)b and the quantified DP in (64)c, but cannot receive the same quantificational interpretation—it cannot be interpreted as equivalent to *something that will be served for lunch* or *some of the things that will be served for lunch*.

(64) a. [Bagels] have to be boiled before being baked.
b. [Bagels] are boiling right now.
c. [Some bagels] are boiling right now.
d. [What will be served for lunch] is boiling right now.

[53] See Le Bruyn et al. (2017) for a recent overview.
[54] But see Dayal (2013) for a recent proposal to handle this behavior of bare nominals without assuming that they trigger existential quantification.

The fact that this semantic restriction is observed in language after language across language families shows that it must be deeply rooted in some basic mechanism of human language. In particular, it must be a restriction governing the mapping between morpho-syntax and the logic behind human language—the logic that makes interpretation and meaning possible. This restriction could be something like a ban against quantification without an overt marker of quantificational force. If the basic meaning of a [−H]RC with no D or FC marker is a set of individuals (e.g., the set of things on the table in (64)d), then quantifying over that set is a semantically contentful operation that affects the core informational content—the truth conditions of the sentence. The situation in which only some of the things on the table are expensive and the situation in which all of them are expensive are crucially different. On the other hand, going from the set of things on the table to the corresponding maximal individual does not change the informational content, but only packages it differently. Intuitively, we are still considering a situation with the same individuals: instead of looking at them as a collection (a set), we conceptualize them as one plural individual.

Semantic speculations aside, the general point we want to make here is that [−H]RCs are important characters on the human language stage and should receive the attention and space they deserve. By studying them, we can learn more about the languages and language families in which they occur and the general principles of human language that govern them. We started making the case for [−H]RCs in this chapter by introducing them in general terms, painting their features with broad strokes. The next chapters present our main characters in their full, rich, and complex language-specific life. It's time to leave the stage to our languages and their scholars.

**Acknowledgments**

Special thanks to Judith Aissen and Christine Bartels for extensive, detailed, and enlightening comments. Thanks to Scott AnderBois, Claudine Chamoreau, Jessica Coon, September Cowley, Philip T. Duncan, Lucero Flores-Nájera, Néstor Hernández-Green, Wendy López Márquez, Enrique L. Palancar, Gilles Polian, and Roberto Zavala Maldonado for precious feedback. The author is solely responsible for all remaining mistakes.

## References


Abney, S. (1987). *The English noun phrase in its sentential aspect*. Ph.D. thesis, MIT, Cambridge, MA.

AnderBois, S., Chan Dzul, M. O., Coon, J., & Vásquez Álvarez, J. J. (2019). Relativas libres en ch'ol y maya yucateco y la tipología de cláusulas relativas sin núucleo. In *Proceedings of Form and Analysis in Mayan Linguistics (FAMLi)*. https://escholarship.org/uc/item/4d837983 5

Andrews, A. (2007). Relative clauses. In T. Shopen (Ed.) *Language Typology and Syntactic Description-Complex Constructions*, 2nd ed., pp. 132–223. New York: Cambridge University Press.

Aschmann, E. (1984). The relative clause in Highland Totonac. *SIL Mexico Workpapers*, 1–27.

Beck, D. (2016). Relative clauses in Upper Necaxa Totonac: Local, comparative, and diachronic perspectives. *Linguistic Discovery*, *14*(1), 1–45.

Beller, R. & Beller, P. (1979). Huasteca Nahuatl. In R. Langacker (Ed.), *Studies in Uto-Aztecan grammar: Volume 2, Modern Aztec grammatical sketches*, pp. 199–306. Arlington, TX: The Summer Institute of Linguistics.

Boudreault, L. (2009). *A grammar of Sierra Popoluca (Soteapanec, a Mixe-Zoquean language)*. Ph.D. thesis, University of Texas, Austin.

Brockway, E. (1979). North Puebla Nahuatl. In R. Langacker (Ed.), *Studies in Uto-Aztecan grammar: Volume 2, Modern Aztec grammatical sketches*, pp. 141–198. Arlington, TX: Summer Institute of Linguistics.

Campbell, L. (1997). *American Indian languages. The historical linguistics of Native America*. New York: Oxford University Press.

Campbell, L. (2014). Mesoamerican Indian languages. *Encyclopedia Britannica*. https://www.britnnica.com/topic/Mesoamerican-Indian-languages (accessed on January 25, 2020)

Campbell, L, Kaufman, T. & Smith-Stark, T. T. (1986). Meso-America as a linguistic area. *Language, 62*(3), 530–570.

Caponigro, I. (2003). *Free not to ask: On the semantics of free relatives and wh-words cross-linguistically*. Ph.D. thesis, University of California Los Angeles.

Caponigro I. (2004). The semantic contribution of wh-words and type shifts: evidence from free relatives cross-linguistically. In R. B. Young (Ed.), *Proceedings of Semantics and Linguistic Theory (SALT) XIV*, pp. 38–55. Ithaca, NY: CLC Publications, Cornell University.

Caponigro, I. (2019). In defense of what(ever) free relative clauses they dismiss: A reply to Donati and Cecchetto (2011). *Linguistic Inquiry, 50*(2), 356–371.

Caponigro, I., & Fălăuş, A. (2018). Free choice free relative clauses in Italian and Romanian. *Natural Language & Linguistic Theory, 36*(2), 323–363.

Caponigro, I. & Fălăuş, A. (2020). Unveiling multiple *wh*- free relative clauses and their functional wh-words. *Semantics & Pragmatics, 13*.

Caponigro, I. & Polinsky, M. (2011). Relative embeddings: A Circassian puzzle for the syntax/semantics interface. *Natural Language & Linguistic Theory, 29*(1), 71–122.

Caponigro, I., Torrence, H., & Cisneros. C. (2013). Free relative clauses in two Mixtec languages. *International Journal of American Linguistics, 79*(1), 41–96.

Carrasco Zúñiga, A. (2006). *Los procesos morfofonológicos de la lengua me̲’ph̲a̲a̲*. M.A. thesis, CIESAS, Mexico City.

Castro, S. (2013). Cláusulas relativas en Purépecha de Carapan, Michoacán. *Cuadernos Lingüística de El Colegio de México, 1*, 26–55.

Chierchia, G. & McConnell-Ginet, S. (2000). *Meaning and grammar*. 2nd ed. Cambridge, MA: MIT Press.

Citko, B. (2004). On headed, headless, and light-headed relative clauses. *Natural Language & Linguistic Theory*, *22*(1), 95–126.

Collins, J. (2015). Ilokano free relatives with *man*. In A. Camp, Y. Otsuka, C. Stabile & N. Tanaka (Eds.), *AFLA 21: The Proceedings of the 21st Meeting of the Austronesian Formal Linguistics Association*, Asia-Pacific Linguistics, pp. 33–50. http://hdl.handle.net/1885/95329

Comrie, B. & Estrada-Fernández, Z. (Eds.). (2012). *Relative clauses in languages of the Americas.* Amsterdam; Philadelphia: John Benjamins.

Dayal, V. (1996). *Locality in wh quantification: Questions and relative clauses in Hindi.* Dordrecht: Kluwer.

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

Dayal, V. (2013). On the existential force of bare plurals across languages. In I. Caponigro and C. Cecchetto (Eds.), *From grammar to meaning: The spontaneous logicality of language*, pp. 49–80. Cambridge, UK: Cambridge University Press.

Dayal, V. (2016). *Questions*. Oxford: Oxford University Press.

Dryer, M. S. (2013a). Order of relative clause and noun. In M.S. Dryer & M. Haspelmath (Eds.) *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/90, accessed on September 19, 2019.)

Dryer, M. S. (2013b). Position of negative morpheme with respect to subject, object, and verb. In M. S. Dryer & M. Haspelmath (Eds.) *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/144, accessed on September 25, 2019.)

Dryer, M. S. & Haspelmath, M. (Eds.) (2013). *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at

http://wals.info, accessed on September 14, 2019.)

Epps, P. (2012). Between headed and headless relative clauses. In Z. Estrada and B. Comrie (Eds.), *Typological and empirical studies of relative clauses*, pp. 191–211. Amsterdam: John Benjamins.

Farkas, D. (2018). Nominal reference. *Oxford Research Encyclopedias – Linguistics*. doi: 10.1093/acrefore/9780199384655.013.328

von Fintel, K. (2000). Whatever. In B. Jackson and T. Matthews (Eds.), *Proceedings from Semantics and Linguistic Theory* (*SALT) X*, 27–39. Ithaca, NY: CLC Publications, Cornell University.

Flores-Nájera, L. (2019). *La gramática de la cláusula simple en el náhuatl de Tlaxcala*. Ph.D. thesis, CIESAS, Mexico.

García Matzar, P. & Rodríguez Guaján, J. O. (1997). *Rukemik ri Kaqchikel Chi'*: *Gramática Kaqchikel*. Guatemala City: Editorial CHOLSAMAJ.

Giannakidou, A. & Lai-Shen Cheng, L. (2006). (In)definiteness, polarity, and the role of wh-morphology in Free Choice. *Journal of Semantics, 23*, 135–183.

Grosu, A. (2004). The syntax-semantics of modal existential wh-constructions. In O. Mišeska Tomić (Ed.), *Balkan syntax and semantics*, pp. 405–438. Amsterdam: John Benjamins.

Grosu, A. (2013). Relative clause constructions and unbounded dependencies. In C. Dobrovie-Sorin and I. Giurgea (Eds.), *A reference grammar of Romanian*, pp. 597– 662. Amsterdam: John Benjamins.

Guarcax González, J. C. (2016). *Las cláusulas relativas en el kaqchikel de Solola*. M.A. thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social, Mexico City.

Gutiérrez Bravo, R. (2012). Relative clauses in Yucatec Maya: Light heads vs. null domain. In B. Comrie & Z. Estrada-Fernández (Eds.), *Relative clauses in languages of the Americas*, pp. 253–268. Amsterdam/Philadelphia: John Benjamins.

Gutiérrez Bravo, R. (2013). Free relative clauses in Yucatec Maya. *Language Typology and Universals (Sprachtypologie und Universalienforschung/STUF), 66*, 22-39.

Gutiérrez Bravo, R. (2015). *Las cláusulas relativas en maya yucateco*. Mexico: El Colegio de México.

Haspelmath, M. (1998). How young is Standard Average European? *Language Sciences*, *20*(3): 271–87.

Heim, I. 2011. Definiteness and indefiniteness. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.), *Semantics*, pp. 996–1025, Berlin/Boston: de Gruyter.

Heine, B. & Kuteva, T. (2006). *The Changing Languages of Europe*. Oxford: Oxford University Press.

Hill, J. H. & Hill, K.C. (1981). Variation in relative clause construction in modern Nahuatl. In F. E. Karttunen (Ed.), *Nahuatl studies in memory of Fernando Horcasitas*, pp. 89–104. Austin: Texas Linguistic Forum 18.

Izvorski, R. (1998). Non-indicative wh-complements of possessive and existential predicates. In P. N. Tamanji & K. Kusumoto (Eds.), *NELS 28: Proceedings of the 28th Annual Meeting of the North East Linguistic Society*, pp. 159–173.

Jacobson, P. (1995). On the quantificational force of English free relatives. In E. Bach, E. Jelinek, A. Kratzer, and B. H. Partee (Eds), *Quantification in natural languages*, pp. 451–486. Dordrecht: Springer.

Jiménez Jiménez, S. (2019). Estudios del la gramática de la oración simple y compleja en zoque de San Miguel Chimalapa. Ph.D. thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social, Mexico City.

Johnson, G. (2015). The morphosyntax of whatever in free relatives: Variation and optionality in Appalachian English. Handout from presentation at 2015 LSA Annual Meeting.

Kotek, H. & Yoshitaka Erlewine, M. (2016). Unifying definite and indefinite free relatives: Evidence from Mayan. In C. Hammerly & B. Prickett (Eds), *Proceedings North East Linguistic Society (NELS) 46*, Vol. 2, pp. 241–254, Amherst, MA: GSLA.

Langacker, R. (ed.). (1979). *Studies in Uto-Aztecan grammar. Volume 2: Modern Aztec grammatical sketches*. Arlington, TX: Summer Institute of Linguistics.

Le Bruyn, B., de Swart, H., & Zwarts, J. (2017). Bare nominals. *Oxford Research Encyclopedia of Linguistics*. https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-399.

Link, G. (1983). The logical analysis of plurals and mass terms: A lattice- theoretical approach. In R. Bauerle, C. Schwarze & A. von Stechow (Eds.), *Meaning, use and the interpretation of language*, pp. 303–323. Berlin: Walter de Gruyter.

Lipták, A. (Ed.). (2009). *Correlatives Cross-linguistically*. Amsterdam: John Benjamins.

Mantenuto, I. & Caponigro, I. (2020). Free relative clauses in Teramano. Manuscript, University of California Los Angeles and University of California San Diego.

McCloskey, J. (2002). Resumption, Successive Cyclicity, and the Locality of Operations. In S. Epstein & D. Seeley (Eds.), *Derivation and explanation*, pp. 184–226, Oxford: Blackwell.

Montague, R. (1970). English as a formal language. In B. Visentini et al., *Linguaggi nella società e nella tecnica*, pp. 189–223. Milan: Edizioni di Comunità. [Reprinted in R. Thomason (Ed.), *Formal Philosophy*, pp. 188–221. New Haven, CT, and London: Yale University Press. 1974.]

Moore, D. (2016). Relative clauses in Coahuitlán Totonac. *Oklahoma Working Papers in Indigenous Languages*, 2, 22–40.

Palancar, E., Zavala Maldonado, R. & Chamoreau, C. (Eds). (Forthcoming). *Relative clauses in Mesoamerica and beyond*. Amsterdam: Brill.

Patterson, G. & Caponigro, I. (2016). The puzzling degraded status of who free relative clauses in English. *English Language and Linguistics 20*(2), 341–352.

Plann, S. (1980). *Relative clauses in Spanish without overt antecedents and related constructions*. Berkeley; Los Angeles: University of California Press.

Polian, G. (2013). *Gramática del tseltal de Oxchuc*. Mexico City: Casa Chata.

Quine, W. V. O. (1960). *World and object.* Cambridge, MA: MIT Press.

Ramos, P. (2015). *La fonología y morfología del zapoteco de San Pedro Mixtepec*. Ph.D. thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social, Mexico City.

van Riemsdijk, H. (2017). Free relatives. In Martin Everaert & Henk van Riemsdijk (Eds.), *The Wiley Blackwell companion to syntax*, 2nd ed., Ch. 116. Wiley Blackwell. https://doi.org/10.1002/9781118358733.wbsyncom116.

Sadler, L. G. & Camilleri, M. (2018). Free relatives in Maltese. *Brill's Journal of Afroasiatic Languages and Linguistics,* 10, 115–159.

Salzmann, M. (2018). Revisiting the NP vs. DP debate. Ms., University of Leipzig. Available at https://ling.auf.net/lingbuzz/003823.

Sharvy, R. (1980). A more general theory of definite descriptions. *The Philosophical Review*, *89*(4), 607–624.

Šimík, R. (2011). *Modal Existential wh-Constructions*. Ph.D. thesis, University of Groningen, Holland; published by LOT – Netherland Graduate School of Linguistics.

Šimík, R. (2017). Existential Wh-Constructions. In M. Aronoff, *Oxford bibliographies in linguistics.* doi: 10.1093/obo/9780199772810-0162

Šimík, R. (Forthcoming). Free relatives. In D. Gutzmann, L. Matthewson, C. Meier, H. Rullmann, & T. Ede Zimmermann (Eds.), *The Semantics Companion Wiley–Blackwell.*

Sischo, W. (1979). Michoacán Nahual. In R. Langacker (Ed.), *Studies in Uto-Aztecan grammar: Volume 2, Modern Aztec grammatical sketches*, pp. 1–141. Arlington, TX: Summer Institute of Linguistics.

Szabolcsi, A. (1986). From the definiteness effect to lexical integrity. In W. Abraham & S. de Meij (Eds.), *Topic, focus, and configurationality*, pp. 321–348. Amsterdam: John Benjamins.

Szabolcsi, A. (2010). *Quantification.* Cambridge, UK: Cambridge University Press.

Torrence, H., Mateo Pedro, P. & Duncan, P. (2016). *Non-interrogative wh-expressions in Kaqchikel*. Ms., University of California Los Angeles and University of Kansas at Lawrence.

Tuggy, D. H. (1979). Tetelcingo Nahuatl. In R. Langacker (Ed.), *Studies in Uto-Aztecan grammar: Volume 2, Modern Aztec grammatical sketches*, pp. 1–141. Arlington, TX: Summer Institute of Linguistics.

Voigtlander, C. & Echegoyen, A. (1985). *Luces Contemporáneas del Otomí, gramática del otomí de la sierra*. 2nd ed. Serie Gramáticas de Lenguas Indígenas de México, núm. 1. Mexico City: Instituto Lingüístico de Verano. [1st ed. 1979, without author names]

de Vries, M. (2018). Relative Clauses in Syntax. In *Oxford Research Encyclopedia of Linguistics*. doi: https://dx.doi.org/10.1093/acrefore/9780199384655.013.56

Westley, D. (1991). *Tepetotutla Chinantec syntax*. Arlington, TX: Summer Institute of Linguistics.

Willett, T. (1991). *A reference grammar of Southeastern Tepehuan*. Arlington, TX: Summer Institute of Linguistics.

Wolter, L. (2009). Demonstrative in philosophy and linguistics. *Philosophy Compass*, *4*(3), 451–468.