Syntactic and Semantic Aspects of Supplementary Relative Clauses

in English and Sōrānī Kurdish

Inauguraldissertation

zur Erlangung des Grades eines Doktors der Philosophie

im Fachbereich Neuere Philologien

der Johann Wolfgang Goethe-Universität

zu Frankfurt am Main

vorgelegt von

Shene Hassan

aus: Sulaimaniyah, Irak

2020

Acknowledgements

It is a pleasure to acknowledge the people who have supported me, encouraged me and provided me with the help I needed to write this thesis.

I want to extend my gratitude to my supervisor Manfred Sailer for guiding me and providing me with every possible support I needed to make this thesis come to being. I would also like to express my most sincere appreciation to my second supervisor Gert Webelhuth for giving me his time and most elaborated comments on the main elements of the thesis.

I also want to thank Frank Richter and Janina Radó who have provided me with the most insightful suggestions and comments, without which the thesis would have been incomplete.

I want to express my heartful appreciation to my dear friend, Hiwa Asadpour for his constant help with the Kurdish chapter.

I am very grateful to Dianne Jones for her valuable remarks and suggestions. I also thank my colleague Sascha Bargmann, Christopher Goetze and Nehir Aygul who have giving me their time while they were the busiest.

I was also very privileged to have received very insightful comments and suggestions from Thomas Jügel, Doug Arnold, Saloumeh Gholami, Pollet Samvelian and Berthold Crysmann.

I am forever grateful to DAAD for giving me one of the best opportunities in my life to come to Germany and pursue my academic ambitions and to accomplish this major goal. Granting me this scholarship has changed my life forever.

Contents

Abbreviations

General abbreviations

COCA	Corpus of Contemporary American English, https://www.english-corpora.org/
DPL	Dynamic Predicate Logic
D_{rel}	discourse relation
DRT	Discourse Representation Theory
HPSG	Head-Driven Phrase Structure Grammar
IRC	integrated relative clause
LRS	Lexical Resource Semantics
NPI	negative polarity items
RC	relative clause
ROA	Radical Orphanage Approach
SRC	supplementary relative clause
UDC	unbounded dependency construction

Abbreviations used in glosses

1	first person
2	second person
3	third person
ACC	accusative
ADD	additive
CONTN	continuous form
COP	copula
DEF	definite

DER	derivational affix, for example for conversion relation
DEM	demonstrative
EZ	Ezafé
INDEF	indefinite
IPFV	imperfective aspect
KE	Kurdish relativizer *ke*
NEG	negative
NOM	nominative
PL	plural
POST	postposition
PRS	present tense
PST	past tense
PTC	particle
SG	singular

Abbreviation of HPSG features names and sort names

CFORM	COMPLEMENTIZER-FORM
EXCONT	EXTERNAL-CONTENT
EZ	EZAFE
INCONT	INTERNAL-CONTENT
LRS	LEXICAL-RESOURCE-SEMANTICS

Chapter 1

Introduction

In this thesis, I examine and analyse *supplementary relative clauses* (SRCs), also known as *non-restrictive relative clauses*. SRCs have received considerably less attention in the study of relative clauses than *integrated*, or *restrictive*, relative clauses (IRCs). The (surface) syntactic structure of the two types of relative clauses (RCs) is largely identical – see (1). Therefore, it is not straightforward to determine where to locate the difference in the interpretation between IRCs and SRCs.

(1) a. Kim finally found a dissertation topic which lies at the interface of biology and physics. (IRC)

b. Kim finally found a dissertation topic, which lies at the interface of biology and physics. (SRC)

To address this question, I focus on two types of English SRCs: determiner-*which* RCs, and SRCs introduced by *that*, see (2) and (3) respectively. Determiner-*which* RCs can only be interpreted as SRCs. *That* RCs are usually claimed to be restricted to IRCs, though *that* SRCs are mentioned in some descriptive grammars

such as Huddleston & Pullum (2002).

(2) Maybe people will attack me, in which case I will attack back. (COCA)

(3) The Patas monkey, that spends almost all of its time in open grassland, adopts just such tactics. (Huddleston & Pullum, 2002: 1052)

These two types of SRCs have received little attention in the description of English and they have been largely ignored in the formal linguistic literature. This in itself makes their study a suitable enterprise for a dissertation. In addition, I investigate how far an analysis of the two constructions allows me to address the question outlined above.

According to Huddleston & Pullum (2002), in IRCs the information expressed in an IRC is an integral part of the matrix clause, while SRCs only provide additional information about the antecedent. They conclude from this that, contrary to IRCs, SRCs are not integrated in the sentence, neither syntactically nor semantically. A counter argument to this has been put forward by Arnold (2004, 2007), who argues that IRCs and SRCs are not much different syntactically and only differ with respect to their type of semantic contribution. I agree with Arnold and use his argument as a basis for my analysis. I will here propose a single syntactic analysis for both IRCs and SRCs.

Consequently, I agree with Arnold (2007) that there is a difference in the semantics of the relative clause construction, but not in their syntax. I analyze this difference as a difference in the semantics of their syntactic head, a relativiser. This relativiser can be phonologically empty as in the case of *wh*-relatives and bare relatives. Alternatively, it can have the phonology *that*. Both, the empty and the non-empty forms of the relativiser are in principle compatible with an IRC and an SRC semantics. Further, I set my analysis within the framework of *Head-Driven*

Phrase Structure Grammar (HPSG, Pollard & Sag 1994a) combined with *Lexical Resource Semantics* (LRS, Richter & Sailer 2004).

At the same time, certain relative words, such as determiner-*which*, contribute a semantics that is only compatible with one of these constructions. My answer, consequently, relies on the interaction of the contributions of both, the relativiser and the relative word. At the same time, the similarities between IRCs and SRCs are equally captured.

Previous HPSG approaches built on the generalization that *that* RCs cannot be SRCs. Hence there is no HPSG analysis for relative *that* in SRCs. This assessment of the data, however, does not seem to be tenable. For example, Huddleston & Pullum (2002) state that this structure is relatively natural in present-day English, providing examples such as (4)

(4) She had long been accustomed to the solitary nature of her son's instincts, that I have tried - - - and failed - - - to stifle.

(Huddleston & Pullum, 2002: 1052)

The inclusion of *that* SRCs into the picture adds to the plausibility of my analysis of having the same syntactic structure for IRCs and SRCs, and a cross-classification of phonologically empty and non-empty relativisers with an IRC and an SRC semantics.

I extend my discussion beyond English by looking at relative clauses in Sorānī Kurdish. I argue that RCs in Sorānī Kurdish share essential properties with English bare RCs and *that* RCs, though Sorānī Kurdish has no equivalent of *wh*-RCs. As shown in (5), RCs are introduced by a function word *ke*, which is often optional in IRCs as in (5a), but obligatory in SRCs, see (5b). I will argue that the conditions on the occurrence of RC-initial *ke* are largely parallel to those for

English relative *that*.

(5) a. ewe guł-eke=ye (ke) ew pé=y dā=m
DEM *flower*-DEM=3.SG KE *he to*=3.SG *give*.PST=1.SG
'This is the flower that he gave me.' (Kim, 2010: 88) (IRC)

b. guł-eke=m dā be ānnā, *(ke) haw-pol=y
flower-DEF=1SG *give*.PST *to Anna* KE *same-class*=3.SG
to=ye
you=COP.PRES.3.SG
'I gave the flower to Anna, who is your colleague.' (SRC)

Including *that* SRCs into the core domain of English relative clauses will allow me to model the Kurdish data in a parallel way, while still leaving room for some language-specific properties.

The structure of this thesis is as follows. In **Chapter 2** I summarize generalizations about English SRCs and discuss some previous analyses. First, I present an overview of English RCs, largely following Huddleston & Pullum (2002) in terminology and in the structure of my presentation. I examine RCs from a formal or syntactic perspective and from a relational or semantic perspective. In order to do this, I distinguish the four types of RCs (supplementary RCs integrated RCs, Cleft Relatives and Fused Relatives). I look at the internal structure of relative phrases and I address relative words and their syntactic status in an RC. After this, I summarize the main differences between IRCs and SRCs as observed in the literature. I will provide twenty distinguishing criteria, some of which are crucial to the analysis of the types of SRCs this thesis is concerned with. In the last section, I discuss the syntactic and semantic integration of SRCs by giving a synopsis of three approaches: a radical orphanage approach, a non-radical orphanage approach, and a syntactically integrated approach. I will opt for the syntactically integrated approach for my own analysis.

Chapter 3 is devoted to the framework I use in my analysis. I will use a combination of *Head-Driven Phrase Structure Grammar* (HPSG, Pollard & Sag 1994a) and *Lexical Resource Semantics* (LRS, Richter & Sailer 2004) to model the syntax and the syntax-semantics interface.

Given that English RCs contain two Unbounded Dependency Constructions – one for the fronting of a relative constituent and one within the relative constituent, I present Pollard & Sag's (1994a) percolation mechanism for non-local features. I then show Pollard & Sag's (1994a) treatment of English RCs, illustrating how they deal with gaps and how they account for them inside RCs. I also present the more recent, constructional approach to RCs in HPSG, as was taken as the syntactic basis in Arnold (2004).

Finally, I put forward a presentation of the basic ideas and techniques used in LRS, which I later employ in analysing the semantics of SRCs.

In **Chapter 4** I systematically investigate the occurrences of determiner-*which*. This is necessary because determiner-*which* SRCs are not a very common phenomenon and there is no clear empirical basis for them in the literature. In this chapter, I investigate the potential complexity of relative phrases with determiner-*which*. Huddleston & Pullum (2002) present six types of upward percolation, also known as *pied piping*, within English relative constituents. The process of upward percolation refers to the observation that, in English RCs, often not just the relative word is fronted, but a larger constituent containing it. In example (6), determiner-*which* is embedded inside an NP, *which wagon*, which is part of an PP, *in which wagon*, and further embedded inside an NP.

(6) It took them half a day to reach the nearest village on their mule-drawn carriage, [the wheels [of [which wagon]]] were wooden. (COCA)

I investigate whether it is possible for determiner-*which* to occur with exactly the same percolation structures found with pronoun-*which*. I conducted a questionnaire study to explore this question. I conclude from my study that the upward percolation found in relative phrases with determiner-*which* are exactly the same as those with pronoun-*which* SRCs. Consequently, there seems to be no empirical motivation for adding special constraints that would make the percolation of determiner-*which* more restricted than that of other relative words. I conclude from this that the description of an SRC containing determiner-*which* should be very close to that of a SRC containing pronoun-*which*.

In **Chapter 5**, I propose a modelling of determiner-*which* in HPSG. Given the empirically attested strong similarities between SRCs and IRCs, I adapt the HPSG analysis that Pollard & Sag (1994a) provide for IRCs to SRCs with some amendment. Using the theory of Unbounded Dependencies in Levine & Hukari (2006), I can assume a unified treatment of subject- and non-subject gaps. This leads to a considerable simplification of the theory of relative clauses as well. In particular, there is no need to assume a different treatment for subject gaps and non-subject gaps. I introduce the semantic representations that I will use for SRCs in general and discuss some special semantic properties of determiner-*which* SRCs. Finally, I present an HPSG analysis of pronoun-*which* SRCs and determiner-*which* SRCs, including both the syntactic and semantic aspects.

Chapter 6 consists of two main parts. In the first part, I present an overview of the construction of SRCs with relative *that* along with an exploratory questionnaire study. The questionnaire study provides empirical support that *that* SRCs should be included in a formal modelling of English RCs. The second part contains a syntactic and semantic analysis of relative *that* with SRCs in HPSG. I treat relative *that* as a relativiser in all its occurrences, independently of whether

it occurs with a subject gap or a non-subject gap inside the RC, and whether it occurs with an IRC or an SRC.

In **Chapter 7** I provide an overview over the internal and external properties of RCs in Sorānī Kurdish. In my analysis, I concentrate on how RCs are introduced in Kurdish. I argue that RCs in Kurdish share essential properties with English bare RCs and *that* RCs. Consequently, I will model the Kurdish data in parallel to the analysis I propose for English in Chapter 6.

Chapter 8 summarizes the main results of this thesis.

Chapter 2

Previous Approaches to Supplementary Relative Clauses

This chapter gives a summary of previous approaches in analysing SRCs. Section 2.1 is a general introduction to RCs; where three main differences between RCs and *content* clauses are discussed. Next is an overview of English RC in section 2.2 where I explain the two main classifications of RCs into types – the formal/syntactic and the relational/semantic perspectives. In this section's subsections, I characterize the four types of RCs in 2.2.1, relative phrases in 2.2.2, and relative words and their syntactic status in the RC in 2.2.3.

In section 2.3, twenty criteria distinguishing IRCs from SRCs are presented. Finally, I discuss the syntactic integration of SRCs in 2.4 and I summarize three approaches: a radical orphanage approach, a non-radical orphanage approach, and a syntactically integrated approach. In section 2.5 I explain how SRCs are semantically integrated into their main clauses. In section 2.6 I give a summary of the chapter.

2.1 A general introduction to relative clauses

When it comes to defining a criterion or a construction, one can see the diversity among linguists in terms of their preferences to provide the construction a typological, semantic, and syntactic definitions or combinations thereof. One semantic/functional definition of an RC is 'We consider any syntactic object to be a relative clause, if it specifies a set of objects (perhaps a one-member set) in two steps: a large set is specified and called a domain of relativization, and then restricted to some subsets of which a certain sentence or the restricting sentence is true.' (Keenan & Comrie, 1977: 63-64)

However, a synchronic typologist such as Lehmann (1984) gives a definition of RCs which is an amalgamation of semantic, functional and formal principles. 'A relative construction is a construction consisting of a nominal (or a common NP, in the term of categorical grammar) (which may be empty) and a subordinate clause interpreted as attributive modifying the nominal. The nominal is called the head and the subordinate clause the relative clause. The attributive relation between head and relative clause is such that the head is involved in what is stated in the clause.' (Lehmann, 1984: 247).

What needs to be clarified at this point is that there are three major differences between relative clauses and other subordinate clauses, in particular *content clauses*. First, in contrast to an RC, a content clause doesn't have to contain an anaphoric reference. Sometimes, ambiguity arises when the verb in the sentence can be used both transitively and intransitively. In this case, the sentence can contain an RC in one reading and the relative sense can be excluded in the other reading. Example (1) is not ambiguous while example (2) is ambiguous as we can imagine that there is a gap in the RC that has *idea* as its antecedent. On the

other hand, *that* introduces a clause which functions as a complement to the noun *idea*.

(1) I bought the game that you recommended.

(2) They rejected the idea that we have developed.

Second, concerning NP structures, RCs act as their modifiers while content clauses act as their complements.

The final distinction is that we can't replace *that* by *which* or any other relative pronouns in content clauses, which would otherwise prevent ambiguity in an RC as in example (2), without a change in the core meaning of the sentence. Here, I repeat example (2) and change the relative word to *which*, this eliminates the ambiguity mentioned above.

(3) They rejected the idea which we have developed.

So far one general type of RCs has been introduced and a further type is a correlative RC, an adjoined RC. This type, in contrast to the embedded RCs, is adjoined. As such, a construction can be left-adjoined or right-adjoined. However, Hendery (2007) states that many linguists only consider left-adjoined RCs to be true correlative RCs. The reason is that true correlatives, left-adjoined RCs, must include a demonstrative-marked phrase or a demonstrative pronoun while the right-adjoined RCs need not. Hendery (2007: 14) proposes that right-adjoined RCs can contain multiple heads while this is not the case with the left-adjoined RCs.

No more will be mentioned about this type as it is not the concern of this thesis.

2.2 English relative clauses: An overview

According to Huddleston & Pullum (2002), there are two major classifications of RCs in English depending on their constructions; namely:

1. Formal types: These depend on the internal syntax of the RC. The formal types are distinguished according to whether they contain one of the special relative words *who*, *which*, etc., or the subordinator *that*, or simply a gap, a missing constituent.

2. Relational types: These types depend on the external syntax. These differences are based on the relationship between the RC and the bigger constituent that contains it. This will be discussed in detail in section 2.3 of this chapter.

   The formal types according to Pollard and Sag (1994) are:

   (a) *Wh*-RCs

   (4) The chocolate which you bought for the party was expired.

   - Subject *wh*-relative

   (5) The chocolate which was expired was used.

   - Non-subject-relative

   (6) The chocolate which you used was expired.

   (b) Non-*wh*-RCs

   - *that* RCs

   (7) The chocolate that was expired was used.

- *that*-less RCs (finite).

(8) The chocolate you used was expired.

The *wh*-relative phrases are usually followed by

1. A clause, as in (9).

(9) This is the man whose car I repaired.

2. Or sometimes they are followed by a finite VP, as in (10).

(10) She is the one whose Secretary sent the invitation letter to us.

Huddleston & Pullum (2002: 1034) define RCs as "they are related by their form to an antecedent. They contain, within their structure, an anaphoric element whose interpretation is determined by the antecedent. This anaphoric element may be overt or covert".

Overt anaphora: this is where the relative words are present; this type is called *wh*-RCs.

Covert anaphora: there are two sub-types of this type, *that* relative and bare relative. This type is called non-*wh* RCs.

Huddleston & Pullum's justification for using *relative construction* to refer to the phrase instead of the traditional term *relative clause* is clarified through the following example:

(11) a. I agree with most of the things that your father was saying. (clause)

b. I agree with most of what your father was saying. (NP)

The structure in example (11a) is considered a clause; however, example (11b) contains an NP that can simply be replaced by the pronoun *it* and both are

considered RCs. I find their argument quite reasonable; however, due to the fact that the term *relative clause* is widely used in the literature, and that I am concentrating on clauses, I decided to use the term *relative clause*.

Additionally, a gap occurs in all of the types of RCs as is clear from all of the examples above. Similar to the other types of RCs, a gap can occur with relative *that* as in (13), as well as with bare-relatives as in (12).

(12) I accept the advice my neighbor gave me.

(13) I accept the advice that my neighbor gave me.

In example (13), the omission of the relative *that* is optional as the RC is in the object position; however, this is not permitted when it is in the subject position as in (14).

(14) The advice my *(that) was given to me was useful.

2.2.1 The four types of relative clauses

There are four types of RCs in English: *Integrated RCs* (IRCs), *Supplementary RCs* (SRCs), *Cleft Relatives* and *Fused Relatives*.

IRC: This type is often discussed in the literature on RCs. The antecedents that occur with this type of RC are mostly nominal as shown in example (15):

(15) The person who fixed her car was an engineer.

The information in the RC *who fixed her car* is an integral part of the matrix clause in the sense that the RC delimits the set denoted by the antecedent. There are occasions where the clause can modify heads other than NPs. It can sometimes

modify a superlative adjective as in example (16) or an interrogative preposition as in example (17).

(16) John is now the richest he has ever been.

(17) When can we meet that is convenient for both of us?

SRC: According to Huddleston & Pullum (2002: 1034), SRCs only provide additional information about their antecedents and they are not integrated to the sentence whatsoever. In section 2.4, this claim of considering SRC as not at all integrated into the main clause and a counter argument, a strict criticism, on this claim, is discussed. Huddleston & Pullum (2002) further discuss that the antecedent in example (18) is not *astronaut*, it rather is the whole NP *the astronaut*; hence, the RC does not identify the antecedent but only provides additional information about it.

(18) The astronaut, who planned the journey, was a genius.

This type of RC allows a wider range of antecedents than IRCs do. An entire clause can be the antecedent of SRCs as in:

(19) She didn't pass the test, which was clear from the beginning.

Here, the relative word *which* refers back to the whole clause *she didn't pass the test*. SRCs can also have a complete adjective phrase as their antecedents as in example (20).

(20) The children were interested in his performance, which I believe the teachers were too.

This property of SRCs is discussed in detail in section 2.3.

Cleft relatives: In this type of RCs, there are *foregrounding* and *backgrounding* constructions. In example (21), *the defender* is foregrounded and *asked the judge for the second trail* is backgrounded.

(21) It was the defender who asked the judge for a second trail.

Fused relatives: This type is always of *wh*-RC type. Here, the antecedent is fused with the rest of the RC.

(22) What you heard was a volcanic eruption.

These two last types will not be discussed in this thesis.

2.2.2 Relative phrase

Relative phrases are only found with *wh*-RCs. They occur either in subject position, see (23a), or in prenuclear position, i.e. in the position before the subject, as in (23b).[1]

(23) a. I took the food [which was offered to me].

b. I took the food [which they offered me].

Relative phrases are of two types:

1. Simple relative phrases: this type only contains a relative word. Most of the above mentioned examples are of this type.

2. Complex relative phrases: in this type there are other elements besides the relative word, as in example (24).

[1]In Huddleston & Pullum's (2002) terminology, the subject and the predicate form the nucleus of the clause, so the prenuclear position is the position preceding this nucleus.

(24) This is the family [[whose daughter] our son married].

The example in (24) contains a complex relative phrase as the relative word forms a constituent with *daughter*. In this sentence, the gap is not co-indexed with *whose*. However, the word *family* is the antecedent of *whose* while the gap is rather co-indexed with *whose daughter*. This can also be illustrated with constructions such as the one in (25):

(25) This is the mixture [[with which] you can make the party cookies].

The noun *mixture* is the antecedent of the relative pronoun, *which*. Inside the RC, the gap is co-indexed with the prenuclear relative phrase, which is the PP *with which*.

In an attempt to classify the construction of complex phrases, Huddleston & Pullum (2002: 1043) introduce six patterns which involve *upward percolation*, also known as *pied piping*. Upward percolation is when a *wh*-phrase is moved it might also drag along an NP or a PP hosting the *it*.

It is important to introduce these constructions here because one of the main structures that is investigated in this thesis is relative *which* as a determiner, determiner-*which*, and this particular structure involves upward percolation. The types of upward percolation are:

1. From the complement of preposition to PP (Comp of preposition to PP): this upward percolation is optional.

   (26) a. Kim was hiding behind the curtain.

   b. The curtain behind which Kim was hiding. (Huddleston & Pullum, 2002: 1040)

2. From PP to adjective (PP to AdjP): this type is only limited to SRCs. In some cases, it is optional in others it is obligatory. In the following example this construction is obligatory.

   (27) Many students applied for the program, best among which will be selected.

3. From PP complement of a noun to NP (PP to NP): This type is mostly seen in SRCs.

   (28) They put forward a proposal, the problem of which was time limitations.

4. From NP to non-finite verb.

   (a) From NP to infinitival

   (29) They have started a project, to develop which, they had to go around the world.

   (b) From NP to gerund-participial

   This type is only restricted to SRCs.

   (30) She has started a degree, finishing which involves six years of study.

5. From genitive *whose* to NP

   (31) The agent, whose equipment we borrowed, is missing.

6. From determiner-*which* to NP: here *which* is a determiner not a pronoun, hence; upward percolation is obligatory. This is only found in SRCs.

(32) Alan Hartman proposed to cut down the transportation's expenses, which solution the congress found sufficient.

(33) I said that it might be more efficient to hold the meeting on Saturday morning, which suggestion they all enthusiastically endorsed. (Huddleston & Pullum, 2002: 1043).

Determiner-*which* with these types of upward percolation is one of the main concerns of this thesis and it is thoroughly investigated in chapter 4.

2.2.3 The relative words and their syntactic status in the relative clause

The relative words that are mainly used in constructing RCs are: *that*, *which*, and *who*. Nonetheless, other *wh*-words occur in constructing RCs, occasionally. *Which* may occur as a determiner when it is in SRCs only, and of course, it is a pronoun in the other types of RCs. The major differences between *who* and pronoun-*which* is that the latter is more commonly used to refer to inanimate objects and animal antecedents. Nonetheless; there are cases where *which* is used to refer to animate beings other than animals. This is when the antecedent is the complement of the auxiliary verb *be* in SRCs as in (34a) or *have* as in (34b).

(34) a. They thought we were the rescuers, which we were.

b. They have a great leader, which we don't have.

In example (34b), in the SRC, the antecedent is a predicative NP. The NP does not refer to an individual, but to a property, which is *being the rescuer* or *being a leader*. It is rather the type of the leader than just a leader alone that they lack.

Relative *who* is commonly used to refer to human beings, but it is not very uncommon to use *who* to refer to animals, especially pets or collective animals[2].

(35) My cat Betty, who is very greedy, would never share her toys.

Whose is used with both personal and non-personal antecedents.

(36) Ameen's car, whose battery was dead, stopped working in the middle of the street.

(37) Ameen, whose wife is Armenian, was in our group.

2.3 Criteria distinguishing IRCs and SRCs

In this section, most common criteria of the two constructions are presented as they are noticed by Ross (1967), McCawley (1982), Sells (1985), Fabb (1990) and Huddleston & Pullum (2002). The criteria are divided into four categories: First, the differences in the two structures' prosody, second, the antecedents they take, third, the syntactic differences, and fourth the semantic differences.

Regarding the differences in the prosody of the two structures the following criteria are noticed.

Criterion 1, the prosody of SRCs is completely different from that of the matrix clause, in a sense that the intonation contour of SRCs is not of that of the matrix clause; there usually is a short pause before the utterance of SRCs in speaking. This difference in prosody is also reflected in writing as SRCs are separated from

[2]The relative word *where* acts like an adjunct of spatial location, a goal complement, or a complement of locative preposition. There are many compounds of *where* + preposition such as whereby, whereas, whereupon, whereon, whereto, whereof, wherein, etc. However, only few of them are common namely: *whereby*, *wherein* and *whereupon*. *When* is generally considered an adjunct of temporal location. *Why* is only used with IRCs that have a reason as their antecedent. *whence* is used in formal discourses, it refers to spatial sources.

their matrix by a comma preceding the clause and if the clause is not final, a comma follows it as well.

(38) The report, which I was supposed to finish tonight, is still on my desk in the office.

(39) I finished the report, which you asked for.

On the other hand, IRCs are a part of the intonation contour of their matrix clause. IRCs are not separated from their matrix by a comma or any other punctuation marks.

(40) A report which I was supposed to finish tonight is still on my desk in the office.

Criterion 2, a reduction of a vowel sound in the pronunciation of a word is possible with IRCs, on the contrary; it is not permitted in SRCs.

(41) The people who will/'ll be there tomorrow.

(42) John, who will/*'ll be there tomorrow. (Fabb, 1990: 71).

The significance of this type of differences, the differences in prosody, will not be discussed in this thesis. Their presence here is merely to show one of the major differences between the two constructions.

The next set of criteria address the differences between the antecedents the two constructions attach to.

Criterion 3, there is a difference in their external structure as there is a limitation of the kinds of antecedents these two types can take. Only SRCs can take a whole clause as their antecedent; Fabb (1990: 60) states that SRCs take all of the sentence elements as their antecedents: namely verbs, projections of prepositional phrases, nouns, adjectives, adverbs and even sentences, as was discussed in sec-

tion 2.2.1. As for proper names, unless they are preceded by a determiner, they cannot be antecedents of IRCs; they only are antecedents of SRCs. The following examples, except for example (47), are from Fabb (1990: 60).

(43) Peter put it under the table, where I had put it earlier. (projection of P)

(44) The cheese was bought by John, which was fortunate. (sentential antecedent)

(45) John luckily escaped, which I unluckily didn't. (VP antecedent)

(46) Bill is drunk all the time, which is probably how you would like to be. (AdjP antecedent)

(47) a. The school dismissed Ellen Harper, who pretended to be the substitute teacher. (SRC with a proper name)

b. The school dismissed the Ellen Harper who pretended to be the substitute teacher. (IRC with a proper name preceded by a determiner)

More on this criteria will be discussed in chapter 5.

Additionally, when an IRC has a definite NP as its antecedent, it is the noun without the determiner that is considered as the antecedent not the whole NP including the determiner (Huddleston & Pullum, 2002: 1038). However, if the noun is preceded by any other element which comes after the determiner, this element is included in the antecedent.

(48) The interesting book that you have read was in French, but the interesting book that I have read was in Spanish.

In example (48), *interesting book* serves as the antecedent of the following RC in both conjuncts. However, in SRCs, it is the whole NP including the determiner that is the antecedent of the RC. Consequently, there is just one interesting book. This makes the SRC-version of example (48), given in (49), infelicitous.

(49) ?The interesting book, which you have read, was in French, but the interesting book, which I have read, was in Spanish.

Criterion 4, Huddleston & Pullum (2002: 1060) state that relative pronouns in IRCs must be co-referential with another element in the antecedent that precedes the RC containing the relative pronoun. Concerning IRCs, because they only take nominal antecedents, the case is easy, for example, the subject or the object in the antecedent is co-referential with the relative pronoun in the RC.

However, the case is a bit different with SRCs as they can take an entire sentence as their antecedent. With reference to NPs and prepositional phrases, the referential issue is straightforward as it is similar to that of IRCs. The rest of the elements that SRCs take as antecedents, such as adjective phrases, adverbial phrases and VPs, can also be co-referential with the relative pronoun in the RC as these elements are originally predicates.

Criterion 5, Ross (1967: 313) observed that SRCs cannot have quantified NPs as their antecedents. They also cannot have pronouns in them that are bound by an externally quantified NP. On the other hand, the antecedent of IRCs can be part of a quantified NP and also an extra quantified NP can bind pronouns that may occur inside the IRCs. This generalization is opposed by Sells (1985: 2). He shows that SRCs can have quantified NPs as their antecedents, His example is given in (50).

(50) Every chess set comes with a spare pawn, which you will find taped to the top of the box.

A discussion on this is presented in chapter 5.

Criterion 6, Huddleston & Pullum (2002: 1060) observe that noun phrases with quantifiers such as *no*, *any* and *every* tend to have a non-specific reference. This

group of noun phrases includes combinations with words such as *-thing*, *-one* and *-body*, but also combinations with ordinary nouns. According to Huddleston & Pullum (2002), such combinations cannot be antecedents of SRCs, see (51a). However, they are possible with IRCs, as in (51b).

(51) a. * No candidate, who scored 40% or more, was ever failed.

b. No candidate who scored 40% or more was ever failed.

(Huddleston & Pullum, 2002: 1060)

The syntactic differences between the two constructions are as follows:

Criterion 7, Arnold (2007: 278) states that the SRCs must be finite and contain a *wh*-relative word. This makes them syntactically more constrained than IRCs. The content of SRCs is separate from the content of their host clauses, and their relative pronouns are in fact 'normal' anaphoric pronouns. Due to the earlier mentioned facts, Arnold (2007: 276) states that the antecedent of SRCs are the same as those of a normal anaphoric pronoun. In return, this gives SRCs the privilege of having a wider range of antecedents. Additionally, the relative constituents in SRCs can be replaced by pronouns without a change in their meanings.

(52) a. I liked this book, which you recommended.

= I liked this book. You recommended it.

b. I liked this book that you recommended.

≠ I liked this book. You recommended it.

Criterion 8, Huddleston & Pullum (2002: 1061) state that stacking is only possible with IRCs. Stacking is when an RC can combine with an antecedent to become an antecedent of another RC in the same sentence. The examples are from Huddleston & Pullum (2002: 1061).

(53) I like the tie you wear that your sister knitted for you.

(54) *They've given the job to Max, who has no qualifications, who starts next month.

This claim is again criticized by the High Syntactic Attachment Analyses and Syntactically Integrated Analyses Approaches of SRCs which are discussed in sections 2.4.3 and 2.4.4.

Criterion 9, it has been previously assumed that SRCs are not syntactically related to the matrix clause. This claim is disputable and it is comprehensively discussed in section 2.4. However, IRCs are syntactically linked to their antecedents and they post-modify them. Both together form a big NP that contains a post modifier which is the IRC in this case. Fabb (1990: 57) states that SRCs are not syntactically related to the sentences they occur in. The only relation is between the relative pronoun and the antecedent. On the contrary, IRCs are completely syntactically linked to the sentences containing them.

Criterion 10, Arnold (2007: 274) explains that SRCs are only adding information about the phrases, the antecedents, whose meanings will not otherwise, without SRCs, be affected. Henceforth, SRCs can be easily deleted in the sentence and still the main meaning of the sentence will not be dramatically changed.

(55) Lilly, who failed the exam, has dropped out of school.

(56) Lilly has dropped out of school.

Again, this claim has been strongly criticized and this criticism and discussion will be presented in section 2.4 in detail. Arnold adds, in contrast, IRCs are interpreted intersectively, this means that they restrict the denotation of the nouns they are modifying.

(57) My sister who failed the exam has dropped out of school.

(58) My sister has dropped out of school.

Criterion 11, there is a difference between the constructions in the types of formal relatives they take. IRCs can be equally seen through all the three types of formal relatives. Contrarily, only *wh*-relatives are found freely with SRCs. Relative *that* is only marginal, that is, it is rare and a matter of discussion according to native speakers. This latter issue has been tested and the results of which are presented and discussed in chapter 6. Finally, bare-relatives are not found at all in SRCs.

Criterion 12, Huddleston & Pullum (2002: 1061) state that, conventionally, RCs have the form of declarative clauses but SRCs can have the form of non-declaratives as well. This means that SRCs can also have the form of imperatives or *closed*-interrogatives and/or *open*-interrogatives. In the declarative types, question tags can be attached, as shown in (59).

(59) I didn't get much response out of Ed, who seemed rather out of sorts, didn't he? (Huddleston & Pullum, 2002: 1061)

Criterion 13, *wh*-movement in a sentence moves the whole phrase in IRCs. On the contrary, *wh*-movement cannot move SRCs. The examples are from Fabb (1990: 70).

(60) We taught the boys, some of whom were deaf, French.

(61) Who did we teach, some of whom were deaf, French?

(62) * Who, some of whom were deaf, did we teach French?

Criterion 14, RCs can also occur in extraposed positions in a sentence. This position is mostly restricted to IRCs. The following examples are from Fabb

(1990: 59).

(63) I met the man yesterday who is leaving for Spain.

(64) I met John yesterday, who I like a lot.

Nonetheless, Fabb (1990) explains that extra-position, in general, is not usual with RCs, he states that the antecedents in both IRCs and SRCs must precede the RCs; additionally, the relative pronoun must be closest to its antecedent in both of the cases. Fabb asserts that, merely in the case of an adverbial interruption, it is possible for the RC to be separated from its antecedent.

Criterion 15, McCawley (1982: 96) observes the following contrast between IRCs and SRCs in VP ellipsis.

(65) Anna found the girl who saved her life, and so did Soma. (IRC)

There are two possible interpretations for the above example; first, it can mean *Soma found the girl who saved her, Soma's, life too.* Second, it can mean *Soma has also found the girl who saved Anna's life.* However, in example (66) this ambiguity is eliminated in the SRC because, even though the SRC occurs next to its VP-internal antecedent, *her biological father*, it is interpreted outside the VP.

(66) Anna found her biological father, who abandoned her ten years ago, and so did Soma. (SRC)

In example (66) the ambiguity that is mentioned in example (65) does not occur. McCawley concludes that ambiguities arise with the interpretation of IRCs and not SRCs due to the fact that the VP-internal SRCs don't form part of the VP; however, the VP-internal IRCs are a part of the main VP.

The semantic differences between the two constructions are:

Criterion 16, concerning their choices of the relative words and phrases, IRCs are more restricted in this case than SRCs. IRCs cannot have, namely: *where-upon*, determiner-*which*, and upward percolation such as from PP to adjective and gerund participles. SRCs can have all of the mentioned structures. One more difference in this aspect is that SRCs cannot have *why* as a relative word, but IRCs can. The following examples are from (Huddleston & Pullum, 2002: 1050)

(67) a. That is the main reason why they won't help us. (IRC)

b. * That is the main reason, why they won't help us. (SRC)

Criterion 17, according to Arnold (2007: 275), the content of SRCs fall outside the scope of negation, and, of course, this is reversed in IRCs, the content of which falls into the scope of negation.

(68) a. My sister has a friend who is a sailor.

b. My sister doesn't have a friend who is a sailor.

c. My sister has a friend, who is a sailor.

d. ? My sister doesn't have a friend, who is a sailor.

Example (68d) cannot have a reading where the indefinite antecedent *a friend* is interpreted in the scope of the negation. Therefore, it is considered odd, which is indicated with the marking '?'.

Criterion 18, another significant difference between these two constructions is that SRCs cannot have 'externally licensed' negative polarity items, NPIs, conversely IRCs can contain them (Arnold, 2007: 277). NPIs, such as English *any* and *ever* are very common in many languages 'Their hallmark property is exclusion from positive assertions with simple past (i.e., episodic sentences that make reference to a single specific event)' (Giannakidou & Quer, 1997: 1661).

(69) a. She didn't propose a solution which was of any use to us.

b. * She didn't propose a solution, which was of any use to us.

The interpretation of the RC as an independent contribution to the discourse is discussed in criteria 19 and 20.

Criterion 19, Burton-Roberts (1999) states that SRCs can contain a discourse relation, to their hosts which otherwise is impossible for IRCs. In examples (70), the SRC in (70b) has a *consequence* relation to its host while this is not possible with the IRC in (70a).

(70) a. * Anna takes good care of her electronic devices which therefore last longer. (IRC)

b. Anna takes good care of her electronic devices, which therefore last longer. (SRC)

Criterion 20, SRCs can show further events in a narrative series as they have a continuative use. The meaning conveyed with continuative SRCs can easily be interpreted as the form of *and* + an IRC.

(71) I gave Anna the book, who returned it to the library.

(72) * I gave the girl the book who returned it to the library.

Example (71) means, I gave the book to Anna and then she returned it to the library.

In this section the concentration has mostly been on the most common differences between the two structures, the purpose behind this is that in the coming section, section 2.4, I point out most of these differences and try to test whether in fact SRCs are much different from IRCs.

To sum up, the following differences between IRCs and SRCs were presented: the differences in their prosody, the kind of antecedent they take, the syntactic and semantic differences. However, the differences and the criteria that I will be concentrating on later for my arguments are the following: the differences in their meanings, their relations to their antecedents and the types of antecedents they take, their positions in sentences and their syntactic constructions.

2.4 The syntactic integration of SRCs

2.4.1 The three prominent approaches to the grammar of SRCs

Here, I shortly discuss the three common approaches and later will only concentrate on the Radical Orphanage Approach and its counter argument, Syntactically Integrated Analyses, which form the basis of my analyses throughout the thesis.

1. Radical Orphanage Approach (henceforth ROA). Fabb (1990), Espinal (1991), Burton-Roberts (1999) and Huddleston & Pullum (2002), among others, propose that SRCs are not syntactically a part of the matrix clause.

2. High Syntactic Attachment Analyses; under this approach, the element is still an orphan but the analysis is a non-radical one. Ross (1967), Emonds (1979), McCawley (1981) and Demirdache (1991) state that SRCs are attached high to the root sentence, S in the tree. They propose that they might be attached to even higher nodes.

3. Syntactically Integrated Analyses : Kempson (2003) and Arnold (2004) oppose the latter and former proposals by stating that SRCs are syntactically connected to their host clauses.

2.4.2 Radical orphanage approaches (ROA)

Fabb (1990) proposes that unlike an IRC, an SRC is not a modifier and 'in fact has no syntactic relation to its host/antecedent' (Fabb, 1990: 57). Fabb states that since SRCs are not modifiers, their relation to their host is not one of predication unlike for IRCs. Hence, SRCs are not coindexed with the noun they are commenting on, the antecedent. Jackendoff (1977: 169), who also is a supporter of the ROA, clarifies that SRCs are not inside the noun projection and this is why they have no link to the syntax of their host clause. He provides the following structure to the syntax of SRCs.

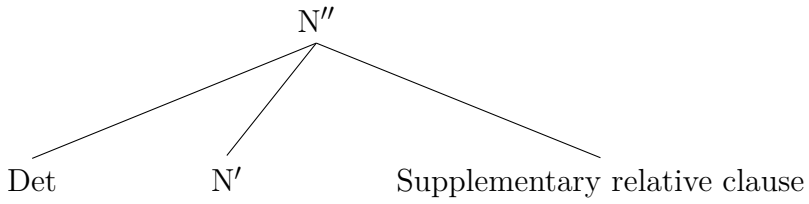


Figure 2.1: The syntactic structure of SRCs

Cinque (1982: 260) analyzes SRCs the same way he analyzes parentheticals, as he assumes that English, French and Spanish SRCs behave and function precisely like parentheticals. Parentheticals, as Dehé (2009: 569) states, are 'expressions of varying length and complexity, syntactic category and function, which are interpolated into the current string of the utterance […] they are linearly, but not hierarchically, integrated into the structure of the host clause.' Dehé (2009: 569) also suggests that parentheticals such as SRCs, are 'obligatorily phrased in a separate intonational phrase' and are syntactically outside the representation of the embedding sentence. Dehé's main idea is that a structure that is loosely linked to the syntax of its host cannot be prosodically integrated into it.

Additionally, Espinal (1991), introduces a term *disjunct constituents* and she

proposes that SRCs are constituents of this type. Under this term she introduces constituents in which the grammatical structures related to the surface string have some separate trees or separate *planes* or *tiers*. These *planes* or *tiers* only relate to each other at the most superficial levels such as Phonological Form (PF) and Conceptual Structure (CS). Nevertheless, the operations that relate elements from the planes PF and CS are limited, this means, they merely links them together. The first tree in Figure 2.2 illustrates an ROA analysis of example (73), and the second tree shows its syntactically integrated analysis.

(73) Anna, who I admire, told Soma, who I hate, my secret.

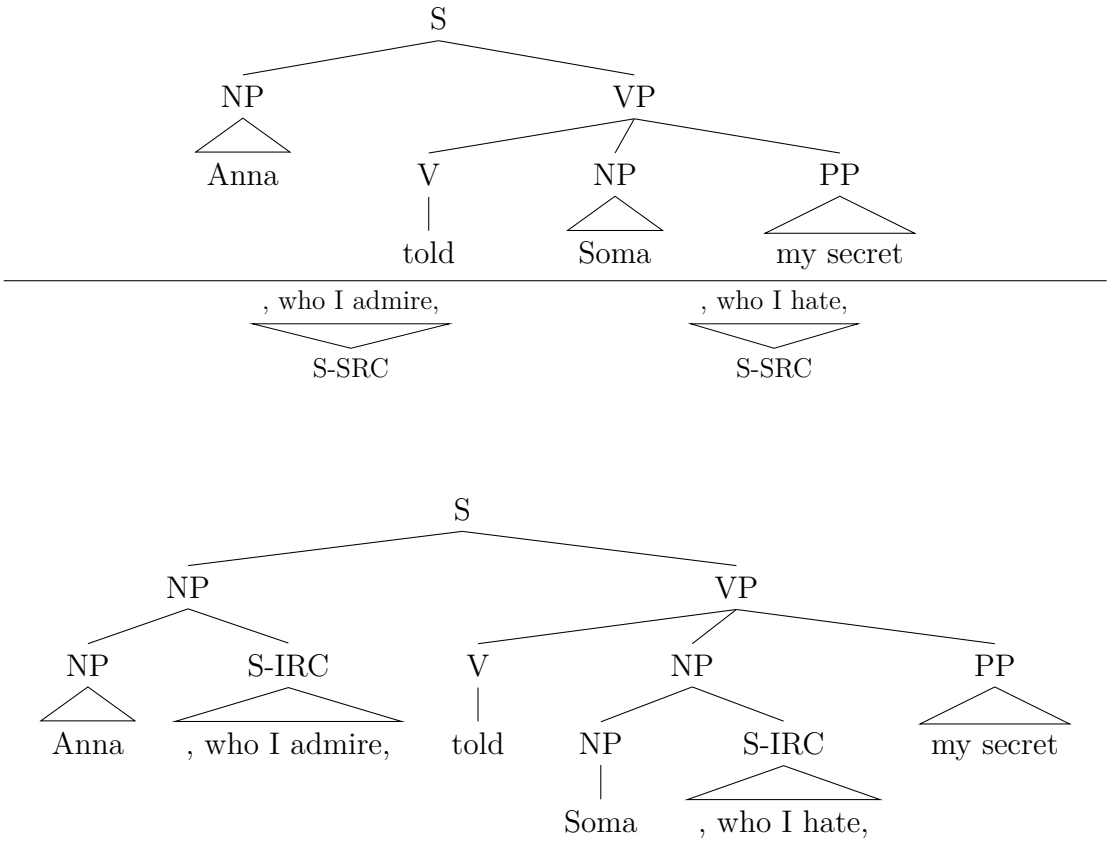


Figure 2.2: An ROA analysis vs syntactically integrated analysis

According to the ROA, the planes are separate; hence, the matrix clause and SRCs are largely independent of one another. This results in having the elements of one plane not being able to interact grammatically with the elements of another plane. This implies that SRCs are treated as separate clauses from the maximal clause hosting them. The ROA accounts for the following properties of SRCs:

1. SRCs should not be in the scope of a negation in the matrix clause. (Criterion 17)

2. SRCs should not be a part of VP ellipses. (Criterion 15)

In the following section, two analyses, which are both against the ROA, are presented. The first analysis is de Vries's (2006) analysis in section 2.4.3, his view is an example of The High Syntactic Attachment Analyses of SRCs. The second analysis is Arnold (2004) and (2007). Arnold's analyses are examples of Syntactically Integrate Analysis of SRCs.

2.4.3 de Vries's (2006) analysis of SRCs

de Vries (2006) states that SRCs can be considered as pairs with appositions as they both are conjuncts to their antecedents and they specify the antecedents' meaning. Here, I mention the types of RCs that are presented in de Vries (2006). de Vries divides RCs into three major types: appositive, restrictive and maximalizing RCs. The last type, maximalizing RCs, was introduced in Carlson (1977) and was later studied in depth by Grosu & Landman (1998). An example of this type is:

(74) I spilled the coffee that there was in the pot. (de Vries, 2006: 229)

The substance degree relative in example (74) is that the whole amount of the coffee in the pot was spilled, and there is no distinction between this coffee and other coffees. Moreover, the RC is essential to the meaning of the sentence. This type of RCs will not be further discussed as it is not the main concern of this thesis.

In this section, I present a short summary of de Vries's (2006) approach to an analysis of SRCs. de Vries argues that SRCs and appositions can be considered the third type of coordination, which expresses a specification relation between the conjuncts. His main claim is that an SRC is a semi-free-relative which has a pronominal head that is usually empty. He also proposes that SRCs have the same syntax as IRCs. de Vries mentions two hypotheses, and he argues for one of them. The hypotheses are:

1. The Main Clause Hypothesis: this was first introduced in Ross (1967), who argues that SRCs are main clauses and they are coordinated with the matrix clause at deep structure.

2. The Subordinate Clause Hypothesis: this is the opposite of Main Clause Hypothesis. This hypothesis was first discussed in Jackendoff (1977), who proposes that SRCs are subordinate clauses and are embedded in the maximal projection of the antecedent; hence, the SRC with its antecedent form one constituent.

Here, it is important to highlight that there are two variants of an orphanage approach. The first one is the ROA, which means that SRCs are not part of the syntactic structure of the matrix clause. The second one is the Non-radical Orphanage Approach, NROA. In the NROA, SRCs are still not generated with their antecedents. The difference between the NROA and the ROA is that, in the

NROA, the SRC is syntactically present in the sentence. In brief, according to the the Main Clause Hypothesis, an SRC is a clause that is conjoined with the matrix clause and the structures interfering between the antecedent and the SRC are all right side extraposed in the sentence.

de Vries (2006) proposes that SRCs are combined into one constituent with their antecedents by coordination. In other words, he supports The Subordinate Clause Hypothesis. A thorough explanation of the Subordinate Clause Hypothesis, concerning how SRCs act like subordinate clauses and how they lose their properties as main clauses, i.e. they must not express imperative or interrogative forces, is presented in this section.

According to de Vries, one major difference between IRCs and SRCs is the scope of the determiner or the quantifier. In IRCs, the quantifier or the determiner that is added to the antecedent noun can take scope over the noun and the RC as well. On the contrary, in SRCs they take scope merely over the noun as in the following examples. This difference has been discussed in criterion 3, in Section 2.3.

(75) a. Every student who took the exam passed.

b. Every student, who took the exam, passed.

In example (75a) *every* has the meaning of the noun *students* and of the RC in its restrictor; conversely, in (75b) *every* has the meaning of the noun *student* but not the RC in its restrictor. de Vries argues that the restrictor of a determiner is determined by its c-command domain; hence, it is clear that IRCs are attached in the tree below the determiner or as a sister of the determiner, on the contrary, SRCs are not. To be more precise, IRCs are adjoined to the right side of the NP while SRCs are adjoined to the right side of the DP as shown in Figure 2.3

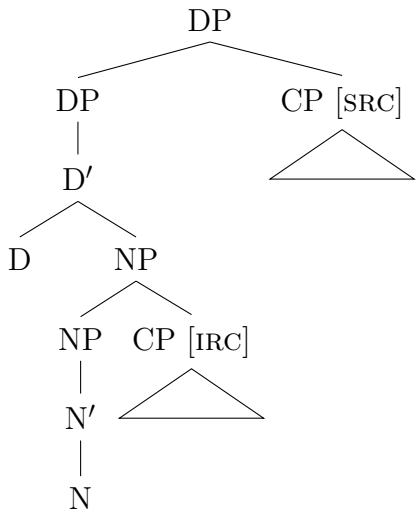


Figure 2.3: IRCs are adjoined to the right side of the NP while SRCs are adjoined to the right side of the DP

At first, de Vries (2006) investigates the idea that SRCs are coordinated, which was also proposed by Koster (2000: 22). Second, he studies the idea of considering SRCs as free-relatives that was first proposed by Canac-Marquis & Tremblay (1998), who suggest this as a variant of the ROA. Canac-Marquis & Tremblay (1998) essentially assumed that SRCs are free-relatives, and like ordinary RCs, they can stand adjacent to their antecedents. de Vries' conclusion is that the similarities between SRCs and IRCs are rooted in the relative part of the constructions, whilst their differences stem from the fact that SRCs are connected to the syntactic context through specifying coordination.

de Vries defines appositions as non-restrictive postnominal DP modifiers, an instance of which would be *my daughter* in the following example.

(76) Anna, my daughter, was selected by the committee.

One major goal that de Vries tries to achieve throughout his work on SRCs is that, when it comes to pining down the syntactic status of appositions, he argues that they should be analyzed as coordinated constituents.

de Vries (2006: 238) agrees with Quirk et al. (1985) in considering SRCs similar to coordinations as they both coordinate units as well as they link constructions of equivalent ranks. Quirk et al. (1985) state that coordinators such as *and* and *or* are sometimes used to explicitly mark appositions. de Vries confirms their statement and provides examples such as (77a) and (77b).

(77) a. The room at the right side of the house, or the study room, has the most beautiful view.

b. The room at the right side of the house, which is the study room, has the most beautiful view.

In example (77a) *or* is used to introduce the appositive. This suggests that SRCs are formed by coordination. There are three types of coordination, and the coordinators determine these types: conjunction, disjunction and specification. Sometimes a specifying phrase introduces the specification. The second conjunct is separated by a comma and a low intonation.

Specifications were first introduced by Kraak & Klooster (2000) and the definition they gave can also be used to define SRCs. They define specifications as asymmetric, in the sense that the first conjunct is specified by the second one, and the phonology of the two conjuncts is set off by low intonation. Klein (1976) assumed that an apposition is a reduced RC. See the equivalence of the examples in (78):

(78) a. Anna, my daughter, finished the project.

b. Anna, who is my daughter, finished the project.

Hence, if appositions are specifying coordination, the intuition here would be that SRCs are also specifying coordination to their antecedents. This implies that the antecedent forms a constituent together with the RC. A support for this

claim is that the constituent can be topicalized the same way an apposition or a conjunction can be. See example (79)

(79) a. Sandy, who I'm sure you remember, I see regularly

b. *Sandy, I see, who I'm sure you remember, regularly.

Another observation that confirms the claim is that SRCs, as with appositions and coordination, are capable of stacking, as shown in example (80)

(80) John, who hands in all his work on time, who is on the student council, who even likes Math, will certainly get a good job.

This goes against Huddleston & Pullum's (2002) claim allowing stacking only with IRCs as was mentioned in criterion 8, in section 2.3.

In brief, appositions and SRCs can both be considered as the third type of coordination both labeled specifiers.

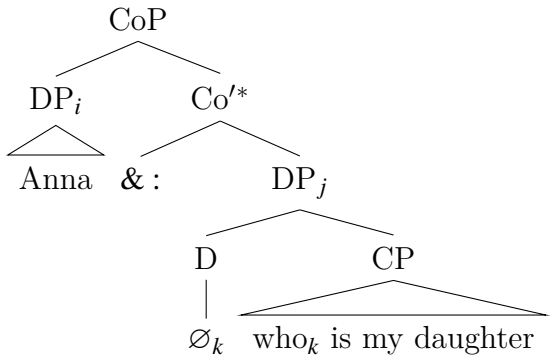
Additionally, de Vries tries to explain the claim he makes for considering SRCs as free-relatives as given in (81).

The sign &: means specifying coordinative connection in example (81)

(81) [[$_{DP1}$ Anna]&: [$_{DP2}$ [$_{CP}$ who is my daughter]]

A free-relative can be coordinated with a DP because it functions as a DP itself. Hence, a syntactically balanced coordination is present in the formation of SRCs. In example (80), the first DP is specified by the second one. A free-relative modifies an empty pronominal head in the second conjunct which can be spelled out as in:

(82) Anna, or she who is my daughter, cannot come to the meeting today.



(de Vries, 2006: 243)

Figure 2.4: The structure of *Anna, who is my daughter*

In example (82) the place of &:, the specifying coordination connection, is filled by *or*, and *she* fills the empty pronominal position $\varnothing_k$. This $\varnothing_k$ has the same case as the antecedent.

Similar to overt free-relatives, in the second conjunct an abstract raising of an NP is implemented. To explain the above claim, de Vries (2006) discusses the promotion theory that is developed by Vergnaud (1974), Kayne (1994) and Bianchi (1999). The benefit of the promotion theory, as he discusses it, is that it describes the connectivity effect between the antecedent and the gap in IRCs.

This theory has three hypotheses

1. Originally, the head noun is formed in the CP, the relative, and then it is raised.

2. The relative CP is the antecedent's determiner's complement.

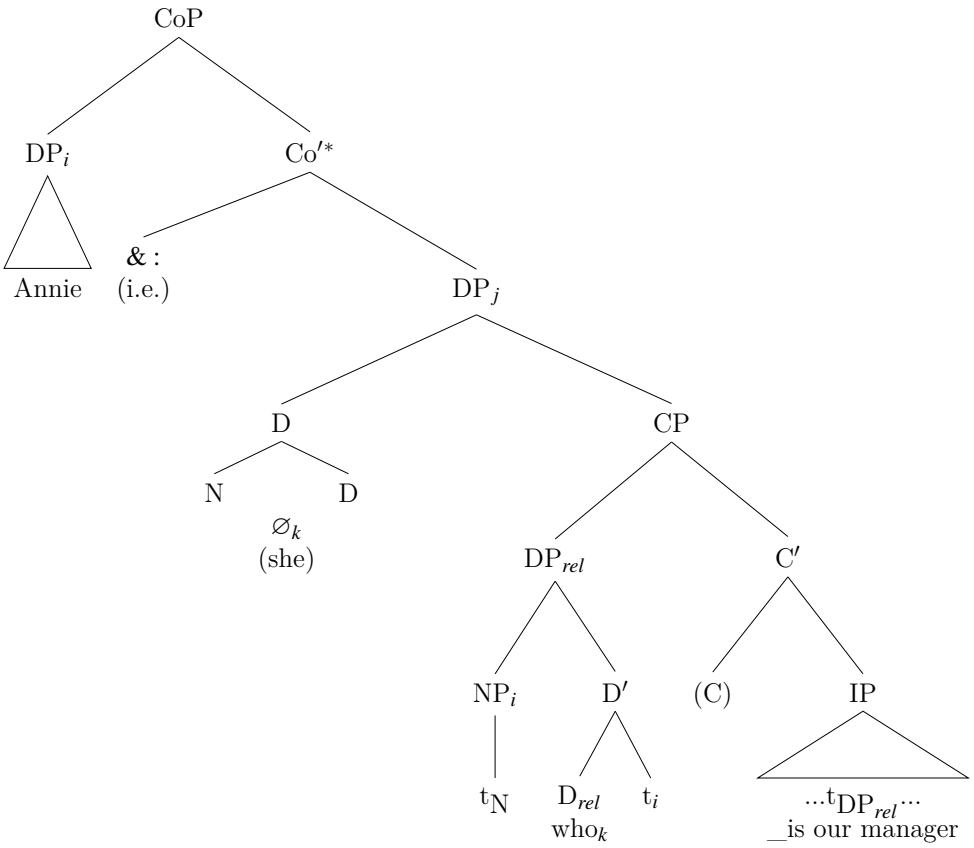
3. A relative pronoun is a determiner.

   In (83) the relative pronoun is represented by D_{rel}.

(83) $[_{DP}$ D $[_{CP}$ (C) ... $[_{DP-rel}$ D_{rel} NP] ...]]

According to de Vries, a thorough observation of the structure of free-relatives would suggest that SRCs are actually not true free-relatives, they rather are false-free-relatives, in that SRCs' pronoun doesn't contain an implied antecedent. The final intuition would be that SRCs have an empty pronominal head as they are a specifying conjunction, and they are false-free-relatives. Figure 2.5 shows details of the internal structure of an SRC together with its antecedent.

(84) Annie, (i.e. she) who is our manager, prepared the speech.



'Annie, (i.e. she) who is our manager'

Figure 2.5: Internal structure of an SRC together with its antecedent, de Vries (2006: 248)

In example (84), the derivation of the restrictive postnominal is the same as the derivation of the second at the lowest level and for the NP to check agreement

with D_{rel} it has to move to SpecDP_{rel}. This explains why a relative pronoun is a bound pronoun. In Figure 2.5, there is a co-indexing between ϕ and *who*. Additionally, the DP_{rel} moves to SpecCP for *wh*-feature-checking. D selects the CP. For reasons of agreement and abstract case checking, N moves to the empty external D. Finally, The complex [N+D] is an abstract personal pronoun which is ∅k.

Another reason why SRCs are not true free-relatives is that upward percolation between them is different. Upward percolation is not possible in true free-relatives as in (85a), while in false free-relatives and SRCs, it is, as in (85b).

(85) a. *I sent it to with whom you talked yesterday. (free-relative)

b. I sent it to Anna, with whom you talked yesterday. (SRC)

According to de Vries, unlike IRCs, SRCs do not fall into the scope of the antecedent's determiner or quantifier.

(86) a. All the students who took the exam passed.

b. ?All the students, who took the exam, passed.

The NP *All the students* is specified by the SRC. In the first conjunct of the specifying coordination phrase, the whole antecedent with its either determiner or specifier is a part of the antecedent. The structure is as the following:

$[_{CoP}\ [_{DP}$ D NP] &: $[_{DP}$ [SRC]]] (de Vries, 2006: 251)

The determiner cannot take scope over the SRC because it is embedded in the overt antecedent; hence, it cannot c-command the SRC. The reason SRCs can be deleted without the loss of acceptability, according to de Vries (2006), is that SRCs are complex appositions, which means that they are specifying conjuncts.

This makes them not essential to the meaning of the matrix clause and therefore IRCs can also be deleted without the loss of acceptability; however, it will affect the truth condition of the sentence.

Jackendoff (1977: 171) states that SRCs follow IRCs not vice versa and he provides the following example.

(87) a. The man that came to dinner, who was drunk, fainted.

b. *The man, who was drunk (,) that came to dinner fainted.

This supports the claim that de Vries previously made stating that it is within the first conjunct that IRCs are embedded in the maximal projection of the antecedent, DP. However, SRCs are in the second conjunct. The structure looks like the following:

[[$_{DP}$ D NP IRC] &: [$_{DP}$ [SRC]]] (de Vries, 2006: 251)

Coordination allows structures consisting of more than two conjuncts, and if SRCs are a type of coordination, then they should allow stacking as well. de Vries (2006: 252) states that there are no syntactic constraints that disallow stacking of SRCs; however, he assumes that pragmatic force might play a role in restricting some SRCs from forming stacking structures. On the contrary, Jackendoff (1977: 171) argues that SRCs, unlike IRCs, are not capable of stacking because stacking involves an asyndetic combination of units but SRCs are combined by overt coordinations. Example (88a) is considered grammatical by de Vries; however, it is considered ungrammatical by Jackendoff. Jackendoff finds examples such as example (88b) grammatical as the two SRCs are combined by an overt conjunct and together form one complex specification that adds to the antecedent. According to Jackendoff, the constructions such as the one in example (88c) is preferable

as we need two intonations to utter the sentence while we need three in example (88a).

(88) a. ?the man, who came to dinner, who hated Lox (SRC)

b. the man, who came to dinner and who hated lox (SRC)

c. the man who came to dinner and who hated lox (IRC) (de Vries, 2006: 253)

However, as mentioned in section 2.4, the proponents of the Syntactically Integrated Analyses have their own take on this subject and they believe that SRCs and IRCs are equally capable of stacking and IRCs can precede SRCs in stacking the same way SRCs can precede IRCs..

Further evidence that de Vries provides to show that SRCs have the same syntax as IRCs is that he states that extraposition of SRCs is possible since they are only apposition to an argument not an argument to the matrix predicate. The argument of the predicate of the matrix is not allowed to extrapose, as shown by the constrast in (89).

(89) a. I talked to the principal yesterday, who is my neighbor.

b. * I talked yesterday to the principle.

Nonetheless, Emonds (1979: 234) and Vergnaud (1974: 181) state that SRCs cannot be extraposed. The examples they provide contain a cause/effect reading, such as example (90), in which it is generally agreed that the sequence of the information is interrupted and the result doesn't stay meaningful; however, not all of the discourses of SRCs involve cause and effect. In other discourses, it generally is acceptable that SRCs, similar to IRCs, can be extraposed.

(90) ? The student came to my office, who you have been complaining about.

It's worth mentioning that de Vries states that a relative pronoun, be it in IRCs or SRCs, is a bound pronoun

(91) a. The principal$_a$ talked to the parent$_b$ who$_b$/*$_a$/*$_c$ brought a big cake. (IRC)

b. The principal$_a$ talked to Anna$_b$, who$_b$/*$_a$/*$_c$ brought a big cake. (SRC)

de Vries states that, in the case of an IRC, it is the raising analysis that applies; however, in the case of SRCs, the reason is that the link to the overt antecedent is not direct. The following structure makes this clearer

(92) [$_{CoP}$ DP$_i$ &: [DP$_j$ [N+D] [$_{CP}$ [$_{DP_{rel}}$ [$_{NP}$ t$_n$] D$_{rel}$ t$_{NP}$] (C) [$_{IP}$ ... t $_{DP_{rel}}$ ...]]]$_j$]
Mary$_i$ $\emptyset_k$ who$_k$ (de Vries, 2006: 258)

Similar to the case in IRCs, the implied antecedent of the free-relative, $\varnothing_k$, is syntactically linked to the relative pronoun *who*$_k$. The empty operator, $\varnothing_k$, denotes the overt antecedent DP$_i$, which is the first conjunct. Since the antecedent does not c-command the second conjunct, it cannot be established syntactically. Sells (1985) and Demirdache (1991) state that in SRCs the connection between the antecedent and the referring element, namely, $\varnothing_k$ must be stated as specification. Here, the right interpretation is made because of the notion of specifying coordination. If in the structure above $\varnothing_k$ in the matrix denoted unrelated entity DP$_x$ such as the *the principal*, in this case $_j$ wouldn't have the same referent as $_i$. For that particular reason, DP$_j$ is not interpreted as a specification of DP$_i$ because it would cause a semantic anomaly. We can infer that only when the empty element in SRCs are specifiers with the present antecedent, the SRC is considers a specifying conjunct.

This approach of analyzing the syntax and semantics of SRCs by de Vries contains some innovations and interesting points; however, I will not be using all

of his analyses in my data. As it is clear by now, de Vries does not account for all of the syntactic and semantic differences that I pointed out in section 2.3 between IRCs and SRCs that are essential for the main point of discussion of this thesis. I will later make a use of some of his syntactic analyses in my final semantic analysis of the data in chapters 5 and 6.

2.4.4 Arnold's (2004) and (2007) analysis of SRCs

Arnold (2007) states that the fact that there are more differences between SRCs and normal parentheticals than there are between IRCs and SRCs is evidence against the assumption that SRCs are parentheticals, which underlies the ROA. One of the major reasons to oppose considering SRCs parenthetical by ROA is that while SRCs can act like parentheticals in some cases, they are completely restrained in others, the following examples clarify these differences. Taking into account the relative position of SRCs/ parentheticals within the matrix clause, it becomes obvious that in the first set of examples, in (93), (94) and (95) the SRCs act the same as the parentheticals.

(93) The president has caused the country so many problems, as we all now admit.

(94) The president has caused the country so many problems, and we all now admit it.

(95) The president has caused the country so many problems, which we all now admit.

On the other hand, in the following set of examples, in (96a), (96b) and (96c), the SRCs do not act like the two parenthetical structures, as parentheticals can freely occur inside their host clause; nonetheless, SRCs cannot, because in the following examples they need to follow their antecedents exactly like IRCs do.

(96) a. The president has caused the country, as we all now admit, so many problems.

b. The president caused the country, and we all now admit it, so many problems.

c. *The president caused the country, which we all now admit, so many problems. (SRC)

The same is true for IRCs.

According to Arnold (2007: 287), another noticeable disparity between parenthetical structures and SRCs is that parentheticals can easily precede their antecedent, while SRCs cannot.

(97) As you might by now have noticed, she broke up with me.

(98) *Which you might by now have noticed, she broke up with me.

By considering these distinctive properties of SRCs, Arnold (2007: 284) argues that SRCs are closer to IRCs as they mostly occur directly after their antecedents similar to IRCs. Nonetheless, this is not the case with all SRCs for some of them can occur distanced from their antecedents.

According to the ROA, another paramount distinction between SRCs and IRCs is that SRCs don't form a constituent with their antecedents. Arnold (2007) presents many indications that refute this prediction of the ROA. The first relevant phenomenon is the possessive marking *'s*. This clitic occurs at the right edge of the NP it marks. The ROA predicts that an SRC would appear to the right of the clitic. Example (99a) shows that the SRC cannot appear in this position. Instead, the clitic must follow the SRC, see (99b). Consequently, the possessive marking provides positive evidence that the SRC forms a constituent together

with its antecedent. This is an adequate evidence to argue that SRCs form a constituent with their antecedents. The examples in (99) can clarify this

(99) a. *King Alphonso's, who ruined the party, mother left early.

b. King Alphonso – who ruined the party – 's mother left early. (Arnold, 2007: 284)

Arnold (2007: 285) also presents a movement-based argument to show that SRCs form a constituent together with their antecedents. In the following examples, the antecedents of the SRCs have been moved to the beginning of the clauses. SRCs must occur adjacent to the surface position of the antecedents, as in example (100b). SRCs cannot stand in the position from which the antecedents have been moved. The following examples are all from Arnold (2007: 286–287).

(100) a. Sandy, who I'm sure you remember, I see regularly.

b. *Sandy, I see, who I'm sure you remember, regularly. [Topicalisation]

(101) a. It is Sandy, who I'm sure you remember, that I see regularly.

b. *It is Sandy that I see, who I'm sure you remember, regularly. [Cleft-formation]

(102) a. Sandy, who I'm sure you remember, always seems helpful.

b. *Sandy always seems, who I'm sure you remember, helpful. [Raising]

(103) a. Sandy, who I'm sure you remember, was vilified by the press.

b. *Sandy, was vilified, who I'm sure you remember, by the press. [Passive]

(104) a. Sandy, who I'm sure you remember, is hard to please.

b. *Sandy is hard to please , who I'm sure you remember. [Tough-movement]

A further similarity between IRCs and SRCs poses a problem for the ROA analysis whereby the complement of N can extrapose over both IRCs and SRCs equally. The following examples are from Arnold (2007: 287)

1. IRC

   (105) Sam pointed out the two teachers [of Basque] [that you mentioned].

   (106) Sam pointed out the two teachers [that you mentioned] [of that strange and beautiful exotic language of North West Europe].

2. SRC

   (107) Sam claims to have found a proof of one of the most famous conjectures in the history of higher mathematics, which many believed could not exist.

   (108) Sam claims to have found a proof, which many believed could not exist, of one of the most famous conjectures in the history of higher mathematics.

In the above examples of SRCs, in (107) and (108), the SRC is not right-adjacent to its antecedent, it rather occurs inside it. ROA would need separate mechanisms to describe the parallel behaviour of IRCs and SRCs with respect to extraposition inside an NP.

Jackendoff (1977) only considers IRCs capable of stacking, which he takes as a support for an ROA analysis. However, the following grammatically correct sentences of SRCs that involve stacking show that this generalization is not valid. A discussion on this is presented in criterion 8, in section 2.3 of this chapter.

(109) John, who hands in all his work on time, who is on the student council, who even likes Math, for chrissake, will certainly get a good job. (Arnold, 2007: 288)

A further proof that IRCs and SRCs are very similar is that sometimes SRCs like IRCs can be extraposed. Although it is not possible for SRCs to be extraposed in example (111), it is possible to do so in example (113). The examples are from Arnold (2007: 288)

(110) Someone came who Mary knew. (IRC)

(111) *John came, who Mary knew. (SRC)

(112) I saw someone yesterday that I hadn't seen for years. (IRC)

(113) I saw my mother yesterday, who I hadn't seen for years. (SRC)

In this case, the SRC is distanced from its antecedent, and the sentence is grammatical. There are even cases where it is necessary for SRCs to be extraposed in the sentence. The examples are from Arnold (2007: 289):

(114) a. There were only thirteen senators present, which number was too few for a quorum.

b. *There were only thirteen, which number was too few for a quorum, senators present.

IRCs and SRCs share the same structural property with regard to restrictions on antecedent-anaphor relations. This is illustrated below, and it poses a further challenge for the ROA. Usually, a pronoun can precede its antecedent but not c-command it. This is also true if the pronoun or the antecedent occurs inside an RC. This restriction accounts for the examples in (116) and (118). Under an

ROA-analysis, it is impossible to apply this generalization to SRCs because they are not part of the same syntactic structure as the rest of the clause.

1. IRCs

   (115) The investors who met her$_i$ were all impressed by Anna's$_i$ invention.

   (116) *She$_i$ impressed the investors who met Anna$_i$.

2. SRCs

   (117) The investors, who met her$_i$, were all impressed by Anna's$_i$ invention.

   (118) *She impressed the investors, who met Anna.

As mentioned in section 2.3, in criterion 15, VP ellipsis work the same with both SRCs and IRCs. Material in one RC can be deleted if it is identical with material in another RC. This is another criticism against ROA for treating IRCs and SRCs differently while in this and many other cases, as mentioned earlier, they behave similarly.

(119) Someone that [supports the war] insulted Kim, who doesn't.

(120) Sandy, who [supports the war], insulted someone that doesn't.

(121) Sandy, who [supports the war], insulted Kim, who doesn't. (Arnold, 2007: 290)

Right node raising, RNR, can occur in SRCs in the same way it does in IRCs. RNR means that a raised element moves to the right of the main clause and this element is correlated with the gap positions in conjuncts of the coordinated sentence. The small squares □ indicate the location of the RNR gaps in the sentences.

(122) I met someone [who supports □], and Shano met someone [who opposes □], the referendum that is supposed to be held next month. (IRC)

(123) I met Sam, [who supports □], and Alan, [who opposes □], the referendum that is supposed to be held next month. (SRC)

Arnold (2007) shows that RNR is also possible if the two conjuncts differ with respect to the type of RC which contains the RNR gap. In the following examples, the sentences contain both an IRC and an SRC both containing a gap. The following examples are from Arnold (2007: 291)

(124) Sam knows someone [that buys □], and Kim knows Leslie, [who sells □], the pretentious garbage that some call Modern Art.

(125) Sam knows Sandy, [who buys □], and Kim knows someone [that sells □], the pretentious garbage that some call Modern Art

Arnold (2007: 291) disagrees with Ross's (1967) claim that unlike IRCs, SRCs can neither have quantified NPs as their antecedents nor can they have pronouns in them that are bound by an external quantified NP. Example (126) will contradict his claim as the antecedent of the SRC is a quantified NP, QNP. This is mentioned earlier in section 2.3: criterion 5.

(126) Two tutors will be assigned to each student, who are then responsible for getting his papers to the Dean's office on time. (Arnold, 2007:292)

Furthermore, it is not the universal quantifier per se that can be the antecedent of SRCs, there are many other varieties of quantifiers that can fill in this position.

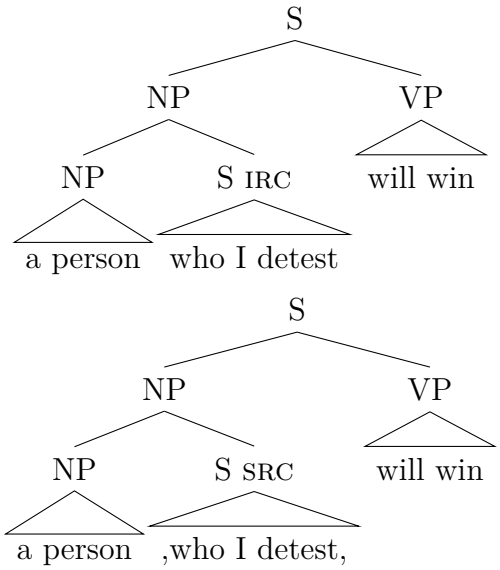
(127) Some properly qualified linguist, who would have been taught phonetics as part of their training, would have got that right. (Existential Quantifier)

(128) No properly qualified linguist, who would have been taught phonetics as part of her training, would have made that mistake (Arnold, 2007: 292). (Negative Quantifier)

Arnold captures the syntactic similarity between IRCs and SRCs in the following structures. We can see that they are both analyzed in the same way.

(129) a. A person who I detested will win.

b. A person, who I detested, will win.



(Arnold, 2007: 293)

Figure 2.6: An HPSG tree of *A person(,) who I detested (,) will win*

The syntactically integrated analysis of SRCs in Figure 2.6 is similar to that of IRCs'. This captures the observed similarities between the two constructions.

Eventually, it becomes clear that the ROA and NROA, as well as de Vries's analysis and Arnold's analyses result in the observation that the arguments in favor of non-integratedness of SRCs are semantic rather than syntactic. Arnold's

syntactic analyses of SRCs and his arguments against the non-integratedness of SRCs are the most valid ones to me and I will be depending on them and using them in analyzing my data throughout.

2.5 SRCs are semantically integrated

In this section, I will argue that SRCs are also semantically more integrated than has often been assumed. This means that there are both syntactic and semantic reasons for treating SRCs as part of the matrix clause and not as orphans. The arguments for the semantic connection between the matrix clause and the SRC have been presented in AnderBois et al. (2015), which I will summarize in this section.

AnderBois et al. (2015) define the at-issue contents as excluding all presuppositions and possible implicatures. They propose that the main two differences between at-issue and non-at-issue contents are (i) only the at-issue content can resolve questions under discussions (ii) the non-at-issue contents can outscope operators that occur in the at-issue contents, such as negation and modals.

To capture these differences between at-issue and non-at-issue contents, Potts (2005) proposes *multidimensional semantic representations*, i.e., semantic representations that are 'sets or tuples of whatever sort of denotations one assigns to (subparts of) simple at-issue assertions' (AnderBois et al., 2015: 94). The proposed multidimensionality goes hand in hand with the assumption of a complete separation between at-issue and non-at-issue contents.

As it is common to consider intra- and inter-sentential anaphora part of semantics proper, AnderBois et al. (2015) argue that the idea of multidimensionality needs to be amended. Examples such as (130) show that an element introduced

in a constituent that is interpreted as non-at-issue – such as *the man* inside the SRC – can serve as the antecedent of a pronoun *him* in the main clause, which is interpreted as at-issue. In (130) I underline the antecedent and I put the discourse pronoun in small caps.

(130) Anna, who helped the man, took HIM to a shelter.

Based on such data, AnderBois et al. (2015) argue that, instead of considering an SRC semantically separate from the main clause containing it, there should be an approach that can account for possible connections between them. Consequently, AnderBois et al. reject a Potts-style multidimensional semantics and return to what they call a *unidimensional semantics*. Rather than using two dimensions of semantic representations, they locate the difference between at-issue and non-at-issue contents in how these are integrated into the discourse.

AnderBois et al. (2015) show that there is a systematic interaction between an SRC's content and the at-issue content that goes in both ways. They illustrate this with three phenomena: presuppositions, anaphora, and ellipsis. I will summarize their discussion briefly, looking at examples of each of these phenomena.

First, AnderBois et al. (2015: 97) show that a presupposition trigger and its satisfying antecedent can be split over an SRC and the main clause in either way. In (131a) the trigger – which is marked in small caps – is included in the SRC and its – underlined – antecedent occurs in the matrix clause. In (131b) we find the opposite constellation.

(131) a. John, who saw Mary, saw SUSAN TOO.

b. John saw Mary, who saw HIM TOO.

Second, the bi-directional boundary crossing is also applicable in interpreting

some types of anaphora. This is illustrated in (132).[3] In (132a), the pronoun in the apposition finds its antecedent in the matrix clause. In (132b), the pronoun *them* occurs in the matrix clause, whereas its antecedent is in the SRC.

(132) a. Jones gave detailed feedback on the final paper of each student, which THEY had submitted only a few days ago.

b. Jones, who graded each student's final paper, gave THEM detailed feedback.

Third, the same phenomena, bi-directionality of boundary-crossing, is found in ellipsis, as well – both in NP-ellipsis and VP-ellipsis. Example (133a) is an NP-ellipsis, and example (134a) is an VP-ellipsis. Both examples are from AnderBois et al. (2015: 99). In (133a) and (134a), the antecedents are in the SRCs and accessible for the ellipsis in the main clause. In (133b), which is an NP-ellipsis, and (134b), which is an VP-ellipsis, the antecedents are in the main clauses and are equally accessible to ellipsis inside the SRCs.

(133) a. The 1980's were dominated by the Lakers, who won five championships, and by the Boston Celtics, who won THREE. (COCA)

b. 'When we've got four or five guys hitting threes,' said guard Pat Bradley, who made THREE, … (COCA)

(134) a. So Lalonde, who was the one person who could deliver Trudeau , DID. (COCA)

b. Mr. Gore at first believed the president, and even defended him to Tipper and his daughters, who DID NOT. (COCA)

[3]Example (132b) is taken from (AnderBois et al., 2015: 98). Example (132a) is my own variation on this example, as the example corresponding to (132b) in AnderBois et al. (2015) does not contain a SRC.

From what we have seen so far, it becomes clear that the three processes (anaphora, presupposition, and ellipsis) don't differentiate between the at-issue and SRCs contents. Hence, AnderBois et al. (2015) conclude that the at-issue content and the SRC-content should be represented within the same semantic dimension, i.e., that there should be a *unidimensional* analysis. In the following, I will sketch their approach.

While the three processes look different, they are, semantically, all instances of anaphoric relations. In all cases, there is a trigger (a pronoun, a word like *too*, or an ellipsis) that requires an already introduced element for its interpretation. The similarity of the three phenomena is made transparent in the use of the small caps and the underlining in the examples above. Frameworks of dynamic semantics, such as *Discourse Representation Theory* (DRT, Kamp & Reyle 1993) or *Dynamic Predicate Logic* (DPL, Groenendijk & Stokhof 1991) are designed to capture discourse anaphoric relations. Since Geurts (1999), it is well established that presuppositions can be considered a discourse anaphoric relation and, consequently, analyzed in a way analogous to pronouns within a dynamic semantic framework.

AnderBois et al. (2015) formulate their analysis in a version of DPL, which has the advantage over DRT that it largely uses the syntax of classical predicate logic and just assigns a dynamic interpretation to its expressions. The choice of a dynamic semantic framework is a first step for capturing the reported data on the semantic interaction between SRCs and their matrix clauses.

The second component of AnderBois et al.'s (2015) analysis consists in their modelling of the difference between at-issue and non-at-issue contents. These differ in the way in which they enter the commitments of the participants in a conversation. These commitments are called the *Common Ground* CG or *Context*

Set, CS. Not-at-issue content is entered into the CS immediately. In contrast to this, the at-issue content will first be *proposed* or *put on the table*, which means that it can potentially be rejected in a conversation. Only when accepted will it be put into the CS.

To model this, AnderBois et al. (2015) use two propositional variables, p^{cs} for the context set, and p for the proposal, i.e. the at-issue content. In the interpretation of an SRC, its content will be intersected with p^{cs} as soon as it occurs. At-issue contributions of a clause will first be integrated into the proposal and only be intersected with p^{cs} once the sentence is finished. Each predicate is parametrized with a propositional variable. This makes it possible to interpret SRC-internal predicates with respect to CS and predicates from the matrix clause with respect to the proposal. Individual variables, as used for the discourse referents of proper names, for example, are not parametrized in such a way. This allows for a constant interpretation of such variables across the proposal and the CS.

I will illustrate this with the example in (135), in which I slightly adjust AnderBois et al.'s (2015) notation to the one I will be using later. I will, however, use their notation $[x]$ to indicate the introduction of a discourse referent.

(135) John, who nearly killed a woman with his car, visited her in hospital. (AnderBois et al., 2015: 110)

a. New proposal: $[p] \wedge p \subseteq p^{cs} \wedge$

b. At issue: $[x] \wedge x = \mathbf{john}' \wedge$

c. Appositive: $[y] \wedge \mathbf{woman}'(p^{cs}, y) \wedge \mathbf{nearly\text{-}kill}'(p^{cs}, x, y) \wedge$

d. At issue: $\mathbf{visit}'(p, x, y) \wedge$

e. Proposal accepted: $[p^{cs}] \wedge p^{cs} = p$

In (135a), a new proposal, p, is introduced and required to be compatible with the CS. The word *John* triggers the introduction of another discourse referent, x, that is assigned the referent of the constant **john′**, see (135b). The interpretion of the SRC is given in (135c). There, a discourse referent, y, is introduced. It is used as an argument of the constant **woman′**. Note that this constant is intepreted with respect to the CS. It is also required that the relation **nearly-kill′** holds of x and y in p^{cs}.

The material from the matrix clause, the meaning of *visited her in hospital*, is interpreted with respect to the proposal p. As there is just one semantic representation instead of using two dimensions, the arguments of the predicate **visit′** can be discourse referents that are introduced in the matrix clause, such as x, or in the SRC, such as y. Finally, accepting a proposal means that the proposal is turned into the CS. This is done in (135e).

The essential ingredient of AnderBois et al.'s (2015) analysis is that the proposal and the CS are treated just as ordinary variables in a dynamic system. This makes it possible to interpret the sentence in an incremental way, instead of having to interpret the matrix clause and the SRC separately. At the same time, it can still be captured that the SRC content cannot be in the scope of quantifiers or negation in the matrix clause.

Henderson's (2014) suggests a modification to this analysis, which is based on the observation that not all SRCs are necessarily interpreted as non-at-issue. He shows that final SRCs can be at-issue in the sense that they can be used to answer a question under discussion and they can be rejected in a direct reply, see (136).

(136) a. A: Who did you invite for the colloquium and where are they from?
B: I invited Mary, who is an American. (Henderson, 2014)

b. A: We should invite Mary, who is an American.

B: No (, she is Australian).

B′: No (, we should invite Susan)

In (136a), there is a multiple question and the clause-final SRC can be used to answer one part of the question. Example (136b) shows that a speaker can react with *no* to the SRC's content as well as to the content of the matrix clause. This suggests that clause-final SRCs can be interpreted as non-at-issue or as at-issue.[4]

The data presented in this section show that there is an interaction between a matrix clause and an SRC contained in it with respect to discourse anaphoric relations. Such an interaction is not expected if the SRC is interpreted independent of the matrix clause. Instead, the SRC needs to be taken into consideration semantically in its surface position. Nonetheless, the kind of semantic interaction is restricted to relations which we typically find in discourses and cannot lead to an interpretation of the SRC in the scope of an operator from the matrix clause. This is modelled in AnderBois et al. (2015) by using a dynamic semantic framework and explicit propositional variables for the at-issue content and the background commitments in a conversation.

In my analysis of SRCs, I will not incorporate the ideas from AnderBois et al. (2015) explicitly. However, my proposal will be compatible with an immediate integration of the content of an SRC.[5] I will briefly return to Henderson's observation on the possible at-issue status of final SRCs, though.

[4]The main point in Henderson's (2014) is that different types of appositions differ with respect to whether they can be interpreted as at issue or not.

[5]See Arnold et al. (2018) for an integration of AnderBois et al. (2015) into the syntactic and semantic framework that I will be using.

2.6 Summary

To sum up, this chapter gives a summary of the previous literature and analyses of RCs in general and SRCs in particular. A thorough comparison between SRCs and IRCs shows that SRCs are different from IRCs in their prosody, the type of antecedents they take, and their semantics and syntax. However, these differences do not mean that they should be treated completely differently and should receive different analyses.

In analyzing SRCs, de Vries (2006) makes some interesting claims by considering SRCs false-free-relatives and complex appositions. By the latter he implies that SRCs are specifying conjuncts. This view of de Vries will not be reflected in my analysis.

In analysing SRCs, Arnold (2004) and (2007) states that SRCs are not parentheticals because, firstly, parentheticals can freely occur inside their host clause; nonetheless, SRCs cannot. Secondly, parentheticals can easily precede their antecedents, while SRCs cannot. Arnold argues that SRCs are not very different from IRCs because of the followings: SRCs and IRCs both occur directly after their antecedents. Similar to IRCs, SRCs form a constituent with their antecedents. Complements of Ns can extrapose over both IRCs and SRCs equally. Similar to IRCs, SRCs are capable of stacking. SRCs like IRCs can be extraposed. VP ellipses work the same with both SRCs and IRCs. SRCs also can have quantified NPs as antecedents.

In analyzing SRCs, AnderBois et al. (2015) state that SRCs are not at-issue. They believe that at-issue contents and SRCs are both unidimensional. They also claim that there is a systematic interaction between SRCs and at-issue contents that goes in both ways, because these three phenomena: presuppositions,

anaphora, and ellipsis treat both constructions the same.

They add, at-issue and SRCs are different because, first, SRCs are different in the way they enter *Context Set*, CS. Second, at-issue meanings are proposals and SRCs contents are imposed on the CSs.

Concerning the way SRCs answer questions, AnderBois et al. observed the following: SRCs cannot answer questions under discussion, QUD, since they are not semantically at-issue. On the contrary, the at-issue contents are capable of that. Concerning the types of the answer the addressee can give to an SRC, when SRCs are final, it is possible to deny it. However, when the SRC is not in the final position, it takes extra linguistic effort to deny it or comment on it.

Chapter 3

The Framework

3.1 Introduction

In this chapter, the frameworks, HPSG and LRS, which model the syntax-semantics interface in the thesis, are presented. The two frameworks are presented separately. I present HPSG in the form of a synthesis of various major versions of the framework, mostly form Pollard & Sag (1994a), Sag (1997), and with some references to Ginzburg & Sag (2000). The version of HPSG used in this thesis is largely identical to that of Walker (2017). In 3.2.1 a short summary of the formal background of HPSG is given. This is done by mainly discussing HPSG's two mechanisms, namely the *signature* and the *theory*. I will introduce the main aspects of the grammar I am using in Section 3.2.2.

I present a summary of Pollard & Sag (1994a)'s Unbounded Dependency Constructions, UDCs, in 3.3.1. Island constraints are discussed in 3.3.2. Following this, there is a conspectus subsection, 3.3, on Pollard & Sag's (1994a) analysis of English RCs, shedding some light on how they treat traces and how they account for them inside RCs. I present a general summary of RCs in Constructional HPSG

in 3.3.4. Finally, I show Arnold's (2004) constructional analyses of SRCs in 3.3.5.

In the LRS section, I present the framework's background in 3.4.1, and I illustrate the theory of the framework in 3.4.2. In 3.4.3, I give the implementations of the framework on an example sentence, showing the gradual combination of its part in an AVM.

3.2 HPSG

3.2.1 Background to HPSG theory

HPSG was originally influenced by previous frameworks and grammar theories such as Government and binding theory of Chomsky (1981), Generalized Phrase Structure Grammar GPSG, Lexical Function Grammar and also Categorical Grammar. The theory of HPSG was first introduced in Pollard & Sag (1987), then it was gradually developed and amended in Pollard & Sag (1994a). Later, Sag slightly changed some theoretical aspects in Sag (1997). Finally, some new mechanisms were introduced by Ginzburg & Sag (2000).

First, Pollard & Sag (1987) introduced the fundamentals of HPSG, which essentially was a cohesive theory of natural language syntax and semantics. They reflected on some linguistic facts, such as sub-categorization, the distinction between complements and adjuncts, constituent ordering and inflection. Pollard & Sag (1994a) intended to account for more linguistic facts, such as long distance dependencies and syntactic constraints on the relationship of pronouns and their antecedents, etc.

HPSG has two mechanisms:

1. The *signature*: here, the main ontology of a language is presented. The

existence of the kinds of objects and their properties are displayed. The signature consists of a type hierarchy and conditions of appropriateness; these condition outlines the appropriate features of the types and the appropriate type of value for each feature.

2. The *theory*: according to Levine & Meurers (2006: 244), HPSG theory is a set of constraints which distinguishes a subset of grammatical objects from the objects that are matching the *signature*. Linguistic objects are permissible in the theory when they satisfy all of the descriptions in the theory.

In HPSG, a description language attribute value matrix representation, AVM, is used to denote sets of objects. The descriptions are:

1. Type descriptions: these distinguish specific types of objects, for example, objects of type *words*, from the other types of objects.

2. Attribute-Value Pairs: these define objects with a specific property, for example: CASE *accusative*.

3. Tags: these identify token identity. They are given in a form of re-occurring boxed numbers or letters within the same description, such as [1] or [B].

In HPSG, utterances are modelled as feature structures of type *sign*. All feature structures of the type *sign* have the features PHONOLOGY and SYNSEM defined on them. The SYNSEM value bares both semantic and syntactic information of the *sign*. Within the SYNSEM value of a *sign*, there is much more detailed and specified information that this feature structure is able to show.

The type *sign* has two immediate subtypes: *word* and *phrase*, which model words and phrases respectively. The lexicon is perceived as a set of constraints on

feature structures of type *word*. These constraints are referred to as *lexical entries*. Analogously, the inventory of grammatical constructions is a set of constraints on feature structures of type *phrase* – referred to as *(immediate dominance) schemas*. I introduce the schemas relevant for my thesis in section 3.2.2. The tree in Figure 3.1, which is provided by Walker (2017: 97), shows the type hierarchy and the appropriateness conditions of the type *sign*.

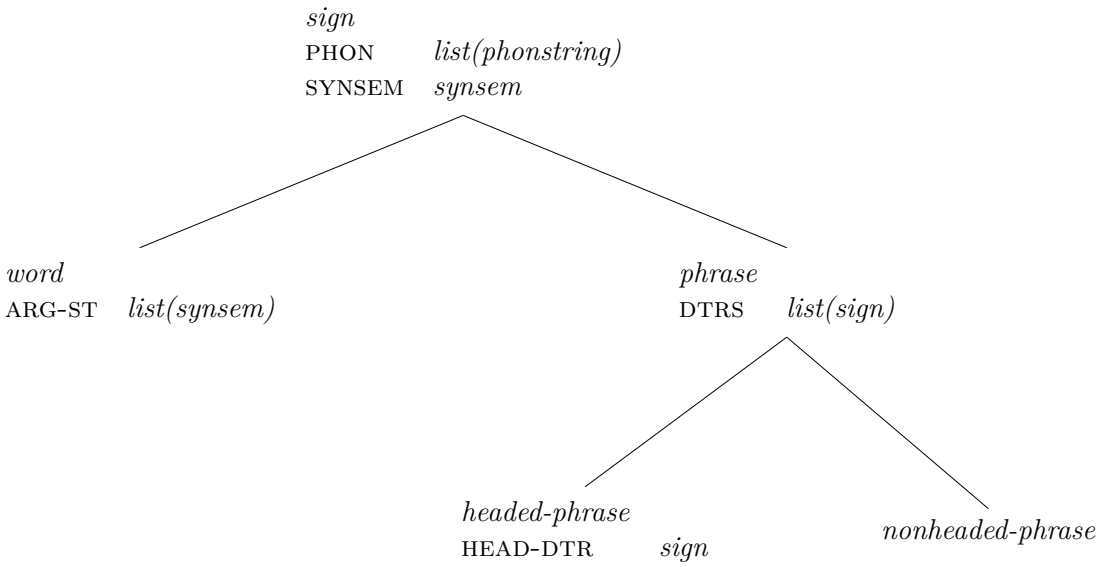


Figure 3.1: The type hierarchy and the appropriateness conditions below the type *sign*

We can now introduce two important grammatical principles, constraints, on phrases. These constraints are:

(1) Head Feature Principle: .

$$\textit{headed-phrase} \longrightarrow \begin{bmatrix} \text{SS}|\text{LOC}|\text{CAT}|\text{HEAD } \boxed{1} \\ \text{HEAD-DTR } \begin{bmatrix} \text{SS}|\text{LOC}|\text{CAT}|\text{HEAD } \boxed{1} \end{bmatrix} \end{bmatrix}$$

The Head Feature Principle guarantees that a phrase and its head daughter have the same value for the feature HEAD.

(2) The Head Daughter Principle

$$\textit{headed-phrase} \longrightarrow \begin{bmatrix} \text{DTRS} & \left\langle \ldots, \boxed{1}, \ldots \right\rangle \\ \text{HEAD-DTR} & \boxed{1} \end{bmatrix}$$

The principle in (2) guarantees that the head daughter is among the overall daughters.

Now that the main mechanisms of HPSG and two major constraints on phrases have been discussed, it is time to present the HPSG schemas that are used in the analyses of the structures throughout this thesis.

3.2.2 The schemas

In this section, I present five schemas, *Head-subject Schema*, *Head-specifier Schema*, *Head-complement Schema*, *Head-functor Schema* and *Head-filler Schema*. The schemas are adopted from van Eynde (2006) and Walker (2017). Changes are made where needed to make the schemas inline with the analysis I use. For instance, in all schemas the MARKING value is added. This value is the same on the mother and the head daughter in all of the schemas except for *head-functor* schema. In this section, I define the schemas without much explanation. I will elaborate on those schemas that are important in the rest of the thesis in more detail in the following sections.

1. The Head-subject Schema

   This schema licenses saturated phrases with a head daughter and complement daughter, non-head daughter, in which the SYNSEM value of the non-head daughter is token-identical with the element that is in the VAL lists of the head daughter. The Head-subject Schema is defined in Figure 3.2.

$$\begin{bmatrix} \text{SS}|\text{LOC}|\text{CAT} & \begin{bmatrix} \text{MARKING } \boxed{3} \\ \text{VAL} & \begin{bmatrix} \text{SUBJ } \langle\rangle \\ \text{SPR } \boxed{\text{A}} \\ \text{COMPS } \boxed{\text{B}} \end{bmatrix} \end{bmatrix} \\ \text{DTRS } \left\langle \left[\text{SS } \boxed{1}\right], \boxed{2} \right\rangle & \\ \text{HD-DTR } \boxed{2} & \begin{bmatrix} \text{SS}|\text{LOC}|\text{CAT} & \begin{bmatrix} \text{MARKING } \boxed{3} \\ \text{VAL} & \begin{bmatrix} \text{SUBJ } \langle \boxed{1} \rangle \\ \text{SPR } \boxed{\text{A}} \langle\rangle \\ \text{COMPS } \boxed{\text{B}} \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 3.2: The description of *Head-subject* schema

2. The Head-specifier Schema

   This schema licenses phrases which have a head daughter and only one non-head-daughter.

$$\begin{bmatrix} \text{SS}|\text{LOC}|\text{CAT} & \begin{bmatrix} \text{MARKING } \boxed{3} \\ \text{VAL} & \begin{bmatrix} \text{SUBJ } \boxed{\text{A}} \\ \text{SPR } \langle\rangle \\ \text{COMPS} \boxed{\text{B}} \end{bmatrix} \end{bmatrix} \\ \text{DTRS } \left\langle \left[\text{SS } \boxed{1}\right], \boxed{2} \right\rangle & \\ \text{HD-DTR } \boxed{2} & \begin{bmatrix} \text{SS}|\text{LOC}|\text{CAT} & \begin{bmatrix} \text{MARKING } \boxed{3} \\ \text{VAL} & \begin{bmatrix} \text{SUBJ } \boxed{\text{A}} \\ \text{SPR } \langle \boxed{1} \rangle \\ \text{COMPS } \boxed{\text{B}} \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 3.3: The description of *Head-specifier* schema

3. The Head-complement Schema

   This schema licenses phrases in which all the complements are realized as sisters to the lexical head. The structure of the phrases this schema licenses will be such as in Figure 3.4.

4. The Head-functor Schema

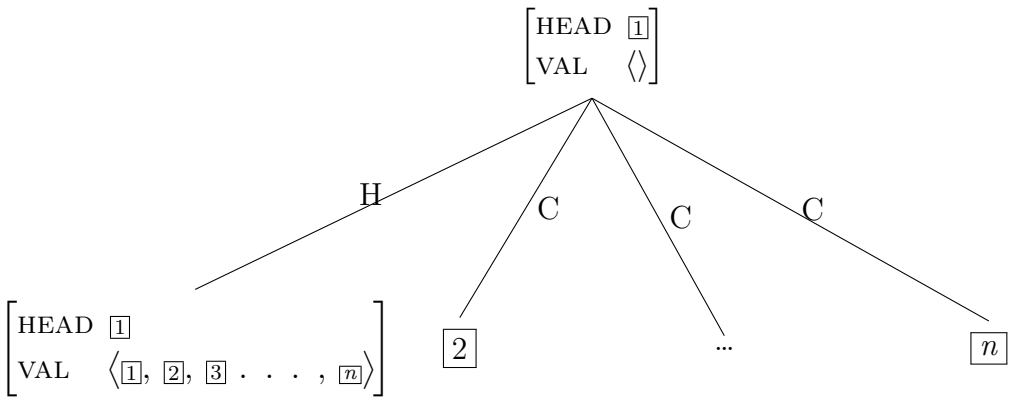


Figure 3.4: The structure of phrases licensed by *The Head-complement* Schema

$$\left[\begin{array}{l} \text{SS}|\text{LOC}|\text{CAT}\left[\begin{array}{ll}\text{MARKING} & \boxed{3}\\ \text{VAL} & \left[\begin{array}{ll}\text{SUBJ} & \boxed{\text{A}}\\ \text{SPR} & \boxed{\text{B}}\\ \text{COMPS} & \langle\rangle\end{array}\right]\end{array}\right] \\ \text{DTRS}\left\langle \boxed{1}, \left[\text{SS } \boxed{2}\right],\ldots,\left[\text{SS } \boxed{\text{n}}\right]\right\rangle \\ \text{HD-DTR } \boxed{1}\left[\text{SS}|\text{LOC}|\text{CAT}\left[\begin{array}{ll}\text{MARKING} & \boxed{3}\\ \text{VAL} & \left[\begin{array}{l}\text{SUBJ } \boxed{\text{A}}\\ \text{SPR } \boxed{\text{B}}\\ \text{COMPS } \left\langle \boxed{2},\ldots,\boxed{\text{n}}\right\rangle\end{array}\right]\end{array}\right]\right] \end{array}\right]$$

Figure 3.5: The description of *Head-complement* schema

This schema is used in van Eynde (1998, 2003a,b, 2006) and Walker (2017). Phrases of type *head-functor-phrase* have SELECT feature as their head features. The feature SELECT replaces Pollard & Sag's (1994a) features, MOD and SPEC. In describing SELECT feature, van Eynde (2006: 164) states 'In contrast to the valence features, which capture the constraints, which a head imposes on its dependents, the SELECT feature captures the constraints which a non-head daughter imposes on its head sister.'

5. The Head-filler Schema

This schema is used to introduce a constituent into the pre-nuclear position

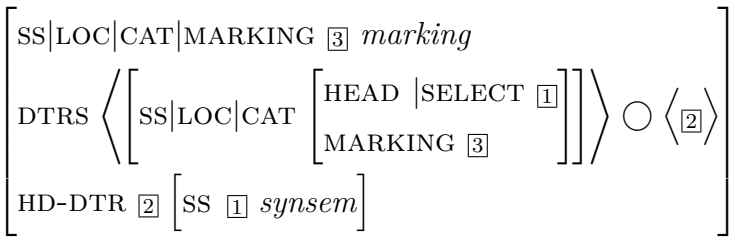


Figure 3.6: The description of *Head-functor* schema

of a sentence. This constituent will be the non-head daughter; the head daughter needs to be a finite clause. The schema defines how the constituent in pre-nuclear position is connected to this clause. The details of the schema will be discussed and motivated in Section 3.3.1.

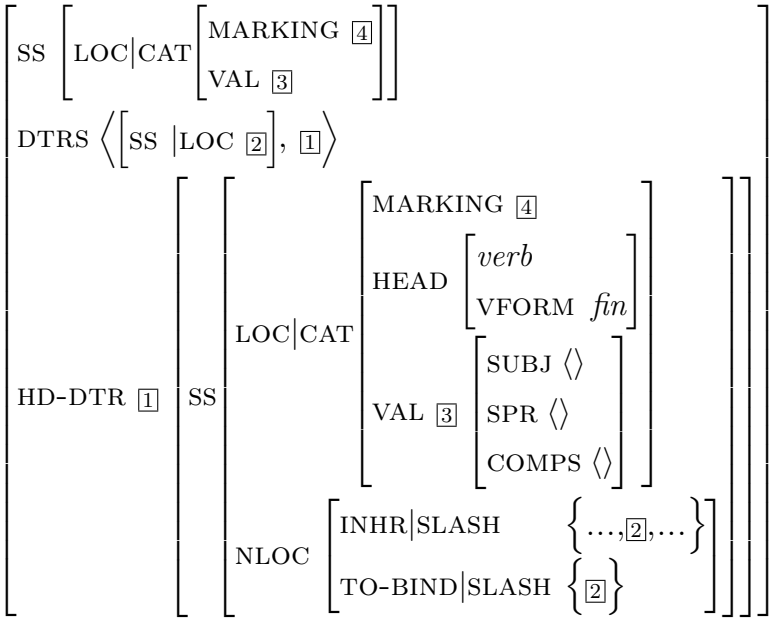


Figure 3.7: The description of *head-filler* schema

3.3 Relative clauses in HPSG

In this section, I will start by introducing Unbounded Dependency Constructions in classical HPSG in section 3.3.1. I illustrate the island constraints that I need

in my data analysis, especially in chapter 7, in 3.3.2. In 3.3.3 I present integrated relative clauses in HPSG, in Pollard & Sag (1994a). I discuss the function of the antecedent in **??**. Later, I briefly present RCs in constructional HPSG in 3.3.4. Finally, I summarize Arnold's (2004) constructional analysis of SRCs in 3.3.5.

3.3.1 Unbounded dependency constructions in classical HPSG

RCs are instances of *unbounded dependency constructions*, UDCs, a term that was introduced in Gazdar (1981). In this section, I motivate and introduce the elements of the HPSG analysis of UDCs. UDCs are grouped into two classes, as described in the following.

In the first class, Gazdar includes constructions that have overt constituents in non-argument position, this can be a topic of any expressions containing *wh*-phrase that are the fillers of the gaps. This is also called the prenuclear position. The phenomena that fit into this class include: topicalization, *wh*-RCs, *wh*-questions (an example of this is given in (3)), *it*-clefts and pseudo-clefts. This type is called *strong* UDCs.

(3) I wonder who Sandy loves. (*wh*-question)

In the second class, there is no overt constituent in a non-argument position. Instead, there is a constituent that is co-referential with the gap. The phenomena that belong to this class include: purpose infinitives, *tough*-constructions and non-*wh*-RC. This type is called *weak* UDCs, the relationship between the argument and the gap is considered co-indexing. In example (4) below, the subscript, "$_1$", shows this co-indexing.

(4) This is [the topic]$_1$ Sandy discussed $\_\__1$. (non-*wh*-RC)

Gazdar (1981) and Pollard & Sag (1994a) state that an UDC is unbounded because the dependency may extend through an arbitrary number of embedded clauses. There is a dependency, because, in cases of strong UDCs, there is a syntactic relation between the gap and its filler. For instance, example (5) is grammatical while (6) is not.

(5) Him, I like.

(6) *Anna, I relied.

Example (5) is grammatical because the gap has the same syntactic category and function as the filler. In example (5) the gap is in the object position, in accusative case, and the filler is in accusative form. On the contrary, in example (6) the gap is in a prepositional phrase's position as imposed by the verb *rely*, but the filler is not a prepositional phrase, which makes the sentence ungrammatical.

In HPSG, fillers are introduced into the structure by the *head-filler* schema, the principle of which is presented in schema 5. The filler-gap dependency means that all the LOCAL information, herby LOC, such as HEAD feature, sub-categorization and CONTENT are percolated from the gap position to the filler. This is illustrated by example (7), for which I show the detailed structure in Figure 3.8.

(7) Him, she said they know I like.

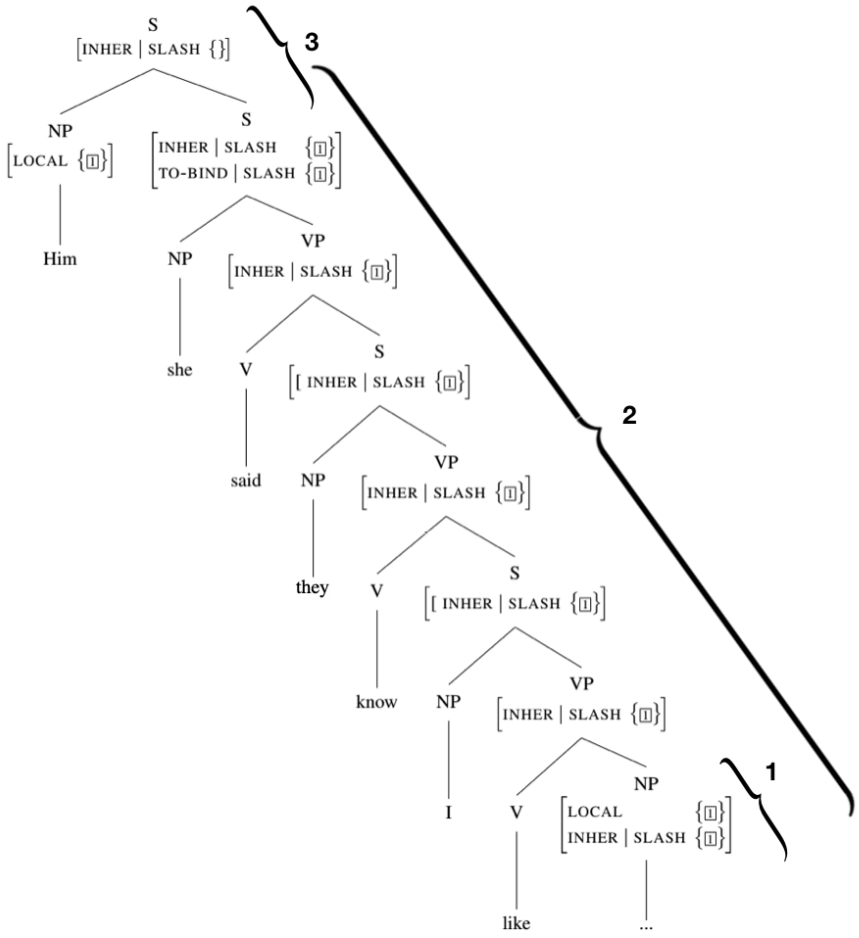


Figure 3.8: An HPSG tree of the sentence *Him, she said they know I like.*

Unbounded dependencies consist of three parts, the top, the middle and the bottom part. The final node, the bottom part, is where the dependency introduced by a *sign* whose NLOC, feature has a *nonempty* value. The middle part, numbered 2 in the diagram, is the structural distance between the gap and the filler. Here, information on the gap (its LOC value) is passed on from the daughters to the mother in each local tree. The NLOC Feature Principle is the mechanism responsible for this percolation. The top part, which is the part numbered 3 in Figure 3.8, is where the dependency resolves. The SLASH value, which has per-

colated from the bottom part is bound off in this part. This is made through the identification of the LOC feature of both the filler in the top part and the gap in the bottom part. It is in this part that *strong* UDCs and *weak* UDCs are distinguished, as only strong UDCs have an overt filler.

In HPSG, a trace is considered as a special lexical item. Pollard & Sag use traces to model gaps throughout their (1994a)'s HPSG except for chapter eight of their book. The trace's phonology is the empty list $\langle\rangle$ and it has no LOC feature. Its *nonempty* value is the SLASH value which is identical with the LOC feature of the filler. Its attribute value matrix AVM is as the following (preliminary version) and it has a non-empty INHR|SLASH value, which contains an element which is identical with the LOC value of the trace, a simplified lexical entry of a trace is given in Figure 3.9.

$$\begin{bmatrix} \text{PHON} & \langle\rangle \\ \text{SS} & \begin{bmatrix} \text{LOC} & \boxed{1} \\ \text{NLOC} & \begin{bmatrix} \text{SLASH} & \left\{ \boxed{1} \right\} \\ \text{QUE} & \{\} \\ \text{REL} & \{\} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 3.9: The description of a trace in HPSG

There have been alternative proposals to the analysis of the bottom of an UDC. Most prominently Sag's (1997) proposes a *traceless* analysis. In his analysis, the relevant INHER|SLASH specification is not introduced through a separate constituent, but by the lexical head selecting the gap, such as the verb in the case of a complement extraction. As I will be mainly concerned with the top part of UDCs in my thesis, I remain agnostic as to whether the bottom should be

analyzed by means of a trace or through the gap-selecting element.[1]

The NLOC feature principle is the mechanism with which the gradual percolation of the information of the UDC is carried out. Pollard & Sag's (1994a)'s simplified definition for this mechanism is: 'The value of each non-local feature on the phrasal sign is the union of the value on the daughters' (Pollard & Sag, 1994a: 162).

The *Head-filler* schema, given above in section 3.2.2 in schema 5, is the HPSG rendering of Gazdar (1981). Here is how the *Head-filler* schema is gradually introduced in two versions in Pollard & Sag (1994a). The description in Figure 3.10 is the preliminary version of *Head-filler* schema, schema 5, by Pollard & Sag (1994a: 164)

$$\left[\begin{array}{l}\text{DAUGHTERS} \left[\begin{array}{l}\text{FILLER-DTR}\left|\text{SS}\left[\begin{array}{l}\text{LOC } \boxed{1} \\ \text{NLOC } \langle\rangle\end{array}\right]\right. \\ \text{HEAD-DTR}\left|\text{SS}\left[\begin{array}{l}\text{LOC}\left|\text{CAT}\left[\begin{array}{ll}\text{HEAD} & \left[\begin{array}{l}\textit{verb} \\ \text{VFORM } \textit{fin}\end{array}\right] \\ \text{VAL} & \left[\begin{array}{l}\text{SUBJ } \langle\rangle \\ \text{SPEC}\langle\rangle \\ \text{COMPS } \langle\rangle\end{array}\right]\end{array}\right]\right. \\ \text{NLOC}\left|\text{TO-BIND} \mid \text{SLASH}\left\{\ldots, \boxed{1}, \ldots\right\}\right.\end{array}\right]\end{array}\right]\end{array}\right]$$

Figure 3.10: The description of *head-filler* schema (preliminary version)

When we have a closer look at the topicalization type of UDCs, a problem arises. The problem occurs when there is no rule to prevent the SLASH feature

[1]My indifference on this issue is also motivated by the fact that the two languages I will look at in this thesis, English and Sōrānī Kurdish, differ with respect to the way in which the gap is marked in a clause, see Section 7.2.3.

of the gap to percolate the tree even when it is already bounded off by the filler. The SLASH feature can still propagate higher to the top of the tree, which results in ungrammatical sentence constructions, as in example (8):

(8) *Him, she said him I like

To solve this problem, Pollard & Sag reorganized the NLOC value of a *sign* in the way shown in Figure 3.11. The INHER(ITED) value contains the NLOC specifications that can percolate upward. The TO-BIND value contains the NLOC specifications that will be blocked from further percolation.

$$\begin{bmatrix} \text{INHERITED} & \begin{bmatrix} \text{QUE} & \textit{(set of npros)} \\ \text{REL} & \textit{(set of refindices)} \\ \text{SLASH} & \textit{(set of loc-structure)} \end{bmatrix} \\ \text{TO-BIND} & \begin{bmatrix} \text{QUE} & \textit{(set of npros)} \\ \text{REL} & \textit{(set of refindices)} \\ \text{SLASH} & \textit{(set of loc-structure)} \end{bmatrix} \end{bmatrix}$$

Figure 3.11: The structure of a *sign*'s NLOC value

The difference between the INHER and TO-BIND values is used in the final version of the NLOC FEATURE PRINCIPLE, which is given in (9). In the new version of this principle, the INHER values of the daughters only percolate to the mother if they are not explicitly bound by the head daughter, i.e., mentioned in the head daughter's TO-BIND value.

(9) NLOC FEATURE PRINCIPLE (Pollard & Sag, 1994a: 164)

For each NLOC feature, the INHER value on the mother is the set union of the INHER value of the daughters minus the TO-BIND value on the head daughter.

Obviously, now the *Head-filler* schema needs to be adjusted. The last version, which includes the NLOC features from Figure 3.11, looks similar to schema 5 which is presented in section 3.2.2.[2]

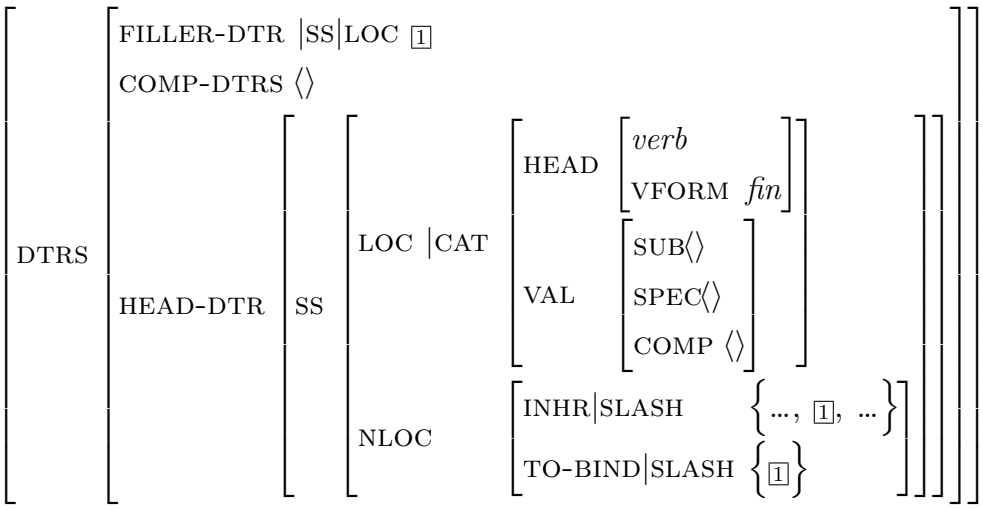


Figure 3.12: Head-filler schema adjusted from Pollard & Sag (1994a) (final version)

The resulting analysis of the sentence *Him, she said they know I like* is shown in Figure 3.13, on page 78.

[2]Some differences can be noticed between Pollard & Sag's description of *Head-filler* schema and the one I have presented including the added MARKING feature and the replacement of SUBCAT with VAL.

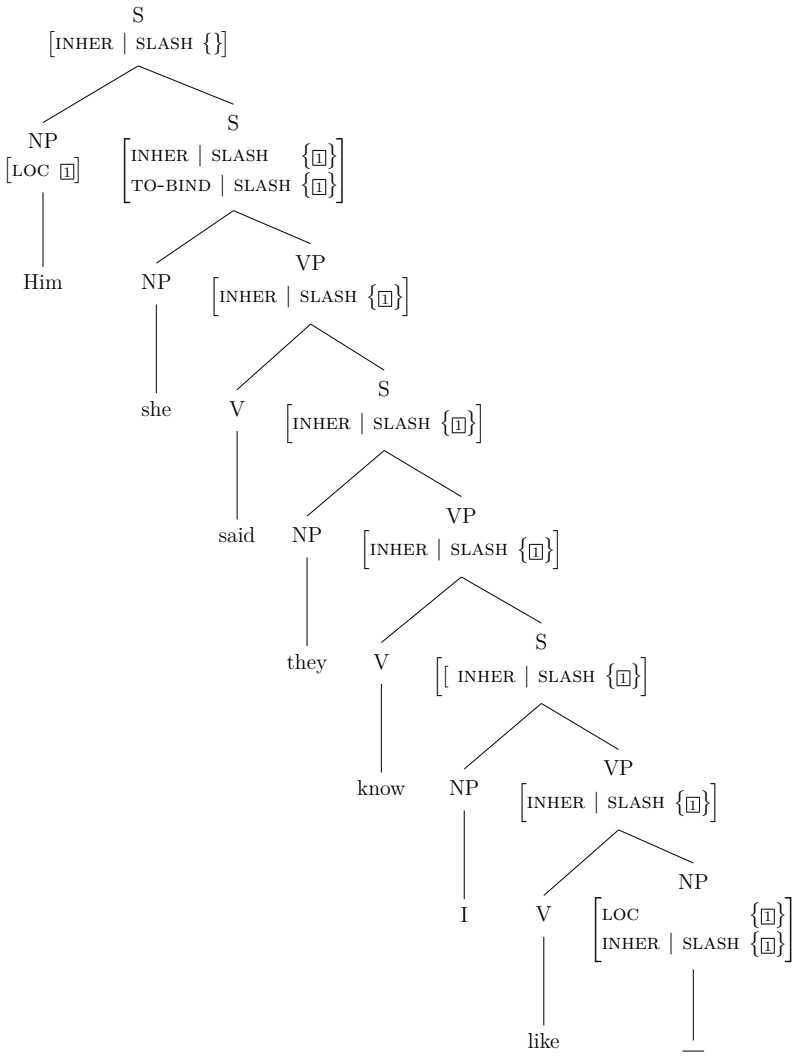


Figure 3.13: An HPSG tree of *Him, she said they know I like*

So far, only strong UDCs are presented, as in all of the above examples the antecedent of the gap is located in prenuclear position. There are other cases where the gap is located elsewhere, as in non-*wh*-RCs. In this case, the gap is directly related to the antecedent of the RC rather than being related indirectly through the relative word or phrase inside the RC as in example (10).

(10) I have the copy$_1$ [you printed $\_\__1$]. (weak UDCs)

A common example of a weak UDC is the *tough*-construction. As in all weak

UDCs there is no filler to bind the gap off in this construction. There is rather an element in an argument position that is co-referential with the gap. In weak UDCs, the element that is coreferential with the gap does not necessarily have the same case as the gap. This shows that the relation between the binder and the gap is rather weak in these constructions, and this is the main reason they are called weak UDCs. In example (11) there is no filler to bind with the gap and the subject that is co-referential with the gap is nominative while the gap is accusative.

(11) This problem.NOM is hard to solve __.ACC.

Pollard & Sag explain this by claiming that some predicates, such as *tough* and *easy*, have infinitive complements that contain an accusative NP gap which is co-referential with their subjects. Figure 3.14 is a partial representation of SYNSEM value of *easy* adjusted from Pollard & Sag (1994a: 167).

$$\left[\begin{array}{l} \text{LOC}|\text{CAT}\left[\begin{array}{l} \text{HEAD}\ \textit{adjective} \\ \text{VAL}\ \left[\begin{array}{l} \text{SUBJ}\ \left\langle \text{NP}_{\boxed{1}},\ (\text{PP}[\textit{for}]) \right\rangle \\ \text{SPEC}\ \langle\rangle \\ \text{COMPS}\left\langle \text{VP}\left[\textit{inf},\ \text{INHER}|\text{SLASH}\left\{ \ldots\boxed{2}\ \text{NP}[\textit{acc}]\text{:}\textit{ppro}\boxed{1},\ldots \right\}\right]\right\rangle \end{array}\right] \end{array}\right] \\ \text{NLOC}|\text{TO-BIND}|\text{SLASH}\left\{\boxed{2}\right\} \end{array}\right]$$

Figure 3.14: A partial representation of SYNSEM value of *easy*

The SLASH percolation is stopped due to the TO-BIND value specification. The co-referentiality is guaranteed by the co-indexing of the NP on the VAL lists and the NP inside the INHER|SLASH value of the selected VP.

The analysis of a sentence with a weak UDC, as in example (12), is shown in Figure 3.15.

(12) Kim is easy to please.

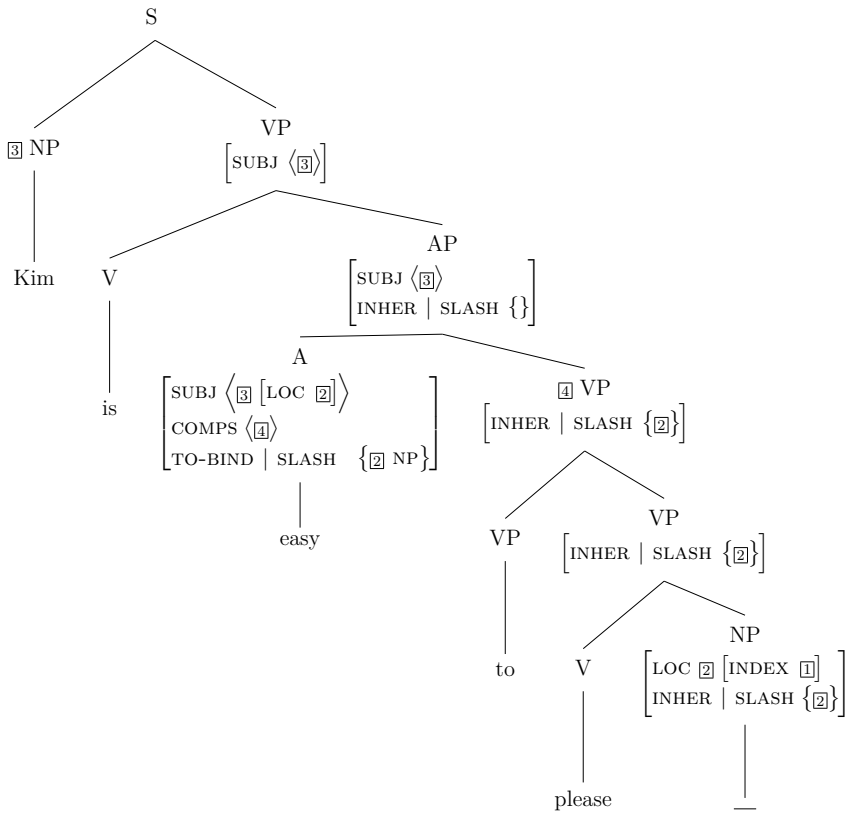


Figure 3.15: HPSG analysis of *Kim is easy to please*

3.3.2 Island constraints

Ross (1967) has raised the question of where gaps can appear in a sentence and formulated restrictions on gaps as so-called *Island Constraints*. In this section, I discuss some of these constraints and show how they are treated in Pollard & Sag's (1994a). Island constraints are considered important diagnostics for classifying a

phenomenon as an UDC. This fact will play an important role in my discussion of RCs in Kurdish in Chapter 7.

Pollard & Sag capture a number of island constraints as a restriction on traces in the form of the Trace Principle in (13).

(13) Every trace must be strictly sub-categorized by a substantive head. (Pollard & Sag, 1994a: 173)

The *Trace Principle* restricts the occurrence of traces to the complement position. It captures a number of *Island Constraints*, as the extraction of subjects or modifiers are more restricted than that of complementizers.

However, Levine & Hukari (2006) show that a more uniform theory of extraction is possible. They allow for a freer occurrence of traces and show that the observed restrictions can be accounted for once different valence lists are introduced, i.e., once valence saturation is not monitored by a simple SUBCAT list but by valence lists for individual grammatical functions, such as SUBJ and COMPS.

I will adapt Levine & Hukari's (2006) approach. This means, that I will not assume the Gap Principle. Instead, I allow for gaps in non-selected positions as well. However, there might be different constraints on gaps on the SUBJ list than on the COMPS list.[3]

Other island constraints have played an important role in the discussion of UDCs as well. I will briefly look at the *Complex NP Constraint* and the *Coordinate Structure Constraint*, as I will refer to them in chapter 7. Chaves (2020) provides an overview of various island constraints and their treatment in HPSG – as well as a critical discussion of their status as grammatical constraints on UDCs.

[3]I will show how restrictions on the occurrence of *that* in RCs can be accounted for in a system that assumes subject gaps, see chapter 6.

According to the Complex NP Constraint, extraction out of a clause that is embedded inside a noun phrase is excluded. This constraint is illustrated in (14) from Chaves (2020).

(14) a. *What does Robin know [someone who has __]?

b. *[Which book] do you believe the claim [that Robin plagiarized __]?

(Chaves, 2020: his examples (14) and (15))

The second island constraint that I will be using later in this thesis is the *Coordinate Structure Constraint*. According to this constraint, no element may be moved out of a conjunct unless it is moved out of all conjuncts, compare (15) and (16)

(15) * Which celebrity did you see [Priscilla and __]?

(Chaves, 2020: his example (4))

(16) Which celebrity did you buy [[a picture of __ and a book about __]]?

(Chaves, 2020: his example (10))

The status of these island constraints as grammatical principles are doubtful – see Chaves (2020). Chaves provides, among others, the following grammatical examples which violate the Complex NP Constraint and the Coordinate Structure Constraint respectively,

(17) Which rebel leader did you hear [rumors [that the CIA assassinated __]]?

(Chaves, 2020: his example (17b))

(18) Which celebrity did you see [the brother of __ with Priscilla]?

(Chaves, 2020: his example (8a))

In the light of the current discussion of island constraints, they should not be formalized as constraints of the grammar, but rather be considered the result of

semantic, pragmatic, and processing factors. Nonetheless, they can still serve as a diagnostic for identifying a UDC phenomenon.

3.3.3 Integrated relative clauses in Pollard & Sag (1994a)

Pollard & Sag (1994a) state that *wh*-RCs act like filler-gap dependencies in the sense that:

1. A *wh*-phrase such as *whom* can construct a super-ordinate phrase similar to when an NP gap constructs a phrase [INHR|SLASH{*NP*}].

2. The body construction of the *wh*-relative dependency is exactly the same as that of the filler-gap dependency, in the sense that they both contain bottom, middle and top. .

3. The *wh*-relative dependencies contain a *non-empty* INHER|REL value that will be inherited. Therefore, we can call the *wh*-relative a *wh*-dependency for it acts similar to the filler-gap dependencies.

All these can be better seen in the lexical entry of the relative pronoun *whom* in Figure 3.16.

The index of the relative phrase *whom*, is structure shared with the INHER|REL value. Since this relative phrase refers to humans, the value of CONTEXT|BACKGROUND|RELN is *human*.

English RCs depend on the parochial constraint, which states that the cardinality of the value of INHR|REL is at most 1.

In English RCs, unbounded dependencies can be introduced in a sentence that contains a *wh*-RC as in:

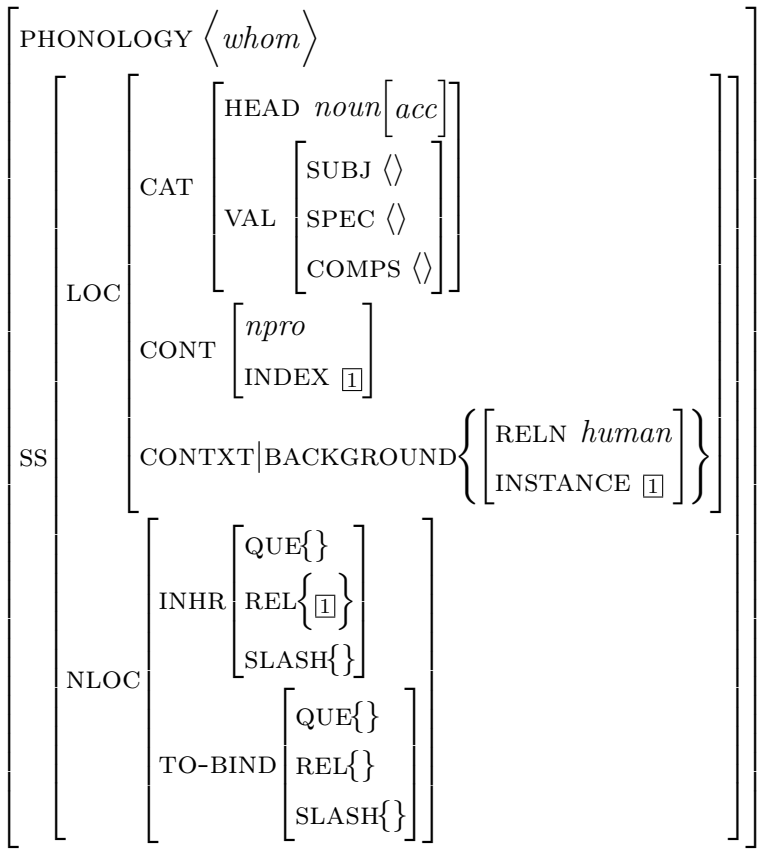


Figure 3.16: The lexical entry of the relative pronoun *whom* adapted from (Pollard & Sag, 1994a: 2010)

(19) Here's the rabbi [[[whose brother's] Bar Mitzvah] we attended]. (Pollard & Sag, 1994a: 212)

In example (19), one UDC is the relative dependency between the relative word and its antecedent NP, with which it shares an index. The other UDC is the filler-gap dependency.

One more characteristic that the filler-gap dependency and relative dependency share is that the relative dependency is unbounded inside the sentence as well as the filler-gap dependency. An instance of a large relative constituent is shown (20).

(20) I will show you the man [[at [the party [of [whose daughter]]]] your student

got the prize].

Additionally, the mechanism by which the inheritance in the REL specification is triggered is not thoroughly constraining. Due to this mentioned fact, multiple relative words can appear in a construction as in the following sentence

(21) * My boss whose wife remarried whose brother after their divorce is completely devastated.

To prevent the appearance of such constructions, more constrains need to be introduced. In this case, the *Relative Uniqueness Principle* might be the answer. This principle states that 'A member of the INHR|REL set on a headed constituent maybe inherited from at most one daughter' (Pollard & Sag, 1994a: 212).

Since RCs are nominal modifiers, they have a non-empty value for MOD, with this mentioned, the followings need to be true.

1. The index of the daughter NP is the same as the REL value of the RC.

2. The INHER|REL value on the mother is empty.

3. The restriction set on the mother's content includes the content of the RC.

Pollard & Sag (1994a) introduced the *null-relativizer*, which acts as the head of an RC. This element is classified under a part of speech that is called *relativizer*. Words of this part of speech have nonempty VAL lists and are specified for the attribute MOD[4].

In Figure 3.17, the lexical entry is subcategorizing for an S [INHR|SLASH{XP}] and an XP with *non-empty* INHR|REL value. In this case, the relativizer firstly, by schema 2, combines with the S [INHR|SLASH{XP}] and then by schema 1 it

[4]I will not be using this feature, MOD, in my analysis, instead, as mentioned in section 3.2.2 of this chapter, the feature SEL will be used. This feature is going to be replaced in all of the following descriptions and AVMs

$$\left[\begin{array}{l}\text{SS}\left[\begin{array}{l}\text{LOC}\left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{ll}\text{HEAD} & \left[\begin{array}{l}\textit{rltvzr}\\ \text{SEL}\ \text{N}'\left[\text{TO-BIND}|\text{REL}\ \left\{\boxed{1}\right\}\right]:\left[\begin{array}{l}\text{INDEX}\ \boxed{1}\\ \text{RESTR}\ \boxed{3}\end{array}\right]\end{array}\right]\\ \text{VAL} & \left[\begin{array}{l}\text{SUBJ}\left\langle \text{LOC}\ \boxed{4},\ \text{INHER}|\text{REL}\left\{\boxed{1}\right\}\right\rangle\\ \text{SPEC}\ \langle\rangle\\ \text{COMP}\left\langle \text{S}[\textit{fin},\textit{unmarked},\text{INHER}|\text{SLASH}\ \left\{\boxed{4}\right\}]:\boxed{5}\right\rangle\end{array}\right]\end{array}\right]\\ & \text{CONT}\left[\begin{array}{l}\textit{npro}\\ \text{INDEX}\ \boxed{1}\\ \text{RESTR}\ \left\{\boxed{5}\right\}\cup\ \boxed{3}\end{array}\right]\end{array}\right]\\ \text{NLOC}\left[\text{TO-BIND}|\text{SLASH}\ \left\{\boxed{4}\right\}\right]\end{array}\right]\end{array}\right]$$

Figure 3.17: The description of the *null-relativizer* from Pollard & Sag (1994a: 216)

combines with the construction with a non-empty INHR|REL value. Due to the fact that the null-relativizer needs to have a subject whose INHR|REL value is *non-empty*; hence, instances such as (22) will be avoided as the NP that must function as the subject of the relativizer doesn't contain a relative word.

(22) *I know the man letters Anna sent to whom.

This REL percolation explained in this section also accounts for complex relative constituents, i.e. for the five different percolation structures, which will be tackled in chapter 5.

According to Huddleston & Pullum (2002: 1082), a gap cannot take all the grammatical functions in English. A gap in UDCs can be either a post-head dependent or a subject in a clause structure. It can be a complement of a verb, complement of preposition, adjunct of a verb, etc. Gaps, when they are in an object position, usually appear in direct object position. It is less likely for them

to appear in indirect object position.

3.3.4 Relative clauses in constructional HPSG

Grammar sets restrictions on what type of feature structure should be in a language and what types shouldn't exist. It again decides on how these types should be systematized into a hierarchy, in return, grammar decides on the supertypes of each of these types. Furthermore, it is the grammar of a language that specifies the features that fit into a specific grammatical type and the type of the value that belongs to each feature.

There are three main differences between the analysis of RCs in Pollard & Sag (1994a) and Sag's (1997) constructional HPSG analysis. First, Sag abandons the idea of attaching RCs to an empty relativisor. He proposes that RCs are headed by the highest verbs in the clause. He sketches a *Multidimensional Organization of Phrasal Types* (MOP), in which he divides phrases into two main categories: *clausality* types and *headedness* types. He believes that each phrase types inherit from both main types. He assumes that this MOP allows him express generalizations about phrases in a way similar to how generalizations are expressed in hierarchical lexicons.

Second, he adds three new types of RCs to the other types introduced in Pollard & Sag (1994a).

The three types are:

1. Reduced RCs

   (23) The worst part is waiting for the results of the test; [being both worried and nervous]; in the lab.

2. infinitival RCs

   (24) She mentioned the exam, [for which to study hard], is the only way to pass.

3. Simple infinitival RCs

   (25) The only person [(for us) to visit] [whose kids Dana is willing to put up with] is Pat. (Sag, 1997: 470)

As stated by Sag (1997: 471), the first type, reduced RCs, might be attached to a nominal projection or an NP. This type is a subtype of *head-complement-phrase*. The last type, simple infinitival RCs, is a subtype of non-*wh*-RC which also is a subtype of *head-complement-phrase*.

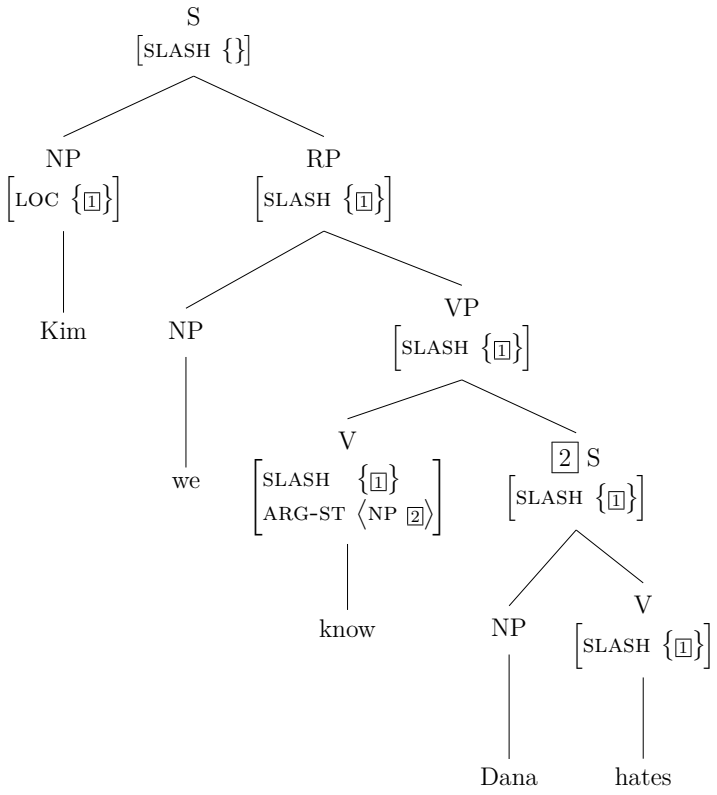
Third, he considers *that* as only a relative pronoun unlike in Pollard & Sag (1994a), where they mention two types of *that*: complementizer *that* and relative pronoun *that*. A discussion on this will be presented in chapter 6. In Sag's analysis, properties of grammar are not correlated with grammatical or lexical formations, they are correlated with constructions. According to Sag, verbs have a non-trivial MOD value when used as the highest verb in a RC. Sag (1997: 444) also distinguished between the types and the constructions of antecedents that the two main types of RC can take. He explains that non-*wh*-RCs take nominal projections that have non-empty values for the attribute SPECIFIER while the *wh*-RCs take constituents that have empty values for the attribute SPECIFIER. Sag's argument for this is, in the case of stacking of RCs, the *wh*-RCs always follow the bare RCs. Hence, the bare RCs are attached to the nominal projections while the *wh*-RCs are attached to NPs.

Sag divides phrases depending on their dimensions into the following types:

1. Clausality
   (a) Clauses
      - Interrogative clauses
      - Declarative clauses
      - RCs
      - Imperative clauses
   (b) Non-clauses
2. Headedness
   (a) Headed-phrases
      - Head-subject-phrase
      - Head-complement-phrase
      - Head-filler-phrase
      - Head-adjunct-phrase
      - Head-marker-phrase
   (b) Nonheaded-phrases

Additionally, Sag (1997: 442) eliminates all non-branching structures such as VPs and NPs; for instance, according to him, a sentence such as *Anna left* consists of only an N and an V, not an NP and an VP.

Sag discusses two types of dependencies: *extraction*, between the filler and gaps and *upward percolation*. In the analyses of filler-gap dependencies, Sag abandons the two attributes introduced in Pollard & Sag (1994a), namely INHR|TO-BIND. His analysis for the phrase *Kim we know Dana hates* is as follows:



(Sag, 1997: 445)

Figure 3.18: An HPSG tree of *Kim, we know Dana hates*

Sag (1997) has been criticized by many, among whom are Müller (2007) and Walker (2017). Müller criticizes Sag for considering the meaning of RCs analogous to the meaning of verbs inside the sentence. Walker agrees with Müller and expresses further criticism, such as the absence of accounting for RC extraposition analysis in Sag (1997). Further references of this constructional HPSG will be mentioned in the following chapters, chapter 5 and 6, when I find it necessary to use it for analyzing my data.

3.3.5 Arnold's (2004) constructional analysis of SRCs

Arnold (2004) tries to provide both semantic and syntactic analyses for SRCs. For the syntactic analysis, he mostly depends on Sag's (1997) treatment of IRCs. For the semantics analysis of SRCs, he depends on the analyses previously provided for ordinary discourse anaphors for he believes that non-restrictive relative pronouns behave similarly to normal anaphoric pronouns. Arnold (2004: 28) argues that the main semantic difference between SRCs and IRCs is that IRCs have an intersective interpretation. SRCs, on the other hand, have what he calls a *global scope* interpretation. This intersective interpretations of IRCs allows the possibility of having a contrast set, an example of this is (26).

(26) Anna has two musical instruments which are broken. The others are fine.

However, SRCs don't allow this as they have a totality interpretation. The antecedent modified by an SRC is wholly modified and there is no room for separating a subset from it, for this reason, example (26) repeated as (27) with an SRC is considered ungrammatical.

(27) Anna has two musical instrument, which are broken. *The others are fine.

Arnold (2004: 30) argues that normal anaphoric pronouns behave the same.

(28) Anna has two musical instrument. They are broken. * The others are fine.

He, additionally, connects the ability of SRCs to take a wider range of antecedents, such as taking split antecedents and a whole clause as its antecedent, to its similar behavior with normal anaphoric pronouns as in example (29).

(29) I fixed her coffee machine, and Karim cleaned the garden, which she noticed after two days.

Arnold states that, similarly to pronominal anaphora, SRCs can contain parts of idioms unlike IRCs.

(30) a. *The headway, which the students made last week, was amazing.

b. The headway which the students made last week was amazing. [IRC]

c. *The headway was amazing. The students made it last week.

Arnold states that previously Sells (1985) suggested that SRCs can have both wide scope with *negative*, *interrogative operators*, and *quantifiers*, and narrow scope with *quantifiers*. SRCs usually take wide scope as they act like usual anaphora. They can have narrow scope when they fall into the scope of a quantifier.

However, Arnold believes that SRCs can have both wide and narrow scope simultaneously.

(31) Sam believes every chess set comes with a spare pawn, which Kim thinks is usually taped to the top of the box with its base uppermost. (Arnold, 2004: 32)

The most natural interpretation suggests that a *spare pawn* is in the scope of *every chess set* and they both are in the scope of *believes*. However, this is not the interpretation of what the sentence in (31) means. The sentence does not suggest that *what Sam believes Kim thinks*. Semantically, the SRC does not fall into the scope of the verb *believe* in the main clause. Arnold uses *Discourse Representation Theory*, DRT, to interpret SRCs semantically. The DRS of the sentence *Sams doesn't own a car, which she wouldn't be able to drive* is as Figure 3.19.

The top box, *LT*, contains the contents of both the main clause and the SRC, and it says that there is an individual s who is *Sam*. The L1 box say that there is an individual c and c is a car and s who is *Sam* owns this c which is *a car*. The

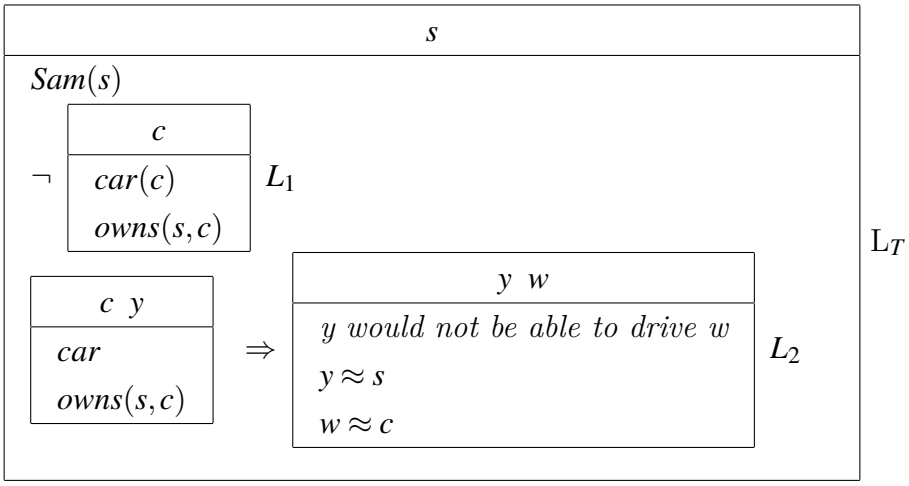


Figure 3.19: The semantic interpretation of *Sam doesn't own a car, which she wouldn't be able to drive* (Arnold, 2004: 37)

box on the lower left says that there are individuals c and y; c is *a car* and y is the anaphoric reference of s, and s owns c. Finally, the box at the lower right, L2, is where the RC is interpreted. In this box conditions are introduced that connect the variables to their antecedents such as y to its antecedent *Sam* and w to its antecedent car.

In solving the problem of which type of RC can function as SRCs or IRCs, and which types of antecedents SRCs and which IRC can take, Arnold (2004: 43) proposes two constraints.

(32) $(\textit{rel-cl} \wedge \textit{intersective-sem}) \rightarrow \left[\textsc{head}|\textsc{mod}|\textsc{head}\ \ \textit{noun}\right]$

(33) $(\textit{rel-cl} \wedge \textit{global-scope-sem}) \rightarrow \left[\textsc{head}|\textsc{mod}|\textsc{comps}\ \langle\rangle\right]$

The constraint in (32) restricts IRCs to nominal antecedents. The constraint in (33) has the effect that SRCs must attach outside complements, i.e., they must adjoin to maximal projections. However, this constraint does not yet restrict the form of the SRC itself. For this reason, Arnold introduces the constraint in (34), which ensures that SRCs are finite and contain a head-filler constellation.

(34) (*rel-cl* ∧ *global-scope-sem*) → *fin-head-filler-phrase*

(Arnold, 2004: 44)

Due to the fact that Arnold uses Sag's (1997) analysis according to which *that* is a relative pronoun, this constraint does not exclude *that* SRCs. However, in his discussion of the data, Arnold considers *that* SRCs ungrammatical. I will sketch how Arnold's (2004) excludes *that* SRCs in chapter 6.

Arnold believes that, syntactically, IRCs behave very similarly to SRCs therefore he applies the analyses Sag (1997) provided for IRCs on SRCs. He provides a list of similarities between these two constructions. However, because this subject has been excessively discussed in chapter 2, I will only mention one main similarity that Arnold mentions. Arnold believes the processes of raising and topicalization prove that similarly to IRCs, SRCs form a constituent with their heads.

(35) a. The parcel, which she sends every month, they deliver on time.

b. *The parcel, they deliver, which she sends every month, on time. (Topicalization)

(36) a. Anna, who had the surgery, seems very pale.

b. *Anna seems, who had the surgery, very pale. (Raising)

Now that all of the principles and mechanisms that are necessary for the HPSG analysis for the syntactic aspects of my data is presented, it is time to introduce the semantic principles and constrains needed for the analysis. Next section of this chapter is dedicated to the semantic analysis I use throughout my thesis.

3.4 LRS

The semantic framework, LRS, which I adopt in analysing the semantics of my data is presented in this section. The section consisted of three parts: firstly, I give the framework's background in 3.4.1. Secondly, I present a summary of the theory of the framework, including the framework's mechanism and main principles in 3.4.2. Finally, I illustrate the implementations of the framework in shapes of AVMs on the gradual combinations of structures in making a sentence in 3.4.3.

3.4.1 The background

The semantic framework LRS is developed by Frank Richter and Manfred Sailer. First, in Richter & Sailer (2001a), LRS was based on analyzing interrogatives in German. Later, Richter & Sailer (2001b) used LRS to analyze negation in Polish. They developed LRS to form a semantic theory that is harmonious with Pollard & Sag's (1994a) grammar theory.

Richter & Sailer (2004) present a comprehensive description of the foundation and the implimention of LRS. Richter & Sailer (2004: 87) argue that since the logical languages are derived from standard mathematical systems, they must be straightforwardly comprehensible by linguists, regardless of the grammatical frameworks they are used for. However, the eccentric use of semantic representations in HPSG framework might seem inadequate to some semanticists. Richter & Sailer aimed to design a new semantic framework as they argued that the semantic representations that are used in HPSG are to serve some purposes which are only useful for the HPSG framework and for non-experts in that framework the representations might seem idiosyncratic.

Richter & Sailer (2004) adopted the language Two-Sorted Type Theory, Ty2, from Gallin (1975) for their semantic representations. Ty2 is analogous to *Intentional Logic* (Montague, 1974); nevertheless, Richter & Sailer (2004) chose Gallin's (1975) Ty2 as they argue that it has technical advantages for their theory.

The use of Ty2 representations in HPSG is independent of the particular choice of the system of semantic combinatorics. For example, Sailer (2003) arrives at such representations using Lexicalized Flexible Ty2, a combinatorial system based on functional application. Another option, which I will pursue here, is LRS. This is a constraint-based mechanism that countenances underspecified representations. Underspecification at the description level is the most prominent example of underspecification in HPSG.

? utilize lexical descriptions to make the notion of underspecification in HPSG more clear. They state that a lexical entry can be considered underspecified if the value of some attributes are not exclusively fixed in the lexical entry. Richter & Sailer argue that the value of the INVERTED attribute of auxiliary verbs in English, in HPSG, is underspecified while this attribute's value is plus or minus in every sentence. A similar example of underspecification in the semantic description of a lexicon in HPSG is seen in the description of the verbs *to* and *be*. The CONTENT values of these two verbs are identical to that of the verbal projection they select. Considering the syntactic context they appear in, this shared value means that they have several CONTENT values instead of a particular identified one.

3.4.2 The theory

According to Penn & Richter (2004: 425), the guiding assumptions behind LRS are:

1. All semantic and syntactic idiosyncrasies are lexical.

2. There is no non-functional semantic contribution from outside of the lexicon.

Sailer (2004) suggests that in order to understand the semantic representation in LRS, it is necessary to know that the framework accounts for lexical/local semantics and compositional semantics separately. The values of the attribute CONTENT are objects of the sort *content*. Inside the LOC value of *sign*, the value of the CONTENT attribute encodes the lexical/local semantic attribute which involves the linking theory, semantic selection, and binding phenomena. The CONTENT value consists of two attributes. One attribute is INDEX, which is the grammatical index. The other attribute stipulates the meaning expression the *sign* contributes, this attribute is called MAIN. INDEX has two values; the first is the discourse referent associated with a *sign*, DR; and second is the is a feature PHI, whose value specifies the phi features (person, number, gender) of a sign. The former and the latter values are added to the original INDEX of HPSG so that all parts of speech can have INDEX not only nouns, adjectives and determiners as it originally was the case in Pollard & Sag (1994a).

This need for the separation of the morphosyntactic function and the semantic function into a set of phi-features and a semantic expression in traditional HPSG is mentioned by Levine et al. (2012). To include these aspects in the value of the INDEX feature, the sort *extended-index* is introduced. The appropriateness conditions of this sort are given in (37).

(37) Sort *extended-index*

extended-index

PHI *phi-index*

DR *meaningful-expression*

This additional sort is shown in the content of the noun *baby* in Figure 3.20.

$$\begin{bmatrix} \textit{content} & \\ \text{INDEX} & \begin{bmatrix} \textit{extended-index} & \\ \text{DR} & x \\ \text{PHI} & \begin{bmatrix} \textit{index} & \\ \text{PERS} & \textit{third} \\ \text{NUM} & \textit{sg} \\ \text{GEN} & \textit{neutr} \end{bmatrix} \end{bmatrix} \\ \text{MAIN} & \textbf{baby}' \end{bmatrix}$$

Figure 3.20: The value of CONTENT for the noun *baby*

The value of the attribute LOGICAL-FORM, LF, of signs, is where the semantic combinatorics are described. Because the combinatorial semantics is not in the SYNSEM value of *sign*, it is not accessible for semantic and syntactic selection by heads. Some kinds of interaction, for instance, the syntactic head's lexical selection of arguments' semantic variables, are permitted as several specific parts of the compositional semantics are linked to some particular constituents of the local content by an interface theory.

In LRS, the composition of semantics distinguishes between three attributes namely; INCONT, EXCONT and PARTS. In a *sign*, the INCONT encodes the scopely lowest expression that the *sign*'s semantic head contributes. The EXCONT is the *sign*'s maximal syntactic projection's contribution to the meaning of the overall expression. Finally, the attribute PARTS contains all semantic contributions of a *sign*.

The appropriateness conditions of the sort *lrs* are given in (38).

(38) The sort *lrs* (LF value of signs)

lrs	
EX(TERNAL-)CONT(ENT)	*me*
IN(TERNAL-)CONT(ENT)	*me*
PARTS	*list(me)*

3.4.3 The implementation of the framework

I assume it is crucial here to illustrate the assembly of the attribute and values of semantic representations in descriptions of lexical entries in 3.4.2. For this purpose, I present the lexical entries of the quantifier *every* in Figure 3.21, the adjective *healthy* in Figure 3.22, the noun *baby* in Figure 3.23 and the verb *cry* in Figure 3.24.

$$\begin{bmatrix} \textit{word} \\ \text{PHON} \ \langle \textit{every} \rangle \\ \text{SYNS} \begin{bmatrix} \text{LOC} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \textit{det} \\ \text{SEL} \ \text{N}\big[\text{INDEX} \ \boxed{3}\big] \end{bmatrix} \\ \text{VAL} \begin{bmatrix} \text{SUBJ} \ \langle\rangle \\ \text{SPR} \ \langle\rangle \\ \text{COMPS} \ \langle\rangle \end{bmatrix} \end{bmatrix} \\ \text{CONT} \begin{bmatrix} \textit{content} \\ \text{INDEX} \begin{bmatrix} \text{PHI} \begin{bmatrix} \text{PER} & \textit{3rd} \\ \text{NUM} & \textit{sg} \end{bmatrix} \\ \text{DR} \ x \end{bmatrix} \\ \text{MAIN} \ \boxed{1} \ \forall \end{bmatrix} \end{bmatrix} \\ \text{LF} \begin{bmatrix} \textit{lrs} \\ \text{EXCONT} \ \textit{me} \\ \text{INCONT} \ \boxed{2} \ [\forall x : \alpha](\beta) \\ \text{PARTS} \ \langle x, \boxed{1}, \boxed{2} \rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 3.21: The description of the lexical entry of the quantifier *every*

The INCONT value of the quantifier *every* specifies parts of the expression that are being described rather than exclusively identifying an expression of the semantic representation language. It expresses the universal quantifier $\forall$ that binds the variable x and has a restrictor, α, and a scope, β.

$$\left[\begin{array}{ll}
\text{PHON} & \langle \textit{healthy} \rangle \\
\text{SYNSEM} & \left[\text{LOC} \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll}
\text{HEAD} & \left[\begin{array}{ll} \textit{adj} & \\ \text{SEL} \quad \text{N}' \text{:} & \left[\text{INDEX } \boxed{1}\right] \\ \text{PRD} & - \end{array}\right] \\
\text{VAL} & \left[\begin{array}{l} \text{SUBJ } \langle\rangle \\ \text{SPR } \langle\rangle \\ \text{COMPS } \langle\rangle \end{array}\right] \\
\text{CONT} & \left[\begin{array}{ll} \text{INDEX } \boxed{1} & \left[\text{DR } x\right] \\ \text{MAIN } \boxed{2} & \mathbf{healthy}'(x) \end{array}\right]
\end{array}\right]
\end{array}\right] \\
\text{LF} & \left[\begin{array}{ll}
\textit{lrs} & \\
\text{EXCONT } \boxed{4} & (\alpha \wedge \beta) \\
\text{INCONT } \boxed{3} & \mathbf{healthy}'(x) \\
\text{PARTS} & \langle x, \boxed{2}, \boxed{3}, \boxed{4} \rangle
\end{array}\right]
\end{array}\right]$$

and $\boxed{3} \lhd \beta$

Figure 3.22: The description of the lexical entry *healthy*

The HEAD value of *healthy* is an adjective and it modifies an N. The interpretation of the adjective *healthy* involves unioning the restrictions on the INDEX of *healthy* with those on the INDEX of the N it describes. The semantic constant $\mathbf{healthy}'$ is the value of the adjective's INCONT attribute. And it is applied to the discourse referent, $\mathbf{healthy}'(x)$. The PARTS list shows the overall semantic contribution of the adjective *healthy*. The PARTS list contains the referential index x, the semantic constant $\mathbf{healthy}'$, $\mathbf{healthy}'(x)$, and $\alpha \wedge \beta$.

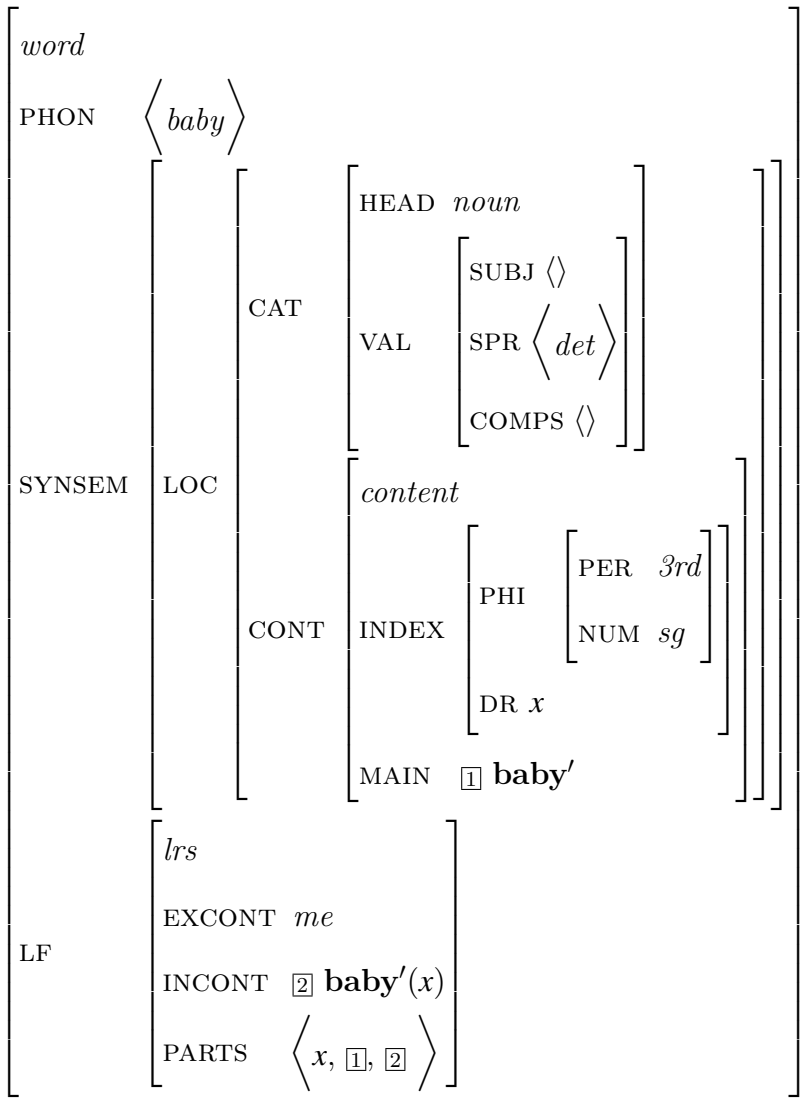


Figure 3.23: The description of the lexical entry *baby*

The value EXCONT of the lexical entry of the noun *baby* is underspecified. The word's EXCONT value is specified as *meaningful-expression* (*me*), i.e., it can be any expression of the semantic representation language Ty2. This implies that when the noun is combined in a phrase then it will receive its specified EXCONT value. The semantic constant **baby′** is the value of the noun's INCONT attribute. It is applied to the discourse referent, **baby′**(x). Finally, the PARTS list demonstrates the overall semantic contribution of the noun *baby*. The PARTS list contains the

referential index x, the semantic constant **baby′**, and **baby′**(x), which is the functional application of the MAIN value to the discourse referent.

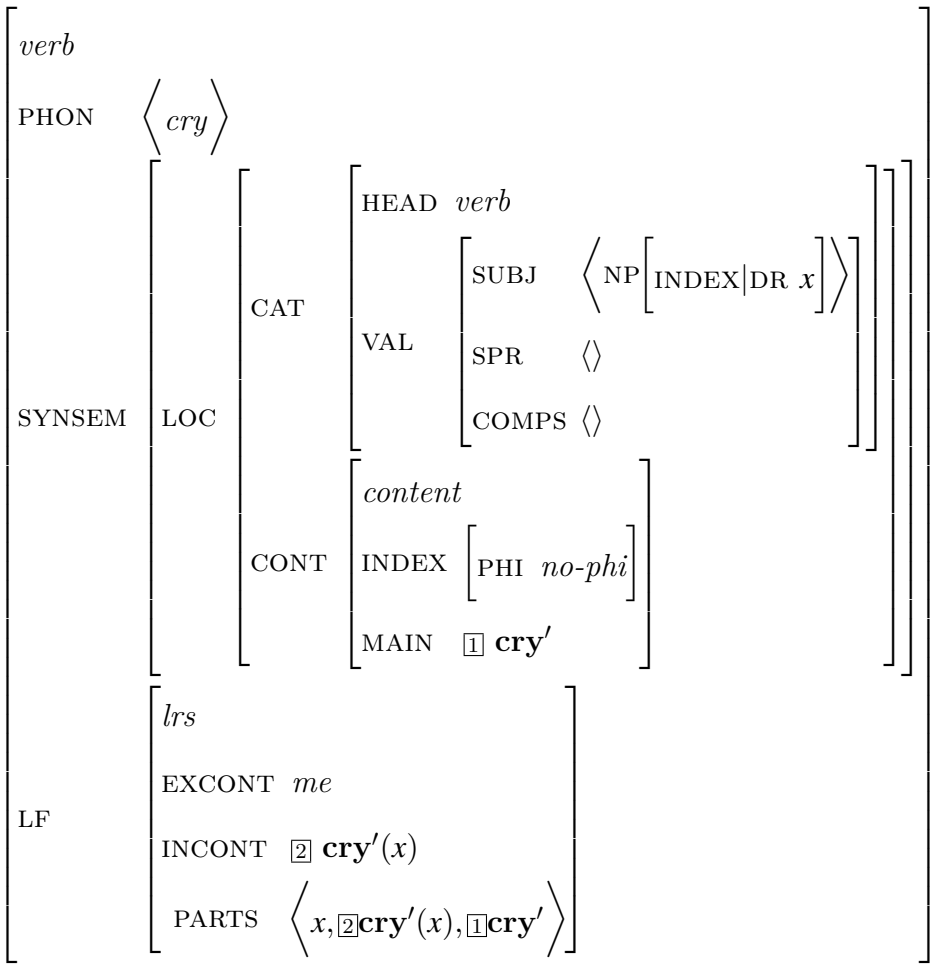


Figure 3.24: The description of the lexical entry *cry*

The core lexical semantics of the verb *cry* is the predicate **cry′**, expressed as the MAIN value. The verb has the PHI value *no-phi*. The PARTS list contains the variable (x), **cry′** and **cry′**(x).

Generally there are four main principles that the LRS anlysis depends on; here, I will make use of the formulation of the principles from Penn & Richter (2004: 230-231). One of these principles is the Incont Principle given in (39). This principle guarantees the INCONT value is included in the PARTS list. Moreover,

the value of INCONT is a component of the EXCONT value.

(39) The Incont Principle

In each *lrs*, the INCONT value is an element of the PARTS list and a component of the EXCONT value.

The EXCONT value of *sign* corresponds to the semantic representation associated with that of the *sign*'s full projection. For an utterance, this means that it is the utterance's overall semantic representation.

The EXCONT Principle, given in (40), consists of two parts: the first concerns phrases which are associated with the EXCONT value of a completed projection. The value of EXCONT of a completed head projection is included in the PARTS list of this sign. The second one concerns utterances. The EXCONT principle declares that all and only the elements of the PARTS list should be components of the EXCONT value. This means that only the meaning contributions by lexical elements in the utterance form the logical form of the utterance.

(40) The EXCONT Principle:

a. In every phrase, the EXCONT value of the NON-HEAD-DTR is an element of the non-head daughter's PARTS list.

b. In every utterance, every subexpression of the EXCONT value of the utterance is an element of its PARTS list, and every element of the utterance's PARTS list is a subexpression of the EXCONT value.

The LRS Projection Principle, given in (41), confirms that, in a phrase, the phrase and its head daughter have the same EXCONT value and the same INCONT value. Additionally, in a phrase, the PARTS list contains all of the elements in the daughter's PARTS lists. This ensures the inclusion of all the semantic contributions

of the signs dominated by the phrase described, it also prevents the introduction of semantic contributions in headed phrases.

(41) The LRS Projection Principle:

In each headed-phrase

a. The EXCONT values of the head and the mother are identical,

b. The INCONT values of the head and the mother are identical,

c. The PARTS value contains all and only the elements of the PARTS values of the daughters.

The Semantics Principle, given in (42), introduces constraints on the possible readings of phrases. These constraints depend on the syntactic and semantic properties of the phrase.

(42) The Semantics Principle:

In each headed-phrase,

a. If the non-head daughter is a quantifier with an INCONT value of the form $[Qx:\rho](\nu)$, then the INCONT value of the head is a component of ρ, and the INCONT value of the non-head daughter is identical with the EXCONT value of the head daughter.

b. If the non-head daughter is a quantified NP with an EXCONT value of the form $[Qx:\rho](\nu)$, then the INCONT value of the head is a component of ν.

c. If the NON-HEAD-DTR is a modifier whose external content is of the form $\alpha \wedge \beta$, then the head's internal content is a component of α and the modifier's external content is a component of the head's external content.

The constraint in (42a) concerns the combination of a head noun with a quantifier. This constraint guarantees that the head noun's INCONT is in the quantifier's restrictor. It also ensures that the value of the head noun's EXCONT equals to the value of the quantifier's INCONT. The constraint in (42b) concerns quantified NP arguments and their combination with a syntactic head. The constraint postulates that the head's INCONT should fall into the quantifier's scope.

To illustrated the above principles, I have provided an AVM of a sentence that consists of the words from the lexical entries shown in Figure 3.21, Figure 3.22, Figure 3.23 and Figure 3.24. The sentence is:

(43) Every healthy baby cries.

$[\forall x : (\mathbf{baby}'(x) \wedge \mathbf{healthy}'(x))](\mathbf{cry}'(x))$

Before presenting the AVM of the whole sentence, I will first show the AVM of the gradual combination of the elements of the sentence. Figure 3.25 shows the AVM of the combination of the noun *baby* with the adjective *healthy*, and figure 3.26 shows the AVM of the combination of the Adj and the N with the determiner *every*. Figure 3.27 shows the AVM of the entire sentence.

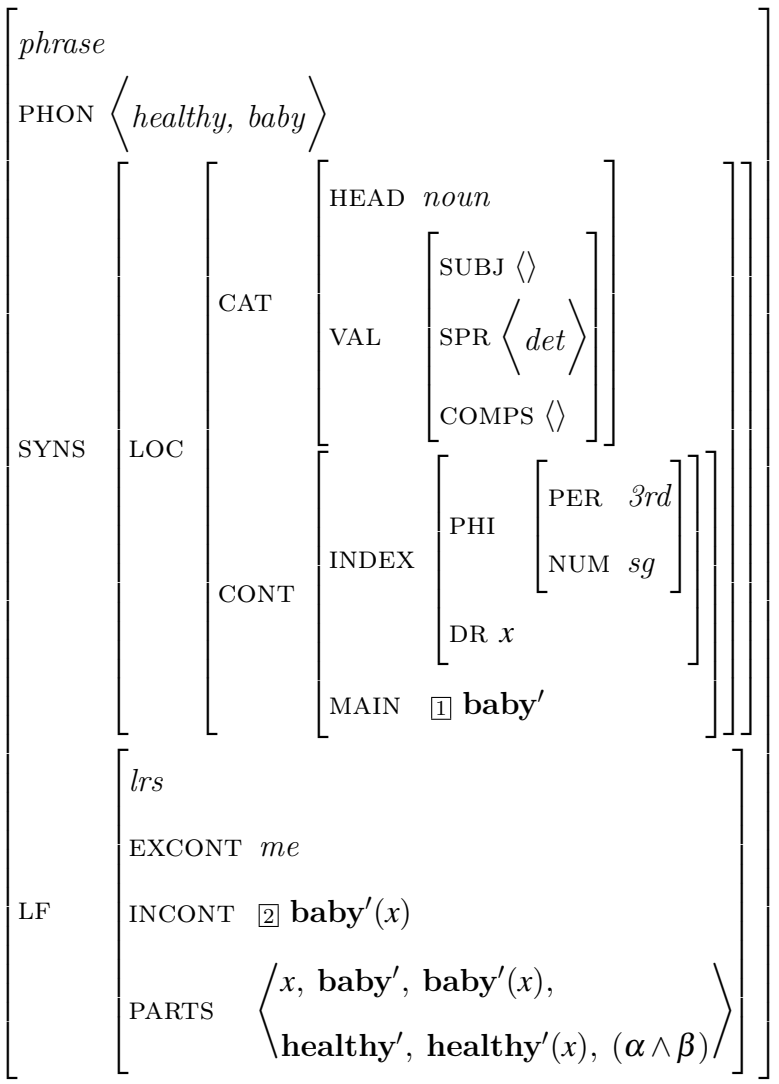


Figure 3.25: *healthy baby*

The HEAD value of the phrase *healthy baby* is the noun *baby*, in a SPR the phrase is looking for its determiner. The semantic constant **baby′** is the value of the phrase's MAIN value. Similarly the value of the INCONT of the phrase is identical with that of the word *baby*, **baby′**(x). The EXCONT value of the phrase is not constrained, as it is neither a non-head daughter of a larger phrase nor a complete utterance. All we know is that it is identical with the EXCONT value of

its head daughter. For this reason, I just write *me* in Figure 3.25. The Semantic Principle requires the INCONT of the word *baby* to be a subexpression of α. The PARTS list is a combination of the PARTS list of both *healthy* in Figure 3.22 and *baby* in Figure 3.23.

$$\left[\begin{array}{ll}
\textit{phrase} & \\
\text{PHON} & \left\langle \textit{every, healthy, baby} \right\rangle \\
\text{SYNS} & \left[\begin{array}{ll}
\text{LOC} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll} \text{HEAD} & \textit{noun} \\ \text{VAL} & \left[\begin{array}{l} \text{SUBJ } \langle\rangle \\ \text{SPR } \langle\rangle \\ \text{COMPS } \langle\rangle \end{array}\right] \end{array}\right] \\
\text{CONT} & \left[\begin{array}{ll} \text{INDEX} & \left[\begin{array}{ll} \text{PHI} & \left[\begin{array}{ll} \text{PER} & \textit{3rd} \\ \text{NUM} & \textit{sg} \end{array}\right] \\ \text{DR} & y \end{array}\right] \\ \text{MAIN} & \mathbf{baby}'(x) \end{array}\right]
\end{array}\right]
\end{array}\right] \\
\text{LF} & \left[\begin{array}{ll}
\textit{lrs} & \\
\text{EXCONT} & [\forall x : (\mathbf{baby}'(x) \wedge \mathbf{healthy}'(x))](\gamma) \\
\text{INCONT} & \boxed{2}\ \mathbf{baby}'(x) \\
\text{PARTS} & \left\langle \begin{array}{l} x,\ \mathbf{baby}',\ \mathbf{baby}'(x), \\ \mathbf{healthy}',\ \mathbf{healthy}'(x),\ (\alpha \wedge \beta) \\ \forall,\ [\forall x : \delta](\gamma) \\ \quad \text{and } \boxed{2} \lhd \delta \end{array} \right\rangle
\end{array}\right]
\end{array}\right]$$

Figure 3.26: *every healthy baby*

The head daughter of the phrase *every healthy baby* is the noun *baby*. The

values of the features MAIN and INCONT are identical on the phrase and the head daughter. The values of the phrase's attribute EXCONT is the same as the value of the phrase *healthy baby*'s EXCONT plus the semantic contribution of the universal quantifier ∀, *every*. The PARTS list is a combination of the PARTS list of *every* in Figure 3.21, *healthy* in Figure 3.22 and *baby* in Figure 3.23.

$$\left[\begin{array}{ll}
\textit{phrase} & \\
\text{PHON} & \left\langle \textit{every, healthy, baby, cries} \right\rangle \\
\text{SYNS} & \left[\text{LOC} \left[\text{CAT} \left[\begin{array}{ll}\text{HEAD} & \textit{verb} \\ \text{VAL} & \left[\begin{array}{l}\text{SUBJ } \langle\rangle \\ \text{SPR } \langle\rangle \\ \text{COMPS } \langle\rangle\end{array}\right]\end{array}\right]\right.\right. \\
 & \left.\left.\text{CONT} \left[\begin{array}{ll}\text{INDEX} & \left[\begin{array}{ll}\text{PHI} & \left[\begin{array}{ll}\text{PER} & \textit{3rd} \\ \text{NUM} & \textit{sg}\end{array}\right] \\ \text{DR} & y\end{array}\right] \\ \text{MAIN} & \mathbf{cry}'\end{array}\right]\right]\right] \\
\text{LF} & \left[\begin{array}{ll}
\textit{lrs} & \\
\text{EXCONT} & [\forall x : (\mathbf{baby}'(x) \wedge \mathbf{healthy}'(x))](\mathbf{cry}'(x)) \\
\text{INCONT} & \boxed{3}\ \mathbf{cry}'(x) \\
\text{PARTS} & \left\langle \begin{array}{l} x,\ \mathbf{baby}',\ \mathbf{baby}'(x),\ \mathbf{healthy}',\ \mathbf{healthy}'(x),\ (\alpha \wedge \beta) \\ \forall,\ [\forall x : \delta](\gamma),\ \mathbf{cry}',\ \mathbf{cry}'(x) \end{array} \right\rangle
\end{array}\right]
\end{array}\right]$$

and $\boxed{3} \lhd \delta$

Figure 3.27: *every healthy baby cries*

The head daughter of the sentence *every healthy baby cries* is the verb *cry*.

Therefore, the semantic constant **cry′** is the value of the phrase's MAIN and INCONT attributes. The EXCONT value of the sentence is also the same as that of the verb. The PARTS list is a combination of the PARTS list of the subject NP and the verb, as given in Figure 3.26 and Figure 3.24 above.

As this is a complete utterance, the second part of the Excont Principle applies, i.e., the EXCONT value of the sentence is constrained to be an element of the sentence's PARTS list that contains exactly all element of the PARTS list as subexpressions. The only candidate for this is the formula given in the AVM.

3.4.4 Summary

To conclude, all the principles and mechanisms that are needed for the HPSG and LRS analyses of my data in chapters, 5, 6 and 7 are illustrated in this chapter. Firstly, the framework of HPSG was presented, based on Pollard & Sag (1987), Pollard & Sag (1994a), Sag (1997) and Ginzburg & Sag (2000). For the analysis of SRCs I mostly rely on Pollard & Sag's (1994a) analysis of IRCs and Arnold's (2004) and (2007) HPSG analysis of SRCs. Secondly: semantic principles and constraints of LRS from Richter & Sailer (2001b) and Penn & Richter's (2004), which were illustrated in this chapter, will provide correct semantic interpretations for the analysis of the types of SRCs discussed in the following chapters.

Chapter 4

Determiner-*which*: A data survey

In this chapter, firstly, I provide a brief background on the usage of determiners as relative words in Section 4.1. Next, in Section 4.2, I present a questionnaire study on testing the naturalness of the occurrences of determiner-*which* with five types of upward percolation, which are introduced in Chapter 2. Section 4.2.2 contains a detailed explanation on the design, methodology, and the predictions. The results alongside a discussion of each type of upward percolation with determiner-*which* and pronoun-*which* are given in 4.2.3. .

4.1 Relative phrases with determiner-*which*

Diessel (2003: 635) states that, in English, interrogatives and demonstratives are the two most used structures in constructing relative words/phrases. In investigating the origin of the interrogative-based RC elements, Hendery (2007: 61) demonstrates that the initial occurrence of *who*-succession as relative pronouns were only *whose*, a possessive pronoun, and *whom*, an oblique pronoun. The usage of *who* as a relative pronoun, only very uncommonly, occurred in the 1500s.

At this time, the usage of *which* as a relative pronoun started to become more common; although, at first it occurred in the early 1200s. According to Grijzenhout (1992: 45), the successive use of *who*-based RCs was initially only for SRCs. Later, this usage incrementally became common with other types of RCs, such as IRCs. Determiner-*which* as a relative pronoun has not been studied prominently in the literature and has only been briefly referred to – rather in reference grammars than in linguistic analyses.[1] One such grammar is Huddleston & Pullum's (2002). As mentioned in section 2.3, Huddleston & Pullum (2002: 1059) state that SRCs can take a wider range of types of relative words than IRCs. One of the relative words that only SRCs take is determiner-*which*.

Huddleston & Pullum present six types of upward percolation in English relative phrases. They illustrate four of the types with prononominal relative words, such as *which* and *who(m)*. They introduce two additional types for the two determiner relative words, possessive *whose* and determiner *which*. For ease of reference, I repeat the list of the types of upward percolation from section 2.2.2.[2]

1. From the complement of a preposition to PP (from NP to PP)

   (1) a. The curtain Kim was hiding behind

   b. The curtain behind which Kim was hiding. (Huddleston & Pullum, 2002: 1040)

2. From PP to AdjP.

   (2) Many students applied for the program, the best among which will be selected.

[1] The same seems to be true for relative *whose*. Even Hendery (2007), who attempts to address all kinds and constructions of relative elements across many languages, including English, does not mention determiners as relative words.

[2] As I will not discuss possessive *whose* in this thesis, I will ignore the percolation type 5.

3. From PP to NP

   (3) She's just sat her final exam, the result of which we are expecting next week. (Huddleston & Pullum, 2002: 1040)

4. From NP to non-finite verb (from NP to V-nonfin).

   a. From NP to infinitival

      (4) They have started a project, to develop which, they had to go around the world.

   b. From NP to gerund-participial.

      (5) She has started a degree, finishing which involves six years of study.

5. From genitive *whose* to NP

   (6) The agent, whose equipment we borrowed, is missing.

6. From determiner-*which* to NP.

   (7) I said that it might be more efficient to hold the meeting on Saturday morning, which suggestion they all enthusiastically endorsed.

Even though Huddleston & Pullum (2002) only mention determiner-*which* in their percolation type 6, they provide an example which shows that this type can be embedded inside a more complex relative phrase which shows percolation type 1 (from NP to PP) at its highest local tree. Their example is given in (8).

(8) He spent all breaks either riding race horses or skying, in which sport he won a European under-18 downhill race. (Huddleston & Pullum, 2002: 1059)

In example (8) the noun *sport* combines with determiner-*which* (percolation type 6). The relative constituent, however, also contains an upward percolation from the complement of a preposition to a PP (percolation type 1).

I investigate here whether it is possible for determiner-*which* to occur with other percolation structures as well. For this purpose, I searched the *Corpus of Contemporary American English*, COCA (Davies, 2008–)[3], to look for instances of these types of upward percolation both with pronoun-*which* and determiner-*which*. I, then, constructed additional examples and asked native speakers of American English about the naturalness of the sentences.

4.2 The questionnaire study

4.2.1 Upward percolation with determiner-*which* vs pronoun-*which*

In this section, I discuss whether determiner-*which* SRCs allow for the same upward percolation patterns that are attested with pronoun-*which*. I will present some examples that I extracted from COCA and show the result of a more systematic study with native speakers. As mentioned above, I intend to study the four general upward percolation types, types 1–4, from Huddleston & Pullum (2002). However, I treated the two subtypes from type 4 (NP to non-finite VP) as distinct types. This resulted in the following five percolation mechanisms, listed here with an example for determiner-*which*.

1. From Comp of preposition to PP (from NP to PP)

   (9) Nurses may also voice concerns about newer electronic documentation

[3]available online at https://www.english-corpora.org/coca/

methods interrupting workflow, in which case they need to become personally involved in workflow design. (COCA)

2. From PP to AdjP

   (10) The many varieties of mammalian skin secretions perform a wide range of functions, prominent among which purposes is sexual attraction. (COCA)

3. From PP to NP

   (11) It took them half a day to reach the nearest village on their mule-drawn carriage, the wheels of which wagon were wooden. (COCA)

4. From NP to infinitival

   (12) In 2017, 13 women were killed in the name of 'honor' in six different Iraqi cities, to condemn which crime the female activists protested for one week in front of UN headquarter. (constructed)

5. From NP to gerund-participial

   (13) They take a rigorous examination, passing which test confers on the students a virtual guarantee of a place at the university. (amended from (Huddleston & Pullum, 2002: 1043))

Before conducting a more systematic exploration of the data, I collected some impressions from individual native speakers. When presenting them with sentences from Huddleston & Pullum's (2002) and some examples from COCA, there were two tendencies: first, examples in which determiner-*which* appeared

inside a PP (percolation type 1) were considered more natural than just having a determiner-*which* NP as relative phrase. Second, examples in which the noun combining with determiner-*which* was abstract or generic, such as (14), were reported as more natural than those with a more specific or concrete nouns, such as (15).[4]

(14) I visited Beau Stubbs on Death Row, where numerous appeals kept him from the gallows for a decade, in which time he used prison to learn a trade and became a highly skilled airline pilot. (COCA)

(15) I was forced at last to open my shirt and show them my stays; which satisfied them very well, for I saw, they believed that I was so locked up in that machine that it was not in my own power to open it, which contrivance they attributed to my husband. (COCA)

Even though I did not built the questionnaire to investigate the second observation, I will address it briefly in the discussion in section 4.2.3.

4.2.2 Method

In this section, I present a questionnaire study that I carried out in 2018 at two universities in the USA: George Mason University in Washington DC and Colorado State University in Colorado. I had 40 consultants who were university students, staff members with at least a university degree, and university teachers. The age range of the participants was between 18 and 57.

The questionnaire had two versions and each participant only received one version of it. Each questionnaire contained 60 items – 40 fillers, and 4 items for

[4]The noun *contrivance* is admittedly low in frequence with only 317 occurrences in COCA, compared with 1,669,166 occurrences of the word *time*, tagged as noun (31.8.2020).

each of the 5 upward percolation mechanisms under investigation. When one item was put with pronoun-*which* in one version, it was put with determiner-*which* in the other version. The versions were counterbalanced in the way that each version contained two items with determiner-*which* and two items with pronoun-*which* for each of the five percolation mechanisms.

An example of an item is shown in (16).[5] The item illustrates the percolation mechanism NP-to-PP. Consultants who saw the pronoun-*which* version were given the relative phrase *during which*. The other consultants who saw the determiner-*which* version had the phrase *during which time*.

(16) I visited Beau Stubbs on Death Row, where numerous appeals kept him from the gallows for a decade, **during which/ during which time** he used prison to learn a trade and became a highly skilled airline pilot.

Most of the items were from COCA. However, because not all of the types of upward percolation are equally common, I had to amend some of the items I extracted from COCA, such as example (17a), to turn them into instances of a particular type of the upward percolation. To turn an example with pronoun-*which* into an example with determiner-*which* I had to add a noun. For example, I changed the attested sentence in (17a) into a sentence with a determiner-*which* RC by adding the noun *cap*, which resulted in (17b).

(17) (from PP to NP)

a. He is wearing a threadbare flannel shirt with a massive blue ink stain beneath the breast pocket, a thrift-store leather jacket, and a fucked-up green trucker's cap, **the color of which** seems to complement perfectly

[5]The highlighting in (16) was not present in the questionnaire.

his hazel eyes, staring out hauntedly from the deep recesses of his unshaven face. (COCA)

b. He is wearing a threadbare flannel shirt with a massive blue ink stain beneath the breast pocket, a thrift-store leather jacket, and a fucked-up green trucker's cap, **the color of which cap** seems to complement perfectly his hazel eyes, staring out hauntedly from the deep recesses of his unshaven face. (COCA amended)

As the items of interest were relatively long and complex, I chose fillers that were of similar length and complexity. One example is given in (18). The list of all items and fillers can be found in Appendix A.1.

(18) While the movie seems simple and amusing in many ways, when approached from the perspective of how it speaks to the role of robots in our society, it, as well, raises some disturbing issues.

The consultants were asked to rate the naturalness of the items in five categories: *very unnatural*, *unnatural*, *I cannot decide*, *natural*, *very natural*. While this instruction was rather vague, the consultants did not seem to have a problem with it.

It is essential for my theoretical modelling in Chapter 5 to find out whether there are any specific constraints on the upward percolation in determiner-*which* RCs. For this reason, I wanted to see whether I can find naturally sounding examples for all of the five upward percolation mechanisms mentioned above.

My hypothesis is that there are no special constraints on the upward percolation for determiner-*which*.

I assume that a given upward percolation mechanism is possible for determiner-*which* RCs in principle if at least one of the two conditions holds: First, at least

half of my consultants rated at least one instance of this pattern as *natural* or *very natural*. Second, there is at least one example for which the corresponding sentence with pronoun-*which* did not receive a much better rating. I consider this second condition to hold if (i) the corresponding sentence with pronoun-*which* was not rated *natural* by at least half of my consultants either, but (ii) both versions of the sentence were judged with the categories *natural* or *undecided* by at least half of my consultants.

In addition to the evaluation of my hypothesis for the five types of upward percolation, I also examine additional properties of the individual sentences, such as properties of the head noun that determiner-*which* combines with or the availability of a corresponding pronoun-*which* example.

4.2.3 Results and discussions

In this section, I present the results of my questionnaire. I received 20 judgments for each of the two versions of my items. As my evaluation is rather coarse grained, I sum up the judgments *natural* and *very natural* into NATURAL and, analogously, I will group the judgments *unnatural* and *very unnatural* into a single category UNNATURAL.

4.2.3.1 From NP to PP

The first upward percolation mechanism is from an NP complement of a preposition to the PP (called here: "from NP to PP"). We saw an instance of this percolation in (8) above. The examples of this percolation type that I included in my questionnaire are given in (19)–(22), together with the judgments for both the determiner-*which* and the pronoun-*which* version. Examples (19) and (20) are

taken directly from COCA, example (21) has been changed slightly, and sentence (22) is a constructed example.

(19) I visited Beau Stubbs on Death Row, where numerous appeals kept him from the gallows for a decade, **during which (time)** he used prison to learn a trade and became a highly skilled airline pilot. (COCA)

determiner-*which*: NATURAL: 15 undecided: 2 UNNATURAL: 3

pronoun-*which*: NATURAL: 12 undecided: 3 UNNATURAL: 5

(20) Nurses may also voice concerns about newer electronic documentation methods interrupting workflow, **in which (case)** they need to become personally involved in workflow design. (COCA)

determiner-*which*: NATURAL: 9 undecided: 6 UNNATURAL: 5

pronoun-*which*: NATURAL: 14 undecided: 2 UNNATURAL: 4

(21) The typical scenario in which employees volunteer on company time is when management sanctions the project. An example was the United Way fall kickoff, **in which (project)** several companies allowed employees to help renovate Amberwood Village Apartments, a rundown complex in east Atlanta's Reynolds town community. (amended from COCA[6])

determiner-*which*: NATURAL: 10 undecided: 5 UNNATURAL: 5

pronoun-*which*: NATURAL: 15 undecided: 1 UNNATURAL: 4

(22) The youngest nephew of Juana Williams won the Jackpot, **with which (money)** she built a hospital in her village. (constructed)

[6]Modified from the following COCA example:

(i) An example was the United Way fall kickoff, **in which area** companies, including IBM, BellSouth, First Union and Northwest Airlines, allowed employees to help renovate Amberwood Village Apartments, a rundown complex in east Atlanta's Reynoldstown community. (COCA)

determiner-*which*:	NATURAL: 9	undecided: 8	UNNATURAL: 3
pronoun-*which*:	NATURAL: 11	undecided: 5	UNNATURAL: 4

In the determiner-*which* variant, examples (19) and (21) are considered NATURAL by at least half of my consultants. For this reason, the hypothesis is met for the percolation (from NP to PP).

The four sentences in this group consisted of two sentences with very general nouns in the relative phrase, *time* in (19) and *case* in (20), and two with more concrete nouns, *project* and *money* in (21) and (22). Both cases included an example that was judged NATURAL by the majority of my consultants and one that wasn't.

The percolation pattern considered in this section shows another important property of determiner-*which* RCs: it is not always possible to replace the determiner-*which*-NP with a simple pronoun-*which*. In my examples, there are probably different antecedents for the RC in sentence (20), depending on the two variants. In the derminer-*which* version, the antecedent of (*in*) *which case* is the voicing of concerns. In the pronoun-*which* version, the antecedent of *which* in *in which* is most likely *workflow*. This makes it impossible to compare the judgments for the two conditions directly.

There are other, even more extreme situations. We find 3,227 occurrences of the sequence *in which case* in COCA (1.9.2020). When used as a relative phrase, it can typically not be substituted with *in which* (or with plain *which*). This is illustrated in (23). Example (23a) is the original COCA hit. There is no synonymous alternative using pronoun-*which*, neither with a preposition or without, see (23b).

(23) a. I'll just be mentally exhausted, **in which case** I'll read something light

like a Newsweek or TIME magazine. (COCA)

b. I'll just be mentally exhausted, ??**in which**/ ??**which** I'll read something light like a Newsweek or TIME magazine.

This suggests that there are cases in which there is no alternative to a determiner-*which* SRC. In the construction of the questionnaire, I could not always ensure that the pronoun-*which* and the determiner-*which* versions were fully synonymous. In those cases, I am not able to compare the observations for the two versions of the sentence.

4.2.3.2 From PP to AdjP

This type of upward percolation is restricted to SRCs. I included the following four items of this type in my questionnaire. The consultants' judgments of each item with pronoun-*which* and determiner-*which* are beneath the items. Example (24) is amended from Huddleston & Pullum (2002: 1042), I added the noun *purposes* to change the pronoun-*which* to determiner-*which*. I constructed examples (25) and (26), and I amended example (27) from COCA by adding the noun *chances* after *which* to change it to determiner-*which*.

(24) The many varieties of mammalian skin secretions perform a wide range of functions, **prominent among which (purposes)** is sexual attraction. (Amended from Huddleston & Pullum (2002: 1042))

determiner-*which*: NATURAL: 10 undecided: 6 UNNATURAL: 4

pronoun-*which*: NATURAL: 9 undecided: 3 UNNATURAL: 8

(25) Up to date, DiCaprio featured in 33 movies, **the best among which (films)** is Titanic. (constructed)

determiner-*which*: NATURAL: 7 undecided: 6 UNNATURAL: 7

pronoun-*which*: NATURAL: 10 undecided: 2 UNNATURAL: 8

(26) Alan Hydros remembers the miseries he has been through while he lived with his biological parents, **the most painful of which (memories)** is when they became homeless for a while. (constructed)

determiner-*which*: NATURAL: 14 undecided: 3 UNNATURAL: 3

pronoun-*which*: NATURAL: 11 undecided: 5 UNNATURAL: 4

(27) During my studies, I became very interested in the field of e-learning. It led me to various other opportunities, **the most exciting of which (chances)** was to work with a local business accelerator to share my story with a new generation of start-ups, helping young entrepreneurs to avoid my mistakes and to communicate best practices. (COCA)

determiner-*which*: NATURAL: 11 undecided: 6 UNNATURAL: 3

pronoun-*which*: NATURAL: 10 undecided: 9 UNNATURAL: 1

At least half of the consultants assessed examples (24), (26) and (27) with determiner-*which* as NATURAL. Similarly, half of the consultants considered three examples NATURAL with pronoun-*which*, namely (25), (26) and (27). I conclude from this that this percolation type is available for determiner-*which* RCs.

4.2.3.3 From PP to NP

I included examples (28)-(31) in the questionnaire as instances of this type of percolation. I was not able to extract examples for this percolation type with determiner-*which* in COCA. Therefore, I constructed examples (28) and (29), and I amended examples (30) and (31) from COCA – the noun *cap* is added after *which* in example (30) and the noun *wagon* is added after *which* in example (31).

(28) They are members of an association, **the first and most precious prin-**

ciple of which (society) is mutual trust. (constructed)

determiner-*which*: NATURAL: 6 undecided: 4 UNNATURAL: 10

pronoun-*which*: NATURAL: 12 undecided: 7 UNNATURAL: 1

(29) Ala asked the HR manager if the boss was planning to dismiss her this month, **the answer to which (question)** I already knew. (Constructed)

determiner-*which*: NATURAL: 7 undecided: 5 UNNATURAL: 8

pronoun-*which*: NATURAL: 0 undecided: 2 UNNATURAL: 18

(30) He is wearing a threadbare flannel shirt with a massive blue ink stain beneath the breast pocket, a thrift-store leather jacket, and a fucked-up green trucker's cap, **the color of which (cap)** seems to complement perfectly his hazel eyes, staring out hauntedly from the deep recesses of his unshaven face. (amended from COCA)

determiner-*which*: NATURAL: 11 undecided: 5 UNNATURAL: 4

pronoun-*which*: NATURAL: 12 undecided: 4 UNNATURAL: 4

(31) It took them half a day to reach the nearest village on their mule-drawn carriage, **the wheels of which (wagon)** were wooden. (amended from COCA)

determiner-*which*: NATURAL: 8 undecided: 6 UNNATURAL: 6

pronoun-*which*: NATURAL: 7 undecided: 6 UNNATURAL: 7

I consider this percolation possible with determiner-*which* in principle, as there is at least one item which was considered NATURAL by half of my consultants, (30).

4.2.3.4 From NP to infinitival

As examples of percolation from NP to infinitival, I included examples (32)-(35) in the questionnaire. Example (34) is modified from Huddleston & Pullum (2002) by adding the noun *judgment* after *which*. The other examples of this type are

constructed because I couldn't find instances of this type in COCA. Huddleston & Pullum (2002: 1043) state that this type of percolation "is rare and very largely confined to purpose adjuncts and catenative complements that are semantically somewhat similar." I tried to model the constructed examples in such a way that they would satisfy Huddleston & Pullum's characterization.

(32) In 2017, 13 women were killed in the name of 'honor' in six different Iraqi cities, **to condemn which (crime)** the female activists protested for one week in front of UN headquarter. (constructed)

determiner-*which*: NATURAL: 5 undecided: 7 UNNATURAL: 8

pronoun-*which*: NATURAL: 11 undecided: 4 UNNATURAL: 5

(33) The sexual assault allegation had shaken Jackson's famous ground, **to repudiate which (claims)** he spent millions of dollars. (constructed)

determiner-*which*: NATURAL: 6 undecided: 5 UNNATURAL: 9

pronoun-*which*: NATURAL: 8 undecided: 2 UNNATURAL: 10

(34) I became disturbed by a 'higher criticism' of bible, **to refute which (judgment)** I felt the need of a better knowledge of Hebrew and archeology. (amended from Huddleston & Pullum (2002: 1043))

determiner-*which*: NATURAL: 6 undecided: 5 UNNATURAL: 9

pronoun-*which*: NATURAL: 8 undecided: 2 UNNATURAL: 10

(35) Jennifer Lawrence won an Oscar last year, **to accept which (award)** she has to be present at the ceremony. (constructed)

determiner-*which*: NATURAL: 6 undecided: 5 UNNATURAL: 9

pronoun-*which*: NATURAL: 8 undecided: 2 UNNATURAL: 10

More than half of the consultants found three of the example of this type of percolation UNNATURAL with both pronoun-*which* and determiner-*which*, namely examples (33), (34) and (35). Only in example (32) half of the consultants found

the example NATURAL only with pronoun-*which*. Even the example from Huddleston & Pullum's (2002), the pronoun-*which* version of (34), was not considered NATURAL by the majority of my consultants. This confirms the assessment of this construction by Huddleston & Pullum as rather rare.

The second criterion from section 4.2.2 allows me to do a careful comparison between the judgments of the two versions of the sentences. In the sentences (33)–(35), both versions of the sentence received a low number of NATURAL judgments. However, for both versions, at least half of my consultants did not consider it UNNATURAL either. According to my criteria, I consider the judgments on determiner-*which* similar enough to the ones on pronoun-*which*. Consequently, the percolation mechanism from NP to infinitival VP seems to be equally available for both relative constituents with pronoun-*which* and determiner-*which*.

4.2.3.5 From NP to gerund/participle

Huddleston & Pullum (2002: 1043) believe this type is not common and highly formal in style. I haven't found any instances of it in COCA. Except for example (36), which is amended from Huddleston & Pullum (2002) by adding the noun *test* after *which*, I have used constructed examples of this percolation type in my questionnaire.

(36) They take a rigorous examination, **passing which (test)** confers on the students a virtual guarantee of a place at the university. (amended from (Huddleston & Pullum, 2002: 1043))

determiner-*which*: NATURAL: 8 undecided: 3 UNNATURAL: 9

pronoun-*which*: NATURAL: 4 undecided: 4 UNNATURAL: 12

(37) The parents of the victims insisted on having a face-to-face discussion with

the mayor, **arranging which (meeting)** was my responsibility. (constructed)

determiner-*which*: NATURAL: 3 undecided: 9 UNNATURAL: 8

pronoun-*which*: NATURAL: 9 undecided: 5 UNNATURAL: 6

(38) Afshin Kaki composed 'the dew drops' when he was 22, **performing which (piece)** at the Grammys in 2015 made him the famous man he is now. (constructed)

determiner-*which*: NATURAL: 11 undecided: 4 UNNATURAL: 5

pronoun-*which*: NATURAL: 11 undecided: 2 UNNATURAL: 7

(39) The employees had to write a proposal for a small project of their choice, **refusing which (task)** would decrease their salaries. (constructed)

determiner-*which*: NATURAL: 10 undecided: 3 UNNATURAL: 7

pronoun-*which*: NATURAL: 11 undecided: 2 UNNATURAL: 7

At least half of the consultants rated examples (38) and (39) NATURAL with both pronoun-*which* and determiner-*which*. Consequently, I consider this percolation type to be available for relative phrases with determiner-*which*.

4.2.4 Summary of the results

To sum up, each of the discussed percolation mechanisms should be considered to be possible for determiner-*which* RCs in principle. In all cases, except for the percolation from PP to an infinitival VP, this result could be achieved directly as the questionnaire contained examples that were considered NATURAL by at least half of my consultants. In the problematic case, I observed that the versions of the items with pronoun-*which* did not receive a much better rating than the corresponding versions with determiner-*which*. Since this percolation mechanism is among those that Huddleston & Pullum (2002) classify as rare and highly

restricted in their occurrences, this result is not surprising. It follows, then, that I could not find evidence that any of the upward percolation mechanisms discussed in Huddleston & Pullum (2002) is available for pronoun-*which* SRCs but not for determiner-*which* SRCs.

4.3 Conclusion

The questionnaire study presented in this chapter had a number of shortcomings. An important reason for this lies in the generally problematic type of the data under investigation.

As indicated in the presentation of the data used in my questionnaire, I was not able to find instances of some of the percolation patterns in the COCA corpus. This shows that it might not be enough to look at corpus occurrences to show that the percolation mechanisms are possible with determiner-*which* in principle.

Similarly, even examples for pronoun-*which* taken directly from Huddleston & Pullum (2002) were not necessarily considered NATURAL by my consultants. This shows that, in this area of the grammar of English, even carefully crafted examples from reference grammars are not unanimously considered natural by native speakers. Given this situation, I tried to provide a larger pool of possible examples and native speaker judgments to get a clearer picture about which types of complex relative constituents should be licensed in SRCs by the grammar of English.

It is known that not all non-local features of Pollard & Sag's (1994a) behave the same. van Eynde (2004) shows in some detail that interrogative constituents in English are much more restricted than what we find for relative constituents. For example, there is no upward percolation of the type "from PP to NP" for

interrogative phrases, see (40a). A corresponding example with a homophonous relative constituent, such as (40b) is, however, grammatical.

(40) a. *I do not know [[(the) friends of whom] they have invited].

(van Eynde, 2004: 22)

b. I know Alex, [[(the) friends of whom] they have invited].

c. I know Alex, (the) friends of which student they have invited.

It could have been the case that we find a more restricted behaviour for determiner-*which* relative phrases, which would, then, require additional constraints - or, in analogy to the analysis in van Eynde (2004), even a new feature. However, as shown in the results of my study, the relevant percolation mechanism is available for determiner-*which* relative phrases, see also (40c).

In the discussion of the first percolation mechanism, "from NP to PP," we saw that the use of a determiner-*which* SRC allows a clearer identification of the antecedent, and, in some cases, allows for abstract antecedents that would not be available for a pronoun-*which* relative constituent. This can be found with uses of determiner-*which* combining with head nouns of a very general, abstract meaning, such as *in which case*, *during which time*, …. This shows that determiner-*which* SRCs enrich the possibility of attaching an RC to a clause in English. There is, consequently, some functional motivation for determiner-*which* SRCs in Present Day English. I will not be able to pursue this line of thought further in the current thesis.

This chapter provides the basis for my formal analysis of determiner-*which* SRCs in chapter 5. The results that show that there is no difference in the upward percolation mechanisms between determiner-*which* and pronoun-*which* makes it possible to use the same percolation mechanism for both cases, which will just be

the standard REL percolation of Pollard & Sag's (1994a).

Chapter 5

Determiner-*which*: A syntactic and semantic analysis

In this chapter, I provide determiner-*which* with a HPSG analysis by mostly adopting from Pollard & Sag (1994a). I start by showing the lexical entry of determiner-*which* that I gave in Figure 5.1. This chapter consists of two main parts: first, in section 5.1, for the syntactic analysis, I present the six types of percolation with determiner-*which* that I discussed in the previous chapter, chapter 4, and sketch their HPSG analyses. Second, I provide a semantic analysis using the framework *LRS* presented in chapter 3, section 3.4.

Before presenting HPSG and LRS analyses of determiner-*which*, it is important to see what the lexical entry of determiner-*which* in HPSG looks like. In Figure 5.1, I present the lexical entry of determiner-*which*, which is based on the lexical entry of determiners and the lexical entry of relative pronouns that are given in Pollard & Sag (1994a).

In Figure 5.1, the HEAD value is a determiner. It selects a noun, namely N', and its VAL lists are empty. Its CONTENT is identical with its RESTR. Finally, it

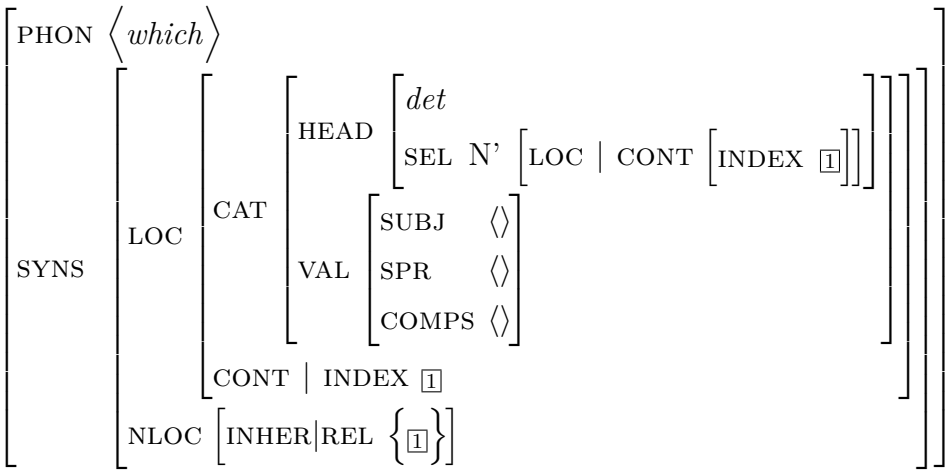


Figure 5.1: Description of determiner-*which*

has a REL value in its NLOC|INHR, and the value of its INDEX is the same as its REL value.[1]

In analysing *wh*-RCs, I assume a null- relativizer[2] which functions as the head of the RC. The core clause of the RCs, which contains the INHR|SLASH value, functions as the complement of the null-relativizer. The relative phrase containing the INHR|REL value is the subject of the null-relativizer. The null-relativizer, first, combines with the complement and then with the subject to construct the RC.

5.1 An HPSG analysis of upward percolation with determiner-*which*

Concerning upward percolation with RCs, the percolation of the REL value plays a vital role. I presented the REL percolation mechanisms of Pollard & Sag (1994a) and Sag (1997) in Chapter 3, Section 3.3. As my analysis is primarily based

[1]The lexical entry in Figure 5.1 is a simplification. In particular, it does not reflect the split of the INDEX value used in LRS, which I will employ in my analysis as well. In particular, I will assume that, in a SRCs, a relative word shares its PHI value with the element in its INHER|REL set, but not its DR value. This will be shown once the LRS-specific attributes are added in sections 5.3 and 5.4.

[2]A detail on this is in chapter 3, section 3.3.3.

on Pollard & Sag (1994a); I will adopt their REL mechanism throughout for my analysis of determiner-*which*.

Huddleston & Pullum (2002: 1040) introduced six types of upward percolation as mentioned in chapter 2 and chapter 4. Here, I illustrate these types of percolation with determiner-*which* and sketch their HPSG analyses. I present the types in a different order here to show the graduate embedding of the determiner-*which* inside the relative constituent.

Upward percolation from determiner-*which* to NP Huddleston & Pullum (2002: 1043) state that when *which* functions as a determiner, and not a pronoun, upward percolation is mandatory.

(1) I was informed yesterday that I had even been appointed the HR department, so that I could go home early, which appointment I declined. (COCA)

(2) Metadorus argues otherwise, claiming a universe populated by numberless worlds since their causes must be numberless, which contention we reject as sophistry. (COCA)

Figure 5.2 is the HPSG analysis of the RC in example sentences (1) and (2). The sentence containing the gap combines with the null-relativizer first, and then percolates up the tree. This is where the SLASH value binds off with the value of the LOCAL at the left side of the tree as their values are identical. In other words, the null-relativizer first combines with its complement, the S which contains the SLASH value, to form a R′. The R′ combines with its subject, the NP which has a INHER|REL{$\boxed{1}$}, to form a RP.

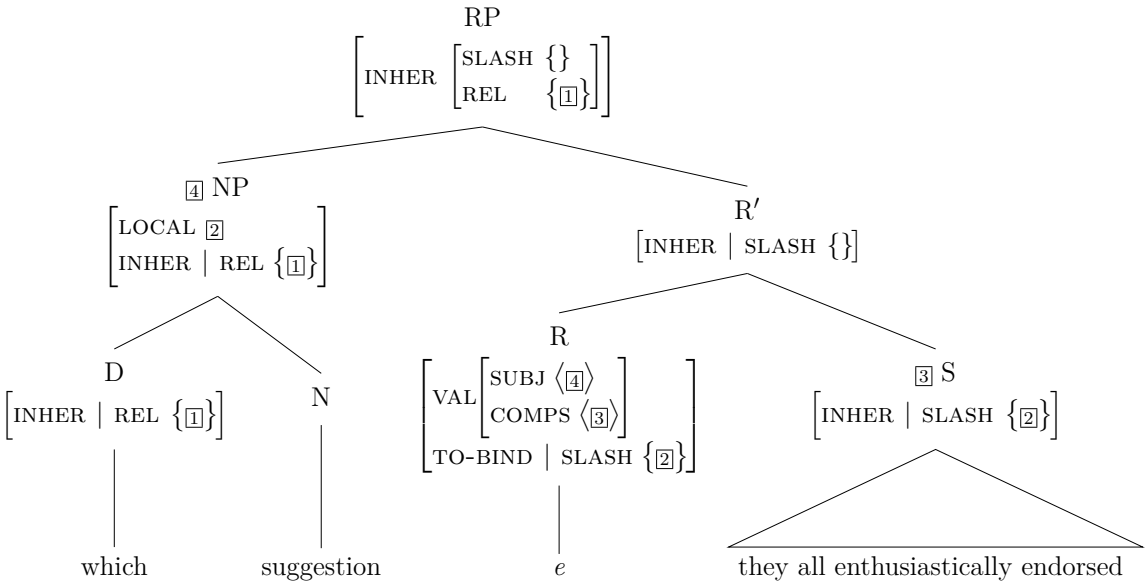


Figure 5.2: The analysis of *which appointment I declined* given in example sentence (1)

As for the subject of the null-relativizer, the head noun combines with its determiner; namely the relative pronoun-*which*, which percolates the REL value up the tree to the NP above and even higher to the S. The NON-LOCAL FEATURE PRINCIPLE makes sure that the REL value percolates up the tree.

From Comp of a preposition to a PP The first type of percolation showed how the REL value percolates from a specifier daughter to the mother. The second type is another instance of percolation from a non-head daughter to the mother, but from a complement daughter, in this case. For comparison, I first show an HPSG analysis for an example of a RC with this type of percolation that contains a relative pronoun-*which* and later I discuss cases with determiner-*which*.

(3) Chavez and the mayor had a long history, of which I knew just enough to not want to know more. (COCA)

Figure 5.3 is the HPSG analysis of the RC in example sentences (3).

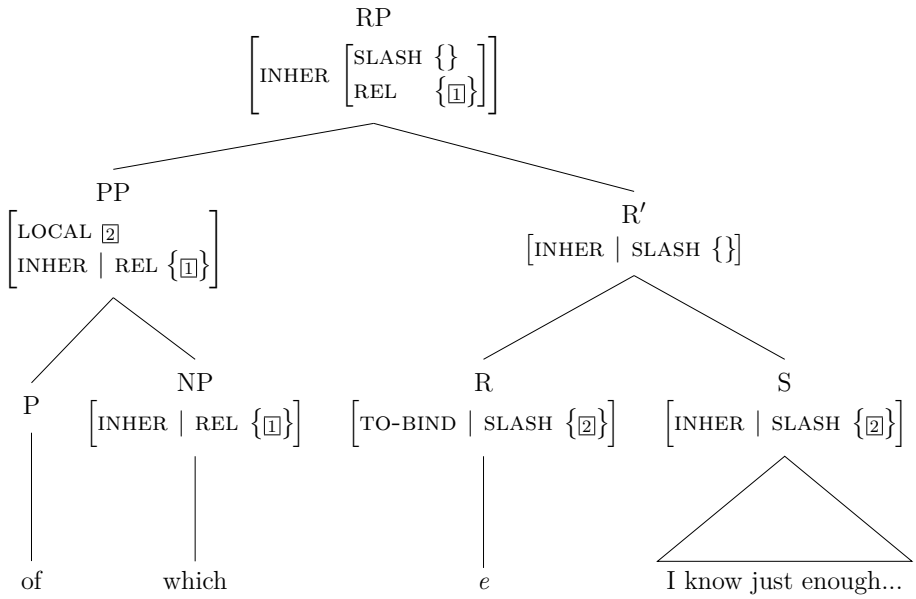


Figure 5.3: The analysis of *of which I know just enough not to want to know more* given in example sentence (3)

At the lower left, the pronoun-*which* combines with its head P *of* to form a PP that carries the REL value of pronoun-*which*. Similar to the tree in Figure 5.2, at the lowest level of the tree on the right side, the null-relativizer first combines with its complement, the S which contains the SLASH value, to form a R′. The R′ combines with its subject, the PP which has a INHER|REL{1}, to form a RP

(4) Maybe people will attack me, in which case I will attack back. (COCA)

Figure 5.4 is the HPSG analysis of the RC in example sentences (4).

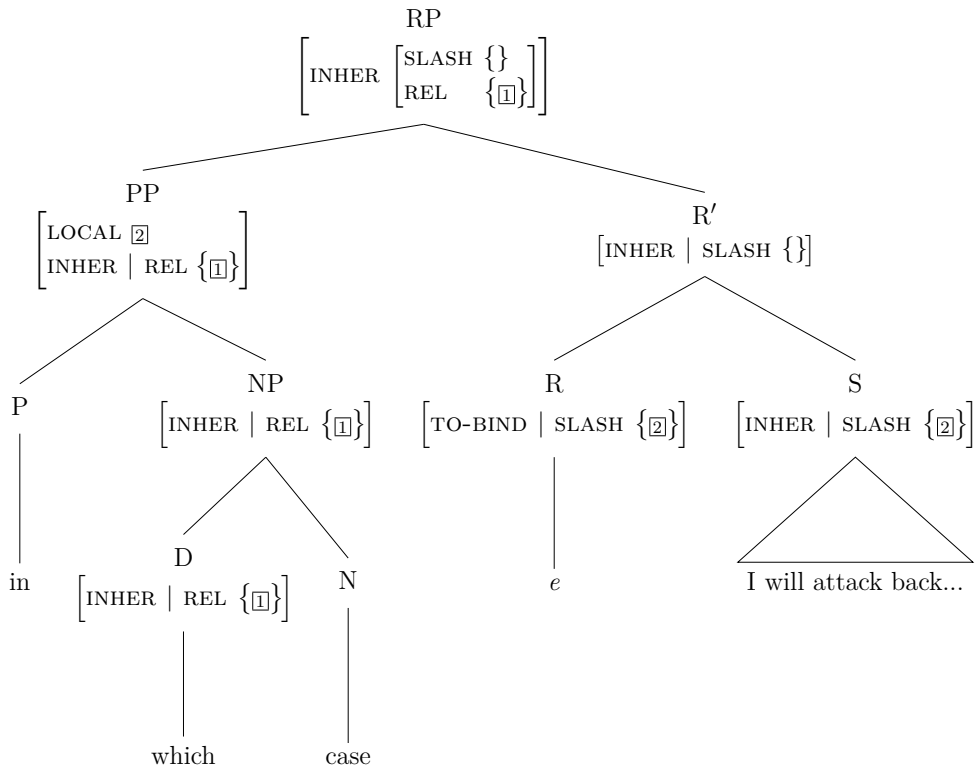


Figure 5.4: The analysis of *in which case I will attack back* given in example sentence (4)

In Figure 5.4, within the fronted constituent, the noun *case*, first, combines with its specifier *which* to form a NP. In this local tree, REL percolates up the tree as illustrated in Figure 5.4. The NP then acts as a complement to the preposition *in* to form a PP. In this local tree, again the REL value percolates from the non-head daughter to the mother. As in example (3), the REL value percolates further up to the RP-node.

From now on, for the analyses and the HPSG trees of further examples of the types of upward percolation, I concentrate only on the details on the left side of the trees as the right sides are all identical in all relevant respects.

From PP to AdjP In the next percolation pattern, the REL value also percolates from a complement daughter to a HEAD, but, in this type, it is from the complement-PP *among which achievements* to the AP headed by the Adj *prominent*. In example (5), determiner-*which* is embedded inside the relative constituent yet one level deeper.

(5) She received many awards in her life, prominent among which achievements are a golden globe and two academy awards. (constructed)

Figure 5.5 is the HPSG analysis of the RC in the example sentences (5).

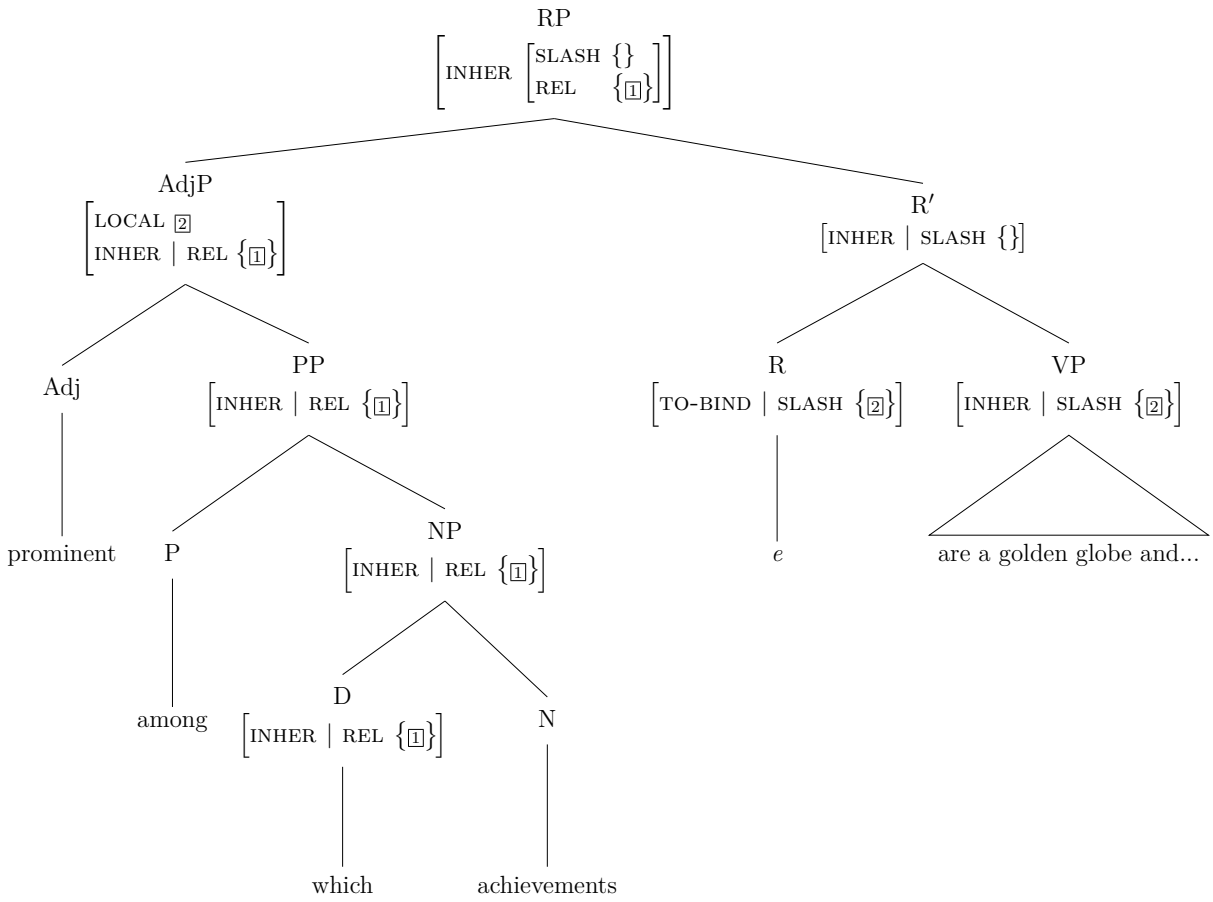


Figure 5.5: The analysis of *prominent among which achievements are a golden globe and two academy awards* given in example sentence (5)

On the lower left, the N, *achievements*, is combined with its specifier D, *which*, to form an NP which is combined with its head P, *among*, to form a PP. The PP is combined with its head Adj, *prominent*, to form an AdjP which carries the REL value of the determiner-*which*.

From a PP to a NP In the fourth case to be considered, the highest node in the relative constituent is not an AdjP but an NP, and apart from this, the case is analogous to the previous one. I did not find instances of percolation from a PP to a higher NP in COCA. I modified an example from COCA to fit this percolation pattern for my questionnaire, and I use this modified example for illustration here. Note that it had been rated as relatively acceptable in my study, see chapter 4, section 4.2.3.

(6) He is wearing a threadbare flannel shirt with a massive blue ink stain beneath the breast pocket, a thrift-store leather jacket, and a fucked-up green trucker's cap, the color of which cap seems to complement perfectly his hazel eyes, staring out hauntedly from the deep recesses of his unshaven face. (COCA amended)

Figure 5.6 is the HPSG analysis of the RC in example (6).[3]

[3]As mentioned in section 3.3.2, I assume non-empty SLASH value for extracted subjects, following Levine & Hukari (2006).

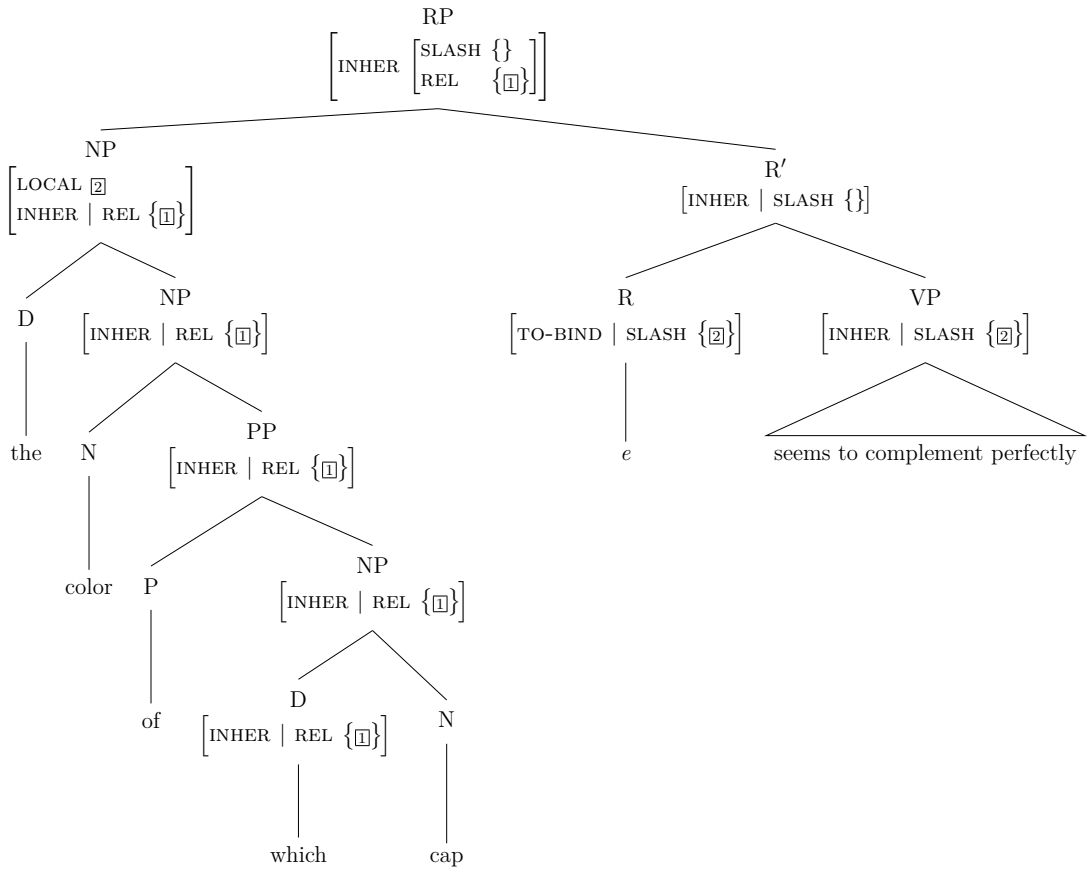


Figure 5.6: The analysis of *the color of which cap seems to complement perfectly* given in example sentence (6)

At the of the relative phrase, the percolation inside the PP of *which cap* is as shown in Figure 5.6. The noun *cap* is combined with its specifier D, *which*, to form a NP, which is combined with the preposition *of* to form a PP. This PP inherits the REL specification. The PP combines as a complement daughter with the noun *color*. The resulting nominal projection inherits the REL value from the non-head. In the next step, the determiner is added. In the highest NP tree, the REL specification is now percolated from the head to the phrase. Finally, as in all of the previous examples, the REL value percolates from the relative phrase to the overall RC. Example (6) shows REL percolation along five local trees. It contains

the only instance of REL percolation from the head daughter to the mother so far.

From an NP to an infinitival VP In this type of percolation pattern, the REL value also percolates from a complement daughter, a base VP, to a head, an infinitival VP.

I did not find many instances of percolation from an NP to an infinitival VP in COCA. Still two of the examples I presented in chapter 4, section 4.2, are from COCA. I constructed two more examples which fit this percolation pattern for my questionnaire. I use one of the constructed examples for illustration here. None of the examples of this type were rated natural in the questionnaire study.

(7) The sexual assault allegations had shaken Jackson's famous ground, to refute which accusations he spent millions of dollars. (constructed)

Figure 5.7 is the HPSG analysis of the RC in example sentences (7). At the lower left, the N, *accusations*, is combined with its specifier D, *which*, to form an NP. This NP combines with the non-finite verb, *refute*, to form a VP. The VP inherits the REL value from its complement NP. This VP then combines with the infinitival marker *to* as a complement daughter. The REL specification of determiner-*which* percolates to this infinitival VP.

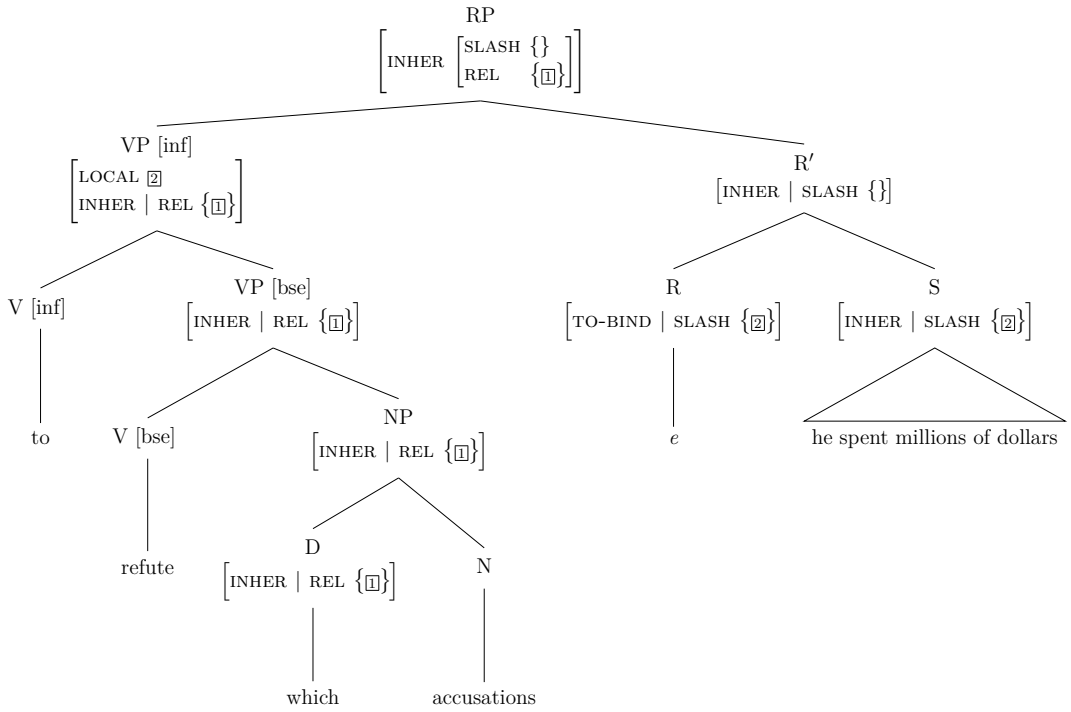


Figure 5.7: The analysis of *to refute which accusations he spent millions of dollars* given in example sentence (7)

From an NP to gerund-participial In this type of percolation pattern, the REL value percolation is similar to the one above with an infinitival VP, except that here the REL value percolated from a complement NP daughter to a head, a gerund VP. Similarly to the cases with infinitival VPs, not many instances of this type can be found in COCA. In chapter 4, section 4.2.3, there are four instances of this type, two of them are from COCA and the other two are constructed. One constructed instance is analysed here. Similar to instances of the infinitival VPs, this type of percolation was also not rated as natural by the participants.

(8) Every knight had to go on a quest, passing which test made him eligible for a place on King Arthur's round table. (constructed)

Figure 5.8 is the HPSG analysis of the RC in example sentences (8).

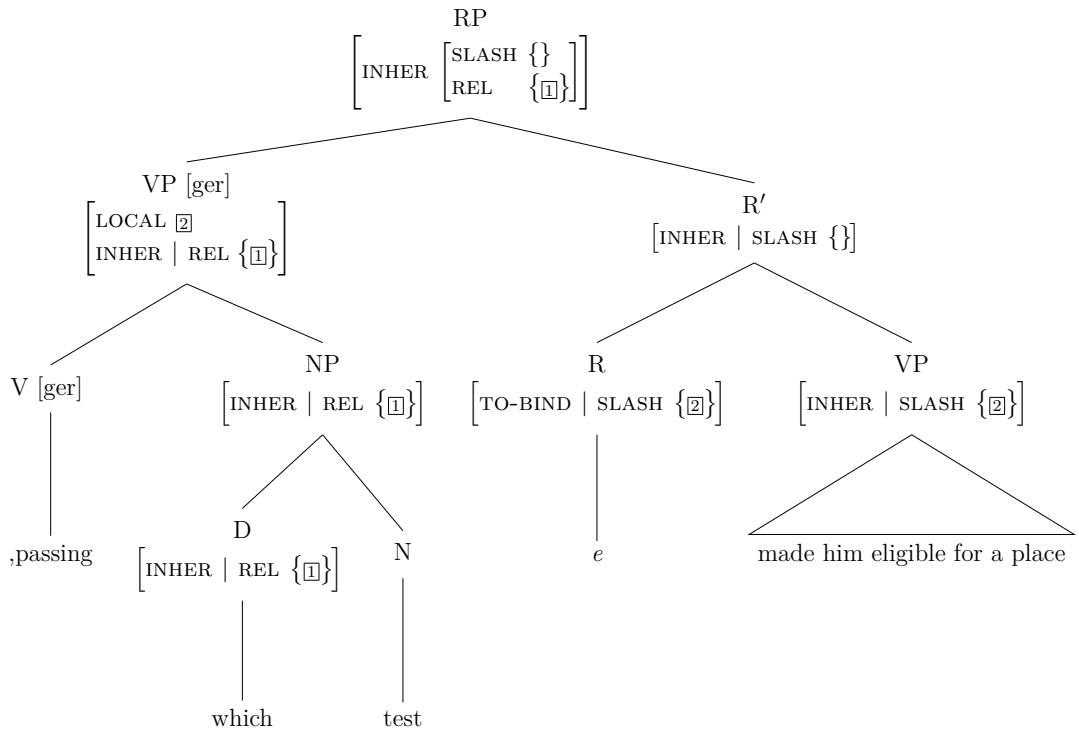


Figure 5.8: The analysis of *passing which test made him eligible for a place* given in example sentence (8)

At the lower left, the N, *test*, is combined with its specifier D, *which*, to form an NP. The NP is combined with its head, a gerund VP, *passing*. The NP and this verb are combined to form a gerund VP that still carries the REL value of the determiner-*which*.

5.2 Semantics of determiner-*which*

In 5.2.1, I introduce the semantic representations of SRCs based on Arnold (2004) and (2007). In 5.2.2, I present properties of determiner-*which* NPs and show how they do not make an at-issue contribution. Additionally, I argue that determiner-*which* NPs have the status of CIs. In 5.2.3, I look at larger relative constituents,

i.e. those in which a more complex upward percolation of the REL feature can be observed. In 5.3, I provide an analysis of SRCs which combines the classical syntactic HPSG analysis from Pollard & Sag (1994a) with LRS. In 5.4, I argue that the semantics of determiner-*which* should be similar to that of other determiners and pronoun-*which*.

5.2.1 Semantic representation of SRCs

Arnold (2004, 2007) uses *Discourse Representation Theory* (DRT, Kamp & Reyle (1993)) in analysing RCs. He provides *discourse representation structures* (DRSs) for SRCs and IRCs. Here, I illustrate the semantic representations proposed by Arnold using his examples as given in (9) and (10), whose DRSs are shown in Figures 5.9 and 5.10.

(9) Sam thinks a person who I detest will win. (Arnold, 2007: 295)

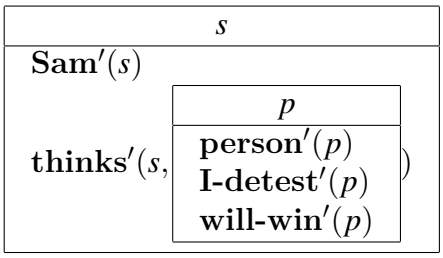


Figure 5.9: DRS of sentence (9)

The DRS in 5.9 says that s, who is an individual, thinks that a person p who is detested by me will win. As can be seen, the content of the IRC *who I detest* appears in the same sub-DRS as the material from clause containing its antecedent.

This contrasts with Arnold's analysis of SRCs. According to him, SRCs have *global-scope* semantics, i.e., they are interpreted similarly to independent clauses

in discourse, unlike IRCs , which have *intersective* semantics.

(10) Sam_s thinks Kim_k, whom I detest, will win. (Arnold, 2007: 295)

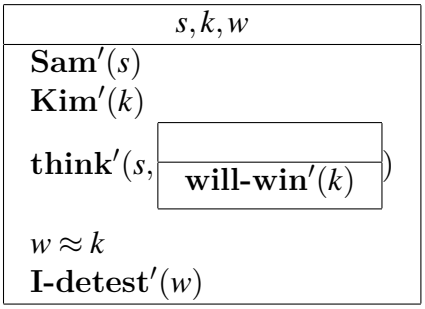


Figure 5.10: DRS of sentence (10)

The representation of example (10) in Figure 5.10 means that there are discourse referents s, k, and w who are all individuals, such that s thinks that k will win; I detest w and w is referentially dependent on k ($w \approx k$). As the SRC content is integrated into the top box, it does not occur in the scope of the intentional predicate *think*. We saw above, in (9), that a corresponding IRC will invariably occur in the scope of *think* if its antecedent does.

In the following, I use a predicate-logic version of Arnold's analysis that I use in my LRS modelling of the syntax-semantic interface as presented in section 3.4. The reason for diverging from his analysis is more conceptual than empirical in nature: Versions of predicate logic are a more common semantic representation language than DRSs, both in general and in LRS publications.[4]

In (11), I provide a straightforward predicate-logic rendering of Arnold's analyis of sentence (9).

(11) Sam thinks that a person who I detest will win. (= (9))

$\textbf{think}'(\textbf{sam}', {}^{\wedge}[\exists p : (\textbf{person}'(p) \wedge \textbf{I-detest}'(p))](\textbf{will-win}'(p)))$

[4]With the exception of Sailer (2007, 2009), where an integration of DRSs into an LRS combinatorics is pursued.

As this formula shows, I use name constants instead of name predicates. This has the purpose of keeping the semantic representations as simple as possible. In (11), the semantic contribution of the RC is integrated into the restrictor of the existential quantifier, together with the semantic contribution of its antecedent, $\mathbf{person}'(p)$. As such, the RC is embedded in the scope of the intensionality operator "$^\wedge$".

An analogous translation of the DRS in Figure 5.10 would look as in (12).

(12) Semantic representation of (10) (preliminary):

Sam thinks that Kim, who I detest, will win. (= (10))

$\mathbf{think}'(\mathbf{sam}', {}^\wedge\mathbf{will\text{-}win}'(\mathbf{kim}')) \wedge \mathbf{I\text{-}detest}'(\iota w : w = \mathbf{kim}')$

In the semantic representation in (12), the semantic contribution of the SRC appears outside the intensionality operator. The SRC contributes a discourse referent, w. This discourse referent is identified with the antecedent, $\mathbf{kim}'$. This is the essential part of the semantic analysis of SRCs in Arnold (2004, 2007) as not being semantically integrated.

My treatment of the relation between the antecedent and the relativized element in an SRCs is, just as in Arnold's proposal, identical to an analysis of coreference in discourse by pronouns. Just like in his approach, and in DRT in general, I assume that the discourse pronoun introduces a new variable whose value is identified with the discourse referent of the antecedent. As I use a definite description for the new discourse referent, this can be considered a version of the analysis of pronouns in Elbourne (2001) in which they are phrases consisting of a definite article and a deleted copy of the antecedent.

In the semantic representation in (12), the relative pronoun is depicted as simply contributing the identity with its antecedent. This is a simplification as

it ignores the contrast between *who* and pronominal *which*. We should, therefore, include the restriction to human – or at least animate – individuals to the semantic contribution of *who*. For pronoun *which*, we could either omit such a restriction or require that the antecedent be not human. I provide the enriched semantic representation in (13).

(13) Semantic representation of (10) (final):

Sam thinks that Kim, who I detest, will win. (= (10))

$\mathbf{think}'(\mathbf{sam}', {}^{\wedge}\mathbf{will\text{-}win}'(\mathbf{kim}'))$

$\wedge \mathbf{I\text{-}detest}'(\iota w : (\mathbf{human}'(w) \wedge w = \mathbf{kim}'))$

From the semantic representation in (12) we can extract the semantic contributions of the individual parts, see (14).

The relative pronoun *who* is translated as a definite description for the one individual that is identical with the antecedent, see (14b). In this simple sentence, the relative constituent is identical with the relative pronoun. Consequently, it has the same semantic contribution, see (14c). The core of the SRC, *I detest* __ contains the semantics of the moved constituent, see (14d). Finally, the empty SRC-relativizer will only contribute a coordination, as shown in (14e).

(14) a. Main clause: $\mathbf{think}'(\mathbf{sam}', {}^{\wedge}\mathbf{will\text{-}win}'(\mathbf{kim}'))$

b. Relative pronoun: $(\iota w : w = \mathbf{kim}')$

c. Relative constituent: $(\iota w : w = \mathbf{kim}')$

d. Core RC: $\mathbf{I\text{-}detest}'(\ldots)$

e. Empty relativizer: $\ldots \wedge \ldots$

f. Entire SRC: $\ldots \wedge (\mathbf{I\text{-}detest}'((\iota w : w = \mathbf{kim}')))$

After this characterization of the semantic representations of SRCs in general,

I present and motivate the representation I assume for determiner-*which* SRCs. I use example (15), which is a simplified version of an example from Arnold (2004: 45). In it, the relative phrase consists just of the determiner-*which* NP *which beer*, i.e., it has the simplest possible structure. I discuss more complex cases later in this section.

(15) She refused a beer, which beverage she never touches.

The semantic representation that I will assume for this sentence is given in (16). It admittedly ignores many aspects of the sentence, but it includes the information needed to illustrate the semantic combinatorics of determiner-*which* RCs. For simplicity, I translate the pronoun *she* as a free variable, y.

(16) $[\exists x : \mathbf{beer}'(x)](\mathbf{refuse}'(y,x)$
$\wedge \neg \mathbf{touch}'(y, (\iota w : (\mathbf{beverage}'(w) \wedge w = x))))$

This formula expresses the following truth conditions: The sentence is true if and only if there exists x, such that x is *beer* and y refuses x and, furthermore, y does not touch the w which is a beverage and identical with x.

In (17), I indicate the semantic contribution of the relevant parts of the sentence: the main clause, the noun *beverage*, determiner-*which*, the relative constituent, the core part of the RC, the null-relativizer, and the entire SRC.

(17) a. Main clause: $[\exists x : \mathbf{beer}'(x)](\mathbf{refuse}'(y,x)\ldots)$

b. Head noun in the relative constituent: $\mathbf{beverage}'(w)$

c. Determiner-*which*: $\iota w : (\ldots \wedge w = x)$

d. Relative constituent: $\iota w : \mathbf{beverage}'(w) \wedge w = x)$

e. Core part of the RC: $\neg \mathbf{touch}'(y, \ldots)$

f. Relativizer: $\ldots \wedge \ldots$

g. Entire RC: $\ldots \wedge \neg \mathbf{touch}'(y, (\iota w : \mathbf{beverage}'(w) \wedge w = x))$

Most parts in (17) are analogous to what we found with the relative pronoun *who* as in (14). This is also true for *which beverage* in (17d), which contains the restriction to beverages and the identity link to the antecedent of the SRC, just as the relative constituent in (14c). The only difference lies in the fact that the relative constituent is composed of a determiner and a noun.

In this subsection, I provided the basics for my semantic analysis of SRCs. I took the analysis of Arnold (2004, 2007) as my starting point, reformulated it in a predicate-logical semantic representation language and used an encoding of pronouns proposed in Elbourne (2001). Finally, I illustrated the semantic representation of a determiner-*which* SRC. In the following section, I take a closer look at the semantic/pragmatic status of the *which*-NP in determiner-*which* SRCs.

5.2.2 Semantic properties of determiner-*which* NPs

In this section, I argue that the semantic/pragmatic status of a determiner-*which* NP is different from that of the rest of its SRC. I showed in section 2.5 that, while Potts (2005) considers SRCs as non-at-issue that do not interact with the at-issue content, AnderBois et al. (2015) provide evidence that there is indeed some interaction and that sentence-final SRCs can be interpreted at issue. I will argue in this section that the semantic contribution of a determiner-*which* relative constituent is *never* at-issue. Here, I propose that it is a Conventional Implicature (CI).

I use mainly example in (18) for illustration. The core SRC has the inference that Kim will read LGB soon, see (18a). The determiner-*which* NP comes with the inference that LGB is a book about syntax, (18b).

(18) Alex read *Lectures on Government and Binding*, which book about syntax Kim will be reading next.

a. Inference of the core SRC: Kim will be reading *LGB* next.

b. Inference of the relative constituent: *LGB* is a book about syntax.

I first show that the descriptive content of determiner-*which* relative constituent cannot be at issue. It is essential for this argument to use sentences in which the SRC can be interpreted as at issue. For this reason, I will only use SRCs in clause-final position and I provide contexts which make the at-issue status of the SRC clear. I will show that, while the inference of the core SRC as in (18a) can be rejected, the one of the determiner-*which* NP as in (18b) cannot.

In the mini-discourses in (19), a sentence like (18) is uttered by the first speaker. In (19a), the second speaker rejects the content of the core RC. This is a natural continuation, pointing to the at-issue status of an inference like (18a).

(19) a. A: Alex read *LGB*, which book about syntax Kim will be reading next.
B: No, that's not true – Kim will read *Barriers* next.

b. A: Alex read *LGB*, which book about semantics Kim will be reading next.
B: # No, that's not true – *LGB* is about syntax, not so much about semantics.

In (19b), however, the second speaker rejects the content of the relative constituent, i.e. an inference like (18b) that the antecedent of the SRC is a book about semantics. This is not a felicitous continuation. This shows that this part of the meaning of the entire SRC has a status difference from that of the rest. In particular, it cannot be at issue, even in a context in which the rest of the SRC can be interpreted at issue.

In the next step, I show that the inference related to the descriptive content of the determiner-*which* NP has properties associated with CIs.

First, CIs are never in the scope of negation, quantifiers, or intensional verbs. This is hard to test as SRCs outscope such operators from the matrix clause as well, which was one of the arguments for the orphanage approach. However, SRCs have their own illocutionary force, which can be different from that of the matrix clause. This is shown in the examples in (20). In the first example, the SRC is an imperative, in the second, it is a question.

(20) a. He said he'd show a few slides towards the end of his talk, at which point please remember to dim the lights.

b. She may have her parents with her, in which case where am I going to sleep?

(Huddleston & Pullum, 2002: 1061)

If there is a question in the SRCs, we can test whether the content of the determiner-*which* NP is part of what is being questioned. This is illustrated in (21). In my paraphrase in (21a), both the content of the relative phrase and the core of the SRC are in the scope of the interrogative operator. As *LGB* is not written by Joan Bresnan, the utterance in (21) should be rejected if (21a) was a possible reading. However, (21a) is not a paraphrase of the sentence, as indicated by the marking "≠". We arrive at a paraphrase of the sentence if we take the content of the determiner-*which* NP for granted and place the content of the core SRC in the scope of the interrogative operator.

(21) Alex read *LGB*, which book by Bresnan you talked about recently, didn't you?

a. ≠ I wonder if [*LGB is a book by Bresnan* and you talked about it recently].

b. =*LGB* is a book by Bresnan and I wonder if you talked about it recently.

This shows that the descriptive content of a determiner-*which* NP cannot be in the scope of an interrogative operator.

Next, we want to test whether the content of the determiner-*which* NP can be in the scope of negation. Again, we know that an SRC cannot be in the scope of negation in the matrix clause. In (22), I show that a negation in the SRC cannot take scope over the content of the relative constituent.

(22) Alex read *LGB*, which book by a famous semanticist Kim won't read next.

a. It is not the case [that Kim will read *LGB* next and that *LGB* is a book by a famous semanticist].

b. *LGB* is a book by a famous semanticist and it is not the case that Kim will read *LGB* next.

In (22a), I provide a hypothetical paraphrase in which the negation would outscope both the content of the core SRC and that of the determiner-*which* NP. This is not a possible reading of the sentence. If it were, the SRC would always be true since the second conjunct in the scope of the negation is false. The reading of the SRCs is more correctly displayed in the paraphrase in (22b). Here, the negation only takes scope over the core clause, but not over the relative constituent.

This shows that the determiner-*which* NP cannot be interpreted in the scope of a negation in the SRC.

The next property of determiner-*which* phrases to be tested is whether they can be interpreted in the scope of belief predicates. The example in (23) has the same structure as the one before. The SRC contains the relevant operator, here the predicate *believe*. As shown in (23a), there is no interpretation in which it would

be enough if Sam believes that *LGB* is a semantics book. Rather, the sentence is true even if Sam knows that *LGB* is a syntax book and just believes that Kim will read it next, as long as the speaker commits to *LGB* being a semantics book, see (23b).

(23) Alex read LGB, which book about semantics Sam believes Kim will be reading next.

a. ≠ Sam believes [that Kim will be reading *LGB* next and that *LGB* is a book about semantics].

b. = *LGB* is a book about semantics and Sam believes [that Kim will be reading *LGB* next].

Again, the content of the determiner-*which* NP behaves like a CI in the sense that it projects out of *believe* contexts.

There is one context out of which CIs need not project: the scope of speech act operators. This has been shown in Bach (1999).

A speaker can truthfully report John's statement from (24a) as in (24b), even if the speaker knows that Bresnan is not the author of *LGB*. This shows that the content of the determiner-*which* NP need not scope out of indirect speech contexts. This is exactly the one context which allows for an embedded interpretation of CIs.

(24) a. John: "Mary liked LGB, but I hated this book by Bresnan."

b. "John said that Mary liked LGB, which book by Bresnan he hated."

We have observed that determiner-*which* phrases have the semantic properties attributed to CIs in the literature: They are not at-issue, they cannot be questioned, they project out of negation and belief contexts, but they need not project out of indirect speech contexts.

I would like to show in the rest of this subsection that the semantic representation I proposed for determiner-*which* SRCs in (16) captures these observations. To illustrate this, I provide the semantic representation of a sentence similar to (18) in (25).

(25) Alex read LGB, which book Kim will be reading next.

$$\mathbf{read}'(\mathbf{alex}', \mathbf{lgb}') \wedge \mathbf{read\text{-}next}'(\mathbf{kim}', (\iota w : (\mathbf{book}'(w) \wedge w = \mathbf{lgb}')))$$

The content of the determiner-*which* NP appears inside the ι-operator. The ι-operator is used as usual, i.e. $[\![\iota x : \phi]\!]$ refers to the unique individual that satisfies ϕ, if there is such a unique individual, and is undefined otherwise. This means that the formula ϕ is not part of the at-issue content but independent thereof. The way the ι-operator is defined, it takes the form of a presupposition rather than a CI. However, I side with Coppock & Beaver (2015) and, within HPSG, Sailer & Am-David (2016) in assuming that at least the uniqueness aspect of the formula inside the ι-operator is a CI.

5.2.3 Semantic properties of complex relative phrases

So far, we have only looked at relative phrases that have the form of a determiner-*which* NP. However, we need to consider more complex relative phrases as well, since I have shown in Chapter 4 that there are no clear restrictions on the formation of complex relative phrases containing determiner-*which*. In this section, I will show that the CI-status only applies to the determiner-*which* NP, independent of how deeply it is embedded inside a relative constituent. I will first look at cases of pronoun-*which* and then turn to determiner-*which* SRCs.

It needs to be kept in mind, though, that a RC has two distinct unbounded dependencies: The relation between the gap and the relative constituent and

the relation between the relative pronoun and its antecedent. The first UDC is modelled by the SLASH-mechanism, the second by the REL-mechanism. Whenever the relative constituent consists just of a relative pronoun, as in (12), these two dependencies concern the same element. In other words, in such a case, the gap inside the RC refers to the antecedent of the RC. For this reason, we need to look at an SRC with a complex relative constituent to refine our semantic representation. Such a sentence is given in (26).

(26) Kim likes *War and Peace*, [RC: [a copy of which]$_x$ Kim possesses __$_x$].

$$\mathbf{like}'(\mathbf{kim}', \mathbf{w\&p}') \wedge [\exists x : \mathbf{copy\text{-}of}'(x, (\iota w : w = \mathbf{w\&p}'))](\mathbf{possess}'(\mathbf{kim}', x)))$$

In this sentence, the antecedent is the book *War and Peace* and the gap is a concrete copy of this book. In the semantic representation, we need to separate these two elements: The relation between the relative pronoun and the antecedent is mediated through the variable w, the relation between the gap and the relative constituent through the variable x.

This more complex example shows that the moved constituent is interpreted just as if it was not moved, and the relative pronoun is interpreted as a definite description, expressing an identity with its antecedent.

Interestingly, we can show that a quantifier that occurs inside a complex relative phrase can be interpreted with narrow scope with respect to an operator in the core SRC.

The examples in (27) show that a universal quantifier inside the relative constituent can take wide scope over an indefinite inside the RC, (27a). But it can also take narrow scope, as in (27b). In the latter example, there are two languages, English and French, which each member of the LSA speaks.

(27) a. Alex joined the LSA, [[every member of which] receives a membership badge].

b. Alex joined the LSA, [[every member of which] speaks two languages – English and French].

The semantic representation of the narrow-scope reading of the universal quantifier in (27b) is given in (28).[5]

(28) $\mathbf{join}'(\mathbf{alex}', \mathbf{lsa}')$
$\wedge[\exists_2 y : \mathbf{language}'(y)]$
$([\forall x : \mathbf{member\text{-}of}'(x, (\iota w : w = \mathbf{lsa}'))](\mathbf{speak}'(x, y)))$

We can replace pronoun-*which* in (27) with the determiner-*which* NP *which association*, as in (29). The sentences in (29) have exactly the same readings as the ones with pronoun-*which*.

(29) a. Alex joined the LSA, [[every member of which association] receives a membership badge].

b. Alex joined the LSA, [[every member of which association] speaks two languages – English and French].

I provide the semantic representation of the narrow scope reading of (29b) in (30), where $\mathbf{asso}'$ abbreviates the predicate $\mathbf{association}'$.

(30) $\mathbf{join}'(\mathbf{alex}', \mathbf{lsa}')$
$\wedge[\exists_2 y : \mathbf{language}'(y)]$
$([\forall x : \mathbf{member\text{-}of}'(x, (\iota w : (\mathbf{asso}'(w) \wedge w = \mathbf{lsa}')](\mathbf{speak}'(x, y)))$

We can observe a similar behavior with respect to *believe* predicates. The sentence in (31) allows for both a de-re and a de-dicto reading as indicated in

[5] I use "$\exists_2$" for "there exist two".

(31a) and (31b) respectively.

(31) Alex joined the LSA, [[every member of which] Kim believes to be a spy].

a. De re: For every x such that x is a member of the LSA, Kim believes that x is a spy.

b. De dicto: Kim believes that for every x such that x is a member of the LSA, x is a spy.

Again, we obtain the same readings if we use *which association* as the relative phrase. However, the condition that the LSA is an association is required to be outside the scope of the *believe* predicate, just as observed in section 5.2.2. The available readings are given in (32).

(32) Alex joined the LSA, [[every member of which association] Kim believes to be a spy].

a. De re: The LSA is an association and for every x such that x is a member of the LSA, Kim believes that x is a spy.

b. De dicto: The LSA is an association and Kim believes that for every x such that x is a member of the LSA, x is a spy.

In section 5.2, I began with the semantic representations provided in Arnold (2004, 2007). I showed how their core properties can be expressed in a predicate-logical semantic representation language. Next, I have explored the semantic properties of determiner-*which* NPs, and I have shown that they do not make an at-issue contribution, even if the overall SRC they are contained in does. I have argued that determiner-*which* NPs have the status of CIs. In the next step, I looked at larger relative constituents, i.e. those in which a more complex upward percolation of the REL feature can be observed. Interestingly, the CI-status does

not get inherited by the entire relative phrase but is restricted to the determiner-*which* NP. In the next sections, I integrate the semantic properties of SRCs in general and determiner-*which* SRCs in particular with the syntactic analysis given in section 5.1.

5.3 LRS analysis of SRCs

In this section, I give an LRS analysis of SRCs in general, postponing the analysis of determiner-*which* SRCs to section 5.4. Once again, I want to stress that, similarly to Pollard & Sag (1994a), I am assuming a phonologically null-relativizer, or *null-relativizer*. There should just be one null-relativizer for all *wh*-SRCs. In particular, it should be consistent with both determiner-*which* and pronoun-*which* SRCs.[6]

For ease of reference, I provide an example with a pronoun-*which* SRC in (33), which will serve as our running example in this section.

(33) Alex read *War and Peace*, which Kim recommended.

$$\mathbf{read}'(\mathbf{alex}', \mathbf{w\&p}') \wedge \mathbf{recommend}'(\mathbf{kim}', (\iota w : (\neg \mathbf{human}'(w) \wedge w = \mathbf{w\&p}')))$$

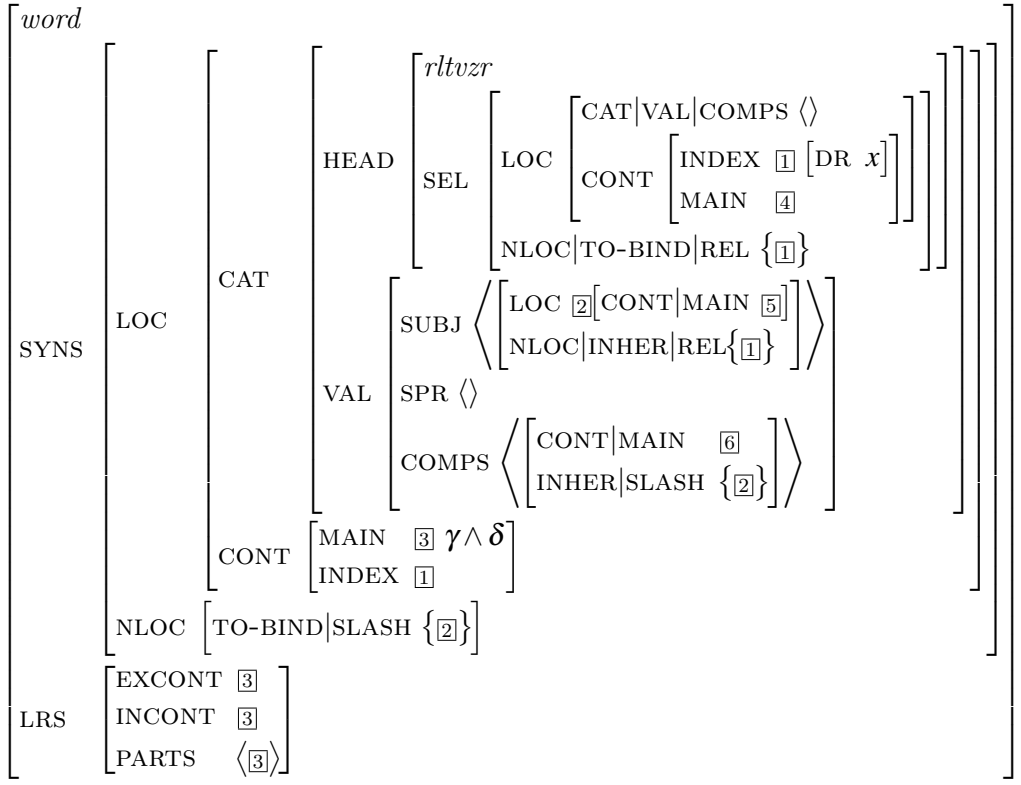
The semantic representation in (33) is as motivated in section 5.2 above. I assume that the relative word *which* contributes the entire ι-expression, whereas the null-relativizers just contributes the coordination of the matrix clause and the SRC.

The description of the null-relativizer is shown in Figure 5.11 below. Syntactically, it looks just like the relativizer for *wh*-IRCs in Pollard & Sag 1994a.

[6] In section 6.4, I will propose a second SRC-relativizer, *that*, for *that*-SRCs.

Semantically, it just contributes a coordination of two formulæ, one being the semantics of the RC.

Shorthand of the semantic contribution: $\gamma \wedge \delta$



and $x \leq \gamma$ and $x \leq \delta$ (discourse referent of the antecendent is in both conjuncts)
and $\boxed{4} \leq \gamma$ (the antecedent is in the 1st conjunct)
and $\boxed{5} \leq \delta$ (the relative constituent is in the 2nd conjunct)
and $\boxed{6} \leq \delta$ (the core SRC is in the 2nd conjunct)

Figure 5.11: Description of null-relativizer for SRCs

Figure 5.11 shows that the HEAD is a relativizer and, in its SEL value, it selects an antecedent. The PHI index of the antecedent, $\boxed{1}$, is identical to the PHI value of the SUB|NP|NLOC|REL, the relative pronoun in the relative phrase. In its SUBJ list, the relativizer is looking for some phrase, the relative phrase. In its COMPS list it is looking for a S, the rest of the RC, which carries the SLASH value. The

relativizer itself just contributes a coordination, $\gamma \wedge \delta$ – as its MAIN, INCONT, and EXCONT value. Its CONTENT is identical with that of the relativized element, [7] in Figure 5.11.

Below the AVM in Figure 5.11, there are additional embedding constraints. These make sure that the content of the matrix clause, including the antecedent occurs in the first conjunct, and that the content of the SRC, including the relative constituent, is a component of the second conjunct.

Shorthand of the emantic contribution:

$$\iota w : (\neg \mathbf{human}'(w) \wedge w = x)$$

$$\begin{bmatrix}
\text{PHON} & \langle \textit{which} \rangle \\
\text{SYNS} & \begin{bmatrix}
\text{LOC} & \begin{bmatrix}
\text{CAT} & \begin{bmatrix}
\text{HEAD} & \textit{noun} \\
\text{VAL} & \begin{bmatrix}
\text{SUBJ} & \langle\rangle \\
\text{SPR} & \langle\rangle \\
\text{COMPS} & \langle\rangle
\end{bmatrix}
\end{bmatrix} \\
\text{CONT} & \begin{bmatrix}
\text{MAIN} & \boxed{4}\ \neg \mathbf{human}'(w) \\
\text{INDEX} & \begin{bmatrix}
\text{DR} & w \\
\text{PHI} & \boxed{2} \begin{bmatrix} \text{PER} & \textit{3rd} \\ \text{GEN} & \textit{neuter} \end{bmatrix}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix} \\
\text{NLOC} & \begin{bmatrix} \text{INHR} \mid \text{REL} \left\{ \begin{bmatrix} \text{DR} & x \\ \text{PHI} & \boxed{2} \end{bmatrix} \right\} \end{bmatrix}
\end{bmatrix} \\
\text{LRS} & \begin{bmatrix}
\text{EXCONT} & \boxed{3}\ (\iota w : (\neg \mathbf{human}'(w) \wedge w = x)) \\
\text{INCONT} & \boxed{4}\ \neg \mathbf{human}(w) \\
\text{PARTS} & \left\langle \begin{matrix} w, \mathbf{human}', \mathbf{human}'(w), \boxed{3}, \boxed{4}, w = x, \\ \neg \mathbf{human}'(w), (\neg \mathbf{human}'(w) \wedge w = x) \end{matrix} \right\rangle
\end{bmatrix}
\end{bmatrix}$$

Figure 5.12: Description of pronoun-*which* in SRCs

Figure 5.12 shows the lexical specification of pronoun-*which*. It is treated as a noun phrase with a non-trivial REL specification. Semantically, it behaves like

a pronoun in discourse. The HEAD is *noun* and its VAL lists are empty. Its MAIN value is the predicate **human′** and its PHI value expresses that the pronoun is for the third person singular, neuter. It has a non-empty INHER|REL value. The LRS|EXCONT is an ι-expression that binds the variable w and has some descriptive content that w is not human, but identical with some x. Finally, in its PARTS list it contains all the parts that pronoun-*which* contributes.

It is essential for my analysis, and maybe somewhat surprising, that the INHER|REL value of the relative pronoun does not contain the pronoun's INDEX value. Instead, it contains the INDEX value of the antecedent. This allows me to include the identity statement $w = x$ in the lexical specification of the pronoun. To see that the x in the lexical entry of the relative pronoun really is the antecedent, it suffices to look again at Figure 5.11. There, the INDEX value of the antecedent is required to be in the INHER|REL set of the relative constituent.

So far, the description of a noun, the null-relativizer and pronoun-*which* have been presented. I will combine them to show the derivation of the NP *LGB, which* ∅ *Kim recommended*. The combination of the null-relativizer with the core SRC is given in Figure 5.13.

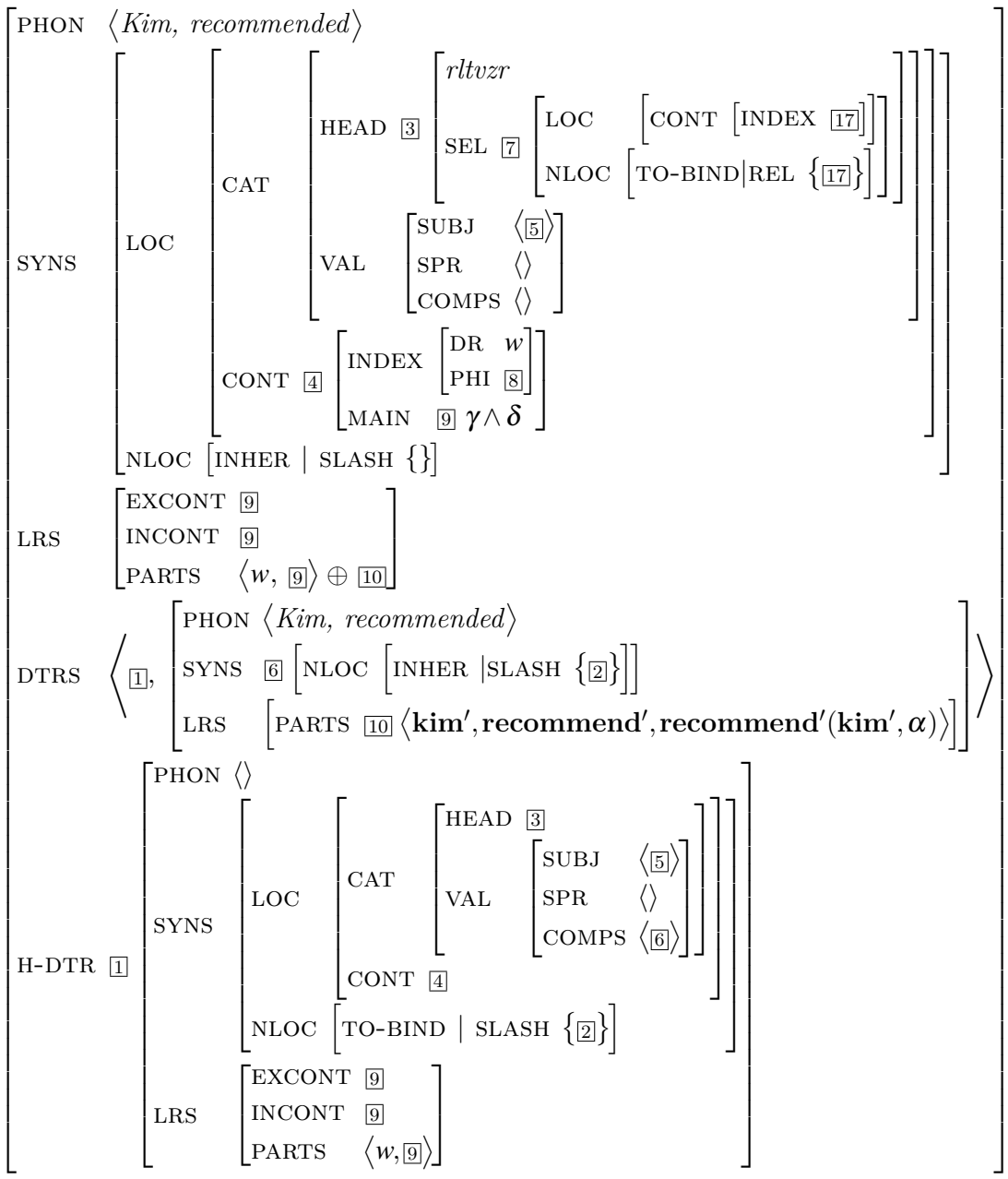


Figure 5.13: Description of null-relativizer combining with *Kim recommended*

Figure 5.14 shows the full SRC. In this structure, the relative pronoun *which* combines with the phrase in Figure 5.13. The LOC value of the pronoun is identical with the slashed element. The pronoun itself introduces a non-trivial INHER|REL value which percolates to the full SRC.

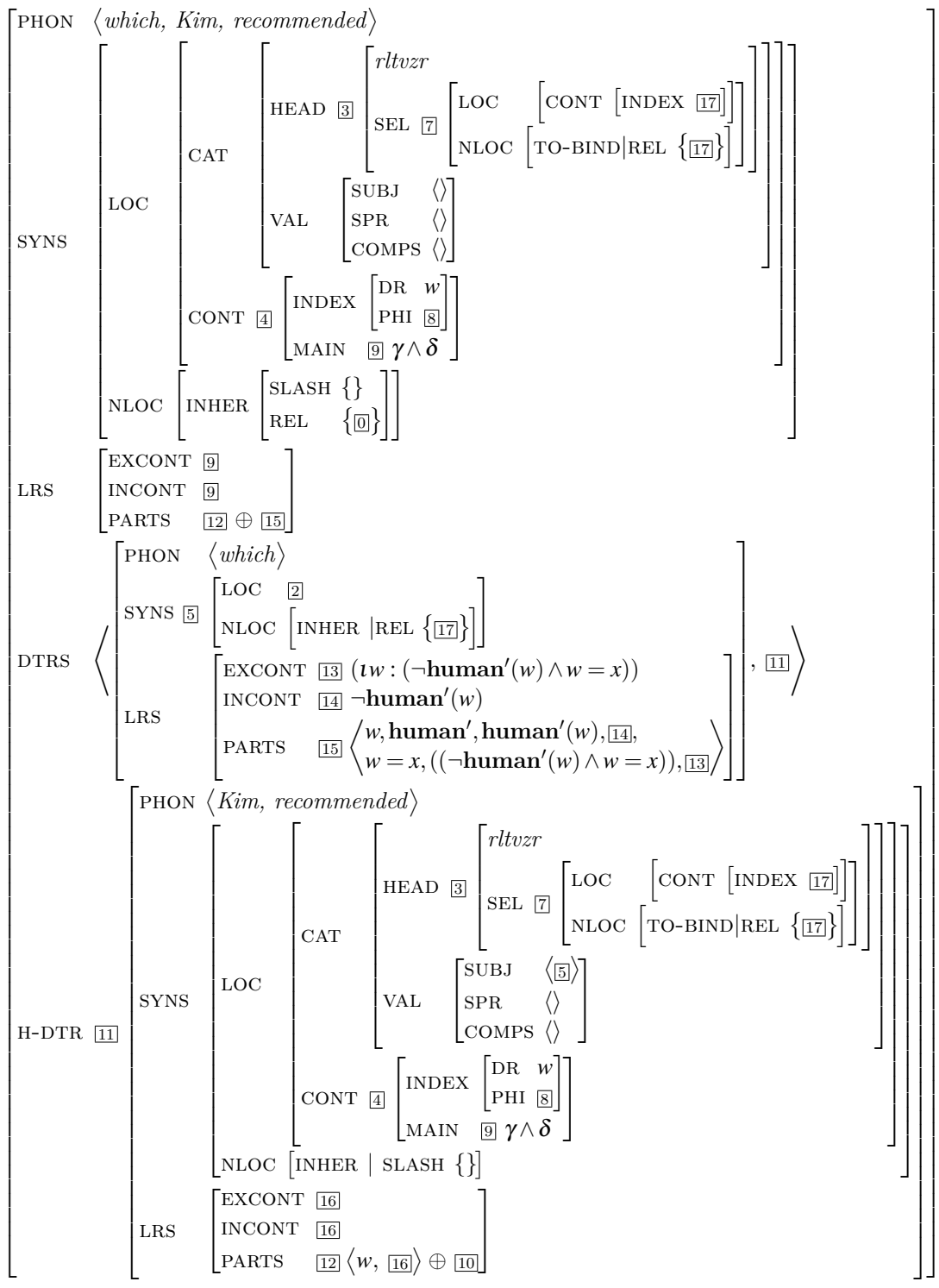


Figure 5.14: Description of the full SRC *which Kim recommended*

The AVM in Figure 5.15, finally, shows the complete NP *LGB, which Kim recommended*. In this NP, the noun *LGB* is the head combining with the SRC via the head-functor schema. It binds the REL value of the SRC.

Note that the semantic contribution of the SRC does not appear in the NP's EXCONT value. Instead, it will be conjoined with the overall matrix clause.

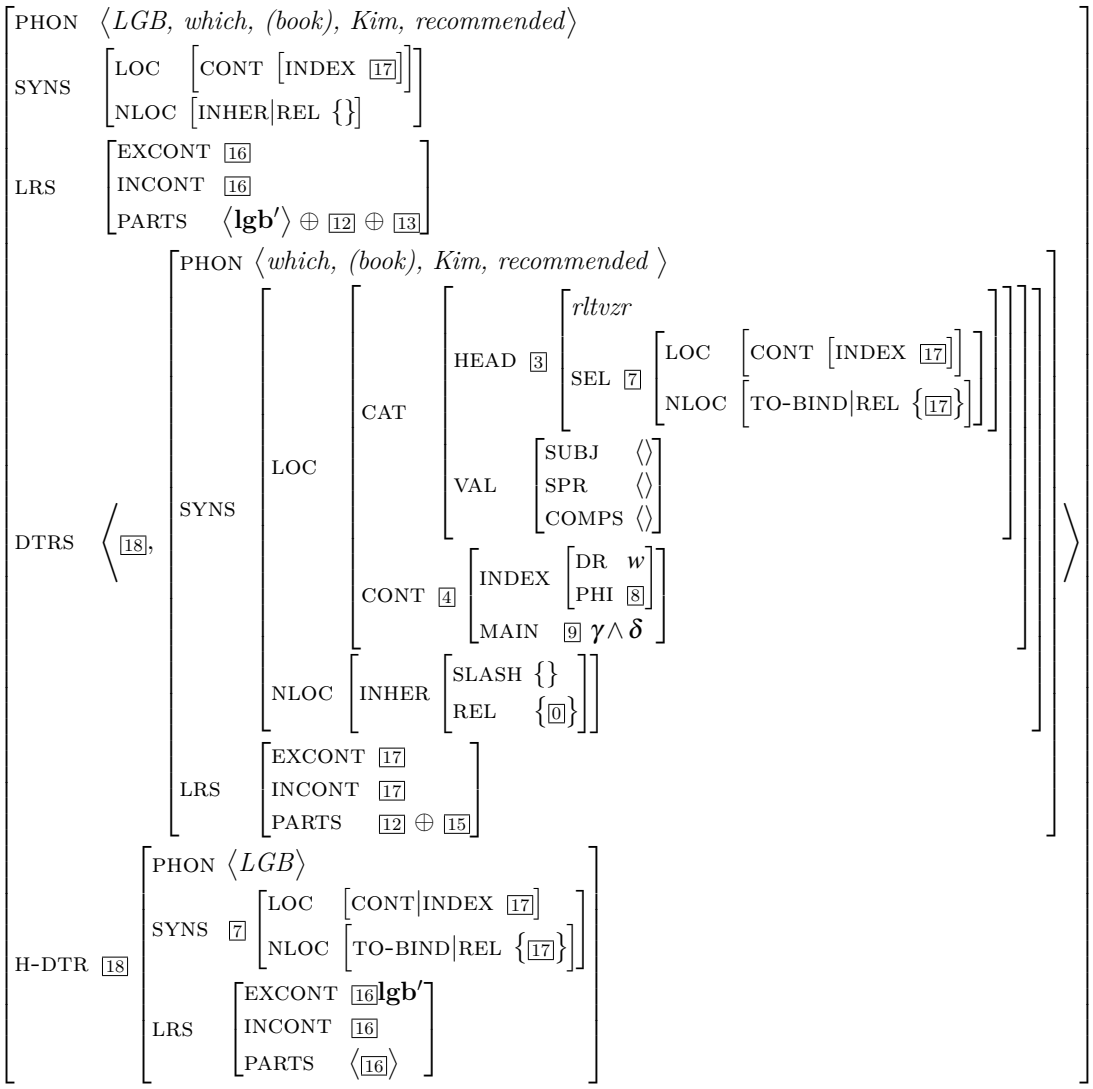


Figure 5.15: Description of the NP *LGB, which Kim recommended*

In this section, I provided an analysis of SRCs which combines the classical

syntactic HPSG analysis from Pollard & Sag (1994a) with LRS. To my knowledge, this is the first LRS analysis of SRCs. In the next section, I extend this analysis to determiner-*which* SRCs.

5.4 Analysis of determiner-*which* in LRS

The semantics of determiner-*which* should be on the one hand as similar as possible to that of other determiners, and, on the other, as similar as possible to that of pronoun-*which*, and I present an analysis that fulfills these two requirements.

Shorthand of the semantic contribution: $\iota w : (\alpha \wedge w = x)$

$$\left[\begin{array}{ll}
\text{PHON} & \langle \textit{which} \rangle \\
\text{SYNS} & \left[\begin{array}{ll}
\text{LOC} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll}
\text{HEAD} & \left[\begin{array}{ll} \textit{det} & \\ \text{SEL} & \text{N'} \left[\text{LOC} \mid \text{CONT} \left[\text{INDEX} \ \boxed{1}\right]\right] \end{array}\right] \\
\text{VAL} & \left[\begin{array}{ll} \text{SUBJ} & \langle\rangle \\ \text{SPR} & \langle\rangle \\ \text{COMPS} & \langle\rangle \end{array}\right]
\end{array}\right] \\
\text{CONT} & \left[\begin{array}{ll}
\text{INDEX} \ \boxed{1} & \left[\begin{array}{ll} \text{DR} & w \\ \text{PHI} & \boxed{2} \end{array}\right] \\
\text{MAIN} & \iota
\end{array}\right]
\end{array}\right] \\
\text{NLOC} & \left[\text{INHER}|\text{REL} \left\{\left[\begin{array}{ll} \text{DR} & x \\ \text{PHI} & \boxed{2} \end{array}\right]\right\}\right]
\end{array}\right] \\
\text{LRS} & \left[\begin{array}{ll}
\text{EXCONT} & \boxed{3} \ \iota w : (\alpha \wedge w = x) \\
\text{INCONT} & \boxed{3} \\
\text{PARTS} & \left\langle w, \iota, \boxed{3}, (\alpha \wedge w = x), w = x \right\rangle
\end{array}\right]
\end{array}\right]$$

Figure 5.16: Description of determiner-*which* for SRCs

The basic form of the lexical entry of determiner-*which* was already given in

Figure 5.1, section 5.1. I enrich this lexical specification by adding values for the semantic features. This is done in Figure 5.16. The HEAD value indicates that the word is a determiner. Through its SEL value, it selects for a noun whose PHI feature value is identical to its own, $\boxed{2}$. The determiner has an ι-operator as its MAIN value.

Being a relative word, determiner-*which* has a non-trivial INHER|REL value. This set contains the index of the antecedent. The PHI values of the relative word and the head noun it selects guarantee the morpho-syntactic agreement between the antecedent and the determiner-*which* NP.

The EXCONT value is the ι-expression that uses this ι-operator. It binds the DR value of the determiner, here the variable w. In the descriptive part of this ι-expression, there is a conjunction consisting of an underspecified slot, $\boldsymbol{\alpha}$, and an identity statement, $w = x$. The identity statement connects the DR value of the determiner, w, and that of the element in the INHER|REL set. In its PARTS list, the determiner has all elements that it contributes, i.e., the variable w, the ι operator, the EXCONT value, the conjunction in the restrictor of the ι-operator, and, finally, the identity statement.

In the next step, we can combine determiner-*which* with a noun, applying the Head-Functor schema from Figure 3.6, in chapter 3. The resulting NP is shown in Figure 5.17.[7] It is very similar to the lexical entry of pronoun-*which*. The only exception being that we have the formula $\mathbf{book}'(w)$ inside the ι-expression, where we had $\neg\mathbf{human}'(w)$ in the case of pronoun-*which*.

When we insert this relative phrase into our example SRC, "*LGB, which book Kim recommended*," we arrive at the semantic representation motivated in section 5.2.2.

[7]See section 3.4 for details of the combination of a determiner and a noun.

Shorthand of the semantic contribution: $\iota w : (\mathbf{book}'(w) \wedge w = x)$

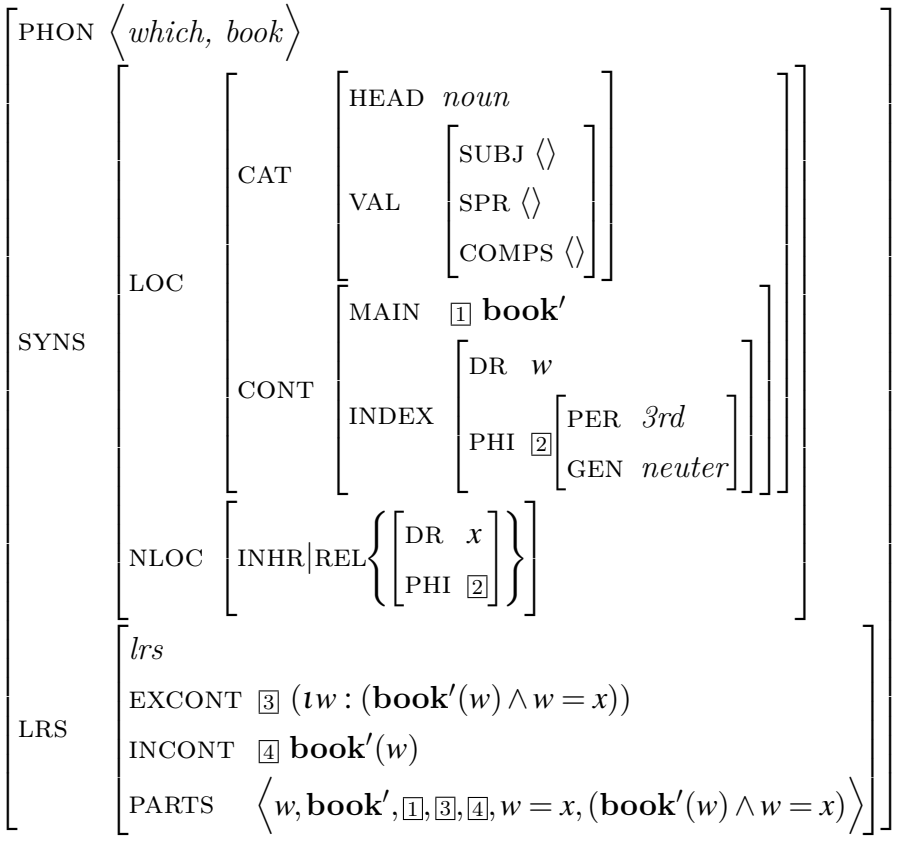


Figure 5.17: Analysis of the SRC relative phrase *which book*

To sum up, the HPSG analysis in Pollard & Sag (1994a) that is provided for IRCs can also be used in analysing SRCs with some amendments. Moreover, looking at the lexical entry of pronoun-*which* in Figure 5.12 and the description of the phrase *which book* in Figure 5.17 makes it clear that the description of a SRC containing a determiner-*which* is very close to that of a SRC containing a pronoun-*which*.

Having seen the HPSG analysis of the types of percolation discussed above, and the results of the data survey in chapter 4, section 4.2.3, the following two generalizations could be drawn. First, the standard REL-percolation accounts for the cases in which the relative phrase contains determiner-*which*. As shown in

the analysis, in all of the types of percolation the REL value percolated up the tree similarly to the cases with pronoun-*which*. Second, as mentioned in chapter 4 section 4.2.3, there seems to be no empirical motivation for adding special constraints that would make the percolation of determiner-*which* more restricted than that of other relative words. The results of the data survey in chapter 4, section 4.2.3 show that the rating of the examples in both cases with pronoun-*which* and determiner-*which* were very similar.

From section 5.3 on, I developed the first LRS-analsis of SRCs. I showed how the DRT-analysis from Arnold (2004, 2007) can be rephrased in a predicate logical semantic representation language. I gave evidence for the special status of the material inside a determiner-*which* NP, which I captured by using an ι-expression in the semantic representation of relative phrases in SRCs. Finally, I provided a full integration of this analysis into LRS. Hereby, I used pronoun-*which* first and showed that determiner-*which* can be modelled in such a way that the determiner-*which* NP looks very much like pronoun-*which*.

Chapter 6

SRCs with *that*

This chapter consists of two main parts. Part one is where an overview of the construction of RCs with relative *that* is given, followed by corpus data and a questionnaire study to get a better impression of the status of *that* SRCs in present day English. Part two is where I discuss the HPSG and LRS analysis of relative *that* in SRCs. In Section 6.1 I give an overview on using demonstratives such as *that* with RCs. I list examples of *that* SRCs that I extracted from COCA in section 6.2. section 6.3 presents the questionnaire study. In the second part of the chapter, I will first provide a general analysis of relative *that* and bare RCs in HPSG (section 6.4). In section 6.5, the analysis of *that* SRCs is given, including their semantics.

6.1 Background on demonstrative-based relative words

Demonstratives as relative words occur in various languages that are not genetically related. Using demonstratives as relative markers is as common cross-

linguistically as using the interrogatives. Hendery (2007: 65) states that demonstrative-based relative words are used in the following language families: Austronesian, Kartvelian, Indo-European, Nigro-Congo, Quechuan, Semitic, Sino-Tibetan, Uralic, Uto-Aztecan, English, Chinese and Hindi.

Hendery adds that the use of demonstratives as relative words might not be due to language contact. She supposes this might rather be due to language-internal processes, for instance, the reanalysis of clause boundaries or nominalization of RCs.

Furthermore, Hendery (2007: 67) proposes that demonstratives should be more common in constructing relative words than interrogatives. Her justification for this claim is that, demonstratives are used when the information is shared by the speaker and the hearer whereas interrogatives are used when the speaker doesn't have the information and therefore asks the hearer for it. Since RCs are anaphorics, i.e. they comment on an already mentioned element in the sentence, demonstratives should be the main ingredients in the construction of relative words.

However, this reasoning is not reflected in the formation of English relative words for there is only one demonstrative-based relative word, *that*, and its usage is, conventionally, limited to IRCs (Huddleston & Pullum, 2002: 1052). The rest of the relative words in English are interrogative-based. It has been assumed that the occurrence of *that*, in English RCs, is limited to IRCs where the RC is identifying, i.e. the RC is vital in the identification of the antecedent, and this consequently means that the information is not shared between the speaker and the hearer.

Furthermore, the occurrence of *that* in SRCs is not traditionally grammatically acceptable; SRCs are not identifying, i.e. they are not vital to the identification of

the antecedent either because the antecedent is generally known or the knowledge about it is, at least, shared between the speaker and the hearer.

Diessel (1999: 123) distinguishes between two types of demonstratives.

1. The pronominal demonstratives; these are the demonstratives that are anaphors. This type can stand alone like relative *that* in English.

2. The adnominal demonstratives; these are the demonstratives that connect an element to its nominal head such as *that exam*.

In modern English, demonstrative *that* occurs in postnominal RCs. This demonstrative-based relative word, *that*, is often not considered the head of the RC similar to the other types of relative pronouns in most of the grammatical frameworks.

Huddleston & Pullum (2002) have already provided some examples which confirmed that some native speakers allow the usage of *that* in SRCs. However, they only present four examples, three of which seem to be taken from literary texts, such as novels and plays. The examples are:

(1) She had long been accustomed to the solitary nature of her son's instincts, that I had tried - - - and failed - - - to stifle.

(2) His heart, that had lifted at the sight of Joanna, had become suddenly heavily at the sight of Ramdez thumping after her.

(3) February, that in other years held intimation of spring, this year prolonged the bitter weather.

Only one example that seems to be from a scientific context is the following.

(4) The patas monkey, that spends almost all of it's time in open grassland, adopts just such tactics (Huddleston & Pullum, 2002: 1052).

Traditionally, it has been assumed that in IRCs there is a semantic difference between *which* and *that*: while *which* is restricted to non-human antecedents only, *that* can be used for all antecedents, as in:

(5) The boy that sold the candy cotton came by. (human antecedent)

(6) The dog that was cut loose was deaf. (animal antecedent)

(7) The tree that was in the middle of our yard was 100 years old. (inanimate antecedent)

(8) The kind of freedom that he talked about does not exist. (abstract antecedent)

The difference between relative *that* and relative *who* is that *who* is more formal and only refers to humans while *that* is more colloquial and is used for all kinds of antecedents.

The acceptance of using *that* in SRCs has only been recently discussed in some blogs and most of the commenters seem to find this construction natural. In one of the blogs by Carey (2013), which is called *sentence first*,[1] the writer declares that he comes across many examples with *that* in SRCs only because he is consciously looking for it. He emphasizes that many examples can be overlooked if we don't carefully pay attention to them. This, to me, is very important, because the fact that if you don't pay enough attention to finding it while reading, it can slip in front of you means that it is natural and acceptable that is why it can easily slip by without getting the reader's attention. In contrast to that, a grammatically incorrect construction can easily attract the reader's attention.

In the same blog Carey provides an example from a British newspaper, the

[1] https://stancarey.wordpress.com/2013/12/15/non-restrictive-that-that-can-be-ambiguous/, accessed 19.02.2020.

The Guardian, where *that* is used in an SRC in one of the sub headings. Here is the example:

Carey (2013)

I conclude here that its occurrence in a well-known newspapers like *The Guardian* is evidence of its acceptance. Carey (2013) also mentions that, earlier usages of *that* in SRCs were common in English literature especially in plays, novels and poetry. An example of *that* in SRCs in a play is found in *Macbeth* by Shakespeare.

(9) Fleance, his son, that keeps him company, whose absence is no less material to me than is his father's. (Shakespeare, 1997: 1206)

Another similar example from Shakespeare's work also from *Macbeth* is

(10) As she is troubled with thick-coming fancies, that keep her from her rest. (Shakespeare, 1997: 1228)

Enright's (2007) *The Gathering* also has several instances of supplementary *that*:

(11) Ada bringing us four red lemonade into a pub, that had a black roof with huge letters of white written across it.

(12) If Ada believed in anything she believed in this persistence, that other people might call the soul.

An instance of *that* in SRCs in poetry can be seen in Tennyson's *In Memorian*

(13) Break, thou deep vase of chilling tears, that grief hath shaken into frost! (Tennyson, 1974)

Carey (2013) also explains that, similar to *wh*-relatives, the problem with using *that* in SRCs is that it can semantically cause a confusion, for example, an SRC could be considered an IRC with an usual comma before *that* or sometimes the SRC could be mistaken for an elliptical clause while the usage of *that* with IRCs can avoid all these confusions. On occasion, the usage of splice commas, which are commas that are wrongly used to coordinate independent clauses, might lead to mistaking the clause to be an SRC; examples of these can be substantially seen in COCA and prose writings as in Kilroy's novel *All Names Have Been Changed*:

(14) He kept writing writers' novels, that was the problem. (Kilroy, 2010: 106)

(15) I'd even heard what the man had to say, that is the class of fool I was then. (Kilroy, 2010: 132)

Carey (2013) states that he is sure that the comma before *that* in (15) is a splice comma and the meaning of (14) is equal to "He kept writing writers' novels. That was the problem".

To test the credibility of Huddleston & Pullum's (2002) claim that the relative word *that* can occur in SRCs and also to settle the argument that runs in linguistics blogs, I decided to search for naturally occurring examples in corpora and to run a questionnaire with examples from different contexts and with different

constructions, both constructed and also from COCA. The results of the questionnaire should shed some light on the truth behind the acceptance of relative *that* in SRCs.

6.2 Occurrences of *that* SRCs in COCA

As the starting point of my research on the occurrences of relative *that* with SRCs, I used COCA to find instances of such a structure. Within my first search, I found many instances of the structure in a variety of discourses and across genres. The following are some of the occurrences of relative *that* with SRCs that I extracted from COCA.

I checked that all these examples have an SRC reading in their occurrence context in COCA. For some cases, an IRC reading might be possible as well – or indistinguishable from an SRC reading. In those cases, I take the separation of the RC by commas as an additional clear indication that they represent SRCs. This is a legitimate assumption as the occurrence of a comma before *that* is highly restricted in English.[2]

(16) The big topic this week was this video that Mitt Romney uploaded on YouTube, **that**, according to reliable sources, had been filmed during a private party of an important sponsor of Romney's 2008 election campaign. (COCA)

(17) Despite her request to take her to a good specialist, her brother-in-law brought some Ayurvedic ointment, **that**, not surprisingly, helped only to spread the sores all over her body, and her condition stayed like that for

[2]To give a very rough idea: only 5% of the occurrence of *that* in COCA are preceded by a comma, whereas 45% of occurrences of *which* follow a comma in COCA.

more than 4 months. (COCA)

(18) Then, after all the stakes were in place, we strung the wire, hanging it from one stake to the next. The actual stringing required a good sense of balance, **that**, luckily, most of us still had, and a considerable amount of physical strength. (COCA)

(19) There were some problems of security, but, Governor, they sold this narrative, **that** I know you are familiar with now.(COCA)

(20) There's been a number of objections, the most telling objections, **that** I know about, have been made by Professor Bill Shots and Professor Ken Nielson. (COCA)

Finding a number of occurrences of relative *that* in SRCs in COCA was reassuring to a large extent. However, I wanted to see if these instances of the structure are acceptable and considered natural by native English speakers nowadays. To do so, I created a questionnaire study to validate the naturalness of the above mentioned examples and some other examples of the structure from the web, literature and also constructed. In the following section, I present the questionnaire study in detail, and the source of the examples tested in the questionnaire are indicated throughout.

6.3 Questionnaire study: *that* vs. *which* SRCs

The data situation on *that* SRCs seems to be even less clear than what I reported on determiner-*which* SRCs. As shown at the beginning of this chapter, *that* SRCs are reported in grammars and we find naturally occurring examples in corpora. Nonetheless, the formal analyses of English RC constructions, so far, all start from

the assumption that *that* RCs are always IRCs. One way to make sense of this apparent discrepancy is that *that* SRCs are not perceived as natural by native speakers and that they can always be replaced by *which* SRCs, i.e. that they are an obsolete construction of English. To test this, I designed an exploratory questionnaire study to validate the naturalness of *that* SRCs in comparison to *which* SRCs.

If *which* SRCs are the unmarked, natural case and *that* SRCs are an obsolete construction, even attested examples of *that* SRCs should be rated rather unnatural and, in any case, less natural than the corresponding *which* SRCs. For this reason, I constructed a questionnaire based on the attested COCA examples listed in Section 6.2, other attested examples from web pages, and cases mentioned in Huddleston & Pullum (2002). I contrast these with their *which* SRC variant in my questionnaire.

It is the sole purpose of this study to determine whether *that* SRCs should be included in the inventory of RC constructions of Present Day English. Since this basic question is not yet resolved, I will not attempt a more detailed investigation as to the specific functional, stylistic or other properties of *that* SRCs – let alone of the contrast between *that* and *which* SRCs. The exploratory nature of my study is reflected in the design of the questionnaire and the formulation of my hypothesis, which do not satisfy the criteria of a systematic, hypothesis-testing experiment. Nonetheless, the study should help reassess the status of *that* SRCs.

6.3.1 Design

In this section, I present the results of a questionnaire that I conducted in 2018 at *George Mason University* in Washington DC and *Colorado State University* in

Colorado. The consultants were all university students, university staff with at least a university degree and some of the lecturers of the universities. The age range of the consultants was between 18 and 60.

The procedure of conducting this questionnaire is similar to the procedure discussed in section 4.2 of chapter 4. The questionnaire had two versions and each participant only received one version of it. On average, the process of filling in the questionnaire took 30 minutes.

Each version of the questionnaire contained 60 items, fifteen of them were relevant for my study, and forty-five were fillers. The crucial items had two variants: one with relative *that*, one with relative *which*. The items that are from COCA or other sources are originally with relative *that*, the relative *that* was changed in the items of the questionnaire to relative *which* for the sake of comparison.

There were two versions of the questionnaire. In version A, eight items were tested with relative *that* and seven items with relative *which*. In version B, the eight items that were tested in version A with relative *that* were tested here with relative *which* and the seven items that were tested in version A with relative *which* were tested here with relative *that*.

Across the comments on the blogs on the examples posted, some linguists and commenters suggested that they prefer the examples that contained parentheticals immediately after *that* in SRCs. During my search in COCA, I retrieved examples that also contained such parenthetical uses of expressions like *surprisingly*, *by the way*, *according to reliable sources*, etc. I wanted to test whether those commenters' intuitions can be confirmed. Therefore, I included seven items that contain a parenthetical (explicitly separated by commas) and eight items without such a parenthetical. Questionnaire A contained 4 items of each type with *which*, 3 parentheticals and 4 non-parentheticals with *that*.

Some of the items were from COCA and from Huddleston & Pullum (2002), others were from the web or constructed items. The informants were asked to rate the naturalness of the items in five categories: *very unnatural*, *unnatural*, *I cannot decide*, *natural*, *very natural*. While this instruction was rather vague, the informants did not seem to have a problem with it.

6.3.2 Hypotheses

While there are attested cases of *that* SRCs, it could still be the case that, overall, native speakers consider this type of SRC unnatural. I created a questionnaire to explore this. I took the examples with *that* SRCs from COCA and Huddleston & Pullum's (2002) as my starting point because I wanted to make sure that the examples meet whatever conditions there might be on *that* SRCs. If *that* SRCs are unnatural in general, this should be reflected in speakers' assessment of these sentences as well.

The discrepancy between the existence of occurrences of *that* SRCs in corpora and the persistent negligence of them in the linguistic literature could be a sign that speakers do not find them natural. If this is the case, I expect the following results for my questionnaire:

> (H1): *That* SRCs are judged as *natural* or *very natural* by less than half of my informants.

It is conceivable that there is a weaker version of this hypothesis. It is possible that the occurrence of a parenthetical between *that* and the core RC helps the acceptance of *that*. A similar assumption was made for the so-called *that*-trace effect. A subject gap in a *that* clause seems to be more acceptable to native

speakers if there is material between *that* and the finite verb, see (21), from Culicover's (1993), with Culicover's judgement.

(21) This is the tree that I said that *(just yesterday) t_i had resisted my shovel.
(Culicover, 1993: 558)

This leads me to a weaker form of the hypothesis:

> (H2): *That* SRCs are judged as *natural* or *very natural* by less than half of my informants whenever there is no parenthetical following relative-*that*.

If any of these hypotheses is confirmed, it would still be necessary to include *that* SRCs in the formal grammar of English. However, such a result would be support for the idea that *that* SRCs are a dispreferred structure in English.

If *that* SRCs are an obsolete structure in English, we expect that the use of a *which* SRC to always be a viable alternative. Therefore, I include a hypothesis on *which* SRCs:

> (H3): *Which* SRCs are judged as *natural* or *very natural* by at least half of my informants.

If this hypothesis is confirmed in the data, this would suggest that *which* SRCs can fully replace *that* SRCs even in cases in which *that* SRCs have been used in the corpus.

It is worth pointing out that I do not formulate a hypothesis that contrasts the two types of SRCs directly. The reason for this is that my primary goal is to determine the status of *that* SRCs in the grammar of English. We know from other types of RCs that the choice between relative *that* and a relative *wh*-word

depends on extra-linguistic factors as well, such as genre, style, etc. As these factors have not been controlled for in my study, I will not address such a direct comparison.

6.3.3 The results

In this section, I present the results of my questionnaire. I received 20 judgments for each of the two versions of my items. As my evaluation is rather coarse grained, I sum up the judgments *natural* and *very natural* into NATURAL and, analogously, I group the judgments *unnatural* and *very unnatural* into a single category UNNATURAL.

The judgements that I received for my data are presented together with the examples in (22)–(36). Items (22), (23), (25), (26), (27),(29) and (31) contain a parenthetical after the relative word, *that* or *which*. Items (22), (27), (29), (30) (33) and (36) were taken directly from COCA in their *that* variant, items (25) and (26) were taken from web pages, and items (23), (24), (28), and (31) were constructed. Items (32), (34), and (35) were examples of *that* SRCs from Huddleston & Pullum (2002).

(22) The big topic this week was this video that Mitt Romney uploaded on YouTube, **that/which**, according to reliable sources, had been filmed during a private party of an important sponsor of Romney's 2008 election campaign. (COCA, parenthetical)

that: NATURAL: 13 undecided: 4 UNNATURAL: 3

which: NATURAL: 16 undecided: 4 UNNATURAL: 0

(23) Our research found significant differences in the way the brains of musicians and non-musicians respond to music, and more subtle differences,

that/which, surprisingly, helped elucidate these different vulnerabilities to earworms, between women and men. (constructed, parenthetical)

that: NATURAL: 18 undecided: 0 UNNATURAL: 2

which: NATURAL: 11 undecided: 4 UNNATURAL: 5

(24) She sent me the draft for her first book, **that/which** she had finished in just two months' time by the way, and I immediately saw that she has all that it takes to become the next best-selling author of our publishing house. (constructed)

that: NATURAL: 12 undecided: 5 UNNATURAL: 3

which: NATURAL: 18 undecided: 2 UNNATURAL: 0

(25) A giant fox, **that/which**, if we can trust the local press releases, is twice the size of a normal specimen, has been captured in Kent, sparking fears that the animals are growing larger because of "easy living" on bins and scraps. (https://www.kentonline.co.uk/, last accessed 24.09.2020, parenthetical)

that: NATURAL: 12 undecided: 2 UNNATURAL: 6

which: NATURAL: 11 undecided: 1 UNNATURAL:8

(26) The key point, **that/which**, in my opinion, all the popular reports missed, is that FOXP2 is a transcription factor.
(https://www.evolutionpages.com/FOXP2_language.htm, last accessed 24.09.2020, parenthetical)

that: NATURAL: 7 undecided: 3 UNNATURAL: 10

which: NATURAL: 10 undecided: 8 UNNATURAL: 2

(27) Despite her request to take her to a good specialist, her brother-in-law brought some Ayurvedic ointment, **that/which**, not surprisingly, helped only to spread the sores all over her body, and her condition stayed like that for more than 4 months. (COCA, parenthetical)

that: NATURAL: 10 undecided: 6 UNNATURAL: 4

which: NATURAL: 10 undecided: 2 UNNATURAL: 8

(28) Anna's wonderful power point presentation, **that/which** I am sure you can download from her web site by now, contains all relevant information on how to run a small consulting company from your living room. (constructed, parenthetical)

that: NATURAL: 14 undecided: 1 UNNATURAL: 5

which: NATURAL: 14 undecided: 2 UNNATURAL: 4

(29) Then, after all the stakes were in place, we strung the wire, hanging it from one stake to the next. The actual stringing required a good sense of balance, **that/which**, luckily, most of us still had, and a considerable amount of physical strength. (COCA, parenthetical)

that: NATURAL: 10 undecided: 4 UNNATURAL: 6

which: NATURAL: 8 undecided: 4 UNNATURAL: 8

(30) There were some problems of security, but, Governor, they sold this narrative, **that/which** I know you are familiar with now. (COCA)

that: NATURAL: 6 undecided: 5 UNNATURAL: 9

which: NATURAL: 8 undecided: 2 UNNATURAL: 10

(31) The teachers emphasized competition between the twins, **that/which**, by the way, the two never took very seriously. (constructed, parenthetical)

that: NATURAL: 12 undecided: 2 UNNATURAL: 6

which: NATURAL: 8 undecided: 6 UNNATURAL: 6

(32) His heart, **that/which** had lifted at the sight of Joanna, had become suddenly heavy at the sight of Ramdez.(Huddleston & Pullum, 2002: 1052)

that: NATURAL: 12 undecided: 2 UNNATURAL: 6

which: NATURAL: 18 undecided: 2 UNNATURAL: 0

(33) Assume further that two taxpayers, April and Bob, **that/which** both live

in State A, that April earns her income in State A whereas Bob earns his income in State B. In this circumstance, Bob will pay more income tax than April solely as he earns income interstate. (amended from COCA[3])

that: NATURAL: 14 undecided: 2 UNNATURAL: 4

which: NATURAL: 10 undecided: 6 UNNATURAL: 4

(34) She had long been accustomed to the solitary nature of her son's instincts, **that/which** I have tried – and failed – to stifle. (Huddleston & Pullum, 2002: 1056)

that: NATURAL: 11 undecided: 7 UNNATURAL: 2

which: NATURAL: 12 undecided: 6 UNNATURAL: 2

(35) The Patas monkey, **that/which** spends almost all of its time in open grassland, adopts just such tactics. (Huddleston & Pullum, 2002: 1052)

that: NATURAL: 15 undecided: 3 UNNATURAL: 2

which: NATURAL: 9 undecided: 5 UNNATURAL: 6

(36) There's been a number of objections, the most telling objections, **that/which** I know about, have been made by Professor Bill Shots and Professor Ken Nielson. (COCA)

that: NATURAL: 14 undecided: 2 UNNATURAL: 4

which: NATURAL: 8 undecided: 5 UNNATURAL: 7

As indicated below the items, there are only two examples, (26) and (30) which were judged as NATURAL by less than half of my consultants. This shows that the first hypothesis, H1, cannot be maintained. In other words, it has not been confirmed that *that* SRCs are generally perceived as rather unnatural by native

[3]Modified from the following COCA example:

(i) Assume further that two taxpayers, April and Bob, both live in State A, but that April earns her income in State A whereas Bob earns his income in State B. In this circumstance, Bob will pay more income tax than April solely because he earns income interstate. (COCA)

speakers.

The same picture emerges when considering the examples without a parenthetical: only one of them, (30), was among the sentences judged as NATURAL by fewer than half of my consultants. This means that H2 can be rejected as well. Note that one of the items with low ratings, (26), contains a parenthetical. This suggests that the occurrence of a parenthetical is neither required nor sufficient for the naturalness of *that* SRCs.

Even though H1 and H2 could not be confirmed in the questionnaire, we can still speak of a preference of *which* SRCs if the versions of the items with *which* were judged as NATURAL throughout. This is not the case, though. Examples (29), (30), (31), (35) and (36) were judged as NATURAL by less than 10 of my informants.

To sum up, none of the three hypothesis was confirmed in my questionnaire. This means that I could not find a dis-preference against *that* SRCs, neither in the sense that the structures are perceived as less natural in general, nor in the sense that *which* SRCs would be perceived as natural throughout, even in sentences in which *that* SRCs would be possible.

6.4 The analysis of *that*

In this section, I extend my analysis of SRCs to *that* SRCs. Previous HPSG approaches built on the generalization that *that*-RCs pattern with bare RCs in that they cannot be SRCs. Consequently, there exists no HPSG analysis of this type of SRCs so far.

I first show how Arnold (2004) excludes *that* SRCs in his approach. Next, I will briefly go through Sag's (1997) arguments for treating relative *that* as a

pronoun. Finally, I present my own analysis, in which I treat *that* as a relativiser.

As mentioned in section 3.3.5, Arnold introduces a constraint to restrict the occurrences of SRCs to types and sub-types of the RCs. The constraint on the possible form of SRCs is the following:

(37) (*rel-cl*$\wedge$ *global-scope-sem*) $\rightarrow$ *fin-head-filler-phrase*

(Arnold, 2004: 43)

This constraint enforces that SRCs be finite. It also requires that SRCs be head-filler phrases, which includes *wh*-RCs. This is, however, not necessarily sufficient to exclude *that* SRCs as *that* could be treated as a relative pronoun following Sag (1997). Arnold uses and modifies specifics of the classical HPSG encoding of semantics to exclude *that* SRCs. He assumes that the REL feature contains objects that are bigger than indices – so called *nominal-objects*, which have the attribute INDEX and a set-valued attribute RESTR. He suggests that there is an empty RESTR value for IRCs, but a non-empty one for SRCs. Arnold excluded *that* SRCs by the additional assumption that *that* has an empty RESTR.

Even without taking *that* SRCs into consideration, Arnold's solution might be problematic. English distinguishes between *who* and *which* in both IRCs and SRCs, where *who* is reserved for animate antecedents. It is not intuitively plausible that this information should be part of the REL specification in the case of SRCs but not for IRCs. If, as it seems, this encoding of the IRC/SRC distinction is mainly motivated to exclude *that* SRCs, it is even more doubtful in the light of the acceptability of *that* SRCs.[4]

Before I present my analysis of *that* SRCs, I would like to discuss the possible status of *that*. Sag (1997) tries to show that *that* in RCs is a pronoun rather than

[4]Arnold (2004: 44) acknowledges that *that* SRCs might be possible, though, but does not show how his analysis would be extended to capture them.

a complementizer. To support this, he states, first, that *that* RCs and *wh*-RCs can be coordinated, see (38a).

(38) a. Every essay [[which/that she's written] and [that/which I've read]] is on that pile.

b. * Every essay [[she's written] and [that/which I've read]] is on that pile.

(Sag, 1997: 463)

In example (38b), the first clause in the coordination does not contain an overt relative word, which leads to an unacceptable sentence. This shows that *that* RCs pattern with *wh*-RCs rather than with bare RCs.

A second argument for the pronoun status of relative-*that* comes from the observation that, in some varieties of English, *that* has a possessive form as in (39).

(39) The pencil [that's lead is broken] (Sag, 1997: 463)

In a footnote, (Sag, 1997: 463) mentions that the availability of *that* SRCs would be further support for treating *that*-RCs on par with *wh*-RCs, as there are no bare SRCs.

However, Sag concedes that one piece of evidence in favour of considering *that* as a complementizer is, unlike most *wh*-relative pronouns that it doesn't allow upward percolation, see (40a). To weaken this objection, Sag points out that there is a *wh*-relative pronoun that doesn't allow upward percolation either, which is the relative pronoun *who*, see (40b).

(40) a. *The university [in that I graduated __] …

b. *The HR manager [to who I sent my C.V. __] …

Sag does not provide an explanation for this restriction but suggests that it could "have to do with case assignment or registered restrictions, or both" (Sag, 1997: 463). Such an explanation might be possible for *who*. This idea has been worked out in Wilcock (1999).[5] Wilcock assumes that accusative *who* is marked as *informal*, whereas PPs with a non-empty INHER|REL value must be marked as *formal*. As all elements in a clause need to have the same register value, this conflict would immediately capture the ungrammaticality of (40b).

This approach is, however, too restricted. In the naturally occurring example (41), two relative clauses – both IRCs – are coordinated.

(41) When I met people [[who I knew], and [with whom I had not had contact since the event]], few were actually able to say anything.
(https://tinyurl.com/y554nvv5, 19.10.2020)

The first contains an occurrence of a fronted, bare, accusative relative pronoun *who*. This would indicate that the sentence is informal. The second conjunct, however, contains the fronted relative phrase *with whom*, which contains a preposition and the form *whom*. This means that the second conjunct would be marked as formal. Consequently, the sentence should be as ungrammatical as (40b). This shows that, at least at present, there is no HPSG proposal encoding Sag's idea of blocking percolation even for *who* based on case and/or register.

Sag's idea and Wilcock's formalization of it do not carry over to *that* RCs. *that* RCs may contain a gap in non-nominative positions. This is shown in (40a) in which the preposition is stranded, see (42). This sentence is well formed. There are no conceivable differences in case assignment to the complement of the preposition between the two examples in HPSG.

[5] I am grateful to Gert Webelhuth (p.c.) for pointing me to Wilcock's analysis.

(42) The university [that I graduated in __] …

This argumentation excludes a register difference between nominative and accusative uses of relative *that*. Consequently, the surface similarity between *that* and *who* does not provide a basis for explaining the fact that *that* cannot be part of a complex relative constituent.

If the ban on upward percolation with *that* is a question of register, it should still be possible to encode it somehow. For example, relative *that* could, in general, be marked as *informal* – remember that Wilcock (1999) postulates that structures with upward percolation are always marked as *formal*. Just as in the case of relative *who*, it is possible to find naturally occurring examples with coordinated RC in which there would be a register conflict under these assumptions. This is illustrated with example (43), taken from COCA. In this example a *that* IRC is coordinated with an IRC with a complex relative constituent.

(43) About the dark [[that disperses] and [in which everything disappears]].
(COCA)

If *that* and upward percolation with *wh*-relative words can occur in the same register, attempts for a register-based account of the constraint on *that* do not seem to be fruitful.

To sum up, the ban on upward percolation is probably the strongest argument against the pronominal status of *that* – even though, this might be a theory-internal argument that is based on the specific percolation mechanisms assumed in HPSG.

As shown in my analysis of *which* SRCs in chapter 5, I follow Pollard & Sag (1994a) in assuming a phonologically empty relativiser for *wh*-RCs. As far as

that RCs are concerned, I deviate from Pollard & Sag (1994a). They distinguish between relative *that* when it occurs as the local subject of the RC and other uses of relative-*that*. In the first case, they analyze it as a pronoun. In the other cases, they treat it as the ordinary complementizer *that* that marks embedded declarative clauses. This allows them a uniform treatment of bare RCs and non-subject *that* RCs: in both cases, there is an empty relativiser that takes a sentence with a non-subject gap as its complement and binds the slash UDC of this gap in the form of a weak UDC.

Their reason for distinguishing these two types, the subject *that* and the non-subject *that*, is convincing to some extent; however, providing two very different analyses for the same relative word is confusing and not optimal. Arnold & Godard (2019) call for a more uniform treatment of relative *that*, "an analysis which provides a uniform treatment of English *that*-relatives is clearly more appealing." (Arnold & Godard, 2019: 28).

These two analyses of *that* IRCs are a reflex of the special treatment of subjects with respect to extraction in Pollard & Sag (1994a). As they do not allow for traces in subject position, any apparent subject extraction needs to be modelled in an alternative way. However, Levine & Hukari (2006) show that subject extraction should be analysed analogously to complement extraction. Once the ban on traces in subject position is removed, it becomes possible to capture *that* RCs in a single analysis.

I propose to achieve this unification in my analysis of relative *that*. I do not follow the pronoun analysis of Sag (1997) for relative *that*, as I do not think it can capture the upward percolation facts directly. I also diverge from Pollard & Sag (1994a). I do not treat relative *that* as an ordinary complementizer. Instead, I assume that relative *that* is a non-empty form of the relativiser, that is, I consider

relative *that* the head of the RC in all such constructions. Such an approach makes it possible to combine properties of a complementizer approach and a pronoun approach.

An additional reason I argue for this analysis is that, in Chapter Sōrānī 7, the Kurdish relative word *ke* is a complementizer and it behaves very similarly to English relative *that*. Thus, I intend to present a similar analysis for both relative words.

The lexical entry of the syntactic properties of *that* as a relativiser is provided in Figure 6.1.[6] This lexical entry is for relativiser *that* in both SRCs and IRCs. In this version its syntax and a simplified semantics are illustrated.

$$\left[\begin{array}{l}\textsc{phon}\ \langle \textit{that} \rangle \\ \textsc{syns}\ \left[\begin{array}{ll}\textsc{loc} & \left[\begin{array}{ll}\textsc{cat} & \left[\begin{array}{ll}\textsc{head} & \left[\begin{array}{ll}\textit{rltvzr} & \\ \textsc{sel}\ \ \mathrm{N} & \left[\begin{array}{l}\textsc{loc}|\textsc{cont}|\textsc{index}\ \ \boxed{1} \\ \textsc{nloc}|\textsc{to-bind}|\textsc{rel}\ \ \{\boxed{1}\}\end{array}\right]\end{array}\right] \\ \textsc{val} & \left[\begin{array}{ll}\textsc{subj} & \langle\rangle \\ \textsc{spr} & \langle\rangle \\ \textsc{comps} & \left\langle \mathrm{S}\left[\begin{array}{l}\textit{fin},\ \textit{unmarked} \\ \textsc{inher}|\textsc{slash}\ \{\boxed{2}\}\end{array}\right]\right\rangle\end{array}\right]\end{array}\right] \\ \textsc{cont}\ \left[\textsc{index}\ \boxed{1}\right]\end{array}\right] \\ \textsc{nloc} & \left[\begin{array}{ll}\textsc{inher} & \left[\begin{array}{ll}\textsc{que} & \{\} \\ \textsc{rel} & \{\boxed{1}\} \\ \textsc{slash} & \{\}\end{array}\right] \\ \textsc{to-bind} & \left[\begin{array}{ll}\textsc{que} & \{\} \\ \textsc{rel} & \{\} \\ \textsc{slash} & \left\{\boxed{2}\left[\textsc{cont}|\textsc{index}\ \boxed{1}\right]\right\}\end{array}\right]\end{array}\right]\end{array}\right]\end{array}\right]$$

Figure 6.1: Sketch of the lexical entry of *that* as a relativiser

The description in Figure 6.1 looks similar to the null relativiser for bare RCs in Pollard & Sag's (1994a): The HEAD value is of sort *relativiser*. In its COMPS list, there is a finite, unmarked sentence with an element, $\boxed{2}$, in its INHER|SLASH value.

[6] In this figure, the LRS-details are not incorporated yet. As a consequence, the INDEX value is not split into its DR and PHI features.

This element, [2], is identical with the element in the relativiser's TO-BIND|SLASH set. The index of [2] is identical with the index of the antecedent of the RC, as such, it is identical with the INDEX value of the nominal projection in the RC's SEL value. There are, however two differences: First, obviously, it has a non-empty phonology. Second, to be commented on below, it has non-empty INHER|REL set. As I allow for traces in subject position, this provides a unified treatment of *that* RCs independent of whether the relativized element is the local subject of the RC or not.

An important property of my analysis of relative *that* is that it introduces an element in its INHER|REL set. This element is also in the TO-BIND|REL value of the antecedent, the noun in the SEL value. As the INDEX value of the gap and the REL value are identical, any possibility of upward percolation is excluded.

Before I add the semantics to the lexical specification of the *that* relativiser, I show how it differs from the empty relativiser for bare RCs. I provide the lexical entry for the relativiser for bare RCs in Figure 6.2.[7]

For bare relatives, the SUBJ list of the relativiser is empty. Consequently, it looks very much like relative *that*. However, in Standard English, bare relatives are not possible when the relativized element is the subject of the RC. To formalize this restriction, the relativiser must have access to the information on the subject of its complement. There is independent evidence that the information on the subject of a clause is available at the clause level, even if the subject is realized inside the clause. Relevant constructions have been discussed in Höhle (2019) (which was originally written in 1994), and Sag (2012). Here, I make use of their ideas in the following way. I adopt the feature name XARG (EXTERNAL-

[7]In this figure, the LRS-details are not incorporated. As a consequence, the INDEX value is not split into its DR and PHI features.

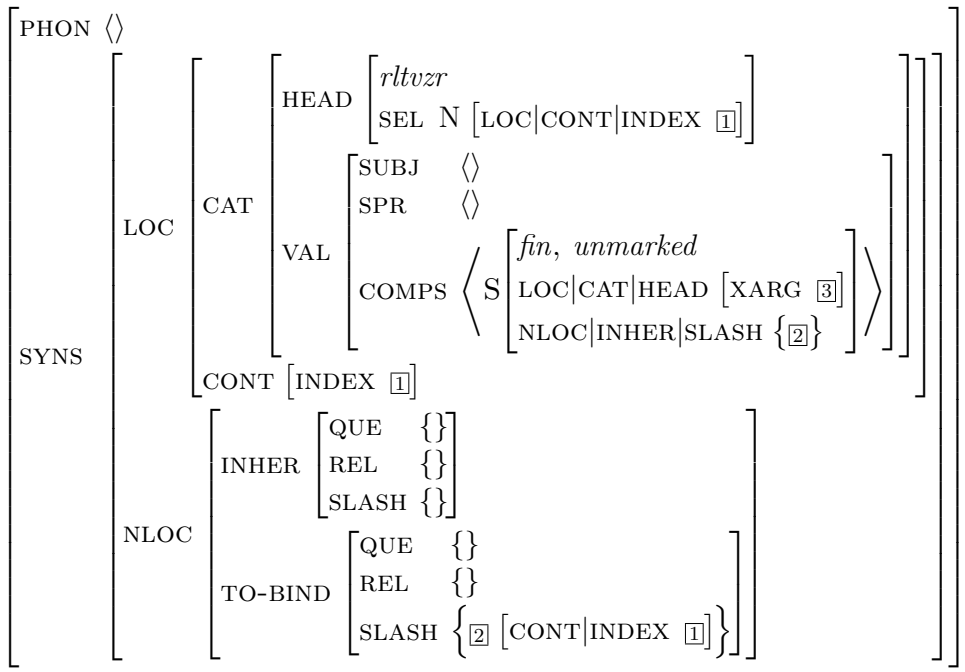


Figure 6.2: Sketch of the lexical entry of the null relativiser for bare RCs

ARGUMENT) from Sag (2012), but, following Höhle (2019), it is a head feature and its value is *none* or a *loc* object. For words with a non-empty SUBJ list, the XARG value is identical with the LOC value of their subject. For other words, XARG has the value *none*.

To exclude bare relatives with a local subject gap, I add the non-identity constraint below the AVM in Figure 6.2: in a bare relative, the subject of the sentence on the relativiser's COMPS list must not be the relativized element, that is the sentence's XARG value 3 must not be identical with the element in its INHER|SLASH set.[8]

The important difference between relative *that* and the null relativiser in Figure 6.2 is that the latter does not introduce an element in its INHER|REL set, whereas

[8]In more permissive varieties that allow for bare RCs with a relativized local subject, this line could simply be omitted.

the former does. This makes the null relativiser for bare RCs look much more like the relativiser I use for *wh*-RCs.[9] However, in the case of *wh*-RCs, the relative constituent will contribute a non-trivial REL value, just as relative *that* does.

At the level of the full RC, *that* RCs pattern with *wh*-RCs: in both cases, there is a non-trivial INHER|REL value. This contrasts with bare RCs, as these always have an empty INHER|REL value. This difference is the basis of an explanation for the coordination facts reported in (38). It is the core assumption of any HPSG approach to coordination that the conjuncts must be of the same – or at least similar – syntactic category. According to Abeillé & Chaves (2019), an identity of HEAD and VAL values is typically assumed.[10] Furthermore, the identity of NLOC features is part of what is required as identity between the conjuncts. This is how Ross's (1967) *Coordinate Structure Constraint* with its *Across-the-Board* exception are modeled in HPSG, thus an element can be extracted from one conjunct if and only if the same element is also extracted from all other conjuncts. In other words, all conjuncts must have the same NLOC features.

I repeat versions of Sag's coordination examples in (44). In (44a), the INHER|REL set of the *which*-RC contains the index value of the antecedent, the noun *essay*. The same is true for the *that* RC. Therefore, both RCs have the same INHER|REL value. In (44b), on the other hand, the bare RC has an empty INHER|REL value. Consequently, the two RCs do not have the same NLOC values and cannot be coordinated.

(44) a. Every essay$_i$ [[which$_i$ she's written] and [that$_i$ I've read]] is on that pile.

[9]It is, therefore possible to unify these two null relativiser into one item with an optional element on the SUBJ list. I refrain from doing this here as such a unified lexical entry would involve disjunctions depending whether or not the SUBJ list is empty. For this reason, it is not obvious whether much is gained from unifying them into one lexical entry.

[10]See there for the merits and problems of assuming this identity for all conjuncts.

b. * Every essay$_i$ [[she's written] and [that$_i$/which$_i$ I've read]] is on that pile.

(Sag, 1997: 463)

In this section, I provided an account of *that* RCs based on treating relative *that* as a relativiser. I showed how this analysis captures the similarities and differences between *that* RCs, *wh*-RCs, and bare RCs. In the next section, I extend my analysis to *that* SRCs.

6.5 The analysis of *that* SRCs

In this section, I provide an analysis of *that* with SRCs. In Figure 6.3, I present the lexical entry of relativiser *that* as used in *that* SRCs. The syntactic analysis is the same as in Figure 6.1, but now, I add its detailed semantics.

The relativiser is the head of the RC. Its SEL value requires an NP that contains its own index, $\boxed{1}$, in its TO-BIND|REL set. The COMPS list of relative *that* has an unmarked sentence. The relativizer has an element, $\boxed{2}$, in its INHER|SLASH value, which it binds in its TO-BIND|SLASH set. The INDEX value of the extracted element is identical with that of the relativiser itself, as indicated with the tag $\boxed{4}$. The antecedent and the relativiser have the same PHI value, $\boxed{3}$, but different DR values: x for the antecedent and w for the relativiser in Figure 6.3.

The LF value of the relativiser has a conjunction as its EXCONT value, $\gamma \wedge \delta$. The INCONT value equates the DR values of the antecedent and the RC, written as $w = x$. The PARTS list contains the relativiser's DR, INCONT, and EXCONT values and, in addition, the definite description that is typical for a discourse pronoun, $\iota w : w = x$, i.e., that it refers to the single element w that is identical with x. Below the AVM, there is the restriction that this definite description occurs inside the

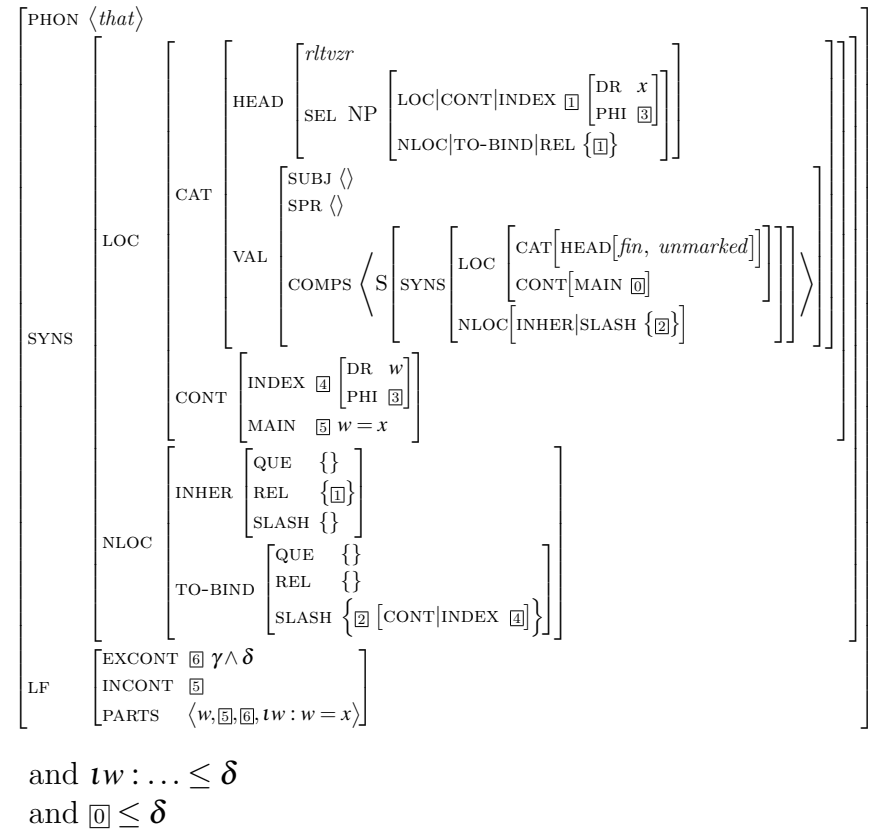


Figure 6.3: Lexical entry of relativiser-*that* with LRS specification

second conjunct of the relativiser's EXCONT value, i.e. as component of δ.

After presenting the syntax and semantics of the relativiser *that* in its lexical entry I show an example analysis of a *that* SRC. Figure 6.4 shows the highest local tree of the SRC in the example in (45).

(45) Alex watched *High Noon*, that Kim had recorded.

$$\mathbf{watch'}(\mathbf{alex'}, \mathbf{HN'}) \wedge \mathbf{record'}(\mathbf{kim'}, \iota w : w = \mathbf{HN'})$$

The relativiser is the head of the tree in Figure 6.4. It takes an unmarked sentence, 8, as its complement. The sentence contains a gap, which means that it has one element in its INHER|SLASH set, indicated with the tag 2. The relativiser has this element in its TO-BIND|SLASH set. Consequently, the INHER|SLASH set of

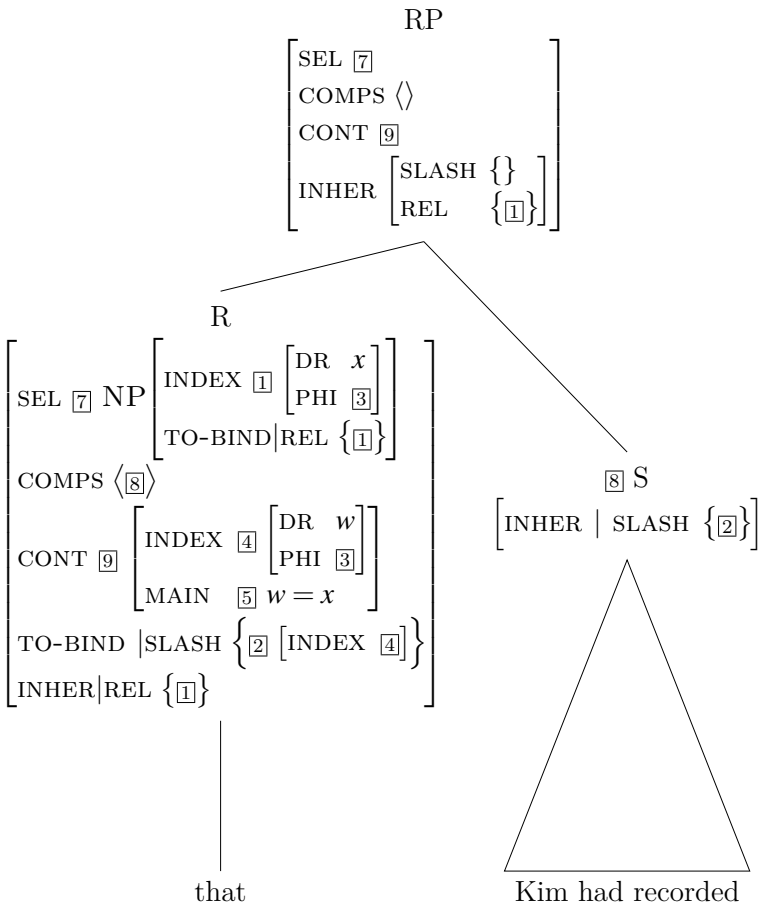


Figure 6.4: Structure of the *that* SRC from example (45)

the overall RC is empty. The INDEX value of the relativiser is identical with that of the gap, 4 – though, crucially, not with that of the antecedent of the relative clause (1).

The relativiser has a *synsem* object as its SEL value, 7. This is the antecedent of the RC. The relativiser also has an element in its INHER|REL set, 1. The externally selected element binds the REL-dependency, i.e., it has the relativiser's INHER|REL in its own TO-BIND|REL set. Furthermore, the INDEX value of externally selected element is identical with the element in the REL sets.

The connection between the index of the antecedent and that of the SRC is

made in two ways: First, they are required to have the same PHI value, indicated with the tag [3] in Figure 6.4. Second, the relativiser introduces an identity statement, $w = x$, where w is the DR value of the RC and x is the DR value of the antecedent.

We can also take a look at the combinatorial semantics of the example sentence. In Figure 6.5, I show the LF value of the core SRC *Kim recorded* __. As it contains a gap, the second argument of the predicate $\mathbf{record'}$ is underspecified. All we know is that the gap's DR value, w, appears inside α. This is expressed by the constraint below the AVM.

$$\left[\begin{array}{ll} \text{LF} & \left[\begin{array}{ll} \text{EXCONT} & \boxed{9}\ \mathbf{record'}(\mathbf{kim'}, \alpha) \\ \text{INCONT} & \boxed{9} \\ \text{PARTS} & \left\langle \mathbf{kim'}, \mathbf{record'}, \boxed{9}, w \right\rangle \end{array}\right] \end{array}\right]$$

and $w \leq \alpha$

Figure 6.5: Value of the LF feature of the core SRC in (45)

The relevant semantic specification of the relativiser is given in Figure 6.6, i.e. its CONTENT and LF values. I use the same tags as in the lexical entry of the SRC relativiser *that* in Figure 6.3 above.

When the two signs combine, we get the semantic specifications in Figure 6.7. The CONT value of the RC is identical with that of the relativiser, and so are the EXCONT and INCONT values. The PARTS list now contains all the elements of the PARTS lists of the daughters. Finally, the two embedding constraints also apply to the overall semantics, which is why I collect them below the AVM.

Given these constraints, there is exactly one way to satisfy them in the seman-

$$\left[\begin{array}{ll} \text{SYNS} \mid \text{LOC} \mid \text{CONT} & \left[\begin{array}{ll} \text{INDEX} & \left[\begin{array}{ll} \text{DR} & w \\ \text{PHI} & \boxed{3} \end{array}\right] \\ \text{MAIN} & \boxed{5}\ w = x \end{array}\right] \\ \text{LF} & \left[\begin{array}{ll} \text{EXCONT} & \boxed{6}\ \gamma \wedge \delta \\ \text{INCONT} & \boxed{5}\ w = x \\ \text{PARTS} & \left\langle w, \boxed{5}, \boxed{6}, \iota w : w = x \right\rangle \end{array}\right] \end{array}\right]$$

and $\iota w : \ldots \leq \delta$

Figure 6.6: Values of the CONT and LF features of the SRC relativiser *that*

$$\left[\begin{array}{ll} \text{SYNS} \mid \text{LOC} \mid \text{CONT} & \left[\begin{array}{ll} \text{INDEX} & \left[\begin{array}{ll} \text{DR} & w \\ \text{PHI} & \boxed{3} \end{array}\right] \\ \text{MAIN} & \boxed{5}\ w = x \end{array}\right] \\ \text{LF} & \left[\begin{array}{ll} \text{EXCONT} & \boxed{6}\ \gamma \wedge \delta \\ \text{INCONT} & \boxed{5}\ w = x \\ \text{PARTS} & \left\langle \begin{array}{l} w, \boxed{5}, \boxed{6}, \iota w : w = x \\ \mathbf{kim}', \mathbf{record}', \boxed{9}, w \end{array} \right\rangle \end{array}\right] \end{array}\right]$$

and $\iota w : \ldots \leq \delta$
and $w \leq \alpha$

Figure 6.7: Values of the CONT and LF features of the SRC *that Kim had recorded*

tic representation of the sentence, repeated in (46). In (46a) and (46b), I show the values of the meta-variables α and δ. These clearly satisfy the embedding conditions imposed by the lexical entries.

(46) $\gamma \wedge \mathbf{record}'(\mathbf{kim}', \iota w : w = \mathbf{HN}')$

a. $\alpha \equiv \iota w : w = x$

b. $\delta \equiv \mathbf{record}'(\mathbf{kim}', \iota w : w = \mathbf{HN}')$

To sum up, the occurrence of relative *that* in SRCs seems to be acceptable in the American English. After having established this new empirical basis, I

presented an analysis of *that* RCs in which *that* is treated as a relativiser. I showed that my approach solves some of the problems that led Sag (1997) to propose a pronoun analysis of *that* RCs. Finally, I combined my syntactic analysis of *that* RCs with the semantics that I proposed for SRCs in chapter 5.

In chapters 4–6, I showed that both determiner-*which* SRCs and *that* SRCs should be accounted for within a formal grammar of English and I presented a lexicalist analysis, based on a lexical entry for determiner-*which* and for relativiser-*that*. This made it necessary to look at *wh*-SRCs in general, bare RCs, and *that* IRCs. This allowed me to show that my analysis fits into a general approach of RCs in HPSG which is a variant of the analyses in Pollard & Sag (1994a). A compact theory, a more simplified version, of the analyses presented in this chapter will be presented in chapter 8.

Chapter 7

Relative clauses in Sōrānī Kurdish

After the discussion of two lesser studied types of English SRCs, I will turn to RCs in Sōrānī Kurdish. I argue here that they share essential properties with English *that* RCs.

In section 7.1, I give an overview of the background of Kurdish. In its first subsection, 7.1.1, I present the three main dialects of Kurdish. In section 7.1.2, I summarize the main phonological elements of the language. In addition, I illustrate the most relevant syntactic aspects of Sōrānī Kurdish in section 7.1.3.

In section 7.2, I give an overview of the properties of Kurdish RCs. I first show that there are barely any restrictions with respect to the antecedent of RCs in Kurdish in section 7.2.1. Furthermore, I discuss the occurrence restrictions of the Sōrānī Kurdish relativiser *ke* in section 7.2.2, showing that it is largely analogous to English relative *that*. In section 7.2.3, I examine how Sōrānī Kurdish marks the relativised element and I show that there is an unbounded dependency in Kurdish RCs, even though they do not contain a gap. In section 7.2.4, I show the

structure, rules and the analysis of the Ezafé marker in Kurdish RCs. In section 7.2.5, I make a comparison between Sōrānī Kurdish IRCs and SRCs.

I address similarities and differences between Sōrānī Kurdish RCs and Persian RCs in section 7.3. My own analysis is presented in section 7.4, and I will restrict my attention to the specification of the relativiser. A summary concludes the chapter in section 7.5.

7.1 The background of the language

Defining a language that, so far, does not have a precise geographic state and a standard or a unified writing system is not straightforward. Sōrānī Kurdish and its various dialects has spread across neighboring geographical areas but remains understudied. Kurdish is an Indo-European language belonging to the Iranian language group.

At present, Kurdish is spoken as a first language by a majority of people in a contingent area that covers parts of Turkey, Iran, Iraq, and Syria. In the Kurdish community, this area is called Greater Kurdistan.

Due to the fact that there is not an independent Kurdistan and speakers are distributed across four countries, there is no standard language or unified writing system for Greater Kurdistan. In three of the four parts of the Greater Kurdistan, the writing system is affected by the standard writing system of the official language of the country in which that particular part is located. The Kurdo-Arabic script is used in Iraq and Iran. The Turkish script with a slight change is used in Turkey and Syria. If it were consistent, the Arabic script would be used by the Syrian Kurds; however, the Syrian Kurds also use the Latin script. This indicates that it is not really the country's majority that triggered the choice

of script, but rather the linguistic affiliation. The Syrian Kurds speak Kurmānji and they use a version of the Latin alphabet that has the Turkish script as a model, but it is not the complete Turkish script that is used; Syrian Kurmānji makes use of more symbols than Turkish.

7.1.1 The dialects of Kurdish

The classification of Kurdish dialects is still an open line of enquiry and new models are being constantly proposed. There is at least one widely acknowledged model by which Kurdish is divides into Northern, Central (Sōrānī), and Southern varieties. This classification is observed in Haig (2007) among others. Öpengin (2013: 1) states that "Central Kurdish or more widely Sōrānī Kurdish is one of the principal Kurdish varieties spoken in Iraq and Iran by the majority of Kurdish population in these two countries. We are not aware of any academic sources citing the number of its speakers." Haig (2007: 165) states that the northern group, Kurmānji Kurdish, is the largest as it encompasses the Kurdish of Turkey, Syria and a part of Iraqi Kurdistan in Zaxo and Dohuk. He adds that the speakers of this Northern variety had language contact with Armenian, Arabic, Persian, Georgian, Turkish and Russian. Kurdish dialects are quite divergent in lexical and grammatical features. As mentioned earlier, there has not been any successful standardization for Kurdish in Greater Kurdistan. In particular, none of its dialects has officially been recognized as a standard. Nonetheless, Sōrānī, the Central Kurdish and especially the dialect of the city of Sulaymaniyah in north-eastern Iraq, might be on its way to being recognized as standard Kurdish. One reason for this could be that South Kurdistan is a semi-independent region in Iraq and Sōrānī is the dominant dialect of the region; Sōrānī Kurdish is the dom-

inant regional language of Iraqi-Kurdistan. This particular dialect, Sulaimaniyah Sōrānī, is investigated in this chapter. In the following, I will refer to this as Sōrānī Kurdish.

7.1.2 The phonology of Sōrānī Kurdish

In this section, I will not present a thorough characterization of the Sōrānī Kurdish sound system, I rather introduce only the transcription system of the language that I will be using throughout the chapter. I use transcription because the dialect I am studying, Sōrānī Kurdish, is written in Arabic script. The system I use is the one presented in Asadpour & Mohammadi (2002) with minor changes.[1] My transcription of the vowels is given in Table 7.1, followed by the diphthongs in Table 7.2 (page 205), and the consonants in Table 7.3 (page 205).

high front vowel (short i)	/i/	/bərsi/ hungry
high front vowel (long i)	/i͡/	/kurdī/ Kurdish
semi-high front vowel (short e)	/e/	/estā/ now
semi-high front vowel (long e)	/e͡/	/sé/ three
back low vowel (short a)	/a/	/dast/ hand
back low vowel (long a)	/ā/	/āzā/ brave
semi-high central vowel (schwa)	/ə/	/pər/ full
semi-high back vowel (short o)	/o/	/otāq/ room
semi-high back vowel (long o)	/ō/	/ōmat/ nation
high back vowel (short u)	/u/	/kurtān/ pack-saddle
high back vowel (long u)	/ū/	/sūr/ red

Table 7.1: Sōrānī Kurdish simple vowels

[1] See for example Thackston (2015: 1–7) for a different system.

/éw/	/sew/ apple
/ey/	/dey/, /ley/ prepositions
/aw/	/šaw/ night
/āw/	/āw/ water
/āy/	/dāyk/ mother
/oy/	/royštən/ to go
/ūy/	/fūy̆léke/ blow

Table 7.2: Sōrānī Kurdish diphthongs

	Bilabial	Labio-dental	Dental, Alveolar	Post Alveolar	Velar	Uvular	Pharyneal	Glottal
Stop	p b		t d		k g	q		ʔ
Affricate				č j				
Fricative	f v	s z	š ž	x ɣ		(ʕ) ħ	h	
Nasals	m		n		ŋ			
Lateral			l ɬ					
Flap			s					
Trill			ɾ r					
Glide	w			y				

Table 7.3: Sōrānī Kurdish consonants

7.1.3 The morphosyntax of Sōrānī Kurdish

Here, I present some prominent morpho-syntactic characteristics of Sōrānī Kurdish.

In contrast to English, Kurdish is an agglutinative language. Consequently, a sentence can consist of one word since subjects and objects in Kurdish can appear in the form of dependent morphemes. This is illustrated in (1).

(1) de-y-xwén-īt.
IPFV-3.SG-*study*.PRES-2.SG
'you are studying it'

Khalid & Hamamorad (2015: 3) state that there even are cases where the subject and/or the object of the sentence become zero morphemes and do not have any phonological exponent in the sentence, see (2).

(2) hāt
come.PST.3.SG
'She/he came'

In Sōrānī Kurdish, there is neither grammatical case nor grammatical gender, and there is no case marking on nominal elements. This is shown in (3)–(5). The form of the first person, *min* 'I' is constant, independently of its grammatical function in the clause (Haig, 2008: 277-279).

(3) min rōš=im
I *leave*.PST=1.SG
'I left'

(4) ew min=ī bīnī
he *I*=3.SG *see*.PST
'He saw me'

(5) min ew=m bīnī
I *he*=3.SG *see*.PST
'I saw him'

Definiteness in Sōrānī is indicated by the suffix -*eke*, and the suffix -*ā(n)* indicates plurality. Usually the plural suffix appears with the definite suffix as in example (6).

(6) kiç-ek-ān
girl-DEF-PL
'the girls'

One morphological rule by which many Sōrānī Kurdish words are linked together is called the *Ezafé (EZ) construction*. In defining Ezafé, Haig (2011: 363) says that the term *Ezafé* "is adopted from Arabic grammar (idāfat), where it means 'addition, supplement.' Within Iranian linguistics it is used to refer to an unstressed vocalic particle which occurs between a noun and an adjective or other nominal modifier."

In the Ezafé construction, a noun in receiving possession is connected to another noun, in a way that the possessed precedes the possessor.[2]

Here I use EZ to gloss the Ezafé morpheme.

The EZ morpheme is *e* for definite NPs, *ī* for indefinite NPs ending in consonants and *j* for indefinite NPs ending in vowels, see (7) and (8).

(7) kič-e baš-eke
Girl-EZ *good*-DEF.SG
'the good girl'

(8) kič-ī šār
girl-EZ *city*
'city girl'

I discuss more about Ezafé in Sōrānī Kurdish in Section 7.2.4.

Sōrānī Kurdish has several types of pronouns: personal, demonstrative, and interrogative pronouns. In this section, I focus on personal pronouns, as they play an important role in Sōrānī Kurdish RC constructions. I will not discuss demonstrative and interrogative pronouns. I address the question as to whether Sōrānī Kurdish has a relative pronoun in section 7.2.2.

In Sōrānī Kurdish, various forms of personal pronouns can be distinguished, such as independent pronouns, clitic pronouns, and verbal inflections. Samvelian

[2]Kim (2010: 18) refers to Ezafé as *Izafe*. He characterizes it as follows: "The *Izafe* links the two parts of a possessive construction and is equivalent to the English 'of.' The *Izafe* is written as "ی" and added directly to **the first part** of the construction" (emphasis in the original).

(2007b) subsumes them under the term *personal morphemes*. I summarize the three types of personal morphemes in Table 7.4, compiled from Samvelian (2007b: 242–243)

Independent pronouns			Clitic pronouns			Verbal endings		
	Sg	Pl		Sg	Pl		Sg	Pl
1	min	eme	1	-(i)m	-mān	1	-(i)m	-īn
2	tō	ewe	2	-(i)t	-tān	2	-ī(t)	-(i)n
3	ew	ewān	3	-ī/-y	-yān	3	-ē(t)/Ø	-(i)n

Table 7.4: Kurdish pronouns (Samvelian, 2007b: 242–243)

As shown in Table 7.4, the language has both independent personal pronouns and clitics which distinguish persons and numbers. An example of the use of an independent pronouns is given in (9).

(9) **min** xełk-ī hewlér=im.
I *people*-EZ *Hewlér*=1SG
'I am from Hewlér.'

The clitic pronouns are added to verbs as well as to nouns and they are mobile, that is, they can appear at the end or in the middle of their host word Öpengin (2013: 290). Haig (2008: 283) states that pronoun clitics express a nominal constituent in some syntactic functions namely: the possessor of a possessed NP as in (10a), the object of a preposition, an indirect participant, an object of transitive verb in present tense, as in (10b), and a subject of a transitive in past tense.

(10) a. kiç-eke**=tān**
girl-DEF=3.PL
'your daughter.'

b. māł=**im** le hewlē=e.
house=1SG *in Hewlē*=COP.PRES.1SG
'My house is in Hewlēr.'

Verbal inflections are affixes that attach to verbal stems and show agreement, as in example (11).

(11) hat-**im** bō hewlér.
come.PST-1SG *to hewlér*
'I came to Hewlēr.'

Sōrānī Kurdish shows split ergativity. Haig (2008: 7) describes this using the parameters grammatical function (S: subject of an intransitive verb, A: subject of a transitive verb, and O: object of a transitive verb) and tense (present and past). As observed in MacKenzie (1962) and Haig (2008), in the present tense, the verb agrees with S and A, see (12a) and (12b) respectively. In the past tense, the verb agrees with S and O, but not with A, as in (13a) and (13b).

(12) a. min de-xew-im.
I IPFV-*sleep*-1.SG.
'I am sleeping' (agreement with S)'

b. tō çī de-ke-yt léra?
you what IPFV-*do*-2.SG.PRES *here*
'what are you doing here?' (MacKenzie, 1962: 4) (agreement with A)'

(13) a. min xewt-bū-m.
I *sleep*-PST.CONTN-1.SG.
'I was sleeping' (agreement with S)'

b. tō çī=t de-kird lére?
you what=2.SG IPFV-*do*.PST *here*
'what were you doing here¿ (agreement with O)'

As I show in section 7.2.3, personal morphemes play an important role in Sōrānī Kurdish RCs, as they are used to mark the relativised element in situ.

7.2 Relative clauses in Sōrānī Kurdish

In this section, first, I show the possible antecedents of Sōrānī Kurdish RCs in 7.2.1. I discuss the function of the relative *ke* and the possibility of omitting it in Sōrānī Kurdish RCs in 7.2.2. I present the resumptive clitics in section 7.2.3, in this section, I also show that there is UDC in Sōrānī Kurdish RCs, and discuss the island constraints regarding Kurdish resumptive clitics. I illustrate the occurrences of Ezafé with IRCs and SRCs in detail in section 7.2.4, and I discuss the similarities and the differences between Sōrānī Kurdish IRCs and SRCs in 7.2.5.

7.2.1 Possible antecedents

Khalid & Hamamorad (2015) present various types of Sōrānī Kurdish subordinate clauses including RCs. They refer to Sōrānī Kurdish RCs as *determinative subordinate clauses*. They give the following characterization: "In complex sentences, sometimes a subordinate clause takes determinative role, and it will define a word in the main sentence. So, this subordinate clause is called determinative subordinate clause. It takes the grammatical role of the defined noun in the main clause." (Khalid & Hamamorad, 2015: 11). They also focus on the grammatical function of the antecedent in the matrix clause.

Khalid & Hamamorad (2015: 11-13) show that the antecedent of RCs can have any grammatical function in the main clause. I repeat their examples in (14)–(18). In the following examples the RCs are introduced by *ke*, which is glossed beneath the examples as KE. A discussion on relative *ke* will be presented in section 7.2.2. The examples are from Khalid & Hamamorad (2015) with my translation and

glossing.[3]

1. RC modifying a subject:

   (14) māzin, ke hāweł-ék-ī dilsoz-ī min bū, koč-ī
   Mazn KE *friend*-INDEF-EZ *loyal*-EZ *I* COP.PST.3.SG *migrate*-EZ
   dwāy kird.
   final *do*.PST.3.SG
   'Mazn, who was my loyal friend, passed away.'

2. RC modifying a complement:

   (15) ew mindāł-ék=ī hénā, ke zōr ru-xoš bū
   (s)he child-INDEF=3.SG *bring*.PST KE *very face-nice* COP.PST.3.SG
   '(S)he brought a child who was very lovely.'

3. RC modifying the complement of a preposition:

   (16) rébwār le šār, ke žīnge pāk n=ye,
   Rebwar in city KE *environment clean* NEG=COP.PRES.3.SG
   bé-zār=e
   without-mouth=COP.PRES.3.SG.
   'Rebwar is sick of cities in which the environment is not clean.'

4. RC modifying a predicative complement:

   (17) gewre-tirīn tāwān-ī rižīm-ī péš-ī érāq ĵénōsāyd
   big-SUPL *crime*-EZ *regime*-EZ *former*-EZ *Iraq genocide*
   bū, ke tāwān-ék-ī
   COP.PST.3.SG KE *crime*-INDEF-EZ
   néw-dewłet-y=e
   inter-country-DER=COP.PRES.3.SG
   'The biggest crime of the former Iraqi regime was genocide, which is an international crime.'

[3]There are no fixed rule for placing commas with RCs in Kurdish. We typically find a comma in front of an SRC, but there is variation with respect to IRCs. In quoted examples, I follow the punctuation of the original. In my own examples, I will not use a comma with IRCs.

5. RC modifying a modifier of a main clause subject:

(18) réše-bā-y slémānī, ke pāytext-ī rōšin-bīr-ī
black-wind-EZ *Suleymaniyah* KE *capital*-EZ *bright-think*-EZ
kurdistān=e, zōr be-héz=e
Kurdistan=COP.PRES.3.SG *very with-power*=COP.PRES.3SG
'The wind of Sulaimanyah, which is the cultural capital of Kurdistan, is very strong.'

The examples illustrate that any NP from the matrix clause can function as antecedent for an RC in Sōrānī Kurdish.

We saw in the discussion of English RCs that English SRCs can take non-nominal antecedents. This is not the case in Sōrānī Kurdish. A relevant English example is given in (19).

(19) Olan arrived late today, which annoyed everyone in the family.

As shown in (20), this cannot be expressed with an RC in Kurdish. Instead, a coordination of two sentences would be used, as in (21)

(20) *olan emro direng geīšt, ke hemw xézen-eke=y
Olan tonight late arrived KE *all family*-DEF=3.SG
bé-zar-kird
without-mouth-make.PST
Putatively: (19)

(21) olān emrō direng geyšt, eme-(š) hemw xézen-eke=y
Olan tonight late arrived this-ADD *all family*-DEF=3.SG
bé-zār-kird
without-mouth-make.PST
'Olan arrived late today, (and) this annoyed everyone in the family.'

It should be noted, though, that while English allows for non-nominal antecedents in SRCs, this is restricted to *wh*-SRCs. An example analogous to (19)

with *that* instead of *which* is only grammatical if *that* is interpreted as a demonstrative pronoun, not as a relative pronoun, see (22). This points to a similarity between English *that* RCs and Kurdish *ke*-RCs that we will encounter throughout this chapter.

(22) *? Olan arrived late today, that annoyed everyone in the family.

In this section, I showed how the antecedents of Kurdish RCs can be of any grammatical function. Additionally I pointed out that Kurdish RCs cannot take a non-nominal antecedent – unlike English *wh*-RCs, but just like English *that* RCs,

7.2.2 The relative marker *ke*

In all examples of Sōrānī Kurdish presented so far, the RC was introduced with the word *ke*, which I glossed as KE. I now show that Kurdish relative *ke* is strikingly similar to English relative *that*, as discussed in Chapter 6.

Fattah (1997: 254) refers to *ke* as a *subordinator*, and in case of an RC, a *complementizer*. He acknowledges that it is called a relative pronoun in some descriptions of Sōrānī Kurdish. However, he argues that it differs substantially from pronouns as it does not show number and person marking and it is syntactically restricted to the initial position of an embedded clause.

Recent publications in which relative *ke* is referred to as a pronoun include Kim (2010) and Saady (2020). As these two are rather descriptive accounts, they do not discuss the exact categorial status of *ke* in detail.

I adopt Fattah's position here. I illustrate his argument about the syntactic difference with other types of pronouns by comparing it with interrogative pronouns. In Germanic and Romance languages, relative pronouns occur at the beginning of a clause, which is also a possible position for interrogative pronouns.

Kurdish has interrogative pronouns, such as *ké* 'who' or *čī* 'what'. As shown in (23) and (24), these interrogative pronouns are realised in situ in Sōrānī Kurdish. For comparison, I provide a non-interrogative clause in (23b) and (24b) in addition.

(23) a. ké em guł-e de-čén-ét?
who DEM *flower*-DEM IPFV-*plant*-3.SG
'Who is planting this flower?'

b. birā-ke=m em guł-e de-čén-ét.
brother-DEF=1.SG DEM *flower*-DEM IPFV-*plant*-3.SG
'My brother is planting this flower.'

(24) a. rōnā čī de-čén-ét?
Rona what IPFV-*plant*-3.SG
'What is Rona planting?'

b. rōnā guł de-čén-ét.
Rona flower IPFV-*plant*-3.SG
'Rona is planting a flower.'

In RCs, there is no way to use relative *ke* in the position of a putative gap as it must be placed at the beginning of the RC, as shown in the corresponding examples in (25).

(25) a. birā-ke=m, ke ew guł-e de-čén-ét, hešt sāł=e.
brother-DEF=1.SG KE DEM *flower*-DEM IPFV-*plant*-3.SG *eight* *year*=COP.PRES.3.SG
'My brother, who is planting that flower, is 8 years old.'

b. ew guł-e-y rōnā de-y-čén-ét zor serinj rākéš=e.
DEM *flower*-DEM-EZ *Rona* IPFV-3.SG-*plant*-3.SG *very interesting* *drag*=COP.PRES.3.SG
'The flower which Rona is planting is very interesting.'

All in all, I agree with Fattah (1997) that relative *ke* should not be treated as a pronoun. This will be reflected in my analysis in section 7.4, where I propose that Sōrānī Kurdish relative *ke* should be analyzed analogously to English relative *that*, as in chapter 6.

A second important aspect of relative *ke* is its optionality in many RCs. Fattah (1997: 255) states that in Sōrānī Kurdish IRCs *ke* is optional, as shown in (26).

(26) ew žin-e-y (ke) xwéndin=ī tewāw-kird qise=y
DEM *woman*-DEM-EZ KE *study*=3.SG *finish-make*.PAS.3.SG *speak*=3.SG
kird.
make.PST.
'The woman who graduated spoke.'

Fattah proposes that relative *ke* can be easily deleted in the sentence as it has no grammatical or syntactic functions. He adds that the clause stays syntactically complete and semantically related to the main clause when *ke* is absent.

Fattah's claim that *ke* can be freely deleted in all IRCs is challenged in Kim (2010). Kim argues that there are types of IRCs in which *ke* is obligatory. In the rest of this subsection, I elaborate further on this point.

Thackston (2015: 73) and Kim (2010: 88) relate the optionality of relative *ke* to the grammatical function of the relativised element. In example (27), where the relativised element is the object of the verb inside the RC, *ke* can be omitted. In example (28), where the relativised element functions as the subject inside the RC, *ke* must be present.

(27) ewe guł-eke=ye ew pé=y dā-m
DEM *flower*-DEF=COP.PRES.3.SG *he* *to*=3.SG *give*.PST-1.SG
'This is the flower he gave me' (Kim, 2010: 88)

(28) a. *qāp-eke-y grān bū šika
plate-DEF-EZ *expensive* COP.PST.3.SG *break*.PST
'The plate that was expensive broke.'

b. qāp-eke-y ke grān bū šika
plate-DEF-EZ KE *expensive* COP.PST.3.SG *break*.PST
'The plate that was expensive broke.'

Kim does not state that sentences such as (28a) are completely ungrammatical. She just observes that relative *ke* is usually present in these cases. Saady (2020) argues that the only way to reduce Kurdish RCs is by omitting the relative *ke*. In example (29), the relativised element is the subject of the RC hosting it. Nonetheless, relative *ke* can be omitted for some speakers according to Saady (2020).

(29) ew šofér-e-y (ke) be héwašy otōmbél le-de-xiwr-ét
DEM *driver*-DEM-EZ KE *with slowly automobile* PRE-IPFV-*drive*-3.SG
selāmet=e.
safe=COP.PRES.3.SG
'The driver who drives cars slowly is safe'

(Saady (2020: 114), my gloss and translation)

The occurrence restriction on relative *ke* looks very similar to what we find for English relative *that* in IRCs. However, Kurdish speakers might be more tolerant in accepting the omission of the relativiser than English speakers.

It is not fully clear if the parallelism to English also holds for cases in which the relativised element is more deeply embedded inside the RC. In English, relative *that* is optional in such cases, independently of the grammatical function of the gap associated with the relative phrase. Huddleston & Pullum (2002: 1047) show that relative *that* is obligatory if the gap is the local subject of the RC, see (30a). Relative *that* is, however, optional if the gap is a subject of an another clause inside the RC, see (30b).

(30) a. I want a car [*(that) __ is safe].

b. I want a car [(that) I know [__ is safe]].

(Huddleston & Pullum, 2002: 1047)

I argue in section 7.2.3 that Sōrānī Kurdish RCs contain an unbounded dependency as well. It seems to me that relative *ke* cannot be omitted when the relativised element is deeply embedded inside the RC. I illustrate this with my own judgments in (31). In both examples, the relativised element originates within a complement clause inside the RC. In (31a), it is the subject, in (31b), it is a complement. For me, only the versions with an overt relative *ke* are possible in either case.

(31) a. amīrā čend žin-ék-y dī [*(ke) līyān pé=y
Amira some woman-INDEF-3.SG *see*.PST KE *Liyan* PRE-3.SG
wā-ye hāz=yān le kitéb=e]
so-POST *like*=3.PL *of book*=COP.PRES.3.PL.
'Amira met women that Liyan thinks $\_\__i$ like books.'

b. amīrā čend žin-ék-y dī [*(ke) līyān pé=y
Amira some woman-INDEF-3.SG *see*.PST KE *Liyan* PRE-3.SG
wā-ye [(ke) alī le zānkō-we de=yān-nās-ét]].
so-POST KE *Ali from university-post* IPFV=3.PL-*know*-3.SG.
'Amira met women$_i$ that Liyan thinks Ali knows $\_\__i$ from university.'

At the moment, I cannot determine whether there is variation with respect to allowing this type of long extraction without relative *ke*. I am not in a position either to explore whether the non-omitability is just a processing constraint or whether it is a constraint on the grammaticality of the sentence. For this reason, I will not attempt to integrate this constraint in my analysis of Kurdish RCs in section 7.4 below.

Let me now turn to a case in which the occurrence of *ke* is obligatory with IRCs for all speakers of Sōrānī Kurdish. RCs can be extraposed to the right side

of the verb. i.e. the verb can separate the antecedent from the RC. When an IRC is extraposed, the relative particle cannot be omitted (Fattah, 1997: 258), as shown in (32). This contrasts with the non-extraposed version of the same sentence, in (33), where *ke*-omission is available.[4]

(32) šuše-ke šika *(ke) tō kirībw=t bō=m.
glass-DEF *break*.PST.3.SG KE *you buy*.PST=2.SG *for*=1.SG
'The bottle broke *(that) you brought.'

(33) šūše-ke-y (ke) tō hénā=t šika.
glass-DEF-EZ KE *you bring*.PST=2.SG *break*.PST.3.SG
'The bottle (that) you bought for me broke.'

This constraint is exactly what we find for *that* relatives, as indicated in the translation below the examples.

As shown in (27), the conditions for omitting *ke* in Sōrānī Kurdish RCs can be met by some IRCs. However, *ke* is required to be present in Sōrānī Kurdish SRCs.

(34) *guł-eke=m dā be ānnā, hāw-pōl-y tō=ye
flower-DEF=1SG *give*.PST *to Anna same-class*-3.SG *you*=COP.PRES.3.SG
'I gave the flower to Anna, who is your colleague'

(35) *peyām-eke=m dā be haeve, to bānghéšt=it
message-DEF=1SG *give*.PST *to Haeve you invite*=2.SG
kird=bū
make=COP.PST.3.SG
'I gave the message to Haeve, who you invited'

I conclude that relative *ke* can be present in all RCs. It can be omitted in IRCs, but many speakers don't accept omission if the relativised element has the

[4]While all speakers agree on the obligatoriness of *ke* in extraposed RCs, the acceptability of extraposition may vary. It is known that the acceptability of extrapositon depends on a number of factors. For example Walker (2017: chapter 3) shows that the acceptability of extraposition in English is influenced by definiteness of the antecedent and the verb class in the matrix clause, but that these factors should be considered soft constraints. The same seems to be true for Sōrānī Kurdish.

role of the grammatical subject of the RC. Relative *ke* is obligatorily present in extraposed RCs. These constraints are strikingly similar to the distribution of relative *that* in English IRCs.

We saw in the discussion of English relative *that* that a major argument in favor of its non-pronominal status concerned its ban on upward percolation. The same can be observed for Sōrānī Kurdish *ke*: This is true for both IRCs as in (36b), and SRCs as in (37b).

(36) a. ew pyāw-e-y ke legeł-y-dā qise=m de-kird
DEM *man*-DEM-EZ KE *with*-3.SG-POST *speak*=1SG IPFV-*make*
'That man whom I was speaking with.'

b. *ew pyāw-e-y legeł ke qise=m de-kird
DEM *man*-DEM-EZ *with* KE *speak*=1SG IPFV-*make*
'That man with whom I was speaking.'

(37) a. ānnā, ke legeł-y qise=m de-kird
Anna KE *with*-3.SG *speak*=1SG IPFV-*make*
'Anna, who I was speaking with'

b. *ānnā, legeł ke qise=m de-kird
Anna *with* KE *speak*=1SG IPFV-*make*
'Anna, with whom I was speaking'

This observation, again, confirms the similarity between English relative *that* and Sōrānī Kurdish relative *ke*.

Before closing this subsection, I would like to mention a final point. Clause-initial *ke* can also mean *when* in Kurdish. Consequently, when a subordinate clause consists of *ke* combined with an VP rather than a sentence that starts with an overt NP, the subordinate clause can be ambiguous, as illustrated in (38). In the following, I will ignore the *when*-readings of *ke* sentences.

(38) ānnā, ke kirās-eke-y kirī, rōyšt bō ferānsā
Anna KE *shirt*-DEF-EZ *buy*.PST.3SG *leave*.PST.3SG *to France.*

Reading 1: 'Anna, who bought the shirt, went to France'

Reading 2: 'Anna, after she bought the shirt, went to France'

7.2.3 The relativised element: Resumptive clitics

The relativised element in a Sōrānī Kurdish RC can be the subject, a direct object, an object of preposition, an adjunct, a genitive or a predicative complement. The first three of these cases are illustrated in (39)–(41).

(39) ew wiše-y-e-y ke [de-m-giryén=ét]
DEM *word*-3.SG-DEM-EZ KE IPFV-1SG-*cry*.PRES=3.SG
'the word that makes me cry' (subject)

(40) ew wāne-y-e-y ke [de-y-xwén=im]
DEM *subject*-3.SG-DEM-EZ KE IPFV-3SG-*study*.PRES=1.SG
'that subject that I am studying' (direct object)

(41) ew kič-e-y ke [legeł-ī-dā wāne=t xwénd]
DEM *girl*-DEM-EZ KE *with*-3.SG-POST *lesson*=2.SG *study*.PST.3.SG
'that girl with whom you studied' (complement of a preposition)

In all of these cases, the core part of the RC, which I have put in brackets above, could also function as an independent clause. The corresponding clauses are given in (42) with their English translation.

(42) a. de-m-giryén-ét. 'It makes me cry.'

b. de-y-xwén-im. 'I am studying it.'

c. legeł-ī-dā wāne=t xwénd. 'You studied with her.'

This shows that the RCs in Kurdish do not contain a gap, contrary to English. Instead of the gap, there is a pronomoninal marking of the relativised element.

In his discussion of the internal structure of Kurdish RCs, Fattah (1997: 254) also elaborates on this point. He states that the relation between the antecedent

and the RC is between a pronominal clitic and the antecedent rather than, as in English, a gap in the RC and the antecedent.

I will follow Fattah (1997: 254) in referring to the pronominal marking of the relativised element in Kurdish RCs as *resumptive clitics*, hereby ResCs.[5] According to Fattah, the term *resumptive* points to the fact that these pronominal elements resume the function and the interpretation of the antecedent and transfer them to the RC. For instance, in (43) the pronominal clitic *yān*, which is marked in boldface, has the same number and person properties as the antecedent, *kitéb-āne* 'books', i.e, third person plural. Leaving this clitic out would make the sentence ungrammatical.

(43) ew kitéb-ān-e-y ke de***(=yān)**-xwén-īt-ewe hī min-in
DEM *book*-PL-DEM-EZ KE IPFV=3.P-*read*-2.SG-PRES *belong I*-3.PL
'Those books that you are reading belong to me.'

If the resumptive clitic is to be realised as a verbal ending in the third person singular, it is done so by an empty allomorph (as indicated in Table 7.4) This is shown in (44a). Here, I mark the empty allomorphe with ∅for illustration, whereas I leave it unmarked in the rest of the examples.

As can be seen in (44b), the marking is overt in the third person plural.[6] Given this pattern, I assume that there is always a resumptive marking, though this marking is phonologically empty in one case, namely for the third person singular verbal ending.

(44) a. ew wāne-ye ke de=t-xiwend-Ø …
DEM *lesson*-DEM KE IPFV=2.SG-*read*.PST-3.SG
'that lesson which you studied'

[5] I am aware that the term *clitic* is not used properly here, as it comprises both the pronominal clitic and the personal verbal endings from Table 7.4, page 208.

[6] I am grateful to Berthold Crysmann and Saloumeh Gholami for bringing my attention to this point.

b. ew wān-āne ke de=t-xwénd=in …
DEM *lesson*-PL KE IMPRFV=2.SG-*read*.PST=3.PL
'those lessons which you studied'

It is obligatory to mark the relativised element with a ResC. Using another overt element in that place, i.e. not having a relativised element in the embedded sentence, is not possible, as shown in (45).

(45) *ew kič-e-y ke min sārā=m xoš-de-w-ét le
DEM *girl*-DEF-EZ KE *I* *Sara*=1.SG *love*-IPFV-*want*.PRES-3.SG *in*
slémānī de-ž-ī
Sulaimaniyah IPFV-*live*.PRES-3.SG
* 'The girl$_i$ that I love Sara$_i$ lives in Sulaimaniyah.'

It is marginally possible to realise the relativised element using an independent pronoun, see (46). In this case, the independent pronoun bears a contrastive stress.

(46) ew kič-e-y ke min ew=m xoš-de-w-ét le
DEM *girl*-DEF-EZ KE *I* *she*=1.SG *love*-IPFV-*want*.PRES-3.SG *in*
slémānī de-ž-ī
sulaimaniyah IPFV-*live*.PRES-3.SG
lit.: 'The girl$_i$ that I love HER$_i$ lives in Suaimaniyah.'

'The girl that I love lives in Sulaimaniyah – though she may not love me.'

As I will not look into details of information structure, I will not take such cases into consideration in the rest of this chapter.

Sōrānī Kurdish RCs pattern with English *that* RCs in that there is no overt filler-like, fronted constituent. However, due to the use of ResCs, they differ from English as there is no gap. In the remainder of this section, I will show that in Kurdish, just like in English *that* RCs, there is an unbounded dependency relating the relativised element to the top node in the RC. I also show that the relativised

element can be deeply embedded inside the RC, but it must not be within an island.[7]

Evidence for the unbounded nature of Sōrānī Kurdish RCs is that clitics in Kurdish, similar to the gaps in English, can be deeply embedded inside the sentence. This is shown in (47), where the resumptive clitic, *=yān*, which is associated with the antecedent, occurs in the complement clause inside the RC.

(47) ew kitéb-ān-e-y [RC:ke tō pé-t-uā=bū [S:ke alī
DEM *book*-PL-DEM-EZ KE *you* PRE-3.SG-*so*=COP.PST.2.SG *ke* *Ali*
de=yān-xwén-ét-ewe estā]], (kitéb-ī) zōr bāš-in
IPFV=3.PL*read*.PRES-3.SG-CONT *now* *book*-EZ *much good*-3.PL
'The books that you think that Ali is reading them right now, are very good.'

The ResCs in Kurdish, exactly like the gaps in English, are sensitive to island constrains. I am aware that island constraints are not strong grammatical constraints, but I am using them as diagnostic. Example (48) is ungrammatical. The relativised element is inside the subject of the RC. This violates an Island Cosntraint as the subject is an island for UDCs.

(48) *ew žin-e-y [RC: ke [NP: guftugō [PP: legel=ī-dā]]
DEM *woman*-DEM-EZ KE *conversation* *with*=3.SG-POST
hemīše zōr xōš=e] bō xuān-ī ewāre bānghéšt=ī
always very pleasant=COP.PRES.1.SG *for meal*-EZ *dinner invite*=3.SG
kird=im
do.PST=1.SG
*'That woman that the conversation with (her) is always very pleasant, invited me for dinner.'

Another island constraint that is applicable to Kurdish ResCs is the Complex NP Constraint. In (49), the hypothetical relativised element occurs inside the

[7]Island Constraints were discussed in Section 3.3.2.

complement clause of a noun within the RC. As indicated, this constellation is not possible.

(49) *ew žin-e-y ke šānō ut=ī guāye ke pédečét alī
DEM *woman*-DEM-EZ KE *Shano say*.PST=3.SG *suspicion* KE *might* *Ali*
xos=ī-bw-ēt māng-y dāhātu deč-ét bō kabul.
love=3.SG-*want*.PRES-3SG *month*-EZ *next* IPFV.*go*-3.SG *to* *Kabul*
*'The woman [RC: that Shano expressed the suspicion [that Ali might be in love with (her)]] will move to Kabul next month.'

Additionally, in Kurdish RCs the coordinate constructions act like islands, that is, we can observe the *Coordinate Structure Constraint* and its *Across-the-Board exception*. A coordination is fine if there is a ResC in all conjuncts, see (50). It is, however, ungrammatical to have a ResC in just one of the conjuncts, as shown in (51), where there is an other direct object, *Ali*, in the second conjunct.

(50) ew kič-e-y [RC: ke [qise=t legeł kird u
DEM *girl*-DEM-EZ KE *talk*=2.SG *with* *do*.PST.2.SG *and*
yārmetī=t-dā]] xušk=im=e
help=2.SG-*give*.PST3.SG *sister*=1.SG=COP.PRES.3.SG.1.SG
'That girl [who [you talked with __ and helped __]] is my sister'

(51) *ew kič-e-y ke qise=t legeł kird u yārmetī alī=t
DEM *girl*-DEM-EZ KE *talk*=2.SG *with* *do*.PST.2.SG *and* *help*.PST *Ali*=2.SG
dā xušmk=m=e.
give.PST *sister*=1.SG=COP.PRES.3.SG.1.SG
*'That girl [who [you talked with __ and you helped Ali]] is my sister.'

This shows that the relation between the relativised element in Kurdish and the overall RC is subject to island constraints. I take this as evidence that it is a UDC, just as found in English RC.

In this subsection I showed that the relativised element is overtly realised in Sōrānī Kurdish in the form of a resumptive clitic. Nonetheless, there is an unbounded dependency relation in Sōrānī Kurdish SRCs.

7.2.4 Ezafé in Sōrānī Kurdish relative clauses

A general introduction, definition and examples of EZ was presented in 7.1.3. Some more notes about EZ that are specific for constructing Kurdish RCs are given in here. I follow Fattah (1997), Karimi (2007: 2160), and Samvelian (2007a: 8) in assuming that the EZ marker is used in Sōrānī Kurdish with RCs.[8] The EZ marker can co-occur with the relative particle to mark RCs in Kurdish, as shown in (52) and in many of the examples of Kurdish RCs above.

(52) ew kič-e-y ke qise=t legeł kird
DEM *girl*-DEM-EZ KE *talk*=2.SG *with* *do*.PST.2.SG
xušk=m=e
sister=1.SG=COP.PRES.3.SG
'That girl that you talked with is my sister'

Fattah (1997) argues that EZ is obligatory with IRCs that occur adjacent to their antecedent, where it is immaterial whether the IRCs is marked with *ke* or not. This is shown in (53) with the EZ-less version of example (52).

(53) *ew kič-e (ke) qise=t legeł kird
DEM *girl*-DEM KE *talk*=2.SG *with* *do*.PST.2.SG
xušk=m=e
sister=1.SG=COP.PRES.3.SG
Putative: 'That girl (that) you talked with is my sister.'

There are three types of cases in which the EZ marker does not occur. I will address them in the following.

First, EZ is only possible if the RC is adjacent to its antecedent. It does not occur with extraposed RCs. I repeat example (32) in (54), now focussing on the EZ marking rather than on the occurrence of *ke*.

[8]Note that in Persian, EZ does not occur with full RCs, see Samvelian (2007a,c).

(54) šuše-ke(*-y) šika ke hénā=t
glass-DEF-EZ *break*.PST.3.SG KE *bring*.PST=2.SG
'The bottle broke that you brought.'

Second, Fattah (1997: 255) says that EZ is excluded when the antecedent is a proper noun, an independent pronoun, or a generically used common noun. This is illustrated for proper nouns in (55).

(55) ānnā(*-y), ke kič-ī min=e, lēre=ye
Anna-EZ, KE *daughter*-EZ *I*=COP.PRES.3.SG, *here*=PRES.3.SG
'Anna, who is my daughter, is here.'

Fattah (1997: 257) explicitly lists the above mentioned three cases as the typical situations in which an RC is interpreted as an SRC. This leads him to conclude that SRCs do not accept an EZ marking on the word preceding it.

The third group of cases in which EZ is absent is harder to characterize. Fattah (1997: 255) names the following four cases:

1. The antecedent ends in the directional enclitic postposition *ewe/dā*.
2. The antecedent ends in the indefinite marker *ēk*.
3. The antecedent ends in a possessive suffix.
4. The antecedent ends in the additive marker -*š* , which means *too*.

These cases do not form a syntactic or semantic class. For this reason, I will discuss them separately. I will start with the clearest case, the lack of EZ marking after the enclitic postpositions *ewe/dā*.

The postposition *dā* occurs, among others, together with the preposition *lew* 'in'. As shown in (56), the combination *lew …-dā* acts as a circumposition around its complement NP.

(56) lew [šwén-e jwāne]-dā
in place-EZ *beautiful*-POST

'in that beautiful place'

Example (57) illustrates Fattah's (1997) point: if an RC follows the postposition, there cannot be an EZ marker.

(57) lew [šwén-e jwāne]-dā(*-y) ke tō=m bini.
in place-EZ *beautiful*-POST-EZ KE *you*=1.SG *see*.PST

'in that place where we first met'

I argue here that the RC in (57) is extraposed, which would immediately account for the absence of the EZ marker. To make this point, we can observe first that adjectives can host the enclitic postposition *dā* as in (56) above, but an RC cannot, see (58). Instead, the RC follows the postposition as shown above in (57).

(58) *lew šwén-e-y [ke tō=m di]-dā
in place-DEM-EZ KE *you=1.sg see*.PST-POST

'in the place where I saw you'

A second important observation is that the relative marker *ke* is obligatory when the RC follows the postposition. Example (59a) shows that *ke* cannot be omitted in sentence (57). Example (59b) shows the same for the postposition *ewe*, which does not allow for EZ marking either. The sentence is grammatical if the relativiser *ke* is present, but ungrammatical otherwise.

(59) a. lew [šwén-e jwāne]-dā *(ke) tō=m bini.
in place-EZ *beautiful*-POST KE *you*=1.SG *see*.PST

'in that place where we first met'

b. guł-eke de-xe-me īnĵān-ek-ewe *(ke) tō čāk=ī
flower-DEF IPFV-*put*.PRES-1.SG *vase*-DEF-POST KE *you repair*=3.SG
de-ke-īt
IPFV-*do*.PRES-2.SG
'I put the flower in the vases that you repair.'

The data on the enclitic postposition shows that they cannot attach to an RC. For this reason, the RC needs to be extraposed from the NP complement of the postposition. Once extraposed, it behaves just like other extraposed RCs in Kurdish, i.e., there is no EZ marking on the antecedent and relative *ke* is obligatory.

I can now turn to the next case, the ban of Ezafé on antecedents with the indefinite marker *ēk*. In her discussion of Ezafé in Persian, Samvelian (2007c) observes a ban on Ezafé for words containing an indefinite marker. Samvelian argues that there are morphological reasons for the lack of an EZ marking, as the EZ marker and the mentioned clitic occupy the same affixal slot. Samvelian, thus, concludes that this case should be treated as instances of *haplology*, i.e., as a case in which only one marking is overtly expressed but it covers the function of two markings. While I think that this type of explanation is also valid for the Sōrānī Kurdish case, it cannot be directly adopted, as the Kurdish data differ in an interesting way from Persian.

Example (60) illustrates the incompatibility of the indefinite marker *ēk* with EZ in Kurdish when there is an RC.

(60) ānnā kitéb-ék(*-ī) (ke) xōy hāz=ī lé=y=bū
Anna book-INDEF-EZ KE *herself like*=3.SG PRE=3.SG=COP.PST.3.SG
kirī
buy.PST
'Anna bought a book (that) she liked.'

As observed in Samvelian (2007a: 11), the indefinite marker is compatible with

EZ in Kurdish. This is illustrated with example (61) from Karimi (2007: 2175). This indicates that the absence of EZ before RCs cannot be due to the same morphological restriction on affix co-occurrence in Kurdish as it is in Persian.

(61) kitéb-ek-ī sur-ī jwān
book-INDEF-EZ *red*-EZ *nice*
'a nice red book' (Karimi, 2007: 2175)

Furthermore, the ban on EZ marking is restricted to the element bearing the indefinite marker. If there is another modifier between the indefinite noun and the RC, the EZ marker can occur, see (62). The example shows that both the noun hosting the indefinite marker and the modifying adjective following this noun can have an EZ marking.[9]

(62) ānnā kitéb-ék-ī jwān-ī ke xōy hāz=ī
Anna book-INDEF-EZ *nice*-EZ KE *herself like*=3.SG
lé=y=bū kirī
PRE=3.SG=COP.PST.3.SG *buy*.PST
'Anna bought a nice book that she liked.'

(63) ew kitéb-e jwān-e-ī ke xōm hāz=m
DEM *book*-DEM *nice*-DEM-EZ KE *myself* PREP=COP-3.SG
lé=ye=tī de-kir-m
PREP=COP=3.SG IPFV-*buy*.PST-1.SG
'I am buying a nice book that I like.'

This is evidence that there is no principled incompatibility of indefinite nominal projections and EZ marking.

[9]Saloumeh Gholami (p.c.) pointed me to a variety of Sōrānī Kurdish in which the object, *ketéb* 'book', in (62) and (63) is followed by a personal clitic. In these cases, (overt) EZ marking is excluded, see (i). However, for the Sulaimaniyah variety of Sōrānī Kurdish discussed here, the personal affix is not possible.

(i) ew kitéb-e jwān-e=m ke xom hāz=m- lé=yetī de-kir=m
DEM book-DEM nice-DEF=1.SG KE myself like=1.SG PRE=COP-3.SG IPFV-buy.PST=1.SG

The Sōrānī Kurdish data show that the indefinite marking restriction cannot be encoded on the RC, as the RC is clearly compatible with indefinite nominal projections. Consequently, it must be a constraint on the nominal head. The nominal head needs to have information on whether it is followed by some dependent and whether this dependent is clausal or not. When the nominal head is marked as indefinite and followed by a clausal dependent, we get the haplology effect assumed in Samvelian (2007c) for Persian.

This rule only applies to the lexical head of nominal projection. This is in line with Samvelian's approach, as she assumes two types introducing the EZ marking: one for EZ marking on the lexical head of a nominal project, and, one for EZ marking on a dependent. This difference can be related to what is traditionally called *primary* and *secondary* Ezafé, and discussed in detail in Samvelian (2007a).[10] Larson & Yamakido (2008: 61) refer to this distinction for Kurmanji Kurdish, where the Ezafé form with definite nouns followed directly by a modifier is different from the Ezafé form elsewhere in an NP, see (64). In (64a) only the definite lexical noun bears primary Ezafé, glossed as 1EZ. The following modifier has secondary Ezafé (2EZ), and so do all EZ-marked elements in the definite NP in (64b).[11]

(64) Primary and secondary Ezafé in Kurmanji, the examples and the glossing are from Larson & Yamakido (2008: 61).

a. kitéb-én bāš-ī nū
book-1EZ.PL *good*-2EZ.PL *new*
'the good new books'

[10]Though, interestingly, Samvelian (2007c) does not refer to this traditional distinction as motivation for her two processes.

[11]While Larson & Yamakido (2008) relate this difference purely to definiteness, Samvelian (2007a) provides evidence showing that this cannot be maintained easily. If my reasoning is correct, the Sōrānī data provide additional support for not treating primary Ezafé as a definiteness Ezafé as we find a special behaviour only with indefinitely marked lexical heads.

b. xānī-n-e bās-ī nū
house-INDEF.PL-2EZ.PL *good*-2EZ.PL *new*
'some good, new houses'

Bögel et al. (2008) criticize Samvelian's splitting of EZ introduction into two rules. At the same time, the language they discuss, Urdu, seems to have only the first type of Ezafé as it does not allow for Ezafé marking on dependents of the head noun (Bögel et al., 2008: 136). Looked at their data this way, their objection can, in fact, be seen as a confirmation of the split between primary, head-marking, and secondary, dependent-marking, Ezafé.

Returning to Sōrānī, we find similar subtle evidence for the distinction between primary and secondary Ezafé: only nominal heads are marked with the indefinite marker *ék*, and only nouns marked in this way show a haplology effect when followed by an IRC. In other words, I assume that the nouns *kitéb-ék* and *kitéb-ék-ī*, in (60) and (62) respectively, have the same Ezafé specification and that the absence and presence of an overt EZ marking is a consequence of an allomorphy of primary Ezafé in Sōrānī.

The next case to be discussed is the non-occurrence of EZ marking after possessive suffixes (Fattah, 1997: 255). In this respect, Sōrānī Kurdish patterns fully with Samvelian's (2007c) Persian data. This restriction is illustrated in (65).

(65) ew kitéb-e=m(*-ī) ke le ser méz-eke dā=m-nā-bu zōr
DEM *book*-DEM=1.SG KE *on head desk*-DEF PREF-1.SG-*put*-PST *much*
serinj-rākéš=e.
attention-drag=COP.PRES.3.SG
'My book that I had left on the desk is very interesting.'

When we look at Sōrānī NPs, we see that the possessive suffix is realised on the last element of an NP, as shown in (66a). It cannot occur on a non-final word in the NP, see (66b).

(66) a. kitéb-e tāze-ke=m
book-EZ *new*-DEF=1.SG
'my new book'

b. *kitéb-e=m tāze-ke
book-EZ=1.SG *new*-DEF
'my new book'

Example (67) shows that relative *ke* cannot be omitted after a possessive. This is evidence that we have a case of extraposition, even though this is a local extraposition, similar to the case of having a relative clause following a clitic postposition, see (59) above.

(67) ānnā u haeve ew kitéb-e=m=yān ke hemīše le nāw
anna and haeve DEM *book*-DEM=1.SG=3.PL KE *always at inside*
jāntā-ke=m-dā hel=ī-de-gir-m bird.
bag-DEF=1.SG-POST *up*=3.SG-IPFV-*catch*-1.SG *take*.PST.3.SG
'Anna and Haeve took a book of mine (that) I always keep in my bag.'

We can account for the blocking of EZ marking on words that bear a possessive clitic following the haplology-based encoding of Samvelian (2007c). Both EZ and the possessive compete for the same affixal slot. I assume that the possessive clitic always needs to be expressed, whereas EZ can remain unexpressed.

Finally, I turn to Fattah's last category of non-EZ marking: antecedents ending in the additive suffix *-š*. I agree with Fattah that the EZ marking may be omitted in this case. However, I accept the occurrence of EZ as well. My jugdments are shared by other speakers that I have consulted. I provide Fattah's example in (68), but I mark the EZ version of the antecedent with "%" to indicate that there are speakers who accept this marking.

(68) ew kič-e-š/ %kič-e-š-ī ke bīn-ī=t kič-ī
DEM *girl*-DEM-ADD/ *girl*-POST-ADD-EZ KE *see*.-3.SG=2.SG *girl*-EZ
min=e
I=COP.PRES.3.SG

'That girl, too, whom you saw is my daughter.'

The additive marker occurs relatively late in an NP. In (69a), for example, it is added to the adjective, not the lexical head.

(69) a. ew kič-e zīrek-e-š kič-ī min=e
that girl-EZ *clever*-DEF-ADD *girl*-EZ *I*=COP.PRES.3.SG

b. *ew kič-e-š-y zīrek=e kič-ī min=e
DEM *girl*-DEM-ADD-EZ *clever*=3.SG *girl*-EZ *I*=COP.PRES.3.SG
'That clever girl, too, is my daughter.'

The data on the additive marker reported so far can be captured if we assume that the additive marker and the EZ marker occupy the same affix slot. Speakers who share Fattah's (1997) judgement show haplology in that only one of the markers can be realised overtly. For speakers who share my judgments either the additive marker or both markers can be realised, and thus the haplology constraint is optional.

Evidence that the two markers are in the same affix slot comes from the possibility of omitting relative *ke*. Relative *ke* can be omitted in both the EZ-less and the EZ-full version of examples (68), as shown in (70).

(70) ew kič-e-š/ %kič-e-š-ī bīn-ī=t kič-ī
DEM *girl*-DEM-ADD/ *girl*-DEM-ADD-EZ *see*.-3.SG=2.SG *girl*-EZ
min=e
I=COP.PRES.3.SG
'That girl, too, whom you saw is my daughter.'

Matters become more complex once the possessive marker enters the picture. When the host of the additive marker also contains a possessive marker, the additive precedes the possessive, as shown in (71),

(71) žin-eke-š=m hāz=ī le gešt=e
wife-DEF-ADD=1.SG *love*=3.SG *to travel*=COP.PRES.3.SG

'My wife too loves to travel.'

The data with a possessive affix are relevant for my purposes, as the possessive is among the elements not compatible with EZ marking, see (65). There I argued that the possessive marker is always at the very end of the NP, which accounts for its incompatibility with EZ and the ban on *ke*-omission. I correctly predict that EZ is not possible with the possessive marker, see (72).

(72) [I have two daughters, but you only know one of them well. My daughter says hello …]

we kič-eke-š=m(*-ī) ke tō bāš
and daughter-DEF-ADD=1.SG-EZ *KE you well*
de=y-nās-īt slāw=t lé de-kā=t
IPFV=3.SG-KNOW.PRES-2.SG *greet*=2.SG *to* IPFV-DO=2.SG

'and my daughter that you know well, too, is saying hello to you.' (i.e., the one of my daughters who you know well)

In the discussion of example (67), we took the non-omissibility of *ke* after possessive pronouns as an indication of a (local) extraposition. From this, we expect that *ke* cannot be omitted if there is an additive marker in addition to the possessive, see (73). However, *ke*-omission seems to be possible in some cases, as in (74). Given the unclear grammaticality status of example (74), I will ignore it for my anaysis and assume that RCs are obligatorily extraposed when there is a possessive affix and that *ke* cannot be omitted. I, therefore, leave it for future research to elucidate the apparently special behaviour of the combination of the additive marker and the possessive affix with respect to *ke*-omission.

(73) [I have two daughters, but you only know one of them well. My husband says hello …]

we kič-eke-š=im *(ke) tō bāš de=y-nās-īt
and daughter-DEF-ADD=1.SG KE *you well* IPFV=3.SG-*know*.PRES-2.SG
slāw=t lé de-kā-t
greet=2.SG *to* IPFV-*do*-2.SG

'and my daughter that you know well, too, is saying hello to you.' (i.e., the one of my daughters who you know well)

(74) [I have two daughter, but you only know one of them well. My daughter says hello …]

ew kič-e-š=im ??(ke) dwéne bīn-ī=t
DEM *daughter*-DEM-ADD=1.SG KE *yesterday see*-3.SG=2.SG
slāw=t lé de-kā-t
greet=2.SG *to* IPFV-DO-2.SG

'and that one of my daughters who you saw yesterday, too, is saying hello to you.'

We have considered the four conditions under which Fattah (1997) observes no EZ marking in front of an RC. I have argued that the contexts do not form a homogeneous class. For enclitic postpositions and possessive affixes, the lack of EZ marking follows from the syntactic position of the RC, as it is extraposed. I consider the other two cases instances of haplology as in all of them the word has all the feature specifications of EZ-marking, but the marking is not overtly realised due to some haplology constraint. For words bearing the indefinite marker *ék*, this is due to a special restriction of primary Ezafé in Sōrānī. For the additive marker, there is some variation: if it appears inside the NP, it shows (optional) haplology. However, it can also co-occur with a possessive marker at the end of the NP; in which case, EZ is not possible.

In this section, I have provided a discussion of the distribution of Ezafé marking in Sōrānī Kurdish. This is interesting in the context of my thesis, as Sōrānī

Kurdish differs from Persian in requiring Ezafé marking on the antecedent of some RCs. However, Sōrānī Kurdish is lacking Ezafé marking for SRCs and for extraposed IRC.

I have pointed out similarities and differences to the Persian data. I argued that an analysis along the lines of Samvelian's (2007c) seems to be most promising for capturing the facts. In the following, however, I will not develop such an analysis, but simply assume that there is some morphosyntactic feature, EZ, whose specification indicates whether or not the antecedent of an RC has an Ezafé marking.

7.2.5 IRCs vs. SRCs in Sōrānī Kurdish

In the previous subsections, I occasionally pointed to differences between IRCs and SRCs in Kurdish. I will elaborate on this in more detail in this section. According to Fattah (1997: 259) there are obvious distinctions between SRCs and IRCs in Sōrānī Kurdish that appear to be similar to English. The most obvious difference between them, parallel to English, is that IRCs are intonationally integrated into the matrix clause, whereas SRCs are not. A further parallelism is that *ke* is optional in most of the cases of IRCs, while it is ungrammatical to omit it in SRCs (see section 7.2.2). This is similar to English, where bare RCs are restricted to IRCs.

A third distinction between IRCs and SRCs, which has no parallel in English, is that EZ markers cannot be used with SRCs while they are obligatory in most IRCs, see section 7.2.4.

Sōrānī Kurdish SRCs behave semantically just like English SRCs. I will illustrate this using examples that are analogous to the once given for English in

chapter 2.

As in English, proper nouns can be used as antecedent of SRCs, but not IRCs, which is shown by the impossibility of using the EZ marker and omitting relative *ke*, as in example (75).

(75) ānnā(*-y), *(ke) kič-ī min=e, lēr=ye
Anna KE *daughter*-3.SG *I*=COP.PRES.3.SG, *here*=COP.PRES.3.SG
'Anna, who is my daughter, is here.'

IRCs are possible with proper nouns only under their common noun use, see (76). As the English translation shows, such uses exist in English as well.

(76) be-rāstī ew alī-ā-y drāwsé=mān-m xōš-de-wé-t, belām
for-real DEM *Ali*-DEM-EZ *neighbour*=1.PL-.SG *nice*-IPFV-*wan*-3.SG *but*
ew alī-e-y xōšewīst-ī tō-ye zōr=m
DEM *Ali*-DEF-EZ *love*-3.SG *you*=COP.PRES.3.SG *much*=1.SG
xōš-nā-wé-t.
nice-NEG-*want*-3.SG
'I really like the Ali who lives next door, but I don't like the Ali you are dating all that much.'

When the antecedent is a quantified NP, an SRC cannot be interpreted in the restrictor of the quantifier. This is shown with the contrast between the IRC in (77) and the SRC in (78). Sentence (78) entails that all students failed.

(77) hemu ew qutābī-ān-e-y ke kewtu=n le
all DEM *student*-3.PL-DEF-EZ KE *fall*=3.PL *in*
tāqī-kirdinewe-ke-dā lyānīkem le dw wāne-dā
test-make.PRES-DEF-POST *at-least* *in two lesson*-POST
āmāde-ne-bun
attend-NEG-COP.PST.3.PL
'All students who failed the exam missed at least 2 class meetings.'

(78) hemu ew qutābī-ān-e, ke kewtu=n le tāqī-kirdinewe-ke-dā
all DEM *student*-3.PL-DEF KE *fall*=3.PL *in test-make*.PRES-DEF-POST

lyānīkem le dw wāne-da āmāde-ne-bu-n
at-least *in two lesson*-POST *attend*-NEG-COP.PST.3.PL
'All students, who failed the exam, missed at least 2 class meetings.'

Another semantic property of SRCs is that the SRC and its antecedent cannot be interpreted in the scope of a negation in the matrix clause. This was discussed for English in chapter 2 using example (68), page 28. In (79) I provide Sōrānī Kurdish translations of the relevant examples. The judgments are the same as for English: Example (79d) cannot be interpreted with the indefinite being in the scope of negation.

(79) a. xušk-eke=m hāwré-yek-ī heye ke
sister-DEF=1.SG *friend*-INDEF-3.SG *have*.PRES.SG KE
keštī-wān=e
ship-DER=COP.PRES.3.SG
'My sister has a friend who is a sailor.' (IRC)

b. xušk-eke=m hāwré-yek=ī n-ye ke keštī-wān
sister-DEF=1.SG *friend*-INDEF=3.SG NEG-*have*.PRES.SG KE *ship*-DER
bét
COP.PRES.3.SG
'My sister doesn't have a friend who is a sailor.' (IRC)

c. xušk-eke=m hāwré-yek=ī h-ye, ke
sister-DEF=1.SG *friend*-INDEF=3.SG *have*.PRES.SG KE
keštī-wān=e
ship-DER=COP.PRES.3.SG
'My sister has a friend, who is a sailor.' (SRC)

d. ? xušk-eke=m hāwré-yek=ī n-ye, ke
sister-DEF=1.SG *friend*-INDEF=3.SG NEG-*have*.PRES.SG KE
keštī-wān bét
ship-DER=COP.PRES.3.SG
? 'My sister doesn't have a friend, who is a sailor.' (SRC)

We saw in section 2.5 that there is some interaction of the content of an SRC

and its matrix clause with respect to discourse anaphoric relations in English. The following examples show that the same is true for Sōrānī Kurdish.

In (80), I repeat my earlier examples from section 2.5, page 55, adding the Sōrānī Kurdish translation.

(80) a. jōn fīdbāk-ī wird=ī le ser rāpōrt-ī kōtāy dā be her
John feedback-EZ *small*=3.SG *on head report*-EZ *final* *give*.PST *to each*
yek le qutābī-ek-ān, ke tenhā čend roj-ék bū péškeš=yān
one of student-DEF-PL KE *only* *some day*-INDEF COP.PST *give*=3.PL
kird-bū.
make-PST
'Jones gave detailed feedback on the final paper of each student, which THEY had submitted only a few days ago.'

b. jōn ke nimre=y le ser rāpōrt-ī kōtāy her yek le
John KE *degree*=3.SG *on head report*-EZ *final* *each one of*
qutābī-ek-ān dānā, fīdbāk-ī wird-ī-š=ī dā-né.
student-DEF-PL *put*.PST *feedback*-EZ *small*-3.SG-ADD=3.SG *give*-3.PL
'Jones, who graded each student's final paper, gave THEM detailed feedback.'

In (80a), there is a discourse pronoun inside the SRC that takes a universally quantified NP from the matrix clause as its antecedent. In (80b), we have the reverse situation: a quantified NP from the SRC can serve as antecedent for a discourse pronoun in the matrix clause.

The data in this section confirmed that Sōrānī Kurdish SRCs have the same semantic properties as English SRCs. While this might not be surprising, it has not been observed systematically for Sōrānī Kurdish so far. In addition, I take the parallelism of the data, both semantically and with respect to the distribution of Sōrānī Kurdish relative *ke* and English relative *that* as motivation for proposing an analysis for Sōrānī Kurdish along the lines of the one I developed for English in chapter 6.

7.3 Taghvaipour's analysis of Persian RCs

At several points in this chapter, I have mentioned similarities and differences between Sōrānī Kurdish and Persian. In this section, I briefly present the analysis of Persian IRCs in Taghvaipour (2004, 2005). While Persian RCs differ in various respects from Sōrānī Kurdish RCs, this analysis can still be taken as a good starting point for my analysis in section 7.4.

An example of a Persian IRC is given in (81)[12] As mentioned in section 7.2.4, Persian does not allow for EZ in front of RCs (Samvelian, 2007a,c). Consequently, there is no EZ marking in (81). The example is taken from (Taghvaipour, 2004: 276) with his glossing.

(81) zæn-i [*(ke) mæn dust+daræm] inja nist.
woman-RES COMP *I* *like*-PRES-1SG *here* NEG-*be*-3SG
'The woman I love is not here.'

The example also shows that Persian RCs are introduced with a relative *ke*. However, contrary to Sōrānī Kurdish, relative *ke* is obligatory in all RCs in Persian. Apart from this obligatoriness, Persian relative *ke* behaves analogously to Sōrānī Kurdish relative *ke* as it is invariant in form and does not allow for any form of upward percolation. Taghvaipour (2004, 2005) treats Persian relative *ke* as an RC complementizer, an analysis that I will adapt for Sōrānī Kurdish.

The marking of the relativised element in Persian RCs is also similar but different from what I described for Sōrānī Kurdish in section 7.2.3. According to Taghvaipour (2004), there is an alternation between a gap and a – non-clitic – resumptive pronoun in Persian.[13] The following example including the glossing is

[12]The translation and glossing is taken from Taghvaipour (2004) directly, just adjusting the font, but not the abbreviations, which might, therefore, be missing from my list of abbreviations on page vii.

[13]Which option is to be chosen depends on the grammatical function of the relativised element

from Taghvaipour (2004: 277).

(82) mærd-i [ke (u ra) diruz molaqat kærdid] aqay-e Bayat
man-RES COMP *he* RA *yesterday* *meet*-PST-2PL *Mr.* *Bayat*
bud.
be-PST-3SG
'The man whom you met (*him) yesterday was Mr. Bayat.'

Taghvaipour (2004, 2005) shows that the relativised element is subject to island constraints, independently of whether the relativised element is a gap or a resumptive pronoun. He concludes from this that there is a UDC in either case.

For Sōrānī Kurdish, I showed that the relativised element is obligatorily marked by verbal inflection or a pronominal clitic, depending on its grammatical function. Neither can it be left out nor can it be realised as an independent pronoun. However, similarly to Persian, the relativised element must not violate island constraints.

A central concern of Taghvaipour (2004, 2005) is to account for the variation in the realisation of the relativised element. As there is no such variation in Sōrānī Kurdish, the details of his analysis are not relevant here. What is important, though, is that he assumes that there is a complete clause (S) following relative *ke* and that this S has a non-empty INHER|SLASH value containing the LOCAL specification of the relativised element. This aspect of the analysis corresponds to what I have assumed for English *that* RCs, and I will make the same assumption for Sōrānī Kurdish RCs.

Taghvaipour (2004, 2005) assumes that Persian relative *ke* is an RC complementizer. I provide his lexical entry in Figure 7.1.[14] The relativiser selects for a clause, [A], that contains a gap, as marked with the tag [4] in the SLASH value of

in IRCs.

[14]The AVM in Figure 7.1 is given as in Taghvaipour (2005), but with no list brackets around the PHON value and with the feature BIND outside the SYNSEM value.

the element in the complementizer's ARG-ST value. Relative *ke* binds this UDC. Since there is no upward percolation in Persian RCs, there is no need to assume a REL feature. The complementizer has a non-trivial MOD value, in which the antecedent of the RC is specified. It is required to be a nominal projection that has the same INDEX value, [1], as the gap. Consequently, the gap is the relativised element. The semantics of the RC, [2], combines intersectively with that of the antecedent, [3].

The analysis of Persian relative *ke* is very similar to my proposal for the English null relativiser for bare RCs in Figure 6.2 on page 193. In particular, this includes the fact that there is no REL percolation involved.[15]

[15]Remember that I showed for English that the absence of a REL percolation is the major difference between the null relativiser and relative *that*.

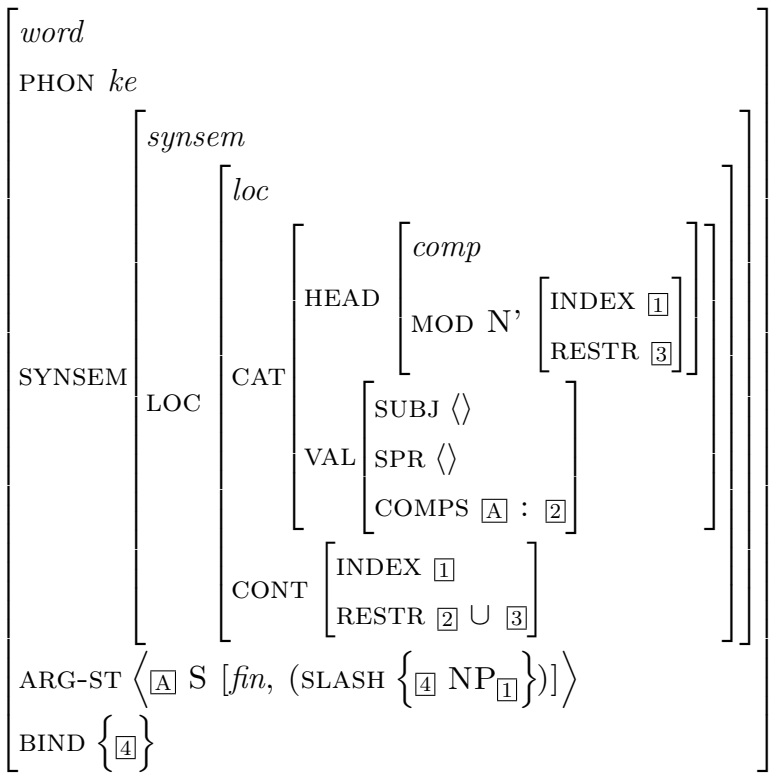


Figure 7.1: Taghvaipour's (2005:166) lexical entry of Persian relative *ke*

In the following section, I show how this analysis can be applied to Sōrānī Kurdish *ke*. The similarities between Sōrānī Kurdish and Persian as well as the similarities between relative *ke* in Sōrānī Kurdish and relative *that* in English justify such a parallel analysis.

7.4 The analysis of Sōrānī Kurdish relative clauses

In this section, I present my analysis of Sōrānī Kurdish RCs. In the previous sections, I have already pointed out the similarities between Sōrānī Kurdish and Persian RCs, and between Sōrānī Kurdish *ke* RCs and English *that* RCs.

Here, I focus on the lexical properties of the relativiser, and rely on the properties of RCs established in the preceding sections. First, I showed in Section

7.2.3 that the core clause of Sōrānī Kurdish RCs has the same structure as ordinary clauses. Second, there is no difference between IRCs and SRCs with respect to the internal structure of the core clause in an RC. Third, as shown in section 7.2.3, the relativised element is overtly realised in the core clause of an RC, either in the form of a verbal ending or as a resumptive clitic. However, the relativised element behaves like a gap in that it introduces an element into the INHERITED|SLASH value. Fourth, I assume a feature EZ(AFÉ), whose value is "−" for nominal projections that are not marked with Ezafé, and "+" for those who are – which includes nominal projections where the Ezafé marking is not visible due to haplology, see Section 7.2.4. With these assumptions in place, I can now turn to the specification of the Sōrānī Kurdish relativiser.

Given the strong similarity between the occurrence of *ke* in Sōrānī Kurdish RCs and the occurrence of *that* in English RCs, I propose largely the same modelling here as I did for English in section 6.5. I assume two relativisers. First, there is *ke*, which is compatible with all RCs. Second, there is a highly restricted phonological empty relativiser: it only occurs in non-extraposed IRCs whose antecedent is marked with Ezafé and whose relativised element is constrained (for some speakers) so that it cannot be the local subject of the RC.

Evidence for the existence of two relativisers comes not only from their differences in distribution but also from coordination data, parallel to what Sag (1997) observed for the categorial difference between bare RCs and *that* RCs in English. In (83) I present variants of the RC-coordination examples from Sag (1997: 463), given in (38) above.[16]

Example (83a) shows that *ke*-less RCs can be coordinated. Similarly, two

[16]Given that Sōrānī Kurdish is a split ergative language, I had to change the tense form in the RCs to ensure that *ke*-omission is possible in principle in them. I also use a plural antecedent so that the resumptive clitic be realised overtly.

ke-RCs can be coordinated, see (83b). However, a *ke*-less and a *ke*-marked RC cannot be coordinated, (83c). This shows that the coordination data on RCs for Sōrānī Kurdish are, in fact, parallel to those reported for English bare and *that* RCs.

(83) a. ew witār-ān-e-y [[ew de=yān-nws=ét] w [min
DEM *essay*-PL-DEM-EZ *she* IPFV=3.PL-*write*=3.SG *and I*
de=yān-xwén=m-ewe]] le-ser ew
IPFV=3.PL-*read*=1.SG-CONTN *on-head* DEM
beste-ye=n
pile-DEM=COP.PST.3.PL
'Those essays$_i$ [[she is writing] and [I will be reading]] are on that pile.'

b. ew witār-ān-e-y [[ke ew de=yān-nws=ét] w [ke min
DEM *essay*-PL-DEM-EZ *KE she* IPFV=3.PL-*write*=3.SG *and KE I*
de=yān-xwén=m-ewe]] le-ser ew beste-y=n
IPFV=3.PL-*read*=1.SG-CONTN *on-head* DEM *pile*-DEM=COP.PL
'Those essays$_i$ [[that she is writing] and [that I will be reading]] are on that pile.'

c. *ew witār-ān-e-y [[ew de=yān-nws=ét] w [ke min
DEM *essay*-PL-DEM-EZ *she* IPFV=3.PL-*write*=3.SG *and KE I*
de=yan-xwén=m-ewe]] le-ser ew
IPFV=3.PL-*read*=1.SG-CONTN *on-head* DEM
beste-ye=n
pile-DEM=COP.PRES.3.PL
'Those essays$_i$ [[she is writing] and [that I will be reading]] are on that pile.'

I modelled the English data in section 6.5 by assuming that *that* RCs have a non-empty INHER|REL value, whereas bare RCs have an empty INHER|REL value. Complex relative phrases in English constitute independent evidence for the need of a REL feature. As we saw in section 7.3, there is no such upward percolation in Persian. For this reason Taghvaipour's (2004) analysis does not need a REL feature. It would be a straightforward step to take the data in (83) as motivation

for a REL feature in Sōrānī Kurdish – with a non-empty value for *ke*-RCs and an empty value for *ke*-less RCs.

However, I am reluctant to propose such an analysis. The reason for this is that *ke*/bare alternation in Kurdish is not fully parallel to what we find in English. Sōrānī Kurdish *ke* is also used as a complementizer for embedded declarative clauses. Just as in English, such clauses can also be realised without a complementizer. This is shown in (84).

(84) pé-m-wā-ye (ke) aléx dil=ī xōš=e.
PRE-1.SG-*so*-POST KE *Alex heart*=3.SG *nice*=COP.PRES.3.SG
'I think (that) Alex is happy.'

In English, *that*-less complement clauses and *that*-complement clauses can be coordinated. This is shown in the COCA example in (85).

(85) Sixty-nine percent of the students surveyed indicated that they think [[∅ homework is meaningful] and [that it reinforces concepts learned in class]], while 64% of the students disagreed with the statement that homework served little or no purpose. (COCA)

This is not what we find in Sōrānī Kurdish. The data in (86) are parallel to the ones in (83), but with coordinated complement clauses rather than RCs. Examples (86a) and (86b) show that coordination of complement clauses with identical marking is possible, though there is a preference for not using a complementizer. However, the coordination of a bare complement clause and a *ke*-marked complement clause is ungrammatical, see (86c).

(86) a. pé=m-wā-ye aléx dil=ī xōš-e u
PRE=1.SG-*so*-POST *Alex heart*=3.SG *nice*=COP.PRES.3.SG *and*
alī-š dil=ī xōš=e.
Ali-ADD *heart*=3.SG *nice*=COP.PRES.3.SG

'I think Alex is happy and Kim is happy too.'

b. ?pé-m-wā-ye ke aléx dil=ī xōš=e u ke
PRE-1.SG-*so*-POST KE *Alex heart*=3.SG *nice*=COP.PRES.3.SG *and* KE
alī-š dil=ī xōš=e.
Ali-ADD *heart*=3.SG *nice*.COP.PRES
'I think that Alex is happy and that Kim is happy too.'

c. *pé-m-wā-ye aléx dil=ī xōš=e u ke
PRE-1.SG-*so*-POST *Alex heart*=3.SG *nice*=COP.PRES.3.SG *and* KE
alī-š dil=ī xōš=e.
Ali-ADD *heart*=3.SG *nice*=COP.PRES.3.SG
'I think Alex is happy and that Kim is happy too.'

There is no motivation for using a REL feature, or any other nonlocal feature for the complement clauses in (86). Nonetheless the coordination of a *ke*-less and a *ke*-marked subordinate clause is equally impossible for complement clauses and RCs. Consequently, the difference between the two types of subordinate clauses should not be captured in terms of REL.

For this reason, I will pursue a different approach, which is, however, speculative. I propose a feature CFORM (COMPLEMENTIZER-FORM). This feature is analogous to PFORM. It is a head feature and defined on complementizers and relativisers. In Sōrānī Kurdish, CFORM can have (at least) the values *ke* – for *ke*-marked RCs and complement clauses – and *none* for the null relativiser and the null complementizer. Conjuncts must agree with respect to their HEAD values.

With these assumptions, we can now look at the two relativisers I propose for Sōrānī Kurdish. I will start with the lexical entry of the more flexible relativiser, *ke*, in Figure 7.2.

This AVM is the adjusted form of the English relativiser from Figure 6.1. The relativiser has the phonology *ke* and its CFORM value is *ke*. It selects a nominal projection, N, via its SEL value. It requires a finite, unmarked clause

as its complement. This clause must contain the relativised element as it has a non-empty INHER|SLASH value, 2. The relativiser has this element in its TO-BIND|SLASH value. The INDEX value of the extracted element, 1, has the same PHI value as the antecedent of the RCs. Just as in Persian, we do not need a REL feature, as there is no upward percolation in Sōrānī Kurdish RCs.

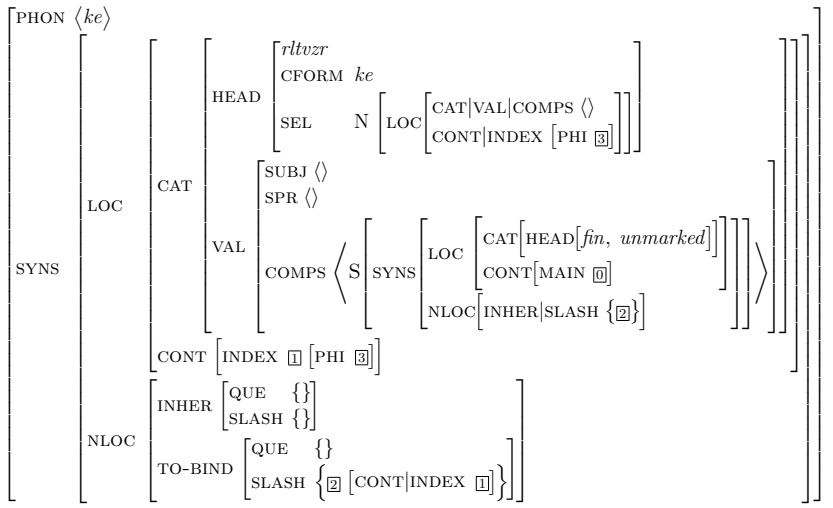


Figure 7.2: Sketch of the lexical entry of the Sōrānī Kurdish relativiser *ke*

As *ke* RCs can occur with Ezafé-marked antecedents but also with antecedents without this marking, there is no constraint on Ezafé in the description of the selected element.

I also leave out the semantics in the lexical entry as relative *ke* occurs with both IRCs and SRCs. As we saw in section 6.5, in either case, the PHI values of the antecedent and the relativised element are identical. However, the DR values are only identical for IRCs but not for SRCs. This is expressed in the constraint in Figure 7.3. The constraint leaves exactly two interpretive options for relativisers expressed in the form of disjuncts.

The first disjunct specifies the case of an IRCs. Here, the antecedent and the relativised element have fully identical INDEX values, written as 1. The external

$\left[\text{SYNS}|\text{LOC}|\text{CAT}|\text{HEAD}\ \ \textit{rltvzr}\right] \rightarrow$

(IRC-rltvzr)

$$\left[\begin{array}{ll}
\text{SYNS} & \left[\text{LOC}\ \left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{ll}\text{HEAD}|\text{SEL} & \text{N}\ \left[\text{LOC}|\text{CONT}|\text{INDEX}\ \boxed{1}\right]\\ \text{VAL}|\text{COMPS} & \left\langle\left[\text{LOC}|\text{CONT}|\text{MAIN}\ \boxed{2}\right]\right\rangle\end{array}\right]\\ \text{CONT} & \left[\text{INDEX}\ \boxed{1}\ \left[\text{DR}\ x\right]\right]\end{array}\right]\right]\\
\text{LRS} & \left[\begin{array}{ll}\text{EXCONT} & \alpha\wedge\beta\\ \text{INCONT} & \boxed{2}\\ \text{PARTS} & \left\langle\alpha\wedge\beta,\ x\right\rangle\end{array}\right]
\end{array}\right]$$

and $\boxed{2} \leq \beta$

or

(SRC-rltvzr)

$$\left[\begin{array}{ll}
\text{SYNS} & \left[\text{LOC}\ \left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{ll}\text{HEAD}|\text{SEL} & \text{NP}\ \left[\text{LOC}|\text{CONT}|\text{INDEX}\ \left[\text{DR}\ x\right]\right]\\ \text{VAL}|\text{COMPS} & \left\langle\left[\text{LOC}|\text{CONT}|\text{MAIN}\ \boxed{0}\right]\right\rangle\end{array}\right]\\ \text{CONT} & \left[\begin{array}{ll}\text{INDEX} & \left[\text{DR}\ w\right]\\ \text{MAIN} & w=x\end{array}\right]\end{array}\right]\right]\\
\text{LRS} & \left[\begin{array}{ll}\text{EXCONT} & \alpha\wedge\beta\\ \text{INCONT} & w=x\\ \text{PARTS} & \left\langle\alpha\wedge\beta,\ w,\ \iota w:w=x,\ w=x\right\rangle\end{array}\right]
\end{array}\right]$$

and $\iota w : \ldots \leq \beta$

and $\boxed{0} \leq \beta$

Figure 7.3: Constraint on the relativiser semantics

content is a conjunction, $\alpha \wedge \beta$, just as for simple intersective adjectives. The internal content of the relativiser is identical with that of the MAIN value of its complement, $\boxed{2}$. This content is constrained below the AVM to occur in the second conjunct of the external content.

The SRC-reading is specified in the second disjunct of the constraint in Figure 7.3. In this case, only the PHI values of the antecedent and the relativiser are identical. This does not need to be stated in this constraint, as it is already specified in the lexical entry of the relativiser. The antecedent and the relativiser have distinct DR values, x and w respectively. The MAIN value of the relativiser is an identity statement between these two discourse referents. The external content is, again, a conjunction $\alpha \wedge \beta$. Its internal content is the identity statement $w = x$. The PARTS list has, furthermore, an iota expression on it which is a subexpression of the second conjunct of the external content. In addition, the iota operator binds the variable w.

This analysis of relative *ke* is as parallel to English relative *that* as possible. There is a difference with respect to the REL specification, which is absent from Sōrānī Kurdish. Instead, I use the feature CFORM to be able to distinguish between relative *ke* and the null relativiser. Relative *ke* is compatible with both an IRC and an SRC reading – just like English relative *that*. As it is, however, the only relativiser for Sōrānī Kurdish SRCs, such SRCs are not perceived as (stylistically) marked.

Let me now turn to the null relativiser. I provide its lexical entry in Figure 7.4. This lexical entry differs in various ways from that of relative *ke* in Figure 7.2. First, obviously, it has an empty PHON list, and its CFORM value is *none*. It requires that its antecedent be marked for Ezafé. This has two consequences: first, there will be such a marking on the antecedent, and, second, this will exclude

extraposition of the RC.

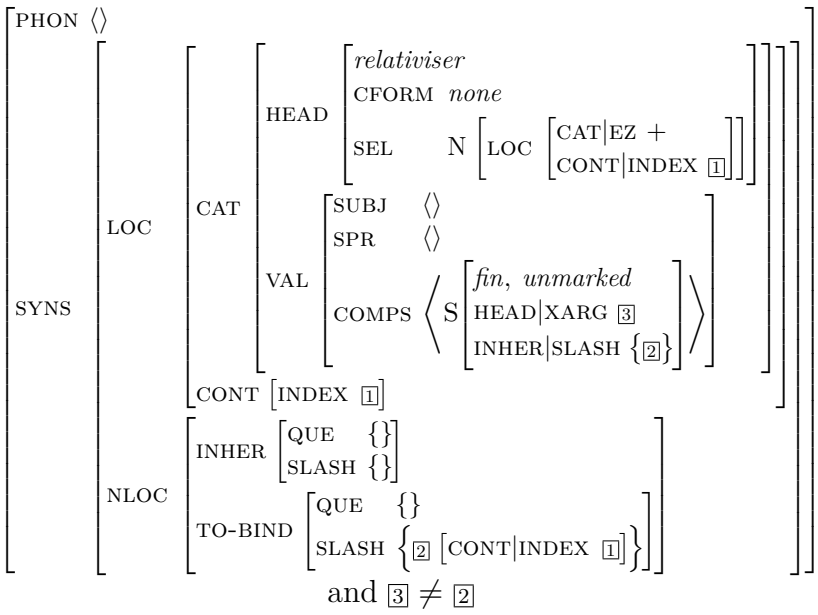


Figure 7.4: Sketch of the lexical entry of the Sōrānī Kurdish null relativiser

The entire INDEX value of the antecedent is structure shared with the relativiser, 1. As a consequence, this lexical entry is only compatible with the first disjunct in the constraint on relativiser semantics in Figure 7.3. In other words, the null relativiser can only be used with IRCs.

The constraints expressed inside the AVM seem to be valid for all Sōrānī Kurdish speakers. This might not be the case for the constraint that I added below the AVM. It states that the element extracted from the core clause, 2, must not be identical to the local subject of that clause, that is, to the clause's XARG value. This constraint seems to be subject to inter-speaker variations. As mentioned above, Saady (2020) reports cases in which bare SRCs are possible even when the relativised element is the local subject of the RC – see his example (29), repeated as (87). Speakers who accept the *ke*-less variant of (87) simply do not have the additional constraint below the AVM in Figure 7.4.

(87) ew šōfér-é-y (ke) be héwašy otōmbél le-de-xwr-ét
DEM *driver*-DEM-EZ KE *with slowly* *automobile* PRE-IPFV-*drive*-3.SG
selāmet=e.
safe=COP.PRES.3.SG
'The driver who drives cars slowly is safe'

(Saady (2020: 114), my gloss and translation)

On the other hand, if the data on long distance relativization that I reported in (31) are a consistent pattern for speakers, there would be another, more complex, restriction on the empty relativiser. We would, then, need to enforce that the relativised element come from the local clause and be not embedded inside another clause within the RC. However, as observed above, more data work is necessary to establish whether such a more complex constraint is needed.

7.5 Summary

In this section, I showed how the antecedents of Sōrānī Kurdish RCs can have any grammatical function, be it a subject, a complement, the complement of an adjective, a predicative complement, or a modifier of a main clause subject. Additionally I pointed out that unlike English RCs, Sōrānī Kurdish RCs cannot take a non-nominal antecedent.

In analysing the relative *ke*, I showed the similarity of *ke* to the distribution of relative *that* in English IRCs. I agreed with Fattah (1997) who states that *ke* should not be treated as a pronoun, and therefore he treats it as a complementizer. I concluded that relative *ke* can be present in all types of RCs. It can be omitted in IRCs, but not when it is the grammatical subject of the RC. I confirmed that, in extraposed RCs, the presence of *ke* is obligatory. Finally, I showed how similar these constraints are to the distribution of relative *that* in English IRCs.

Later, I discussed the fact that the construction of Sōrānī Kurdish RCs depends heavily on ResCs as the relativised element in Sōrānī Kurdish is realised in the form of a ResC. Further, I showed that Sōrānī Kurdish RCs have UDCs.

Under this analysis of the distribution of Ezafé marking in Sōrānī Kurdish, I concluded that Sōrānī Kurdish differs from Persian in requiring Ezafé marking on the antecedent of some RCs; however, Sōrānī Kurdish is lacking Ezafé marking for SRCs and for extraposed IRC. I assumed that there is some morphosyntactic feature, EZ, whose specification indicates whether or not the antecedent of an RC has an Ezafé marking.

In an attempt to provide Sōrānī Kurdish RCs with an HPSG analysis, and due to the fact that the distribution of *ke* in Sōrānī Kurdish RCs is very similar to *that* in English RCs, I argued for the same modelling for Sōrānī Kurdish as I did for English. Since there is no upward percolation in Sōrānī Kurdish RCs, I did not use the REL feature. I pursued a different approach by which I assume a feature called CFORM (COMPLEMENTIZER-FORM). It is a head feature and defined on complementizers and relativisers. I argued for two relativisers: the relative *ke*, and the null-relativiser which only occurs in non-extraposed IRCs whose antecedent is marked with Ezafé and whose relativised element is constrained (for some speakers) and cannot the local subject of the RC.

Chapter 8

Conclusion

In this thesis, I have explored syntactic and semantics aspects of supplementary relative clauses (SRC). I concentrated on types of SRCs that have not received much attention in the literature: determiner-*which* SRCs and *that* SRCs. In addition, I included Sorānī Kurdish relative clauses, for which we also lack a formal description. The underlying theoretical question was how the differences and similarities between SRCs and integrated relative clauses (IRCs) can be captured in a constraint-based, formal model of grammar. Given that the structures discussed in this thesis have not been thoroughly studied so far, I also tried to establish an empirical basis for my analysis.

In this concluding chapter, I will look at different aspects and results of my thesis. I will summarize the main differences between SRCs and IRCs (section 8.1), and the main empirical properties of determiner-*which* SRCs and *that* SRCs in section 8.2. I will, then point to the central aspects of my analysis of English relative clause constructions (section 8.3). I will put my research on Sorānī Kurdish relative clauses into the overall context of my thesis (section 8.4). Finally, in section 8.5, I will reformulate the individual ingredients of my theory to arrive at

a compact representation of English relative clause constructions that avoids the duplication of lexical entries for elements that can occur in both integrated and supplementary relative clauses.

8.1 The nature of the IRC/SRC difference

My overview of the descriptive and theoretical works on RCs in general, and SRCs in particular revealed a close similarity between SRCs and IRCs. While SRCs differ from IRCs in their prosody, the antecedents they take, their semantics and syntax, these differences are not sufficient to impose different treatments and different analyses for the structures.

Arnold (2007) showed convincingly that, with respect to many syntactic phenomena, SRCs pattern with IRCs rather than with other constructions, such as parentheticals. From this observation, it follows that there are many conclusive indications showing that SRCs are in fact integrated to the matrix clause hosting them. Orphanage approaches, radical or not, reject such an integration and are; therefore, unable to provide an adequate analysis of SRCs.

Instead, the differences between SRCs and IRCs seem to be based on the way in which their semantics combines with that of the matrix clause. In particular, the relative pronoun in an SRC is interpreted as a discourse pronoun and SRCs are (usually) not part of the at-issue content.

Nonetheless, AnderBois et al. (2015) discuss phenomena showing that the interpretation of SRCs and their host clause need to go hand in hand. For example, there is an interaction between a matrix clause and an SRC contained in it with respect to discourse anaphoric relations. Furthermore, the SRC needs to be taken into consideration semantically in its surface position. Such an interaction would

not be possible if the SRCs were interpreted completely independently of the matrix clause.

However, the kind of semantic interaction is restricted to relations which we typically find in discourses and cannot lead to an interpretation of the SRC in the scope of an operator from the matrix clause. This confirms the insights formulated in Arnold (2007).

8.2 The status of determiner-*which* and *that* SRCs

The main empirical contribution of the discussion of English SRCs consisted in a systematic collection of corpus examples for determiner-*which* SRCs and *that* SRCs, which have been complemented by two exploratory questionnaire studies.

My questionnaire on determiner-*which* did not reveal strong differences to SRCs with pronoun-*which* with respect to various percolation mechanisms. I concluded from this that determiner-*which* SRCs can and should be treated in the same way as pronoun-*which* SRCs and that there is no empirical reason for imposing additional, special constraints on them.

My second questionnaire study was designed to evaluate native speakers' (dis-)preference of the occurrence of relative *that* in SRCs. Speakers did not consistently rate *that* SRCs unnatural nor did they assign corresponding pronoun-*which* SRCs better ratings. I concluded from this that there are no grammar-related differences between *that* SRCs and pronoun-*which* in SRCs. This means in particular that I think *that* SRCs need to be included in attempts to arrive at a full characterization of the English relative clause system.

My two questionnaire studies provide evidence that both determiner-*which* SRCs and *that* SRCs should be considered part of the grammar of present day

English. This system is rich in apparently freely interchangeable structures – such as the classical choice between bare IRCs and *that* IRCs, but also the newly discussed free variation between determiner-*which* SRCs and pronoun-*which* SRCs as well as that between *that* SRCs and pronoun-*which* SRCs. In this thesis, I only attempted to establish the existence of this free variation. This is clearly just a first step and it would be important to identify differences between those alternating structures. However, such differences seem to lie in areas other than the syntactic structure and the interpretation.

8.3 Concluding results of the HPSG and LRS analyses of the structures

My analysis is based on the empirical and theoretical insights on SRCs of Arnold (2004, 2007). Arnold formulates his theory in terms of *constructional HPSG* using a DRT semantics. I proposed an alternative formalization which is more in the spirit of Pollard & Sag (1994b) in that the properties of relative clauses are projected from function words, which include potentially empty relativisers.

Based on my assessment of determiner-*which* SRCs and *that* SRCs as bona fide members of the English relative clause system, this thesis is also the first explicit analysis of these two structures in HPSG.

The percolation mechanism found in *wh*-relative phrases played an important role in my thesis. First, I argued that the percolation mechanism proposed for *wh*-relative phrases need not be changed to accommodate determiner-*which* SRCs. Second, the absence of upward percolation with relative *that* in English is certainly among the strongest arguments in favor or treating it as a relativiser rather than a pronoun. Nonetheless, I still make use of the relevant feature, REL, in my

analysis of *that* relative clauses. In particular, I assume that relative *that* and the relativiser for bare relative clauses differ in important respects, such as in their REL specification. This allows me to capture the observation that *that* relative clauses and *wh*-relative clauses can be coordinated – which was used as a major argument against a relativiser analysis of *that*.

My analysis of determiner-*which* is, eventually, very similar to that of the SRC-use of pronoun-*which* – with the exception that determiner-*which* needs to combine with a nominal head, of course.

I would also like to comment on some semantic aspects that played a role in my thesis. One of the important generalizations in Arnold (2004, 2007) is that the connection between an SRC and its antecedent should be treated semantically as discourse anaphoricity. I adopt this generalisation and show how it can be integrated into my syntactic analysis. Even though I treat relative *that* as a relativiser, I assign it the semantic contribution of a discourse anaphoric pronoun.

In my discussion of the semantic properties of determiner-*which* SRCs, I showed that determiner-*which* NPs trigger a conventional implicature, even in cases in which the overall SRCs can be interpreted at-issue.

I formulate the semantic aspects of my approach in some version of predicate logic as my semantic representation language. This allows me to propose a combinatorial semantics based on Lexical Resource Semantics (LRS). To my knowledge, the present work is the first systematic application of LRS to supplementary relative clauses.

The integration of determiner-*which* SRCs and *that* SRCs into a formal analysis of English relative clauses seems to fill a gap in the present HPSG literature. I will try to present the sketch of a more complete, systematic picture of the English

relative clause system in the final section of this conclusion, Section 8.5.

8.4 Concluding results of analysing Sorānī Kurdish RCs

I included a chapter on relative clauses in Sorānī Kurdish because they provide additional support for my analysis of English *that* relative clauses. At the same time, I discussed some basic feature of Sorānī Kurdish relative clauses, though I concentrated on their left periphery – which is either bare or shows the particle *ke*.

Sorānī Kurdish relative clauses differ from English relative clauses in general in that, in most cases, there is no gap in Sorānī Kurdish but a resumptive element. Nonetheless, I showed that Sorānī Kurdish relative clauses are sensitive to the same island constraints as English relative clauses. Consequently, they should be modelled as unbounded dependencies as well.[1]

I argue that the bare/*ke* distinction shows essential similarities with the bare/*that* contrast in English: in both languages, bare relative clauses can only be interpreted as IRCs, they need to be adjacent to their antecedent and cannot be extraposed, and are disprefered when the relativised element is in the position of the local subject of the relative clause.

Similarly, Sorānī Kurdish *ke* and English *that* relative clauses share important properties: both occur with an IRC and SRC reading, they can be extraposed, and they impose no further restrictions on the relativised element within the relative clause.

A relativiser analysis for relative *ke* is much more natural than an analysis

[1]Note that I do not commit to a particular status of island constraints in grammar, but use them as diagnostics for unbounded dependency constructions.

as a pronoun. It is not homonymous with any pronoun and lacks the number and person variation found with pronouns in Sorānī Kurdish. It is, however, homonymous with the complementiser used to introduce declarative complement clauses. The strong similarities between English relative *that* and Sorānī Kurdish relative *ke* support a parallel analysis for the two languages, i.e., an analysis as relativisers.

The Sorānī Kurdish relative clause system is considerably simpler than the English system, as there is no equivalent to *wh*-relative clauses in Sorānī Kurdish. For this reason, there is no upward percolation in Sorānī Kurdish relative clauses, and, consequently, no need for a REL feature in my HPSG analysis of Sorānī Kurdish.

In my HPSG analysis of Sorānī Kurdish relative clauses I focus exclusively on the modelling of the bare/*ke* alternation. My analysis should, however, be compatible with any suitable treatment of the phenomena that occur inside relative clauses, such as the above mentioned resumptive marking of the relativised element, the split-ergativity found in the realization of grammatical functions, and the Ezafé marking on modified nominal projections.

As I include IRCs into my theorising in this chapter much more strongly than in the earlier chapters, the differences in the modelling of the two types of relative clauses become clear. A basic difference lies in the fact that IRCs share their entire INDEX value with their antecedent, whereas SRCs and their antecedents only share their INDEX|PHI features. This difference is a consequence of the discourse anaphoric treatment of the SRC-antecedent relation proposed in Arnold (2007) and can be expressed directly in the structure of the INDEX value that has been proposed for LRS independently.

8.5 A compact theory of English relative clauses

In this final section, I will summarize the theory of relative clauses that emerges from this thesis. The difference between IRCs and SRCs is primarily semantic. Leaving the details of LRS aside, the main difference is that there is full index identity between the relativized element and the antecedent in IRCs. In SRCs, there is only the PHI features that are shared between the relativised element and the antecedent. The DR value is different.

I proposed two relativisers for Sorānī Kurdish, a phonologically empty one and *ke*. Only the latter is compatible with an SRC semantics. The analysis of Sorānī Kurdish is identical to that of English bare RCs and *that* RCs.

The English system of RCs is, however, considerably more complex, as it includes *wh*-RCs in addition. In my analysis in chapter 5, I assumed an empty relativiser with an SRC semantics – presupposing the need for an empty relativiser with an IRC semantics. Similarly, I provided lexical entries for relative words, such as pronoun-*which*, with an explicit SRC semantics. Again, this suggests that I assume analogous lexical entries for those relative words that can occur in IRCs.

I will now show how the syntax and semantics of IRCs and SRCs can be factored out to achieve an analysis as alluded to in the introductory chapter. My final proposal, therefore, has the following components:

- a single constraint on elements with IRC or SRC semantics (Figure 8.1)
- a single lexical entry for relative *that* (Figure 8.2),
- a single lexical entry of a bare RC relativiser (Figure 8.3),
- a single lexical entry for the relativiser for *wh*-RCs (Figure 8.4)

$\left[\textsc{syns}|\textsc{loc}|\textsc{cat}|\textsc{head}\ \textit{relativiser}\right] \rightarrow$

(IRC-relativiser)

$$\left[\begin{array}{ll}\textsc{syns} & \left[\textsc{loc}\ \left[\begin{array}{ll}\textsc{cat} & \left[\begin{array}{ll}\textsc{head}|\textsc{sel} & \mathrm{N}\ \left[\textsc{loc}|\textsc{cont}|\textsc{index}\ \boxed{1}\right]\\ \textsc{val}|\textsc{comps} & \left\langle\left[\textsc{loc}|\textsc{cont}|\textsc{main}\ \boxed{2}\right]\right\rangle\end{array}\right]\\ \textsc{cont} & \left[\textsc{index}\ \boxed{1}\ \left[\textsc{dr}\ x\right]\right]\end{array}\right]\right]\\ \textsc{lrs} & \left[\begin{array}{ll}\textsc{excont} & \alpha\wedge\beta\\ \textsc{incont} & \boxed{2}\\ \textsc{parts} & \left\langle\alpha\wedge\beta,\ x\right\rangle\end{array}\right]\end{array}\right]$$

and $\boxed{2} \leq \beta$

or

(SRC-relativiser)

$$\left[\begin{array}{ll}\textsc{syns} & \left[\textsc{loc}\ \left[\begin{array}{ll}\textsc{cat} & \left[\textsc{head}|\textsc{sel}\ \mathrm{XP}\ \left[\textsc{loc}|\textsc{cont}|\textsc{index}\ \left[\textsc{dr}\ x\right]\right]\right]\\ \textsc{cont} & \left[\begin{array}{ll}\textsc{index} & \left[\textsc{dr}\ w\right]\\ \textsc{main} & w=x\end{array}\right]\end{array}\right]\right]\\ \textsc{lrs} & \left[\begin{array}{ll}\textsc{excont} & \alpha\wedge\beta\\ \textsc{incont} & w=x\\ \textsc{parts} & \left\langle\alpha\wedge\beta,\ w,\ \iota w:\gamma,\ w=x\right\rangle\end{array}\right]\end{array}\right]$$

and $(\iota w:\gamma) \leq \beta$ and $(w=x) \leq \gamma$

Figure 8.1: Constraint on the relativiser semantics

- a single lexical entry for each relative word (Figure 8.5)

The basic difference between IRCs and SRCs is stated in Figure 8.1, which is slightly adjusted from Figure 7.3.

In English, we have three relativisers, which correspond to the three formal types of RCs: relative *that*, an empty relativiser for bare RCs, and an empty relativiser for *wh*-RCs. The first and the last of them are compatible with both types of relative clause semantics.

The lexical entry of relative *that* is very much like that of Sorānī Kurdish *ke*: it has a non-empty phonology and is compatible with both an IRC and and SRC semantics. However, it introduces an element in its INHER|REL feature. I provide

the lexical entry in Figure 8.2.

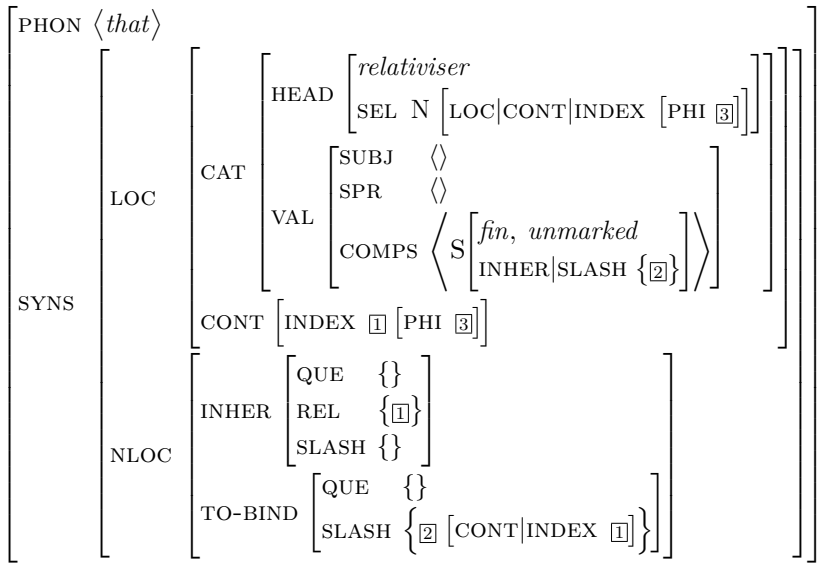


Figure 8.2: Sketch of the lexical entry of the English relativiser *that*

I will, next, turn to the empty relativiser for English bare relatives. Its lexical entry is give in Figure 8.3. It is the only relativiser that is restricted to IRCs. This is captured by requiring full identity of the INDEX values of the antecedent, the relative clause, and the gap it contains. Below the AVM, I express the constraint that a bare RC cannot have a gap in its local subject position. Note, that bare RCs have an empty INHER|REL set.

We can, now, turn to *wh*-RCs. I provide the lexical entry of the relativiser in Figure 8.4. It selects the extracted element on its SUBJ list and the core relative clause on its COMPS list. The semantics of the extracted element is integrated into the overall semantics as part of the second conjunct. It is compatible with both an IRC and an SRC semantics.

I will, next, provide the lexical entry of pronoun *which*. The relative word can occur in IRCs and SRCs. In chapter 5, Figure 5.12, I only considered a version of this pronoun that would appear in SRCs. Now, however, I will present

$$\begin{bmatrix}\text{PHON } \langle\rangle \\ \text{SYNS} \begin{bmatrix}\text{LOC} \begin{bmatrix}\text{CAT} \begin{bmatrix}\text{HEAD} \begin{bmatrix}\textit{relativiser} \\ \text{SEL } \text{N}' \left[\text{LOC}|\text{CONT}|\text{INDEX } \boxed{1}\right]\end{bmatrix} \\ \text{VAL} \begin{bmatrix}\text{SUBJ } \langle\rangle \\ \text{SPR } \langle\rangle \\ \text{COMPS } \left\langle \text{S} \begin{bmatrix}\textit{fin},\ \textit{unmarked} \\ \text{LOC}|\text{CAT}|\text{HEAD } \left[\text{XARG } \boxed{3}\right] \\ \text{NLOC}|\text{INHER}|\text{SLASH } \left\{\boxed{2}\right\}\end{bmatrix} \right\rangle\end{bmatrix}\end{bmatrix} \\ \text{CONT } \left[\text{INDEX } \boxed{1}\right]\end{bmatrix} \\ \text{NLOC} \begin{bmatrix}\text{INHER} \begin{bmatrix}\text{QUE } \{\} \\ \text{REL } \{\} \\ \text{SLASH } \{\}\end{bmatrix} \\ \text{TO-BIND} \begin{bmatrix}\text{QUE } \{\} \\ \text{REL } \{\} \\ \text{SLASH } \left\{\boxed{2} \left[\text{CONT}|\text{INDEX } \boxed{1}\right]\right\}\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

and $\boxed{2} \neq \boxed{3}$

Figure 8.3: Sketch of the lexical entry of the null relativiser for bare RCs

a version that is compatible with either an IRC or an SRCs. This lexical entry is much simpler, as the IRC/SRC semantics will be contributed exclusively by the relativiser.[2]

I will illustrate this briefly. Example (1) contains a *wh*-IRCs with pronoun-*which* as its relative word.

(1) Every book which Kim read was interesting.

$$\forall x((\mathbf{book}'(x) \wedge (\mathbf{non\text{-}human}'(x) \wedge \mathbf{read}'(\mathbf{kim}', x))) \rightarrow \mathbf{interesting}'(x))$$

For this sentence, the combination of the *wh*-relativiser with the constraint on IRCs will be chosen. The result is given in Figure 8.6. Because of the full INDEX identity in IRCs between the antecedent and the relativized element, the figure is

[2] I also simplify the lexical semantics from $\neg_{\text{human}}(x)$ to using a constant **non-human**′ that denotes any object that is not human.

$$\left[\begin{array}{ll}
\textsc{phon} & \langle\rangle \\
\textsc{syns} & \left[\begin{array}{ll}
\textsc{loc} & \left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll}
\textsc{head} & \left[\begin{array}{ll} \textit{relativiser} & \\ \textsc{sel} \ \ \mathrm{N} & \left[\begin{array}{l} \textsc{loc}|\textsc{cont}|\textsc{index} \ \ \boxed{4}\left[\textsc{phi}\ \boxed{3}\right] \\ \textsc{nloc}|\textsc{to-bind}|\textsc{rel} \ \ \{\boxed{4}\} \end{array}\right] \end{array}\right] \\
\textsc{val} & \left[\begin{array}{ll} \textsc{subj} & \left\langle\left[\begin{array}{ll} \textsc{loc} & \boxed{2}\left[\textsc{cont}|\textsc{main}\ \boxed{5}\right] \\ \textsc{nloc} & \left[\textsc{inh}|\textsc{rel}\ \{\boxed{4}\}\right] \end{array}\right]\right\rangle \\ \textsc{spr} & \langle\rangle \\ \textsc{comps} & \left\langle \mathrm{S}\left[\begin{array}{l} \textit{fin},\ \textit{unmarked} \\ \textsc{inher}|\textsc{slash}\ \{\boxed{2}\} \end{array}\right]\right\rangle \end{array}\right] \end{array}\right] \\
\textsc{cont} & \left[\textsc{index}\ \boxed{1}\left[\textsc{phi}\ \boxed{3}\right]\right]
\end{array}\right] \\
 & \textsc{nloc}\left[\begin{array}{ll}
\textsc{inher} & \left[\begin{array}{ll} \textsc{que} & \{\} \\ \textsc{rel} & \{\boxed{4}\} \\ \textsc{slash} & \{\} \end{array}\right] \\
\textsc{to-bind} & \left[\begin{array}{ll} \textsc{que} & \{\} \\ \textsc{rel} & \{\} \\ \textsc{slash} & \{\boxed{2}\} \end{array}\right]
\end{array}\right] \\
\textsc{lf} & \left[\textsc{excont}\ \ \alpha \wedge \beta\right]
\end{array}\right] \\
\text{and } \boxed{5} \leq \beta
\end{array}\right]$$

Figure 8.4: Sketch of the lexical entry of the English relativiser for *wh* RCs

slightly simpler than the earlier one.

When this relativiser combines with the core relative clause from example (1), we arrive at the following LRS value.

(2) LRS value of combining the *wh*-relativiser and the core RC from (1), ∅ *Kim read*:

$$\left[\begin{array}{ll}
\textsc{excont} & \boxed{6}\ \alpha \wedge \beta \\
\textsc{incont} & \boxed{5}\ \mathbf{read'} \\
\textsc{parts} & \left\langle \mathbf{kim'}, x, \mathbf{read'}, \mathbf{read'}(\mathbf{kim'}, x), \alpha \wedge \beta \right\rangle
\end{array}\right]$$
$$\text{and } \mathbf{read'} \leq \beta$$

Next, we add the relative pronoun from Figure 8.5 to the combination of the *wh*-relativiser and the core relative clause with the relative pronoun. This leads

Shorthand of the semantic contribution: $[\textbf{non-human}'(w)] \wedge \delta[w]$

$$\left[\begin{array}{ll}
\textsc{phon} & \langle \textit{which} \rangle \\
\textsc{syns} & \left[\begin{array}{ll}
\textsc{loc} & \left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll} \textsc{head} & \textit{noun} \\ \textsc{val} & \left[\begin{array}{l} \textsc{subj}\ \langle\rangle \\ \textsc{spr}\ \langle\rangle \\ \textsc{comps}\ \langle\rangle \end{array}\right] \end{array}\right] \\
\textsc{cont} & \left[\begin{array}{ll} \textsc{main} & \textbf{non-human}' \\ \textsc{index} & \left[\begin{array}{ll} \textsc{dr} & w \\ \textsc{phi} & \boxed{2}\left[\begin{array}{ll} \textsc{per} & \textit{3rd} \\ \textsc{gen} & \textit{neuter} \end{array}\right] \end{array}\right] \end{array}\right]
\end{array}\right] \\
 & \textsc{nloc}\ \left[\textsc{inhr}|\textsc{rel}\left\{\left[\textsc{phi}\ \boxed{2}\right]\right\}\right]
\end{array}\right] \\
\textsc{lrs} & \left[\begin{array}{ll}
\textsc{excont} & \boxed{4}\ (\textbf{non-human}'(w) \wedge \delta[w]) \\
\textsc{incont} & \boxed{3}\ \textbf{non-human}'(w) \\
\textsc{parts} & \left\langle w, \textbf{non-human}', \boxed{3}, \boxed{4} \right\rangle
\end{array}\right]
\end{array}\right]$$

Figure 8.5: Description of pronoun-*which* for RCs

to the semantic specification in (3).

(3) LRS value of full IRC in (1), *which* ∅ *Kim read*:

$$\left[\begin{array}{ll}
\textsc{excont} & \boxed{6}\ \alpha \wedge \beta \\
\textsc{incont} & \boxed{5}\ \textbf{read}' \\
\textsc{parts} & \left\langle \begin{array}{l} \textbf{kim}', x, \textbf{read}', \textbf{read}'(\textbf{kim}', x), \alpha \wedge \beta, \\ x, \textbf{non-human}', \textbf{non-human}'(x), (\textbf{non-human}'(x) \wedge \delta[x]) \end{array} \right\rangle
\end{array}\right]$$

and $\textbf{read}' \leq \beta$ and $\textbf{non-human}'(x) \leq \beta$

Only the formula in (4) satisfies all the constraints on the reading collected in the relative clause.

(4) Semantic representation of the relative clause in (1), *which* ∅ *Kim read*:

$$\alpha[x] \wedge (\textbf{non-human}'(x) \wedge \textbf{read}'(\textbf{kim}', x))$$

$$\left[\begin{array}{ll}
\textsc{phon} & \langle\rangle \\
\textsc{syns} & \left[\begin{array}{ll}
\textsc{loc} & \left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll}
\textsc{head} & \left[\begin{array}{ll} \textit{relativiser} & \\ \textsc{sel} \ \mathrm{N} & \left[\begin{array}{l} \textsc{loc}|\textsc{cont} \left[\textsc{index} \ \boxed{4}\right] \\ \textsc{nloc}|\textsc{to-bind}|\textsc{rel} \left\{\boxed{4}\right\} \end{array}\right] \end{array}\right] \\
\textsc{val} & \left[\begin{array}{ll} \textsc{subj} & \left\langle \left[\begin{array}{ll} \textsc{loc} & \boxed{2} \\ \textsc{nloc} & \left[\textsc{inh}|\textsc{rel} \left\{\boxed{4}\right\}\right] \end{array}\right] \right\rangle \\ \textsc{spr} & \langle\rangle \\ \textsc{comps} & \left\langle \mathrm{S} \left[\begin{array}{l} \textit{fin}, \ \textit{unmarked} \\ \textsc{main} \ \boxed{5} \\ \textsc{inher}|\textsc{slash} \left\{\boxed{2}\right\} \end{array}\right] \right\rangle \end{array}\right] \end{array}\right] \\
\textsc{cont} \left[\textsc{index} \ \boxed{4}\right]
\end{array}\right] \\
\textsc{nloc} & \left[\begin{array}{ll}
\textsc{inher} & \left[\begin{array}{ll} \textsc{que} & \{\} \\ \textsc{rel} & \left\{\boxed{4}\right\} \\ \textsc{slash} & \{\} \end{array}\right] \\
\textsc{to-bind} & \left[\begin{array}{ll} \textsc{que} & \{\} \\ \textsc{rel} & \{\} \\ \textsc{slash} & \left\{\boxed{2}\right\} \end{array}\right]
\end{array}\right]
\end{array}\right] \\
\textsc{lrs} & \left[\begin{array}{ll} \textsc{excont} & \boxed{6} \ \alpha \wedge \beta \\ \textsc{incont} & \boxed{5} \\ \textsc{parts} & \left\langle \boxed{6}, \ x \right\rangle \end{array}\right]
\end{array}\right]$$

and $\boxed{5} \leq \beta$

Figure 8.6: Sketch of the IRC version of the English relativiser for *wh*-RCs

For the sake of completeness, I will show the analogous derivation of an SRC with pronoun-*which*, as in example (5).

(5) LGB, which Kim read, is interesting.

$$\mathbf{interesting'}(\mathbf{lgb'} \wedge \mathbf{read'}(\mathbf{kim'}, (\iota w : (\mathbf{non\text{-}human'}(w) \wedge w = \mathbf{lgb'}))))$$

The combination of the *wh*-relativiser with the constraint on SRCs leads to the description in Figure 8.7.

Combining the *wh*-relativiser from Figure 8.7 with the core relative clause results in the LRS value given in (6).

$$\begin{bmatrix}\text{PHON } \langle\rangle \\ \text{SYNS} \begin{bmatrix}\text{LOC} \begin{bmatrix}\text{CAT} \begin{bmatrix}\text{HEAD} \begin{bmatrix}\textit{relativiser} \\ \text{SEL N} \begin{bmatrix}\text{LOC}|\text{CONT}|\text{INDEX} \ \boxed{4} \begin{bmatrix}\text{DR } x \\ \text{PHI } \boxed{3}\end{bmatrix} \\ \text{NLOC}|\text{TO-BIND}|\text{REL } \{\boxed{4}\}\end{bmatrix}\end{bmatrix} \\ \text{VAL} \begin{bmatrix}\text{SUBJ } \left\langle \begin{bmatrix}\text{LOC } \boxed{2}\left[\text{CONT}|\text{MAIN } \boxed{5}\right] \\ \text{NLOC } \left[\text{INH}|\text{REL } \{\boxed{4}\}\right]\end{bmatrix} \right\rangle \\ \text{SPR } \langle\rangle \\ \text{COMPS } \left\langle \text{S}\begin{bmatrix}\textit{fin},\ \textit{unmarked} \\ \text{INHER}|\text{SLASH } \{\boxed{2}\}\end{bmatrix} \right\rangle\end{bmatrix}\end{bmatrix} \\ \text{CONT} \begin{bmatrix}\text{INDEX } \boxed{1}\begin{bmatrix}\text{DR } w \\ \text{PHI } \boxed{3}\end{bmatrix} \\ \text{MAIN } w = x\end{bmatrix}\end{bmatrix} \\ \text{NLOC} \begin{bmatrix}\text{INHER} \begin{bmatrix}\text{QUE } \{\} \\ \text{REL } \{\boxed{4}\} \\ \text{SLASH } \{\}\end{bmatrix} \\ \text{TO-BIND} \begin{bmatrix}\text{QUE } \{\} \\ \text{REL } \{\} \\ \text{SLASH } \{\boxed{2}\}\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \text{LRS} \begin{bmatrix}\text{EXCONT } \alpha \wedge \beta \\ \text{INCONT } w = x \\ \text{PARTS } \langle \alpha \wedge \beta, w, \iota w : \gamma, w = x \rangle\end{bmatrix}\end{bmatrix}$$

and $\boxed{5} \leq \beta$ and $\iota w : \gamma \leq \beta$ and $w = x \leq \gamma$

Figure 8.7: Sketch of the SRC version of the English relativiser for *wh*-RCs

(6) LRS value of combining the *wh*-relativiser and the core RC from (5), $\emptyset$ *Kim read*:

$$\begin{bmatrix}\text{EXCONT } \alpha \wedge \beta \\ \text{INCONT } w = x \\ \text{PARTS } \left\langle \begin{matrix}\mathbf{kim}', w, \mathbf{read}'(\mathbf{kim}', w), \\ \alpha \wedge \beta, w, \iota w : w = x, w = x\end{matrix} \right\rangle\end{bmatrix}$$

and $\mathbf{read}' \leq \beta$ and $\iota w : \gamma \leq \beta$ and $w = x \leq \gamma$

In the last step, we add the relative pronoun-*which*. This results in the LRS

value in (7).

(7) LRS value of the full SRC from (5), *which* ∅ *Kim read*:

$$\left[\begin{array}{ll} \textsc{excont} & \alpha \wedge \beta \\ \textsc{incont} & w = x, \\ \textsc{parts} & \left\langle \begin{array}{l} \mathbf{kim}', w, \mathbf{read}'(\mathbf{kim}', w), \\ \alpha \wedge \beta, w, \iota w : w = x, w = x \\ \mathbf{non\text{-}human}'(w), \mathbf{non\text{-}human}'(w) \wedge \delta[w] \end{array} \right\rangle \end{array}\right]$$

and $\mathbf{read}' \leq \beta$ and $\iota w : \gamma \leq \beta$ and $w = x \leq \gamma$ and $\mathbf{non\text{-}human}' \leq \beta$

The semantic representation compatible with the cosntraints in (7) is given in (8).

(8) Semantic representation of the SRC in (5), *which* ∅*Kim read*:

$$\alpha \wedge \mathbf{read}'(\mathbf{kim}', \iota w : (\mathbf{non\text{-}human}'(w) \wedge w = x))$$

The variable x in (8) is the DR value of the antecedent. In our example sentence, this is the constant $\mathbf{lgb}'$.

So far, I could show that a single lexical entry for a *wh*-relativiser and a single lexical entry for the relative pronoun-*which* are sufficient to account for their use both in IRCs and SRCs. It only remains to be shown how a relative word, such as determiner-*which* can be excluded from combining with the IRC version of the *wh*-relativiser.

Again, the basic technique to be used here is the split in the INDEX value: in the case of pronoun *which*, I did not commit to identity or non-identity of the DR values of the pronoun itself and the element in its INHER|REL value. Consequently, the pronoun was compatible with both IRCs and SRCs. In the case of determiner-*which*, I will require that there is a non-identity between the two DR values. This

results in the adapted lexical entry of determiner-*which* in Figure 8.8. Note that the symbol "$\neq$" is used to express the non-identity of the variable symbols, w and x.

$$\left[\begin{array}{ll}
\text{PHON} & \langle \textit{which} \rangle \\
\text{SYNS} & \left[\begin{array}{ll}
\text{LOC} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll}
\text{HEAD} & \left[\begin{array}{ll} \textit{det} & \\ \text{SEL} \;\; \text{N'} & \left[\text{LOC} \mid \text{CONT} \left[\begin{array}{ll} \text{INDEX} & \boxed{1} \\ \text{MAIN} & \boxed{3} \end{array}\right]\right] \end{array}\right] \\
\text{VAL} & \left[\begin{array}{ll} \text{SUBJ} & \langle\rangle \\ \text{SPR} & \langle\rangle \\ \text{COMPS} & \langle\rangle \end{array}\right]
\end{array}\right] \\
\text{CONT} & \left[\begin{array}{ll} \text{INDEX} & \boxed{1}\left[\begin{array}{ll} \text{DR} & w \\ \text{PHI} & \boxed{2} \end{array}\right] \\ \text{MAIN} & \boxed{3} \end{array}\right]
\end{array}\right] \\
\text{NLOC} & \left[\text{INHER}\mid\text{REL} \left\{\left[\begin{array}{ll} \text{DR} & x \\ \text{PHI} & \boxed{2} \end{array}\right]\right\}\right]
\end{array}\right] \\
\text{LRS} & \left[\begin{array}{ll} \text{EXCONT} & \boxed{4} \;\; \gamma[w] \wedge \delta[w] \\ \text{INCONT} & \boxed{3} \\ \text{PARTS} & \langle w, \boxed{3}, \boxed{4} \rangle \end{array}\right]
\end{array}\right]$$

and $w \neq x$ and $\boxed{3} \leq \gamma$

Figure 8.8: Description of determiner-*which* for SRCs

In this version, the lexical entry of determiner-*which* does not contain an ι operator. Instead, it simply integrates the semantics of the head noun that it combines with into the first conjunct of its EXCONT value. This has the consequence, that a relative constituent of the form *which book* will have the EXCONT value $\mathbf{book}'(w) \wedge \delta$, i.e., an EXCONT value that looks very much like that of pronoun-*which*.

This concluded the presentation of the final version of my theory of English

relative clauses. The proposal captures the syntactic similarities of IRCs and SRCs in English. At the same time, it respects the semantic differences. It also allows to constrain both relativisers and relative words to only one type of RCs, if needed. This is necessary for the relativiser for bare RCs, which can only occur with IRCs, and for determiner-*which*, which is restricted to SRCs.

Appendix A

Questionnaires

A.1 Questionnaire 1: Determiner-*which*

Respondent's Personal Details .....

Gender:

1. Female
2. Male
3. Prefer not to say

Age: .....

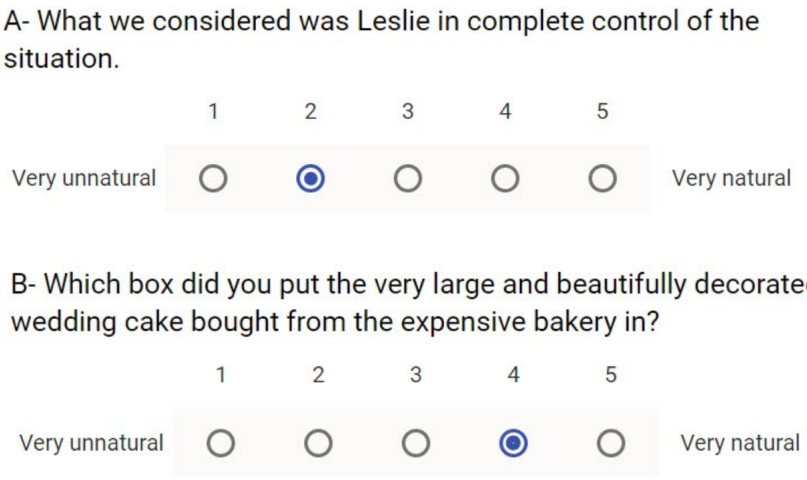
First Language(s): .....

Highest Qualification: .....

Instructions: Please check the scale underneath each sentence to rate the examples from sounding very unnatural English to you to very natural based on the following scales:

Very unnatural	(1)
Unnatural	(2)
I can't decide	(3)
Natural	(4)
Very natural	(5)

Sample Answer The following examples are answered to help you understand the questions better



Items

- From the complement of preposition to PP

    1. Nurses may also voice concerns about newer electronic documentation methods interrupting workflow, in which **case** they need to become personally involved in workflow design.

    2. I visited Beau Stubbs on Death Row, where numerous appeals kept him from the gallows for a decade, during which **time** he used prison to learn a trade and became a highly skilled airline pilot.

    3. The typical scenario in which employees volunteer on company time is when management sanctions the project. An example was the United Way fall kickoff, in which **project** several companies allowed employees to help renovate Amberwood Village Apartments, a rundown complex in east Atlanta's Reynoldstown community.

4. The youngest nephew of Juana Williams won the Jackpot, with which **money** she built a hospital in her village.

- From PP to Adj

1. The many varieties of mammalian skin secretions perform a wide range of functions, prominent among which **purposes** is sexual attraction.

2. Up to date, DiCaprio featured in 33 movies, the best among which **films** is 'Titanic'.

3. Alan Hydros remembers the miseries he has been through while he lived with his biological parents, the most painful of which **memories** is when they became homeless for a while.

4. During my studies, I became very interested in the field of e-learning. It led me to various other opportunities, the most exciting of which **chances** was to work with a local business accelerator to share my story with a new generation of start-ups, helping young entrepreneurs to avoid my mistakes and to communicate best practices.

- From PP to a NP

1. They are members of an association, the first and most precious principle of which **society** is mutual trust.

2. Ala asked the HR manager if the boss was planing to dismiss her this month, the answer to which **question** I already knew.

3. He is wearing a threadbare flannel shirt with a massive blue ink stain beneath the breast pocket, a thrift-store leather jacket, and a fucked-up green trucker's cap, the color of which **cap** seems to complement perfectly his hazel eyes, staring out hauntedly from the deep recesses of his unshaven face.

4. It took them half a day to reach the nearest village on their mule-drawn carriage, the wheels of which **wagon** were wooden.

- From NP to infinitival VP

1. In 2017, 13 women were killed in the name of 'honour' in six different Iraqi cities, to condemn which **crime** the female activists protested for one week in front of UN headquarter.

2. The sexual assault allegation had shaken Jackson's famous ground, to repudiate which **claims** he spent millions of dollars.

3. I became disturbed by a 'higher criticism' of bible, to refute which **judgment** I felt the need of a better knowledge of Hebrew and archeology.

4. Jennifer Lawrence won an Oscar last year, to accept which **award** she had to be present at the ceremony.

- From NP to gerund participial

1. They take a rigorous examination, passing which **test** confers on the students a virtual guarantee of a place at the university.

2. The parents of the victims insisted on having a face-to-face discussion with the mayor, arranging which **meeting** was my responsibility.

3. Afshin Kaki composed 'the dew drops' when he was 22, performing which **piece of music** at the Grammys in 2015 made him the famous man he is now.

4. The employees had to write a proposal for a small project of their choice, refusing which **task** would decrease their salaries.

Fillers

1. The deleterious effects of the transmission line were found to be statistically significant, seemingly small at first glance although.

2. Sherman-Grant note, this is a 15-cent banknote, depicting Union Army generals William Tecumseh Sherman and Ulysses S. Grant, dated 1866 and intended as part of the fractional currency introduced to the United States following the American Civil War.

3. I remember that mighty well, too. But when the Big Bubble busted back in' 29 , seems like it took half the magic in the world with it when it went.

4. I burst into tears looking at the snowflake-spotted pictures of my mother in her hoop skirt costume for the May Day parade. nd the returns in the final column of Exhibit 4, most core and value-add funds are unlikely to have achieved their performance targets.

5. This means the manager will be paid performance fees for any return below

the hurdle rate, but only after achieving this rate, hence the name "catch-up".

6. Arkansas Attorney General Leslie Rutledge immediately appealed that order to the U.S. 8th Circuit Court of Appeals, and as of Monday afternoon, seem determined to hold her ground.

7. Clinic visits include a medical record review, physical examination and laboratory screening to test for hepatitis B and C, syphilis, tuberculosis and HIV.

8. These supermarkets are the cheapest ones in the country, and you can find them in every corner.

9. It's everything else about our machinery of death, what else could it be, that has been thrown starkly into doubt.

10. They pick up on our biases, and they can learn biases that we didn't even know we had. That is going to be, I think, one of the most immediate detrimental effects of machine learning on our daily lives.

11. The iconic car of the electric age, as it is obvious to everyone, must be a Volkswagen. Volkswagen's original Type 1 Microbus debuted in 1950 and had just 30 horsepower

12. The bathroom door was broken, to me a sign of struggle it was, but to them it just was a result of her falling over it!

13. What if one identifies all the sources of energy dissipation in the existing structure and works to reduce or eliminate such losses through redesigning

the guitar or substituting different materials, could one build a better, louder acoustic guitar?

14. Every night I had to go out, collect the chickens, and lock them back in, though it was the hardest job of the day!

15. The first hurdle to doing so lay in articulating the objective. Because " uniformity " or " consistency " might mean different things to different people, the referees devoted considerable time to discussing what type of uniformity would be feasible without placing undue constraints on referees.

16. It had probably been a maid's room in the old days of the house, but we believe it to be Marissa who had decorated it in white on white and used it for watching TV.

17. These three components resulted in the Maryland tax scheme imposing an unreasonable burden on interstate commerce as compared to intrastate commerce

18. As most of the intended recipients had died by then, and relations with France had improved, the medals were never struck, though modern-day editions have been made for sale to collectors.

19. It wasn't how I thought it would look, but here it was and it was my mother, not God, who giving it to me. I thought then of the cost of such a chance.

20. Teachers emphasize competition rather than creativity and self-actualization, that their students will move one step closer to success.

21. There are a great number of factors might work in practice to influence an offender's sentence in any given case beyond the occurrence of a statutory

harm, such as his past criminal history.

22. It will all be well by dawn, she told her this, but Emily could sense the uncertainty in her voice, asking herself, what can happen to set everything right?! A miracle maybe!

23. A Waterloo Medal was designed by sculptor Benedetto Pistrucci. Commemorating the Battle of Waterloo (18 June 1815), those medals were commissioned by the British Government in 1819 on the instructions of George IV.

24. Nobel laureates in Physics are determined annually by the Royal Swedish Academy of Sciences; the Nobel Prize in Physics is awarded to scientists in the various fields of physics.

25. The man, on the bus we saw, turned out to be from my home town.

26. The bathroom door was broken, to me a sign of struggle it was, but to them it just was a result of her falling over it!

27. By the time they reported it to the police, the whole village were out in the woods looking for them. No one could sleep even that night.

28. So! Even this was not enough for you! Listen to me very well, he said, you are the kind of people, who are not happy in your own skin! I guess you need to let it go!

29. The school teaches in both Spanish and English, however, I believe the most concentration is on the English language due to the fact that most of their students are international.

30. The three weeks of holiday that every employee has in the year will be extend to 30 days starting from February.

31. Assume further that two taxpayers, April and Bob, both live in State A, that April earns her income in State A, whereas Bob earns his income in State B. In this circumstance, Bob will pay more income tax than April solely as he earns income interstate.

32. This article proposes a solution to this long-standing conceptual problem. We begin by introducing a dichotomy between two broad and exhaustive categories of ideological justifications in punishing criminal offenders.

33. Chemical profde in IRIS presents an acceptable numerical dose that, if ingested, inhaled, or absorbed through the skin, could cause cancer, could respiratory illness, brain damage, or other adverse health effects.

34. While the movie seems simple and amusing in many ways, when approached from the perspective of how it speaks to the role of robots in our society, it , as well, raises some disturbing issues.

35. It turns out that Frank's health is Robot's top priority, superseding all other considerations including the wellbeing of others, during the course of the movie, we find that Robot has a central role in steering Frank back into the right mood.

36. Rather than striking out against the system or attempting to subvert the rules, inmates may internalize their resentment, enact violence to themselves, or employ other coping mechanisms.

37. Franzen combats this shame as he continues his treatise on " The Problem

of Hard-to-Read Books " with a plea to literary critics, saying, " Think of the novel as lover: Let's stay home tonight and have a great time.

38. Some theorists have also tried to justify the inclusion of results into a retributive theory of punishment by analogizing punishment to a moral lottery.

39. The employees had to write a proposal for a small project of their choice, refusing which task would decrease their salaries.

40. Giving the time period and the returns in the final column of Exhibit 4, most core and value-add funds are unlikely to have achieved their performance targets.

A.2 Questionnaire 2: SRCs with *that*

The items containing SRCs with *that/which*

1. The big topic this week was this video, that/which Mitt Romney uploaded on YouTube, that, according to reliable sources, had been filmed during a private party of an important sponsor of Romney's 2008 election campaign.

2. Our research found significant differences in the way the brains of musicians and non musicians respond to music, and more subtle differences, that/which, surprisingly, helped elucidate these different vulnerabilities to earworms, between women and men.

3. She sent me the draft for her first book, that/which she had finished in just two months' time by the way, and I immediately saw that she has all that it takes to become the next best-selling author of our publishing house.

4. A giant fox, that/which, if we can trust the local press releases, is twice the size of a normal specimen, has been captured in Kent,sparking fears that the animals are growing larger because of "easyliving" on bins and scraps.

5. The key point, that/which, in my opinion, all the popular reports missed, is that FOXP2 is a transcription factor.

6. Despite her request to take her to a good specialist, her brother-in-law brought some Ayurvedic ointment, that/which, not surprisingly, helped only to spread the sores all over her body, and her condition stayed like that for more than 4 months

7. Anna's wonderful power point presentation, that/which I am sure you can download from her web site by now, contains all relevant information on how to run a small consulting company from your living room.

8. Then, after all the stakes were in place, we strung the wire, hanging it from one stake to the next. The actual stringing required a good sense of balance, that/which, luckily, most of us still had, and a considerable amount of physical strength.

9. There were some problems of security, but, Governor, they sold this narrative, that/which I know you are familiar with now.

10. The teachers emphasized competition between the twins, that/which, by the way, the two never took very seriously.

11. His heart, that/which had lifted at the sight of Joanna, had become suddenly heavy at the sight of Ramdez.

12. Assume further that two taxpayers, April and Bob, both live in State A, that/which earns her me in State A whereas Bob earns his income in State B. In this circumstance, Bob will pay more income tax than April solely as he earns income interstate.

13. She had long been accustomed to the solitary nature of her son's instincts, that/which I have tried – and failed – to stifle.

14. The Patas monkey, that/which spends almost all of its time in open grassland, adopts just such tactics.

15. There's been a number of objections, the most telling objections, that/which I know about, have been made by Professor Bill Shots and Professor Ken

Nielson.

16. The deleterious effects of the transmission line were found to be statistically significant, seemingly small at first glance although.

17. Sherman-Grant note, this is a 15-cent banknote, depicting Union Army generals William Tecumseh Sherman and Ulysses S. Grant, dated 1866 and intended as part of the fractional currency introduced to the United States following the American Civil War.

18. I remember that mighty well, too. But when the Big Bubble busted back in' 29 , seems like it took half the magic in the world with it when it went.

19. I burst into tears looking at the snowflake-spotted pictures of my mother in her hoop skirt costume for the May Day parade.

20. This is not right I am telling you, whenever something would happen, he came up there and have to make a pronouncement and fuel the life of the spotlight on him.

21. Giving the time period and the returns in the final column of Exhibit 4, most core and value-add funds are unlikely to have achieved their performance targets.

22. This means the manager will be paid performance fees for any return below the hurdle rate, but only after achieving this rate, hence the name "catch-up".

23. Arkansas Attorney General Leslie Rutledge immediately appealed that order

to the U.S. 8th Circuit Court of Appeals, and as of Monday afternoon, seem determined to hold her ground.

24. Clinic visits include a medical record review, physical examination and laboratory screening to test for hepatitis B and C, syphilis, tuberculosis and HIV.

25. These supermarkets are the cheapest ones in the country, and you can find them in every corner.

26. It's everything else about our machinery of death, what else could it be, that has been thrown starkly into doubt.

27. They pick up on our biases, and they can learn biases that we didn't even know we had. That is going to be, I think, one of the most immediate detrimental effects of machine learning on our daily lives.

28. The iconic car of the electric age, as it is obvious to everyone, must be a Volkswagen. Volkswagen's original Type 1 Microbus debuted in 1950 and had just 30 horsepower

29. The bathroom door was broken, to me a sign of struggle it was, but to them it just was a result of her falling over it!

30. What if one identifies all the sources of energy dissipation in the existing structure and works to reduce or eliminate such losses through redesigning the guitar or substituting different materials, could one build a better, louder acoustic guitar?

31. Every night I had to go out, collect the chickens, and lock them back in, though it was the hardest job of the day!

32. The first hurdle to doing so lay in articulating the objective. Because " uniformity " or " consistency " might mean different things to different people, the referees devoted considerable time to discussing what type of uniformity would be feasible without placing undue constraints on referees.

33. It had probably been a maid's room in the old days of the house, but we believe it to be Marissa who had decorated it in white on white and used it for watching TV.

34. These three components resulted in the Maryland tax scheme imposing an unreasonable burden on interstate commerce as compared to intrastate commerce

35. As most of the intended recipients had died by then, and relations with France had improved, the medals were never struck, though modern-day editions have been made for sale to collectors.

36. At our first anniversary, I gave her a beautiful golden necklace, which, by the way, she sold just a few days after we broke up, and we went to have dinner at our favorite Mexican restaurant.

37. It wasn't how I thought it would look, but here it was and it was my mother, not God, who giving it to me. I thought then of the cost of such a chance.

38. The other thing is as other countries back way from this, as the medical profession backs away, you're now seeing, this guy was basically a lab experiment.

39. There are a great number of factors might work in practice to influence an offender's sentence in any given case beyond the occurrence of a statutory

harm, such as his past criminal history.

40. The employees had to write a proposal for a small project of their choice, refusing which task would decrease their salaries.

41. A Waterloo Medal was designed by sculptor Benedetto Pistrucci. Commemorating the Battle of Waterloo (18 June 1815), those medals were commissioned by the British Government in 1819 on the instructions of George IV.

42. Rather than striking out against the system or attempting to subvert the rules, inmates may internalize their resentment, enact violence to themselves, or employ other coping mechanisms.

43. Nobel laureates in Physics are determined annually by the Royal Swedish Academy of Sciences; the Nobel Prize in Physics is awarded to scientists in the various fields of physics.

44. The man, on the bus we saw, turned out to be from my home town.

45. It took them half a day to reach the nearest village on their mule-drawn carriage, the wheels of which wagon were wooden.

46. The bathroom door was broken, to me a sign of struggle it was, but to them it just was a result of her falling over it!

47. By the time they reported it to the police, the whole village were out in the woods looking for them. No one could sleep even that night.

48. It will all be well by dawn, she told her this, but Emily could sense the uncertainty in her voice, asking herself, what can happen to set everything right?! A miracle maybe!

49. In addition to providing hints for its consumption in print ads, the company published a booklet of recipes that, in the worldly-wise phrasing of one text, was available on a moment's notice, by the way , that moment is a Mexican moment.

50. The school teaches in both Spanish and English, however, I believe the most concentration is on the English language due to most of their students are international.

51. The three weeks of holiday that every employee has in the year will be extend to 30 days starting from February.

52. This article proposes a solution to this long-standing conceptual problem. We begin by introducing a dichotomy between two broad and exhaustive categories of ideological justifications in punishing criminal offenders.

53. Chemical profde in IRIS presents an acceptable numerical dose that, if ingested, inhaled, or absorbed through the skin, could cause cancer, could respiratory illness, brain damage, or other adverse health effects.

54. While the movie seems simple and amusing in many ways, when approached from the perspective of how it speaks to the role of robots in our society, it , as well, raises some disturbing issues.

55. It turns out that Frank's health is Robot's top priority, superseding all other considerations including the wellbeing of others, during the course of the movie, we find that Robot has a central role in steering Frank back into the right mood.

56. Jennifer Lawrence won an Oscar last year, to accept which award she has to be present at the ceremony.

57. The Division of Pulmonary, Allergy and Critical Care Medicine (DPACCM), which has been directed very successfully by Victor Thannickal, M.D., for the last five years, received widespread media attention for the creation of a 'mini lung' to study the effect of pulmonary fibrosis drugs.

58. Franzen combats this shame as he continues his treatise on " The Problem of Hard-to-Read Books " with a plea to literary critics, saying, " Think of the novel as lover: Let's stay home tonight and have a great time.

59. During my studies, I became very interested in the field of e-learning. It led me to various other opportunities, the most exciting of which chances was to work with a local business accelerator to share my story with a new generation of start-ups, helping young entrepreneurs to avoid my mistakes and to communicate best practices.

60. Some theorists have also tried to justify the inclusion of results into a retributive theory of punishment by analogizing punishment to a moral " lottery

Bibliography

Abeillé, Anne & Rui Chaves. 2019. Coordination. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, Berlin: Language Science Press. Prepublished version.

AnderBois, S., A. Brasoveanu & R. Henderson. 2015. At-issue proposals and appositive impositions in discourse. *Journal of Semantics* 32. 93–138.

AnderBois, Scott, Adrian Brasoveanu & Robert Henderson. 2010. Crossing the appositive/at-issue meaning boundary. *SALT* 20. 328–346.

Arnold, Doug. 2004. Non-restrictive relative clauses in construction based HPSG. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Leuven 2004*, 27–47. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/5/arnold.pdf.

Arnold, Doug. 2007. Non-restrictive relative clauses are not orphans. *Journal of Linguistics* 43. 271–309.

Arnold, Doug, Sascha Bargmann & Manfred Sailer. 2018. Update constraints on idioms and non-restrictive relatives. Presentation at the *European Workshop on HPSG* 2018.

Arnold, Doug & Danièle Godard. 2019. Relative clauses in HPSG. In *Head-driven phrase structure grammar: The handbook*, Berlin: Language Science Press.

Asadpour, Hiwa & Maryam Mohammadi. 2002. A comparative study of phonological system of kurdish varieties. *Journal of Language and Cultural Education* 2(3). 1339–4045.

Bach, Kent. 1999. The myth of conventional implicature. *Linguistics and Philosophy* 22(4). 327–366.

Bianchi, Valentina. 1999. *Consequences of antisymmetry: Headed relative clauses*. Berlin: Mouton de Gruyter.

Burton-Roberts, Noel. 1999. Language, linear precedence and parentheticals. in the clause in English, Peter Collins and David Lee (eds.). *Studies in Language Companion Series* 45. 33–51.

Bögel, Tina, Miriam Butt & Sebastian Sulger. 2008. Urdu ezafe and the morphology-syntax interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG08 conference*, 129–149. Stanford: CSLI Publications. http://web.stanford.edu/group/cslipublications/cslipublications/LFG/13/papers/lfg08boegeletal.pdf.

Canac-Marquis, Rejean & Mireille Tremblay. 1998. The wh-feature and the syntax of restrictive and non-restrictive relatives in French and English. In José Lema & Estela Trevini (eds.), *Theoretical analyses on Romance languages*, vol. 53, 127–142. Amsterdam: John Benjamins.

Carey, Stan. 2013. Non-restrictive 'that', that can be ambiguous. Blog

entry from 2013 with more recent comments. https://stancarey.wordpress.com/2013/12/15/non-restrictive-that-that-can-be-ambiguous/.

Carlson, Greg N. 1977. Amount relatives. *Language* 53. 520–542.

Chaves, Rui. 2020. Island phenomena and related matters, no pagenumers yet, 55 pages. Berlin: Language Science Press. https://hpsg.huberlin.de/Projects/HPSG-handbook/PDFs/islands.pdf. Prepublished version.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Cinque, Guglielmo. 1982. *On the theory of relative clause and markedness*. Cambridge: Cambridge University Press.

Coppock, Elizabeth & David Beaver. 2015. Definiteness and determinacy. *Linguistics and Philosophy* 38. 377–435.

Culicover, Peter W. 1993. Evidence against ECP accounts of the *that-t* effect. *Linguistic Inquiry* 24(3). 557–561.

Davies, Mark. 2008–. The Corpus of Contemporary American English: 450 million words, 1990-present. Available online at https://www.english-corpora.org/coca/.

Dehé, Nicole. 2009. Clausal parentheticals, intonational phrasing, and prosodic theory. *Journal of Linguistics* 45. 569–615.

Demirdache, Hamida. 1991. *Resumptive chains in restrictive relatives, appositives and dislocation structures*: MIT University dissertation.

Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*, vol. 42 (Typological Studies in Language). Amsterdam: John Benjamins.

Diessel, Holger. 2003. The relationship between demonstratives and interrogatives. *Studies in Language* 27. 635–655.

Elbourne, Paul. 2001. On the semantics of pronouns and definite articles. In Karine Megerdoomian & Leora Anne Bar-el (eds.), *WCCFL 20: Proceedings of the 20th West Coast Conference on Formal Linguistics*, 164–177. Somerville, MA: Casadilla Press.

Emonds, Joseph. 1979. Appositive relatives have no properties. *Linguistic inquiry* 10. 211–243.

Enright, Anne. 2007. *The gathering*. New York: Grove Press.

Espinal, Maria-Teresa. 1991. The representation of disjunct constituents. *Language* 67. 726–762.

Fabb, Nigel. 1990. The difference between English restrictive and nonrestrictive relative clauses. *Journal of Linguistics* 26. 57–78.

Fattah, Muhammad Maruf. 1997. *A generative grammar of Kurdish*: University of Amsterdam dissertation.

Gallin, Daniel. 1975. *Intensional and higher-order modal logic: With applications to Montague semantics*. Amsterdam: North Holland.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12. 155–184.

Geurts, Bart. 1999. *Presuppositions and pronouns*. Oxford: Elsevier.

Giannakidou, Anastasia & Josep Quer. 1997. Long-distance licensing of negative indefinites doi:M10.1075/cilt.155.06gia.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations. the form, meaning, and use of english interrogatives*. Stanford: CSLI Publications.

Grijzenhout, Janet. 1992. The change of relative that to *who* and *which* in late seventeenth-century comedies. *Language Evolution* 20. 33 – 52.

Groenendijk, Jeroen & Martin Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 14. 39–100.

Grosu, Alexander & Fred Landman. 1998. Strange relatives of the third kind. *Natural Language Semantics* 6(2). 125–170.

Haig, Geoffrey. 2007. Grammatical borrowing in Kurdish (Northern group). *Mouton de Gruyter* 9(4). 165–184.

Haig, Geoffrey. 2008. *Alignment change in Iranian languages. A construction grammar approach*. Belin. New York: Mouton de Gruyter.

Haig, Geoffrey. 2011. Linker, relativizer, nominalizer, tense-particle: On the Ezafe in West Iranian. In Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), *Nominalization in Asian languages: Diachronic and typological perspectives*, 363–390. John Benjamins.

Henderson, Robert. 2014. The dynamics of nominal apposition. Handout to a presentation University of Frankfurt Research Unit 1783 "Relative Clauses" guest lecture, Frankfurt.

Hendery, Rachel. 2007. The diachronic typology of relative clauses. Phil. Dissertation (2007). 301 The Australian National University. https://www.academia.edu/378945/The_diachronic_typology_of_relative_clauses.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Höhle, Tilman N. 2019. Featuring creatures of darkness. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur deutschen Grammatik. Gesammelte Schriften von Tilman N. Höhle*, 553–569. Berlin: Language Science Press 2nd edn. doi:M10.5281/zenodo.2588379. https://langsci-press.org/catalog/book/234. Previously unpublished article of a talk given 7.9.1994.

Jackendoff, Ray. 1977. *X-bar syntax: A study of phrase structure*. Cambridge: MIT Press.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer Academic Publishers.

Karimi, Yadgar. 2007. Kurdish Ezafe construction: Implications for DP structure. *Lingua* (117). 2159–2177. doi:Mdoi:10.1016/j.lingua.2007.02.010.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, Massachusetts: MIT Press.

Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1). 63–99.

Kempson, Ruth. 2003. Nonrestrictive relatives and growth of logical form. *WCCF* 22. 301–314.

Khalid, Hewa Salam & Atta Mostafa Hamamorad. 2015. New function and form of subordinate relative clause in Kurdish. *International Journal of Kurdish Studies* 1(2). 1–14. doi:MDOI:10.21600/ijks.31501.

Kilroy, Claire. 2010. *All names have been changed*. London: Faber and Faber.

Kim, Danielle. 2010. A basic guide to kurdish grammar: A self-study reference and practice book. URL: http://www.clik2rock.org.

Klein, Maarten. 1976. Appositionele NP's in het Nederlands. *De Nieuwe Taalgids* 69. 139–153.

Koster, Jan. 2000. *Extraposition as parallel construal*. University of Groningen.

Kraak, Remmert & Wim Klooster. 2000. *Syntaxis*. Culemborg: Stam-Kemperman.

Larson, Richard & Hiroko Yamakido. 2008. In Louise McNally & Christopher Kennedy (eds.), *Adjectives and adverbs: Syntax, semantics, and discourse*, 43–70. Oxford: Oxford University Press.

Lehmann, Christian. 1984. *Der Relativsatz*. Tübingen: Narr.

Levine, Robert D. & Thomas E. Hukari. 2006. *The unity of unbounded dependency constructions*. Stanford: CSLI Publications.

Levine, Robert D. & Walt Detmar Meurers. 2006. Head-Driven Phrase Structure Grammar: Linguistic approach, formal foundations, and computational realization. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, vol. 5, 237–252. Oxford: Elsevier 2nd edn. http://purl.org/dm/papers/ell2-hpsg.html.

Levine, Robert D., Frank Richter & Manfred Sailer. 2012. Formal semantics: An empirically grounded approach. Manuscript Columbus and Frankfurt a.M.

MacKenzie, D.N. 1962. *Kurdish dialect studies*. London: Oxford University Press.

McCawley, James. 1982. Parentheticals and discontinuous constituent structure. *Linguistics Inquiry* 13(1). 91–106.

McCawley, James D. 1981. The syntax and semantics of English relative clauses. *Lingua* 53. 99–149.

Montague, Richard. 1974. English as a formal language. In Richmond H. Thomason (ed.), *Formal philosophy. Selected papers of Richard Montague*, 188–221. Yale: Yale University Press.

Müller, Stefan. 2007. *Head-Driven Phrase Structure Grammar. Eine Einführung.* Tübingen: Stauffenburg.

Penn, Gerald & Frank Richter. 2004. Lexical resource semantics: From theory to implementation. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on HPSG 2004, Leuven*, 423–443. Stanford: CSLI Publications. cslipublications.stanford.edu/HPSG/5/penn-richter.pdf.

Pollard, Carl & Ivan Sag. 1987. *Information based syntax and semantics. vol.1: Fundamentals.* Stanford: CSLI Publications. CSLI Lecture Notes 13.

Pollard, Carl & Ivan A. Sag. 1994a. *Head-driven phrase structure grammar.* Chicago and London: University of Chicago Press.

Pollard, Carl & Ivan A. Sag. 1994b. *Head-driven phrase structure grammar.* Chicago and London: University of Chicago Press.

Potts, Christopher. 2005. *The logic of conventional implicatures.* Oxford: Oxford University Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language.* London: Longman.

Richter, Frank & Manfred Sailer. 2001a. On the left periphery of German finite sentences. In W. Detmar Meurers & Tibor Kiss (eds.), *Constraint-based approaches to Germanic syntax*, 257–300. Stanford: CSLI Publications.

Richter, Frank & Manfred Sailer. 2001b. Polish negation and lexical resource semantics. In Geert-Jan Kruijff, Lawrence S. Moss & Richard T. Oehrle (eds.), *Preproceedings of FGMOL 2001*, This paper has been superseded by Richter & Sailer (2002).

Richter, Frank & Manfred Sailer. 2002. Polish negation and lexical resource semantics. In Lawrence S. Moss & Richard T. Oehrle (eds.), *Electronic notes in theoretical computer science*, vol. 53, 309–321. Elsevier.

Richter, Frank & Manfred Sailer. 2004. Basic concepts of lexical resource semantics. In Arne Beckmann & Norbert Preining (eds.), *Esslli 2003 – course material i*, vol. 5 (Collegium Logicum), 87–143. Vienna: Kurt Gödel Society Wien.

Ross, John Robert. 1967. *Constraints on variables in syntax*: MIT dissertation. Published as Ross (1986).

Ross, John Robert. 1986. *Infinite syntax*. Norwood, New Jersey: Ablex Publishing.

Saady, Dilovan Sayfuddin. 2020. Reduced clauses in English and Kurdish. *Journal of Tikrit University for Humanities* 27(5). 106–118.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33. 431–483.

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar*, 69–202. Stanford: CSLI Publications.

Sailer, Manfred. 2003. Local semantics in HPSG. Paper presented at the *Colloque de Syntaxe et S@mantique*@ Paris, Paris, October 2003.

Sailer, Manfred. 2004. Past tense marking in Afrikaans. In Cecille Meier & Matthias Weisgerber (eds.), *Proceedings of sinn und bedeutung 8*, 233–248. Konstanz. www.ub.uni-konstanz.de/kops/volltexte/2004/1383.

Sailer, Manfred. 2007. NPI licensing, intervention and discourse representation structures in HPSG. In Stefan Müller (ed.), *Proceedings of the 14th international conference on hpsg, stanford, 2007*, 214–234. Stanford: CSLI Publications. `cslipublications.stanford.edu/HPSG/14/sailer.pdf`.

Sailer, Manfred. 2009. A representational theory of negative polarity item licensing. Habilitation thesis, Universität Göttingen.

Sailer, Manfred & Assif Am-David. 2016. Definite meaning and definite marking. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar, Polish Academy of Sciences, Warsaw, Poland*, 641–661. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2016/headlex2016-sailer-am-david.pdf.

Samvelian, Pollet. 2007a. The Ezafe as a head-marking inflectional affix: Evidence from Persian and Kurmanji Kurdish. In S. Karimi, V. Samiian & D. Stillo

(eds.), *Aspects of Iranian linguistics: Papers in honor of Mohammad Reza Bateni*, 339–361. Cambridge: Cambridge Scholars. halshs-00673182.

Samvelian, Pollet. 2007b. A lexical account of Sorani Kurdish prepositions. In Stefan Müller (ed.), *The proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 235–249. CSLI Publications. http://cslipublications.stanford.edu/HPSG/8/.

Samvelian, Pollet. 2007c. A (phrasal) affix analysis of the Persian Ezafe. *Journal of Linguistics* 43. 605–645.

Sells, Peter. 1985. *Restrictive and non-restrictive modification*. Stanford: CSLI Publications.

Shakespeare, William. 1997. *The complete works of Shakespeare*. Longman: Grove Press.

Taghvaipour, Mehran A. 2004. An HPSG analysis of Persian relative clauses. In Stefan Müller (ed.), *Proceedings of the HPSG04 Conference, Leuven*, 275–293. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2004/taghvaipour.pdf.

Taghvaipour, Mehran A. 2005. Persian relative clauses in Head-driven Phrase Structure Grammar, .

Tennyson, Alfred. 1974. *In memoriam*. London: J.M. Dent.

Thackston, Wheeler M. 2015. *Sorani Kurdish: A reference grammar with selected readings*. https://sites.fas.harvard.edu/ iranian/Sorani/sorani$_c$omplete.pdf.

van Eynde, Frank. 1998. The immediate dominance schemata of hpsg. a deconstruction and a reconstruction. In P.A. Coppen, H. van Halteren & L. Teunissen (eds.), *Computational linguistics in the netherlands 1997. selected papers from the eighth clin meeting*, 119–133. Amsterdam/Atlanta: Rodopi. http://www.ccl.kuleuven.ac.be/Papers.html.

van Eynde, Frank. 2003a. Morpho-syntactic agreement and index agreement in Dutch NPs. In Tanja Gaustad (ed.), *Computational linguistics in the Netherlands 2002. Selected papers from the thirteenth CLIN Meeting*, 111–127. Amsterdam: Rodopi.

van Eynde, Frank. 2003b. On the notion 'determiner'. In Stefan Müller (ed.), *Proceedings of the HPSG-2003 Conference, Michigan State University, East Lansing*, 391–396. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/4/.

van Eynde, Frank. 2004. Pied piping is a local dependency. In Stefan Müller (ed.), *Proceedings of the HPSG-2004 Conference, Center for Computational Linguistics, Katholieke Universiteit Leuven*, 313–334. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/5/.

van Eynde, Frank. 2006. NP-internal agreement and the structure of the noun phrase. *Journal of Linguistics* 42(1). 139–186.

Vergnaud, Jean Roger. 1974. *French relative clauses*: MIT dissertation.

de Vries, Mark. 2006. The syntax of appositive relaticisation;on specifying coordination, false free relatives, and promotion. *Linguistic inquiry* 37. 229–270.

Walker, Heike. 2017. *The syntax and semantics of relative clause attachment*: University Frankfurt dissertation. https://hds.hebis.de/ubffm/Record/HEB399426221s.

Wilcock, Graham. 1999. Lexicalization of context. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 373–387. Stanford: CSLI Publications. http://www.ling.helsinki.fi/ gwilcock/Pubs/CSLI-98.ps.

Öpengin, Ergin. 2013. *Clitic/affix interactions: A corpus based study of person marking in the Mukri variety of Central Kurdish*: University of Sorbonne Nouvelle & University of Bamberg dissertation.