RIGHT NODE RAISING
IN PARSING AND GENERATION

Jonas Kuhn, Christian Rohrer and Sina Zarrieß

Universität Stuttgart

Proceedings of the LFG10 Conference

Miriam Butt and Tracy Holloway King (Editors)

2010

CSLI Publications

http://csli-publications.stanford.edu/

Abstract

We discuss a grammar engineering account of non-constituent coordination in the clausal syntax of German as it occurs in real corpus data, generalizing the approach of Forst and Rohrer (2009) to right node raising cases not captured so far. We compare our account to the unimplemented metagrammatical technique by Maxwell and Manning (1996) and point out that many of their ideas can be captured using only the standard LFG formalism and making minimal adjustments to an existing large-scale grammar. We show that information-structural constraints on right node raising can be simulated in f-structures and note that they are well-behaved in a generation scenario.

1 Introduction

Syntactic coordination phenomena above the straightforward levels of NPs and PPs continue to be a challenge for broad-coverage parsing (and generation) systems. As the corpus study in Forst and Rohrer (2009) showed, a non-negligible proportion of German corpus instances involving coordination at the verbal/clausal level are cases of "non-constituent coordination", including a fair number of right node raising instances. (1) is a list of examples from Maxwell and Manning (1996) indicating the different constructions typically assumed under the umbrella term non-constituent coordination. This paper focusses on the phenomena in (1a-b).

(1)	a.	*Conjunction reduction*	Bill gave the girls spades and the boys recorders.
	b.	*Right node raising*	Bill likes, and Joe is thought to like cigars from Cuba.
	c.	*Gapping*	Bill gave a rhino to Fred, and Sue a camera to Marjorie.
	d.	*Ellipsis*	Bill likes big cars, and Sally does too.
	e.	*Non-symmetric coordination*	Bill went and took the test.

Overall, each variant of these coordination constructions is in the low frequency spectrum in real corpus data, which explains in part why broad-coverage parsing approaches tend to ignore the problem. For statistical approaches, the training data for the numerous variants can be expected to be too sparse to support the picking up of effective generalizations; moreover, the effort to make a parser capable of dealing with the phenomena is unlikely to pay off in terms of a significant gain in evaluation results on standard data sets.

Rule-based approaches on the other hand could in principle address these issues at the right level and capture the morphosyntactic constraints that drive the

*The work reported in this paper was supported by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) in SFB 732 *Incremental specification in context*, project D2 (*Combining Contextual Information Sources for Disambiguation in Parsing and Choice in Generation*; principal investigators: Jonas Kuhn and Christian Rohrer).

phenomena. LFG, with its elegant set-based representation of coordination phenomena (Kaplan and Maxwell, 1989/95), is a particularly promising framework for such an approach. Based on a straightforward c-structure treatment of coordinations, each conjunct contributes its f-structure information in the form of an element of a set, following the annotation in the schematic coordination rule in (2); information from "outside" the coordination construction is then simply distributed over the set elements, making sure, for instance, that *Bill* in (3) ends up as the f-structural subject of the two f-structures for *sing* and *play*.

(2) $\text{XP} \rightarrow \underset{\downarrow \in \uparrow}{\text{XP}} \quad \text{CONJ} \quad \underset{\downarrow \in \uparrow}{\text{XP}}$

(3) Bill sings and plays the guitar.

Moreover, with Maxwell and Manning (1996), there exists the outline of a metagrammatical approach (on top of standard LFG parsing) to dealing with certain non-constituent coordination phenomena, which extends ideas of the standard set-based coordination analysis. (Compare Milward (1994) for a discussion of different computational strategies for non-constituent coordination.)

But although LFG provides a good basis for tackling non-trivial coordination cases from real corpus data in a broad-coverage approach, it turns out that it is very hard to tell ahead of time whether a treatment of certain low-frequency non-constituent coordination phenomena will be more helpful than harmful. This may sound counterintuitive, but it has to do with the trade-off one has to make in grammar engineering between goals that are not fully compatible in practice, as sketched in Figure 1. Linguistic quality will normally warrant an advantage in engineering goals (e.g., it is easier to maintain, extend or adapt a systematic analysis; hence, one can over time achieve higher coverage if the analyses chosen are linguistically valid and general). With low-frequency phenomena however, adding a linguistically appropriate analysis to the grammar typically has the effect that the search space for many other constructions is also increased. This can lead to a loss of efficiency – not just for the critical data, but potentially for every input sentence. In any practical system, this will have an indirect negative effect on the system's

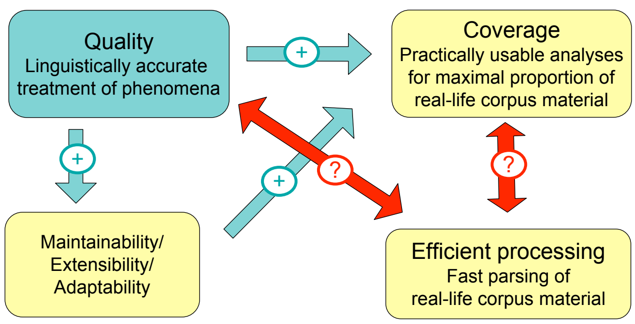


Figure 1: Trade-off between grammar engineering goals

coverage, for the following reason: in parser application, there is a maximal time that the application can wait for a parsing result, the timeout period. If a change in the grammar increases the search space, this can mean that more sentences will time out. This way, the gain from increased linguistic quality in the analysis of a low-frequency phenomenon may be outweighed by the increase of timeouts in unrelated material.

The treatment of coordination (above the NP/PP level) is a particularly tricky area in grammar engineering, because any extension or relaxation of the c-structure rules will readily interact with a large number of other analyses: any added ambiguity may extend the search space dramatically. In the present study, we discuss recent advances that we made by careful conservative extensions to the rule set of the German ParGram grammar (Butt et al., 2002; Dipper, 2003) in order to provide an appropriate analysis for the bulk of right node raising examples occurring in real corpus data. This work is an extension of Forst and Rohrer (2009).

Our analysis is inspired by Maxwell and Manning's (1996) stack-of-automata approach, but is formulated in terms of standard LFG c-structure rules. We take advantage of the fact that an account of German clausal structure following the endocentric principles of Bresnan (2001) goes along with fairly compact rule networks, such that the effect of Maxwell and Manning's idea can be emulated reasonably well on the basis of independently motivated c-structure rules.

We believe that the complex of interacting factors that have to be dealt with in an implementation of non-constituent coordination is typical of the situation in advanced grammar engineering work. The existence of a highly developed engineering platform like XLE and long-term efforts like the ParGram project have brought about relatively mature broad-coverage grammars for which the trade-off sketched in Figure 1 can be addressed at a more realistic level than it was possible a few years ago. The present contribution is meant to convey some of the considerations that have to be taken in this trade-off. We make no claims that the proposed analysis is linguistically justified outside the grammar engineering context; but we do believe that the view from the grammar engineering perspective can give some inspiration for a realistic linguistic account.

Section 2 introduces the technique of Maxwell and Manning's (1996) account in some detail, since it is important for the discussion of our approach. Section 3 introduces the strategy of Forst and Rohrer's (2009) approach, extending it to cover a broader set of non-constituent coordination cases. In section 4, we briefly discuss an overgeneration issue and the role of information structure in restricting the choices. Section 5 offers a short conclusion.

2 Maxwell and Manning's (1996) approach

The coordination constructions usually called "conjunction reduction" and " right node raising" (or "left-deletion") can be characterized as follows: Material which could in principle be realized twice and which would occur at the left or right periphery in both conjuncts, occurs only a single time (on the left edge of the left-most conjunct, or on the right edge of the right-most conjunct).

There are (at least) two views on this type of data. They could be seen as a result of deleting identical material from the left or right conjunct (or in less transformational terms, as the result of some economy principles), as sketched in (4). Or they can be viewed as coordination of special units that are not normally considered constituents, illustrated in (5), and that can be completed by the identical (left or right) material (which does not have to form a canonical constituent either).

(4) a. Bill [[gave the girls spades] and [~~gave~~ the boys recorders]].

b. [[Bill likes ~~cigars from Cuba~~] and [Joe is thought to like cigars from Cuba]].

(5) a. Bill gave [[the girls spades] and [the boys recorders]].

b. [[Bill likes] and [Joe is thought to like]] cigars from Cuba.

Putting aside the issue of the unjustified non-constituent units, the latter approach is very much in the spirit of the standard LFG treatment of coordination (Kaplan and Maxwell, 1989/95). Since the material occurring just once (shown underlined in (5) and all following examples) is outside the technical coordination, the usual distribution mechanism will fire, in the same way as the subject is distributed for (3).

2.1 C-structure coordination as a meta-grammatical "process"

Creating rules for all possible non-constituent units just for the purpose of coordination is of course an unattractive prospect. This is where the main idea of Maxwell and Manning's (1996) "meta-grammatical" account comes in: they posit a general mechanism, living outside the grammar proper, which provides the required units for the purposes of non-constituent coordination, as needed.

We use the graphical illustration in Figure 2 to suggest that partial versions of the regular constituents are involved when analyzing such coordinations. A partial constituent is required both on the outside of the coordination (the beginning of the VP triangle above "gave"), and inside the coordination – here in two copies, the two hatched parts of the VP triangle for "the girls spades" and "the boys recorders".

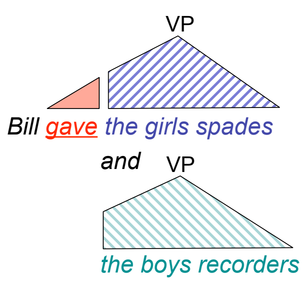
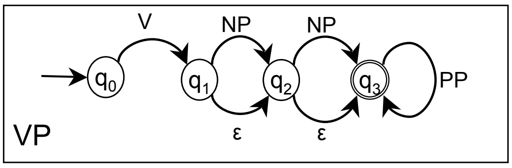


Figure 2: Non-constituent analysis for (4)

C-structurally, the outside and (each of the) inside parts have to complement one another to form a regular constituent. In the LFG parser, the right-hand side

(RHS) of a c-structure rule (which is a regular expression that may contain a Kleene star, optionality brackets etc.) is defined as a transition network or finite-state automaton. Maxwell and Manning ensure the correct assembly of partial constituents in the parser by keeping track of how much of the RHS of a c-structure rule has been consumed in terms of the states of the RHS automaton. The automaton for the VP rule in (6) is given in (7).

(6) VP → V (NP) (NP) PP*

(7)


In Figure 2, the coordination comprises the c-structure material starting after the initial V, so the subautomaton from the initial state q_0 to q_1 is used only once, outside of the coordination, whereas each of the conjuncts covers the subautomaton from q_1 to the final state q_3.

Technically, the analysis of non-constituents is assumed to be a meta-grammatical "process" (which can be perfectly viewed in a non-derivational sense, just like automata are not a derivational model in plain LFG) that can pass a sub-automaton of a rule twice (or more times), using the automaton states to keep track of the rule coverage. The meta mechanism will also ensure that the conjuncts are separated by (commas and) a conjunction like *and/or*. The abstract process driving the analysis in Figure 2 is sketched in the flow-chart in Figure 3, suggesting a "parallel" analysis of the two conjuncts, covering the stretch of the automaton between q_1 and q_3.

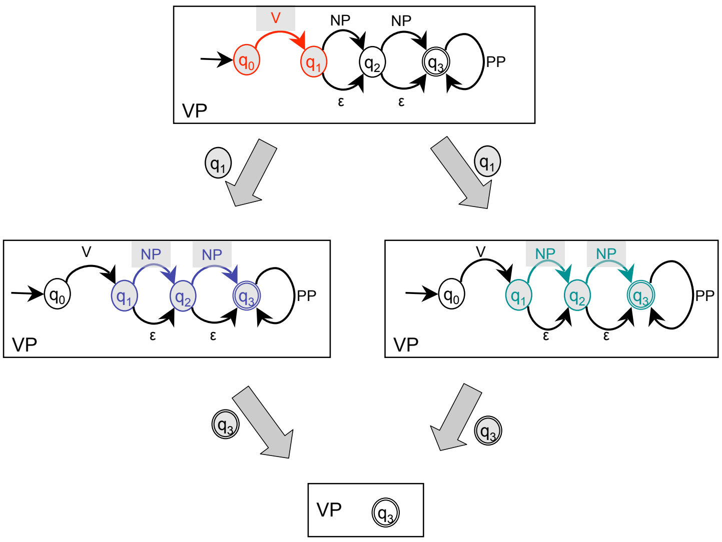
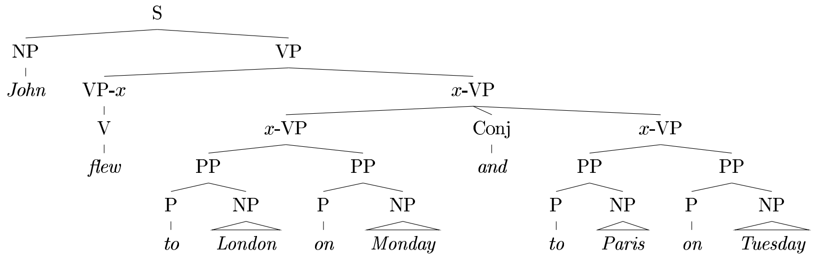


Figure 3: "Split" parsing process for Figure 2

Maxwell and Manning use the notation in (8) to draw the resulting c-structures. The internal automaton states are indicated by variables $x, y, \ldots$, which have to match across conjuncts and with the material outside of the coordination. In the resulting c-structures, subtrees do not necessarily correspond to constituents licensed by the LFG base grammar – so the coordination of "non-constituents" can be derived in an elegant way, without altering the grammar proper.

(8)

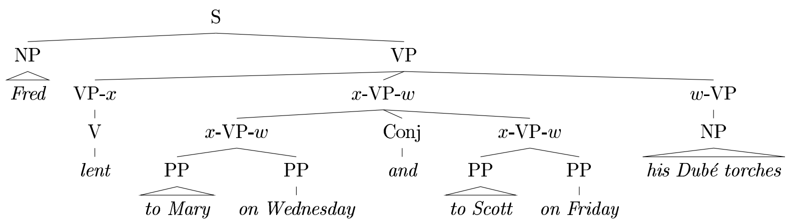


Since the mechanism keeps track of subautomata covered, the two conjuncts do not have to include the same categorial material, nor the same number of (sub-)constituents, as long as the alternatives are each licensed by a rule automaton. This licenses cases like (9).

(9) You can call me [[directly] or [after 3pm through my secretary]].

Monitoring of partial rule consumption can also occur on both sides of the coordinated sub-structures, as is shown in (10):

(10)



This example could not be derived with the VP rule/automaton in (6)/(7), since it involves a shift of the object NP *his Dubé torches*. But based on a more general VP rule automaton as in (11), the example could be analyzed as in Figure 4. Note that on both ends of the original VP rule, material occurs only a single time (and is thus underlined in our notation), whereas the middle part between q_1 and q_4 is duplicated.

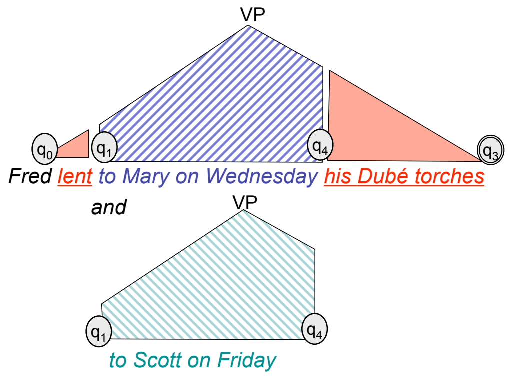
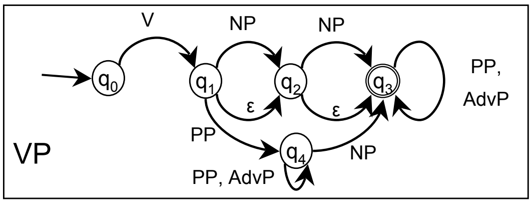
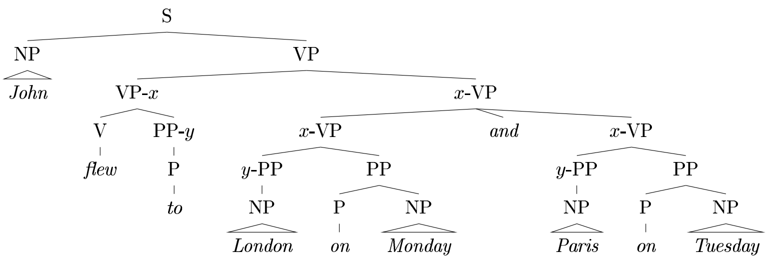


Figure 4: Analysis of (11) keeping track of states

(11) 

Examples like (12) suggest that the mechanism is even more general. Here, the non-coordinated material *flew to* does not originate from a single rule automaton, but essentially cuts across hierarchically embedded LFG rules.

(12) 

The flow-chart of automaton calls in Figure 5 illustrates how the mechanism proceeds. The dotted box at the top displays the parsing status after reading *flew to*, having consumed the beginning of the VP, plus part of the PP. The rest of the PP and the continuation of the VP (using the PP loop from q_3) is again processed in parallel for the two "conjuncts" *London on Monday* and *Paris on Tuesday*, before

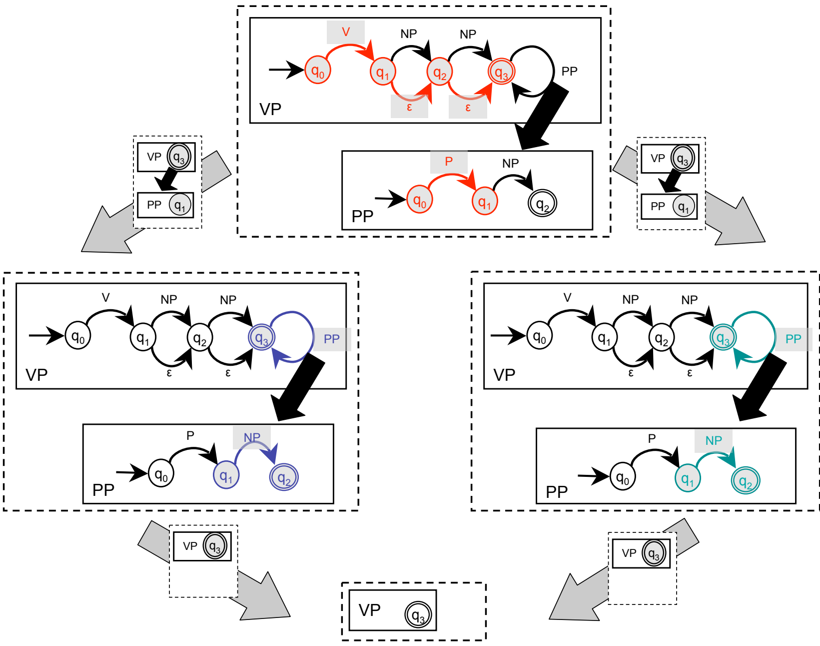


Figure 5: Coordination meta-mechanism for (12)

jumping back to the final state of the VP rules.

Note that for such hierarchically embedded cases, more than a single automaton state is needed for keeping track of the consumption status. In Figure 5, this is indicated by the boxes labelling the gray flow-chart arrows, which can include state recordings from several embedded rule automata. Maxwell and Manning assume a *stack* of automaton states from (potentially) different rules as the representation controlling the meta mechanism.

2.2 F-structure treatment and discussion

So far, we did not mention the f-structure treatment that goes along with the meta-grammatical mechanism of non-constituent coordination. The mechanism was designed the way it was so that the standard LFG coordination f-annotation ($\downarrow \in \uparrow$ on each of the conjuncts) can be directly used. So, in addition to keeping track of c-structure rule consumption, the meta-mechanism will simply add this annotation to each conjoined "pseudo constituent" arising from the process.

We note that Maxwell and Manning's (1996) approach provides a very elegant solution to the main problems that conjunction reduction and the right-node raising phenomena pose: no unjustified pseudo constituents have to be added to the grammar since the meta-grammatical mechanism generates the partial constituents on the fly, as needed – both for the coordinated material and the "outside" material occurring only once. F-structural distribution of the outside material follows the standard LFG coordination treatment: since the $\downarrow \in \uparrow$ annotations coerce the f-structures involved in the coordination to be set elements, the standard "across the board" distribution mechanism is triggered (at least in the simple configurations).

To our knowledge, Maxwell and Manning's coordination approach has never been implemented on a larger scale (Zarrieß and Seeker (2008) present an experimental implementation which uses finite-state technology to compile an XLE grammar into a version following Maxwell and Manning's approach). This presumably has to do with a fundamental practical issue of the approach: when the coordination mechanism is assumed to interact with the rule automata as they are used in large-scale LFG grammars, a combinatorial explosion of potential "rule splitting" points will arise – in particular through the hierarchical rule stacking. To see this intuitively, note that for a typical attachment ambiguity example like *John read a book and an article on Semantics*, the meta-grammatical mechanism would add the possibility of coordination of partial higher-level constituents; in this case, *a book* could be a partial VP, possibly open for taking in more material at the embedded NP level (namely, the PP *on Semantics*), besides numerous other splitting alternatives.

Moreover, the combinatory possibilities of rule splitting are multiplied by alternatives in the plain c-structure rules. The VP automaton in (11), which would be required to capture example (10), is still a very reduced sample automaton. Yet it gives an idea of the degree of optionality that broad-coverage rules have to offer at the level of c-structure (compare Kuhn and Rohrer (1997)). With the constraining role of f-structure this is not a practical problem in standard LFG, as the impressive parsing speed of the XLE system shows. However, pairing this c-structure level optionality up with a highly liberal, c-structure based non-constituent coordination approach is likely to be computationally intractable.

In fact there seems to be a tendency – at least for the broad-coverage grammar of German – that due to optionality of most rule parts (essentially following the principle-based approach of Bresnan (2001)), many argument phrases and sequences of argument phrases can legally be analyzed, for instance, as VPs (without the verb being realized). It is the broader c-structure rule context for using such VPs, plus f-structure that constrains this c-structural freedom, striking a balance between coverage of real corpus material and efficient processing. Allowing for arbitrary partial constituents to capture non-constituent coordination would almost definitely break this balance.

Independent of these complexity problems, there is an issue with the apparent simplicity of Maxwell and Manning's coordination approach that tends to escape attention when one first looks at it. What we said above about the f-structure treatment and the automatic effect of distributing "outside material" over the conjunct f-structures is actually not quite true. In a hierarchical-stacking situation, the outside material would not receive the correct f-structural interpretation if the standard f-annotations were not changed. In case of asymmetrically "deep" coordinations the f-annotation will even need to contain a functional uncertainty equation. Consider example (13):

(13) [[John drove to] and [Sue is planning to bike to]] San Francisco.

In order for *San Francisco* (as part of a partial x-VP with an embedded partial y-PP) to receive the correct interpretation, the f-annotation has to introduce an optional XCOMP embedding in order to capture the situation in both conjuncts. Similar examples with different embeddings are possible too.

Since the standard f-annotation of PP in the VP rule does not come with this functional uncertainty, it turns out that the meta-grammatical coordination mechanism does have to make some more substantial alternations to the plain LFG grammar, which are motivated in a purely technical way.[1]

3 Generalized constituent approach

Since the direct implementation of Maxwell and Manning's approach is not an option for broad-coverage grammar engineering, we discuss here a feasible alternative strategy for dealing with non-constituent coordination in the context of the German ParGram grammar. The goal is to find an analysis that avoids introducing new special "constituents" as much as possible, but instead to use independently

[1]This particular aspect of the f-structure distribution problem is not discussed in the (Maxwell and Manning, 1996) paper.

motivated rules for the coordination case, making only very few special additions to rules.

The treatment of a number of VP coordination cases was already proposed in Forst and Rohrer (2009) (see Section 3.1). In the present paper, we discuss the underlying strategy and extend it to a number of additional cases of right node raising (Section 3.2).

3.1 Coordination configurations in the verbal projection spine

In the remainder of this paper, we discuss modifications of the German ParGram grammar. Since a discussion of changes in the grammar code would be hard to convey, we use schematic illustrations of the c-structure analysis of the clausal spine to highlight the most important aspects of the analysis.

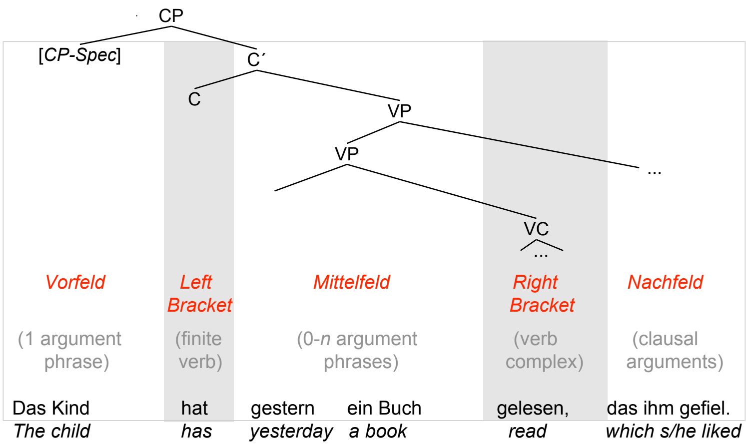


Figure 6: Schematic illustration of German clausal syntax

Figure 6 illustrates the basic clause structure of German, overlaying the classical descriptive field model over a c-structure scheme which by and large follows Bresnan's (2001) endocentric schema. We use the shaded areas to highlight (i) the left verbal bracket, containing the complementizer in embedded clauses and the finite verb in main clauses, and (ii) the right verbal bracket, containing the verb complex, with the least embedded verb appearing at the right edge. (In main clauses containing just a single verb, the verb complex will be empty.) The two verbal brackets delimit the *fields* in the German clausal structure: *Vorfeld, Mittelfeld, Nachfeld*.

The problem posed by typical VP coordination cases in corpus data is illustrated by the two examples in Figure 7: Under the analysis chosen for the German clause structure in the grammar, both cases involve conjuncts not corresponding to constituents in the grammar. In both cases, it would be possible to modify the clausal analysis in a way that would turn it into a constituent coordination: for the first example, a strictly right-branching VP analysis could be assumed; for the second example, the verb complex analysis that bundles up all verbs under VC could be replaced with a recursive VP embedding analysis. The problem is however that

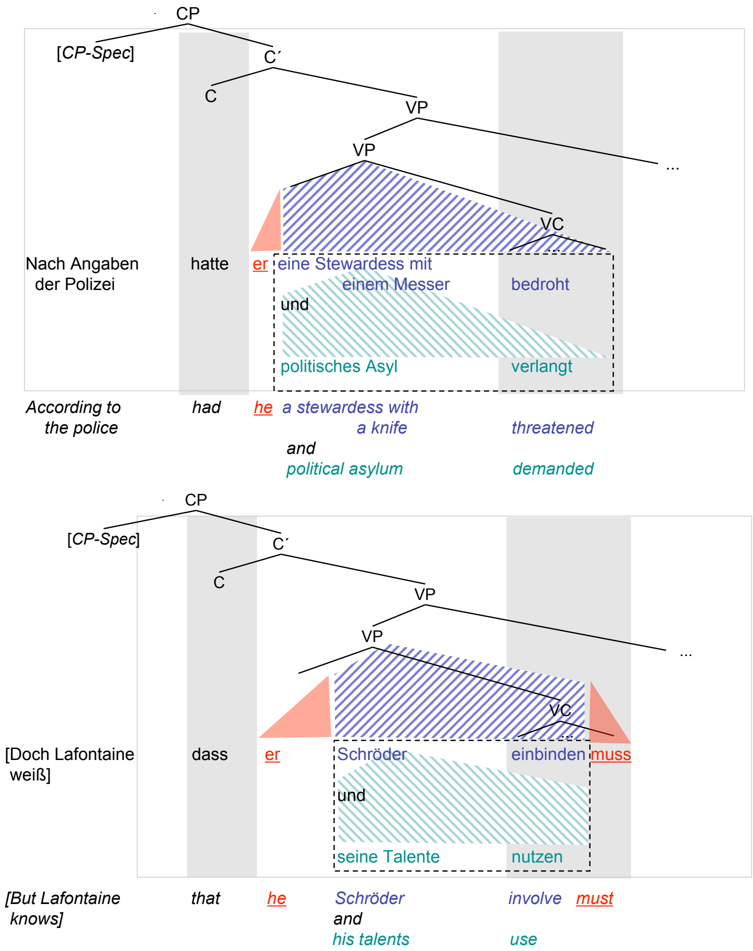


Figure 7: VP coordination issues involving coordinated verbs

the analyses would lead to spurious ambiguity in all non-coordination examples, if they were allowed along with the standard analysis. Figure 8 is a case where the coordinated material does not include any verb; hence, none of the above options would work.

Forst and Rohrer (2009) propose a solution for all three cases, following the strategy of making minimal modifications to the standard clausal backbone, specifically for the coordination situation. So, the rules are not relaxed in arbitrary situations but only in the presence of the coordinating conjunction. This guarantees that the search space is extended very carefully and the analysis stays computationally manageable.

The generalized constituent analysis simply allows the conceivable variants of the standard clausal analysis mentioned above *for the coordination case*. This way, the two “non-constituent coordination” cases from Figure 7 receive the constituent coordination analysis in Figure 10 with minimal grammar adjustment. The coordination-specific use of categories (here, VP) is indicated by the dotted box.

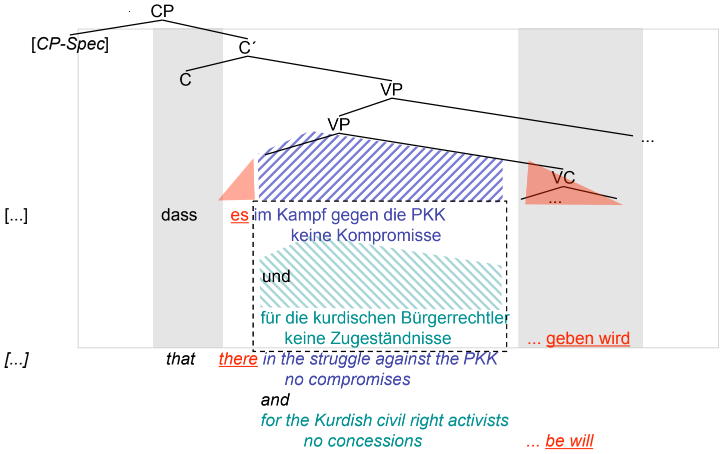


Figure 8: A VP coordination issue with a "raised" verb

Figure 8 is the only type of construction for which a new special category has to be posited – combining a sequence of two or more verbal arguments (or adjuncts) into a "pseudo-constituent" VPargs, as shown in Figure 9. (One could also view this pseudo-constituent as a VP in which the optional V is not realized – but calling it by a special name has the advantage that its use can be heavily restricted.)

When we compare the "generalized constituent approach" to Maxwell and Manning's, we note that since (a) the combination of verbal constituents typically makes them f-structure co-heads, and (b) most material in c-structure rules is actually optional, the effect of combining partially consumed rules (from Maxwell and Manning's mechanism) can be simulated quite well by combining several instances of "full" verbal constituents in which the optional subparts are left unrealized, just as required. In particular, the material outside of the coordination uses canonical rules with standard f-annotations in all cases that we have seen so far (obtaining

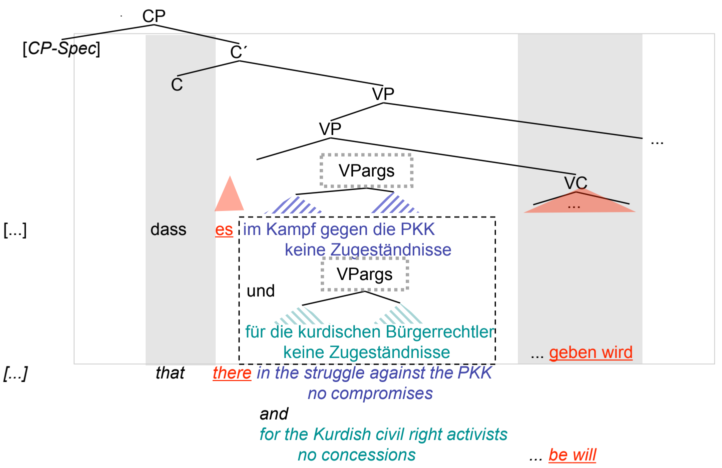


Figure 9: Forst and Rohrer's (2009) solution for Figure 8

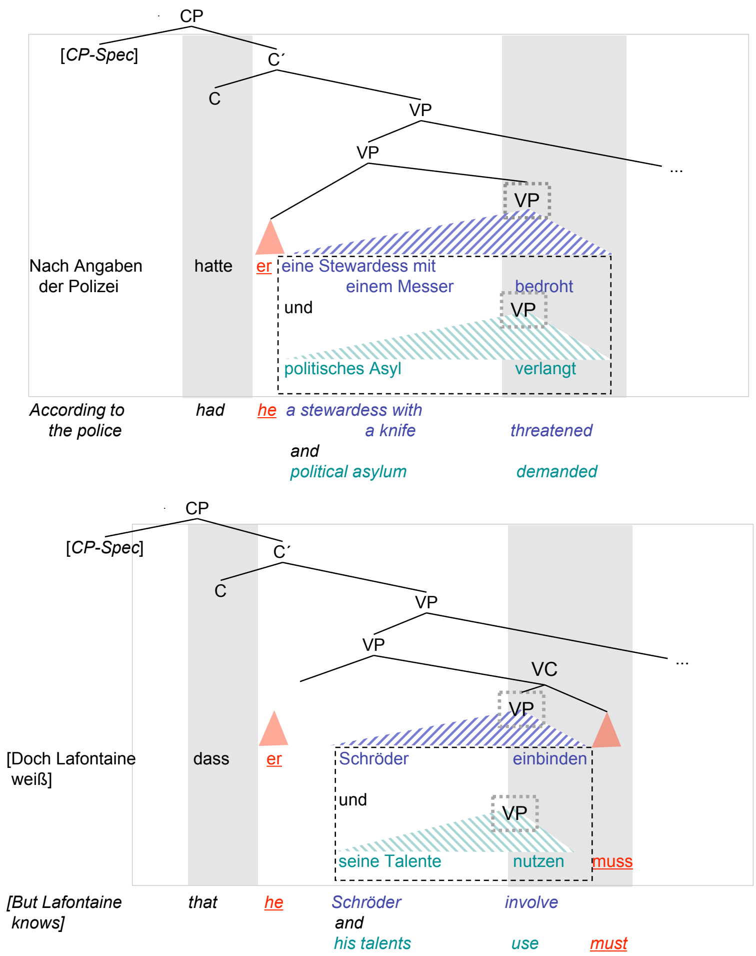


Figure 10: “Generalized constituent analysis” for Figure 7

the f-structure distribution effect with the standard LFG mechanism of coercing the coordinated f-structure into a set).

The strict bookkeeping of c-structural splitting points of Maxwell and Manning (through the stack of states) is lost in this simulation. However, with rules similar to the rule scheme “XP $\rightarrow$ { YP | ZP | ... }* (X)” this is not too critical in practice: There are many alternative paths between the rule automaton states, and f-structural constraints play the most important role in constraining the options of combination.

Of course the generalized constituent approach will only cover cases of non-constituent coordination that were foreseen in the coordination-specific relaxation of the rules. Given the subtle balance between coverage and efficiency (through limited search space), this is actually a desirable property of the account.

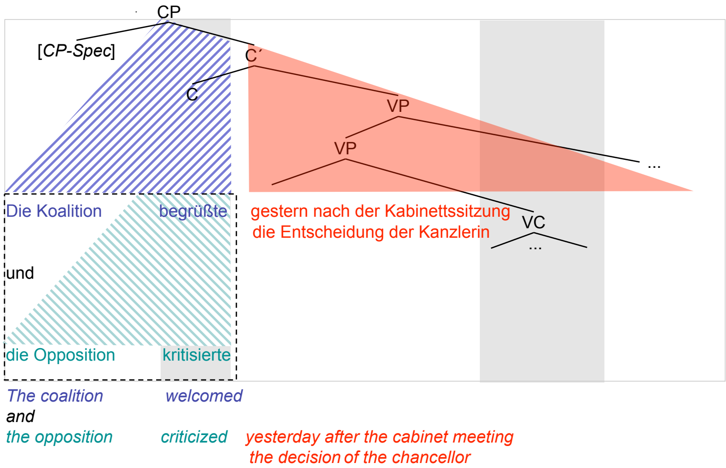


Figure 11: Coordination problem involving "low" non-coordinated material

3.2 Extending the generalized constituent strategy

In all cases we discussed so far, the material "outside" the coordinated stretch of the clause was either the head (verb) or a "high" argument phrase like the subject. Is it possible to generalize the idea further to also allow for cases in which the "outside material" appears to be much "lower", as in Figure 11?

It turns out that we can take Forst and Rohrer's philosophy of minimal grammar extension one step further. Figure 12 shows a canonical case of an extraposed clause, appearing in the *Nachfeld* region of the clause. Since the verb selecting it may be embedded under a modal or some raising verb, the f-annotation for CPs in this position has to contain a functional uncertainty equation, providing an optional embedding under XCOMP's.

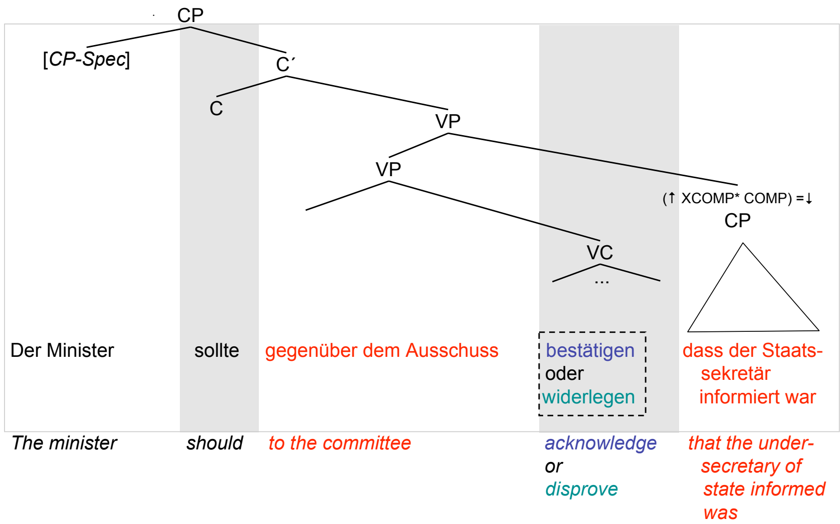


Figure 12: Standard *Nachfeld* analysis

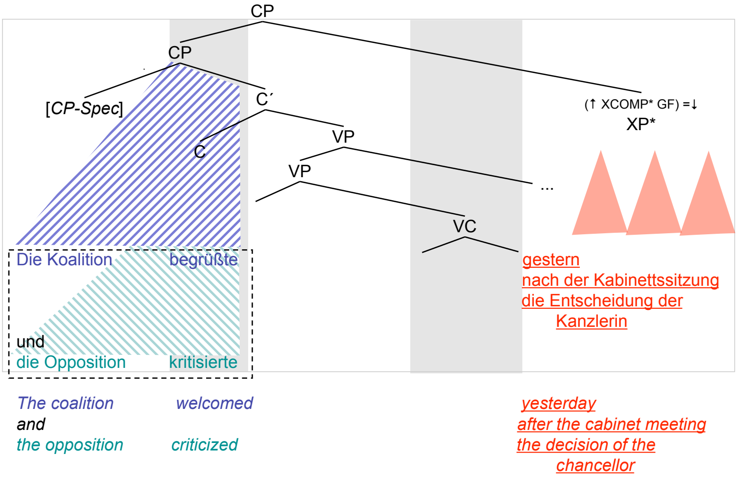


Figure 13: Extended *Nachfeld* analysis for Figure 11

Now, it turns out that the existing *Nachfeld* position can very easily be used in order to capture cases like Figure 11 in terms of a constituent coordination, as is seen in Figure 13. If we assume that the adjunction of the *Nachfeld* can be at the root CP level and we relax the categorial restriction to also allow for nominal arguments (as they also appear extraposed in spoken German), the material that appeared to occur "low" inside the constituents can quite naturally be analyzed as appearing "high", outside the coordinated CP constituent. The fact that the entire *Mittelfeld* and the right verbal bracket can be empty is nothing new: these parts of the clausal structure have to be optional anyway in order to capture sentences like *Maria lacht* ("Maria laughs") as instances of verb second, where *Maria* appears in Spec-CP, and *lacht* in C. The variant of Figure 13 in Figure 14 shows that the "splitting point" can easily be moved around.

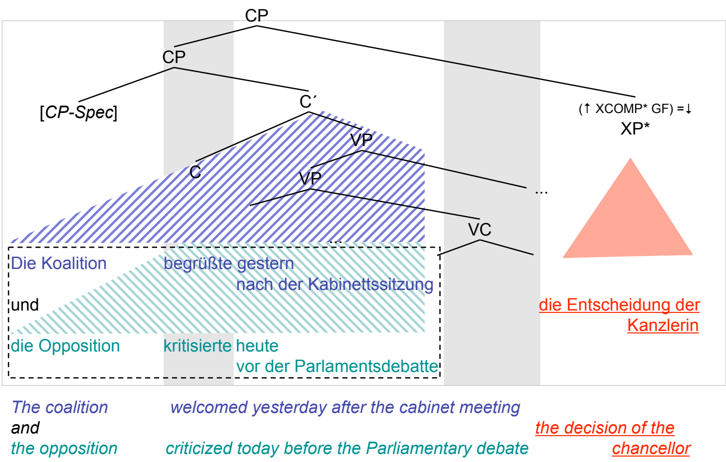


Figure 14: Variant of Figure 13 with a smaller overlap in conjuncts

In our analysis, the unwanted duplication of such optional constituents (generating them both inside the coordination and as part of in the "outside material") or the complete skipping of required material is excluded on the basis of standard f-structure mechanisms like uniqueness and completeness/coherence.

However, we have to enforce that the "gaps" in the constituents (resulting from optionality) are indeed peripheral. In Maxwell and Manning's account this is ensured by keeping track of the automaton state up to which material was covered. Without further provisions, our account would overgenerate, for instance allowing the ungrammatical example (14) in which a complex verb form *hat … begrüßt* is used, which requires the VC to be filled, such that the *Nachfeld* material is no longer adjacent to the *Mittelfeld* "gap" in the conjuncts. In our grammar, this restriction is implemented as a simple c-structural constraint that forbids right node raising constructions in contexts where the VC is filled.

(14) *[[Die Koalition hat gestern nach der Kabinettssitzung begrüßt] und
The coalition has yesterday after the cabinet meeting welcomed and
[die Opposition hat heute vor der Parlamentsdebatte kritisiert]] die
the opposition has today before the Parliamentary debate criticized the
Entscheidung der Kanzlerin
decision of the chancellor

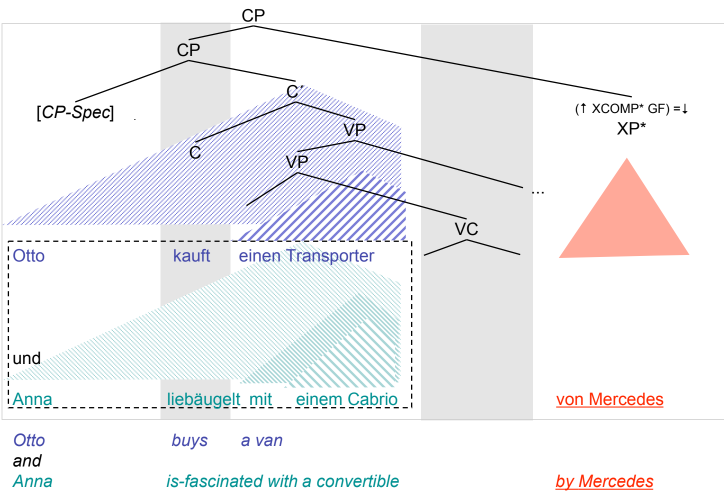


Figure 15: Right node raising involving asymmetric hierarchical embeddings

The extended *Nachfeld* approach is also directly compatible with a situation where the "raised material" originates from constituents involving a complex hierarchical embedding (Figure 15), replicating Maxwell and Manning's idea that involves stacks – with the proviso that any "partial" rules one would like to use for this either have to coincide with available full categories usable in the *Nachfeld* or must be especially added.[2]

[2]Note that the application of such hierarchically embedded partial rules can no longer be con-

We note that the philosophy of applying minimal extensions to the canonical grammar can also be applied to cover right-node raised argument phrases. The functional uncertainty path needed to ensure distribution into embedded structures is even independently motivated from the canonical *Nachfeld* treatment (contrary to the situation in Maxwell and Manning's approach discussed at the end of Section 2.2, where the functional uncertainty annotations needed to be stipulated).

However, there are limitations to what can be realistically covered without opening up the search space too much. In particular, the addition of otherwise unmotivated embedded "partial constituents" as would be required for examples like (12) (which are however not too common in German) leads to serious efficiency problems. As a general strategy, we did not include hierarchical embeddings which would descend more than two constituent levels in our large-scale LFG. This is typically the case in right node raising constructions which involve partial PPs or DPs, illustrated in example (15).

(15) Bei den Wahlen von 1990 verlor $[_{\text{VP}}$ $[_{\text{VP}}$ $[_{\text{DP}}$ die CDU] $[_{\text{DP}}$ die rechten]]
In the elections of 1990 lost the.NOM CDU the right.ACC
und $[_{\text{VP}}$ $[_{\text{DP}}$ die SPD] $[_{\text{DP}}$ die linken]]] $[_{\text{DP}}$ Wähler].
and the.NOM SPD the left.ACC voters

"In the 1990 elections, the CDU lost the right-wing voters and the SPD lost the left-wing voters."

4 Information structure, prosody and generation

In our generalized constituent approach, the "partial" constituents required in cases of non-constituent coordination are in fact using up entire rules, mostly by not realizing one or more optional subconstituents at the edge. As mentioned above, our approach cannot enforce an "automatic" c-structural periphery condition as it is included in Maxwell and Manning's meta mechanism, which is entirely controlled by the c-structure rule consumption.

Considering the linguistic constraints at play, it appears however that it is not c-structural adjacency alone that drives the construction. As Maxwell and Manning point out too, prosody and information structure place important additional constraints. Right node raising examples typically involve a pronounced contrastive focus on the two (right-peripheral subconstituents of) the conjuncts, which is marked by pitch accents and the intonational phrasing. Furthermore, the two focus elements have to be parallel (typically at the level of grammatical functions), excluding cases like (16).

strained straightforwardly, since the rule in which they are called does not directly include the coordination. One could in principle use the parametrized rule mechanism of XLE to propagate the information of being inside a coordination (compare for instance Kuhn (1999)), but this would involve highly complex grammar engineering, unless one could automatize rule compilation based on a cross-classification of rule meta-principles.

(16) ?[[Peter schenkte seiner Tochter] und [ihrem Sohn schenkte Maria]] ein Buch

?[[Peter	schenkte	seiner	Tochter]	und	[ihrem	Sohn	schenkte	Maria]]	ein	Buch
Peter.NOM	gave	his.DAT	daughter	and	her.DAT	son	gave	Maria.NOM	a.ACC	book

"Peter gave to his daughter and to her son gave Mary a book."

Following the main ideas of Féry and Hartmann (2005), it is in fact relatively easy to simulate these information-structural constraints in an LFG account using only f-structure, by introducing a special discourse function RNR-FOCUS (standing in for a full account involving an information structural projection), as indicated in Figure 16.

The coordination rules can generally be formulated in a way so they enforce (a) that each of the conjuncts ends in an element introducing the RNR-FOCUS (leaving the appropriate "gap" to its right), and (b) that the two (or more) peripheral elements introduce the same grammatical function.

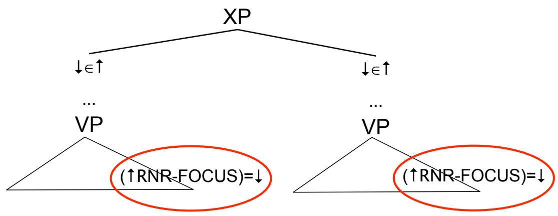


Figure 16: Information structural constraints

In the grammar engineering context, these parallelism constraints are particular relevant for using the grammar in generation tasks. For instance, Cahill and Riester (2009) apply the German LFG in a surface realisation ranking task, where they exploit the reversible XLE engine to generate all possible sentences from a given input f-structure. In such a task, a grammar without the above focus constraint on RNR would generate a lot of infelicitous word order variations that independently permute the constituents in the conjuncts, as in example (16).

The RNR focus restriction constrains the generation of word order variations to cases where the constituents inside the conjuncts are permuted in parallel to each other. Thus, if the generator is supplied with the f-structure for the sentence in (17) (without fixing the discourse functions to particular grammatical functions), it produces, among others, the output alternations in (18). Note that the order of the nominative and accusative arguments in the conjuncts is always the same. Infelicitous cross-over configurations are not generated due to the grammatical constraint.

(17)

Gestern	hat	[[der	Attentäter	einen	Polizisten]	und	[der	Demonstrant	einen	Politiker]]	mit	einer	Pistole	bedroht
Yesterday	has	the	assassin.NOM	a	police man.ACC	and	the	demonstrator.NOM	a	politician.ACC	with	a	gun	threatened.

(18) *Surface alternations generated for (17)*

a. Gestern hat [[einen Polizisten der Attentäter] und [einen Politiker der Demonstrant]] mit der Pistole bedroht.

b. Mit der Pistole hat [[einen Polizisten der Attentäter] und [einen Politiker der Demonstrant]] gestern bedroht.

c. Bedroht hat [[der Attentäter einen Polizisten] und [der Demonstrant einen Politiker]] gestern mit der Pistole.

However, the generation constraints on word order become more difficult to formulate in the case of right node raising that involves partial DPs or PPs. As an example, consider the following word order alternations of varying fluency/acceptability that the grammar generates for sentence (15), where the "raised" noun figures before the coordination (the latter being an instance of the Split NP construction (Kuhn, 2001)):

(19) *Surface alternations generated for (15)*

a. ?Bei den Wahlen von 1990 verlor Wähler [[die CDU die rechten] und [die SPD die linken]].

b. Wähler verlor bei den Wahlen von 1990 [[die CDU die rechten] und [die SPD die linken]].

5 Conclusion

We discussed a practical account of non-constituent coordination cases in the clausal syntax of German as they occur in real corpus data, generalizing the approach of Forst and Rohrer (2009). In particular, we compared the account to the meta-grammatical technique proposed by Maxwell and Manning (1996), which captures most aspects of the construction in a very elegant way, but is presumably impossible to implement on a larger scale.

(20) [[[Bíll]$_T$ [lìkes]$_F$ ~~cigars from Cuba~~] and [[Jóe]$_T$ [is thòught to like]$_F$ ~~cigars from Cuba~~]] cigars from Cuba.

The main ideas behind our account – sketched in an inutive notation in the English example (20), a repetition of (4) – are the following: (i) to use canoncial rules as much as possible and take advantage of optionality to generate a variant of the constituent involving what looks like "deletion" in the periphery of the conjuncts; (ii) to extend/relax existing rules (such as extraposition rules) in order to host the raised or "outside material", using the standard LFG mechanism of f-structural distribution and in particular taking advantage of the independently motivated functional uncertainty annotation in order to capture asymmetric hierarchical embeddings in right node raising examples; (iii) to approximate information-structure constraints behind right node raising, which enforce an opposition of two pairs of contrastive topic/contrastive focus, along with the appropriate prosodic (phrasal and pitch) marking.

We make no claims that the approach captures the linguistic constraints in the most appropriate way since the overarching motivation comes from grammar engineering considerations, i.e., the goals of maximizing (1) coverage and (2) quality of analysis, relative to corpus frequency of the phenomena.

However, we believe that linguistically grounded grammar engineering work, as in the LFG-based research paradigm, is now at a point where the efficiency of the parsing system (XLE) and the broad coverage of the grammars allow us to address questions that are interesting both from a linguistic and an engineering point of view, and find engineering answers that can at least be related in an interesting way to linguistic considerations.

References

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Butt, Miriam, Helge Dyvik, Tracy Holloway King, Hiroshi Masuichi, and Christian Rohrer. 2002. The Parallel Grammar project. In *Proceedings of COLING-2002 Workshop on Grammar Engineering and Evaluation*, Taipei, Taiwan.

Cahill, Aoife, and Arndt Riester. 2009. Incorporating information status into generation ranking. In *Proceedings of ACL-IJCNLP 2009*, pp. 817–825, Singapore.

Dipper, Stefanie. 2003. *Implementing and Documenting Large-Scale Grammars — German LFG*. PhD thesis, IMS, University of Stuttgart. appeared in Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung (AIMS), Vol. 9(1).

Féry, Caroline, and Katharina Hartmann. 2005. Focus and prosodic structures of German Gapping and Right Node Raising. *The Linguistic Review* 22:67–114.

Forst, Martin, and Christian Rohrer. 2009. Problems of German VP Coordination. In *Proceedings of the LFG '09 Conference*, pp. 297–316, Cambridge, England.

Kaplan, Ronald, and John Maxwell. 1989/95. Constituent coordination in Lexical-Functional Grammar. In M. Dalrymple, R. M. Kaplan, J. T. Maxwell, and A. Zaenen (eds.), *Formal Issues in Lexical-Functional Grammar*. Stanford, CA: CSLI Publications.

Kuhn, Jonas. 1999. Towards a simple architecture for the structure-function mapping. In M. Butt and T. H. King (eds.), *Proceedings of the LFG99 Conference, Manchester, UK*, CSLI Proceedings Online.

Kuhn, Jonas. 2001. Resource sensitivity in the syntax-semantics interface and the German split NP construction. In D. Meurers and T. Kiss (eds.), *Constraint-Based Approaches to Germanic Syntax*, pp. 177–215. Stanford: CSLI Publications.

Kuhn, Jonas, and Christian Rohrer. 1997. Approaching ambiguity in real-life sentences – the application of an Optimality Theory-inspired constraint ranking in a large-scale LFG grammar. In *Proceedings of DGfS-CL Heidelberg*.

Maxwell, John, and Christopher Manning. 1996. A theory of non-constituent coordination based on finite-state rules. In M. Butt and T. H. King (eds.), *On-line Proceedings of the First LFG Conference*, Grenoble.

Milward, David. 1994. Non-constituent coordination: Theory and practice. In *Proceedings of the fifteenth International Conference on Computational Linguistics (COLING-94)*, pp. 935–941.

Zarrieß, Sina, and Wolfgang Seeker. 2008. Finite-State Rule Deduction for Parsing Non-Constituent Coordination. In *Finite-state methods and natural language processing : 6th International Workshop, FSMNLP 2007*, Potsdam.