Heinrich-Heine Universität Düsseldorf
Institut für allgemeine Sprachwissenschaft

Interpretation of the Topological Field Model of the German clause in Role and Reference Grammar

Masterprüfung in Linguistik integrativ (revised version)

von

Arne Ruhnau
Kontakt: arne.ruhnau@flir.com

Contents

1 Introduction **1**

2 The prefield in German declarative sentences **3**
2.1 The Topological Field Model and its Role in Syntactic Analyis 3
2.2 Syntactic candidates for the prefield 7
2.2.1 Singular elements in the prefield 8
2.2.2 Multiple elements in the prefield 13
2.2.3 Other candidates for the prefield 15

3 A short introduction to RRG **16**
3.1 Syntactic Representation 17
3.1.1 Consituent projection 18
3.1.2 Operator projection 21
3.1.3 Focus Projection 23
3.2 Semantic Representation 25
3.3 Linking Syntactic and Semantic Representation 28

4 Contemporary Role and Reference Grammar analyses of the prefield **30**
4.1 The original account of verb second languages 31
4.2 Prefield as obligatory precore slot: Diedrichsen (2008) 34

5 A non-configurational analysis of the prefield **38**
5.1 The prefield is an information unit 38
5.2 Word order constraints in simple sentences 44
5.2.1 Linearization constraints in German 46
5.2.2 Filling the prefield 54
5.3 The discontinuous properties of the Core 58
5.3.1 Partial 'VP' fronting 59
5.3.2 'tun'-Periphrasis 62
5.3.3 Matrix coding constructions 62
5.3.4 Multiple arguments in the prefield 65
5.4 On seemingly discontinuous NPs and the prefield 66
5.5 The prefield and the case of apparent subordination 73

6 Conclusion and Outlook **80**

1 Introduction

When an introductory course to linguistics at a German university arrives at the point where the syntactic structure of German is discussed, there are typically two statements made: First: The verb is in the second position in declarative sentences. Second: It is more economic to analyse German as a verb-final language. Both views suffer from the observation that the placement of the verb depends on the sentence type: yes/no questions put it first, *wh*-questions second, in subordinated clauses it is final, … One model attempts to formalize these topological regularities: the topological field model. Its basic terminology dates back to Drach (1937), and it has proven to be a good basis for a description of German syntax.

In this thesis, we examine one particular position of the topological field model, the prefield. Coming back to our first statement of verb-placement: Is it indeed true that the verb is in second position? If so, what is in the first position? When first confronted with tests for constituency, linguists learn that in German, the so-called "Vorfeldtest" can be used to determine whether a group of words is a constituent. Similarly, grammatical theories like the Minimalist Program or Head-Driven Phrase Structure Grammar state that the first position is a syntactic constituent of the clause.

In Role and Reference Grammar (RRG), the topology of German clauses has received more attention through the proposal of Diedrichsen (2008), which also equals the first position with a constituent. This thesis sets out to explore the problem domain of sentence topology in German. Our hypothesis is that the constituent-focused treatment of the prefield leads into the wrong direction, and cannot adequately account for observable word orderings. Especially, we take a look at the interplay of discourse information and sentence topology.

The problem with a constituent-focused account in RRG can be seen clearly in the following data:

(1) a. *Peter hat dem Mann das Buch gegeben*
peter.NOM has the man.DAT the book.ACC given
'Peter has given the man the book'

b. *Gegeben hat Peter dem Mann das Buch*

c. *Das Buch gegeben hat Peter dem Mann*

d. *Dem Mann gegeben hat Peter das Buch*

e. *Dem Mann das Buch gegeben hat Peter*

Which of the constituents postulated by RRG is in the prefield? What is the role that focus structure plays in these constructions?

This thesis starts by presenting the variation with respect to the prefield (section 2). As we will see, we can have one, more than one or, as it seems, only part of a constituent in this position. Afterwards, we present Role and Reference Grammar, focusing on the concepts relevant to our discussion (section 3). We then take a detailed look at the two ways RRG currently handles the prefield: First, there is the approach used in van Valin (2005) and van Valin & LaPolla (1997) (section 4.1). Second, we take a detailed look at the argumentation provided by Diedrichsen (2008) (section 4.2).

Section 5 is devoted to our proposed explanation for verb-second placement in German. As it will turn out, treating it as an information unit - seperate from the constituent projection of RRG - provides a lot of advantages. One advantage is that this step appears to integrate very well with contemporary work regarding word order, a topic which we will explore in section 5.2). Afterwards, we focus on three constructions in German: Discontinuous CORES (section 5.3), apparently discontinuous RPs (section 5.4) and sentence intertwining (section 5.5).

2 The prefield in German declarative sentences

In German, the usual word order in declarative sentences has the finite verb as the second constituent in the clause. Since Drach (1937), this observation has been captured in the so called topological field model, which provides us with labels for various coherent regions in the surface order of German sentences. In this section, we will present this descriptive model and explore the properties of the prefield in detail. The discussion is based mainly on Höhle (1986) and Pafel (2009).

2.1 The Topological Field Model and its Role in Syntactic Analyis

Drach (1937) introduced the term "Feld" (field) for three regions in German sentences. He identified three of these fields, namely the "Vorfeld" (prefield), "Mittelfeld" (middlefield) and the "Nachfeld" (postfield). Together, these fields can be used to describe regularities in the word order of German sentences.

From this basis, the topological field model was developed, which also describes the placement of verbal elements: In declarative sentences, the finite verb appears between the prefield and the middlefield (the so-called left bracket), whereas non-finite verbal elements are placed between the middlefield and the postfield (the so-called right bracket). We can also change our perspective on these terms: the two brackets would be our basic elements, and the position of the three fields are given relative to them.

(2) gives example sentences and their respective description[1] within the topological field model.

[1]The following abbreviations will be used to reference certain positions in the topological field model: COORD = coordination position, VVF = "Vorvorfeld", VF = prefield, LB = left bracket, MF = middlefield, RB = right bracket, NF = postfield

(2) a. *Peter hat geschlafen*
VF LB RB
Peter has slept
'Peter has slept'

b. *Peter hat Karl den Ball gestern gegeben*
VF LB MF RB
Peter has Karl the ball yesterday given
'Peter has given the ball to Karl yesterday'

c. *Peter hat Karl den Ball gegeben der blau war*
VF LB MF RB NF
Peter has Karl the ball given which blue was
'Peter has given Karl the ball that was blue'

These examples demonstrate the following properties of the topological field model:

1. Not all positions are filled in each sentence. (2 a) does not contain a middlefield even though both brackets are present.

2. Topological fields do not necessarily correlate with the constituent structure of a sentence. The middlefield in (2 b) contains both direct and indirect object as well as an adverbial, but excludes the predicate. This does not entail that these two form a single constituent.

3. The topological field model is recursive: A prefield can contain a complex clause. (2 c) contains a relative sentence in the postfield that itself can be divided into fields and braces (cf. Reis (1980)).

The slots defined by the topological field model also apply to subordinated clauses. In these, the left bracket is filled by a subordinating complementizer (3a), a relative phrase (3b), or an interrogative phrase (3c). Additionally, there are two constructions in German that involve a special marker (3d,e). In each of these examples, all verbal forms appear in the right bracket.

(3) a. *(Peter wollte wissen,)* ***ob*** *die Erde flach ist*
(Peter wanted know,) whether the earth flat ist
'Peter wanted to know whether the earth is flat

b. *(Peter liest das Buch,)* ***das*** *John Grisham geschrieben hat*
(Peter reads the book) that John Grisham written has
'Peter is reading the book John Grisham has written'

c. *(Peter wollte wissen,)* ***wer*** *das Buch geschrieben hat*
(Peter wanted know,) who the book written has
'Peter wanted to know who has written the book'

d. ***Je*** *mehr Geld ich verdiene, desto mehr kann ich kaufen*
the more money I earn, the more can I buy
'The more money I earn, the more things I can buy'

e. ***So*** *schön das Bild auch ist, ich kann es mir*
However beautiful the picture also is, I can it 1Sg.Dat
nicht leisten
not afford
'However beautiful the picture is, I cannot afford it'

Besides these five basic slots, more are required for a complete topological description of German. A look at the data in (4c-e) reveals that the right bracket has internal structure that defines the word order in the verb cluster. The right bracket itself can be separated into an "Oberfeld" and an "Unterfeld", depending on how verb form government "flows" through the cluster. In (4c-e), a verb V_n governs the form of V_{n+1}. If, following the linear order of verbs, n increases, the verbs belong to the "Oberfeld", else they form the "Unterfeld" (cf. Bech (1983)) (4e) is non-canonical ordering and stylistically marked, but nevertheless acceptable. In (4e), both "Oberfeld" and "Unterfeld" contain two verb forms. It is ungrammatical to have crossing dependencies between the verb forms, as (4f) demonstrates. Being based solely on the surface order of the German clause, the topological field model easily captures these regularities.

(4) a. *Peter wird das Haus sehen*
Peter will the house see
'Peter will see the house'

b. *Peter wird das Haus gesehen haben wollen*
Peter will the house seen have want
'Peter will have wanted to have seen the house'

c. *…dass der Fall erledigt worden sein muss*
V_4 V_3 V_2 V_1
…that the case closed been be must
'…that the case must have been closed'

d. *…dass der Fall muss erledigt worden sein*
V_1 V_4 V_3 V_2

e. *…dass der Fall muss sein erledigt worden*
V_1 V_2 V_4 V_3

f. * *dass der Fall muss erledigt sein worden*

Another point where the topological field model is extended revolves around left dislocation constructions in German. Left displaced elements are set of from the rest of the sentence by an intonation break, and typically introduce a new referent. In the remained of the sentence, a resumptive pronoun appears. Such phrases appear before the prefield, in the so-called "Vorvorfeld". Together with the reduced prosodic integration, the syntactic integration of the referents introduced through left disclocation varies. While in (5a) the displaced noun phrase bears the same case as the corresponding resumptive pronoun, nominative case as in (5b) is possible as well.

(5) a. *Den Apfel$_i$, den$_i$ habe ich gegessen*
the apple it have I eaten
'The apple, I ate it'

b. *Der Apfel$_i$, den$_i$ habe ich gegessen*

As the final extension to the topological field model, we need an additional slot which precedes left disclocated elements. This so-called coordination position can host, as the name implies, coordinating conjunctions. (6) is an example where all slots of the topological field model are filled.

(6) *…und* *dein Privatleben* *das* *hat* *sich doch auch* *verändert*
COORD VVF VF LK MF RK
nach dem Einbruch
NF
'…and your private life, it certainly has changed after the burglary'

How should we integrate these topological findings into a grammar model such as RRG? When working with the terms of the topological field model, one always has to be aware of two facts. First, this model is a mere *descriptive apparatus*. It does not strive to explain linguistic phenomena: Instead of defining what are correct and/or incorrect sentences, it only assigns a surface-based description. Questions like agreement, case assignment or how to represent control verbs are not part of the topological field model. Nevertheless, we can expect a reasonable linguistic theory to also come up with generalizations that allow equally elegant descriptions - and explanations. Second, this model is tailor-made to suite the German language. There are not many cross-linguistic insights to be gained from it. From the point of view of a linguistic theory, this might prove as a major hurdle. How can we reformulate the findings of the topological field model without introducing too much language-specific machinery into our grammar? Are there more general mechanisms of language at work that, by being composed correctly, explain the specific properties of German?

A case in point where this proves to be possible is the correspondence between the "Vorvorfeld" and the LDP in RRG. The question is, if there is an equally well suited element in RRG that corresponds to the prefield.

2.2 Syntactic candidates for the prefield

We will now take a closer look at the prefield in German sentences. What syntactic elements can appear there? At this point, this is a rather exploratory approach, and does not intend to be exhaustive. In the end, we will have come up with a list of elements that can appear in the prefield that has enough variation to prevent a too

rigid analysis later.

The presentation of the data follows a rather controversial categorization, since we aim to be as theory-neutral as possible. Given the available data, though, this means that we cannot follow the traditional (and too often unquestioned) thinking that the prefield consists of a single constituent. In order to not cloud our judgement, though, we must set this intuition aside. For the purpose of the present categorization, we therefore look only at the "first level" on constituents, meaning PPs, NPs, adverbs and "bare" verbs. We do not presuppose any theory of a 'VP' constituent. Of course, the syntactic dependencies between a verb and its arguments are obvious, but whether or not to capture these in terms of constituency must be left undecided for now. As a result, we categorize "verb-second" sentences in those cases where a single constituent appears before the finite verb, and those where there is more than one.

Another critical point for this data needs to be adressed: To a large amount, they are the result of introspection. Where this is not the case, the original source has been given. Finally, a large amount of the examples rely on rather specific discourse contexts, and some of them may even be considered ungrammatical by some speakers. For most of these cases, an analysis is presented later in this paper.

2.2.1 Singular elements in the prefield

First, let us examine cases with a single constituent in the prefield.

The prefield can be filled by **any one of the arguments** of the predicate[2]:

(7) a. *Der Mann kauft das Buch*
the man buys the book
'The man is buying the book'

b. *Das Buch kauft der Mann*

c. *Der Mann gibt der Frau die Blume*
the man gives the woman the flower

[2]Glosses are omitted in cases of mere linear reorderings of constituents.

'The man gives the flower to the woman'

d. *Der Frau gibt der Mann die Blume*

e. *Die Blume gibt der Mann der Frau*

The prefield can be filled by **an adverbial**, such as an adverb or a prepositional phrase:

(8) a. *Der Mann liest das Buch langsam*
the man reads the book slowly
'The man reads the book slowly'

b. *Langsam liest der Mann das Buch*

(9) *Seit geraumer Zeit liest der Mann das Buch*
for fairly long time reads the man the book
'For a fairly long time, the man is reading the book'

The prefield can be filled by **the predicate**. Since a finite verb has to appear in the left bracket, we need an auxiliary verb to free the predicate from this position. Such an auxiliary verb can be introduced by, e.g., a change of tense or voice of the clause. An adequate translation of such an example varies depending on the intonation of the original sentence:

(10) a. *Der Mann hat das Buch gelesen*
the man has read the book
'The man has read the book'

b. *Gelesen hat der Mann das Buch*
read has the man the book
'It was the man who has read the book' (if "der Mann" is stressed)

c. *Gelesen wurde das Buch*
read was the book
'It was the book that was read' (if "das Buch" is stressed)

The prefield can be filled by **a discourse connector**. Traditionally, such words are subsumed under the cover term *adverb*. However, their semantics are not limited to the clause they appear in, but relate two propositions to each other.

(11) a. *Diese Beobachtung ist allerdings unbrauchbar*
this observation is however useless
'This observation, however, is useless'

b. *Allerdings ist diese Beobachtung unbrauchbar*
however is this observation useless
'However, this observation is useless'

In *wh*-questions, **the question word** is in the prefield:

(12) a. *Wer gab der Frau die Blume*
who gave the woman the flower
'Who gave the flower to the woman?'

b. *Wem gab der Mann die Blume*
who gave the man the flower
'To whom did the man give the flower?'

c. *Was gab der Mann der Frau*
what gave the man the woman
'What did the man give to the woman?'

d. *Wann gab der Mann der Frau die Blume*
when gave the man the woman the flower
'When did the man give the flower to the woman?'

An expletive pronoun can be in the prefield. As reflected in the translation, there is no equivalent construction in English:

(13) a. *Es gab der Mann der Frau die Blumen*
it gave the man the woman the flowers
'The man gave the woman the flowers'

A subordinated clause can be in the prefield:

(14) a. *Nachdem der Regen aufgehört hatte kam der Sturm*
after the rain stopped had came the storm
'After the rain had stopped, the storm came'

b. *Der Sturm kam nachdem der Regen aufgehört hatte*

A separable **verb prefix** can be in the prefield. This always entails contrastive focus on the prefix:

(15) a. *An habe ich ihn gefahren, nicht um*
at have I him driven, not down
'I touched him (with my car), I did not knock him down'

How do these possibilities interact with matrix coding constructions? As it turns out, the prefield seems to be rather orthogonal to these issues. The examples in (16) demonstrate the behaviour for raising verbs by adding "scheinen" (seem). (17) shows the possibilities of control constructions, exemplified with "versuchen" (try).

In (18), raising-to-object data is presented with "sehen" (see). The prefield can be filled by either any of the arguments of the matrix verb, or by those of the embedded one. When an adverbial is in the prefield, the case becomes a little less clear. Depending on the scope of the adverbial, it can be interpreted as modifying the embedding predicate (henceforth "matrix reading"), or the embedded predicate (henceforth "embedded reading"). For the embedded reading, the prefield element needs to be stressed.

(16) a. *Der Mann schien das Buch langsam zu lesen*
the man seemed the book slowly to read
'The man seemed to read the book slowly'

b. *Das Buch schien der Mann langsam zu lesen*

c. *Langsam schien der Mann das Buch zu lesen*

d. *Zu lesen schien der Mann das Buch langsam*

e. *Es schien der Mann das Buch langsam zu lesen* cf. (13a)

f. *An schien ich ihn zu fahren, nicht um* cf. (15a)

(17) a. *Der Mann versuchte der Frau die Blume langsam zu geben*
the man tried the woman the flower slowly to give
'The man tried to slowly give the flower to the woman'

b. *Der Frau versuchte der Mann die Blume langsam zu geben*

c. *Die Blume versuchte der Mann der Frau langsam zu geben*

d. *Zu geben versuchte der Mann der Frau die Blume langsam*

e. *Langsam versuchte der Mann der Frau die Blume zu geben*

f. *Es versuchte der Mann der Frau die Blume langsam zu geben* cf. (13a)

g. *An versuchte ich ihn zu fahren, nicht um* cf. (15a)

(18) a. *Der Polizist sah den Mann der Frau die Blume geben*
the police officer saw the man the woman the flower give
'The police officer saw the man give the flower to the woman'

b. *Den Mann sah der Polizist der Frau die Blume geben*

c. *Der Frau sah der Polizist den Mann die Blume geben*

d. *Die Blume sah der Polizist den Mann der Frau geben*

e. *Langsam sah der Polizist den Mann der Frau die Blume geben*

f. *Es sah der Polizist den Mann der Frau die Blume geben* cf. (13a)

g. *An sah mich der Polizist ihn fahren, nicht um* cf. (15a)

The scope ambiguities of adverbials in the prefield also happen when **a subordinated clause** is in the prefield.

(19) a *Nachdem der Regen aufgehört hatte schien der Sturm zu*
After the rain stopped had seemed the storm to
kommen
come
'It was after the rain had stopped that the storm seemed to come' or
'After the rain had stopped, the storm seemed to come'

b *Nachdem der Regen aufgehört hatte versuchte der Mann nach*
After the rain stopped had tried the man to
Hause zu gehen
home to go
'It was after the rain had stopped that the man tried to go home' or
'After the rain stopped, the man tried to go home'

c *Nachdem der Regen aufgehört hatte sah der Polizist den*
After the rain stopped had saw the police officer the
Mann der Frau die Blume geben
man the woman the flower give
'It was after the rain had stopped that the police officer saw the man give the flower to the woman.' or 'After the raind had stopped, the police officer saw the man give the flower to the woman'

2.2.2 Multiple elements in the prefield

In the previous section, we listed possibilities for a single element in the prefield. There are, however, also cases for the prefield where more than one constituent appears there.

A **predicate and one or more of its arguments**, exluding the subject, can appear in the prefield. (20a) demonstrates the "basic" word order:

(20) a. *Der Mann hat der Frau die Blume gegeben*
the man has the woman the flower given
'The man has given the woman the flower'

b. *Die Blume gegeben hat der Mann der Frau*
the flower given has the man the woman
roughly: 'As for the flower, the man gave it to the woman'

c. *Der Frau gegeben hat der Mann die Blume*
the woman given has the man the flower
roughly: 'As for the woman, the man gave her the flower'

d. *Der Frau die Blume gegeben hat der Mann*
the woman the flower given has the man
roughly: 'What the man did was give the woman the flower'

e. *Die Blume der Frau gegeben hat der Mann*

The facts on a **predicate and the subject** (as well as more arguments) are less clear:

(21) a. *Das Auto kaputtgegangen ist*
DEF.N.SG.NOM car.SG break.down.PSTP be.3SG.PRES
John noch nie.
John(DAT) yet never
'It never happened to John that his car broke down.'

b. *Das Auto kaputtgegangen ist John*
the car break.down.PSTP be.3SG.PRES John(Dat)
'The car broke down on John'

c. **Das Auto kaputtgegangen ist*

d. *Autos repariert wurden in dieser*
Car.PL repair.PSTP become.3PL.PAST in DEM.F.SG.DAT
Firma schon viele
firm.SG already many
'Lots of cars have been repaired in this firm already.'

e. ?? *Autos repariert wurden in dieser Firma schon*

f. *Kinder gespielt haben hier noch nie.*
child.PL.NOM play.PSTP have.3PL.PRES here yet never
'It never happened here that children played.' (= Diedrichsen 2008, (24))

g. ?? *Kinder gespielt haben noch nie.*

h. *Kinder Fußball gespielt haben hier noch*
child.PL.NOM football play.PSTP have.3PL.PRES here yet
nie.
never
'It never happened here that children play football.' (= Diedrichsen 2008, (25a))

We can also add an adverbial to the predicate in the prefield:

(22) a. *Mit Hingabe gegeben hat der Mann der Frau die Blume*
with devotion given has the man the woman the flower
roughly: 'With devotion, the man has given the woman the flower'

b. *Der Frau mit Hingabe gegeben hat der Mann die Blume*

c. *Mit Hingabe der Frau gegeben hat der Mann die Blume*

We can even have an adverbial and an argument, or two arguments, but not the predicate, in the prefield (for these and other combinations, see Müller (2003), Müller (2005)).

(23) a. *Vermutlich vom gleichen Täter wurden zwei Tankstellen in*
Probably by the same offender were two gas stations in
Hemsbach und Heidelberg überfallen
Hemsbach and Heidelberg robbed
'Two gas stations in Hemsbach and Heidelberg were probably robbed by the same offender' (= Müller 2003, 7a)

b. *Zum zweiten Mal die Weltmeisterschaft errang Clark*
For the second time the world championship achieved Clard
1965…
1965
'Clark achieved the world chamionship for the second time in 1965' (= Müller 2003, 14e)

c. *Hauptberuflich als Anwalt hat er nur kurz gearbeitet.*
full-time as advocate has he only shortly worked
'He has worked full-time as an attorney only for a short time' (= Müller 2003, 42c)

d. *Nicht der Anna einen Brief hätte er schreiben sollen,*
Not the Anna a letter would have he write should
sondern der Ina eine Postkarte
but the Ina a post card
'He should have written not a letter to Anna, but a post card to Ina' (= Müller 2003, 62a)

e. ?? *Der Anna einen Brief hätte er schreiben sollen, nicht der Ina eine Postkarte*

2.2.3 Other candidates for the prefield

Our current distinction between singular and multiple constituents in the prefield is challenged by examples like (24).

(24) a. *Syntaktiker kenne ich drei/viele/keine*
syntactitians know I three/many/none
'As for syntacticians, I know three/many/none'

b. *Polizeiwagen kenne ich nur grüne*
police cars know I only green
'As for police cars, I only know green ones'

This construction is commonly known as "split topicalization", or - at least for (24a) - as "floating quantifiers". Superficially, an argument of the predicate seems to be realized by two distinct constituents. One of them appears in the prefield, the other as the last element in the middlefield. In addition to quantifiers and attributive adjectives, this last element can also be an NP[3]:

[3]Thanks to Prof. van Valin, who pointed these examples out to me (p.c.).

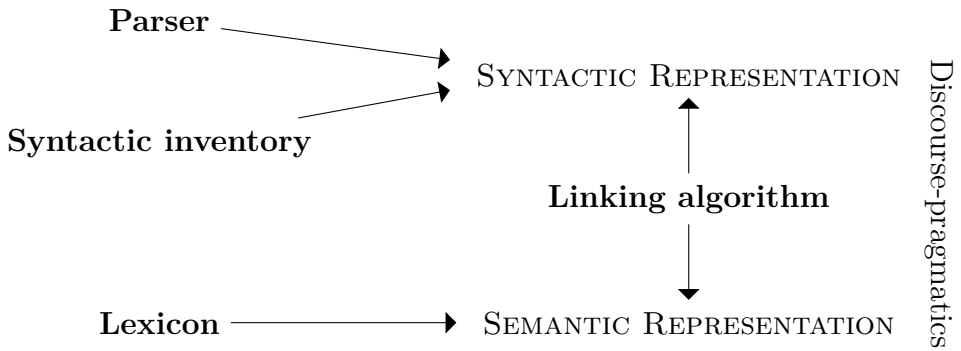

Figure 1: The architecture of RRG

(25) a. *Syntaktiker kenne ich nur Chomsky*
syntacticians know I only Chomsky
'As for syntacticians, I only know Chomsky'

b. *Syntaktiker kenne ich nur Chomsky, Sternefeld und Uszkoreit*

We have to answer two questions to account for this data: 1) What is the syntactic status of the NP in the prefield? 2) What is the relation between this NP and the floating element?

For now, we leave it at asking these questions; section 5.4 presents an RRG account of these constructions.

3 A short introduction to RRG

Role and Reference Grammar (RRG) is a non-generative grammar theory that aims for cross-linguistics applicability and generality. The most recent version of the theory is presented in van Valin (2005), which builds upon the version described in van Valin & LaPolla (1997). What sets RRG apart from other contemporary grammar theories is the explicit use of a linking algorithm - a finite list of steps that associate the syntactic and the semantic representation of a sentence. The overall architecture of the theory can be visualized as in figure 1.

As the figure shows, RRG is composed of several modules that target a specific aspect of language: The semantic representation, or logical structure, represents the lexical meaning of a sentence - the source of this representation is the lexicon. The syntactic representation represents the surface form of a sentence. In order to include the "generation" as well as the "understanding" parts of human language processing, the syntactic representation is either created by a parser, or it is constructed by combining elements from the syntactic inventory. The bidirectionality of language processing is also the core feature of the linking algorithm - it is applied in either of the two directions. The linking algorithm can be influenced, language-specifically, by constructional templates.

In RRG, a prime use of the linking algorithm is to handle displacement scenarios. Since RRG denys the necessity of phonetically null elements (besides zero-marking on the morphological level), such constructions need to be accounted for by another mechanism. The RRG analysis of, e.g. *wh*-extraction, combines syntactic and pragmatic properties to link syntax to semantics.

As van Valin (2005) states in the introduction, the two motivating questions for RRG are

1. What would linguistic theory look like if it were based on the analysis of languages with diverse structures such as Lakhota, Tagalog and Dyirbal, rather than on the analysis of English?

2. How can the interaction of syntax, semantics and pragmatics in different grammatical systems best be captured and explained?

The following sections present the modules of RRG in more detail.

3.1 Syntactic Representation

The syntactic representation of a sentence in RRG is tied very closely to its surface form. By comparision with, e.g., the minimalist program, this has rather drastic

consequences: There are no phonetically null, or "abstract" syntactic elements. RRG decomposes the surface form of a sentence into three seperate projections that are more or less tightly coupled to each other. The common denominator for each of these projections are the actual words of an utterance. These three projections are:

1. The constituent projection, to represent the hierachical organization of constituents

2. The operator projection, to represent dependencies with respect to grammatical features

3. The focus structure projection, to represent the influence of information structural concepts on the sentence

3.1.1 Consituent projection

The constituent projection is the syntactic counterpart of the semantic representation. It is concerned with the syntactic realization of predicates and their arguments, and it is on this projection where part-of-speech and similar notions are relevant. The units that make up this projection are nevertheless semantically motivated. The most prominent element is the syntactic counterpart of the semantic predicate, termed nucleus (Nuc). Not all nuclei predicate, and to account for this fact, a predicating nucleus dominates a Pred element in the constituent projection. The nucleus projects into the Core, which additionally contains the arguments of the predicate. Above the core is the Clause, which hosts the Core as well as non-arguments (termed the Periphery).

The crucial next step now is the insight that the above correspondences are but prototypical. It is possible (and sometimes necessary) to realize an argument *not* in the Core but outside of it. For the concept of the Periphery, the case is quite similar: RRG does not use the Periphery only at the Core level. Instead, the Nuc, the Core and the Clause all can have peripheral modifiers.

The units discussed so far apply in each language. RRG also provides us with four syntactic units that are not present in each language. On the CLAUSE level, these are the precore slot (PRCS) as well as the postcore slot (POCS). These two slots are used for displaced elements that are still integrated in the CLAUSE, such as *wh*-question words in English (which appear in the PRCS, cf. (26) or focused arguments in Dhivehi (which, if displaced, appear in the POCS, as 27) shows):

(26) *What did you eat today?*

(27) a. *Alī bunī* KĪKE *ta?*
Ali say.PAST.FOC what Q
'What did Ali say?' (= van Valin 2005, (1.3a))

b. *Alī kīke bunī ta?*
Ali say.PAST.FOC what Q
'What did Ali say?' (= van Valin 2005, (1.3b))

The other two language-dependant units are the left-detached position (LDP) and the right-detached position (RDP). As the name already implies, these positions are used for material that is not integrated in the CLAUSE. German has both of these additional positions, as (28) shows. The crucial distinction between a detached position and a pre- or postcore slot is that the detached positions are set off by an intonation break from the rest of the CLAUSE.

(28) a. *Den Film, den habe ich nicht gesehen*
the movie.ACC, that.ACC have I not seen
'That movie, I haven't seen it'

b. *Ich habe ihn nicht gesehen, den Film*

Figure (2) shows the general form of the constituent projection in RRG, with the optional PERIPHERY left out.

The association of a PERIPHERY to its layer is shown in figure 3. We have avoided to specify any nodes below the PRED on purpose. RRG explicitly claims that there is not a correlation between specific parts of speech, their semantic function and their

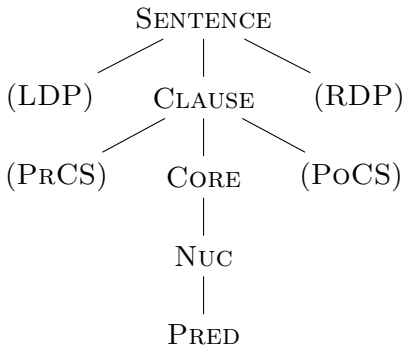

Figure 2: The constituent projection of RRG

syntactic potential. For a predicating nucleus, we typically have a verb, but this is not necessarily always the case. In a sentence like (29), the syntactic predicate is the NP "a violent vigilante". We can explain the occurence of the copula by assuming that English does not allow predicating nuclei without a verbal element.

(29) Batman is a violent vigilante

The category-agnostic nature of the layered structure of the clause carries over to the representation of arguments. Here, RRG provides us with another cross-linguistically valid syntactic category: the referential phrase (RP, van Valin (2008)). The RP provides a syntactic abstraction over language-specific parts of speech, and by the same time it removes the need for an endocentric category for referential arguments. For a language like English, an RP typically corresponds to a noun phrase. However, other lexical categories can fulfill the role of an RP as well. An example from German is given in (30)

(30) *Der* *Lange* *ist* *eingeschlafen.*
the.M.SG.NOM tall be.3SG.PRES fall asleep.PASTPART
'The tall one has fallen asleep.' (= van Valin 2008, (4))

For nominal phrases, RRG assumes a layered structure as well. Why should such a representation be necessary? Consider the noun phrase in (31):

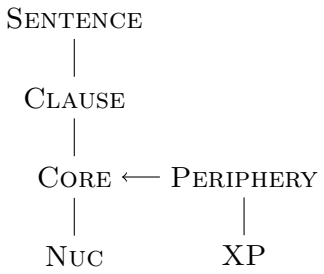

Figure 3: Representing a Periphery_{Core} in RRG

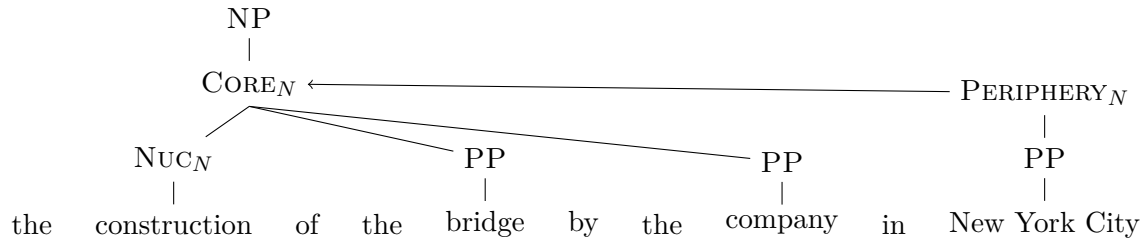

Figure 4: The layered structure of the NP in English

(31) the construction of the bridge by the company in New York City (taken from van Valin 2005, figure 1.18)

The noun *construction* takes two arguments: *bridge* and *company*. Additionally, there is nominal periphery for the locative prepositional phrase. From this, it is only natural to deduce that a noun projects into a nucleus and into a core. The nominal core hosts the arguments and is modified by the PP in a Periphery_N. A graphical representation of these relations is given in figure 4 (again taken from van Valin 2005, figure 1.18).

3.1.2 Operator projection

If the constituent projection represents the dependencies between the content words of a language, the operator projection represents the function words. The operator projection is a mirror image of the constituent projection, meaning that it follows the layered structure of the clause. In RRG, operators are essentially modifiers, and

therefore not part of the constituents involved.

RRG formulates a universal principle with respect to operators: The proximity to the nucleus is iconic of the scope of the operator. For example, if there is both a nuclear and a core operator at the right of the nucleus, RRG predicts that the nuclear operator is always closer to the nucleus than the core operator. A more graphical version of this principle is that there are no crossing branches on the operator projection.

This scope hierarchy results in the following inventory of operators in RRG:

Nuclear Operators Aspect, negation and directionals modifying the orientation of an event

Core Operators Directionals modifiying the orientation of motion of a participant, event quantification, modality and narrow scope negation

Clausal Operators Status, tense, evidentials and illocutionary force

This inventory of operators is not claimed to be universal: not every language has every operator. Every language uses the *concept* expressable by these operators, but the decisive criterion is whether a given has developed a grammatical category for it.

Since the constituent projection also proposes a layered structure for noun phrases, the operator projection also defines a set of noun-related operators. For the nominal nucleus, this is "nominal aspect", i.e. classifiers or count-mass distinction. The $\textsc{Core}_N$ can be modified by number, quantifiers and negation, whereas on the top-level NP, we find definitess and deixis. As was the case for "verbal" operators, the operators in the layered structure of the NP are ranked: deixis is the outmost operator, and nominal aspect is the closest.

3.1.3 Focus Projection

As a third projection, RRG represents the information structure of a sentence in the focus projection. Unlike the operator projection, the focus structure projection is rather losely coupled with the constituent projection. In fact, there is - apart from the actual words of the sentence - only one part of the other projections that is relevant for the focus structure projection: The CLAUSE node that is modified by the illocutionary force operator.

We will return to the role this operator plays later. First, we shall present the relevant terminology used by RRG with respect to information structure. The terms used by RRG are based on the work of Lambrecht (1994), who introduces them as follows[4]:

> PRAGMATIC PRESUPPOSITION: The set of propositions lexicogrammatically evoked in an utterance which the speaker assumes the hearer already knows or believes or is ready to take for granted at the time of speech. (52)
>
> PRAGMATIC ASSERTION: The proposition expressed by a sentence which the hearer is expected to know or believe or take for granted as a result of hearing the sentence uttered. (52)
>
> FOCUS, or FOCUS OF THE ASSERTION: The semantic comopnent of a pragmatically structured proposition whereby the assertion differs from the presupposition. (213)
>
> FOCUS STRUCTURE: The conventional association of a focus meaning with a sentence form. (222)

Based on these notions, one can derive several well-defined *focus types*. The first criterion one can check is whether there is one or more constituent in focus (or in the *focus domain*). If it is one constituent, it is a case of *narrow focus*. For multiple

[4]The numbers behind the definitions refer to the respective page in Lambrecht (1994).

constituents, the term *broad focus* is used. Lambrecht (2000) provides us with two types of broad focus: *predicate focus*, which follows a topic - comment scheme, and *sentence focus*, which lacks any presupposed topic.

The following English examples (taken from van Valin 2005, p. 70f) show, in order, narrow focus, predicate focus and sentence focus:

(32) Q: I heard your motorcycle broke down?

A: No, my CAR broke down

(33) Q: How's your car?

A: My car / it broke DOWN

(34) Q: What happened?

A: My CAR broke down

For each of these three focus types, we can now decompose an utterance in context into the presupposition and the assertion. From there, we can derive the focused information, and relate it to the elements of the sentence that constitute the focus domain.

The last piece of information needed for our purposes is the distinction between the actual and the potential focus domain in RRG. The potential focus domain, as the name implies, defines where a focused constituent in the CLAUSE *can* occur and is a syntactic notion. Naturally, the actual focus domain is that part of the sentence that is in focus.

Visually, the potential and actual focus domains are a third projection from the surface form of the sentence. We deviate slightly from the way van Valin (2005) represents focus structure, as figure 5 shows. Also, in cases where all three projections are to be represented, this thesis blends the operator projection and the focus structure projection. This is merely an aesthetic preference.

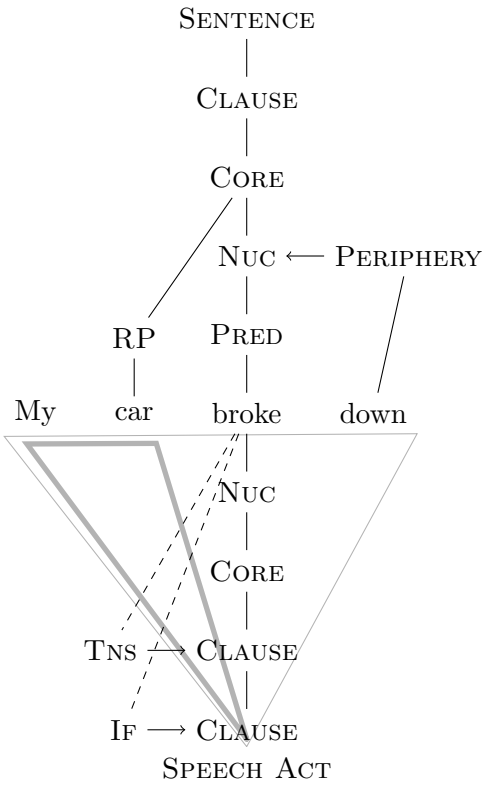

Figure 5: The graphical representation of focus structure

3.2 Semantic Representation

The previously discussed projections dealt with the surface form of a sentence. RRG assumes a distinct level of representation that defines a sentence's semantic content: the logical structure (LS). LS basically distinguishes three types of expressions: predicates, arguments, and semantic operators. Additionally, one can specify the semantic counterpart of syntactic operators. Since the remainder of this thesis is mainly concerned with the interactions between the syntactic projections, the discussion of the semantic representation will be rather coarse. For details on the subjects presented here, see e.g. van Valin (2005), ch. 2, which also forms the basis for the current section.

While the motivation for predicates and arguments is similar to, e.g., propositional logic, the semantic operators argued for by RRG need further justification.

At the heart of the semantic representation of a state of affairs lies the *Aktionsart* classification initially proposed by Vendler (1967). Extending his set of verb classes, RRG associates each *Aktionsart* with a basic building pattern for a logical structure (cf. van Valin 2005, table 2.3). The two basic blocks are *state* and *activity*:

STATE	**predicate'**(x, (y))
ACTIVITY	**do'**(x, [**predicate'**(x, (y))])

The predicate appears in boldface, and is always part of a semantic metalanguage, not of the language discussed. The argument slots, though, are filled by the actual linguistic expression used in the sentence at hand. The paranthesis around the second argument of the predicate indicate optionality, i.e. a state can have either one or two arguments.

These two patterns do not have any semantic operator but only consist of one or two logical predicates as well as one or two arguments. A more complex pattern is used for *achievement*, *semelfactive* and *accomplishment* verb classes:

ACHIEVEMENT	INGR {state or activity pattern}
SEMELFACTIVE	SEML {state or activity pattern}
ACCOMPLISHMENT	BECOME {state or activity pattern}

The added uppercase elements are the aforementioned semantic operators and express the inherent semantic contribution of the *aktionsart*. The two last patterns for LS express *active accomplishment* and causality. As we can see, active accomplishments can have either two or three participants (denoted by x, y and z).

ACTIVE ACCOMPLISHMENT	**do'**(x, [**predicate**$_1$**'**(x, (y))])
	& INGR **predicate'**$_2$(z,x) or (y)
CAUSATIVE	α CAUSE β where α and β are
	logical structures of any type

With such a classification scheme at hand, we need a way to decide, for a verb, which *aktionsart* it belongs to. First, each *aktionsart* can be decomposed into four

binary features: [± static], [± dynamic], [±telic] and [±punctual]. A state, for example, describes a static, atelic non-dynamic und not punctual state of affairs, while an active accomplishment is not static, dynamic, telic and not punctual. A complete list of these decompositions is given in van Valin (2005), (2.4). An additional semantic operator is DO, which represents lexicalized agency.

Additionally, there are a number of language-specific tests that one can employ to check the value of any of the four features in a given context. For English, one can for example exclude *activity* as an *aktionsart* if we cannot add "in an hour" to the sentence, as example (35) shows:

(35) a. The soldiers marched in the park

b. #The soldiers marched in the park in an hour

The next concept that applies to LS is that of semantic macroroles, which are a central component of RRG. A semantic macrorole is an abstraction over situation-dependent thematic relations: *Actor* hides the differentiation of agent, experiencer or instrument, while *Undergoer* encompasses patient, theme or recipient. Such a layer of abstraction proves to be useful to describe a voice-construction like passive in semantic terms: The actor will be optionally realized as an obliquely marked RP, and the undergoer is promoted to the subject of the sentence. RRG goes one step further and states that, beyond descriptive simplification, the semantic macrorole of an argument can, to some extent, be *predicted* by its place in the logical structure along the actor-undergoer hierarchy:

ACTOR				UNDERGOER
arg. of DO	1st arg of **do'**	1st arg. of **pred'**(x,y)	2nd arg. or **pred'**(x,y)	arg. of **pred'**(x)

Once assigned, macroroles are used as the explanatory basis for various linguistic phenomena. First, they are the basis for the notion of a priviledged syntactic argument (PSA) of a construction. In contrast to other contemporary theories, RRG does not posit any grammatical relation, such as "subject" to be universal. Instead,

such relations are cases where a differentiation between two semantic roles is neutralized in a given context. Hence, a grammatical relation is also referred to as a *restricted neutralization* (cf. van Valin 2005, ch. 4.2).

As a voice-construction such as passive shows, a speaker can choose which macrorole should be the PSA. In an accusative system, the default choice is the actor (resulting in active voice), and the undergoer is the marked choice. RRG generalizes contrasts such as this into the priviledged syntactic argument selection hierarchy (= van Valin 2005, (4.14)):

arg. of DO > 1st arg. of **do'** > 1st arg. of **pred'**(x,y) > 2nd arg. of **pred'**(x,y) > arg. of **pred'**(x)

This hierarchy is used to determine case assignment, as the following general rules demonstrate (= van Valin 2005, (4.25, 4.26)):

1. Case assignment rules for accusative constructions:
   (a) Assign nominative case to the highest ranking macrorole argument.
   (b) Assign accusative case to the other macrorole argument.
2. Case assignment rules for ergative constructions:
   (a) Assign absolutive case to the lowest ranking macrorole argument
   (b) Assign ergative case to the other macrorole argument

The choice of the PSA in a given construction, as well as the cas marking principles, are two steps in the linking algorithm of RRG, which will be discussed in the following section.

3.3 Linking Syntactic and Semantic Representation

Now that we have taken a look at both the surface form and its semantic counterpart, we turn to the linking algorithm of RRG. The task for the linking algorithm is to associate the semantic elements to the syntactic ones, *and vice versa*. This means

that this algorithm provides a bidirectional process to generate one representation out of the other. We do not present the linking algorithm in full detail, but only outline its basic conceptual steps.

A cornerstone of the linking algorithm is the *completeness constraint* (= van Valin 2005, (5.1)). This constraint ensures that no syntactic arguments are left out from LS, and in turn that every argument on LS receives an overt realization:

> All of the arguments explicitly specified in the semantic representation of a sentence must be realized syntactically in the sentence, and all of the referring expressions in the syntactic representation of a sentence must be linked to an argument position in a logical structure in the semantic representation of the sentence.

While the semantic representation is generated in the lexicon, the syntactic structure is composed of several syntactic templates. RRG proposes one universal principle to select a syntactic template (= van Valin 2005, (5.2)):

> **Syntactic template selection principle**
>
> The number of syntactic slots for arguments and argument-adjuncts within the core is equal to the number of distinc specified argument positions in the semantic representation of the core.
>
> **Language-specific qualifications**
>
> 1. All cores in the language have a minimum syntactic valence of 1
> 2. Argument-modulation voice constructions reduce the number of core slots by 1
> 3. The occurence of a syntactic argument in the pre/postcore slot reduces the number of core slots by 1 (may override (1) above)

For the linking from semantics to syntax, the abbreviated steps of the linking algorithm are as follows (for a detailed representation of the algorithm, see van Valin 2005, (5.5)):

1. Construct the semantic representation of the sentence, based on the logical structure of the predicator.
2. Determine the actor and undergoer assignments following the actor-undergoer-hierarchy.
3. Determine the morphosyntactic coding of the arguments (i.e. PSA selection, case assignment and agreement).

4. Select the syntactic template(s) for the sentence.

5. Assign arguments to the positions in the syntactic representations of the sentence.

When linking from syntax to semantics, the algorithm has to infer the macroroles and core arguments from the surface form. The detailed algorithm can be found in van Valin (2005), (5.10). Here, we only give an abbreviated form:

1. Determine the macrorole(s) and other core argument(s) in the clause (depending on transitivity, voice and other factors).

2. Retrieve from the lexicon the logical structure of the predicate in the nucleus of the clause, and execute step 2 from the linking of semantics to syntax (subject to constraints).

3. Link the arguments determined in step 1 with the arguments in step 2 until all core arguments are linked.

4. Predicative adpositional adjuncts are inserted into the LS. The core becomes the second argument, and the object of the adposition becomes the first argument.

5. Handle, if any, elements in the pre- or postcore slot by linking them either to the free argument slot in LS, or treat it as a predicative apposition (see step 4).

The important part in the linking from syntax to semantics is that once the surface form has been encoded, we can apply the same linking rules that are used when linking from semantics to syntax. With respect to the precore slot, we find both directions contain an explicit step for this position. When targeting syntax, step 5 decides whether an argument is placed there. If we link to the semantics, step 5 is sensitive to an element in this slot.

4 Contemporary Role and Reference Grammar analyses of the prefield

In this section, we outline the analysis of verb second languages like present in RRG. We start the discussion with the information contained in both van Valin & LaPolla (1997) and van Valin (2005), and then present the proposal of Diedrichsen (2008). For the first two, due to their nature of being not language but theory

specific, there is no explicitly formulated "V2 hypothesis". However, the discussion of various Icelandic and German data is consistent with one another and draws a good picture of the actual underlying hypothesis.

4.1 The original account of verb second languages

van Valin (2005) handles some of the data we presented in section 2 explicitly as cases of a PrCS, namely:

- *wh*-questions
- Clause-initial non-subject NPs
- Clause-initial subordinated clauses

The data to support the first two items is repeated in (36) (= van Valin 2005, 5.6a-c):

(36) a. *Der* *Mann* *ha-t* *der* *Frau*
the.MsgNOM man have-3gPRES the.FsgDAT woman
den *Hut* *geschenkt.*
the.MsgACC hat give.PSTP
'The man gave the hat to the woman [as a gift].'

b. *We-m* *ha-t* *der* *Mann* *den* *Hut*
who-DAT have-3sgPRES the.MsgNOM man the.MsgACC hat
geschenkt?
give.PSTP
'To whom did the man give the hat?'

c. *Den* *Hut* *ha-t* *der* *Mann* *der*
the.MsgACC hat have-3sgPRES the.MsgNOM man the.FsgDAT
Frau *geschenkt.*
woman give.PSTP
'The hat the man gave to the woman.'

The *wh*-word in (36b) and the accusative NP in (36c) are linked directly to the precore slot. While the question formation is uncontroversial and follows directly

from the RRG approach to displaced *wh*-words, there are two open questions to note:

1. The precore slot linking of the accusative NP lacks, at this point, argumentative support

2. There is no mention of the linking of the nominative NP in (36a).

Given the similarity of word order flexibility between German and Icelandic, we can find clues on how to answer these questions in (van Valin & LaPolla 1997, p. 38). For expository purposes, the data is repeated in (37)[5]

(37) a. Henni ***hef-ur*** *alltaf þótt* *Ólaf-ur*
3FsgDAT have-3sgPRES always think.PSTP Olaf-MsgNOM
leiðinleg-ur.
boring-MsgNOM
'She has always considered Olaf boring'

b. *Ólafur* ***hefur*** henni *alltaf þótt leiðinlegur.*

c. **Ólafur* henni ***hefur*** *alltaf þótt leiðinlegur.*

d. Hún ***haf-ð-i*** *unn-ið* *að* *brúarsmíði* *í*
3FsgNOM have-PAST-3sg work-PSTP at bridge.building in
sumar.
summer
'She worked at bridge-building in the summer.'

e. *Í sumar* ***hafði*** hún *unnið að brúarsmíði.*

'In the summer she worked at bridge-building.'

f. **Í sumar* hún ***hafði*** *unnið að brúarsmíði.*

g. *Hvenær* ***haf-ð-i*** hún *unn-ið* *að* *brúarsmíði?*
when have-PAST-3sg 3FsgNOM work-PSTP at bridge.building
'When did she work at bridge-building?'

h. **Hvenær* hún ***hafði*** *unnið að brúarsmíði?*

[5]The data in these examples is originally taken from Maling & Zaenen (1981). Boldface marks the finite verb, normal font indicates the subject.

The discussion of this data is based a) on the fact that in Icelandic, the finite verb must be in second position in the clause, b) on the fact that Icelandic has a rather free constiuent order and c) on the hypothesis that Icelandic has a default position for subjects before the finite verb.

Now, "[w]hen a non-subject phrase appears in initial position, […] the finite verb must immediately follow it, and the subject can no longer appear in its default position […]." (van Valin & LaPolla 1997, p. 39). Why is this consistent with a precore slot linking of non-subject NPs? The assumptions behind this argumentation seems to be that a) the default subject position is *not* the precore slot and b) that there are in fact two positions before the finite verb: One is core internal and reserved for the subject, and the other one is the precore slot[6]. In Icelandic, these positions are mutually exclusive due to the strict verb second constraint. In English, however, both can be filled at the same time:

(38) *Star Wars I have never seen in my life*

Having clarified the assumptions behind the linking of non-subject NPs in Icelandic, it is evident that German data has been analysed on the same basis. Therefore:

1. The accusative NP in (36c) is linked to the PrCS because it is not the subject

2. The nominative NP in (36a) is linked to the Core

Let us now consider the case of clause-initial subordinated clauses:

(39) *Wenn es regn-et, geh-e ich nicht raus.*
if 3Nsg rain-3sgNPSG, go-1sgNPST 1sgNOM not out
'If it rains, I don't go out' (= van Valin 2005, 6.20, p. 193)

In the original analysis, the clause-initial adverbial occupies the PrCS. This is based on the same assumptions as above: Since the adverbial is not the subject of

[6]The same structural model is assumed for Danish by Jensen (2000). She does not qualify this assumption in any way, though.

the clause, it cannot occupy the preverbal core-internal subject position. The only other candidate position that appears before the finite verb is the PrCS. The verb-second requirement then makes the subject appear after the finite verb. Changing our angle on the problem, "[...] the finite verb and privileged syntactic argument are inverted, signalling that the initial conditional clause is in the precore slot." (van Valin 2005, p. 193).

To sum up, the original RRG analysis of German clause structure is: The position before the finite verb is core internal if and only if the (non-*wh*) subject occupies it. In any other case, the constituent under question is placed in the precore slot.

4.2 Prefield as obligatory precore slot: Diedrichsen (2008)

Diedrichsen (2008) argues for the prefield to be identical to the precore slot[7]. Therefore, the subject is not treated differently from other constituents in the constituent projection. As a side-effect, the precore slot loses pragmatic specificity and reduces to a purely configurational position.

Diedrichsen gives two supportive arguments:

1. The semantics of epistemic modal verbs require a precore slot
2. The linking algorithm can be simplified by having an explicitly labeled position before the finite verb to link to

There are two reasons as to why the following review goes into more detail, compared to the previous section:First, there are actual arguments for the proposal available. Second, the article discusses the same problem as this thesis. Irrespective of the quality of the arguments we find in the article, the general thrust of the hypothesis is indeed appealing, since it reduces the complexity inherent in the original verb second analysis.

[7]The idea of an obligatory precore slot was applied in van Valin & Diedrichsen (2006) and originates, as noted there, from personal communication with Ina Bornkessel and Matthias Schlesewsky.

Consider the data in (40)[8]:

(40) *John* *muss* *das* *Auto* *waschen*
John must.3sg.PRES DEF.N.SG.ACC car.SG wash.INF
'John has the obligation to wash the car' (deontic)
'There is some obligation/strong reason to assume that John is washing the car.' (epistemic)

Diedrichsen analyzes the deontic reading as a core operator, while the epistemic modal is a clause operator due to its increased scope. Additionally, she states that "[w]hile with a deontic reading, the modal verb seems to point to the right and modify the action that is stated in the non-finite verb, with an epistemic meaning the modal verb rather points to the left, where the Vorfeld-element is located." (Diedrichsen 2008, p. 208).

As additional evidence, she shows that specific temporal adverbials in the prefield can exclude either a deontic or epistemic interpretation of the modal verb:

(41) a. *Morgen* *muss* *John* *das* *Auto* *gewaschen* *haben*
tomorrow must.3sg John DEF.N.SG.ACC car.SG wash.PSTP have.INF
'John must have washed the car tomorrow' (= Diedrichsen 2008, (12))

b. *Gestern* *muss* *John* *das* *Auto* *gewaschen* *haben*
yesterday must.3sg John DEF.N.SG.ACC car.SG wash.PSTP have.INF
'John must have washed the car yesterday.' (= Diedrichsen 2008, (13))

In (41a), an epistemic reading is not possible, (41b) can only have an epistemic reading. The reason for this is a semantic incompatibility: "Speakers cannot make an assumption about the truth of something which is to take place in the future. The deontic reading is acceptable, however, because it is possible that a speaker know about something that another one has to do in the future." (Diedrichsen 2008,

[8]This example is an abbreviated representation of example (10) in Diedrichsen (2008).

p. 209). Likewise, in (41b) "[i]t is not possible to talk about the obligation that somebody else has, referring to a past day." (ibid.).

The problem with this argumentation is that it is not compelling. The scope differences between epistemic and deontic readings are evident and justify an adequate representation as clausal/core operators. But: Diedrichsen goes one step further and claims that the semantic representation of the prefield element is more prominent than that of the remainder of the clause. This statement refers to the concept of "aboutness", a notion relevant on the focus structure projection. There are three points for objection here:

First, whatever insights semantic scope reveals, they are first and foremost represented in the logical structure of a sentence. The constituent projection is not required (or able) to be a mirror image of it. Displaced *wh*-question words are one of many examples of constructions with a clear mismatch between constituent structure and semantic scope. Second, the notion of topicality, or "aboutness" need not be represented configurationally, and is explicitly not part of the logical structure. Third, a semantic incompatibility need not be explained by the topological position of the adverbial. Quite the opposite is true: In the case at hand, we observe the same effect when the adverbial in question is not in the prefield at all. (42a) cannot have a deontic reading, and in (42b) the epistemic reading is blocked:

(42) a. *John muss gestern das Auto gewaschen haben*

b. *John muss morgen das Auto gewaschen haben*

Differences between epistemic and deontic readings do not convincingly argue for an obligatory precore slot. Let us therefore take a look at the second argument of Diedrichsen (2008): a simplification of the linking algorithm.

The linking algorithm is responsible to place verbal elements into the correct slot of the CORE. Under this assumption, an obligatory precore slot makes this step easy to formulate. The relevant steps from (Diedrichsen 2008, p. 211) are:

1. If the nucleus is finite, assign it to the first position in the core.

2. If the nucleus is non-finite, assign it to the last position in the core (default) or the precore slot (subject to focus structure restrictions).

3. If the nucleus is non-finite, place the finite auxiliary before the first slot in the core; non-finite auxiliaries are placed after the nucleus. If the nucleus is in the PrCS, non-finite auxiliaries can either be placed adjacent to it or after last position in the core, but need not be adjacent to each other.

This correctly positions the verbal elements in sentences like (43):

(43) *Gewaschen* *haben* *muss* *John* *das* *Auto*
wash.PSTP have.INF must.2SG.PRES John DEF.N.SG.ACC car.SG
gestern
yesterday
'John must have washed the car yesterday.' (= Diedrichsen 2008, (17))

Without the precore slot, however, positioning the verbal elements in the prefield becomes tedious. How should the position of the finite verb "muss" be specified? If it is "after the first element in the core", this contradicts (43). To compensate, an exceptional rule needs to be formulated that is sensitive for this particular construction.

At this point, the sole argument for an obligatory precore slot is theory internal. The linguistic evidence reviewed was inconclusive. Still, the topological field model challenges RRG to come up with a proper treatment of the prefield. The readily available notion of a precore slot is tempting, but seems to be hard to argue for.

In the next section, we will propose an explicitly non-configurational account for the prefield in German. Since this proposal runs against the core assumption of Diedrichsen (2008), we will contrast the two approaches where appropriate.

5 A non-configurational analysis of the prefield

Having presented two possible models for the prefield in RRG, no convincing candidate has been found. Treating clause-initial subjects different from other elements fails to capture the generality of the topological facts in German clauses. Additionally, there are no arguments for this dichotomy. As we stated, in a language like English, both a fixed subject position and an independant position before it is needed to account for topicalization constructions, the same is not true in German. We discussed one proposal that seeks to capture the generality of the prefield constraint, but resorting to an obligatory configurational position lacked convincing linguistic evidence, and we are left with theory-internal advantages.

The rest of this section is structured as follows: Section 5.1 gives a short overview over the analyses proposed in different grammatical frameworks. From these attempts, combined with the analytical devices, we formulate a an abstract template for the German clause that defines the prefield as an information unit. As this approach requires both CORE-internal linearization as well as a decision which elements appear in the prefield, section 5.2 explores these problem domains. The remainings sections are devoted to linguistic evidence in support of the proposed interpretation of the topology of German. Section 5.3 shows how we can account for various CORE related constructions. Going up one layer, section 5.4 shows that some constructions in the German clause indeed justify a PRCS. Finally, 5.5 presents an analysis of a complex sentence construction involving two CLAUSES.

5.1 The prefield is an information unit

Accounting for the topological facts of German has resulted in basically two types of solutions. The first type is a strictly configurational approach, proposed in Generative Grammar and for HPSG: In Generative Grammar, there is no other means than to use the configurational structure created through the course of the derivation to

arrive at the correct surface structure. The initial analysis given by den Besten (1983) decomposes V2 into two movement operations[9]. *V* moves to C^0 (yielding a verb-initial structure) and another constituent moves to [Spec,CP]. The elements at the left edge of the clause are preposed "by a rule which is similar in effect to Wh-Movement" (den Besten 1983, p. 55). In cases where it is not the subject that moves to [Spec,CP], den Besten, this movement is "uncontroversial" (ibid.) topicalization. The superficial similarity of verb-second in both subject- and non-subject-initial clauses is then treated as a general instance of "Constituent Preposing" (ibid.).

Zwart (1993)'s analysis of Dutch shows that wh-movement, topicalization and subject-initial sentences should not be folded into a single syntactic category due to their individual properties, stating that "verb movement in main clauses in Dutch targets different heads in each type of construction. However, the mechanism explaining the verb movement is by and large the same in each case." (Zwart 1993, p. 286). Each of these three possibilities correlates with a distinct functional category with a (for Dutch) strong N feature that triggers movement of a nominal element. An extended and generalized view on functional categories that capture information structural properties is given by Rizzi (1997). On this foundation, Speyer (2004) interprets his Optimality Theory (Prince & Smolensky (1993)) based account as targeting different functional heads.

A similar, configurational view on the prefield is presented for Head-Driven Phrase Structure Grammar (Pollard & Sag (1994)) in Müller (2007). From a verb-final syntagma, the finite verb is extracted to the front, and subsequently one element is chosen to be put in the first position of the sentence. Although the formal details differ, the HPSG account and the view of Generative Grammar both relate their theory of constituent structure to the topological properties of German. For multiple prefield elements, Speyer (2008a) can assume a violation of a OT constraint that demands only a single constituent in the prefield, resulting in more than one

[9]The original names of the relevant categories have been carefully adopted to that used, e.g., by Vikner (1995).

functional category in the left periphery to have an overt realization. Müller (2003) assumes that there still is a single constituent fronted, namely a VP with its head extracted previously.

The other type of solution is to dispose of configurationality and assume a radically flat hierarchy. For Lexical Functional Grammar (Bresnan (2001)), Choi (1999) is a representative of this option, using feature decomposition to explain the topological potential of scrambling. Between this and the configurational extreme, the work of Kathol (2000) argues for an explicit tier of linguistic structure on which linear order is represented that exists side-by-side with a configurational structure.

With this continuum between strictly configurational and strictly flat, where do we expect to find a suitable analysis in the framework of RRG? We already have a candidate for the strictly configurational view on the prefield, and found the argumentation provided in Diedrichsen (2008) not compelling. The single advantage, as it appears, is the fact that the precore slot is already "at the right spot" in the constituent projection. The problem that a configurational account of topological properties for German has, on a higher level, is that it relies strictly on *extraction* to model *displacement*.

Why is this problematic? Displacement into the PrCS is handled by the linking algorithm (and can be delegated to a constructional template). Placing a non-peripheral element in the prefield has the usual consequence that the Core template is reduced by one argument slot. If we decide to place a peripheral element into the prefield, however, nothing changes for the Core template. Instead, the syntactic representation of the scope of the periphery needs to be removed. Otherwise, we would end up with an element that is both a daughter of the PrCS node and linked to its target layer.

For arguments, the linking algorithm already correctly "undoes" the extraction, but there is no equivalent step to recover the actual syntactic relation an extracted periphery is in. While it may be possible to come up with an appropriately updated

linking algorithm for German, this is not the only additional work it will have to do. For each of the possible candidates in the prefield, we need to check (and, if not present, enhance accordingly) the algorithm. Intuitively, this goes against its cross-linguistic validity, and may very well result in a considerable number of steps to handle a word order phenomenon.

There is one alternative to extraction, which given the history of the phenomenon might appear radical: What if we assume that German, in fact, allows for a discontinuous CORE? On the plus side, the linking algorithm would no longer have to contain the word order rules of German to handle extraction scenarios. However, we need to find another way to constrain the CORE (and, possible, the CLAUSE as well)[10].

We carefully avoided to state that each instance of the prefield is explainable by discontinuity. The model proposed in this thesis does not take sides in the dichotomy of extraction versus linear order. Depending on the construction at hand, one can (and should) argue for either approach. Extraction, e.g., is necessary when an argument is coded in a higher-level CLAUSE or in cases of argument sharing (e.g. in CORE coordination). Nevertheless, we propose a unified approach to the prefield.

Before we present our resolution, let us recapitulate the properties of the prefield:

1. It does not necessarily correlate with topic or focus.

2. The pragmatic function, if any, applies to the prefield as a whole.

3. The prefield is not sensitive to a specific syntactic category.

4. It is not restricted to a single constituent.

If we combine these observations with the architecture of RRG outlined in section 3, one way to consistently represent the prefield is as an information unit.

[10]It should be noted that we would have to do the same work if we relied on the linking algorithm for topological phenomena. We expect, though, that keeping this problem domain seperated from the linking algorithm will be beneficial for both sides.

Information units are the primitive elements of the focus structure projection. They function independant of the predicate/argument distinction made by the semantic representation, and are also independant of any configurational structure expressed in the constituent projection. Of course, information units are formally tied to the surface structure of a clause, and carry with them the meaning of the elements contained, but not every information unit has a direct counterpart in either of the two other representations.

Once we take the step to use the focus structure projection to model the prefield, the other properties listed above follow almost at once. Since focus structure is concerned with the discourse status of referents, and hence with the actual focus domain and topicality, we should easily be able to capture the fact that everything in the prefield has the same pragmatic function. Also, given that it's not closely tied to the constituent projection, we can claim that an information unit can consist of more than one constituent. We go one step further and formulate more precisely: An information unit can consist of other information units, or it corresponds to a single constituent. To some extent, this recursive nature of information units is already used in RRG with respect to focus structure in complex sentences (van Valin 2005, p. 214).

With the resulting hierarchical structure, the actual focus domain can be represented as a single information unit (marked as being focused), which in turn expands into its "constituents". It might appear that we are currently only re-inventing constituent structure, but section (5.3) will show that this is not the case. Also, it is not merely a notational variant to have recursive information units: Our proposed integration of the prefield into RRG crucially relies on such a representation.

The discussion as of now has, on a high level, lead us to a solid grasp of the prefield. The next step is how we approach the problem of verb-second placement. Diedrichsen (2008) uses the linking algorithm to determine the correct position of the nucleus: if it is finite, it follows the PrCS, if it is non-finite, is linked to the end

of the CORE. We do not need to use the linking algorithm, though. Again, a switch of perspective on the problem leads to a cleaner solution: The crucial property of verb-second placement is not related to the constituent projection, but to the operator projection. The finite verb is always in the same position - following the clause-initial information unit. To be more precise: The operator expressing tense is in second position.

At first, this seems to be only a change of terminology. However, now that we have factored out verb-second placement as a constraint on the operator projection, the correct linking of nuclei comes (almost) for free. Non-finite nuclei are placed at the right edge of the CORE, and a finite nucleus, in German, is at the same time the tense operator of the CLAUSE. Therefore, the rule on operator placement already places finite nuclei at the correct spot.

As it stands, we should be able to collapse the two topological rules in a syntactic template. We need to make a small extension to the information that can be contained in such a template. While van Valin (2005) acknowledges that a template may also describe constraints on the focus struture projection, the operator projection has not been included in them.

A syntactic template for the German CLAUSE in declarative sentences is given in figure 6. An explanation of its properties is in order: The "X" represents the place of the tense operator and has no other meaning. It may seem as if the illocutionary force operator is missing, but this representation is intentional. We need it for example to handle sentence intertwining (section 5.5) correctly. Also, this makes it possible to encode the illocuationary force of a *wh*-question in the syntactic template for clause-initial *wh*-questions. Finally, the link between the constituent CORE and the operator CORE does not imply an elliptic construction. It is required because of the strong tie between the constituent and the operator projection.

Obviously, this template alone is too abstract to be usable in and of itself. To represent the CORE, we propose that instead of various CORE templates with a

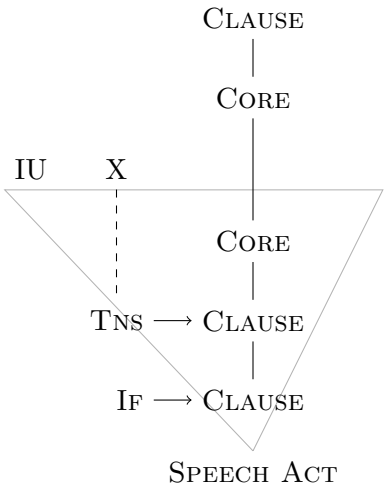

Figure 6: A syntactic template for verb-second sentences in German

fixed order of its daughters, German has a Core template that is completely unconstrained with respect to the linear order of its daughters. Combining such a Core template with the proposed Clause template forms the basis to account for all grammatical orderings in simple German clauses. As noted above, the spot for the tense operator and the predicate in the Core will interact correctly upon template combination to result in the correct placement of the finite verb.

In the following sections, we will now fill in the missing pieces. First, we have not yet shown how we intend to account for word order: We have an unconstrained Core template, and do not want to burden the linking algorithm with too much topological knowledge. Second, we present several constructions of German that show the feasibility of a non-configurational approach to the prefield, relying on the focus structure projection.

5.2 Word order constraints in simple sentences

The IU model already predicts the correct verb placement, but lacks a proper set of rules to govern the linearization of arguments or adjuncts. This problem does not

arise because of the way we analyse the prefield - the PrCS model has to come up with linearization rules as well. The IU model, however, introduces another twist to word order constraints - we need to ask ourselves in which domain they apply. In the context of the topological field model, word order constraints are always limited to the middlefield, given that the prefield hosts only a single constituent. The PrCS model was able to capture this topological fact by having the Core be equal to the middlefield - we had a configurational node that defined the domain for linearization. With the IU model, we do not seem to have a single element that defines the linearization domain.

This apparently becomes even more problematic if we take the linear order of elements in the prefield into account. As has been noted (Müller 2003, p. 23), the order of multiple elements in the prefield follows the same regularities as those in the middlefield. The most direct way to account for these facts is that there are in fact two linearization domains in the German clause, seperated by the tense operator. If we want to give this observation a configurational interpretation, we can claim that the Core defines two domains for linearization: one before the (verbal) tense operator, and one before the (possible empty) first verbal element *after* it. An alternative to this approach is to assume another level of representation for this exact purpose (following Kathol (2000)), but as long as our preliminary formulation suffices, we refrain from adding another level of description[11].

With two domains for linearization constraints situated properly, there are two questions we need to answer: First: What exactly are those linearization constraints? Second: Which criteria decide whether an element is placed in the information unit of the prefield or in the middlefield? Let us approach these two questions in order.

[11]Instead of introducing another level of representation, the obvious candidate seems to be the focus structure projection. Since the prefield is already defined here, one might consider an equivalent definition of the middlefield. If, and how, this is possible is a topic for future research in that area.

5.2.1 Linearization constraints in German

Word order in German, specifically in the middlefield, is a very complex topic in and of itself. For the purpose of this thesis, we shall only outline, by example, some of the factors that play a role here, among which there are:

1. The case marking on the arguments
2. The distinction between pronominal and non-pronominal RPs
3. Definiteness of the RPs involved
4. Animacy of the arguments
5. The *aktionsart* of the verb
6. Information structure, i.e. the marking of focal or topical referents

As a starting point, consider the following linearization constraints taken from van Valin & Diedrichsen (2006)[12]:

1. General constraints: pronoun > other, NP > PP
2. Case-based argument ordering constraint: NOM > DAT > ACC (default)
3. If ACC = pronoun, then ACC > DAT (default)

Lenerz (2001) gives the following ordering rules (based on Lenerz (1977) and Büring (1996)):

1. [±def IO] > [±def DO]
2. [+def DO] > $[\text{IO}]_F$
3. *$[\pm \text{ DO}]_F$ > IO
4. *[-def DO] > $[\text{IO}]_F$

[12]In van Valin & Diedrichsen (2006), these constraints apply to all arguments except the one linked to the PrCS, i.e only the middlefield.

The first rule defines the "unmarked order", while the second one allows a focused indirect object to be placed after a definite direct object. The last two rules show that is is ungrammatical to scramble a focused direct object, and that an indefinte direct object must not precede a focused indirect object. We already begin to see a non-trivial interaction of case/grammatical function, definiteness and focus marking. Instead of adding additional complexity to this rule set, we shall at first take a step back and discuss methods to derive such word order rules.

In their discussion of adverbial placement in German, Frey & Pittner (1998) explore six methods to determine base positions of arguments. Since their work is framed in the context of Generative Grammar, we will not be able to transfer each method to an RRG based analysis. The respective methods deal with:

1. Word order in sentence focus constructions
2. The theme-rheme-condition of Lenerz (1977)
3. Effects due to binding principle C
4. Fixed positions of existentially interpreted *wh*-phrases
5. Complex prefield fillings
6. Quantifier scope in *verum focus* sentences

Let us quickly recapitulate these methods, as well as their findings, while reinterpreting them (to some extent) in the framework of RRG.

The first method to identify a default ordering of arguments is to look at sentence focus constructions (capitalization indicates stress).

(44) *Was ist geschehen?*
What is happened?
'What happened?' (= Frey & Pittner 1998, (6))

a. *Gestern hat ein Kollege einer Dame ein GeDICHT vorgetragen*
yesterday have.3SG a colleague.NOM a woman.DAT a poem.ACC recited
'Yesterday, a colleague recited a poem to a woman'

b. *#Gestern hat ein Kollege ein Gedicht einer DAme vorgetragen*

While the first answer can have sentence focus, the second one has narrow focus on the stressed constituent, and is therefore not acceptable in the given discourse context. The deviation from the default ordering induces focus structural effects that render a sentence focus interpretation impossible. For a verb like "vortragen", the default word order uncovered by this method is indeed NOM > DAT > ACC.

The same result is achieved by employing the theme/rheme condition of Lenerz (1977) for (45). In this approach, we force narrow focus in the answer by asking a *wh*-question:

(45) *Wem hat Otto heute ein Gedicht vorgetragen?*
Who.DAT has Otto.NOM today a poem.ACC recited?
'To whom did Otto recite a poem today?' (= Frey & Pittner 1998, (7))

a. *Otto hat heute einer KolLEGin ein Gedicht vorgetragen.*
Otto.NOM has today a colleague.DAT a poem recited.
'Otto recited a poem to a colleague today.'

b. *Otto hat heute ein Gedicht einer KolLEGin vorgetragen*

(46) *Was hat Otto heute einer Kollegin vorgetragen?*
What.ACC has Otto.NOM today a colleague.DAT recited?
'What did Otto recite to a colleague today?' (= Frey & Pittner 1998, (8))

a. *Otto hat heute einer Kollegin ein GeDICHT vorgetragen.*
Otto.NOM has today a colleague.DAT a poem recited.
'Otto recited a poem to a colleague today.'

b. ?? *Otto hat heute ein GeDICHT einer Kollegin vorgetragen*

In (45), both ACC > DAT and DAT > ACC is possible, the actual focus domain being the dative RP in both cases. Hoever, as (46), when the actual focus domain is the accusative NP, ACC > DAT is highly marked. This effect is explained by the theme-rheme-condition of Lenerz (1977)[13]:

> Die Abfolge BA kann gegenüber der Abfolge AB dadurch eingeschränkt sein, daß in ihr B (bei thematischem A) nicht Rhema sein kann.

Being less constrained, the order DAT > ACC can be seen as more basic.

We now turn to the effects of principle C violations. Since RRG accounts for binding effects not on the configurational level but on the logical structure, we cannot directly deduce an underlying word order pattern. If anything, such effects provide insight into the semantic representation of the predicate at hand. The invariability of existentially interpreted *wh*-phrases, on the other hand, does help in identifying default argument positions, and hence a default word order: The data in (47) shows that nominative RPs are per default realized before accusative RPs:

(47) a. *weil ein Professor wen beleidigt hat*
because a professor.NOM someone.ACC insulted has
'because a professor has insulted someone.' (= Frey & Pittner 1998, (12))

b. * *weil wen ein Professor beleidigt hat*

The possibility of complex prefield fillings must be set aside for our discussion of basic word order, since the arguments presented by Frey & Pittner (1998) assume a 'VP' constituent. In our account of the prefield, though, the acceptability of partial 'VP's in the prefield cannot be a question of underlying word order but is instead one of a "possible information unit". Likewise, taking scope facts into account does not necessarily help us uncover basic word order, again due to the way

[13] "The order BA can be constrained with respect to the order BA in the sense that B (in the case of a thematic A) cannot be rheme."

RRG analyses quantifier scope. The approach by Frey & Pittner (1998) might prove helpful nonetheless, since they restrict their data to *verum focus* sentences - neither of the quantifiers should therefore be focal:

(48) a. *Gestern HAT er mindestens einer Dame fast*
Yesterday has he.NOM at least one lady.DAT almost
jedes Gedicht vorgetragen
every poem.ACC recited
'Yesterday, he DID recite almost every poem to at least one lady.' (= Frey & Pittner 1998, (15))

b. *Gestern HAT er fast jedes Gedicht mindestens einer Dame vorgetragen*

(48a) is unambiguous, but (48b) is ambiguous. What can this, in the framework of RRG, tell us about default word order? For quantifier scope, RRG assumes that, as a default, topical quantifiers take scope over focal quantifiers (cf. van Valin 2005, p. 82). Consequently, the deviation in word order in (48b) can have repercussions in the information structure of the clause, while (48a), having only a single interpretation, appears to be the unmarked case. Again, DAT > ACC comes out as the unmarked order (which was to be expected given the results of the other methods discussed so far).

Is NOM > DAT > ACC always the basic word order for ditransitive verbs? In Frey & Pittner (1998), we find two deviations from this ordering. First, let us consider the verb "unterziehen" (to subject)[14]. In sentence focus constructions, the default order appears to be NOM > ACC > DAT:

(49) a. *Gestern hat der Peter das Auto dieser PRÜfung*
yesterday has the Peter.NOM the car.ACC this check.DAT
unterzogen
subjected
'Yesterday, Peter subjected the car to this check.' (based on Frey & Pittner 1998, (18c,d))

[14]Frey & Pittner (1998) discuss the behaviour of "unterziehen" in the context of complex prefield elements. As said before, we cannot make us of this method since we lack a configurational concept of 'VP'. If we leave restrict ourselves to middlefield order, we still observe the same effects, though.

b. *#Gestern hat der Peter dieser Prüfung das AUto unterzogen*

(49b) is not ungrammatical, but does not allow sentence focus interpretation when the preverbal element is stressed. One might argue that this is due to the influence of the demonstrative determiner, but changing it to a definite article does not change the impossibility of sentence focus (50a). Even more, if we use an indefinite article, NOM > DAT > ACC becomes unacceptable to the point of being ungrammatical (50b). For "normal" ditransitive verbs like "geben", the equivalent displacement is less marked (51):

(50) a. *#Gestern hat der Peter der Prüfung das AUTO unterzogen*

b. **?Gestern hat der Peter einer Prüfung das AUTO unterzogen*

c. *Gestern hat der Peter das Auto einer PRÜFUNG unterzogen*

(51) a. *Gestern hat der Peter ein Buch dem MANN gegeben*

b. *Gestern hat der Peter das Buch einem MANN gegeben*

As it stands, the dative argument of "unterziehen" does not have the same syntactic status as the non-macrorole core argument of a verb like "give" (geben). The following sketches a possible semantic explanation for this: The selectional restrictions for the dative RP occuring with "unterziehen" are quite strong, suggesting a different logical structure. Let us assume that the predicate in the above sentence is in fact not only "unterziehen", but "einer Prüfung unterziehen". The "incorporated" argument is non-referential in (50c), while the dative RP in (51b) is not. When a speaker *does* want to refer to a specific check, the complex predicate is dissolved in the logical structure, resulting in an additional core argument. On the constituent projection, the "incorporated" argument is realized inside the nucleus. Its case is handled correctly by RRG, since dative is the default case. The NOM > DAT > ACC order in (50b) is ungrammatical because, by assumption, German does not allow discontinuous nuclei (cf. figure 7).

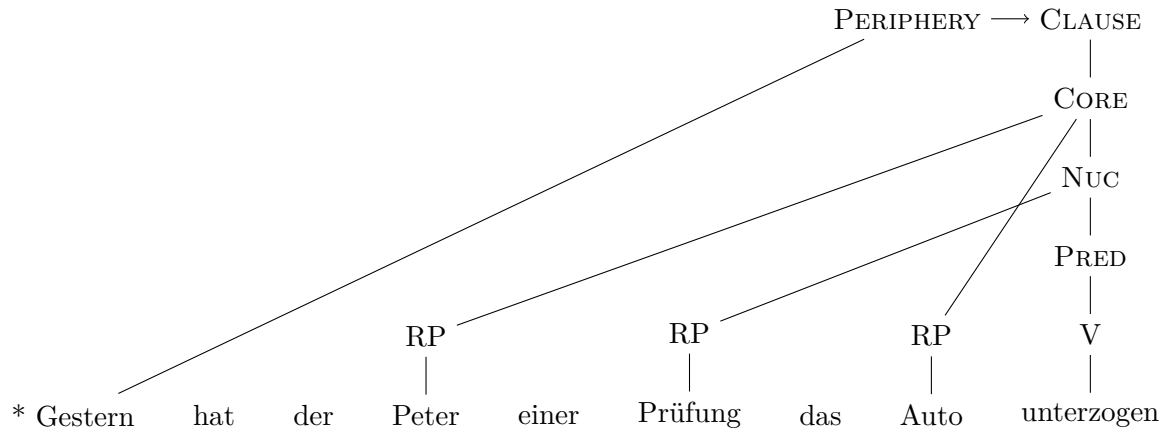

Figure 7: Constituent projection for "einer Prüfung unterziehen"

Another deviation from the NOM > DAT > ACC pattern is given by non-agentive readings of the verb "beeindrucken" (to impress):

(52) *Gestern hat ein Professor eine StuDENtin beeindruckt*
yesterday has a professor.Nom a student.Acc impressed
'Yesterday, a professor impressed a student.' (= Frey & Pittner 1998, (17b))

"beeindrucken" can have two different meanings here. In the "agentive" reading, the professor consciously acts in such a way that the student is impressed by him. The non-agentive reading, however, implies that the professor is not necessarily aware of the effects that his actions have. In this latter interpretation, the sentence in (52) can only have narrow focus, *contra* what we would expect if only case played a role. The non-agentive reading is very similar to that of object-experiencer verbs like "ängstigen" (to frighten) or "gefallen" (to please):

(53) a. *Gestern hat ein Film eine StuDENtin geängstigt*
yesterday has a movie.Nom a student.Acc frightened
'Yesterday, a movie frightened a student'

b. *Gestern hat eine Studentin ein FILM geängstigt*

c. *Gestern hat ein Film einer StuDENtin gefallen*
yesterday has a movie.Nom a student.Dat pleased
'Yesterday, a movie pleased a student.'

d. *Gestern hat einer Studentin ein FILM gefallen*

The pattern for object experiencer verbs deviates from the default case ordering constraints we arrived at before. The nominative argument appears after the dative/accusative argument in sentence focus constructions. We do not intend to go into more detail on word order patterns in sentence focus constructions for different classes of verbs here. For the current discussion, the key element is that the *aktionsart* of a predicate contributes to the basic word order by defining the default case-based ordering. What remains is the influence that focus placement can have. In other words: What are possible places for the actual focus domain in a German clause?

While discussing the theme/rheme condition of Lenerz (1977), we already encountered two variants for placing a focused constituent (see examples (45) and (46)). If the focused RP is dative, it is either realized *in situ*, or it is moved to the right edge of the middlefield. An accusative RP can either be realized *in situ*, or it can switch places with the dative RP (yielding a highly marked order). The markedness of this last option can be explained by the fact that the displacement is superfluous. For an accusative RP, being realized *in situ* and at the right edge of the middlefield require no reordering (in contrast to the middlefield-final placement of a dative RP). What can we observe when a nominative RP is in the actual focus domain?

(54) Q: *Wer hat gestern dem Mann das Buch gegeben?*
who.NOM has yesterday the man.DAT the book.ACC given?
'Who has given the book to the man yesterday?'

A: (a) *Gestern hat der PETER dem Mann das Buch gegeben*

(b) ? *Gestern hat der PETER das Buch dem Mann gegeben*

(c) ?? *Gestern hat dem Mann der PETER das Buch gegeben*

(d) ?? *Gestern hat das Buch der PETER dem Mann gegeben*

(e) *Gestern hat dem Mann das Buch der PETER gegeben*

(f) ? *Gestern hat das Buch dem Mann der PETER gegeben*

All of the six possible argument orderings in the middlefield are grammatical. The introspective judgements on most of them need to be explained, though. The first answer leaves the focused RP *in situ*, and the other arguments follow in the order predicted by their case marking. To explain the decreased acceptability of the second answer, we need to be aware of two deviations: first, the order does not correspond to the one given in the question. Second, the order also deviates from the default case-based ordering, resulting in a subtle difference with respect to the information conveyed by the answer: Not only do we have completive focus on the nominative RP, but the dislocation of the accusative marks it as the topic of the answer (for a more general discussion of fronting in the middlefield in the context of information structure, see Musan (2002)).

In the third (and fourth) answer, the effect is similar: the unstressed, displaced RP at the start of the middlefield is topical. What makes these cases worse is that the focused RP does neither occur *in situ*, nor in the alternative, middlefield final position. The fifth answer is, again, better since the case ordering is preserved among the non-focused arguments, and the focused RP occurs at the end of the middlefield. Finally, the sixth and last possibility again displaces the accusative RP, marking it as topical in addition to the completive focus on the subject.

As this short exploration has shown, the placement of the actual focus domain interacts with at least two other ordering systems: The case-based ordering (which is typically NOM > DAT > ACC) and the marking of topics by realizing them at the start of the middlefield. Additionally, we started to uncover that the case hierarchy depends on the *aktionsart* of the predicate involved.

5.2.2 Filling the prefield

The next problem we have to solve is: which elements are placed in the prefield? The view adopted in this thesis is that the prefield is filled by the "best" candidate,

given a set of factors. With a focus structure based account on this topological position, we already acknowledge that it is not only the morphosyntactic status of a constituent that plays a role here, but that in addition discourse context can have an impact as well.

Recently, using a competitive approach has been proposed for the middlefield by Pafel (2009) on the basis of the topological field model. For the prefield, another proposal is made by Speyer (2004) and his later work. For a critic evaluation on the competitive aspect, as opposed to a collaboration of multiple levels, see Lenerz (2001). In the following, we outline the model of Speyer and show how its insights can be applied to our analysis.

The theoretical framework underlying the competitive model of Speyer (2004) is Centering Theory (CT, Grosz et al. (1995), Walker et al. (1998)). As argued for in Speyer (2007), CT - especially its concepts of discourse coherence - is vital to predict and explain which constituents are placed in the prefield of a German clause. The discourse model used by RRG as of now relates a sentence only to its previous context, enabling us to factor out its pragmatic presupposition and the pragmatic assertion. Centering Theory, as we will see, adds the "future" of the discourse to this view.

Considering two sentences, A and B, where B immediately follows A in the discourse, we can define three types of centers: Each referential element of sentence A is a *forward-looking center* and might be referenced again in sentence B. Depending on its syntactic type, a referential element is more or less likely to be referenced again. Defined by a language-specific preference hierarchy, one of the forward-looking centers of sentence A is the *preferred center*. In sentence B, the referential element that picks up a center of sentence A is the *backward-looking center* (also simply called *center*).

Based on these terms, we can come up with the following prediction[15]: "Man er-

[15]"It is to be expected that the Center changes in the next sentence for sentences where the Backward-looking Center is not simultaneously the Preferred Center."

wartet, dass sich das Center im nächsten Satz verändert in Sätzen, wo das Backward-looking Center nicht gleichzeitig auch das Preferred Center ist." (Speyer 2007, p. 87). The "ideal" discourse, under this view, always realizes the preferred center of sentence A as the preferred center of sentence B. A grammaticalized version of this principle are topic chains, or coordination reduction constructions: The subject of all but the first conjunct is dropped since they are coreferential. In case where the default voice of a sentence would result in a non-preferred center, passive occurs (cf. van Valin 2005, p.103f).

CT predicts four different ways to change the center in discourse:

Continue keeps the backword-looking center of A, and realizes it as the preferred center in B.

Retain keeps the backword-looking center of A without it being the preferred center in B.

Smooth Shift changes the backword-looking center, and realizes it as the preferred center in B.

Rough Shift does neither keep the backword-looking center, nor realize the new center in the preferred way in B.

It appears reasonable to assume that the prefield is used to maximize the coherence between to sentences. To this end, the center as the crucial factor of coherence should appear in it. In the terminology of RRG, we would expect that topics are the most prominent candidate for the prefield[16].

However, based on a corpus study Speyer (2007) comes to the conclusion that this is not the case[17]. Instead, his data suggests that the prefield is the preferred

[16]Although not covered by Centering Theory (which can only handle referential elements), the discourse connectors listed in (11) would also be expected to occur in the prefield frequently. The corpus study of Dipper & Zinsmeister (2009) provides evidence that this is not the case.

[17]Dipper & Zinsmeister (2009) also states that "[c]ontrary to our working hypothesis, the majority of Vorfeld constituents in our corpus is not related to the prior context (by a relation reference or discourse)." (Dipper & Zinsmeister 2009, p. 77)

place for either scene-setting or focal information. Speyer (2008a) adds an additional constraint that only a single constituent is allowed in the prefield, resulting in the following ranking of thesse possibilities in his OT-based analysis:

1-VF » Scene-setting VF » Contrast-VF » Topic-VF

Hence: "if a sentence contains more than one phrase conforming to the conditions stated in the vorfeld-constraints, the optimal candidate has the phrase in the vorfeld that conforms to the conditions of the highest-ranked relevant constraint" (Speyer 2008b, p. 287).

While our proposal does explicitly deny the 1-VF constraint, the information structure based notions appear to integrate very well with our clausal template. One problem of this approach, as well as the model we propose, is the occurence of a dummy "es" in the prefield. A detailed account of its properties from an RRG perspective can be found in Kretzschmar (2006). For the constraint model, Speyer (2009) proposes a way for integration.

Why is the occurence of the dummy "es" in the prefield problematic? In our account, we assume that the prefield is filled by an information unit. Quite the opposite is true, though: the element in question does not contain any information at all. For a consistent representation of the clause, we need to assume that the dummy "es" *does* have a representation on the focus structure projection, but that this information unit is void. As this is an exceptional situation for an information unit, we should represent this property in a constructional template. This is exactly what Kretzschmar (2006) arrives at. She makes the interesting observation "dass Sätze mit einem Vorfeld-*es* stets dem Sentence Focus zuzuordnen sind" (Kretzschmar 2006, p. 71). Therefore, the dummy "es" acts as an illocutionary force operator for the Clause. Instead of a last-resort mechanism, dummy "es" becomes an integral part of a constructional template (cf. Kretzschmar 2006, (II-4), p. 76). All we need to do is to add the exceptional status with regard to the focus structure projection representation to this constructional template.

Is this view compatible with the constraint model discussed so far? The expectation of a constraint-based model is that this pronoun only occurs if no other suitable candidate can be determined: There should be neither a scene-setting, nor a contrasting, nor a topical element available to be the center of the clause. From this perspective, the dummy "es" appears to be more a last-resort mechanism than an explicit construction of German. However, the properties of sentence focus constructions in fact *predict* the possibility of a dummy "es": By definition, there is a) no topical element available and b) we have not a single referential entity focused but the whole clause. None of the information structure-based constraints can select a suitable candidate. Hence, the two analyses of the dummy "es" do not contradict each other but simply approach the same effect from two different perspectives.

At this point, we set the issues of word ordering and prefield filling aside. We have proposed two different linearization domains inside the Core, which are split by the tense operator. For both domains, we explored various factors that affect word order. Given the complex interactions that are relevant, we can argue *against* using the linking algorithm to handle the details of word order in German sentences directly. Instead, we need an additional step in the linking algorithm that delegates this work to a language-specific set of rules. In addition, we looked at one way to decide which elements are placed in the prefield. The notions involved in this decision were always related to information structure, further supporting our analysis of the prefield as an information unit. The subsequent sections will now apply our proposal to various constructions of German.

5.3 The discontinuous properties of the Core

We have argued in section 5.1 that the Core in German can be discontinuous. In this section, we will look at several constructions that involve the Core and compare our information unit based approach against the configurational PrCS model.

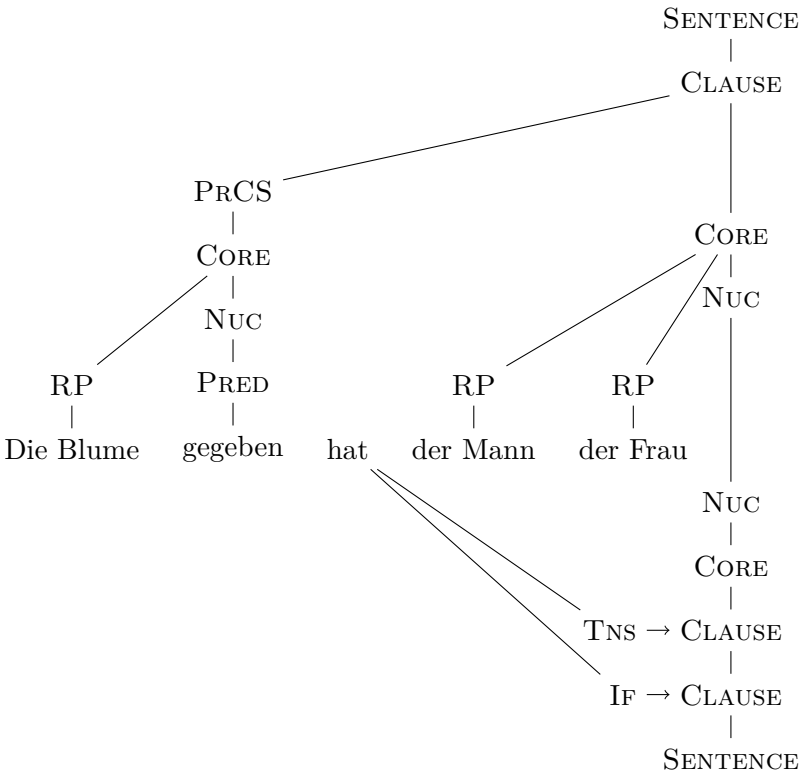

Figure 8: With an obligatory PrCS, partial VP fronting requires ellipsis

5.3.1 Partial 'VP' fronting

The first construction we look at can be called "partial VP fronting". Consider the sentence in (55), repeated from (20).

(55) *Die Blume gegeben hat der Mann der Frau.*

The prefield is occupied by the nucleus and one of the arguments of the Core. The two remaining arguments occur after the finite verb. Under the assumption that the prefield equals the PrCS, we have a problem. The PrCS can host only a single constituent. If we decide to adhere to this restriction, we need a single element - a Core in this case - in the PrCS that contains both the nucleus and the argument At the same time, we need to have the remaining arguments belong to a Core as well. The only viable solution for this dilemma is to analyse this second Core as an ellipsis, as is shown in figure 8.

An elliptic constituent projection might be justified for this construction. Closer scrutiny, however, reveals that this is not the case. There is a crucial difference between this elliptic structure and that of a sentence like (56):

(56) *Kim is eating an ice cream cone, and Sandy is, too.* (= van Valin 2005, (7.10c))

The reason for a configurational ellipsis in this sentence is the presence of a clause operator in the second conjunct. Therefore, the constituent projection of the second conjunct needs to include a CLAUSE. As there is no lexical predicate present, the CORE and NUC nodes required below the CLAUSE remain empty, yielding the elliptic structure. Additionally, the occurence of such an elliptic CLAUSE needs to be licensed by the immediate discourse context. In (56), this context is provided by the first conjunct[18].

To the German example, this reasoning does not apply too well. It would require us to set up a CLAUSE-local discourse context out of which the second CORE draws its interpretation. The only reason to introduce an elliptical structure in the second "half" of the sentence is a direct consequence of an obligatory PRCS node. There is no *linguistic* argument that leads to an elliptic constituent projection.

Furthermore, the operator projection of such an elliptic structure would no longer represent the morphosyntactic dependencies directly. To retain a minimal amount of consistency, the tense as well as the illocutionary force operator apply to the elliptic CLAUSE, further shadowing the relation between the auxiliary "hat" and the participle form of the predicate. To overcome all this complexity, we could attempt to formulate a constructional template for partial 'VP' fronting that encodes the deviations noted so far. Since our initial motiviation for an obligatory precore slot was a consistent treatment of the prefield, this price in added complexity appears to be too high.

[18]For a possible analysis of discourse 'VP' ellipsis, see (van Valin 2005, p. 233).

What happens if we loosen the configurational constraints on the PrCS? As Diedrichsen (2008) proposes in her discussion of 'VP' fronting: "It is claimed that this position has to be filled with elements that belong to the same pragmatic domain, while the number of 'constituents' is not generally restricted." (Diedrichsen 2008, p. 221). For (55), this means that the PrCS immediately dominates the Nuc and RP. Even then, we would be confronted with an elliptical Clause without any compelling evidence.

For the template selection, partial 'VP' fronting poses a challenge as well: If treated as extraction, we are forced to assume Core ellipsis, and we have to formulate a rule that describes the pragmatic motivation behind the displacement. From the point of view of the template selection step in the linking algorithm, we now have to choose two Core templates. Each of these templates needs to have the correct number of slots for arguments, since the arguments of the predicate are distributed between an overtly realized Core and an elliptic one[19].

How does the IU model overcome these difficulties? Apparently, there are none to overcome: There is only a single Core, whose daughters are not realized continuously (since the tense operator intervenes). What remains is to argue for a single information unit in front of the finite verb.

For the case at hand, we assume the following context for (55)[20]: There is a man who is given two different tasks: he shall give a rare blue flower to someone and he shall steal the purse of someone. He is accompanied by an observer who shall later give a detailed account of the man's actions. After all tasks are done, the observer is questioned on what the man did.

The information in (55) is then seperated as follows: In the prefield, we find an information unit carrying contrastive focus. There are two tasks available for

[19]If we follow Diedrichsen (2008), the PrCS will not host a Core, but only the Nuc and any number of RPs. In this case, we need to select one Core template, appropriately reduced by the Nuc and whichever arguments appear in the prefield.

[20]As van Valin (2005) proposes, one can use Discourse Representation Theory to formalize this problem.

description, and the observer chooses to start with the "flower" task. The man is topical, and finally the new information is the narrow focus on "der Frau": it was not clear beforehand who the target of the "flower" task would be.

The sentence in (55) is nevertheless marked, which can be explained by the requirements on the discourse context. It was necessary to have a topic, a completive focus and a contrastive focus, all at the same time. Nevertheless, in this specific context the sentence appears to be a valid option to encode these three distinctions simultaneously, which is why we can safely say that the IU model of the clause is justified.

5.3.2 'tun'-Periphrasis

Diedrichsen (2008) shows that her approach to the prefield correctly explains 'tun'-periphrasis constructions as seen in (57).

(57) *Waschen* *tut* *er* *das* *Auto* *nie.*
wash.INF do.3SG.PRES 3m.SG.NOM def.N.SG.ACC car.SG never
'He never washes the car.' (= Diedrichsen 2008, (32a))

The dummy verb "tun" (to do) is required in her model since the topical "waschen" could not go into the PRCS if it was finite. The IU model has a slightly different perspective on this construction: The predicate is the clause-initial information unit, therefore we lack a tense operator. As a last-resort mechanism, "tun" is inserted. We cannot argue for or against either of the two models with this construction, but can only note that both explain it adequately.

5.3.3 Matrix coding constructions

The example in (58) is a case of core coordination[21]:

(58) *Ich* *sah* *den* *Polizisten* *oft* *der* *Frau* *helfen*
I saw the police officer.ACC often the.DAT woman help

[21] (58) is actually ambiguous. The adverbial can modify either of the two CORES. For the present discussion, this issue is irrelevant, hence we only look at one of the two interpretations.

'I often saw the police officer help the woman'

There are other orderings possible, subject to focus structure restrictions:

(59) a. *Den Polizisten sah ich oft der Frau helfen*

b. *Helfen sah ich den Polizisten der Frau oft*

c. *Der Frau sah ich den Polizisten oft helfen*

If we apply the PrCS model, we would arrive at the following situations: In (59a), the matrix Core is reduced by one slot and the matrix-encoded argument is linked to the PrCS. (59b) is a case of partial 'VP' fronting, yielding an elliptic, coordinated second Core. In (59c), the coordinated Core is reduced by one slot (it now has zero syntactic arguments), and its non-macrorole core argument is linked to the PrCS.

Now, consider the variation in (60).

(60) *Den PoliZISTen helfen sah ich OFT.*
the police officer help saw I often
'I saw the police officer help often'

(60) is constructed along the lines of (21e). Both put a predicate and its actor in the prefield, with the difference that this time the syntactic encoding of the actor is different. For the PrCS model, such data raises an important question concerning the constituency of the precore slot. In the case of 'VP' fronting, we were able to have an option - either the PrCS hosts a Core or it hosts multiple constituents. Now, though, there is no single element in the constituent projection that could be put below the PrCS node, since this would predict the wrong constituent structure of this matrix coding construction. We *must* assume that the PrCS is not a single constituent. Once we have commited ourselves to this, we need to state the precise conditions under which more than one constituent can appear in it.

Besides this issue, we also lose a general representation of Core junctures. The usual representation of core coordination is a Clause that directly dominates two

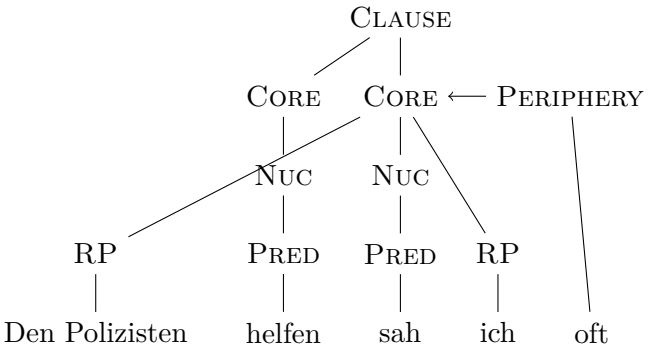

Figure 9: The constituent projection for core coordination

Cores. With the PrCS model, however, sometimes one of the coordinated Cores is "trapped" in the precore slot. Like before, this problem would propagate to the template selection step of the linking algorithm, which must select different templates depending on the linear arrangement of the clause.

The IU model handles this case in the same way as it previously handled partial 'VP' fronting. We can, again, argue for an information unit on the grounds of the *semantic* representation of the sentence. The predicate of the second Core and its sole argument form the topic of (60). The fact that said semantic argument is a syntactic argument of the first Core is irrelevant to information structure. Such behaviour is actually *predicted* by the assumption that a) the Core in German can be discontinuous and b) that the prefield is an information unit.

The same reasoning applies to raising and control constructions. As the examples (16) and (17) have shown, any one of the constituents in a raising or control construction can be put into the prefield. The discontinuity of the Core directly accounts for this data, while the PrCS model would have to assume extraction out of either of the Cores involved.

5.3.4 Multiple arguments in the prefield

In section (2.2.2), we presented one case of multiple arguments, without the predicate, in the prefield. The relevant data is repeated in (61) from (23d,e).

(61) a. *Nicht der Anna einen Brief hätte er schreiben sollen, sondern der Ina eine Postkarte*
Not the Anna a letter would have he write should but the Ina a post card
'He should have written not a letter to Anna, but a post card to Ina' (= Müller 2003, 62a)

b. ?? *Der Anna einen Brief hätte er schreiben sollen, nicht der Ina eine Postkarte*

The problems that the configurational approach of the PrCS model faces here have been discussed already: We would have to assume either a single constituent hosting two arguments without their predicate, or multiple constituents without proper explanation. Similar to 'VP' fronting and core coordination, in (61a) we have a contrastive focus in the prefield. The partitioning of information works as expected: In this specific context, the presupposition is that "he" wrote a letter to Anna. Since this was the wrong thing for "him" to do, the speaker contrasts the violating parts of the presupposition with their correct counterparts: Both the undergoer and the non-macrorole core argument are false, hence subsumed under the actual focus domain and marked by the focus particle "nicht". The actual focus domain of the first clause can then be placed into the prefield as usual. The second clause in (61a) consists of a VP ellipsis, which depends on the "correct" part of the presupposition. The new, corrected information is in this clause's actual focus domain, being completive focus.

Why is the variant in (61b) marked? In this example, the completive focus is put into the prefield of the first clause, and the contrastive information follows later. We suggest that there are two reasons for the decreased acceptability of this order. First, this way of presenting information is marked: Stating what is correct depends

on stating what was wrong, and therefore the contrastive focus should come first. Second, both focus domains have no direct counterpart in either the constituent projection or the logical structure, and are hence more difficult to process. The addition of the focus particle in (61a) very early indicates that what follows is a single information unit. In (61b), no such marker is present, which is why the speaker expects a tense operator right after the first contituent.

Again, the IU account is on the one hand flexible enough to handle such marked constructions, while at the same time explaining said markedness: For (61), none of the modules of RRG directly represents the unit before the tense operator. On the constituent projection, we are faced with two seperate constituents that are displaced from their core. The logical structure also does only contain a predicate containing both arguments seperately. Even on the discourse level, a framework like DRT lists both information seperately. It is only because both are contrastive at the same time that this construction is possible - which is exactly what our account would predict.

5.4 On seemingly discontinuous NPs and the prefield

We now return to the "other" data that can be present in the prefield of a German clause. Section (2.2.3) presented samples of a construction that is commonly referred to as "split topicalization". The relevant data is repeated as (62) from (24):

(62) a. *Syntaktiker* *kenne* *ich* *drei/viele/keine*
syntactitians know I three/many/none
'As for syntacticians, I know three/many/none'

b. *Polizeiwagen* *kenne* *ich* *nur* *grüne*
police cars know I only green
'As for police cars, I only know green ones'

At first sight, we might be tempted to treat this construction similar to the analysis of the clitical pronoun *ne* in Italian (cf. van Valin 2005, p. 175ff). *ne* is

analyzed as the topical head of a discontinuous NP, with the other part constituting a focal quantifier:

(63) Q: *Quanti* *student-i* *sono* *venu-t-i* *alla* *festa?*
how.many student-pl be.3PL come-PSTP-3MPL to.the party
'How many students came to the party?' (= van Valin 2005, (5.38))

A: *Ne* *sono* *venu-t-i* *ventiquattro.*
ne be.3PL come-PSTP-3MPL twenty-fourth
'Twenty four of them came'

The split between topical and focal information is uncontroversially present in the German data as well. Even more, German appears to be more flexible: The topical element can be a bare plural, and the focal element can be any restrictive modifier (viz. the use of an adjective in (62b)). Also, an adjacent realization of the NP is grammatical as well, further supporting the hypothesis that these are NPs that have been split up due to focus structure requirements.

However, if we start to take a closer look at this phenomenon, a configurational analysis of the split loses a solid foundation. If there indeed was a discontinuous NP involved, we would expect each NP to be able to realize its head in the pre-field, while everything else stays the same. However, there is a) a morphosyntactic difference between the "discontinuous" variant and the adjacent realization, and b) there is no agreement in number between the topical and the focal element for the "discontinuous" case. This is shown for quantifiers in (64), but holds for attributive adjectives as well:

(64) a. *Bücher* *kenne* *ich* *drei*
book.3PL know I three
'I know three books'

b. *Ich kenne drei Bücher*

c. **Ich kenne drei Buch*

d. *Ich kenne ein Buch*

e. **Buch kenne ich ein*

f. *Bücher kenne ich eins*

g. **Bücher kenne ich ein*

h. **Ich kenne eins Buch*

i. **Ich kenne eins Bücher*

(64a) shows the apparently discontinuous NP - the topical part is marked for plural, and the quantifier agrees with it. The adjacent case is (64b), (64c) shows that we have indeed a morphosyntactic dependency - the quantifier is inherently marked for plural, and hence cannot modify a singular NP. If we reduce the quantity of books involved to one, we get a different picture. (64d) is the grammatical, adjacent case. If we realize the topical part in the prefield, the sentence becomes ungrammatical (64e). Instead, we have to change two parts of the sentence: the topical element must become referentially unspecific, which we achieve by using a bare plural, and the quantifier is marked with strong inflectional morphology (64f). If we change only one of these, the result is ungrammatical (64g,h). Also, the adjacent realization of a strongly inflected quantifier is ungrammatical (64i,j). To overcome these arguments and make a discontinuity analysis viable, we would have to assume that the quantifier is somehow sensitive for a dislocated head. It is questionable whether this added complexity is necessary.

Similar arguments against a syntactic discontinuity can be made on the basis of examples as (65):

(65) a. *Syntaktiker kenne ich nur Chomsky*
syntacticians know I only Chomsky
'As for syntacticians, I only know Chomsky'

b. ** Ich kenne nur Syntaktiker Chomsky*

c. ** Ich kenne nur Chomsky Syntaktiker*

d. *Ich kenne nur den Syntaktiker Chomsky*

e. *Ich kenne nur die Syntaktiker Chomsky, Sternefeld und Uszkoreit*

f. ** Ich kenne nur Syntaktiker Chomsky, Sternefeld und Uszkoreit*

g. *Syntaktiker kenne ich nur Chomsky, Sternefeld und Uszkoreit*

The information-structural properties are the same as with the quantifiers and attributive adjectives discussed above: The nominal element in the prefield is topical, and at the end of the clause we have a restrictive focal element. Interestingly enough, this time the restricting element is a proper name, which we can safely assume to be an RP. Intuitively, the adjacent variant should be an apposition, with the topical element as the syntactic head. The naive attempt is given in (65b,c), which are both ungrammatical (as expected), and (65d) appears to be the correct adjacent variant. By extension, (65e,f,g) show the effect plural marking has on the grammaticality of this construction.

For appositions, it might not be clear which of the nominal elements involved is the syntactic head. We can use the fact that proper names are not marked for any case but the genitive as a diagnostic tool. Albeit rare, German has verbs that assign genitive case to one of its arguments:

(66) a. *Am Morgen gedachte ich des Syntaktikers Chomsky*
in.the morning commemorated I the.GEN syntactician.GEN Chomsky
'In the morning, I commemorated the syntactician Chomsky.'

b. *Am Morgen gedachte ich Chomskys*

c. **Am Morgen gedachte ich des Syntaktikers Chomskys*

As expected by the pattern of the construction discussed so far, in the adjacent case "Syntaktiker" is the syntactic head, as shown in (66). How can we account for this construction on the constituent projection? Instead of syntactic discontinuity, we propose that there are actually two separate RPs present in the constituent projection. We cannot place both of these RPs into the CORE, since we do not have

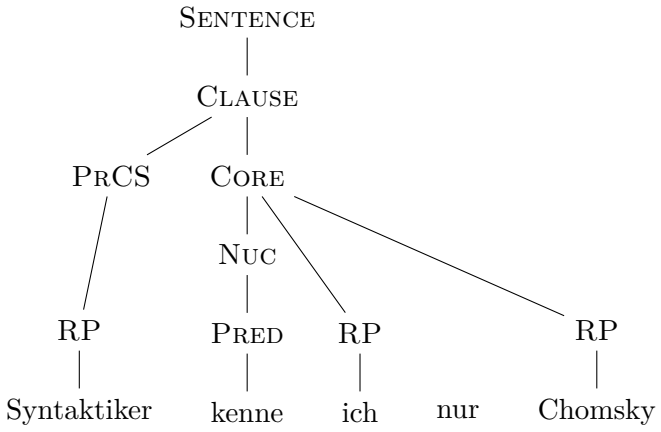

Figure 10: The constituent projection for split arguments in German

enough slots available: The number of argument slots in the Core is determined by the number of arguments available from the logical structure. In combination with the tight constraints on the information structure of this construction, the PrCS appears to be an ideal candidate (see figure 10).

There are a number of additional advantages of this analysis: By using two distinct RPs, we are immediately able to explain the change of the inflection class observed for quantifiers, since the trigger here is the lack of a seperate definiteness operator. At the same time, we predict that attributive adjectives change their inflection class as well - although this does not have any visible consequences due to the inflectional paradigm involved. The "split" induced by the topic/focus distinction appears to happen in the semantic representation: the complex representation of a single semantic argument is realized by two independant syntactic units.

If we look at case assignment, we find that only the focal part is in the argument slot of the predicate, as (67) shows:

(67) a. **Syntaktikern helfe ich dreien*
syntacticians.Dat help I three.Dat
'As for syntacticians, I help three (of them)'

b. *Syntaktiker helfe ich dreien*

"helfen" (to help) assigns dative case to its non-macrorole core argument. If the topical RP in this construction would also be part of said argument, it should receive dative case as well. As (67a) shows, this is not grammatical. Instead, the RP appears without an overt case marker[22]. This raises an interesting question: Is the topical RP indeed part of the logical structure of the clause? If so, where does it reside? We can argue that the RP in the PrCS is *not* subject to the completeness constraint, because it is not referential. Instead, we can license it by the immediate discourse context, similar to the analysis of discourse ellipsis proposed for Mandarin (cf. van Valin 2005, p.173ff).

The information partitioning exemplified by split topicalization is used in other constructions as well. However, the syntactic reflexes of this partitioning differ, as the next construction shows. Consider the superficially similar example sentences illustrated in (68):

(68) a. *Ich mag nur Kuchen mit Schlagsahne*
I like only cake with whipped cream
'I only like cake with whipped cream'

b. *Kuchen mag ich nur mit Schlagsahne*

(68a) shows the "normal" word order. Apparently, we can split the argument, put the topical element in the prefield and leave the focal restrictive modifier in the default focus position (68b). Such an analysis, however, would predict that a prepositional phrase in German can itself constitute a referential phrase. This prediction is incorrect, as shown by the contrast to attributive adjectives when the topical element is not realized in the sentence but instead inferred from discourse context:

(69) Q: *Was für Lieder hörst du?*
what for songs.Acc listen-to you
'What kind of songs are you listening to?'

[22]Presumably, the RP is in the nominative, but due to the inflectional paradigm this case cannot be differentiated from accusative in this construction.

A: (a) *Ich höre nur besonders lange*
I listen-to only especially long-ones
'I only listen to especially long ones'

(b) *[*] Ich höre nur von Peter Maffay*
I listen-to only of Peter Maffay
intended: 'I only listen to those of Peter Maffay'

If (68) were the same construction, (69b) should be grammatical. Since it is not, it follows that the focal constituent in (68b) is *not* a restrictive modifier for the topical RP. Changing the lexical elements paves the way towards a more suitable analysis:

(70) a. *??Ich lese nur Zeitungen mit Brille*
I read only newspapers with glasses
'It is only with newspapers with glasses that I read'

b. *Zeitungen lese ich nur mit Brille*
Newspapers read I only with glasses
'As for newspapers, I only read them with glasses'

c. *Ich lese Zeitungen nur mit Brille*
I read newspapers only with glasses
'As for newspapers, I only read them with glasses'

d. *??Zeitungen lese ich nur die mit Brille*
Newspapers read I only those with glasses
'As for newspapers, I only read those with glasses'

The problem in the original data are the multiple functions of "mit" (with) in German as well as extra-linguistic knowledge. Among other things, "mit" marks the instrumental case (70b,c) on an argument, as well as NP-internal relations (70a,d). As indicated, the NP-internal reading is semantically odd, but we can split the argument in two separate RPs nonetheless. It is our extra-linguistic knowledge about whipped cream and cake that leads us to assume that in (68b) we still have restrictive modification. Instead, as (70b) indicates, we simply have an additional oblique core argument. In fact, if we remove the focus particle from the sentences, they are a classic example of PP attachment ambiguities.

The commonality of the split argument construction and the PP fronting are captured by their information structure properties: the topical element appears in the prefield, while the focal element is put into the default focus position. For the split argument construction, this linear order is mandatory - fronting the PP is a choice the speaker makes:

(71) a. *Nur mit Schlagsahne mag ich Kuchen*

b. **Nur Chomsky kenne ich Syntaktiker*

c. **Nur grüne kenne ich Polizeiwagen*

d. **Drei kenne ich Bücher*

These word order facts cannot be accounted for by information structure alone, since in all four cases the focused element is in the prefield. The least we have to do is to state, in the constructional template for split arguments in German, that, if present, the topical element must be placed in the prefield. For this case, the use of the PrCS appears to be the simplest solution: It is in the right topological spot, and it supports only a single constituent.

5.5 The prefield and the case of apparent subordination

The next construction we want to take a look at revolves around predicates that take a whole proposition as their argument. Syntactically, this is typically reflected by a subordinated clause:

(72) *Der* *Peter* *sagte* *dass* *der* *Polizist* *dem* *Mann*
the Peter.Nom said that the police-officer.Nom the.Dat man.Dat
geholfen *hat*
helped has
'Peter said that the police officer has helped the man.'

The constituent projection that our proposal predicts for this sentence is given in figure (11). If we followed the PrCS model, the only variation would be that the

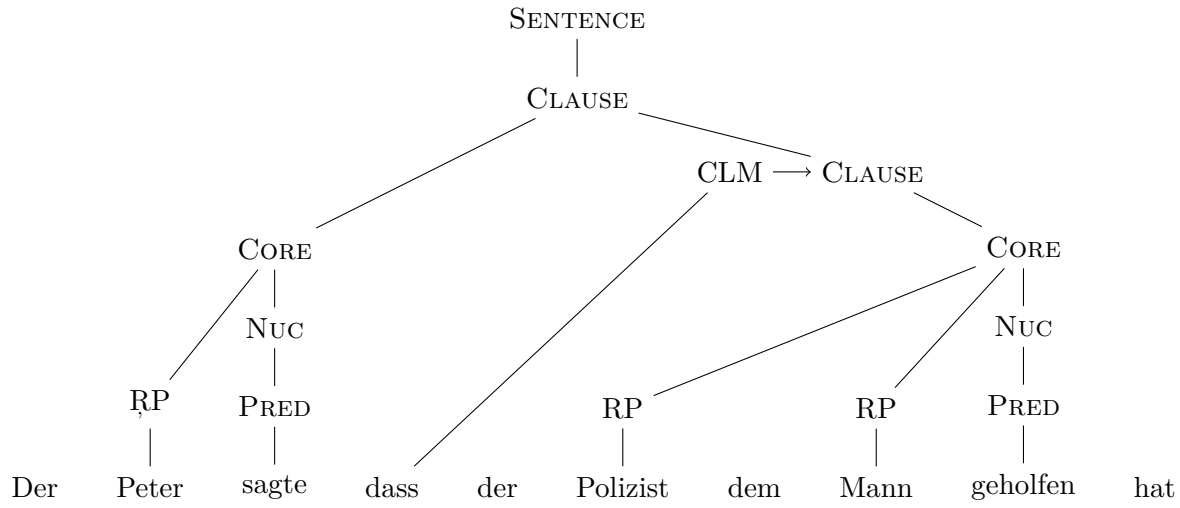

Figure 11: Constituent projection for an object-complement *dass*-clause

matrix subject was not part of the CORE but instead occupied the precore slot. Since we deal with a subordinated CLAUSE, long-distance *wh*-questions are possible (73):

(73) a. *Wem sagt der Peter dass der Polizist*
who.DAT said the Peter.DAT that the police-office.NOM
geholfen hat?
helped has
'Who did Peter say that the Policeman helped?'

b. ? *Wer sagte der Peter dass dem Mann geholfen hat?*

For these cases, having a PRCS accounts for the verb-second placement (since there is no room in front of the tense operator) and is consistent with the general treatment of *wh*-questions in RRG. For an explanation of the markedness of subject extraction in (73b), see (van Valin 1998). The constituent projection assigned by the PRCS model is, of course, identical. There is a variation on this kind of sentence formation that challenges a configurational analysis of the prefield: Instead of using a verb-final subordinated clause, we can also use a verb-second clause as an object complement. Historically, such cases of verb-second subordination have been marked on the verb

(74a), but the unmarked version is grammatical as well (74b).

(74) a. *Der Peter sagte der Polizist habe dem Mann geholfen*

b. *Der Peter sagte der Polizist hat dem Mann geholfen*

So far, this looks like a typical case of clausal subordination. However, in stark contrast to the *dass*-subordination above, displacement out of the second clause into the prefield is unmarked.

(75) a. *Der PoliZIST sagte der Peter hat dem Mann geholfen*

b. **Der PoliZIST sagte der Peter dass dem Mann geholfen hat*

c. *Dem MANN sagte der Peter hat der Polizist geholfen*

d. ??*Dem MANN sagte der Peter dass der Polizist geholfen hat*

The grammaticality judgements on extraction out of subordinated clauses in declarative sentences are my own, and appear to differ from those given by e.g. Frey (2006). There, we find the following example of extraction, which for him is grammatical and becomes, as he notes, a "perfect surface structure" (Frey 2006, footnote 1) once the embedded clause is extraposed:

(76) a. *Den Hans hat Karl dass Maria morgen*
the Hans.ACC has Karl.NOM that Maria.NOM tomorrow
treffen wird behauptet
meet will claimed
'Hans has Karl claimed that Maria will meet tomorrow' (= Frey 2006, (2b))

b. *Den Hans hat Karl behauptet dass Maria morgen treffen wird*

From my judgement, the deep embedding in (76a) is simply ungrammatical, while the extraposition variant in (76b) is only slightly marked. As van Valin (1998) shows, these contrasts in acceptability can be related to properties of information structure - an option the derivational account of extraction Frey (2006) uses does not have.

Semantically, there is no relation between the fronted element and the following verb, as (77) shows:

(77) *Morgen* *stand* *gestern* *in* *der* *Zeitung* *kommt* *der*
tomorrow was-written yesterday in the newspapers comes the
Weihnachtsmann
santa clause
'Yesterday, the newspapers wrote that tomorrow, Santa Clause is coming'

If there was a semantic relation, the only reading possible would be that the clause-initial adverbial defines the temporal origin of the second adverbial. In this interpretation, the proposition would refer to "today". The more likely interpretation, as indicated by the translation, is not blocked, though.

We can account for this difference by analyzing the verb-second examples in (75a,c) not as clausal subordination but as clausal cosubordination. Following the terminology of Jensen (2000), we will call this construction sentence intertwining, although the technical details of our analysis differ. To refer to the two clauses involved, we use "embedding clause" vs. "embedded clause", referring to the semantic relation between them. We know that we deal with two CLAUSEs, since both are marked for tense independently. The shared operator is illocutionary force, resulting in two clauses in the same potential focus domain. Therefore, elements of both clauses are equally available to displacement, explaining the unmarkedness of preposing the subject of the second clause (cf. (75a) vs. (75b)).

While it is evident that this construction is not clausal subordination, we need to justify why it is not clausal coordination. Is the illocutionary force operator indeed shared by both clauses? As it stands, we have no conclusive answer to this question. Our choice of cosubordination is the simplest option to explain that two clauses are sensitive to the same focus structure projection constituents.

An additional property of sentence intertwining is that the embedding clause can be placed more freely in the sentence than the subordination equivalent. It can always be realized immediately after the actual focus domain, but this is not

obligatory:

(78) a. *Dem Mann hat der PoliZIST sagte der Peter*
the man.DAT has the police-officer.NOM said the Peter.NOM
geholfen
helped
roughly: "Peter said it was the police officer that has helped the man."

b. *Der Polizist hat dem MANN sagte der Peter*
the police-officer.NOM has the man.DAT said the Peter.NOM
geholfen
helped
roughly: "Peter said it was the man that the police officer has helped."

c. *Dem Mann sagte der Peter hat der PoliZIST*
The man.DAT said the Peter.NOM has the police-officer.NOM
geholfen
helped
roughly: 'Peter said who helped the man was the police officer."

d. *Der Polizist sagte der Peter hat dem MANN*
the police-officer.NOM said the Peter.NOM has the man.DAT
geholfen
helped
roughly: 'Peter said who the police officer has helped was the man."

How can we explain the verb-second placement in both of the clauses? One option is to accept the PrCS, which for the sake of the argument at hand we will do for now. Verb-second placement is explained by a precore slot before the finite verb. Looking at (75c), this predicts the constituent projection in figure 12. The constituent projection would become even worse for the data in (78). Additionally, we lose the representation of the clausal cosubordination present in sentence intertwining. One way out of this would be to introduce an exception for this construction, allowing the precore slot to be empty in the embedded clause. This ad-hoc solution only solves part of the problem, since in (78a) the embedding clause *does* have a PrCS. Clearly, a configurational account of verb-second placement cannot handle sentence intertwining adequately.

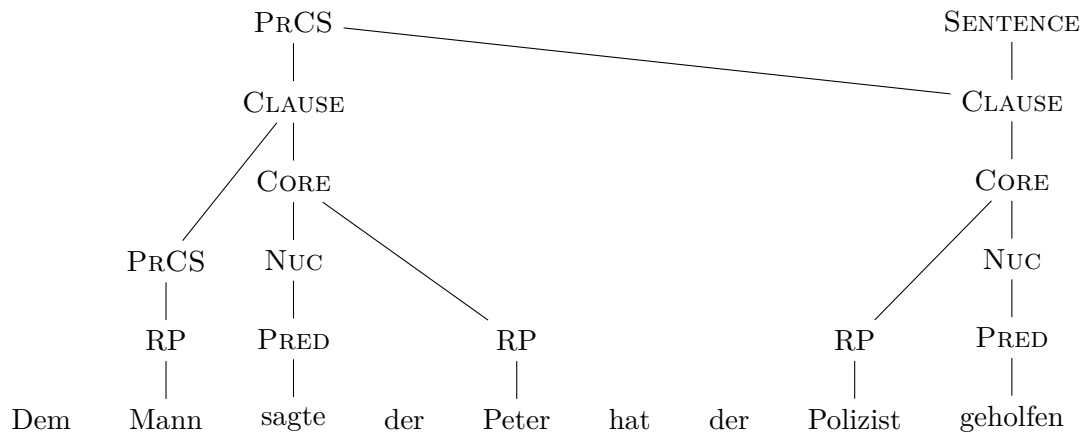

Figure 12: The constituent projection predicted by the PrCS model for sentence intertwining

The IU model proposed in this thesis explains the verb-second placement directly. The syntactic template we developed specifies that for each clause, the tense operator must follow an information unit. In sentence intertwining, both clauses share the illocutionary force operator and hence are in a single potential focus domain. This explains why a syntactic argument of the embedded clause can fill the prefield of the embedded clause: Not because it is a constituent of the latter, but because the focus structure projection of the two clauses are identical.

The example in (78a) allows us to be more precise: We assume that we can access information structure from the discourse context throughout the linking algorithm. When linking from semantics to syntax, we will have two Clause templates to fill. In both templates, we now assign an information unit to the slot before the tense operator. The information unit was computed from the semantic representation of the sentence. In example (78a), we choose the topical information unit for the embedded clause, and the focal information unit for the embedding clause. Due to constructional properties, we apply the cosubordination template, which joins the two previously independent potential focus domains at the level of juncture.

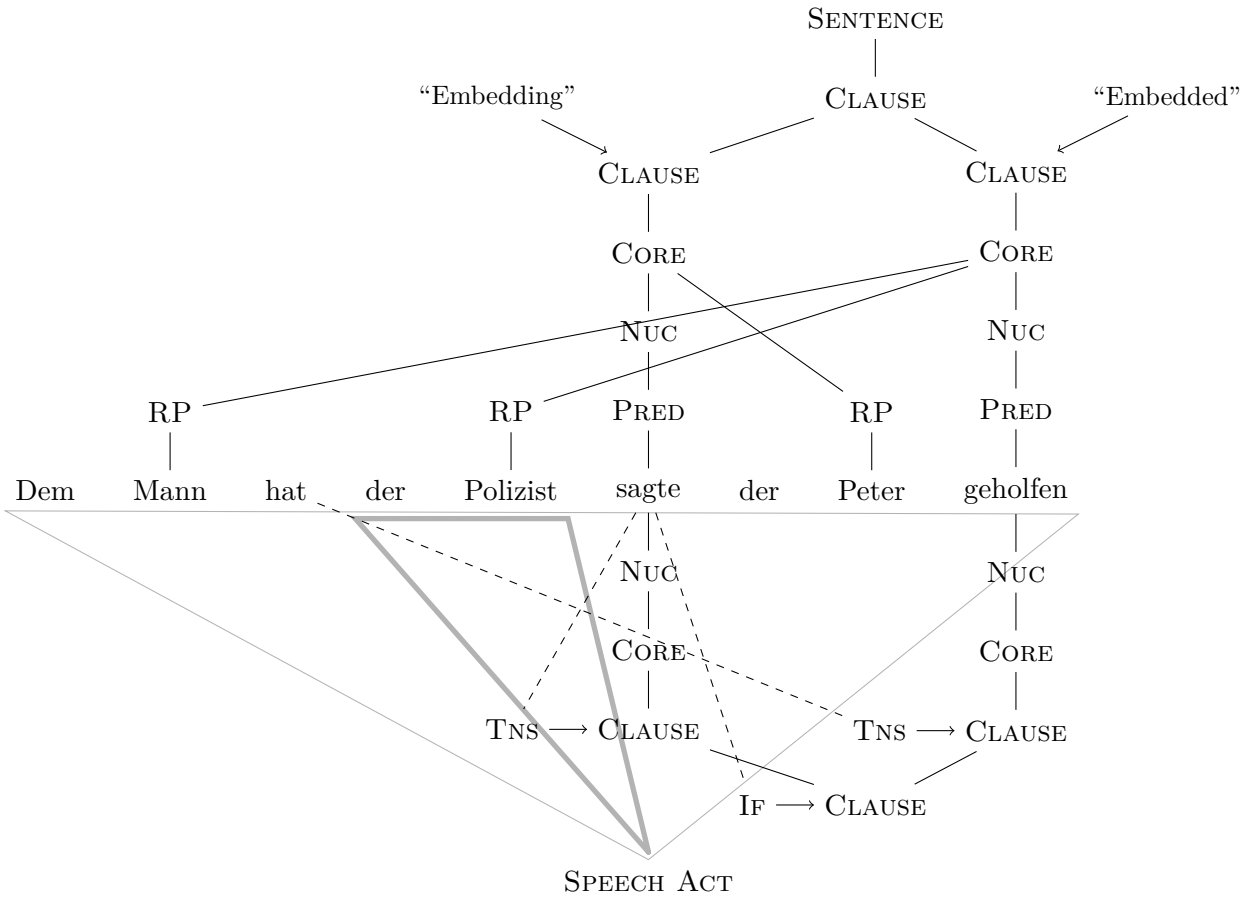

Figure 13: The syntactic structure of sentence interwining in German

Later on, the linearization of both clauses will result in the observed surface order. From the perspective of the embedded clause, the focal constituent is placed in the default focus position. Since the embedding clause is already committed to use it as its prefield, it intervenes between the predicate and its focal argument. The other arguments are realized as expected: the topical RP is already in the prefield of the embedded clause, and the core argument of the embedding predicate is placed in the only spot left free.

As it appears, the IU model in fact *predicts* sentence intertwining, as long as discontinuous Cores are possible. Figure 13 shows the syntactic structure of (78a).

6 Conclusion and Outlook

In this thesis we have attempted to argue for an account of the topology of German sentences that is based on the focus structure projection. While we focused on the prefield, our exploratory discussion of word order in German has shown that information structure relations also play a role in serializing the middlefield. At the same time, we were not able to formulate precise rules that could be incorporated in an RRG grammar for German. This topic remains controversial, but we think that the seperation of the CORE into two linearization domains is inevitable if one aims to account for a large set of data.

The CLAUSE template proposed, on the other hand, has proven to be a valuable component in analyzing more complex constructions in German. In contrast to a purely configurational account of sentence topology, it has proven vital to embrace the view of RRG that language is the result of multiple modules interacting with each other. One branch that we could not adequately follow in this thesis was the choice a speaker has for *wh*-questions. Sentence intertwining allows him to use discontinuity instead of extraction. In subordinated clauses, speakers can also opt to keep the *wh*-phrase in its CLAUSE and add a dummy *wh*-phrase in the matrix clause. Given the work that the linking algorithm has to perform to undo extraction, we could consider this a marked construction. Are there any unmarked non-local extraction constructions remaining?

Also, this thesis has been focused very heavily on the syntactic side of each construction. For sentence intertwining, it would be useful to uncover whether there are any discourse-semantical differences between subordination and sentence intertwining. How do they fit into the interclausal relations hierarchy?

Despite these unanswered questions, we have found valuable evidence *against* a configurational interpretation of the topological field model. Further work in this area will show which other traditional views on German clause structure need to be dropped to arrive at a sound analysis.

References

Bech, Gunnar (1983) *Studien über das deutsche Verbum infinitum*. Tübingen:Niemeyer.

Bresnan, Joan (2001) *Lexical functional syntax*. Oxford:Blackwell.

Büring, Daniel (1996) *Towards an economy-theoretic treatment of GermanMittelfeld word order*. Master's thesis, Universität Frankfurt und Köln.

Choi, Hye-Won (1999) *Optimizing Structure in Context. Scrambling and Information Structure*. CSLI:Stanford.

den Besten, Hans (1983) On the interaction of root transformations and lexical deletive rules. In *On the formal syntax of the Westgermania*, Werner Abraham, ed., Amsterdam:John Benjamins, 47–131.

Diedrichsen, Elke (2008) Where is the precore slot? Mapping the layered structure of the clause and German sentence typology. In *Investigations of the Syntax-Semantics-Pragmatics Interface*, John Benjamins:Amsterdam, 203–224.

Dipper, Stefanie & Heike Zinsmeister (2009) The role of the German Vorfeld for local coherence: A pilot study. In *From Form to Meaning: Processing texts automatically. Proceedings of the Biennial GSCL Conference 2009*, Christian Chiarcos, Richard Eckard de Castilho, & Manfred Stede, eds., Tübingen:Narr, 69–79.

Drach, Erich (1937) *Grundgedanken der deutschen Satzlehre*. Frankfurt am Main.

Frey, Werner (2006) How to get an object-es into the german prefield. In *Form, structure and grammar - a Festschrift presented to Günther Grewendorf on occasion of his 60th birthday*, Berlin:Akademi Verlag, 159–185.

Frey, Werner & Karin Pittner (1998) Zur Positionierung der Adverbiale im deutschen Mittelfeld. *Linguistische Berichte* **176**: 489–534.

Grosz, Barbara, Aravind K. Joshi, & Scott Weinstein (1995) Centering: A Framework for modelling the local coherence of discourse. *Computational Linguistics* **21**: 203–225.

Höhle, Tilman N. (1986) Der Begriff 'Mittelfeld': Anmerkungen über die Theorie der topologischen Felder. In *Kontroversen, alte und neue: Akten des 7. Internationalen Germanistenkongresses*, A. Schöne, ed., Tübingen:Niemeyer, 329–340.

Jensen, Anne (2000) Sentence intertwining in danish - a challenge to the RRG account? URL `http://linguistics.buffalo.edu/people/faculty/vanvalin/rrg/vanvalin_papers/Jensen.pdf`.

Kathol, Andreas (2000) *Linear Syntax*. Oxford:Oxford University Press.

Kretzschmar, Franziska (2006) *Zum expletiven und pronominalen es im Deutschen. Syntaktische, semantische und varietätenspezifische Aspekte*. Master's thesis, Fachbereich Germanistik und Kunstwissenschaften der Philipps-Universität Marburg.

Lambrecht, Knud (1994) *Information structure and sentence form*. Cambridge:Cambridge University Press.

Lambrecht, Knud (2000) When subjects behave like objects: a markedness analysis of sentence focus constructions across languages. *Studies in Language* **24**: 611–682.

Lenerz, Jürgen (1977) *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen:Narr.

Lenerz, Jürgen (2001) Word order variation: competition or co-operation? In *Competition in Syntax*, Gereon Müller & Wolfgang Sternefeld, eds., Berlin: de Gruyter, 249–281.

Maling, Joan & Annie Zaenen (1981) Germanic Word Order and the Format of Surface Filters. In *Binding and filtering*, F. Heny, ed., London:Croom Helm, 255–278.

Müller, Stefan (2003) Mehrfache Vorfeldbesetzung. *Deutsche Sprache* **31**(1): 29–62.

Müller, Stefan (2005) Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* **203**: 297–330.

Müller, Stefan (2007) *Head-Driven Phrase Structure Grammar. Eine Einführung*. Tübingen: Stauffenburg Verlag.

Musan, Renate (2002) Informationsstrukturelle Dimensionen im Deutschen. Zur Variation der Wortstellung im Mittelfeld. *Zeitschrift für Germanistische Linguistik* **30**: 198–221.

Pafel, Jürgen (2009) Zur linearen Syntax des deutschen Satzes. *Linguistische Berichte* **217**: 37–79.

Pollard, Carl & Ivan A. Sag (1994) *Head-driven phrase structure grammar*. CSLI, fifth ed.

Prince, Alan & Paul Smolensky (1993) Optimality Theory - Constraint Interaction in Generative Grammar. Tech. rep.

Reis, Marga (1980) On justifying Topological Frames: "Positional Field" and the Order of Nonverbal Constituents in German. *Documentation et Recherche en Linguistique Allemande Contemporaine* **22/23**: 59–85.

Rizzi, Luigi (1997) The fine structure of the left periphery. In *Elements of Grammar*, Liliane Haegeman, ed., Kluwer Academic Publishers, 281–337.

Speyer, Augustin (2004) Competing Constraints on Vorfeldbesetzung in German. In *Proceedings of the Dislocated Elements Workshop*, Benjamin Shaer, Werner Frey, & Claudia Maienborn, eds., no. 35 in ZASPiL, 519–541.

Speyer, Augustin (2008a) Doppelte Vorfeldbesetzung im heutigen Deutsch und im Frühneuhochdeutschen. *Linguistische Berichte* **216**: 455–486.

Speyer, Augustin (2008b) German Vorfeld-filling as constraint interaction. In *Constraints in discourse*, Anton Benz & Peter Kühnlein, eds., Amsterdam:Benjamins, 267–290.

Speyer, Augustin (2009) Das Vorfeldranking und das Vorfeld-es. *Linguistische Berichte* **219**: 323–353.

Speyer, Augusting (2007) Die Bedeutung der Centering Theory für Fragen der Vorfeldbesetung im Deutschen. *Zeitschrift für Sprachwissenschaft* **26**: 83–115.

van Valin, Jr., R. D. (1998) The acquisition of wh-questions and the mechanisms of language acquisition. In *The new psychology of language*, Michael Tomasello, ed., London:Lawrence Erlbaum Associates, 221–249.

van Valin, Jr., Robert D. (2005) *Exploring the syntax-semantics interface*. Cambridge University Press.

van Valin, Jr., Robert D. (2008) RPs and the nature of lexical and syntactic categories in Role and Reference Grammar. In *Investigations of the Syntax-Semantics-Pragmatics Interface*, Jr. R. D. Van Valin, ed., Amsterdam:Benjamins, 161–178.

van Valin, Jr., Robert D. & Elke Diedrichsen (2006) Bonsai Grammar for German (a fragment of a Role and Reference Grammar-based analysis of German syntax). ms. (http://www.coli.uni-saarland.de/∼tania/CMGD/RRG_BonsaiGrammarGerman.pdf).

van Valin, Jr., Robert D. & Randy J. LaPolla (1997) *Syntax. Structure, meaning and function.* Cambridge University Press.

Vendler, Zeon (1967) *Linguistics in philosophy.* Ithaca:Cornell University Press.

Vikner, Sten (1995) *Verb Movement and Expletive Subjects in the Germanic Languages.* Oxford University Press.

Walker, Marilyn A., Aravind K. Joshi, & Ellen F. Prince (1998) Centering in naturally occurring discourse: an overview. In *Centering Theory in Discourse*, Marilyn A. Walker, Aravind K. Joshi, & Ellen F. Prince, eds., Oxford University Press, 1–28.

Zwart, Jan-Wouter (1993) *Dutch syntax. A minimalist approach.* Ph.D. thesis, University of Groningen.