

HPSG Analysis of German

Stefan Müller and Walter Kasper

DFKI GmbH, Saarbrücken, Germany

1 Introduction

The goal was to develop a grammar for German based on multi-lingual principles in the common framework of HPSG (cf. chapter 8) for deep syntactic and semantic analysis.

The core is a general domain independent competence grammar based on the notion of grammatical correctness. To be usable for Verbmobil it had not only to be extended with domain specific constructions for the Verbmobil domains of appointment scheduling and travel planning, but also to deal with specific properties of spontaneous speech. Especially the latter task presents not only a challenge of how to accommodate it with criteria of grammatical correctness but also gives rise to apparently conflicting requirements. On the one hand, the grammar should be very restrictive in order to filter from word hypothesis graphs (WHGs) the correct string to compensate recognition errors. This strategy presupposes that most of the time speakers use language grammatically correct. On the other hand, the analysis must be permissive enough to be able to understand even strictly “ungrammatical” utterances when they occur. This seems to call for some kind of relaxing grammatical constraints. But in using the grammar for generation also a restrictive policy with respect to correctness is required. An additional complication arises from the fact that the utterances are not necessarily “sentences” as usually described by a formal grammar but *turns* and *turn segments* delimited prosodically and not necessarily coinciding with phrases in a strict linguistic sense. Efficiency in using the grammar in a real time system poses additional restrictions.

Some of these questions regarding the analysis of spontaneous speech are answered in Verbmobil by distribution of labor: so the repair module (cf. chapter 8) accounts for e.g. restarts and auto corrections, other types of ungrammaticality are treated in the robust semantics module (cf. chapter 8).

In the following we describe the setup and coverage of the German grammar related to these issues and with respect to the Verbmobil data. In the first sections we describe lexicon and grammar. Finally, we will present results on the coverage of the grammar.

2 The Lexicon

In theories like HPSG the lexicon is the most important part of the grammar since it contains not just morphological information about the form of words and their syntactic category or part of speech but also all the information about what types of expressions can be combined with it in order to yield full phrases.

For (0) the lexicon contains the information that *machen* has to be combined with *aus* and two noun phrases to yield a sentence. *einen Termin* and *wir* are possible instantiations of the required NPs. In the lexical entry of *Termin*, it is encoded that this noun needs a determiner to be a full NP.

- (1) Wir machen einen Termin aus.
 we make an appointment PART
 ‘We schedule an appointment.’

Besides syntactic and morphological information the lexical entries also specify the lexical semantics of the word. Each kind of information is specified as a type lattice. Multiple type inheritance allows to integrate these into a linguistic *sign*. Interaction of constraints between information types is specified by coreference. The objects in the lexicon are instances of specific lexical signs defining their class. In addition to types, the German lexicon employs parameterized templates. Templates allow the specification of only those properties which distinguish a lexical item from others in its class. The following entry for the noun *Termin* shows an example. The value for the morphological stem (STEM), the semantic relation that is introduced (RELN) and for the sort in the Verbmobil ontology (SORT) are specified as parameters of the template for `count-noun`.

- (2) TERMIN-NN := @count-noun(\$STEM = 'termin,
 \$RELN = 'termin,
 \$SORT = time_sit_poly).

The use of templates facilitates a specification of the lexicon that is independent of a certain syntactic theory or a special feature geometry. This approach makes it easy to maintain and extend the lexicon even for a non-specialist in HPSG since no specific knowledge about the type definition in the grammar is required.

The template definition specifies the lexical class of an entry and the paths of the parameter values:

- (3) count-noun(\$STEM,\$RELN,\$SORT=anything) :=
 count-noun-lex &
 [MORPH <! [STEM \$STEM] !>,
 SYNSEM.LOC.CONT [INDEX.SORT \$SORT,
 KEY.PRED \$RELN]].

In the list of template parameters, the SORT parameter has a default specification. Defaults make the specification of the lexicon still easier as one has only to specify parameters for exceptional cases. If the value of SORT is not provided in a description like (2), the value `anything` is inserted. The parameters are inserted as values under certain paths in the feature structure that is described by the type `count-noun-lex`. The type `count-noun-lex` inherits from a type that establishes the interface to the morphology component and from two types that describes the syntactic and semantic properties of a count noun, respectively. Figure 1 on the facing page shows a feature description that corresponds to the noun *Termin*. Due to space restrictions certain features have been suppressed and the type `det-synsem` has not

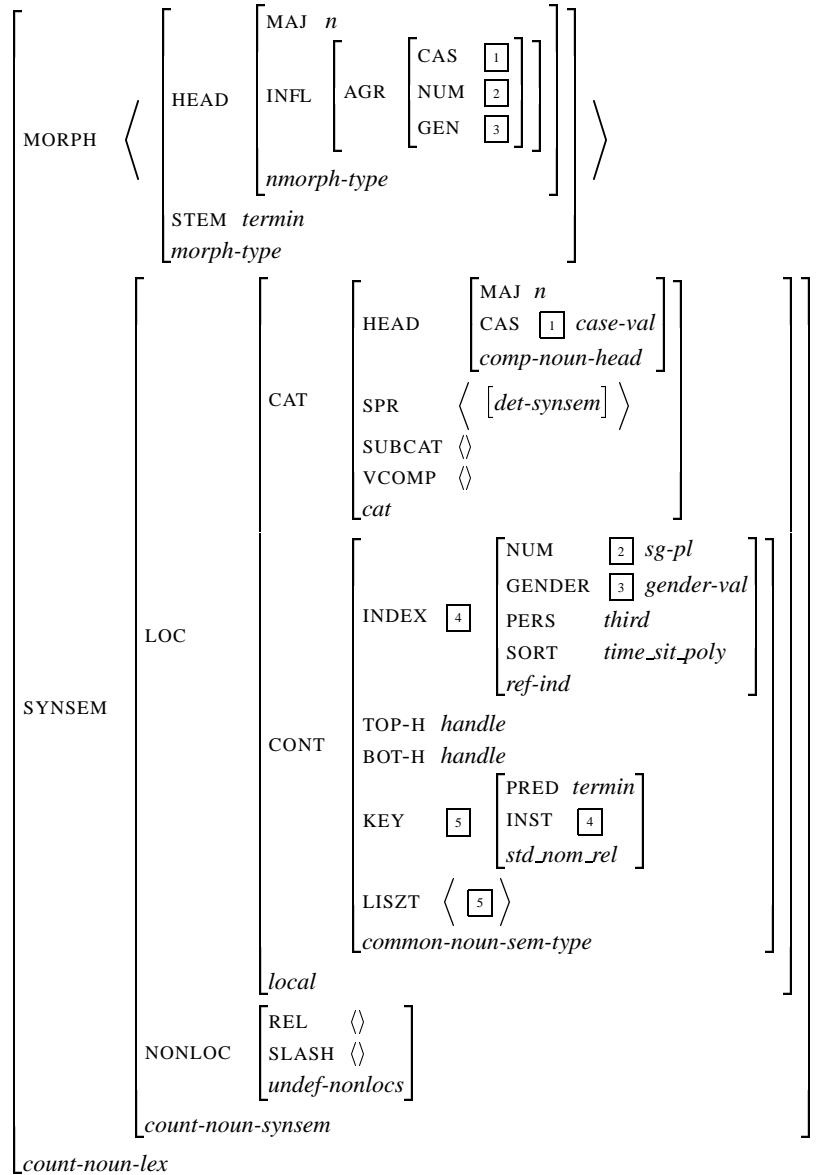


Figure1. Feature description of *Termin*

been expanded. The type refers to another structure of type *synsem*. It is easy to see that a change in the definition of lexical type usually does not require any change to the lexical database itself. One can even extend the list of template parameters without necessarily having to edit the lexicon when new parameters have a default value satisfied by the already existing entries. To each lexical class there belongs a template. Currently, 300 fine-grained lexical classes are defined, thus exceeding the number of standard parts of speech considerably.

One will notice that in the example only the morphological stem is specified. The agreement features *gender*, *number*, and *case* are not specified in the lexical entries. These are instantiated by the morphological analysis of a word form under MORPH. The interface between morphology and syntax/semantics is represented by the structure sharings in Figure 1 which specify that some of the agreement features — *gender* and *number* — represent semantic properties and only *case* is a syntactic property. By this use of morphology, one needs to specify only one entry to cover all forms, e.g. *Termin*, *Termine*, *Termins*. This reduces the size of the lexicon. The recognition word list of 10000 word forms in Verbmobil is covered by about 5900 stem. Because of paradigmatic completeness actually the stem lexicon even exceeds this wordlist. Due to syntactic and semantic ambiguities, the overall size of the stem lexicon is 9900 entries.

A second component of the lexicon besides the stem lexicon is a *multi-word lexicon* containing fixed word combinations and collocations with idiomatic meanings, such as *Auf Wiedersehen* (= 'good bye'). Their meaning cannot be derived compositionally from the meanings of the parts.

The third component in the lexicon deals with unknown words. For each possible type of unknown word — in Verbmobil several classes of names — a generic lexical default entry is defined from which the parser can derive additional instances when unknown words are encountered.

From instances that are listed in the lexicon other lexical entries can be computed by lexical redundancy rules (Flickinger et al., 1985; Flickinger, 1987). There are 20 lexical rules in the grammar. For example, one rule relates verbs with a valency representation for active sentences to verbs with a valency representation for passive sentences (Bresnan, 1982; Pollard and Sag, 1987). Lexical rules as we use them here are fully integrated into the HPSG formalism. They are also described by feature structures and resemble unary rules (Copestake and Briscoe, 1992; Krieger and Nerbonne, 1993) but are only applied in lexicon lookup and not in later stages of parsing.

Lexical entries with the same stem as well as grammar rules can be assigned relative weights in order to specify a ranking of readings for tuning the grammar and the lexicon towards a domain and application.

3 The Grammar

Since most of the syntactic information resides in the lexicon the core grammar requires only a small number of general rule schemata (in contrast, e.g., to pure phrase

structure grammars). It consists of 21 rule schemata, 7 special rules that introduce verbal projections for verb movement, and 17 special domain specific rules. All these rule schemata are unary or binary branching. As HPSG schemata, they specify coreference constraints between the constituents and the projection of features on the syntactic and semantic level thereby automatically building the semantic representation. The grammar defines not only full sentences as possible utterance but allows also for other kinds of maximal projections as full utterances, such as individual NPs or PPs as elliptical utterances. In addition, it specifies in a transparent way what kind of phrases count as partial analyses and good fragments which are passed to the robust semantics component in case a full grammatical analysis is not possible.

4 Phenomena

The grammar consists of three layers: first, there is the domain independent base grammar, covering phenomena that have to be dealt with by every grammar for German. This layer is described in detail in Müller (1999a). Second, there are domain specific rules covering constructions and expressions typical for a specific domain or application. This second part is a monotonous extension of the base grammar, No rule in the base grammar is overwritten or relaxed, just new rules are added. Thirdly, there are rules for analyzing phenomena of spontaneous speech. We achieved to set up these also as a plain extension of the base grammar. In the following, these three layers will be discussed in more detail.

4.1 Base Phenomena

The following list is an overview of the covered phenomena.

- declarative clauses (‘Ich komme am Montag.’)
- questions
 - *wh* questions (‘Welchen nehmen wir?’)
 - yes/no questions (‘Meinen Sie den Montag?’)
- imperatives (‘Schlagen Sie einen vor!’)
- embedded clauses (‘Ich würde sagen, das machen wir so / daß wir das so machen.’)
 - relative clauses (‘der Termin, den wir ausgemacht haben’)
 - adjunct clauses (‘Ich rufe an, weil wir noch einen Termin ausmachen müssen.’)
 - infinitival clauses (‘Ich rufe an, um noch einen Termin mit Ihnen auszumachen.’)
- phrases as full utterances (‘Gut.’, ‘Am Montag?’)
- complementation / adjunction
- participle constructions (‘den für den nächsten Monat geplanten internationalen Kongress’)
- active / passive alternations
- relatively free constituent order

- predicate complexes
- local extraposition of PPs and clausal elements
- symmetric coordination ('Der Montag und der Donnerstag sind schon belegt.')

In order to account for the sentence types listed above, the grammar has to account for verb first, verb second, and verb last sentences. The analysis of verb first sentences is a mixture between analyses that have been proposed by Kiss and Wexler (1991), Netter (1992), and Kiss (1995): A lexical rule introduces a lexical entry for a finite verb that subcategorizes for a verbal projection that contains exactly the complements that are required by the verb that is the input to the lexical rule. The verbal projection is introduced by unary projections that project possible complements of verbs. Figure 2 shows an example. The connection between the empty

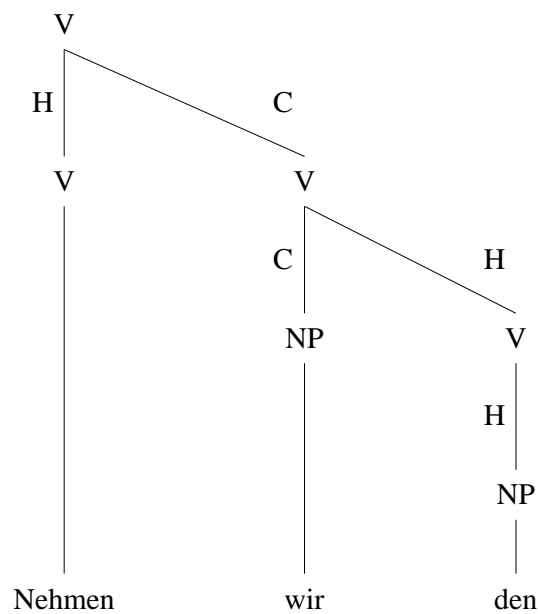


Figure 2. Analysis of *Nehmen wir den*.

projection in the verb final position and the verb first verb is established via a head feature which accounts for the locality of verb movement.

Fixed phrases and collocations are handled by the multi-word lexicon.

4.2 Domain-Specific Rules

A number of domain specific sub-grammars had to be developed for Verbmobil. These include rules for date and time expressions, complex temporal adverbials, complex names, spelling, and price information.

4.3 Spontaneous Speech

The processing of speech input poses a specific set of problems: first the grammar has to be applied to the analysis of word hypothesis graphs. The graphs contain hypotheses of different quality and the grammar may not be overly permissive and provide analyses for hypotheses that represent utterances that were not spoken and in fact are ungrammatical. On the other hand, one finds a lot of ungrammatical utterances in spoken language and the grammar has to be designed in such a way that it provides reasonable output in those cases.

Other problems arise from the segmentation of an utterance. If the speaker makes pauses within a constituent, e.g., between a preposition and its complement NP, the parser might not see the complete phrase in a segment and so be unable to analyze it as a constituent.

As mentioned above, many of these problems with ungrammatical input are treated in *Verbmobil* by collaboration with the robust semantics component. This module prefers to get rather good phrases and fragments from a restrictive grammar rather than having to deal with too many too wrong analyses due to a too permissive analysis.¹ Also, relaxing constraints increases the search space considerably, such impeding performance.

Some of these problems are related to and indistinguishable from the uncertainty in the result of speech recognition and thereby difficult to describe linguistically in a systematic way. Therefore, we rather selected a restricted set of high frequency phenomena in the corpus which are much less important in written input and are capable of a linguistic analysis. Examples are *Vorfeldellipse* (also called ‘topic drop’, note that in (4d) an expletive is dropped, which cannot be a topic) (4a), preposition stranding (4b), turn initial discourse particles and final echo tags (4c), determiner drop with count nouns (4d), and sloppy case (4e).²

- (4) a. Klingt gut.
sounds good
- b. (*Da*) bin ich *mit* einverstanden.
there am I with agreed
‘I agree with this.’
- c. Ja, das paßt.
yes this suits
‘Yes, this suits me.’
- d. Es gäbe eines mit Sauna; ja, gibt auch Hallenbad.
it_{expl} gives one with sauna yes gives also swimming.pool
- e. Anfang nächste(r) Woche
beginning next week

¹ The statistical grammar (cf. chapter 8) appears to suffer a bit from its permissiveness.

² See also Block and Schachtl (1995) for an overview of important phenomena in speech systems.

(4a) is handled by a special rule for *Vorfelldellipse* and preposition stranding is handled by a special lexical rule. The combination of discourse particles with complete phrases as well as the appending of an echo tag are also handled by special rules.

These phenomena are not handled in the base grammar but as extensions since they either do not occur at all in written language or occur with a rather low frequency. Another reason for not describing a phenomenon in the base grammar is that the phenomenon is specific for certain text sorts and stylistic variation. An example is the *Vorfelldellipse* which occurs in certain text styles only (Fries, 1988, p. 27). It is also not obvious that, e.g., topic drops should be generated.

4.4 Other Grammar Extensions

For a small number of phenomena, we came to the conclusion that it is preferable to leave them outside the Verbmobil run time system, either because they are very rare and at the same time very costly, or because it can increase the number of false analyses in WHG parsing significantly and is better handled by the robust semantics component. The latter concerns especially some non-local extraposition phenomena which presuppose certainty with respect to the boundary of the utterance for finding a correct analysis.

An example of the former kind are some special and rare forms of passive:

- (5) Sie bekommen dann die Unterlagen von uns zugeschickt.
 you get then the documents by us PART (to).sent
 ‘We will send you the documents then.’

The example in (5a) is an instance of the so-called *bekommen* passive. The pronoun *Sie* is the dative object of *zuschicken*, but surfaces as nominative in the *bekommen* passive. In (Müller, To Appear b; In Preparation) it is argued that passive should be handled as a lexical rule.

- (6) a. Wir haben einen Termin ausgemacht.
 we have an appointment scheduled
 ‘We scheduled an appointment.’
 b. Der Termin wurde ausgemacht.
 the appointment got scheduled

Since the form of the perfect participle is always identical to the passive participle this leads to an increase of the lexical ambiguity. The dative passive is possible with almost all verbs that allow for a normal passive and that govern a dative. So, having a rule for the *bekommen* passive in the grammar results in an enormous increase in the number of lexical hypotheses. The same is true for the *lassen* passive.

- (7) a. Sie warten am besten an der Pforte und lassen anrufen.
 you wait at best at the entrance and have call
 ‘You best wait at the entrance and have the doorman call.’

- b. Sie lassen den Pförtner anrufen.
 you have the doorman call

The sentence in (7a) uses a passive form of *anrufen*. The subject of *anrufen* is suppressed. *lassen* embeds an infinitive without *zu* and having an appropriate lexical rule in the grammar to handle cases like (7a) slows down the processing of all sentences that contain a bare infinitive of a verb that can be passivized.

A related phenomenon with the *zu* infinitive are modal infinitives.

- (8) a. Ich habe nix gegen Sonntag Abend einzuwenden.
 I have nothing against Sunday evening to.have.an.objection
 ‘I have no objection to Sunday evening.’
 b. Gegen Sonntag Abend ist nix einzuwenden.
 against Sunday evening is nothing to.have.an.objection
 ‘There is no objection to Sunday evening.’

There are active and passive modal infinitive constructions.

The last phenomenon we want to mention here is left dislocation (Altmann, 1981) as in (9).

- (9) Der Montag, der paßt mir gut.
 the Monday it suits me well
 ‘The Monday suits me fine.’

In these constructions initial phrases are doubled which requires in some sense that an argument role could be filled by two different constituents.

5 Semantics

The basic construction of semantic representations in the framework of MRS is part of the grammar as described in chapter 8. We will confine ourselves here to some special issues: the use of *lexical underspecification* for verbs and particles and the treatment of certain *relational constraints*.

Besides types of structural ambiguities such as *scope* and *attachment* ambiguities there is also a large amount of lexical ambiguity in the meaning of words even in restricted domains like those of Verbmobil. Especially certain verbs, adverbs and discourse particles give rise to problems of their correct disambiguation in the absence of information from discourse context.

As an example we will discuss the highly frequent verb *gehen* which occurs in the Verbmobil data with at least eight different basic senses, such as *to go, suit, act, feel, last, concern*, not even counting its uses with separable verbal particles. In general, these senses can be distinguished and disambiguated by their differing subcategorization frames and by selectional restrictions on their arguments. Problems arise when arguments are only optional and so can be absent, when the arguments are pronominal and so the selectional restrictions cannot be evaluated locally, or when only a partial analysis is available. Instead of having to enumerate all possible

senses or making an arbitrary choice in each such case, underspecified verb lexemes were introduced. The possible argument frame for these entries comprises the argument frames of the possible individual senses. Disambiguation is delayed until discourse semantics (cf. chapter 8) has resolved pronouns and so has established e.g. the sort of the argument and robust semantics had a chance to combine partial analyses into more complete analyses. Also, the parser benefited from this treatment of underspecification as it reduced the number of lexical hypotheses it had to deal with.

Another problem in using unification frameworks arises with the use of relational or functional constraints which have to be invoked when feature values cannot be established by simple unifications, e.g. between the mother node and structures in its daughters. Because of great control and efficiency problems such constraints are often excluded from the formalism or parser; in the Verbmobil-parser, too. On the other hand, they cannot be avoided in semantics. Besides the problems of having to deal with non-local contexts such as discourse, there are even much simpler problems. In the Verbmobil-domains, it is pretty difficult to deal with dates and times without doing some — even numerical — calculation. E.g., an expression like *a quarter to nine* should represent the time 8:45, which is easily computed as *9 o'clock minus a quarter of an hour*. In order to provide a declarative interface to such processes in the grammar, a special type of semantic relation `computed_relation` was introduced. These relations specify the name of a function or relation with their parameters from the utterance. The relations are evaluated after parsing to yield a correct representation. They are required to have an inverse for reverse computation.

Another application for such constraints in the grammar is the treatment of *uncertain* contextual or extra-linguistic information such as that from prosody (cf. the next section). As uncertain, the information should not be treated as a hard constraint but rather is recorded as potential conflict which is evaluated when more (non-local) information can be taken into account to resolve the potential conflict if there is one.

6 Prosody

Inclusion of prosodic information in a grammar for analyzing spoken language is important for several reasons (Strom et al., 1997):

- *phrase* and *utterance boundaries* within a turn because there are no punctuation marks as in written language
- sentence mood: whether an utterance is declarative, question or an imperative cannot be determined from grammatical properties alone:

(10) Sie haben Zeit?
 you have time?

is a question despite its declarative sentence form.

- stress can have important semantic functions for marking contrasts, focus, topic and scope, e.g.,

(11) noch einen Termin

can be translated as *yet another appointment* or simply *an appointment* depending on whether *noch* is accentuated or not. In the first case *noch* usually has contrastive meaning, in the second case it is a discourse continuation marker.

In Verbmobil prosodic information of these kinds is available in the WHG with confidence values. If the confidence values are beyond a threshold the information is included in the analysis. Prosodic projection principles as described in Kasper and Krieger (1996) project it as constraints to the phrasal and sentence levels. Potential conflicts between grammatical and prosodic constraints as those concerning sentence mood are resolved after parsing during the VIT-construction as relational constraint.

7 Empirical Results

To evaluate the grammar the transliterated dialogs from several Verbmobil-CDs were segmented by hand, thereby simulating the case of perfect recognition and segmentation. In a manual evaluation of the resulting 24,584 utterances 23,104 were classified as grammatical and 1,480 as clearly ungrammatical. 82.3 % of the grammatical inputs are successfully parsed. 3.8 % successful parses on the ungrammatical inputs indicate a low rate of overgeneration in the grammar.

Table 1 shows the results of a detailed error analysis of failed parses of grammatical utterances from one CD indicating remaining problems in the coverage of certain phenomena. The second column gives the total number of sentences and the third column gives the percentage of all well-formed sentence.

Table1. Error statistics for CD-15: 2231 segments, 2040 grammatical

phenomenon	number	%	phenomenon	number	%
ellipsis	96	4.7	copula	14	0.7
coordination	26	1.4	missing determiner	14	0.7
w-embedding	18	0.9	Extraposition	12	0.6
correlate sentences	17	0.9	parenthesis	7	0.3

The table shows that — except for ellipsis and coordination — in the average the phenomena on which the grammar fails are rare and sometimes exotic. The following example illustrates ellipsis and coordination at once:

- (12) Mir würde der Montag passen und der Dienstag.
 me would the Monday suit and the Tuesday
 ‘Monday or Tuesday would suit me.’

The second conjunct is elliptical. Agreement in (12) and similar sentences shows that *der Montag und der Dienstag* cannot be a (discontinuous) constituent (Hoffmann, 1997, p. 2366). The problem of resolving ellipses cannot be solved by the grammar alone. In the case of (12) partial analysis for *Mir würde der Montag passen* and *und der Dienstag* would be generated, which can be translated separately. So, in this case, the missing complete analysis does not necessarily lead to wrong result.

The second large group of parse fails involves asymmetric or otherwise complicated cases of coordination. Coordination in general is an unsolved problem in all syntactic theories. But here, too, in many cases a separate translation of the conjuncts is possible.

Sentences like (13) are also not handled in the grammar:

- (13) a. *Das* ist ganz interessant, so was mitzukriegen.
 that is quite interesting such what to.experience
 ‘It is interesting to realize such things.’
- b. aber vereinbaren wir *es* auf jeden Fall, daß wir was
 but arrange we it in any case that we somewhat
 Gemütlicheres machen wollen.
 more.informal make want
 ‘but we arrange it in any case that we want to do something more informal.’
- c. Es käme jetzt *drauf* an, ob wir dort noch was
 it came now there.upon PART whether we there yet somewhat
 unternehmen wollen.
 undertake want.to
 ‘It is now dependent on whether we want to undertake something there.’
- d. Das geht eigentlich hier *so* flott, daß ich eigentlich immer auf die
 this goes in.principle here so quick that I in.principle always on the
 Bahn verzichten möchte.
 train do.without want.to
 ‘This goes so quick here in principle that I always want to do without the train in principle.’

In (13a) the extraposed infinitive corresponds to the subject *das* at the beginning of the sentence. In (13b) the *daß* clause is coreferent with the accusative object. In (13c) the *ob* clause corresponds to a nominal part that is contained in the so-called pronominal adverb *drauf*. Such pronominal adverbs consists of the prepositional part (*auf*) and a nominal part (*dr* or *dadr*).

In English syntax it is common to analyze the English equivalents to (13a–b) with the expletive pronoun *it* (Williams, 1980, p. 221). Pollard and Sag (1994, p. 149–150) suggested a lexical rule that applies to lexical entries that subcategorize for a clause and produce another lexical entry that subcategorizes for an expletive pronoun and a clause. It is clear that such proposals cannot deal with sentences

like (13d) where the sentence refers to an adjunct of the verb (*so*). Furthermore this approach fails in cases of apposition, as the following example by Pütz (1982, p. 361) shows.

- (14) Die Firma lehnt es als zu umständlich ab, das Auto in Serie gehen
 the company refuses it as to awkward PART the car in series go
 zu lassen.
 to let

The *es* is a complement of *ablehnen* and of *umständlich*. If it were an expletive that is introduced by a lexical rule, both predicates had to subcategorize for an expletive. This means that both predicates had to undergo the lexical rule. One would expect two infinitives since there are two predicates with extraposition *es*, but this is clearly not the case.

Since the number of correlate sentences is negligible, these incorrect proposals were not implemented.

The grammar fragment does account for local extraposition only.

- (15) Dann würde ich doch sagen, daß wir den Zug nehmen.
 then would I but say that we the train take
 'I would say that we take the train.'

In (15) the complement clause of *sagen* is extraposed. This is described by simply linearizing the clause to the right of its head. Such a local treatment is not possible for clause extraposition of NP-adjuncts as in (a16).

- (16) a. Ich rufe Sie an damit wir einen *Termin* ausmachen,
 I call you PART there.with we an appointment schedule
 der uns beiden paßt, für die Reise nach Hannover.
 which us both suits for the trip to Hannover
 'I am calling to schedule an appointment for the trip to Hannover that suits both of us.'
- b. Ich rufe dich an, um mit dir unsere Reise zu besprechen nach
 I call you PART to with you our trip to speak.about to
 Hanover.
 Hanover
 'I call to talk with you about our trip to Hanover.'

The relative clause and its antecedent noun are discontinuous. There are extraposition analyses that can deal with such nonlocal dependencies while assuming continuous constituent structures (Keller, 1994; Bouma, 1996; Müller, 1999a, Chapter 13.2), but these are very expensive since neither the kind of material nor their number can be predicted in a local context. Claims that the number of extraposed constituents is restricted to two are wrong (17a), as is the claim that NPs cannot be extraposed (17b).

- (17) a. Ich habe gearbeitet [an diesem Abend] [in der Kneipe] [als
I have worked at this evening in a pub as
Kellnerin].³
bar-maid
'I worked as a bar-maid in the pub that evening.'
- b. Oh, mir täte passen der Samstag.
oh me did suit the Saturday.
'Oh, Saturday would suit me.'

Other phenomena that have not be treated are certain sub-cases of Preposition Stranding (for an analysis see Müller, 1997b), idioms, support verb constructions, and free relative clauses (cf. Müller (1999b)). All these phenomena occur with rather low frequency.

8 Conclusion

A large, broad coverage grammar for German is available as a result of Verbmobil, incorporating a fully developed syntax/semantics interface. Its successful integration and use in the Verbmobil system demonstrates the benefits of integrating deep linguistic knowledge into speech applications. At the same time, it improved and extended the linguistic understanding of speech related phenomena. Its core grammar provides a valuable resource as a basis for extensions to other domains and applications.

References

- Altmann, H. (1981). *Formen der „Herausstellung“ im Deutschen: Rechtsversetzung, Linksversetzung, freies Thema und verwandte Konstruktionen*. Number 106 in *Linguistische Arbeiten*. Tübingen: Max Niemeyer Verlag.
- Block, H.-U., and Schachtl, S. (1995). What a grammar of spoken dialogues has to deal with. In Heyer, G., and Haugeneder, H., eds., *Language Engineering*, Advanced Studies in Computer Science. Wiesbaden: Vieweg.
- Bouma, G. (1996). Extraposition as a nonlocal dependency. In Kruijff, G.-J., Morrill, G., and Oehle, D., eds., *Proceedings of Formal Grammar 96*, 1–14.
- Bresnan, J. (1982). The passive in lexical theory. In Bresnan, J., ed., *The Mental Representation of Grammatical Relations*, MIT Press Series on Cognitive Theory and Mental Representation. Cambridge: Massachusetts, London: England: The MIT Press. 3–86.
- Copestake, A., and Briscoe, T. (1992). Lexical rules in a unification based framework. In Pustejovsky, J., and Bergler, S., eds., *Lexical Semantics and Knowledge Representation*, number 627 in *Lecture Notes in Artificial Intelligence*, 101–119. Berlin: Springer-Verlag.
- Flickinger, D. P., Pollard, C. J., and Wasow, T. (1985). Structure-sharing in lexical representation. In Mann, W. C., ed., *Proceedings of the Twenty-Third Annual Meeting of the Association for Computational Linguistics*, 262–267. Chicago, IL: Association for Computational Linguistics.

³ Spiegel, 23/1997, p. 122

- Flickinger, D. P. (1987). *Lexical Rules in the Hierarchical Lexicon*. Ph.D. Dissertation, Stanford University.
- Fries, N. (1988). Über das Null-Topik im Deutschen. Forschungsprogramm Sprache und Pragmatik 3, Universität Lund-German. Inst., Lund.
- Hoffmann, L. (1997). Koordination. In Eroms, H.-W., Stickel, G., and Zifonun, G., eds., *Grammatik der deutschen Sprache*, volume 7.3 of *Schriften des Instituts für deutsche Sprache*. Berlin, New York: Walter de Gruyter. 2358–2445. <http://www.ids-mannheim.de/gra/grammar.html>. 07.13.99.
- Kasper, W., and Krieger, H.-U. (1996). Integration of prosodic and grammatical information in the analysis of dialogs. In Görz, G., and Hölldobler, S., eds., *KI-96: Advances in Artificial Intelligence. 20th Annual German Conference on Artificial Intelligence. Dresden, September 1996. Proceedings*, 163–174. Berlin: Springer. Also available as Verbmobil-Report 141.
- Keller, F. (1994). Extraposition in HPSG. Verbmobil-Report 30, IBM Deutschland Informationssysteme GmbH, Institut für Logik und Linguistik, Heidelberg.
- Kiss, T., and Wesche, B. (1991). Verb order and head movement. In Herzog, O., and Rollinger, C.-R., eds., *Text Understanding in LILOG*, number 546 in *Lecture Notes in Artificial Intelligence*. Berlin Heidelberg New York: Springer-Verlag. 216–242.
- Kiss, T. (1995). *Infinite Komplementation. Neue Studien zum deutschen Verbum infinitum*. Number 333 in *Linguistische Arbeiten*. Tübingen: Max Niemeyer Verlag.
- Krieger, H.-U., and Nerbonne, J. (1993). Feature-based inheritance networks for computational lexicons. In Briscoe, T., Copestake, A., and de Paiva, V., eds., *Inheritance, Defaults, and the Lexicon*. Cambridge: Cambridge University Press. 90–136.
- Müller, St. (1997a). An HPSG-analysis for free relative clauses in German. In Kruijff, G.-J. M., Morrill, G. V., and Oehrle, R. T., eds., *Proceedings of Formal Grammar 97*, 179–188.
- Müller, St. (1997b). Scrambling in German – extraction into the *Mittelfeld*. Research Report RR-97-06, Deutsches Forschungszentrum für Künstliche Intelligenz, Saarbrücken.
- Müller, St. (1998). Case in German—an HPSG analysis. In Kiss, T., and Meurers, D., eds., *Proceedings of the Workshop Current Topics in Constraint-based Theories of Germanic Syntax*, 113–132.
- Müller, St. (1999a). *Deutsche Syntax deklarativ. Head-Driven Phrase Structure Grammar für das Deutsche*. Number 394 in *Linguistische Arbeiten*. Tübingen: Max Niemeyer Verlag.
- Müller, St. (1999b). An HPSG-analysis for free relative clauses in German. *Grammars* 2(1):53–105.
- Müller, St. (1999c). Parsing of an HPSG grammar for German: Word order domains and discontinuous constituents. In Gippert, J., and Olivier, P., eds., *Multilinguale Corpora. Codierung, Strukturierung, Analyse. 11. Jahrestagung der Gesellschaft für Linguistische Datenverarbeitung*, 292–303. Prag: enigma corporation.
- Müller, St. (1999d). Restricting discontinuity. In *Proceedings of the 5th Natural Language Processing Pacific Rim Symposium 1999 (NLPRS'99)*, 85–90.
- Müller, St. (2000). Continuous or discontinuous? In Hinrichs, E., Meurers, D., and Wintner, S., eds., *ESSLLI'2000 Workshop on Linguistic Theory and Grammar Implementation*.
- Müller, St. (In Preparation). Yet another theory of complex predicates.
- Müller, St. (To appear a). Case in German – towards an HPSG analysis. In Kiss, T., and Meurers, D., eds., *Constraint-Based Approaches to Germanic Syntax*, *Studies in Constraint-Based Lexicalism*. Stanford: Center for the Study of Language and Information.

- Müller, St. (To Appear b). Object-to-subject raising or lexical rule—an HPSG analysis of the German passive. In *KONVENS '2000*. Berlin: VDE Verlag. http://www.dfki.de/~stefan/Pub/e_pass_lr.html. 05.26.2000.
- Netter, K. (1992). On non-head non-movement. An HPSG treatment of finite verb position in German. In Görz, G., ed., *Konvens 92. 1. Konferenz „Verarbeitung natürlicher Sprache“*. Nürnberg 7.–9. Oktober 1992, Informatik aktuell, 218–227. Berlin, Heidelberg, New York: Springer-Verlag.
- Pollard, C. J., and Sag, I. A. (1987). *Information-Based Syntax and Semantics Volume 1 Fundamentals*. Number 13 in CSLI Lecture Notes. Stanford: Center for the Study of Language and Information.
- Pollard, C. J., and Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics. Chicago, London: University of Chicago Press.
- Pütz, H. (1982). Objektsprädikate. In Abraham, W., ed., *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung*, number 15 in *Studien zur deutschen Grammatik*. Tübingen: Gunter Narr Verlag. 331–367.
- Strom, V., Elsner, A., Hess, W., Kasper, W., Klein, A., Krieger, H. U., Spilker, J., Weber, H., and Görz, G. (1997). On the use of prosody in a speech-to-speech translator. In *Proceedings of Eurospeech'97*, 1479–1482.
- Williams, E. (1980). Predication. *Linguistic Inquiry* 11(1):203–238.