Gefördert vom
bmb+f
Bundesministerium für
Bildung, Wissenschaft,
Forschung und Technologie

Verbmobil
Verbundvorhaben

Transfer in the Verbmobil Demonstrator

Bianka Buschbeck-Wolf,
Miriam Butt, Mariana Damova,
Michael Dorna, Kurt Eberle,
Martin Emele, Rita Nübel,
Sabine Reinhard,
Bärbel Ripplinger,
Michael Schiehlen,
Helmut Schmid

IAI Saarbrücken

IBM Deutschland Informationssysteme GmbH

Universität Stuttgart

Eberhard-Karls-Universität Tübingen

Vm

Report 147
August 1995

August 1995

Bianka Buschbeck-Wolf, Miriam Butt, Mariana Damova, Michael Dorna,
Kurt Eberle, Martin Emele, Rita Nübel, Sabine Reinhard, Bärbel Ripplinger,
Michael Schiehlen, Helmut Schmid

Institut für Maschinelle Sprachverarbeitung
Universität Stuttgart
Azenbergstraße 12
D – 70174 Stuttgart

Tel.: (0711) 121 - 1352
Fax: (0711) 121 - 1366
e-mail: `vm@ims.uni-stuttgart.de`

Gehört zum Antragsabschnitt: 11, 12

Die vorliegende Arbeit wurde im Rahmen des Verbundvorhabens Verbmobil vom Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF) unter dem Förderkennzeichen 01 IV 101 U gefördert. Die Verantwortung für den Inhalt dieser Arbeit liegt bei den AutorInnen.

Contents

Chapter 1

Introduction

For the Verbmobil demonstrator we adopted an HPSG sign-based transfer approach. The basic units of transfer are lexical signs which are accessed via a recursive procedure which operates on the HPSG daughters structure of the parsed utterance. Figure 1.1 gives a schematic overview about the MDS demonstrator architecture and about the basic input/output functionality of the transfer component.

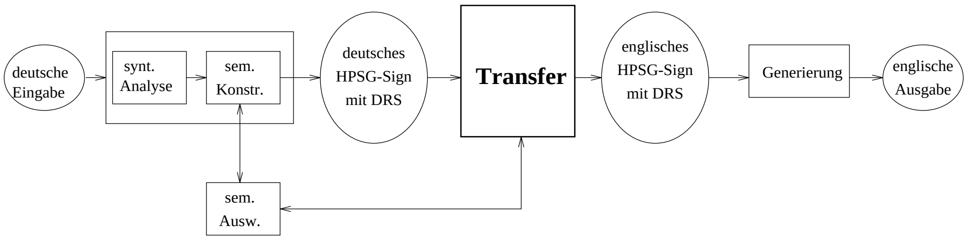


Figure 1.1: Architecture

The input to the transfer module is the result of parsing a German utterance and instantiating its semantics according to the semantic construction process. The transfer process walks over the HPSG daughters structure of the parse tree in a semantic head-driven fashion, applying transfer rules to the lexical signs. Based on the predicate names of the lexical signs, the lexical transfer rules establish translation equivalences between the corresponding lexical sign of the source and target language predicates. The English predicate names serve as access keys to the English semantic lexicon, which had to be built accordingly to the German semantic lexicon by the transfer group due to lack of other responsibility. The MDS semantic construction operations are used to combine the results of transfer into a single DRS which can be input to generation. Additional pragmatic information (e.g speech act, topic/focus information, etc.) from the source sign is passed on to the target sign without further translation.

The functionality of the transfer module was specified in Dorna et al. (1994) and was implemented at the university of Stuttgart. The implementation of the transfer relation is further described in chapter 2. The coding of the lexical transfer relations was done by the different TP-12 transfer sites: IAI Saarbrücken, IBM

Heidelberg, SfS Universität Tübingen and Universität Stuttgart according to the division of the official Verbmobil wordlist into different word classes and semantic phenomena. The different classes will be further described in their respective chapters and subsections of this handbook. Chapter 3, subsections 3.5, 3.6 describe the translation of nouns, pronouns and determiners respectively. It is followed by a description of the verbal translation in chapter 4, continued by a description of the translation of verbal and nominal modifiers in chapter 5. Chapter 6 describes the treatment of idiomatic expressions and particles. Chapter 7 explains the treatment of prepositions and chapter 8 describes the underlying sortal hierarchy. The treatment of tense is described in chapter 9. Finally chapter 10 concludes with a translation example output from dialogue 7 and a description of the overall coverage of the transfer component.

Chapter 2

Implementation of Transfer

We present the basics of the Transfer module implementation realized in the STUF-III language (Momma et al. (1994)) which uses the CUF inference engine (Dörre and Dorna (1993); Dörre et al. (1995)). We explain the definition of the main Transfer predicates used as transfer relation between signs of the source language and a semantic representation of the target language. In this sense, this section is a kind of introduction to the implementation part following in the next sections.

2.1 Preliminary Remarks

In general, we have used high-level ADTs for extracting syntactic information out of a sign (like `intrans_syn/2` or `prep_syn/2`, see below). The same holds for constructing semantic structures (e.g., `intransitive_verb_sem1/5` or `prep_sem1/3`, see below). We mention this because a lot of access/construction predicates were designed and/or implemented by the Transfer group which are now integrated in the Semantic Construction part and documented there. E.g., the splitting into the three layers of lexical construction/access macros was originally developed within a Transfer prerelease. Furthermore, the inauguration of different and unique semantic predicate names for the one and the same lexeme according to different semantic analyses and therefore different semantic structures was motivated and forced by the Transfer group.[1]

For building an *index* over the Transfer rules it was necessary to use atomic predicate names (lemmata). These predicate names served as an interface to Semantic Analysis, Transfer and Generation because they allow for a direct access to the semantic part of a lexeme. We assumed different predicate names for different meanings and therefore different semantic structures to get unambiguous results.

[1]Unfortunately, the semantic lexicon was build on base of lexemes and not on semantic predicates. Hence, no explicit mapping between lexemes (as part of the morpho-syntactical lexicon) and semantic predicates (as part of the semantic lexicon) was defined.

2.2 The Predicate tau/2

The main predicate `tau/2`[2] defining the transfer relation between signs of the source language and a semantic representation of the target language has the following type declaration:

```
@tau(sign,list) => sem_t.
```

The list argument of `tau/2` was used for percolating semantic arguments of some source language predicate to the target language counterpart of *another* source language predicate not incorporating this argument in its argument frame, i.e. for restructuring or also known as structural transfer (see, e.g., the *zu1* example in section 4.2.2). We do not further describe this argument here.

The output of semantic analysis in the Verbmobil implementation was a list of signs. Furthermore, the arguments of a predicate as well as of a modifier were given as a list of signs. Hence we defined the predicate `taus/1` operating on these lists. Each of the sign elements will be translated by `tau/2` into one target semantic.

```
@taus(list) => list.
taus([]) => [].
taus([#F|#R]) => [tau(#F, [])|taus(#R)].
```

The set of transfer rules can be partioned into the set of recursive rules for decomposing the source DRS called `tau/2`, and those rules called `tau_lex/{2,5}` defining lexical rules.

The definition of recursive transfer rules merely based on a recursion on the source DRS for decomposing semantic parts seems to be difficult. This is because there exists no link from the conditions set to a construction history which reflects the parts belonging to one lexical or phrasal semantic unit which was composed with the parts of the other units.

Therefore we prefered an approach which uses the transfer rules to recurse along the hierarchical structure of the phrasal skeleton of the source sign using the semantic projection line. The definition of semantic heads is given by the definition of the Semantics Principle of the HPSG grammar. The recursion along the semantic head projection bottoms out if a tree label carrying lexical information (hence of type `word_s`) is reached. In this case the predicate `tau_lex/2` (see below) is called:

```
tau(word_s & #Sign, #Args) => tau_lex(#Sign, #Args).
```

If the source sign is a phrasal one (of type `phrase_s`), we have to distinguish between different cases for finding the semantic head projection depending on the local constituent structure. We can distinguish the structures which are head argument structures (of type `head_comp_struc`) and those which are not.

1. The non-head argument structures can be divided into those which make a distinction between the syntactic and semantic head and those which do not.

[2]Because of the functional writing of relations in STUF-III we write the predicate `R(X,Y)` of arity 2 as unary one `R(X) & Y` with the implicit argument `Y`, the "result" argument. For a predicate `Pred` with arity `Arity` we write `Pred/Arity` according to functional writing, i.e. we would write `R/1` for the relation above.

(a) Adjuncts are semantic functors, hence the recursion follows the adjunct's daughter which is the value of the feature `adj_dtr`. For topicalized adjuncts in head topic structures (of type `head_topic_struc`) and in the head possessive genitive structures (of type `head_possgen_struc`) the recursion leaves the syntactic head projection, too, following the value of `top_dtr` and `pg_dtr`, respectively.

```
tau(phrase_s &
    dtrs:(head_adj_struc &
          adj_dtr: #SemHead), #Args) => tau(#SemHead, #Args).
tau(phrase_s &
    dtrs:(head_topic_struc &
          top_dtr: #SemHead), #Args) => tau(#SemHead, #Args).
tau(phrase_s &
    dtrs:(head_possgen_struc &
          pg_dtr: #SemHead), #Args) => tau(#SemHead, #Args).
```

(b) For all the other non head argument structures (head filler and head conjuncts structures with types `head_filler_struc` and `head_conjs_struc`, respectively) the semantic head is the same as the syntactic head (the value of `h_dtr`).

```
tau(phrase_s &
    dtrs:((head_filler_struc ; head_conjs_struc) &
          h_dtr: #SemHead), #Args) => tau(#SemHead, #Args).
```

2. In case of head argument structures we have to distinguish between several cases as well because of the verb movement analysis for German in the Verbmobil syntax.

(a) In case of a syntactic head in verb first or second position (type `word_s`, `sent_type(v1;v2)` and no trace, i.e. `phon:cons`[3]), i.e. the verb was moved from its base position, we have to recurse along the semantic projection via the complement to make sure to find possible adjuncts in the Mittelfeld in front of the finite verb trace.

```
tau(phrase_s &
    dtrs:(head_comp_struc &
          h_dtr: (word_s &
                  phon:cons &
                  sent_type(v1;v2)) &
          c_dtrs: [#SemHead]), #Args) => tau(#SemHead, #Args)
```

(b) In all other cases the semantic and the syntactic heads are the same as one would expect for a standard Semantics Principle of HPSG.

i. Verb Final Position (`sent_type(v3)`):

[3] The feature term `phon:cons` is logically equivalent to `phon: ~ []` with negation operator `~` and the empty list `[]`.

```
tau(phrase_s &
    dtrs: (head_comp_struc &
           h_dtr: (word_s &
                   #SemHead &
                   phon:cons &
                   sent_type(v3))),
    #Args) => tau(#SemHead, #Args).
```

ii. Non-verbal Lexical Head (`category(~verb)` and `phon:cons`):

```
tau(phrase_s &
    dtrs: (head_comp_struc &
           h_dtr: (word_s &
                   #SemHead &
                   phon:cons &
                   category(~verb))),
    #Args) => tau(#SemHead, #Args).
```

iii. Phrasal Projection (type `phrase_s`):

```
tau(phrase_s &
    dtrs: (head_comp_struc &
           h_dtr: (phrase_s  &
                   #SemHead)),
    #Args) => tau(#SemHead, #Args).
```

iv. Trace (`phon:[]`):

```
tau(phrase_s &
    dtrs: (head_comp_struc &
           h_dtr: (word_s &
                   #SemHead &
                   phon:[])),
    #Args) => tau(#SemHead, #Args).
```

In the Verbmobil semantic construction there are two special cases of semantic representation, namely time structures (of type `time_struc`) and spelling structures (of type `spell_struc`). In these structures the semantic head information is directly mapped to the maximal phrase projection because of the non-compositional semantic treatment within this constituent, hence we can stop the recursion there.

```
tau(phrase_s &
    #SemHead &
    dtrs:time_struc, #Args) => tau_lex(#Sign, #Args).
tau(phrase_s &
    #SemHead &
    dtrs:spell_struc, #Args) => tau_lex(#Sign, #Args).
```

Obviously, with the `tau/2` predicate defined above we can traverse the phrasal structure of a sign to get to the semantic head of it. At this point the `tau_lex` relation defined in the following plays an important rôle.

2.3 The Predicates tau_lex/{2,5} and sem_lex/2

On the one hand, tau_lex determines for a given lexical target sign a – not necessarily lexical – target DRS representation. On the other, it triggers the recursion with the arguments of the semantic predicate or the functor. In the easiest case we can relate a unique predicate name of the source and the target language, respectively, and additionally formulate local and/or global contextual conditions restricting the application of a rule. We will limit the following description mainly to this case for presentation purposes.

```
@tau_lex(sign,list) => sign.
tau_lex(pred_name(#SourcePred)&#Sign,#Args) =>
        tau_lex(#SourcePred,#TargetPred,#SemArgs,#Args,#Sign) &
        sem_lex(#TargetPred,#SemArgs).
```

The predicate tau_lex/2 extracts the predicate name of the source sign with the ADT pred_name/1. The source predicate name is the first argument of the tau_lex/5 predicate which defines the substantial part of lexical transfer rules. Using tau_lex/5 successfully instantiates the target predicate name in the second and the target predicate's arguments in the third argument. Furthermore it triggers the target language semantic construction by calling sem_lex/2.

In the following sections there are a lot of examples defining tau_lex/5 and sem_lex/2 clauses. E.g. the following example was taken from the transfer of verbs (see section 4 for further explanation).

```
tau_lex(fahren,go,
        [tau_tense(#Tense),tau(#S)], [],
        intrans_syn(#S,#Tense)) => sem_t.

sem_lex(go,#Args) =>
        intransitive_verb_sem1(go,
                               dynamisch_c,agent_rel,person_c,
                               #Args).
```

2.4 The Predicates tau_prep/4 and tau_pred/3

There was a special case of conceptual transfer for prepositions we want to mention here. It was triggered by the following tau_lex/2 clause:

```
tau_lex(category(prep) & #Sign, #Args) =>
        tau_lex(prep_pred,#TargetPred,#SemArgs,#Args,#Sign) &
        sem_lex(#TargetPred,#SemArgs).
```

The corresponding tau_lex/5 clause was

```
tau_lex(prep_pred,
        #TargetPred,
        [#ConcRel,tau(#VP,#Args),tau(#NP)&pred_sort(#NPSort)],
        #Args,
```

```
        prep_syn(#VP&pred_sort(#ExtSort),
                 #NP&pred_sort(#IntSort)) &
        sem:prep_sem1(#SourcePred,#ConcRel,_) &
        tau_prep(#SourcePred,#ConcRel,#ExtSort,#IntSort)) =>
    tau_pred(#ConcRel,#TargetPred,#NPSort).
```

`tau_prep/4` determines a conceptual relation (`#ConcRel`) based on a semantic predicate's name an the semantic sort restrictions of the internal and external arguments of a preposition. On the other hand, `tau_pred/3` uses the sort of the internal argument of the target semantic (`#NPSort`) and the named conceptual relation to determine the target predicate.

Since indexing does not work these concepts defined in type hierarchies it would have been better for efficiency reasons to partially evaluate all `tau_prep/4` and `tau_pred/3` rules into different `tau_lex/5` rules.

Chapter 3

The Transfer of Nouns

Nouns are the least problematic noun category in the Verbmobil translation system. Translation ambiguities within the narrow domain of scheduling of appointments are rare. Problems mainly arise with the words *Termin* and *Essen*.

3.1 Rigid Nouns

Rigid nouns are translated according to the semantic predicate which is assigned to them by the semantic construction (e.g. *dofw* for a name of a weekday). The semantic predicate functions as an interlingua predicate and contains an additional value slot (e.g. which particular day of the week is meant). The value slot is not translated at all; it is just handed over to generation. The transfer rules for the semantic predicate is trivial since it uses the same technical predicate for both German and English, namely `rigid_noun_pred`. The subgoal `tau_rigid_rigid` is used to pass on the rigid noun predicate, the rigid noun value slot and its sort.

```
tau_lex(rigid_noun_pred,rigid_noun_pred,[],[],#SourceSign) =>
        tau_rigid_rigid(#SourceSign).
```

3.2 Common Nouns

Common nouns were the one noun category which contained ambiguous words. Unambiguous common nouns are translated by the simple rule pattern exemplified here for the German word *Büro*. The semantic predicate *buero* is just replaced by the English predicate *office*.

```
tau_lex(buero,office,[],[],#SourceSign) => tau_nn_nn(#SourceSign).
```

Disambiguation of ambiguous nouns is accomplished using sortal restrictions. The noun *Termin* has three possible English translations within the Verbmobil domain: *date*, *appointment* and *slot*. Which one has to be chosen depends on the embedding verb. If the German verb *ausmachen* is translated to *make*, then *Termin* has to be translated by *appointment*. If *fix* is chosen instead as the translation of *ausmachen*, the correct translation of *Termin* is *date*.

Three disjunct sorts from the sort hierarchy were assigned to the different translation possibilities. The relevant part of the sort hierarchy is:

```
temporal_c       = zeit_c || situation_c.
zeit_c           = zeit_punkt_c || zeit_intervall_c.
zeit_intervall_c = tageszeit_c || zeit_periode_c || monate_c || tage_c ||
                   saison_c || feiertag_periode_c || tage_periode_c.
situation_c      = mental_sit_c || funkt_sit_c || kommunikat_sit_c ||
                   termin_sched_sit_c || termin_fix_sit_c || beweg_sit_c ||
                   rest_sit_c.
veranstaltung_c  < funkt_sit_c.
```

The possible sorts assigned to *Termin* are *veranstaltung_c*, *zeit_punkt_c* and *zeit_periode_c*. The transfer rules for *Termin* check to which sort the discourse referent belongs. If it belongs to sort *zeit_punkt_c* e.g., *Termin* is translated into *date*.

```
tau_lex(termin,appointment,      [],[],pred_sort(subsumes(veranstaltung_c))) => sem_t.
tau_lex(termin,date,             [],[],pred_sort(subsumes(zeit_punkt_c))) => sem_t.
tau_lex(termin,appointment_slot,[],[],pred_sort(subsumes(zeit_periode_c))) => sem_t.
```

Additionally, the transfer rules for all words which take the word *Termin* as an argument must specify the sort of their argument in order to disambiguate its translation. Hence, the disambiguation of the argument is a side effect of the translation of the predicate.

```
tau_lex(ausmachen, arrange,
                [tau_tense(#Tense), tau(#S), tau(#O)], [],
                trans_syn(#S, #O &
                sem:cond_sem(basic_cond(termin,marker,[])) &
                pred_sort(veranstaltung_c),
                #Tense)) => sem_t.

tau_lex(frei, free, [tau(#VP,#Args)], #Args,
                adv_syn( #VP &
                sem:cond_sem(basic_cond(haben1,marker,list)) &
                pred_args([##,sem:cond_sem(basic_cond(termin,marker,list)) &
                              pred_sort(zeit_periode_c)]))) => sem_t.
```

3.3 Relational Nouns

The groups of the relational nouns contains *Anfang*, *Ende*, *Mitte* and *Hälfte*. Their predicate is translated exactly like common nouns. They differ from common nouns, in that they subcategorize for an optional genitive argument, which needs to be translated recursively by means of the subgoal `tau_rn_rn`.

```
tau_lex(anfang,beginning,#Tau,[],#SourceSign) => tau_rn_rn(#SourceSign,#Tau).
```

The subgoal `tau_rn_rn` extracts the predicate arguments, i.e. only the genitive argument and constructs the semantic argument list which comprises the external sort of the relational noun, `of_rel` (the genitive case role), the sort of the argument and the recursive translation of the argument.

```
tau_rn_rn(pred_args(#Args) &
          category(~modifier) &
          sem:cond_sem(basic_cond(atom,sort:#Sort,[arg:sort:#ArgSort])),
          [#Sort,of_rel,#ArgSort|taus(#Args)]) => sem_t.
```

3.4 Proper Names

Translation of proper names was unproblematic since they were all unambiguous. The corresponding transfer rule, therefore, just maps the interlingual semantic predicate to the English one:

```
tau_lex(propername,propername,[],[],#SourceSign) => tau_ne_ne(#SourceSign).
```

The subgoal `tau_ne_ne` make use of a translation table for mapping the German proper names into to the corresponding English ones.

3.5 Pronouns

This section will give a short overview of the way pronoun translation was handled in the MDS.

- Pronouns were already recognized and classified in Semantic Construction mainly on the basis of syntactic information.
- Semantic Evaluation of the MDS could not resolve anaphora.
- Therefore Transfer passed the structures it got from Semantic Construction on to Generation.

While the practical implementation is not particularly exciting from a research perspective, some theoretical work was devoted to the question which component should be responsible for triggering anaphora resolution (Eberle et al. 1994). To make a long discussion short, from our point of view anaphora involves so much genuinely syntactic information (like gender) that the Transfer Component, which should not have detailed knowledge about the target language, is unable to do the job. In view of the prototype we think that the problems of anaphora resolution should be settled in close interaction between Semantic Construction, Semantic Evaluation and Generation, but that Transfer will not have much to say about it.

First we list the classes of pronouns provided by Semantic Construction.

Macro	Type	Pred	Examples
deictic_pronoun	std	speaker	ich
deictic_pronoun	reflexive	speaker	mir,mich
deictic_pronoun	std	hearer	du,dir,dich,ihr,sie,ihnen
deictic_pronoun	reflexive	hearer	dir,dich,euch
deictic_pronoun	std	speaker_hearer	wir
deictic_pronoun	reflexive	speaker_hearer	uns
pronoun	demonstr		der,etc.; derjenige,etc.
pronoun	demonstr	loc_near_pred	dieser,etc.
pronoun	demonstr	loc_far_pred	jener,etc.
pronoun	comp		derselbe,etc.
pronoun	anaphor		er,sie,es,ihm,ihn,ihr,ihnen,man
pronoun	reflexive		sich,selbst,selber
ellipsis_pronoun	event_type		das,es
uni_pronoun	no	something_pred	nichts
uni_pronoun	no	person_pred	niemand,etc.
uni_pronoun	every	something_pred	alles,allem
exist_pronoun		something_pred	(et)was
exist_pronoun		person_pred	jemand,etc.
expletive_pronoun	reflexive		sich,mich,uns

While the treatment of pronouns in the MDS may not be too upsetting it is in some respect rather instructive. We tried out two methods of transferring "interlingual" information, as constituted by the semantic analysis of pronouns.

- One method was to have one transfer hook per semantic macro. The information needed for further classification then had to be retrieved by calling some interface predicates.

```
tau_lex(pronoun,pronoun,[#Marker,#AlfaType],[],
        sem:cond_sem(alfa_cond(#Marker,#AlfaType,[]))) => sem_t.
```

- In another case we specified a different transfer hook for each pronoun class. So we could directly create the semantic contribution to the target DRS

```
tau_lex(alles  ,alles  ,[],[],sem:##) => ##.
```

3.6 Determiners

The transfer treatment of determiners in the MDS was again quite plain. We simply took over the classification proposed by Semantic Construction.

Macro	Type	Pred	Examples
exist_det			andere,etc.;ein,etc.;einige,etc.;etliche,etc.
ref_det	anaphor		der,etc.
ref_det	demonstr		derjenige,etc.
ref_det	comp		derselbe,etc.
ref_det	demonstr	loc_near_pred	dieser,etc.
ref_det	demonstr	loc_far_pred	jener,etc.
uni_det	all		alle,etc.
uni_det	every		jeder,etc.;saemtliche,etc.
uni_det	no		kein,etc.
uni_det	many		viele,etc.
uni_det	some		mehrere,etc.
uni_det	few		wenige,etc.
uni_det	wh		welche,etc.
poss_det		speaker/ind	mein,etc.
poss_det		hearer/ind	dein,etc.;ihr,etc.
poss_det		hearer/coll	ihr,etc.;euer,etc.
poss_det		[]/ind	sein,etc.;ihr,etc.
poss_det		[]/coll	ihr,etc.
poss_det		speaker_hearer/coll	unser,etc.

All the features were first peeled out of the German structure and then transmitted to the target semantic lexicon. In line with the general head-driven approach, the determiners trigger transfer of the noun they occur on. A slight complication arose from the treatment of NP to PP raising in Semantic Construction.

Ich komme diesen Dienstag.

These raising rules were modelled in the lexical semantics of the (referential) determiners. Correspondingly Transfer introduced new hooks that allowed a smooth construction on the target side. For raised referential determiners we had to call the English Semantic lexicon twice: once for the referential determiner, and once for the NP to PP raising.

Chapter 4

The Transfer of Verbs

4.1 General Remarks

For the Verbmobil demonstrator, we adopted a sign-based transfer approach. This meant that lexical rules were used to translate lexical verbal heads. Based on the predicate name of a verb (*pred_name*), a lexical transfer rule establishes a transfer relation (*tau_lex*) between the corresponding lexical sign of the source and the target predicate.

The *tau_lex*-relation determines the translation of the lexical signs. It also recursively defines the local constraints for the application of a given relation by determining the valency patterns as well as the sortal constraints on verb arguments to differentiate between readings. Other context conditions can be stated as well as can be seen in the examples below. Verb readings which do not clearly fit in the VM domain have not been accounted for. The general format of basic lexical transfer rules for verbs is illustrated below:

```
tau_lex(#SourcePred,#TargetPred,#SemArgs,#Args,#SourceSign) => sem_t.
```

with the first argument being the source predicate name, the second argument being the goal predicate name, the third argument being a list of the semantic goal arguments, the fourth argument being a list of arguments passed down the tree (cf. some examples below), and the fifth argument being the entire source sign.[1]

Part of the transfer work also consisted in determining the lexicon entries for the English semantic construction. The predicates of the target language semantic construction lexicon are represented by semantic macros. They correspond to those used in the source language (SL) semantic dictionary. Once again, the mapping is based on the predicate name of a given verb, which is considered as commonly defined access predicate between the analysis, transfer, and generation modules.[2] Below, the general format of the entries of the target language (TL) semantic lexicon

[1]The complete number of transfer rules for lexical verbal heads is listed in the file *tau_verb.stuf* of the rule documentation available at the IMS, University of Stuttgart.

[2]The lexicon entries are listed in the file *e_sem_lex_verb.stuf* of the complete IMS rule documentation.

predicates is illustrated for intransitive, transitive, and ditransitive verbs:

a. Intransitive verbs

```
sem_lex(#TargetPred, #Args) =>
        intransitive_verb_sem1(#PredName, #Sort, #SubjRole,
                #SubjSort, #Args).
```

b. Transitive verbs

```
sem_lex(#TargetPred, #Args) =>
        transitive_verb_sem1(#PredName, #Sort, #SubjRole,
                #SubjSort, #ORole, #ObjSort, #Args).
```

c. Ditransitive verbs

```
sem_lex(#TargetPred, #Args) =>
        ditransitive_verb_sem1(#PredName, #Sort, #SubjRole,
                #SubjSort, #ObjRole, #ObjSort, #Obj2Role,
                #Obj2Sort, #Args).
```

In the following, examples for the treatment of main verbs, copula, and modal verbs in the MDS will be given.

4.2 Main Verbs

In this section examples for simple transfer rules will be presented, followed by the illustration of more complex transfer statements. For each example, first the transfer rule is given and then the corresponding entry in the English semantic lexicon. Under the heading 'Context' the conditions to be fulfilled to trigger the respective rules are stated. They are followed by an example from the Verbmobil dialogues.[3]

4.2.1 Standard Cases

Here we give simple transfer rules for intransitive, transitive, and ditransitive verbs. These consist of a simple mapping between the source predicate's argument structure and the target predicate's argument structure.

1. Intransitive Verbs

- **fahren**
  Rule:

```
tau_lex(fahren,go,
        [tau_tense(#Tense),tau(#S)], [],
        intrans_syn(#S,#Tense)) => sem_t.
```

[3]The numbering of the dialogues refers to the original reference numbers. (These were changed in December 1994).

```
sem_lex(go,#Args) =>
        intransitive_verb_sem1(go,dynamisch_c,
                agent_rel,person_c,#Args).
```

<u>Context:</u>

intransitive *fahren* with *agent*.

<u>Example:</u>

X fährt $\longleftrightarrow$ X goes
DD3:05: *Ich wollte am Samstag den siebzehnten Juli in Urlaub fahren.*
ED3:05: *I wanted to go on vacation on Saturday the seventeenth.*

2. Transitive Verbs

- **erledigen**
  <u>Rule:</u>

```
tau_lex(erledigen, settle,
        [tau_tense(#Tense), tau(#S), tau(#O)], [],
        trans_syn(#S, #O, #Tense)) => sem_t.

sem_lex(settle, #Args) =>
        transitive_verb_sem1(settle,dynamisch_c & funkt_sit_c,
                agent_rel,person_c,theme_rel,entitaet_c,#Args).
```

<u>Context:</u>

transitive *erledigen* with *agent* and *theme*.

<u>Example:</u>

X erledigt Y $\longleftrightarrow$ X settles Y
DX4:18: *Ich dachte, das wäre in einer halben Stunde erledigt.*
EX4:18: *I thought that could be settled within half an hour.*

3. Ditransitive Verbs

- **legen**
  <u>Rule:</u>

```
tau_lex(legen, schedule_for,
        [tau_tense(#Tense), tau(#S), tau(#O), tau(#O2)], [],
        ditrans_syn(#S, #O, #O2, #Tense)) => sem_t.

sem_lex(schedule_for,#Args) =>
        ditransitive_verb_sem1(schedule_for,dynamisch_c,agent_rel,person_c,
                theme_rel,temporal_c,for_rel,temporal_c,#Args).
```

Context:

ditransitive *legen* with *agent*, *theme*, and *goal*.

Example:

X legt Y auf Z $\longleftrightarrow$ X schedules Y for Z
DD3:12: *Könnten wir das vielleicht auf den Vormittag legen.*
ED3:12: *Could we perhaps schedule that for the morning.*

4.2.2 Special Cases

So far, we have shown simple transfer cases where no supplementary conditions have to be fulfilled to trigger the respective transfer rule. However, in the Verbmobil dialogues a wide range of more problematic transfer cases occurred which had to be dealt with in the MDS; we give a small number below.[4]

- **ausmachen**
  Rule:

```
tau_lex(ausmachen, arrange,
        [tau_tense(#Tense), tau(#S), tau(#O)], [ ],
        trans_syn(#S, #O & pred_name(termin)) &
                           pred_sort(veranstaltung_c),
                  #Tense)) => sem_t.

sem_lex(arrange,#Args) =>
        transitive_verb_sem1(arrange,dynamisch_c,
                agent_rel,person_c,theme_rel,temporal_c,#Args).
```

Context:

transitive *ausmachen* with *theme* not equal to (zeit_punkt_c, zeit_periode_c, object_c).

Example:

X macht Y aus $\longleftrightarrow$ X arranges Y
DD1:1: *Ja, prima, dann lassen Sie uns doch noch einen Termin ausmachen.*
ED1:1: *Yes, fine, then let's arrange an appointment.*

In order to account for selectional restrictions, the semantic access predicate *pred_sort* has been provided. In the slightly simplified rule given above the sortal restriction *veranstaltung_c* will trigger the correct translation of *Termin* as *appointment* and not as *date* or *appointment-slot*, cf. the corresponding transfer rules for *Termin* below.

Rule:

[4] For a more complete description see Reinhard (1995).

```
tau_lex(termin,appointment,[],[],
        pred_sort(subsumes(veranstaltung_c))& #SourceSign)
=> tau_nn_nn(#SourceSign).

tau_lex(termin,date,[],[],
        pred_sort(subsumes(zeit_punkt_c))& #SourceSign)
=> tau_nn_nn(#SourceSign).

tau_lex(termin,appointment_slot,[],[],
        pred_sort(subsumes(zeit_periode_c))& #SourceSign)
=> tau_nn_nn(#SourceSign).
```

Thus, a mutual disambiguation between the verb and one of its arguments takes place, since the verb itself also has multiple translations, e.g. *arrange for* or *switch off*, etc.

- **einfallen**
  Rule:

```
tau_lex(einfallen, remember,
        [tau_tense(#Tense), tau(#O), tau(#S)], [],
        trans_syn(#S, #O, #Tense)) => sem_t.

lex(fallen) =>
        transitive_prefix_verb_sem(ein,einfallen,dynamisch_c&mental_sit_c,
                theme_rel,entitaet_c,experiencer_rel,person_c).

sem_lex(remember,#Args) =>
        transitive_verb_sem1(remember,statisch_c,
                experiencer_rel,person_c,theme_rel,entitaet_c,#Args).
```

  Context:

  transitive *einfallen* with switched *experiencer* and *theme* arguments.

  Example:

  X fällt Y ein $\longleftrightarrow$ Y remembers X
  DD3:16: *ach da fällt mir was ein - zu der Zeit muss ich zum Zahnarzt.*
  ED3:16: *oh, I remember something - at that time I have to go to the dentist's.*

The translation of *einfallen* involves the treatment of a thematic divergence, i.e. the theme which is the subject in German has to be realized as the object in English. This argument diathesis is shown in the rule above. Here also the entry of the German semantic lexicon is given to show how seperable prefix verbs are dealt with.

- **treffen**
  Rule:

```
tau_lex(treffen1,meet1,
        [tau_tense(#Tense),tau(#S)],[],
        trans_syn(#S & deictic_pronoun(std) & sem:number_sem(collective),
        deictic_pronoun(reflexive) & sem:number_sem(collective),
        #Tense)) => sem_t.

sem_lex(meet1,#Args) =>
        intransitive_verb_sem1(meet1,dynamisch_c,
                agent_rel,person_c,#Args).

tau_lex(treffen1,meet,
        [tau_tense(#Tense),tau(#S),tau(#O)],[],
        trans_syn(#S,
        #O & (deictic_pronoun(~reflexive);pred_name(~deictic_pronoun)),
        #Tense)) => sem_t.

sem_lex(meet,#Args) =>
        transitive_verb_sem1(meet,dynamisch_c,
                agent_rel,person_c,theme_rel,entitaet_c,#Args).
```

  Context:

  transitive *treffen* with x(plural) and REFL(plural).

  Example:

  X treffen Y $\longleftrightarrow$ XandY meet
  DREF:028: *wir treffen uns dann in der Eingangshalle des Czerczinsky mit den Unterlagen.*
  ED3:16: *we will meet in the lobby of the Czerczinsky with the papers.*

Two rules for the translation of *treffen* had to be implemented. This is due to the fact that *treffen* can either be translated with monovalent *meet1*, in case the first argument position is taken by a plural deictic pronoun and the second by a plural reflexive pronoun, or it can be translated by bivalent *meet* when there are different referents for the subject and object positions respectively.

- **kommen**
  Rule:

```
tau_lex(kommen, come,
        [tau_tense(#Tense), tau(#S)], [],
        intrans_syn(#S, #Tense)) => sem_t.

sem_lex(come,#Args) =>
```

```
        intransitive_verb_sem1(come,dynamisch_c,agent_rel,person_c,#Args).

tau_lex(whq_adv,whq_adv,[#WhqType,tau(#Mod,[get1|#Args])],#Args,
        phon:[wie] &
        mod([#Mod&sem:#ModSem&pred_name(kommen)]) &
        sem:whq_adv_sem1(#WhqType,[#ModSem])) => sem_t.


tau_lex(kommen, get1,
        [tau_tense(#Tense), tau(#S)], [get1],
        intrans_syn(#S, #Tense)) => sem_t.

sem_lex(get1,#Args) =>
        intransitive_verb_sem1(get1,dynamisch_c,agent_rel,person_c,#Args).
```

Context:

intransitive *kommen* as *get* in context of *wie*.

Example:

X kommt vs. Wie kommt X to Y? ⟷ X comes vs. How does X get to Y?
DD3:22: *ich komme dann Freitag um elf.*
ED3:22: *I'll come on Friday at eleven.*
DD2:9: *wenn Sie mir noch kurz erklären, wie ich zu Ihnen komme.*
ED2:9: *if you could just briefly explain how I get to your place.*

For *kommen* two slightly different readings have to be distinguished, since they yield two different translations. The first translation is the simple case where *kommen* is translated by *come* (compare the first *tau_lex* rule above). In the second case, *kommen* is translated with *get* in the context of a wh-question, here *wie*; (see the second and third transfer rule above). Note the use of the list on the fourth argument position to percolate an argument (here *wie*) down the tree, as mentioned in the beginning.

This example is quite complex since it requires yet another transfer rule to account for the restructuring needed for the correct translation of *zu Ihnen* into *to your place* and not into *to you.* Thus the personal pronoun has to be mapped onto a possessive pronoun modifying a noun. The corresponding rule is given below:

Rule:

```
tau_lex(zu2, to,
        [##, tau(#VP,#Args),
         poss_det_sem1(of3, of_rel, #Sort,#DiscRole,#SemNum,
         [common_noun_sem1(place,dm_sort)])], #Args,
        sem:cond_sem(basic_cond(kommen,marker,list)) &
        prep_syn(#VP,
                 sem:number_sem(#SemNum) &
```

```
                 deictic_pronoun_sem(std;reflexive,#DiscRole,#Sort)))
=> sem_t.
```

- **in Frage kommen**

  Rule:

```
tau_lex(kommen_in_frage, be1,
        [tau_tense(#Tense),
         sem_lex(possible,
                [pp2np(sem_lex(for,[for_rel,#O])), tau(#S)])], [#O],
        intrans_syn(#S, #Tense)) => sem_t.

sem_lex(be1,#Args) =>
        copula_verb_sem1(be1,#Args).

lex(kommt) =>
        intransitive_prefix_verb_sem(in_frage,kommen_in_frage,statisch_c&
                mental_sit_c,theme_rel,entitaet_c).
```

  Context:

  SVC *in Frage kommen* as SVC.

  Example:

  X kommt in Frage (bei Y) $\longleftrightarrow$ X is possible (for Y)
  DRef:DE011: *als Ausweichmöglichkeit bei mir kommt wieder in Frage zwischen dem fünfzehnten und neunzehnten.*
  ERef:DE011: *as an alternative, between the fifteenth and the nineteenth is possible for me.*

Here, as well as in the next example *einen Vorschlag machen*, we show the translation of German support verb constructions (SVCs). Their translation can yield different target structures depending on various conditions, see also Winhart (1995).

A SVC can be lexicalized and thus have a proper entry in the transfer lexicon since it represents one single translational unit. Or the SVC is not lexicalized and thus a compositional translation of the respective parts results. An example of the first type is the idiomatic *in Frage kommen*. In the source semantic lexicon, *in Frage* is encoded as a separable prefix of *machen2* because it is a non-modifiable constituent of the SVC, compare the entry above. The translation will yield a copula construction, see the transfer rule above. In the translation of the phrase *der Freitag kommt bei mir in Frage*, the adjunct is passed down the tree and inserted into the dimension condition supplied by the adjective.

- **einen Vorschlag machen**

  Rule:

```
tau_lex(machen2, make2,
        [tau_tense(#Tense), tau(#S),
```

```
          tau(#O)& cond_sem(basic_cond(suggestion,marker,[]))], [],
         trans_syn(#S, #O, #Tense)) => sem_t.

tau_lex(machen2, suggest,
         [tau_tense(#Tense), tau(#S), tau(#O) &
              cond_sem(basic_cond(date,marker,[]))], [],
         trans_syn(#S, #O &
              sem:cond_sem(basic_cond(terminvorschlag,marker,[])),
         #Tense)) => sem_t.

sem_lex(make2,#Args) =>
         transitive_verb_sem1(make2,statisch_c,agent_rel,person_c,
                         theme_rel,entitaet_c,#Args).

tau_lex(machen1, do,
         [tau_tense(#Tense), tau(#S), tau(#O)], [],
         trans_syn(#S, #O, #Tense)) => sem_t.

sem_lex(make1,#Args) =>
         transitive_verb_sem1(make1,dynamisch_c,agent_rel,person_c,
                         theme_rel,entitaet_c,#Args).
```

Context:

SVC *(Termin-)Vorschlag machen* verbalized or as SVC.

Example:

X macht einen Vorschlag vs. einen anderen Terminvorschlag $\longleftrightarrow$
X makes a suggestion vs. X suggests another date.
DD10:10: *machen Sie mir doch mal ein' Vorschlag bitte.*
ED10:10: *then why don't you make a suggestion please.*

DD3:17: *können Sie noch einen anderen Terminvorschlag machen.*
ED3:17: *can you suggest another date.*

In contrast, *einen Vorschlag machen* is an example of a non-lexical SVC since the nominal part can be modified in various ways. The unmodified version is realized as a SVC, i.e. *make a suggestion*, also in English, cf. the first rule above. In case of a compound noun with a deverbal second constituent, the SVC can't be translated with an English SVC but with its verbalized form. The incorporated object of the compound noun then takes the object position in the target language. As compounds couldn't be treated compositionally for the MDS, the target object had to be introduced into the transfer rule directly, see above. To capture the difference between *machen* in a SVC such as *einen Vorschlag machen* and *machen* in *eine Geschäftsreise machen*, there are two different lexical entries for static *machen2* and dynamic *machen1* and the English equivalents *make2* and *make1*, cf. the entries above.

4.3 Copula

For the treatment of copula, the predicative and the equality reading have to be distinguished.

- **sein**

Examples of the predicative reading of the copula *sein* which occurred in the Verbmobil dialogues are *schlecht sein*, *da sein*, *okay sein*. These are translated compositionally and covered by the following transfer rule.

Rule:

```
tau_lex(sein1,be1,[tau_tense(#Tense),tau(#O)],[],
        trans_syn(#S,#O & mod([#S]),#Tense)) => sem_t.
```

Example:

X ist schlecht ⟷ X is bad
DD6:10: *Sechzehn Uhr dreissig ist nicht schlecht.*
ED6:10: *Four-thirty is not bad.*

If these phrases take an additional PP complement in argument position, the construction is analyzed as a ditransitive copula, which is covered by the transfer rule below:

Rule:

```
tau_lex(sein1,be1,[tau_tense(#Tense),tau(#Adj)],[],
        ditrans_syn(##,##,#Adj,#Tense)) => sem_t.
```

Example:

X ist schlecht bei/für Y ⟷ X is bad for Y
DD6:05: *Das ist bei mir schlecht.*
ED6:05: *That is somewhat bad for me.*

- **lieber sein**

  Rule:

  ```
  tau_lex(sein1, prefer,
          [tau_tense(#Tense), tau(#O), tau(#S)], [],
          ditrans_syn(#S, #O, ad_dimen_sem(lieb1,dimen_val), #Tense))
   => sem_t.
  ```

  Context:

  copula construction *lieber sein* as main verb plus head switching plus thematic divergence.

  Example:

X ist Y lieber $\longleftrightarrow$ Y prefers X
DD3:10: *Mir wäre Donnerstag der achte Juli eigentlich lieber.*
ED3:10: *I would actually prefer the eighth of July.*

The translation of the source language copula construction *lieber sein* constitutes a particular case since it is translated with a transitive main verb. The category change is combined with head switching plus a thematic divergence between subject and object arguments. The transfer relation thus has to account for these divergencies.

In contrast to the predicative reading, the equality reading of copula *sein* has a noun complement instead of an adjectival one. The transfer rule below shows the treatment of these cases in the MDS.

Rule:

```
tau_lex(sein1,be1,[tau_tense(#Tense),tau(#S),tau(#O)],[],
        trans_syn(#S,#O & np_syn,#Tense)) => sem_t.
```

Example:

X ist Y $\longleftrightarrow$ X is Y
DRef:23: *Ist Allerheiligen nicht ein Feiertag bei Ihnen?*
ERef:23: *Isn't All Saints' Day a holiday for you?*

- **werden**

The copula *werden* has been analyzed as a ditransitive verb when it takes a predicative AP complement and a PP complement in argument position, cf. the second rule below. The *tau*-relation is specified for the predicative complement only; subject and modifier position are shared due to the semantic analysis. The first rule is for copula *werden* taking a predicative AP complement only.

Rule:

```
tau_lex(werden1,be1,[tau_tense(#Tense),tau(#O)],[],
        trans_syn(##,#O,#Tense)) => sem_t.

tau_lex(werden1, be1,
        [tau_tense(#Tense),tau(#Adj)], [],
        ditrans_syn(##,##,#Adj,#Tense)) => sem_t.

sem_lex(be1, #Args) =>
        copula_verb_sem1(be1,#Args).
```

Context:

Ditransitive *werden* with AP and PP complement.

Example:

X knapp/eng (Y) werden bei/für Z $\longleftrightarrow$ X be tight (Y) for Z.
DD1:14: *Das wird etwas knapp bei mir.*
ED1:14: *That will be a bit too tight for me.*

4.4 Modal Verbs

Modals are mainly used to indicate the speaker's attitude towards what he/she is saying, or when the speaker is concerned about the effect of what he/she is saying on the person he/she is speaking or writing to. As modals are the following verbs classified by Semantic Construction:

dürfen, müssen, sollen, können, wollen.

The modality as it is treated in the MDS qualifies the propositional content of an utterance only along the lines of necessity and possibility. As the general rule below shows the structure of the source language modal predicate is directly mapped onto the target language predicate.

```
tau_modal(#Pred,[#Op,tau_tense(#Tense),tau(#VP,#Args)],#Args,
          modal_syn(#VP&sem:#VPSem,#Tense) &
          sem: modal_verb_sem1(#Pred,#Op,[#Tense,#VPSem])) => sem_t.
```

The following example shows the transfer rule for *können*:
Rule:

```
tau_lex(koennen, can, [#Op, tau_tense(#Tense), tau(#VP)], [],
         modal_syn(#VP&sem:#VPSem, #Tense) &
         sem:modal_verb_sem1(koennen, #Op, [#Tense, #VPSem])) => sem_t.
```

whereas the #Op = pos.
Context:

können i.e. to be possible
(cf. the value of #Op above)

Example:

X kann Y tun $\longleftrightarrow$ X can/could do Y
DD1:13: *Können wir uns dann nicht auf halb drei einigen.*
ED1:13: *Can't we agree on half past two then.*

4.5 Subjunctive

Similar to modality the use of subjunctive in our corpus has mainly the purpose to express politeness. The subjunctive introduces a possibility modality and is treated like modal verbs.

4.6 Pragmatics

Although we are dealing with spoken language the pragmatic aspects are not really considered in the demonstrator. The Semantic Evaluation component computes for a set of utterances the corresponding dialogue act. This information will then be transferred to Generation.

This information can be taken to disambiguate some readings (cf. 4.7) and therefore select the correct translation. The following shows an example in which way pragmatics can help to support the translation task.

4.7 Dialogue Verbs

In the demonstrator corpus – the Blaubeuren Dialogues – there is a class of verbs which function as dialogue control verbs, i.e. they are used to express/emphasis on the speaker's attitude towards his/her utterance (like modality). To this class of verbs belong *denken, glauben, finden, meinen, sehen etc.* but also *sagen, wissen etc.* For a sample rule see the rule for *meinen* below.
Rule:

```
tau_lex(meinen1,think1,
          [tau_tense(#Tense), tau(#S), tau(#O)], [],
          trans_syn(#S, #O, #Tense)) => sem_t.
```

Context:

These verbs are classified as "statisch" and "mental_sit" or "kommunikat_sit"; they are transitive and take a sentence complement as an object.

Examples:

X prop_attn_verb Y, whereas Y is a sentence

DD3:19: *Ich glaube das wäre bei mir auch in Ordnung?*
ED3:19: *I think that would be all right with me too?*

DD3:14 *Meinen Sie das reicht uns .*
ED3:14: *Do you think that will be enough.*

DD1:09: *Ich muß sagen, mir wär's dann lieber, wenn wir die ganze Sache ...*
ED1:09: *I must say I'd prefer if we could move the whole thing ...*

Besides their usage as dialogue verbs each verb of these class can also be used as "normal" performative verb as already described in 4.2. The transfer rule does not change but the translation equivalent; for instance, *meinen* is then translated to *mean*, *glauben* to *believe* etc.

Chapter 5

The Transfer of Adverbs and Adjectives

We will discuss our approach to handle *adverbs* and *adjectives* in the Transfer module of MDS together, as these two word classes intuitively show similar semantic properties. Most *particles* were included into the word class of *adverbs*.

5.1 Preliminary Remarks on *Adverbs* and *Adjectives*

The classification of the word list of VerbMobil was based on part of speech distinction, i.e. the lexemes were assigned the standard grammatical word classes[1]. *Adverbs* and *adjectives* were distributed into two word classes. Thus, two groups of properties were necessary to characterize *adverbs* and *adjectives* with respect to their differencies and with respect to their similarities:

1. morpho-syntactic properties which reflect the difference between *adverbs* and *adjectives* as parts of speech
2. semantic properties which reflect the relation between the semantic functions of *adverbs* and the semantic functions of *adjectives*

Some distinguishing marks[2] characterizing adverbs and adjectives with respect to the first group of properties are: a) adverbs are uninflectional, whereas adjectives are inflectional (ex.(2.1a)); b) adverbs and adjectives can be assigned comparison degrees (ex.(2.1b)); c) adverbs can apply to adjectives and verbs and build one constituent in the sentence structure (ex.(2.1c)), whereas adjectives only apply to nouns (ex.(2.1d)).

(5.1) (a) Peter kommt FRÜH(adv). vs. das FRÜHE(adj) Treffen

[1]The standard grammatical subgroups of parts of speech are: verbs, nouns, adjectives, numerals, pronouns, adverbs, prepositions, conjunctions, interjections, particles.

[2]For an exhaustive discussion on the distinguishing marks with respect to the first group of properties see Ralf Steinberger (Steinberger (1994)).

(b) FRÜHER(adv,comp) war Peter GLÜCKLICHER(adj,comp).
(c) Peter KOMMT(v) FRÜH(adv). vs. das WIRKLICH(adv) FRÜHE(adj) Treffen
(d) das FRÜHE(adj) Treffen(n)

The distinguishing marks with respect to the second group of properties show that *adverbs'* and *adjectives'* common structural property is that they apply to different semantic entities and modify them in providing restrictive information for their interpretation. For example, the adverb FRÜH(adv) in (2.1a) applies to the event described by the verb *kommen* and denotes its relative temporal location, whereas the adjective FRÜHE(adj) in (2.1a) applies to the individual refered to by the noun *Treffen* and specifies one of its relevant parameters – relative temporal location. Following the conception adopted by the Semantic Construction module, we define adjectives and adverbs semantically as modifiers.

Adverbs and adjectives in VerbMobil are divided into subclasses after their meaning and their compositional properties, (see Dorna et al. (1994)). We adopted the subclassification of adverbs provided by the Semantic Construction module of VerbMobil:

modal adverbs, focus adverbs, temporal adverbs, pronoun adverbs, intensifiers, interrogative adverbs, negative adverbs, standard adverbs, discourse relations

dimensional adjectives, relational adjectives, rigid adjectives, adjectives in comparision degrees (comparative, superlative), negative adjective

A discussion of each of them takes place in the appropriate subsection below.

We will discuss briefly in the following how the Transfer module works, and sketch the task of treating modifiers in it.

5.2 Connection of Transfer Module with other Modules in MDS

The transfer is realized at level of PredNames[3], where the TAU rule uses the information of two semantic lexicons: a German one and an English one. The German semantic lexicon provides the relation between the German lexeme and the semantic information associated with it: (a) PredNames, and (b) semantic type of the adverb (ex.(2.2a)), whereas the English semantic lexicon contains the relation between the PredNames, their arguments, and the semantic type of the adverb (ex.(2.2b)).

```
(5.2) (a) lex(LEXEME) => SUBCLASS_adv_sem(PredName).
      (b) sem_lex(PredName, #Args) => SUBCLASS_adv_sem1(PredName, #Args).
```

There are German lexemes of the word list of VerbMobil, which were given several semantic subclasses, because of their various semantic propreties. Thus, the number of the lexical entries in the semantic lexicon encreased. For example, *noch* was initially classified as a temporal adverb for uses like in: "*Im July bin ich* NOCH *im Urlaub.*", and as a focus adverb for uses like in: "*Wir machen* NOCH

[3]For detailed argumentation on the transfer strategy see Dorna et al. (1994) and section 2.

einen Termin aus.". Also, *eher* was classified as a comparative of the adverb *bald* for uses like in: "*Peter kommt eher als Suzan*", and as an intensifier for uses like in: "*Peter ist eher klug.*". Thus, some lexemes occur more than once in the German semantic lexicon of VerbMobil. And, it happens that one lexeme gets classified in semantic subclasses normally assigned to distinct word classes. For example, *echt* was classified as an adjective, as a modal adverb, and as an intensifier, *etwas* was classified as an indefinite pronoun, and as an intensifier, *offensichtlich* was classified as an adjective, and as a standard adverb (see sem_lex2.stuf).

5.3 Sketch of the Task

The task of the treatment of adverbs in the transfer module of the MDS was devided into four parts:

1. to find translation equivalents in English of each adverb from the German word list issued by the University of Bielefeld.

2. to determine their readings, and to introduce appropriate (one or more) semantic predicate names (PredNames) into the English Lexicon[4].

3. to formulate conditions for disambiguation[5] :
   a) by means of occurrence of PredNames restrictions
   b) by means of semantic sortal restrictions.

4. to produce transfer rules for each lexical entry

As the scenario of VerbMobil considered the particular domain of appointment scheduling, this set constraints on the exhaustiveness of the possible word meanings of each single adverb. Readings, relevant for the context of VerbMobil, which we determined on the basis of studies of the evidence from the test suite of dialogues, were handled with priority, and the rest of "non relevant" readings for the scenario was not really taken into consideration.

We followed a strategy of choice of PredNames after two principles:

1. PredNames as close as possible to English lexemes – keep in mind the lexical choice in the Generation module

2. as few PredNames as possible – avoid need of disambiguation by introducing equally ambiguous English items[6]

[4]In cases of ambiguous lexemes when a suitable unique candidate for a PredName (i.e. showing complete semantic equivalence between the German and the English lexeme within one PredName) were not available, more than one English PredName was introduced.

[5]The cited in a) and b) means were the two technically available mechanisms for disambiguation provided by Semantic Construction, and possible within CUF. The Transfer module obeyed the uniform output of the Semantic Construction module, which provided some technical constraints on the representation of ambiguities, and on the possible ways to disambiguate. Furthermore, the strict determination of disambiguation conditions was not always possible, because of lack of enough appropriate studies of most of the semantic classes of adverbs in VerbMobil to provide a well elaborated theoretical framework, and the constraints issued from the short time limits of MDS.

[6]We avoided introducing more English PredNames by choosing the most suitable one for all possible readings of the German lexeme, i.e. the one which showed parallel ambiguities.

For example, the lexeme *eigentlich* can be translated into English with: *actually, really, as a matter of fact, in reality, exactly, strictly speaking, originally, frankly.* We only introduced the PredName *actually*, which is a good candidate for PredName, as the rest of the lexicalizations and respectively readings can be associated with it. Thus, only one transfer rule was produced, translating *eigentlich* into *actually.*

Contrastive analyses were pursued to determine the English translation equivalents, and the readings of the German lexemes in order to produce well matching German–English PredNames. The results of some of the analyses did not actually influence the Semantic Construction module of the MDS, but provided a set of descriptive studies and outline of problems which are to be considered for the future development of the system (cf. Section 2.10.6.).

5.4 Data – Sources and Size

The contrastive analysis of adverbs for VerbMobil we pursued was based on consulting three sources of data: 1) Word list with suggested translations of CSLI, based on the German word list of Bielefeld, 2) corpus of German dialogues of the testsuite for VerbMobil translated into English, 3) corpus of German non-translated dialogues of the testsuite for VerbMobil. For theoretical consistency we also looked at bilingual English-German dictionaries, and relevant literature. Most of the results however are issued directly from the empirical studies of the dialogues.

The word list of Bielefeld contains 215 adverbs, distributed in the listed semantic classes. Not all adverbs of the list occur in the testsuite of Dialogues for the demonstrator. The occurring ones will be itemized in the appropriate subsections.

5.5 Transfer Rules for Adverbs in MDS

The transfer rules for adverbs were built on the basis of the semantic predicates (PredNames) by means of lexical transfer rules of the general form, shown in (2.3):

(5.3)
```
tau_lex(SourcePredName, TargetPredName, [tau(#Mod, #Args)], #Args,
  mod([#Mod])) => sem_t.
```

Some groups of adverbs with special semantic status were transfered by means of TAU rules of different form. These rules will be shown in the appropriate subsections, where the particular groups of adverbs will be discussed in greater detail.

As the readings of ambiguous German lexemes were transfered by means of more than one PredName, the number of transfer rules was at the end greater than the number of the German lexical entries from the semantic lexicon. The relation between the number of lexical entries in the German semantic lexicon, and the number of produced transfer rules are given in table 5.1 and table 5.2[7].

[7]**Legend:**

1. No of LE = Number of lexical entries in sem_lex2.stuf
2. No of tau_rules = Number of tau_rules

ADVERB CLASS	No of LE	No of tau_rules
Modal Adverbs	32	63
Focus Adverbs	33	41
Intensifiers	7	10
Standard Adverbs	23	35
Disc_rel Adverbs	27	62
Negativ Adverbs	1	1

Table 5.1: (a) Classes of adverbs and number of rules.

ADVERB CLASS	No of LE	No of tau_rules
Temporal Adverbs	55	4 special transfer rules and 32 tau_rules
Interrogativ Adverbs	20	1 special transfer rule
Pronoun Adverbs	18	1 special transfer rule

Table 5.2: (b) Classes of adverbs and number of rules.

5.6 General Problems for the Transfer of Adverbs

This subsection outlines informal[8] results of the contrastive analyses of adverbs we pursued, without implementing them in the version of MDS. We summarize into classes the problematic cases to handle in the Transfer so far, and discuss the phenomena causing the transfer difficulties.

5.6.1 Lexicalization

- Mismatch in the lexical meaning of the source and the target adverb – mapping of meanings

  Some lexical translations of German adverbs in English did not correspond to their lexical meanings. The lexical choice in these cases depended on the stylistic and the contextual conditions, created by the scenario of VerbMobil. For example, the relevent translation equivalent for the adverb *höchstens* in the context of VerbMobil is *only possibly* (ex.(2.4)), whereas in a standard German-English Dictionary, such as Pons (Terell et al. (1991)) this reading is translated as *not more than, at the most, except.*

  (5.4) HOMN038 :
  wir könnten's HÖCHSTENS so machen, am 7ten und am 14te
  We could ONLY POSSIBLY *do it so that it is on the seventh and the fourteenth.*

  Similar phenomenon can be observed with the adverb *noch.* One of the translation equivalents for its "additive reading" in the dialogues of VerbMobil is *also* (ex.(2.5)), whereas in a standard German-English Dictionnary (Terell et al. (1991)), this reading is translated into *one more*

  (5.5) HOMN009:

[8] "Informal" here is used as "non formalized"

dann war NOCH vorgesehen ein Abteilungsleitertreffen, ne?
Then a meeting of department heads had ALSO *been planned, right?*

- Mismatch in the lexical type of the source and the target adverb

  Some English translation equivalents of German adverbs are discontinuous morphemes[9] or composed adverbial phrases (ex.(2.6a-e)). This is due to the complex lexical meaning of the German adverbs, which is decomposable into smaller semantic units. The cited adverbs in ex.(2.6a-e) contain an anaphoric element, and a further semantic element. Thus, a semantic representation approach similar to the one pursued by the analysis of pronoun adverbs (see Section 2.10.8.) will be suitable.

  (5.6) (a) deswegen → that is why
  (b) infolgedessen → as a result
  (c) trotzdem → for all that
  (d) vorhin → a little while ago
  (e) vorerst → for the time being

  The adverbs cited in ex.(2.7a-b) are also semantically complex, but the semantic units in which they are decomposable cannot be set as clearly as by the examples in ex.(2.6a-e). So, it is not obvious how to approach for their semantic representation.

  (5.7) (a) dennoch → but still
  (b) kurzfristig → at short term

  However, the English word complexes ex.(2.7a-b)) are not well represented by one single semantic predicate (PredName), and a possibility to build complex representations of single German lexemes in order to provide a consistent transfer is requirable.

- Translation into null_lexeme

  Some German particles, classified in the group of adverbs, are regularly omitted in the English translations of the dialogues of VerbMobil, as shown in ex.(2.8a-b).

  (5.8) (a) denn → ε
  MPS1_1_01:
  Wann wär's Ihnen DENN recht?
  When would suit you?

  (b) ja → ε
  MMS4_1_05:
  Da Sie JA wissen, daß ich mittags noch einen Termin hab', bei Dani
  You know that I've got an appointment with Dani at lunchtime.

  The fact that they are not literally translated into English does not mean that these particles are meaningless. It is difficult to determine whether their

[9] Thanks to Arthur Merin for this term. It corresponds to the German *Mehrwortlexeme*

informative value is a semantic one or a stylistic one and also whether it applies to the single expressions, they occur in, or to the entire discourses these single expressions occur in. For example, *doch* in (2.9) does not appear lexicalized in the English translation. Nevertheless, it is obvious that this German particle is not rôleless in the cited expression. The same is true for the particles in the sentences of ex.(2.8a-b).

(5.9) mps1_1_01:
Ja, prima, dann lassen Sie uns DOCH noch einen Termin ausmachen.
Yes, fine, then let's just arrange an appointment.

5.6.2 Meaning

- Context dependent translations

  Lexically ambiguous German adverbs trigger context dependent translations. Thus, only an account for the course of the conversation will help disambiguating the German *also* in the following two occurences (ex.(2.10a-b)) – "*also* of the new topic coming" in 5.10(a)[10], and "*also* consecutive" in 5.10(b)[11].

(5.10) (a) *also* → *well*
HI:1:FBS1_1_02:
ALSO ich dachte, noch in der nächsten Woche –
WELL, *I was thinking of next week already* –

(b) *also* → *so*
HI:6:MMS5_1_09:
ALSO könnten wir etwa ab halb 5 uns treffen
SO *we could meet after half past four*

- Scope ambiguities

  Some adverbs show scope ambiguities, which influence the translation into English. The type of the biggest semantic entity which falls into the scope of such adverbs seems to help disambiguating them. Thus, the sentences in (2.11 a-c) will trigger distinct semantic representations each, because the adverb *überhaupt* has scope over different semantic entities in them: 1) over negation in (2.11a), 2) over a question operator in (2.11b), 3) over a whole expression (rsp. a whole DRS) in (2.11c). Furthermore, their English translations are also distinct, and seem to depend on the scope of *überhaupt*, as shown in (2.11a-c).

(5.11) (a) *überhaupt* with scope over negation → *at all*
MFD1_2_02:
schönen guten Tag Herr Schaaf. leider paßt es mir am 9ten
ÜBERHAUPT nicht weil ich vom 9ten bis zum 11ten außer Hause bin.

[10] The same readings of the German *also* was independnently discussed in VM Memo No 72 by Kerstin Fischer and al.(Schmitz and Fischer (1995)) There it was refered to as the discourse function "uptake".

[11] Thanks to Michael Schiehlen for this term.

(b) *überhaupt* with scope over question → *at all*
MMS1_3_10:
zu welchem Termin innerhalb der nächsten 2 Wochen hätten Sie denn ÜBERHAUPT Zeit?

(c) *überhaupt* with scope over an entire expression → generally
KAE008:
und wir müssen ÜBERHAUPT noch einen Besprechungstermin im Oktober ausmachen.

- Pragmatics in the translation of adverbs

  The adverb *vielleicht* in (2.12a-c) is translated into *maybe* (12a), *perhaps* (2.12b), and *possibly* (2.12c).

  (5.12) (a) MPS1_1_11:
  Gut, wollen wir's dann gleich am Montag den 3ten Mai machen VIELLEICHT um halb 4.
  Good, do we want to do it rightaway on Monday the 3rd of May then, MAYBE *at three thirty.*

  (b) MPS1_3_10:
  ich habe von 9 bis 12 Uhr auch noch eine Vorlesung. wenn wir es VIELLEICHT danach probieren wollen?
  I have another lecture from nine to twelve too. what if we PERHAPS *try it afterwards?*

  (c) MHM1_1_12:
  Können wir VIELLEICHT das auf den Vormittag legen
  Could we POSSIBLY *schedule that for the morning.*

  This fact raises the question whether it is a lexicalisation problem to be solved by the Generation module, or it is a phenomenon with semantic effects, which should be taken into consideration within the Transfer module. German native speakers' intuitions argue that *vielleicht* conveys an unambiguous interpretation, and the three English translation equivalents were just different lexicalizations of one and the same meaning. Thus, one transfer rule would be sufficient to provide the information relevant for the Generation module. Whereas English native speakers' intuitions argue that distinct "things" are conveyed by the expressions in which one of each English lexical equivalent of *vielleicht* occurs. So, one should provide information about these distinctions in the Transfer module to guide the lexical choice of the Generation module. One hypothesis for the interpretation of this phenomenon is that the mode of the dialogue expression[12] gives the relevant conditions for selecting one of the three English translation candidates.

[12] We consider that a "dialogue mode" can be figured out from the combination between the type of the dialogue act of the current expression (for example: *make_a_date_proposal*, vs. *preference_for_a_proposed_date* to distinguish between *maybe* and *perhaps*) and the information about the common knowledge conveyed by the two participants of the conversation (for example: in order to distinguish between *subjective_possibility*, and *objective_possibility* of the appointment scheduling after the circumstances accounted for to select *possibly*).

A salient transfer in such cases depends on the dialogue act of the expression in which it occurs, and the general pragmatic context of the conversation pursued. Such information was not really available in the MDS version.

- Semantic complexes formed by adverbs and other word classes

  Some German adverbs syntactically combine with other word classes, and form one semantic complex. Intuitively, the adverbs seem to get particular readings, only in particular syntactic environment. Example (2.13a) illustrates a semantic effect triggered by the combination between the adverb *so* and the preposition *um*, in which the adverb plays a role of a "preposition modifier", approximating or weakening the meaning of the preposition. Thus, the semantic complex *so um* they form, is translated into the English *around*. Another semantic effect triggered by the syntactic cooccurence of an adverb and a word of a different word class can be observed in example (2.13b) too. The adverb *noch* forms a semantic complex with the indefinite pronoun *(et)was*, in which the individual refered to by the pronoun implies the interpretation of *noch* in its "additive reading". Thus, the meaning of the entire semantic complex *noch etwas* can be expressed with *one more (additional) thing*, which justifies the translation into the English *something else*.

(5.13) (a) *so um* → *around*
FBS1_1_04:
Aber Donnerstag vormittag SO UM 9 wär mir recht.
But Thursday morning AROUND *9:00 would be fine with me.*

(b) *noch etwas* → *something else*
HAH005:
da müssen wir uns ja NOCH WAS anderes überlegen.
we'll have to think of SOMETHING ELSE *then.*

It is transparent that such complexes can be accounted for compositionally, but there are still mechanisms needed to distinguish them in the texts, and to supply correct syntactic and semantic structures.

5.7 Conclusion

It is obvious that to produce a satisfiable semantic representation for adverbs, which will account for the context of the VerbMobil scenario of spoken dialogues it is necessary to conserve a sufficiently abstract apparatus of semantic features. It should be powerful enough in expressiveness to: 1) cope with the mismatches in the translations, talked about, without loss of information, and 2) transfer the semantic and stylistic content and leave open the final decision for lexicalization to the Generation module.

5.8 Pronoun Adverbs

In this section we will understand by *pronoun adverbs* all those adverbs that relate the event they modify with an entity that is not given by the adverb meaning but is still to be determined anaphorically.

The definition is rather broad: In addition to central cases like *damit, hierauf* it includes a range of adverbs with specifically local and temporal interpretations like *daher, da*. For full clarity we give a list of the lexical items we treat under the heading of pronoun adverbs.

> *da, dabei, dadurch, dagegen, daher, dahin, damit, danach, daran, darauf, darin, darüber, darum, davon, davor, dazu, demnach, dort, dorthin, drauf, ebenso, genauso, hier, hierfür, hierhin, hiermit, hierzu, irgendwann, irgendwie, rüber, so*

Another feature of pronoun adverbs is that their meaning can be broken down into a *relational* component (RC) and a *pronominal* component (PC). Arguments supporting this position can be found in semantics, (German) morphology and German-English transfer.

- The RC specifies which relation the pronoun adverb designates, whereas the PC supplies information about the entity that is to be determined from context.
- The dichotomy is also discernible on the morphological level. We give an example.

  damit (therewith), PC: da-, RC: -mit
  daher (thence), PC: da-, RC: -her

- Most importantly, contemporary English overtly splits RC and PC.

  damit → with that (RC+PC)
  daher → from there (RC+PC)

For a range of phenomena it is decisive whether the relation expressed by the pronoun adverb is based on the meaning of a preposition or not. Henceforth, we will call pronoun adverbs with a discernible underlying preposition *Prepositional Pronoun Adverbs* (PPAs) and such without one *Nonprepositional Pronoun Adverbs* (NPAs). The distinctive phenomena are listed below.

- Only PPAs can be used in subcategorisation frames[13].

  Ich warte darauf.

- Translation of PPAs necessitates transfer of the preposition involved. There is no such need for NPAs whose RC is abstract, hence interlingual[14].
- While the PC of PPAs refers to *objects*, the PC of NPAs designates *points*.

[13]In the expression *den Termin dahin legen (z.B. auf Sonntag)* we find a subcategorised NPA, however.

[14]The underlying relation for NPAs is not always mapped to the empty string in English. A relevant exception is *dorther → from there*.

First of all we undertook a contrastive analysis on the basis of the Verbmobil dialogues and several German-English dictionaries. For this investigation we took into account all Verbmobil dialogues, the translated ones as well as those that were available only in German. We then proceeded to devise a strategy for an efficient though general translation of these items and to construct detailed entries for the semantic lexicon. As a result, the actual transfer rules could be held fairly simple.

5.9 Contrastive Analysis of PPAs

The PC is made up of the prefixes *da(r)*, *hier* and *wo(r)*[15]. The RC consists of a preposition. If the preposition begins with a vowel, there exists a colloquial variant of the *da(r)*-prefix (e.g. *drauf*).

We give some statistics of the data as extracted from the corpus. The underneath table reads as follows. The pronoun adverbs are classified according to their PC (columns) and RC (lines). The numbers give the count of the respective item in the corpus. Brackets indicate that the item did not show up in the Verbmobil word list and therefore was only treated in a cursory manner.

	da(r)-	dr-	hier-	wo(r)-
an	11	10	(0)	0
auf	51	49	(0)	0
bei	34		(0)	(24)
durch	1		(0)	0
für	(39)		0	(0)
gegen	1		(0)	0
in	1	22	(0)	0
mit	48		0	1
nach	85		(0)	(0)
über	13	7	(0)	1
um	11	2	(0)	(3)
von	54		(0)	1
vor	37		(0)	(0)
zu	36		0	0
sum	422	90	0	30

A final remark on the table is in order. Often pronoun adverbs are split in German (e.g. *da kann ich nichts mit anfangen*). Such split pronoun adverbs cannot easily be extracted from the corpus since satisfying results can only be achieved by successful disambiguation of word classes, hence successful parsing. Consider the ambiguous sentence

Da stößt er gegen eine Stange.

Not all data we had at our disposal were tagged with the information needed, so we decided to disregard these cases in our analysis.

[15]The adverbs with prefix *wo* do not really fall into the category of pronoun adverbs since the missing entity is not contextually determined but must be supplied by an answer. Nevertheless they can be and were analysed along with the pronoun adverbs proper.

For an overwhelming majority of the examples an acceptable translation can be produced by translating PC and RC separately. We first go into the translation of the PC and then turn to the translation of the RC.

Translation of the Pronominal Component

The following table gives frequent translations of the PC.

da(r) → that, it/them		
hier → this, it		
wo(r) → what		(as interrogative pronoun)
wo(r) → which		(as relative pronoun)

This is, however, not always the right way to go. If the PC refers to a constituent of the same sentence the English translation sounds extremely awkward.

Ich bin darum hier, weil ich mit Ihnen sprechen will.
→ *I am here because of that because I want to speak with you*

Therefore resolution of sentence-internal anaphora is crucial for an adequate translation of pronoun adverbs[16]. We present the constructions involving a sentence-internal antecedent.

- In 3 cases the pronoun referred to a PP with the same preposition in the Nachfeld.

  CD:L009n:SIH067 ich sollte doch jeden Monat schon etwas daran arbeiten, an dem Projekt,

- In 29 cases the pronoun referred to another adverb *da* (or *wo* in relative and interrogative clauses) preceding the pronoun adverb. This construction is very common in spoken language and seems to be linked with the aforementioned splitting construction.

  CD:K007d:FNK042 und die Woche da drauf?

- In 44 cases the pronoun referred to an embedded finite (26) or infinite (7) clause or a conditional clause (12). In the correct translation the pronoun adverb was sometimes simply cancelled:

  KA:30:mdj1_2_10 ich bin überzeugt davon daß Sie es schaffen werden. → *I am convinced you'll manage that.*

  Sometimes the clause was transformed into a gerund which was then given as an argument to the preposition:

  KA:10:mps1_1_07 Was halten Sie davon, wenn wir 's im Januar probieren? → *What do you think about trying January?*

  But more often than not a construction-specific translation had to be chosen:

[16]The lack of this information in the demonstrator only did not make itself felt because the relevant constructions by accident did not occur in the selected dialogues.

> CD:M016n:NCW004 wir sollten darauf achten, daß wir dann nich' irgendwie Freitag anpeilen → *we should make sure that in any case we do not head for Friday*
> SB:54:KS000 ich bin dran, meine Seminararbeit zu schreiben, → *I am writing my seminar paper.*

- In 8 cases the pronoun referred to an embedded interrogative clause. In English wh-clauses can serve as an argument to PPs, so this construction does not make trouble.

  > KA:5:moc1_1_20 das kommt darauf an, wie lange da der Dienstag bei mir durch eine Klausur belegt ist → *that depends on how long the exam on Tuesday will take.*

Translation of the Relational Component

A general objective was to keep the transfer component as simple as possible. Therefore, we tried to handle the translation of the RC of a pronoun adverbs in tandem with the transfer of the underlying preposition whenever possible and thereby avoid a tedious redoubling of information. Prepositions chiefly occur in three environments.

- Sometimes, prepositions contribute their meaning to the overall sentence. In this case there has to be a transfer rule for the preposition. In accordance with our general strategy we used the transfer rule of the preposition also in the transfer rule of pronoun adverbs.
- In other cases prepositions are subcategorised by a lexical item (verb, noun, or adjective). Subcategorisation requirements are checked by syntax. Thus, semantic construction normally can easily distinguish if preposition contribute meaning or not.
- Finally prepositions are used in idioms. In the MDS idioms were usually recognized only by transfer, although an in-depth semantic analysis would also have to recognize idiomatic constructions.

It was, however, not possible in every case to find a way to reduce pronoun adverbs to prepositions.

- The RC of pronoun adverbs can sometimes have meanings that are clearly out for normal prepositional constructions.
- There are idioms that require the presence of particular pronoun adverbs.

In the sequel, we will go through these cases one by one and list the examples that we found in the corpus.

Preposition Carries Meaning

The task here was to identify exactly which meanings of the preposition give rise to a derived pronoun adverb meaning. Not all of them do so, cf. the anomalous discourse

Morgen ist der 16te Februar. * Daran (an diesem Tag) hat Peter Geburtstag. → *Tomorrow is the 16th of February. On it, Peter has birthday.*

This task was greatly facilitated by the preparatory work done on prepositions at IBM Heidelberg and IAI Saarbrücken. In fact we used their classification of preposition meanings as a basis for the choice of admissible prepositional pronoun adverb meanings.

The table below reads as follows:

- The first column gives the German preposition treated. A plus sign indicates that the particular meaning (as given in column 2) was not considered by IBM and IAI, but nevertheless turned out to be relevant in connection with pronoun adverbs.
- The second column lists the names of the meanings. Prepositional meanings are encoded as roles in the MDS formalism.
- The signs in the third column reflect our judgement whether or not a derived pronoun adverb meaning is associated with the respective preposition meaning.
- The fourth column gives the English translations of the meaning.
- The fifth column shows a German example for the meaning, mostly drawn from the studies carried out at IBM and IAI.
- Finally we give the corpus examples.

in	unspec_temporal_inclusion_rel	+	at;in;on;0	in dieser Woche
	spatial_inclusion_rel	+	in	im Gebäude
		–	on (floor, town)	im zweiten Stock, in Berlin
	temporal_spatial_inclusion_rel	–	at	in einer Besprechung sein
	institutional_inclusion_rel	–	at	Treffen in der Universität
	abstract_inclusion_rel	+	in	im Notizblock stehen
	functio_spatial_inclusion_rel	–	at	Treffen im Büro
	spatial_inclusion_goal_rel	–	in;into;to	ins Gebäude
	temporal_spatial_goal_rel	–	to	in die Vorlesung gehen
	temporal_inclusion_goal_rel	–	to	in den März legen
	institutional_goal_rel	–	to	in die Universität kommen
	abstract_goal_rel	–	into	in den Notizblock schreiben
	modality_mood_rel	–	in	in Not sein

- (unspec_temporal_inclusion_rel)
  Tag in Woche
  CD:M003N:JOH001G *in der 46sten Woche, wo auch Buß- und Bettag drin ist*
- (abstract_inclusion_rel)
  Termine im Kalender
  CD:M144D:HIM011 *ich hab' da ein paar andere Termine drin.*

Subcategorised pronoun adverbs

Below we list occurrences of pronoun adverbs in the corpus that we classified as instances of subcategorised argument. Not all of the verbs involved figure in the Verbmobil word list.

- abgesehen von → apart from
  CD:M023N:SID018 *abgesehen davon, daß ich am 27sten bis zum 30sten eine Exkursion habe in Frankfurt.*
- achten auf → pay attention to
  CD:M016N:NCW004 *wir sollten darauf achten.*
- alternative zu → alternative to
  KI:G201A:PRB042 *im Moment sehe ich keine weitere Alternative dazu.*
- anfangen mit → start
  CD:M005N:DOS004E *komm' ich dann wahrscheinlich besser mit rein, wenn ich dann damit anfange.*
- anfangen mit → start with
  CD:L055D:CAO001 *fangen wir damit an.*
- ankommen auf → depend on
  KA:5:MOC1_1_20 *das kommt darauf an, wie lange da der Dienstag bei mir durch eine Klausur belegt ist.*
- sich anschließen an → be after
  KI:G076A:HAH010 *und direkt daran schließt sich 'n Treffen in Hamburg an.*
- arbeiten an → be working on
  CD:L009N:SIH067 *ich sollte doch jeden Monat schon etwas daran arbeiten.*

Idiomatic Use of Prepositions

- etwas anfangen können mit
  BN:4:HAEN050 *können wir damit irgendwas anfangen?*
- Der Teufel ist in X.
  CD:M114D:REA027 *da ist wahrscheinlich irgendwo der Teufel drin.*
- keinen Sinn finden in
  CD:M005N:DOS004B *sonst sind die Vorträge so auseinandergerissen, daß ich dann also auch keinen Sinn mehr drin (das zu machen) finde.*

Special Meanings of Pronoun Adverbs

- darauf → after that
  occurs very often in the dialogues (64 times)
- darum → that's why
  CD:K015D:WOG046 *darum einen ganzen Tag zu belegen.*

Idiomatic Use of Pronoun Adverbs

- dabei haben → have got on oneself
  KA:24:MTS2_1_01 *haben Sie Ihren Termin-Planer dabei?*

- dabei sein (bei) → take part (in)
  KI:G094A:ANS000 *und ich habe auch gehört, daß Sie dabei sind.*

- X ist dabei → you can count X in
  KI:G084A:KAK006 *da bin ich dabei.*

- mit X dabei → with X
  KI:G113A:JAK015 *mit ein bißchen Kaffeetrinken dabei*

- X ist dran → it is X's turn
  KI:G112A:JAK003 *also ich notier' das schon mal, daß Sie dann dran sind.*

- Da ist etwas dran. → There's something in it.
  KI:G315A:SVA011 *da ist was dran.*

- gut dran sein → have got it good
  CD:K017D:RED062 *und wenn wir es in einer Stunde fertig haben, dann sind wir gut dran.*

- drin sein → be possible
  CD:M024N:MPI013 *dann die folgende Woche, ist da der Mittwoch vielleicht drin?*

5.10 Semantics

In the foregoing section we saw that transfer from German to English demands a separation of the RC from the PC of the pronoun adverb. We decided to carry out this separation already within semantic construction in order to reduce the complexity of the transfer task. We tried to reuse existing semantic macros, that had been written with independent motivation, as far as possible. In particular, we used the semantics of prepositions to model the meaning of the RC of pronoun adverbs and likewise the semantics of pronouns for their PC. The semantic macros for these word classes are quoted below.

- **prep_sem2**:
  λ NP λ VP λ E. VP(λ e. NP(λ x. $\underline{\text{Prep}}$(e, $\underline{\text{Rel}}$: x) $\otimes$ E(e))).
  Prep: name of the preposition
  Rel: role of the preposition as coded in the sort hierarchy

- **pronoun_sem2**:
  λ VP. α: $\boxed{\begin{array}{c}[x]\\ \underline{\text{Pred}}(x)\end{array}}$ $\otimes$ VP(x).
  Pred: further restriction (see below)

Via the composition operation ($\odot$) the following semantics can be obtained.

- **pronoun_prep_adv_sem2**:

  λ VP λ E. VP(λ e. α: $\boxed{\begin{array}{c}[\mathrm{x}]\\ \underline{\mathrm{Pred}}(\mathrm{x})\end{array}}$ $\otimes$ $\underline{\mathrm{Prep}}$(e, $\underline{\mathrm{Rel}}$: x) $\otimes$ E(e)).

We associate this semantics with pronoun adverbs.

```
pronoun_prep_adv_sem1(#Pred, #Prep, #AlfaSort, #AlfaType, #Rel,[#VPSem]) =>
   compose(compose(prep_sem2(#Prep,#Rel),
                   pronoun_sem2(alfa_cond(#Inst&sort:#AlfaSort,#AlfaType,
                                          [basic_cond(#Pred,#Inst,[])]))),
          #VPSem).
```

A second lexical entry for prepositions specifies their contribution in case of subcategorisation, which is none. As we saw above, only PPAs can be subcategorised. We therefore need a second entry for them, too, to account for this usage: Subcategorised PPAs correspond to pronouns.

```
pronoun_prep_adv_sem(#Pred, ##, #AlfaSort, #AlfaType, ##) =>
    category(prep) &
    cat_shift(pp,pp) &
            subcat([])&
            mod([])&
            pred_args([])&
            lex_pred(padvcomp)&
            sem:pronoun_sem2(alfa_cond(#Inst&sort:#AlfaSort,#AlfaType,
                                        [basic_cond(#Pred,#Inst,[])])).
```

Basically there are now three parameters that can be fixed in the lexicon: the name of the underlying relation (Prep), the role associated with this relation (Rel) and a restriction on the pronoun (Pred).

- Pred: One of the basic differences between PPAs and NPAs is that the PCs of PPAs are objects and those of NPAs points. We encode this sort difference in the pronoun restriction. Another use of the pronoun restriction is to distinguish between **here** and **there**. **here** specifies that some point is *near* to the speaker, **there** is understood to refer to a point that is *far* away from the speaker. We get the following classification:
  - **loc_far_pred(x)**: x is an object that is located at a point p far away from the speaker (pronoun restriction for *that*). This predicate is used for PPAs beginning with *da(r)*.
  - **loc_near_pred(x)**: x is an object that is located a point p near to the speaker (pronoun restriction for *this*). This predicate is used for PPAs beginning with *hier*.
  - **far_pred(p)**: p is a point or a location that is far away from the speaker (pronoun restriction for *there*). This predicate is used for NPAs belonging to the *there*-family.
  - **near_pred(p)**: p is a point or a location that is near to the speaker (pronoun restriction for *here*). This predicate is used for NPAs belonging to the *here*-family.

An example may help to illustrate the point

hierfür: für(e,x) α: [x] loc(x,p) α: [p] near(p,i) <i,speaker>

dorthin: dir(e,p) α: [p] far(p,i) <i,speaker>

- Prep: For PPAs it is relatively clear what should go into the relation slot: The name of the underlying preposition. But what about NPAs? NPAs express that the event

  - is located at the point that is determined anaphorically (*dort*, *ebenso*),
  - is directed towards it (*dorthin*) or
  - comes from it (*dorther*).

  We encode the first two aspects with the predicates **loc_pred** and **dir_pred**. The matter is not so straightforward with the abstract relation for the source role, since this relation must be translated for English (*dorther* → *from there*). We decided to use the German preposition **von** here.

- Rel: Although roles are not relevant for semantic construction, they are specified in the semantic lexicon for later use by the transfer component. We will therefore go into the technical values of the role description only in the section on transfer.

  The contrastive analysis shows that not all readings of prepositions lead to corresponding readings of pronoun adverbs.

  Ich komme um neun. Kommst Du auch darum? (≠ at this time)

  Therefore, the readings available for pronoun adverbs have to be specified in the (semantic) lexicon.

```
lex(darauf)     => pronoun_prep_adv_sem(loc_far_pred,auf,~person_c,demonstr,
        institutional_inclusion_rel;
        top_region_inclusion_rel;
        % top_region_inclusion nicht mit insel_c
        top_region_goal_rel;
        projection_rel;
        projection_goal_rel;
        temporal_inclusion_goal_rel).
```

On the other hand, some pronoun adverbs allow for additional readings.

> An dem Sonntag geht es nicht. Wie wäre es am Sonntag darauf?
> (* auf dem eben genannten Sonntag)

These meanings are encoded as further entries in the semantic lexicon.

```
lex(darauf)      => pronoun_prep_adv_sem(loc_far_pred,nach1,~person_c,demonstr,
        temporal_span_posterior_rel).
```

5.11 Transfer

In the transfer component nothing has to be done but the translation of preposition for PPAs (call to **tau_prep** and **tau_pred**). Both the pronominal component and the abstract relation underlying NPAs are interlingual.

```
% PPAs
tau_lex(pronoun_prep_adv,pronoun_prep_adv,
         [#Pred,#PrepT,#AlfaSort,#AlfaType,#Rel,tau(#VP,#Args)],#Args,
         adv_syn(#VP&sem:#VPSem&pred_sort(#ExtSort)) &
         sem:pronoun_prep_adv_sem1(#Pred,#PrepS,#AlfaSort,#AlfaType,#Rel,[#VPSem]) &
         tau_prep(#PrepS,#Rel,#ExtSort,#AlfaSort)) =>
    tau_pred(#Rel,#PrepT,#AlfaSort).

% NPAs
tau_lex(pronoun_part_adv, pronoun_prep_adv,
        [#Pred,#Prep,#AlfaSort,#AlfaType,#Rel,tau(#VP,#Args)],#Args,
        adv_syn(#VP&sem:#VPSem) &
        sem:pronoun_prep_adv_sem1(#Pred,#Prep,#AlfaSort,#AlfaType,#Rel,[#VPSem])) => sem_t.
```

Preposition transfer operates on the roles given in the semantic lexicon. German prepositions are first mapped into roles (this is done in the German semantic lexicon) and then English prepositions are generated from them (**tau_pred**). Numerous factors are taken into account, in particular the sorts belonging to the internal (noun) and external (event) argument of the preposition (**tau_prep**).

5.12 Conclusion

We showed a strategy to translate pronoun adverbs on the basis of their semantics. In particular, we strove to keep programming code as simple as possible.

- As far as possible, steps necessary for transfer were already done in semantic construction.
- Interlingual representations were abstract enough for transfer purposes, but we deliberately refrained from making them as detailed as they would have to be for semantic analysis purposes. However, we gave a clear description of what the interlingual predicates were meant to stand for in order to facilitate subsequent elaboration.

Due to time pressure several problems could not be solved. For a full treatment of pronoun anaphora sentence-internal anaphora must be resolved. This task has to be done either in syntax (also for syntax-internal reasons, e.g. in order to explain complement clauses) or in semantics.

> Ich gehe davon aus, daß er kommt.

A neater definition of the sort of the pronominal entity has to be given. We did not check this up against the sorts used in disambiguation of prepositions. We merely stipulated following our intuition that the anaphoric entity cannot be a person.

Another serious problem that often demands context-driven inference is how to disambiguate between literal and idiomatic readings. Consider the example

> CD:M024N:MPI013 *dann die folgende Woche, ist da der Mittwoch vielleicht drin?*

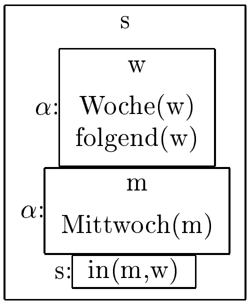
The sentence might occur in (at least) two different contexts. In the first context we prefer the usual translation of the pronoun adverb *drin* (*(included) in it*).

> H: Geht es bei Ihnen am Montag, den 2. Februar?
> S: Ganz unmöglich.
> H: Aha, Montag geht also nicht. Dienstag, Freitag sind bei mir schlecht. Aber Mittwoch würde mir eigentlich ganz gut passen. Mittwoch ist sowieso immer so ein schöner Tag, mitten in der Woche. Würde es bei Ihnen an dem Mittwoch auch gehen?
> S: Entschuldigung, ich habe jetzt nicht ganz aufgepaßt. Der Mittwoch, den Sie da vorschlagen, ist das der Mittwoch in der ersten Februarwoche? Nein? Dann die folgende Woche, ist da der Mittwoch vielleicht drin?

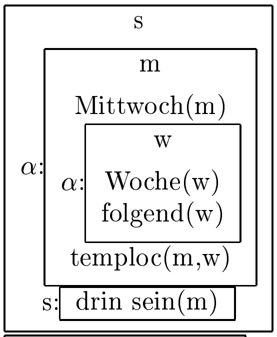
In the second context the adverb must be translated idiomatically (*possible*).

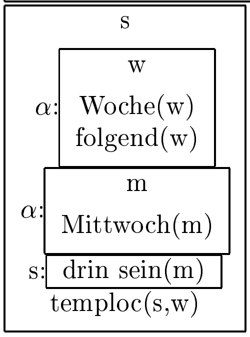
> S: Geht es bei Ihnen am Mittwoch oder am Donnerstag in der ersten Februarwoche?
> H: Leider nein, ich bin die ganze Woche im Urlaub.
> S: Dann die folgende Woche, ist da der Mittwoch vielleicht drin?

We will now retrace the line of argument leading to disambiguation. First, we give a DRS for the literal reading. We assume that anaphora resolution has already taken place correctly.



In the second reading a further ambiguity is lurking: The adverb *da* can be adjoined either to the subject *Mittwoch* or to the verb *ist*[17]. We get the following DRSs for the two options.





We distinguish between four types of context.

- Suppose no Wednesday has been mentioned before (cf. the authentic CD-ROM dialogue). So the α-DRS containing *Mittwoch* cannot be fixed to an antecedent and must be accommodated. Accommodation, however, normally presupposes that the principal DR of the accommodated α-DRS refers to a unique object[18]. Whereas there is a multitude of Wednesdays in general, each week has only one Wednesday, hence Wednesday of the following week is a uniquely determined time interval. The uniqueness requirement is met by the second DRS, but not by the two other ones. We choose the second DRS.

- Another possibility is that some Wednesday m′ figures in the context and we can deduce that either it is in the following week w or it is not (cf. context 2). Then arguably also speaker S knows that m′ is or is not in w, respectively. A general principle states that people do not ask things they know (at least if they do not have ulterior motives, something which can be taken for granted in a goal-oriented dialogue).

  > If S asks at time t whether p then before and at time t S does not know whether p.

[17] Adjunction to the subject is clearly preferred, but not derivable with the demonstrator version of the TrUG system, which only allows adjunction to verbs.

[18] We also can accommodate non-unique definites, but in this case we feel that we do not fully understand the text. A case in question is H's usage of *Mittwoch* in context 1 which prompts S to ask her which day she actually has in mind. By the same principle it is harder to accommodate *the king* than to stipulate the existence of *the king of France*, who is a unique person.

Hence the first DRS is out.

- Suppose a Wednesday m' has been mentioned but we cannot be sure whether or not it is the one of the following week. Further suppose m' turned up in a suggestion made by the discourse partner H (as in context 1). From the fact that H suggested m' speaker S can deduce that m' is a possible date for H.

  > If H suggests m' to S at time t (and the suggestion is felicitous) S knows from t on that m' is possible for H.

  Asking H whether m' is a possible date for her is a senseless thing for someone to do who already knows it is. Thus we have reasons against the second and third DRS and choose the first one.

For correct disambiguation we need several modules:

- a component for anaphora resolution,
- a mechanism for temporal deduction,
- a speech act recognition component and
- a speech act logic.

5.13 Modal Adverbs

This section describes the treatment of modal adverbs in the Transfer module of MDS. When we talk about *modalities*, as this term occurs in the name of the class of *modal adverbs*, we have to distinguish between: 1) the notion of *modality* in the logic, where the possibility, probability or necessity in linguistic expressions are interpreted with respect the truth conditions of these expressions, and 2) the notion of *modality* in the language, where it is defined as a mixed morpho–syntactic and semantico–pragmatic caterogy refering to the relation between the speaker and his predication on one hand and the relation between the predication and the reality on another hand[19]. *Modal adverbs* in VerbMobil are designed after the notion of *modality* in the language.

5.13.1 Modal Adverbs in the Semantic Construction Module of MDS

Modal adverbs are handled uniformly as sentence adverbs by the Semantic Construction module of MDS. They appear at the end of the structure of the analyzed sentences, introduce a Dimension Condition, and refer to the event described by the sentence. The concept of modality is defined as VALUE of the feature DIMENSION (ex.(2.14)).

(5.14) Er kommt EIGENTLICH.

$$\begin{bmatrix} DRS \\ \begin{bmatrix} kommen : \boxed{1} \\ \begin{bmatrix} dimen\_condition \\ dimen\_inst : \boxed{1} \\ dimen\_pred : eigentlich \\ dimension : modality \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

5.13.2 Size of the Data

Modal adverbs in the wordlist issued by the University of Bielefeld (see sem_lex2.stuf) count 32:

echt, natürlich, ganz, notfalls, komplett, prinzipiell, glattweg, schätzungsweise, rundweg, schlimmstenfalls, völlig, sicher, vollends, sicherlich, ziemlich, sowieso, allerdings, strenggenommen, bestimmt, unwahrscheinlich, eigentlich, vielleicht, eventuell, wahrscheinlich, fast, womöglich, gegebenenfalls, zirka, herum, zur_not, möglicherweise, zweifellos

Eight of them occur in the testsuite of dialogues for the MDS:

[19] We adopted the definitions of *modality* given by Lexandowski in Linguistisches Wörterbuch Lexandowski (1994).

echt, eigentlich, sowieso, allerdings, vielleicht, komplett, fast, natürlich

They were investigated in greater detail, than the rest of the wordlist.

5.13.3 Determination of Readings and PredNames

We use the adverb *allerdings* to illustrate how the determination of an English PredName took place. *Allerdings* is ambiguous between two readings: 1) contrastive; 2) confirmative. Their English translation equivalents stay for them in (2.15a-b).

(5.15) (a) allerdings → but
(b) allerdings → indeed

The reading in (2.15b) does not occur in the testsuite of dialogues for the Demonstrator, but in the entire corpus of dialogues for VM. The further examples show that other lexicalizations for the readings of *allerdings* in English are possible.

(5.16) *allerdings* into *but*

MPS1_1_03:
Ja,am Dienstag, den 6ten April hätte ich noch einen Termin frei
ALLERDINGS nur nachmittags.
Yes, I would have time on Tuesday April 6th – BUT *only in the afternoon.*

(5.17) *allerdings* into *however*

BN:1:FLHN037:

am 28sten ALLERDINGS nicht

On the twenty-eighth HOWEVER *not*

(5.18) *allerdings* into *unfortunately*

HI:2:FSP2_1_06:
Da ging es bei mir ALLERDINGS nur am Freitag
UNFORTUNATELY *I am only free on Friday that week. It would only be possible on Friday, however*

Example (2.16) lexicalizes *allerdings* into *but*, example (2.17) lexicalizes *allerdings* into *however*, example (2.18) lexicalizes *allerdings* into *unfortunately.* All these lexicalizations were summarized into PredName *but* (ex.(2.19)), as they all convey a relation of contrast with the previous sentence.

(5.19) allerdings → but2

Example (2.20) illustrates the second reading of *allerdings.* The semantic effect of this adverb is transfered not literally, but through a phrasal change. Its emphasizing effect is lexicalized in English through the focal stress of the redundant *do* in the expression *I do already have.*

(5.20) *allerdings* into *phrasal change and focal stress*
BN:1:HOMN038:
21sten hab ich ALLERDINGS auch einen sehr langen Termin
on the twenty-first I DO *also already have a very long appointment.*

A plausible hypothesis for the process of disambiguation between these two readings of *allerdings* is that prosodic information will be helpful, as an accent on this adverb triggers almost always its *confirmative* reading.

5.13.4 Transfer PredNames of Modal Adverbs in MDS

The strategy of Transfer module was to provide Target PredNames, which are as near as possible to the corresponding English lexemes. We list here informally the German PredNames with their English equivalents (see table 5.3 and tau_adv.stuf).

The shape of the transfer rule for modal adverbs was the standard one, quoted in Section 5 ex. 5.3, repreated here as ex.(2.21).

(5.21)
```
tau_lex(SourcePredName, TargetPredName, [tau(#Mod, #Args)], #Args,
    mod([#Mod])) => sem_t.
```

5.13.5 Classes of Modal Adverbs

This subsection outlines informal results of the semantic analysis of modal adverbs we pursued, without implementing them in the version of MDS. We provided a finer classification of modal adverbs which seems to be consistent with the semantic information conveyed by them on one hand, and fits the strategy to represent semantic information in VerbMobil on another hand.

The classes of modal adverbs were identified after two general criteria:

- structural properties – account for the semantic entities they can apply to
- nature of the modification – account for the semantic effect of the modifier applied

For example, the modal adverbs accounted for show types of modalities with the following structural properties:

- sentence internal modalities. These modalities belong to one sentence. They are to be classified as event modifiers[20].
- sentence external modalities. These modalities belong to different sentences, i.e. they denote relations between two DRS-es.

Two groups of modal adverbs were distinguished.

- on the basis of their structural properties
  - the first group can refer to smaller semantic entities – events, modifiers etc.

[20]Preliminary discussion on types of modifiers can be found in (Eberle (1995))

SourcePred	→	TargetPred
echt1	→	null_pred
echt1	→	really1
ganz2	→	quite
komplett	→	totally
glattweg	→	absolutely1
rundweg	→	flatly
voellig	→	entirely
vollends	→	wholly
ziemlich	→	quite
allerdings	→	but2
bestimmt	→	certainly
eigentlich	→	actually
eventuell	→	maybe
fast	→	almost1
gegebenenfalls	→	if_necessary
herum	→	around1
moeglicherweise	→	if_possible
natuerlich	→	of_course
notfalls	→	if_necessary
prinzipiell	→	basically
schaetzungsweise	→	roughly
schlimmstenfalls	→	at_worst
sicher	→	certainly
sicherlich	→	surely
sowieso	→	anyway
strenggenommen	→	strictly_speaking1
unwahrscheinlich	→	probably_not
vielleicht	→	maybe
vielleicht	→	perhaps
vielleicht	→	possibly
wahrscheinlich	→	probably
womoeglich	→	possibly
zirka	→	approximately
zur_not	→	if_necessary
zweifellos	→	undoubtedly

Table 5.3: Transfer PredNames of the modal adverbs in the MDS.

  - the second group can refer to a whole DRS
- on the basis of the type of their semantic contribution
  - the first group is neutral to the information in the previous discourse (context independent)
  - the second group introduces semantic nuances which can only be accounted for in connection with the previous discourse

We illustrate the first group with the example of the modal adverb *komplett*, and the second group on example of the modal adverb *allerdings*.

5.13.6 Komplett

Komplett is representative of the group of modal adverbs which introduce a modality relevant for the event described by a single sentence. Thus, they are expected to appear in the DRS as event modifiers (ex.(2.22)).

(5.22) Montag können wir KOMPLETT vergessen.

$$\begin{bmatrix} DRS \\ \begin{bmatrix} \text{können: } \boxed{1} \\ \begin{bmatrix} DRS \\ \begin{bmatrix} vergessen: \boxed{2} \\ \begin{bmatrix} dimen\_condition \\ dimen\_inst: \boxed{2} \\ dimen\_pred: \ komplett \\ dimension: \ modality \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Modal adverbs of this group do not relate to eventualities described in the preceding discourse, (or the events described in the previous discourse). Compare example (2.23a) and (2.23b). The presence of the adverb *komplett* in sentence RFD:DE017, example (2.23a), and the absence of the adverb *komplett* in sentence RFD:DE017, example (2.23b), shows that this adverb does not seem to be related to the interpretation of the discourse relations in the whole discourse, as it does not semantically contribute to the coherence of the discourse, but just to the interpretation of the event described by sentence RFD:DE017.

(5.23) (a)

RFD:EL016:*I can't make it at the beginning; I'm on vacation then.*
RFD:DE017: ach, da können wir den Oktober ja KOMPLETT VERGESSEN, aber nicht den November.
RFD:VM018: *oh, so we can* TOTALLY FORGET *about October, but not November.*

$$\boxed{\begin{array}{c} \text{x y e} \\ \text{x = wir} \\ \text{y = Oktober} \\ \text{e: vergessen(x,y)} \\ \text{komplett(e)} \end{array}}$$

(b)
RFD:EL016: *I can't make it at the beginning; I'm on vacation then.*
RFD:DE017: ach, da können wir den Oktober ja VERGESSEN, aber nicht den November.
RFD:VM018: *oh, so we can* FORGET *about October, but not November.*

$$\boxed{\begin{array}{c} \text{x y e} \\ \text{x = wir} \\ \text{y = Oktober} \\ \text{e: vergessen(x,y)} \end{array}}$$

Another property of this class of modal adverbs is that they do not only apply to events, but also to other types of modifiers (Mod), and to DRS–operators (DRSop), as it is shown in ex. (2.24) and (2.25a-c).

(5.24) MFD1_4_04:
das paßt wunderbar. der 16te ist bei mir KOMPLETT FREI(Mod). ich würde sagen wir treffen uns dann um 9 Uhr. Ort entscheiden Sie bitte.

(5.25) (a) Er hat seinen Termin FAST VERGESSEN (E).
(b) Er ist FAST FERTIG (Mod).
(c) Er ist FAST IMMER (DRSop) zu spät.

Other modal adverbs belonging to the group of *komplett* are: *fast, echt, prinzipiell.*

5.13.7 Allerdings

First, the modal adverb *allerdings* does not always apply to the event described by the sentence, as it normally modifies its finite VP. For example, modal verbs in general do not describe an event, but introduce a modal condition, and as the sentences in (2.26), and (2.27) show, *allerdings* applies to the modal verb *können* and not to the described event by the verb *machen.*

(5.26) ALLERDINGS können wir das Treffen montags machen.

$$\begin{bmatrix} \mathit{DRS} \\ \begin{bmatrix} \text{können: } \boxed{1} \\ \begin{bmatrix} \mathit{dimen\_condition} \\ \mathit{dimen\_inst} : \boxed{1} \\ \mathit{dimen\_pred} : \ \mathit{allerdings} \\ \mathit{dimension} : \ \mathit{modality} \end{bmatrix} \begin{bmatrix} \mathit{DRS} \\ \begin{bmatrix} \mathit{machen} : \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

(5.27) ALLERDINGS wollte er montags kommen.

Second, *allerdings* introduces a semantic nuance which semantically relates to the information conveyed by the previous sentence.

(5.28) MPS1_1_19:

Ist bei mir etwa schlecht. ALLERDINGS könnte ich um 11 noch einen Termin reinschieben.

We interpret the semantic role of *allerdings* in example (2.28) in the following way. It is not only a sentence adverb, which refers to a Dimension *Modality*. Its semantic interpretation is connected with the information of the previous discourse. A rhetoric relation of contrast with the previous sentence, similar to the relation introduced by the conjunction *aber*, is conveyed by the sentence containing *allerdings*.

Other modal adverbs belonging to the group of *allerdings* are: *vielleicht, eigentlich*. They also relate to two DRS-s, but they convey other types of rhetoric relations between them.

The reported semantic classes of modal adverbs were not implemented because of the uniform output of the Semantic Construction.

5.13.8 Future Work

It is necessary to determine and represent by means of distinguishing marks in the lexicon on example of the analysis of *komplett* and *allerdings*: 1) what kind of modalities can be expressed by the modal adverbs; 2) what semantic entities can be modified or bound by means of modal adverbs; 3) what are the semantic effects of different types of *Modalities*, which can be introduced by modal adverbs.

5.14 Temporal Adverbs

5.14.1 The Semantics of Temporal Adverbs

The Verbmobil analyses of the temporal adverbs of the demonstrator corpus use the classification of temporally localizing adverbials described in Eberle and Kasper (1994):

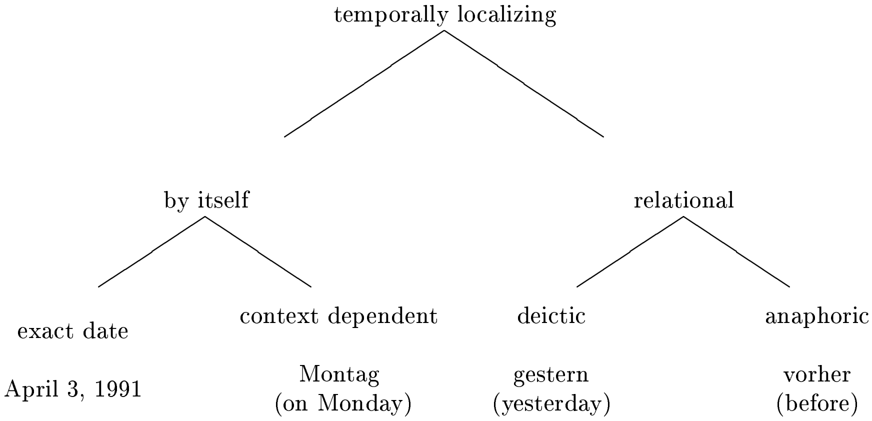


Figure 5.1: The Class of Temporally Localizing Adverbials

The classification distinguishes between modifiers of the verbal phrase that *by itself* introduce a temporal anchor for the event introduced by the VP and modifiers that relate the VP event to a contextually given time. The first class is subdivided into globally *exact dates* and dates that are only locally exact, i.e. with respect to a particular time slice. Since the exactness of this class of dates depends on a contextually given particular time slice, we call them *context dependent* or *pseudo exact* dates (*pexact* for short in the following). Obviously, all exact calendar dates are exact dates, but also events that are cultural common knowledge of the speech community, like *Second World War* or the *reign of Henri VIII.*

The second class consists of the *deictic* expressions, that refer to the temporal parameter of the utterance situation, the *now*, and of the *anaphoric* expressions, that refer to times introduced by the text. Deictic expressions are *jetzt (now), heute (today), gestern (yesterday)* ..., anaphoric expressions are *(am Tag) vorher ((the day) before), danach (after this), währenddessen (during this)* ....

In order to characterize the adverbs along these lines the semantic lexicon of the demonstrator uses the parametrized features as introduced in figure 5.2. [21]

In contrast to Reichenbach's general reference time (Reichenbach (1947)), in figure

[21] Compare Eberle (1995) for a more detailed description of features that characterize the semantic contribution of lexical items.

Feature	Description	
temploc(exact(@SORT,REL-E-L))	exact temporal location with:	
	@SORT:	the domain model sort that is assigned to the location time
	REL-E-L:	the relation between the VP event (E) and the introduced location time (L)
	example:	1994 $\Longrightarrow$temploc(exact(jahre_c,$\subseteq$))
temploc(pexact(@SORT,REL-E-L))	'pseudo' exact temporal location with:	
	@SORT:	the domain model sort that is assigned to the location time
	REL-E-L:	the relation between the VP event and the location time
	example:	montags $\Longrightarrow$temploc(pexact(tage_c,$\subseteq$))
temploc(rel(st,REL-L-R,REL-E-L))	relational location with deictic reference, where:	
	st:	indicates the reference to the speech time and
	REL-L-R:	the relation between the location time and the reference time, which, here, is the speech time
	REL-E-L:	the relation between the VP event and the location time
	example:	heute $\Longrightarrow$temploc(rel(st,$\supseteq$,$\subseteq$))
temploc(rel((rt;pt),REL-L-R,REL-E-L))	relational location with anaphoric reference where:	
	rt ; pt:	indicates the reference to the actual reference time or perspective time and
	REL-L-R:	the relation between the location time and the reference time, and the perspective time respectively
	REL-E-L:	the relation between the VP event and the location time
	example:	später $\Longrightarrow$temploc(rel(rt,$\succ$,$\subseteq$))

Figure 5.2: Temporal Features

5.2 we make use of a fine-grained distinction of location time, reference and perspective time for the analyses of adverbs. This follows suggestions for DRT analyses of tense forms and adverbials as described in Kamp and Rohrer (1985), Eberle (1991), Eberle and Kasper (1994) and others.

The classifying features are used as arguments of the semantic macros that are assigned to the adverbs in the lexicon. The thus classified adverbs, by expanding the macros, are interpreted as partial DRS that, informally, can be rendered as in figure 5.3.

Figure 5.3 shows that adverbs that refer to the speech time introduce an external anchor (cf. Asher (1986) for an early introduction of external anchors into DRT) that links the discourse referent for the speech time to the speech time of the given utterance situation. Adverbs that refer to the reference time (or perspective time) introduce a α-condition. α-conditions specify requirements of presupposition resolution. The introduction of this condition type follows van der Sandt (1992), also in that anaphora resolution is understood as a specific case of presupposition resolution. We emphasize that, in the line of Partee (1973), Kamp and Rohrer (1983) Eberle and Kasper (1989) and others, we conceive the task of finding the temporal coordinates that determine the position of new events and times within the temporal structure of the preceding text as anaphora resolution, and call it *temporal resolution*. The temploc(rel(rt,_,_))-adverbs introduce specific α_{rt}-conditions. That means that the DRF of the α-DRS, Rt, has to be identified with (one of) the actual reference time(s). The α-DRS can specify constraining information about this referent if the adverb or tense provides such information. Accordingly, temploc(rel(pt,_,_))-

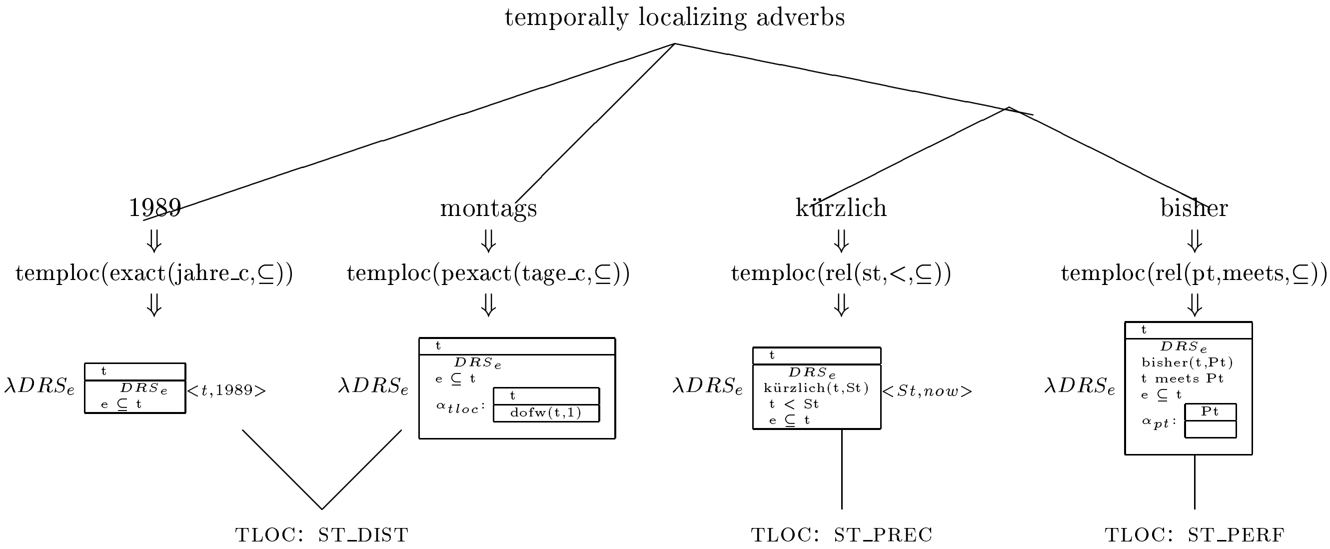


Figure 5.3: The Meaning of Temporal Adverbs

adverbs introduce α_{pt}-conditions that has to be resolved to an actual perspective time. A perspective time is a reference time that, roughly, functions as a transposed *now*: For instance, the point in the past from which a *flashback* starts - a story that is embedded in the main story - is a perspective time with respect to the flashback events. temploc(pexact(_,_))-adverbs introduce α_{tloc}-conditions. The annotation *tloc* triggers the search for a location time of an actual reference event. Normally, in the context of pexcact-adverbs, locations are found that serve as time frame for the unique determination of the α_{tloc}-referent via the sortal description of this referent: For instance, the context provides a week that allows for the exact determination of an adverb whose referent is sorted `tage_c` like *montags*. In figure 5.3 we have used canonical interval relations (cf. Allen (1983)) for describing the relationship between location time and located event, and between location time and the different contextual times in the case of the relational adverbs. In the implementation, however, the domain model provided only a subset of the relevant relations that, since introduced for other purposes, also showed a sometimes misleading naming convention. This and the fact that the semantics omitted to take into account the impact of the adverbs on the Aktionsart of the modified events mark natural extensions of the demonstrator semantics towards a semantics for the verbmobil prototype.

Using the described features, the following entries of the semantic lexicon were specified:

```
lex(morgens) => temporal_adv_sem(morgen1,temploc(pexact(tageszeit_c,temporal_inclusion_rel))).
```

and by the same characterization:

vormittags, mittags, nachmittags, abends

```
lex(montags) => temporal_adv_sem(rpred(montag),temploc(pexact(tage_c,temporal_inclusion_rel))).
```

and by the same characterization:

dienstags, mittwochs, donnerstags, freitags, samstags, sonnabends, sonntags.

(The function `rpred` applied to the days of the week introduces an interlingua description of the day specification, that, therefore, remains constant when the source representation is translated into the target representation, viz. the second partial DRS in figure 5.3).

```
lex(damals) => temporal_adv_sem(damals,temploc(rel(st,temporal_point_posterior_rel,temporal_inclusion_rel))).
```

The other adverbs of the demonstrator corpus classified as related to the speech time were:

gestern, heute, jetzt, kürzlich, morgen, seinerzeit, soeben, übermorgen, vorgestern, vorhin, zur_Zeit.

An entry that exemplifies the relation to a perspective time is:

```
lex(bisher) => temporal_adv_sem(bisher,temploc(rel(pt,temporal_frame_end_rel,dur_temporal_inclusion_rel))).
```

Other adverbs classified as introducing a location time that relates to a contextually given perspective time are:

bislang, eben, einstweilen

The demonstrator class of adverbs with the weakest requirements on the contextual anchor (it must only have the quality of being an actual reference time) consists of

```
lex(bald) => temporal_adv_sem(bald,temploc(rel(rt,temporal_point_anterior_rel,temporal_inclusion_rel))).
```

and of:

bereits, dann, endlich, gleich, inzwischen, nachher, neulich, schließlich, seitdem, sofort, unterdessen, unversehens, vorerst, vorher, zunächst, zwischendurch.

In contrast to exact dates the extension of only locally exact location descriptions does not consist of one single interval or time point but of a set of intervals or time points that is reduced to a unique location time only by means of contextually driven presupposition resolution. So, without the presuppositional contribution, `montags` specifies the introduced location time only as a non further characterized monday. This seems to be the reason why it is possible to comprehend pexact adverbs also as quantiers that claim that each time that is characterized according to the existential reading of the adverb serve as location time for an event that is characterized by the VP representation. The feature `quant` applied to `temploc(pexact(@Sort,Rel))` characterizes this reading. The restrictor of the corresponding duplex condition restricts the variable that is quantified over to @Sort and in the nuclear scope the variable is used as location time of the event, according to `temploc(pexact(@Sort,Rel))`.

```
lex(morgens) => temporal_adv_sem(morgen1,quant(temploc(pexact(tageszeit_c,temporal_inclusion_rel)))).
lex(tags) => temporal_adv_sem(tag1,quant(temploc(pexact(tageszeit_c,temporal_inclusion_rel)))).
lex(montags) => temporal_adv_sem(rpred(montag),quant(temploc(pexact(tage_c,temporal_inclusion_rel)))).
lex(feiertags) => temporal_adv_sem(feiertag,quant(temploc(pexact(tage_c,temporal_inclusion_rel)))).
```

and similarly the other `tageszeit_c`- and `tage_c`-characterizations.
Other quantifying expressions without a similar characterization of the range that is quantified over are:

```
lex(einmal) => temporal_adv_sem(einmal,quant(exist)).
lex(immer) => temporal_adv_sem(immer,quant(univ)).
lex(jedesmal) => temporal_adv_sem(immer,quant(univ)).
lex(nie) => temporal_adv_sem(nie,quant(neg)).
lex(niemals) => temporal_adv_sem(nie,quant(neg)).
lex(dauernd) => temporal_adv_sem(dauernd,quant(pos)).
```

The assumption is that with these characterizations quantifiers over non further specified situations are introduced such that the situations embed the event from the VP description. Next to *dauernd* the other positive temporal quantifiers in Verbmobil are:

häufig, je, kaum, manchmal, mehrmals, meist, oft, selten, so_oft, vielmals, x_mal, zig_mal.

The German present tense is notoriously ambiguous when compared to its possible English translations. Simplifying the data, normally, on the one hand the German present tense introduces an event or state that overlaps with the situational now. In this case the target tense form will be present or present progressive. On the other hand, the event or state reported can be situated after the now. In that case the target tense will be future (cf. Butt (1995b)). Localizing adverbs help to obtain the correct translations. Without going into detail, we mention that exact and 'pseudo' exact dates per default do not overlap with a given *now* and, therefore, anchor the VP-event or -state at a certain distance of the now, what results in a non-present translation of German present tense in such cases. In order to keep track of this information for the transfer, the semantics of the temporal adverbs specify a feature `tloc`. Exact and locally exact dates introduce the value `st_dist`: i.e. the location time does not overlap with the now of the utterance (at least this is the assumption as long as there is no contradicting information from an inference module). Relational adverbs can introduce other values that are relevant for the translation of the tenses (in particular of German present tense). Figure 5.3 shows that *kürzlich* introduces the value `st_prec`. This value marks the present tense of a VP that is modified by a corresponding adverbial as historical present, because in such cases the introduced event or state is located before the now of the utterance. The value `st_perf`, that is introduced by adverbs like *bisher*, means that a location time is introduced that directly relates to the speech time, or that relates to the speech time via identification of perspective time and speech time in case of present tense utterances, in such a way that the present tense of a VP that is modified by a corresponding adverbial normally is translated into English present perfect.

For details of the analysis of the tenses in Verbmobil, compare the section 9 in this handbook and Butt (1995b).

5.14.2 Semantic Representation of Temporal Adverbs

The semantic representation of temporal adverbs followed the classification proposed by K. Eberle in section 5.14.1.

5.14.3 Transfer of Temporal Adverbs

As described above, within the recursive transfer concept the actual formulation of transfer rules for adverbs is fairly straightforward. For most of the temporal adverbs in table 5.14.3, the standard transfer rule, repeated below as ex. 5.29, was used.

SourcePred	→	TargetPred
bald	→	soon
damals	→	then
dann	→	then
einmal	→	once
einstweilen	→	for_the_time_being
gerade/grade	→	right, now, just
gestern	→	yesterday
gleich1	→	right
gleich2	→	right_away
heute	→	today
irgendwann	→	at_some_point
kürzlich	→	recently
mal	→	null_pred
morgen	→	tomorrow
nachher	→	afterwards
neulich	→	the_other_day
seinerzeit	→	at_that_time
seitdem	→	since_then
soeben	→	just_now
sofort	→	right_away
übermorgen	→	day_after_tomorrow
unterdessen	→	in_the_meantime
vorgestern	→	day_before_yesterday
vorher	→	before_then
vorhin	→	a_little_while_ago
zwischendurch	→	in_between

(5.29)
```
tau_lex(SourcePred, TargetPred, [tau(#Mod,#Args)],#Args,
            mod([#Mod])) => sem_t.
```

The determination of transfer equivalences of these adverbs was based on an examination of Verbmobil dialogs available at the time, in particular on the *Blaubeurer Dialoge*. In some cases, additional conditions had to be postulated in order for transfer to take place. These are described below.

5.14.3.0.1 Interesting Cases Given that contextual information is not available for the disambiguation and translation of anaphoric adverbs, it has been necessary to find the most generally applicable translation for highly ambiguous items like *da* and *dann* (Ehrich 1992), *mal*, *gerade* and *gleich*.

Da and Dann The pronominal *da* can function as a locative or temporal adverb, a true locative pronoun, as a conjunction equivalent to the English *as* or *since*, or as a discourse particle. All of these readings, except for the locative, are also possible for *dann*: it can serve as a temporal adverb, as a sequential conjunction equivalent to the English *then*, or as a discourse particle.

Within the current implementation, only one variety of *da* and *dann* each are recognized by the Semantic Construction: locative *da* and temporal *dann*. The task of transfer is simplified in the extreme: *da* is realized in an interlingua representation for locative *there*, and *dann* is always translated as *then*.

For a more in-depth discussion of *da* and *dann* see Hamp (1995) and Butt (1995).

Mal, Gleich, and Gerade The German *mal* was originally analyzed as a quantifier in the semantic lexicon, in analogy to *einmal, zweimal*, etc. However, a close look at its usage shows that *mal* does not have any quantificational effects, but instead functions as a discourse particle with no overt English translational equivalent.

(1) KarlsD7:9 ich bin **mal** durch den Terminkalender gestolpert ...
KarlsE7:9 I stumbled through my appointment-book ...

The adverbs *gleich* and *gerade* are also used similarly as discourse particles which serve to focus, or intensify, a particular part of an utterance ((2)). Both can also be used as temporal adverbs ((3)).

(2) a. D1:11 wollen wir's dann **gleich** am Montag den dritten Mai machen
E1:11 do we want to do it **right** on Monday the third of May then

b. KarlsD2:16 warten Sie **gerade** fünf Minuten auf mich
KarlsE2:16 **just** wait five minutes for me

(3) a. KarlsD2:16 ich sage meiner Sekretärin **gleich** Bescheid
KarlsE2:16 I will tell my secretary **right away**

b. KarlsD7:1 wo bist Du **gerade**
KarlsE7:1 where are you **now**

Another very interesting use of *gerade*, which we have not been able to take into account for the Demonstrator, is shown in (4) and (5).

(4) KarlsD10:5 ja elf Uhr fünfzehn sehe ich **gerade** da bin ich in einer Besprechung
KarlsE10:5 yes, eleven fifteen, I've **just** noticed I am in a meeting then

(5) Ich erkäre ihm **gerade** den Weg.
I am/was **just** telling him how to get there.

The examples in (4) and (5) illustrate an interesting difference between a perfect and a progressive use. Since the progressive is usually analysed as being sensitive to Aktionsart, this possibility was investigated. However, the relevant parameter with respect to *gerade* seems to be *punctuality*, which is not factored into the representation of Aktionsart: in (5) the explanation of a route may take a while, while in (6) the German *einfallen* can only be momentaneous.

(6) ach, da fällt mir gerade was ein
oh, I just remembered something

A more extensive look at the parameter of punctuality is necessary in order to determine whether it should be factored into the model of Aktionsarten, as is sometimes suggested with regard to achievements, or whether it more properly belongs in the realm of world knowledge as suggested by Egg (1994).

5.14.3.0.2 Future Work A more extensive analysis of the adverbs needs to be undertaken, in terms of both a more extensive coverage of data, and an extension of the semantic evaluation of the temporal information that is contributed by these adverbs.

5.15 Focus Adverbs

This section describes the treatment of focus adverbs in the Transfer module of the MDS. We understand under *focus adverbs* the semantic subclass refered to in the literature as *focus particles* (see König (1991)) or *focus sensitive particles* (see Bos (1995)). After Johan Bos these particles do not introduce a focus themselves, but they have the property to apply to constituents, which are focused and prosodically stressed. The focus sensitive particles introduce presuppositions, which depend on the constituents in focus that appear in their scope. Furthermore, Bos claims that "the focus particles do not add anything to the meaning of the sentence, but rather "judge" whether the sentence in which they appear is acceptable in a given context or not." König on the other hand observes that focus particles in German and in English do have a lexical meaning, and outlines general parameters that play a role in their semantic analysis. These parameters are: 1) the scope of the focus particles, 2) alternatives – focus particles set restrictions on the selection of alternatives, in the framework of conception of focus as a relation between the value of a focused expression and a set of alternatives 3) scales – focus particles set selection restrictions for alternatives ordered with respect to the focus value in a certain way, 4) evaluation – focus particles set selection restrictions inducing an order for the value of the focus and the altrenatives under consideration also express an evaluation. He also emphasizes that the contribution made by a particle to the meaning of a sentence depends on the meaning of two components of that sentence: (a) on that of its focus and (b) on that of its scope. It seems that the structural account considered by Bos, and the meaning account considered by König in combination will give an optimal framework to deal with focus adverbs.

The semantic representation of focus adverbs in MDS is discussed in the following section.

5.15.1 Semantic Construction of Focus Adverbs in MDS

Focus adverbs were defined in the semantic construction of MDS as sentence adverbs with scope over the whole sentence. They introduce a *phi_condition* into the semantic structure (see ex.(2.29)), and their scope is designed to be always over constituents describing *states* or *events* (see phi_arg in ex.(2.29)).

(5.30) Er kommt AUCH.

$$\begin{bmatrix} DRS \\ \begin{bmatrix} DRS\_t \\ phi\_condition \\ phi\_arg => & \begin{bmatrix} kommen : \boxed{1} \\ \end{bmatrix} \\ phi\_op : alternative \\ phi\_pred : auch \end{bmatrix} \end{bmatrix}$$

Each focus adverb is assigned a semantic type refering to the presupposition it conveys. A feature FOCUSOP (focus operator) is introduced into the semantic lexicon to make this explicit. Thus, the form of the lexical entry for focus adverbs in the semantic lexicon consists of a lexeme, a semantic subclass, a PredName and a FocusOp (see ex.(2.30)):

(5.31) `lex(LEXEME) => focus_adv_sem(Pred, FocusOp)`

The attributed values of FocusOp in the Semantic Construction of MDS are: *alternative, temp, event, grad* (see table 5.4). They reflect the pursued analysis of the possible semantic effects triggered by focus adverbs[22].

ADVERB	PredName	FocusOp
auch	auch	alternative
denn	denn1	##
erst	erst	temp
gerade	gerade	event
höchstens	höchstens	grad
noch	noch	temp;event

Table 5.4: Values of the feature FocusOp in the MDS with examples of focus adverbs.

5.15.2 Size of the Data

Focus adverbs in the wordlist issued by the University of Bielefeld (see sem_lex2.stuf) count 33:

alternativ auch ausserdem doch denn dann erst gerade grade gleich höchstens insofern ja mal mindestens noch nochmal nur schon selbst sogar sonst wenigstens wieder wiederum wohl zumindest ausnahmsweise überhaupt ausschliesslich frühestens spätestens später

Sixteen of them occur in the Testsuite of Dialogues for MDS:

dann noch auch ja doch denn nur schon nochmal sonst gleich gerade erst wieder ausnahmsweise mal

5.15.3 Determination of Readings and PredNames

The lexically ambiguous adverb *höchstens* will be the example for the strategy adopted in the Transfer module of MDS to chose equally ambiguous English equivalents of German PredNames to be the transfer PredNames.

[22]As the purpose of the present paper is to report about the actual realizations in MDS, we will not comment here further on the transparent possiblities to extent the semantic classification of focus adverbs, and on ways to interpret the defined FocusOp. As *auch*, *ausserdem* and *ausnahmsweise* were given the same FocusOp value: *alternative*, but the character "things" refered to by these alternatives is not intuitively one and the same.

The translations of *höchstens*, and respectively their readings, in Pons (see [Pons83]) are: 1) *not more than* (nicht länger als; nicht mehr); 2) *at the most, at best* (bestenfalls); 3) *except* (außer). *Höchstens* is translated in the testsuite of VerbMobil with the words: *only*, *only possibly*, *possibly* (ex.(2.31),(2.32)). These lexical equivalents correspond to the second reading quoted in Pons: *at the most, at best*. The other two readings do not occur in the context of VerbMobil.

(5.32) *höchstens* → *only possibly*
HOMN038:
mm, den 28sten kann ich auch nicht, da bin ich in Berlin, 21sten hab ich allerdings auch einen sehr langen Termin, da seh ich schwarz, daß ich den verschieben kann. wir könnten's HÖCHSTENS so machen, am 7ten und am 14te
Mm, I can't make the twenty-eighth either, I'll be in Berlin then, on the twenty-first I do also already have a very long appointment, I don't think I can reschedule that. We could ONLY POSSIBLY *do it so that it is on the seventh and the fourteenth.*

(5.33) *höchstens* → *only*
HOMN044:
den 6ten bin ich leider auch außer Haus, da wär, seh ich auch keine Möglichkeit, das zu verschieben. es wär HÖCHSTENS dann 13te, ja , da könnt ich vormittags, und,
eh, I'll be out of town on the sixteenth as well, there is, I don't see a rescheduling possibility there either. There would ONLY *be the thirteenth, yes, I am free in the morning*

A problem was to decide how to represent this reading in PredNames in order to provide enough information to the Generation that the lexical meaning *at the most* is to be lexicalized into *possibly*, *only* in that particular contexts.

Two possibilities to cope with this problem exist:

1. to write one TAU rule which would transfer the predicate *höchstens* into the predicate *only_possibly.*

2. to find the reason why it is possible from the lexical meaning *at the most* to get the lexicalization *only possibly,* and to motivate by structural representation and compositional account the lexical mismatches.

The transfer rule of *höchstens* in MDS consented the first option (see (ex.(2.33))):

(5.34) höchstens → only_possibly

In addition, we provided an analysis of the interpretation of *höchstens* in the context of VerbMobil and informally proposed ways to consider the second option, which are briefly discussed in the following.

In the two examples above *höchstens* occurs in contexts where possibilities about temporal location of one event are described. The possibility is denoted in the first example by the conjunctive, and the meaning of the modal verb *können* itself. In the second example – by the noun *Möglichkeit* (possibility), and the conjunctive of *sein* (to be) – *wäre*. Furthermore, the morphology of *höchstens* is a superlative degree of the adverb *hoch*. The comparison degrees of adverbs and adjectives have particular semantic content. Thus, information about the superlative degree and the scalar character of the adverb should be available in the semantic lexicon, and information about the semantic content of the entity the adverb has scope over (in this particular case – a *possibility*) should be provided. The interpretation of the combination between the meaning of the superlative, the scale and the possibility derives justified lexicalisation of *höchstens* into *only* and *possibly*, which corresponds to the intuitions conveyed by the examples above (ex.(2.31),(2.32)). [23]

In order this to be realized, it is necessary to introduce the relevant distinguishing marks at relevant places in the lexicons of the system, and to make sure that the representation of the right semantic structures with the available mechanisms will be possible.

5.15.4 Disambiguation

The highly ambiguous adverb *noch* will be our example to show the process of disambiguation of focus adverbs in MDS. We considered two readings of this adverb: 1) noch temporal (ex.(2.34)), 2) noch eventive (ex.(2.35)), which reflect the semantic entity each of them has scope over. *Noch* in the first reading has scope over temporal location of an eventuality described, and in the second reading – over the event described. These two readings were felicitously translated into the English *still*, and *just*.

(5.35) *noch* temporal → *still*
KA:11:MPS1_1_03:
tut mir leid, am 13ten April, bin ich NOCH im Urlaub.
I am sorry. on the thirteenth of April I will STILL *be on vacation*

(5.36) *noch* eventive → *just*
HI:2:MPS1_1_09:
Wenn Sie mir NOCH kurz erklären wie ich zu Ihnen komme
If you would JUST *briefly explain how to reach you*

The lexical entry of *noch* contains two disjoint FocusOp relevant each for one of the two cited readings (ex.(2.36)).

(5.37) `lex(noch) => focus_adv_sem(noch,temp;event).`

The semantic construction delivers an ambiguous output, as the value of the phi_op contains two disjoint values (ex.(2.37)).

[23] The lexical entry for *höchstens* should contain the following information: "this is a focus adverb, which is scalar (i.e. it refers structurally to a semantic entity, which can be evaluated through a scale), and refers to the superlative value of the scala. What scale is exactly considered is determined structurally through the scope of the adverb. If the scope is on a semantic entity, which describes a possibility, regardless from the fact whether this information comes from a verb, or from a noun, one looks for the quantitative values of this scale of the possibilities. It is obvious that there is only one possibility which could be at the superlative value of the scale of possibilities. After all this the lexicalisation of *höchstens* into the English *only* can be justified. Moreover, this interpretation corresponds to the intuitions conveyed by the examples above.

(5.38) Ende Juni bin ich *noch* im Urlaub.

$$\begin{bmatrix} DRS \\ \begin{bmatrix} DRS\_t \\ phi\_condition \\ phi\_arg => & \begin{bmatrix} im\_Urlaub\_sein : \boxed{1} \\ \end{bmatrix} & timeloc : Ende\_Juni \\ phi\_op : temp; event \\ phi\_pred : noch \end{bmatrix} \end{bmatrix}$$

The disambiguation of *noch* takes place at the level of Transfer. The marks for disambiguation were the semantic indexes of the sorts of the eventuality, which was in the scope of the adverb. Thus, a stative eventuality triggered the temporal reading of *noch* (ex.(2.38)), and an event triggered the eventive reading of *noch*(ex.(2.39)).

(5.39) Ende Juni bin ich *noch* im Urlaub.

$$\begin{bmatrix} DRS \\ \begin{bmatrix} DRS\_t \\ phi\_condition \\ phi\_arg => & \begin{bmatrix} marker \\ sort : static \\ to\_be\_on\_vacation : \boxed{1} \\ \end{bmatrix} & timeloc : end\_of\_June \\ phi\_op : temp \\ phi\_pred : still \end{bmatrix} \end{bmatrix}$$

(5.40) Ich erkläre *noch* kurz.

$$\begin{bmatrix} DRS \\ \begin{bmatrix} DRS\_t \\ phi\_condition \\ phi\_arg => & \begin{bmatrix} marker \\ sort : dynamic \\ kurz\_erklären : \boxed{1} \\ \end{bmatrix} \\ phi\_op : event \\ phi\_pred : just \end{bmatrix} \end{bmatrix}$$

The conditions for disambiguation are intergrated in the two TAU rules produced to cover the two readings of *noch* (ex.(2.40),(2.41)).

(5.41) `tau_lex(noch, still, [tau(#Mod,#Args)], #Args, mod([#Mod]) & sem:idx_sem(sort:statisch_c)) => sem_t.`

(5.42) `tau_lex(noch, just,[tau(#Mod,#Args)],#Args,mod([#Mod]) & sem:idx_sem(sort: ~statisch_c)) => sem_t.`

5.15.5 Transfer PredNames of Focus Adverbs in MDS

The transfer of focus adverbs was basically carried out at level of PredNames. The shape of the transfer rule was the standard one, quoted in Section 5, ex.(2.3), and repeated here as ex.(2.42).

(5.43) `tau_lex(SourcePredName, TargetPredName,[tau(#Mod,#Args)],#Args, mod([#Mod]) => sem_t.`

The German PredNames with their English equivalents are listed informally in table 5.5.

SourcePred	→	TargetPred
alternativ	→	alternatively
auch	→	too
ausserdem	→	furthermore
dann1	→	then1
denn1	→	null_pred
doch	→	after_all
erst	→	null_pred
erst	→	only
hoechstens	→	only_possibly
ja1	→	null_pred
insofern	→	so_far
mindestens	→	at_least
noch	→	still
noch	→	just
nochmal	→	again
nur	→	only
selbst1	→	even
schon	→	already
sogar	→	even
sonst	→	otherwise
wenigstens	→	at_least
wieder	→	again
wiederum	→	on_the_other_hand1
wohl	→	null_pred
zumindest	→	at_least
ausnahmsweise	→	for_once
fruehestens	→	at_the_earliest
spaetestens	→	at_the_latest
spaeter1	→	later1

Table 5.5: Transfer PredNames of the focus adverbs in the MDS.

The following section outlines the results of the semantic analysis of focus adverbs, which were not implemented in MDS. It addresses the problems of ambiguities, and heuristics for disambiguation in the processing of *focus adverbs*.

5.15.6 Scope Ambiguities and Heuristics by Processing of Focus Adverbs

Focus adverbs were designed in the Semantic Construction module of MDS as particles which can have scope over *states* or *events*. The examples in (2.43),(2.44), show however that this is not always true. *Noch* has scope over the negation in (2.43), and over the individual *Termin* in (2.44). Furthermore, examples (2.43) and (2.44) illustrate two more readings of the adverb, which were not accounted for in MDS. These two additional readings can be seen in their translations into English. *Noch* with scope over negation is translated into *not yet*, and with scope over the individual – into *another*.

(5.44) HI:5:MOC1_1_18:
Das kann ich NOCH nicht sagen.

(5.45) MHK:1_1_02:
aber wir brauchen NOCH einen Termin.

The two sentences in ex.(2.43) and (2.44) conform different semantic representations, from the one given in ex.(2.29). The disambiguation conditions in these two cases are to be found in the type of semantic entity which falls in the scope of the adverb *noch*. In fact if a structurally suitable output of the Semantic Construction will be provided, part of the disambiguation process will take place at that level, and consequently this will make possible the correct transfer of these two occurrences of *noch*.

Moreover, *noch* shows further ambiguities, depending on the prosodically stressed constituent of the sentence, as shown in ex.(2.45a-c). The sentence in (2.45a) describes a situation in which an enumeration of events is presupposed (reading: *among other events*), the sentence in (2.45b) describes a situation in which an enumeration of individuals is presupposed (reading: *one more x*), the sentence in (2.45c) describes a situation in which a final element of a finite enumeration of individuals is presupposed (reading: *only one more x*).

(5.46) (a) aber wir brauchen **noch** einen Termin
(b) aber wir brauchen *noch* einen **TERMIN**
(c) aber wir brauchen *noch* **EINEN** Termin

These three examples would trigger different representations. Information helpful for deciding what semantic structure is relevant can be provided by the output of the prosodic profile of the expression (see ex.(2.46 a-b)).[24]

(5.47) (a) *noch* with scope over individuals → *one more x*
DATEN/N002K/NPS1K002.CPR:
sollen wir gleich im März NOCH **EINEN ANDERN** Termin ausmachen oder wann paßt 's Ihnen am besten

(b) *noch* with scope over a described event → *among other events*
DATEN/N019K/NHK1K002.APR:

[24] The analysis and the interpretation of the readings of *noch* depending on the prosodic information were discussed with and approved by Jorg Meier.

offensichtlich am Telefon aber wir brauchen NOCH einen **TERMIN** soweit ich weiß

Thus, the scope of *noch*[25] and the semantic properties of the entity which is in its scope, are of great importance to determine its reading. A process of semantic evaluation resulting in a structured semantic representation which accounts for the scope of the adverb will provide conditions to decrease the need of disambiguation in the Transfer module. A consideration of the prosodic profile of the expressions in the process of semantic evaluation will furnish necessary information with this respect.

5.15.7 Future work

It is still necessary to decide how to distinguish formally between the different readings of one adverb and then how it is possible to represent them without losing valuable semantic information. Means for structural repesentation of the scope, incorporation of prosodic information, and clearly elaborated account for the semantics of the *focus operators* will be of importance to achieve an efficient future development of the system VerbMobil.

[25] We used this adverb to show some phenomena proper for focus adverbs in general.

5.16 Intensifiers

This section describes the treatment of *intensifiers* in the Transfer module of MDS. The subclass of *intensifiers* has the property to apply to *adverbs* or *adjectives*, i.e. *modifiers* with the semantic effect to introduce additional information on the degree of intensity of the modifiers (ex.(2.47a-b).

(5.48) (a) Peter kommt SEHR(int) SPÄT(adv).
(b) Peter ließt ein SEHR(int) INTERESSANTES(adj) Buch.

Thus, an intensifier and a modifier form one constituent, which relates as a whole to the rest of the expressions they occur in.

5.16.1 Semantic Construction and Transfer of Intensifiers

The intensifiers in MDS are designed as applying to *dimentional adjectives* (see Section 5.19), which occur in expressions as adjectives or as adverbs (ex.(2.48a-b)).

(5.49) (a) Peter kommt SEHR(int) SPÄT(adv).
(b) Wir machen einen SEHR(int) SPÄTEN(adj) Termin aus.

And the intensifiers are represented within the dimension condition introduced by the dimensional adjective. They are refered to with the feature: DIMEN_INTENSITY (see ex.(2.49)).

(5.50) Peter kommt SEHR SPÄT.

$$\begin{bmatrix} DRS \\ \begin{bmatrix} kommen: \boxed{1} \\ \begin{bmatrix} dimen\_condition \\ dimen\_inst: \boxed{1} \\ dimen\_intensity: sehr \\ dimen\_pred: spaet \\ dimension: timeloc \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

A special TAU rule provides the transfer of this semantic group (see ex.(2.50)), which combines the lexical TAU rule with one semantic macro.

(5.51)
```
tau_lex(SourcePred, TargetPred, tau_intens(#Args,#Sign), #Args,#Sign) =>
    sem_t.

    @tau_intens(list,sign) => list.
    tau_intens(#Args, ad_syn(#Mod & pred_name(#SourcePred))) =>
        true(tau_lex(#SourcePred,#TargetPred,[#XPSem],#Args,#Mod)) &
        [sem_lex(#TargetPred,[#XPSem]),#XPSem].

    tau_intens(#Args, ad_syn(#Mod & pred_name(#SourcePred))) =>
        true(tau_lex(#SourcePred,#TargetPred,[#NPSem,#XPSem],#Args,#Mod)) &
        [sem_lex(#TargetPred,[#NPSem,#XPSem]),#XPSem].
```

SourcePred	→	TargetPred
echt2	→	null_pred
echt2	→	really2
durchaus	→	absolutely
eher	→	rather1
etwas1	→	a_bit1
sehr	→	very1
ueberaus	→	extremely
zu3	→	too1

Table 5.6: Transfer PredNames of the intensifiers in the MDS.

5.16.2 Size of the Data and Transfer PerdNames for Intensifiers

The intensifiers in the Wordlist for the Demonstrator (after sem_lex2.stuf) count seven:

echt, durchaus, eher, etwas, sehr, überaus, zu.

Five of them occur in the Testsuite of Dialogues for MDS:

echt, eher, etwas, sehr, zu.

The transfer PredNames of intensifiers are informally listed in Table 5.6.

5.16.3 Future Work

The semantic classes of adverbs and adjectives which can combine with intensifiers are still to be set apart. Furthermore, some of the adverbs classified in the group of standard adverbs can occur in expressions as intensifiers (see ex.(2.54)). Deeper semantic analysis of the semantic and combinatoric properties of intensifiers will affect the present subdivisions of the classes of adverbs, and will provide conditions for even more effective transfer.

5.17 Standard Adverbs

This section describes the treatment of *standard adverbs* in the Transfer module of MDS. The semantic subclass of *standard adverbs* illustrates the initial general conception of the role of adverbs as modifiers in VerbMobil. They have the property to apply syntactically to an entire expression, and to refer semantically to different perspectives or circumstances of the eventuality described by the expression. This structural definition did not account for further details concerning the semantic roles of the modifiers. We will present in the following section the design of *standard adverbs* in the Semantic Construction module of MDS.

5.17.1 Standard Adverbs in the Semantic Construction Module of MDS

Standard adverbs are analyzed as separate entity in the semantic representation, they intorduce a BASIC CONDITION, which only contains explicit a PredName, as shown in ex.(2.51).

(5.52) LEIDER komme ich montags.

$$\begin{bmatrix} DRS \\ \begin{bmatrix} kommen : \boxed{1} \begin{bmatrix} tlocation : montags \\ inst : \boxed{1} \end{bmatrix} \begin{bmatrix} basic\_condition \\ pred : \ leider \\ inst : \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Thus, the semantic lexical entry for *standard adverbs* is of the following form (see ex.(2.52)):

(5.53) `lex(LEXEME) => adv_sem(Pred)`

5.17.2 Size of the Data

Standard adverbs in the Wordlist for the Demonstrator (after sem_lex2.stuf) count twenty two:

beinahe, derart, etwa, gar, genaugenommen, glücklicherweise, halbwegs, hindurch, leid, leider, meinetwegen, nah, offensichtlich, pünktlich, überhaupt, umsonst, unglücklicherweise, weg, zusammen, besonders, insbesondere, insgesamt.

Two of them occur in the Testsuite of Dialogues for MDS:

leid, leider.

Most of the listed adverbs can be intuitively assigned semantic properties of the defined in MDS subcalsses. For example *überhaupt* can be analyzed as focus adverb, as it has scope over different entities (ex.(2.11a-c) in Section 2.10.6.2.). *Leider* and *unglücklicherweise* can be classified as modal adverbs, as they seem to introduce a modality conneceted to previous discourse, as *allerdings* does (see ex.(2.51),(2.53)), and compare with ex.(2.26),(2.27),(2.28)).

(5.54) WIL016/GRA017:
UNGLÜCKLICHERWEISE habe ich am Mittwoch, den achten Juno wieder ab mittags eine Konferenz hier in Hamburg.
UNFORTUNATELY *on Wednesday the eighth I have a conference in Hamburg.*

Beinahe and *etwa* can be in the semantic subclass of intensifiers, as they also apply to other modifiers (see ex.(2.54) and compare with ex.(2.29a-b)).

(5.55) MPS1_1_19:
Ist bei mir ETWA SCHLECHT(adv).
That's not SO *good.*

And *offensichtlich*, which is morphologically an adjective, can be assigned only the semantic class of dimensional adjectives.

The next section considers the determination of PredNames, and the furnishing of transfer rules.

5.17.3 Transfer PredNames of Standard Adverbs in MDS

The determination of the readings and the PredNames of the group of *standard adverbs* was pursued after the adopted strategies in the Transfer module (see Section 2.10). The transfer PredNames of *standard adverbs* are informally listed in Table 5.7.

The transfer rule for standard adverbs is the general one, cited in Section 5, ex.(2.3), and repeated here as ex.(2.55).

(5.56)
```
tau_lex(SourcePred, TargetPred,  [tau(#Mod,#Args)], #Args,
    mod([#Mod])) => sem_t.
```

5.17.4 Future Work

Many of the adverbs classified as *standard adverbs* seem to belong to other semantic subclasses. This makes difficult to motivate the existance of a separate class of standard adverbs. Thus, it is necessary to revise the semantic status of the group of *standard adverbs* in comparison with the other semantic subclasses of adverbs. Furthermore it is necessary to review the listed items as standard adverbs in MDS, and to reclassify them, in order to assign them the appropriate semantic subclass according to their semantic properties.

SourcePred	→	TargetPred
beinahe	→	almost
derart	→	so
etwa	→	a_bit
etwa	→	somewhat
etwas2	→	a_bit
gar	→	at_all
genaugenommen	→	strictly_speaking
gluecklicherweise	→	luckily
halbwegs	→	halfway
halbwegs	→	more_or_less
hindurch	→	through1
hindurch	→	throughout1
leid	→	sorry
leider	→	unfortunately
meinetwegen	→	for_my_part
nah	→	near1
offensichtlich2	→	obviously
puenktlich	→	sharp1
ueberhaupt	→	null_pred
ueberhaupt	→	generally
umsonst	→	for_nothing
ungluecklicherweise	→	unfortunately
weg1	→	away1
zusammen1	→	together
besonders	→	mainly
insbesondere	→	above_all
insgesamt	→	altogether

Table 5.7: Transfer PredNames of the standard adverbs in the MDS.

5.18 Discourse Relations – *DISCREL*

This section describes the treatment of word classes refering to "discourse relations" (henceforth *discrel*) in the Transfer module of MDS[26]. The semantic class of *discel* contained *conjunctions* and *adverbs* with the properties to bind two expressions (see ex.(2.56), (2.57)).

(5.57) MPS1_2_02:
WENN(conj) ich da so meinen Terminkalender anschaue, das sieht schlecht aus. ich kann Ihnen den Dienstag, 6ten April anbieten. ODER(conj) Freitag. den 16ten April.

(5.58) MPS1_1_13:
Montag 12ter April ist bei mir schon voll. BEZIEHUNGSWEISE(adv) da bin ich noch im Urlaub.
Monday the twelfth of April is already tight for me. THAT IS *I am still on vacation then.*

So, lexemes of these word classes cannot be interpreted within one single expression. Their semantic effect can be rhetorical or truth conditional, as they can supply information on how the expressions relate to each other (the role of *bezihungsweise* in ex.(2.57)), or how the eventualities described by these expressions relate to each other (the role of *wenn* and *oder* in ex.(2.56)). The word classes describing discourse relations were devided into subgroups which we will discuss in the following section.

5.18.1 Semantic Construction and Transfer of *Discrel* in MDS

The word classes describing discourse relations were devided into several subgroups according to the type of relation they introduce, and some syntactic criteria:

1. discrel (proper) - this class was assigned to adverbs and conjunctions intorducing particular rhetorical relations in the dialogues (ex.(2.58),(2.59)).

   (5.59) *nämlich* introducing a discourse relation of *explanation*
   MMM4_1_04:
   ja das kombiniert gut. ich hab[e] NÄMLICH im 1ten März-Woche auch nichts Größeres vor und so könnten wir da unseren Termin in der 1ten Woche zwischen Montag dem 1ten und Freitag dem 5ten legen.

   (5.60) *folglich* introducing a discourse relation of *consequence*
   RAL003:
   aber dann hätten wir auch keine 5 Tage mehr, da ich schon am 18ten in Lübeck sein muß. es wär' also FOLGLICH vielleicht günstiger, wir würden am 20sten irgendwie wegfahren. und in der Zeit vom 20sten bis zum 28sten, da hätte ich natürlich viel Zeit.

   The types of relations were refered to in the semantic lexicon with the feature: DISC_REL_VAL (ex.(2.60)), whose values are: *contrast, perspective, elaboration, explanation, consequence, instance, disjunction, purpose.*

[26] The specific transfer rule which will be described in this section were produced by Michel Dorna.

(5.61) `lex(lexeme) => disc_rel_sem(Pred,disc_rel_val,disc_syn_val)`

Adverbs expressing *discourse relations* in the Word list issued by the University of Bielefeld (see sem_lex2.stuf) count twenty two:

aber, als, als_auch, also, beziehungsweise, dagegen, daher, damit, darum, das_heißt, demzufolge, denn, dennoch, deshalb, desto, deswegen, folglich, indem, infolgedessen, jedoch, nämlich, umso, und, und_so_weiter, ob, obwohl, oder, sodass, somit, sondern, soweit, sowie, sowohl, trotzdem, weil, zumal, zwar.

The transfer of this subgroup of *discrel* is executed through a combination of a lexical TAU rule and one semantic macro, in which the syntactic role of the transfered item[27] is taken into consideration (ex.(2.61)).

(5.62)
```
tau_lex(SourcePredName, TargetPredName, #Taus,
    #Args, #Sign) => tau_DISC_SYN_VAL(#Sign,#Args,#Taus).

    @tau_DISC_SYN_VAL(sign,list,list) => sem_t.
    tau_DISC_SYN_VAL(DISC_SYN_VAL_syn(#Mod),#Args,[tau(#Mod,#Args)])
    => sem_t.
```

2. coordination - this class included the conjunctions *und* and *oder* (ex.(2.62)).

   (5.63) `lex(lexeme) => coordination_sem(Pred)`

   The two of them were treated in a distinct way in the Syntaxe and the Semantic Construction modules, as they can bind sentences or just constituents of sentences (ex.(2.63),(2.64),(2.65)).

   (5.64) RAL005:
   Aschermittwoch, dem 24sten bis hinein Anfang März, da könnten wir das doch [*packen*] UND [*wegfahren*].

   (5.65) RAL016:
   bei mir sieht das so aus, daß ich [*am Dienstag*] UND [*am Freitag*] also Zeit hätte.

   (5.66) AEL013:
   mein Name ist Niehmeyer UND ich möchte mit Ihnen ein Treffen vereinbaren.

   Special transfer rules were provided to process these coordination cases. They consisted of combination of lexical TAU rule (ex.(2.66)) and several semantic macros (ex.(2.67)), one of which was to be selected according to the concrete type of syntactic coordination (compare ex.(2.63),(2.64),(2,65)).

[27]The syntactic roles of the item were provided by the feature DISC_SYN_VAL with values: *subord, coord, adverb*, which was made explicit in the semantic lexical entry.

(5.67)
```
tau_lex(und, #Pred, #Taus, [], #Sign) =>
                tau_coordination(#Sign,and1,#Pred,#Taus).
tau_lex(oder, #Pred, #Taus,[], #Sign) =>
                tau_coordination(#Sign,or1,#Pred,#Taus).
tau_lex(dummyconjshead, #Pred, #Taus, [], #Sign) =>
                tau_coordination(#Sign,dummyconjshead,#Pred,#Taus).
```

(5.68)
```
@tau_coordination(sign,atom,atom,list) => sem_t.

% non-raised case
tau_coordination(category(nominal & ~modifier) & conj_syn(#Arg1,#Arg2),
                 #Pred, #Pred, [tau(#Arg1),tau(#Arg2)]) => sem_t.

% raised cases
tau_coordination(category(fc_nom & ~modifier) & conj_syn(#Arg1,#Arg2),
                 #Pred,
                 n2indefnp_pred,[coordination_sem1(#Pred,[tau(#Arg1),
                 tau(#Arg2)])]) => sem_t.

tau_coordination(category(modifier) & conj_syn(#Arg1,#Arg2)&mod([#Mod]),
                 #Pred, det_prep,[#Prep,unspec_temporal_inclusion_rel,
                 tau(#Mod),coordination_sem1(#Pred,[tau(#Arg1)&
                 pred_sort(#NPSort),tau(#Arg2)])]) =>
                 tau_pred(unspec_temporal_inclusion_rel,#Prep,#NPSort).
```

3. boolean conjunction - this class consists of syntactically binary conjunctions – *entweder-oder*, *weder-noch* – (ex.(2.68)) which describe a relation refering to the boolean functions: condtional, disjunction, negation.

   (5.69) MMM4_1_08:
   ja und FALLS unsere Besprechung länger geht können wir das Mittagessen auch noch gemeinsam einnehmen.

   The lexical entry in the semantic lexicon consisted of PredName and a feature BOOL_OP_VAL (ex.(2.69)), which was assigned the values: *if_then, or, not.*

   (5.70) `lex(lexeme) => bool_sem(Pred,bool_op_val)`

   The items belonging to the boolean conjunction semantic subclass count five:

   wenn, entweder, falls, sobald, weder

The transfer of boolean conjunctions is provided by one lexical TAU rule (ex.(2.70)):

(5.71)
```
tau_lex(SourcePredName, TargetPredName, [tau(#Arg1),tau(#Arg2)],[],
                        subord_syn(#Arg1,#Arg2)) => sem_t.
```

Four "discrel" adverbs and one "boolean" conjunction occur in the Testsuite of Dialogues for MDS:

aber, oder, also, deshalb, wenn .

5.18.2 Transfer PredNames of *Discrel* in MDS

As the class of *discrel* was transfered with lexical TAU rules, we provided TargetPredNames for each item of the word list following the strategy adopted in the Transfer module (see section 2.9.2., 2.9.3.). The transfer PredNames of *discrel*, coordinations, and boolean conjunctions are listed informally in the following tables (see table 5.8, and table 5.9).

SourcePred	→	TargetPred	SourcePred	→	TargetPred
aber	→	but	infolgedessen	→	consequently
als	→	when	jedoch	→	however
als_auch	→	both	nämlich	→	namely
also	→	so2	und2	→	and2
beziehungsweise	→	that_is	und_so_weiter	→	and_so_on
dagegen2	→	on_the_other_hand	ob	→	whether
daher2	→	therefore	obwohl	→	although
damit2	→	in_order_to	oder2	→	or2
darum2	→	because_of_it	sodass	→	so_that
das_heißt	→	that_is	somit	→	thus
demzufolge	→	consequently	somit	→	consequently1
denn	→	because	somit	→	hence
denn	→	since1	sondern	→	but
deshalb	→	therefore	soweit	→	as_far_as
deshalb	→	so1	sowie	→	as1
deswegen	→	therefore	sowie	→	just_as
deswegen	→	so1	sowohl	→	as_well_as
folglich	→	thus	trotzdem	→	nevertheless
folglich	→	consequently	weil	→	because
folglich	→	therefore	zumal	→	particularly1
indem	→	however	zwar	→	especially1

Table 5.8: Transfer PredNames of *discrel* and *coordination*

SourcePred	→	TargetPred
entweder_oder	→	either
falls	→	if
sobald	→	as_soon_as
weder_noch	→	neither
wenn	→	if

Table 5.9: Transfer PredNames of *boolean conjunctions*

5.19 Adjectives

This section describes the treatment of adjectives in the Transfer module of MDS. *Adjectives* are semantically modifiers refering to different properties of the modified entities. *Adjectives* are a morphologic class which syntactically applies to nouns, forms a sentence constituent with them and is congruent in number and gender with them. Some adjectives can be applied to verbs and form a sentence constituent with them. Their syntactic function in such cases is adverbial.

5.19.1 Semantic Construction of Adjectives in MDS

The adjectives were devided in several subgroups by the Semantic Construction module of MDS, according to their morphologic and semantic properties: *dimensional, relational, rigid, comparision degrees (comparative and superlative), negative.* They are all designed as introducing *conditions* in the semantic represntation of the sentences they occur in. We are going to discuss the subgroups consiquently in this subsection.

1. Dimensional adjectives

   Adjectives which can be used as noun modifiers (adjectives proper)(ex.(2.71a)), and as verb modifiers (adverbs)(ex.(2.71b)) are defined as *dimensional.*

   (5.72) (a) MPS1_1_18: kein SCHLECHTER(adj) Vorschlag.
   (b) DE008/VM009: das paßt echt SCHLECHT(adv) bei mir.

   They introduce a DIMENSION CONDITION in the semantic structure of the analyzed sentences. The information whether one dimensional adjective forms a constituent with a verb or a noun in the sentence is provided by the syntactic analysis, and the Semantic Construction module builds the semantic representation on the basis of this output (ex.(2.72a-b)).

   (5.73) (a) kein SCHLECHTER(adj) Vorschlag.

$$\begin{bmatrix} DRS \\ \begin{bmatrix} quant\_condition \\ qtr : no \\ restr : \begin{bmatrix} drs\_t \\ conds : \begin{bmatrix} basic\_condition \\ pred : vorschlag \\ inst : \boxed{1} \end{bmatrix} \begin{bmatrix} dimen\_condition \\ dimen\_inst : \boxed{1} \\ dimen\_pred : schlecht \\ dimension : quality \end{bmatrix} \end{bmatrix} \\ var : \boxed{1} \end{bmatrix} \end{bmatrix}$$

(b) das paßt echt SCHLECHT(adv) bei mir.

$$\begin{bmatrix} \mathit{DRS} \\ \begin{bmatrix} \mathit{drs\_t} \\ \mathit{conds}: \begin{bmatrix} \mathit{basic\_condition} \\ \mathit{pred}: \mathit{passen} \\ \mathit{inst}: \boxed{1} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \mathit{drs\_t} \\ \mathit{conds}: \begin{bmatrix} \mathit{dimen\_condition} \\ \mathit{dimen\_inst}: \boxed{1} \\ \mathit{dime\_intensity}: \boxed{1} \\ \mathit{dimen\_pred}: \mathit{schlecht} \\ \mathit{dimension}: \mathit{quality} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Dimensional adjectives are semantically interpreted as refering to types of properties which were conceived in the Semantic Construction module of MDS as dimensions. A feature DIMENSION was introduced in the semantic lexicon with the values: *quality, quantity, mood_dim, onto_type, complexity, timespan, timeloc, volume, modality, sympathy, age* to make explicit the semantic dimension in which an adjective modifies the entity to which it applies (see the feature *dimension* in ex.(2.72a-b). An exhaustive list of the defined dimensional values with examples is given in table 5.10.

DimAdj	→	DimVal
angenehm	→	quality
belegt	→	quantity
betrüblich	→	mood_dim
botanischen	→	onto_type
einfach	→	complexity
früh	→	timeloc
kurzfristig	→	timespan
laut	→	volume
nett	→	sympathy
neu	→	age
offensichtlich	→	modality

Table 5.10: Examples of *dimensional adjectives* with their *dimensional values.* (DimAdj stays for *dimensional adjective*, and DimVal stays for *dimansional value*)

The lexical entry in the semantic lexicon of MDS consists of a *lexeme*, the *semantic subgroup*, a *PredName* (refered to with *Value*), and a *Dimensional value*(refered to with *Dimension*) (ex.(2.73)).

(5.74) `lex(lexeme) => ad_dimen_sem(Value,Dimension).`

Dimensional adjectives in the word list issued by the University of Bielefeld count fifty one:

angenehm, ausgeschlossen, ausgezeichnet, belegt, betrüblich, bezüglich, botanischen, ehrlich, einfach, einverstanden, entschieden, ernst, fest, frei, früh, ganz, gebraucht, genau, gern, gerne, geschickt, gewiss, gut, hervorragend, knapp, *kurz, kurzfristig, lange, langfristig, laut, lieber, metaphysischen, mittelfristig, möglich, nett, neu, offensichtlich, okay, prima, recht, richtig, ruhig, schlecht, schön, spät, ungern, ungeschickt, ungünstig, unmöglich, voll, wunderbar.*

Thirteen of them occur in the testsuite of dialogues for MDS:

hervorragend, recht, richtig, früh, schlecht, voll, ganz, lieber, spät, knapp, möglich, ungünstig, betrüblich.

2. Relational adjectives

   Adjectives and pronouns which can connect the discourse referents or conditions they actually apply to with previously introduced discourse referents or conditions issuing the semantic effect to induce a relation of particular order between them are defined as *relational* in the Semantic Construction module of MDS[28] Thus, the sentence in ex.(2.74a) induces the existance of at least one previously made appointment (*Termin*) between the participants of the dialogue, because of the pronoun *anderen*(another), and the sentence in ex.(2.74b) refers to the relation between the temporal location of the conversation and the temporal location of the scheduled appointment in the dialogue item, because of the adjective *nächste*(next), modifying *Termin*.

(5.75) (a) Wir machen einen ANDEREN Termin aus.
(b) Wir treffen uns NÄCHSTE Woche.

The relational adjectives introduce a BASIC CONDITION in the semantic representation of MDS (ex.(2.75)). An interlingua approach is applicable for the interpretation and for the translation of this group of adjectives.

(5.76) Wir machen einen ANDEREN Termin aus.

$$\begin{bmatrix} DRS \\ \begin{bmatrix} drs\_t \\ conds : \begin{bmatrix} basic\_condition \\ pred : termin \\ inst : \boxed{1} \end{bmatrix} \begin{bmatrix} basic\_condition \\ pred : andere \\ inst : \boxed{1} \end{bmatrix} \begin{bmatrix} basic\_condition \\ pred : ausmachen \\ inst : \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The semantic lexical entry of relational adjectives consists of a *lexeme*, the *semantic subclass*, and a *PredName* (ex.(2.76)).

(5.77) `lex(lexeme) => rel_adj_sem(Pred).`

Relational adjectives in the word list issued by the University of Bielefeld count ten:

andere, darauffolgende, folgende, kommende, letzte, nächste, übernächste, vorhergehende, vorherige, vorige.

Two of them occur in the testsuite of dialogues for MDS: *andere, nächste.*

[28]For more details see Eberle (1995).

3. Rigid adjectives

   The word class normally refered to with *cardinal* and *ordinal* NUMERALS was defined in the semantic class of *rigid adjectives* in the Semantic Construction of MDS. These adjectives modify nouns by adding a RIGID CONDITION to their referential argument (McGlashan (1994)).

   The semantic lexical entry of the rigid adjectives consists of a type of the rigid predicate (cardinal - *card*, or ordinal - *ord*), and of a rigid designator (ex.(2.77)).

   ```
   (5.78) lex(lexeme) => ad_rigid_sem(RigidPred,RigidDes).
   ```

   The real number refered to by the rigid predicate (1,2,3,...n) is made explicit through the rigid designator (ex.(2.78a-b)).

   ```
   (5.79) (a) lex(erste) => ad_rigid_sem(ord,1).
          (b) lex(eins)  => ad_rigid_sem(card,1).
   ```

4. Comparision Degrees

   Adjectives in *comparative* and in *superlative* degree are assigned separate semantic classes in the Semantic Construction module of MDS. The lemmata in comparative degrees were listed as lexical entries in the semantic lexicon (ex.(2.79)).

   ```
   (5.80) lex(lexeme_in_comparative_degree) => COMPARATIVE_DEGREE_ad_sem
                     (PredName_of_adjective_in_positive_degree,Dim,Op).
   ```

   The comparative degrees of adjectives trigger a complexe structure in the Semantic Construction module of MDS. The interpretation of the degrees is represented apart from the interpretation of the adjective in positive degree, which carries out the role of modifier in the entire representation of the analyzed sentence. We discuss the structures conveyed by the presence of a comparative or a superlative degree of an adjective in a sentence in the following two items.

   - Comparative Degree

     The comparative degree of adjectives triggers a complexe structure in the Semantic Construction of MDS. It introduces an ALFA CONDITION. The later is satisfied by a QUANTIFICATION CONDITION, consisting of the semantic representation of the adjective in positive degree, and the interpretation of the comparative itself by means of *comp_ condition*(ex.(2.80)). The structure in ex.(2.80) also illustrates how the distinct elements of this representation are unified with each other.

(5.81) Wir treffen uns SPÄTER.

$$\begin{bmatrix} \mathit{DRS} \\ \begin{bmatrix} \mathit{drs\_t} \\ \mathit{conds}: \\ \begin{bmatrix} \mathit{basic\_condition} \\ \mathit{pred}: \mathit{treffen} \\ \mathit{inst}: \boxed{1} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \mathit{dimen\_condition} \\ \mathit{dimen\_inst}: \boxed{1} \\ \mathit{dime\_value}: \boxed{3} \\ \mathit{dimen\_pred}: \textit{spät} \\ \mathit{dimension}: \mathit{timeloc} \end{bmatrix} \begin{bmatrix} \mathit{alfa\_conition} \\ \mathit{alfa\_type}: \mathit{comp} \end{bmatrix} \begin{bmatrix} \mathit{quant\_condition} \\ \mathit{qtr}: \mathit{every} \\ \mathit{conds}: \begin{bmatrix} \mathit{dimen\_condition} \\ \mathit{dime\_value}: \boxed{5} \\ \mathit{dimen\_pred}: \textit{spät} \\ \mathit{dimension}: \mathit{timeloc} \end{bmatrix} \\ \begin{bmatrix} \mathit{drs\_t} \\ \mathit{conds}: \begin{bmatrix} \mathit{comp\_condition} \\ \mathit{comp\_arg1}: \boxed{3} \\ \mathit{comp\_arg2}: \boxed{5} \\ \mathit{comp\_inst}: \boxed{11} \\ \mathit{comp\_op}: \mathit{greaterthan} \end{bmatrix} \end{bmatrix} \\ \mathit{var}: \boxed{5} \end{bmatrix} \end{bmatrix}$$

Eight adjectives in comparative degree occur in the testsuite of dialogues for MDS:

angenehmer, besser, eher, früher, lieber, mehr, schöneres, später.

- Superlative Degree

  The superlative degree of adjectives also triggers a complexe structure in the Semantic Construction of MDS. The superlative degree is interpreted in the combination of a *negative condition*, and the included in its conditions *comp_condition* (see ex.(2.81)). The way the unification between the distinct elements of the semantic structure takes place can be seen in ex.(2.81).

  (5.82) Wir machen den besten Termin aus.

$$\begin{bmatrix} \mathit{DRS} \\ \begin{bmatrix} \mathit{drs\_t} \\ \mathit{conds}: \\ \begin{bmatrix} \mathit{basic\_condition} \\ \mathit{pred}: \mathit{termin} \\ \mathit{inst}: \boxed{1} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \mathit{dimen\_condition} \\ \mathit{dimen\_inst}: \boxed{1} \\ \mathit{dime\_value}: \boxed{3} \\ \mathit{dimen\_pred}: \mathit{gut} \\ \mathit{dimension}: \mathit{quality} \end{bmatrix} \begin{bmatrix} \mathit{neg\_condition} \\ \mathit{neg\_arg}: \\ \begin{bmatrix} \mathit{drs\_t} \\ \mathit{conds}: \\ \begin{bmatrix} \mathit{dimen\_condition} \\ \mathit{dime\_value}: \boxed{5} \\ \mathit{dimen\_pred}: \mathit{gut} \\ \mathit{dimension}: \mathit{quality} \end{bmatrix} \begin{bmatrix} \mathit{comp\_condition} \\ \mathit{comp\_arg1}: \boxed{5} \\ \mathit{comp\_arg2}: \boxed{3} \\ \mathit{comp\_inst}: \boxed{11} \\ \mathit{comp\_op}: \mathit{greaterthan} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

  Two adjectives in superlative degree occur in the testsuite of dialogues for MDS: *besten, frühesten.*

5. Negative adjective

   The negative article (*kein*) in German is defined as *negative adjective* in the Semantic Constuction module of MDS. The lexical entry of *kein* in the semantic lexicon of MDS consists of the *lexeme*, the semantic *subgroup of generalized quantifiers*, and the *PredName*(ex.(2.82)).

   (5.83) `lex(kein) => gen_qtr_sem(no).`

   It is designed as introducing a NEGATIVE CONDITION in the semantic structure of the sentence described (ex.(2.83)), and is unified with the noun it modifies.

(5.84) Ich habe KEINE Zeit.

$$\begin{bmatrix} DRS \\ \begin{bmatrix} neg\_condition \\ qtr : no : \\ \begin{bmatrix} restr : \begin{bmatrix} drs\_t \\ conds : \begin{bmatrix} basic\_condition \\ pred : zeit \\ inst : \boxed{1} \end{bmatrix} \end{bmatrix} \\ scope : \begin{bmatrix} drs\_t \\ conds : \begin{bmatrix} basic\_condition \\ pred : haben1 \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ var : \boxed{1} \end{bmatrix} \end{bmatrix}$$

5.19.2 Transfer Rules for Adjectives

The structures issued by the Transfer module are parallel to the representations of the Semantic Construction module.

1. *Dimensional adjectives.* The dimensional adjectives are transfered with two lexical TAU rules: one for the adjectival interpretation (ex.(2.84a)), and one for the adverbial interpretation (ex.(2.84b)). The information of the type of syntactic constituent which they belong to determines the selection of the right transfer rule. This is made explicit by the inclusion of argument types into the TAU – ad_syn(#NP) and adv_syn(#VP) (ex.(2.84a-b)).

   (5.85)
   ```
   (a) tau_lex(einfach,simple,[tau(#NP,#Args)],#Args,ad_syn(#NP)) => sem_t.
   (b) tau_lex(einfach,simply,[tau(#VP,#Args)],#Args,adv_syn(#VP)) => sem_t.
   ```

2. *Relational adjectives.* The relational adjectives are transfered with one lexical TAU rule (ex.(2.85)), as they do not occur as adverbs in sentences. The syntactic constituent which they belong to is also made explicit in the structure of the rule (see adj_syn(#NP) in ex.(2.85)).

   (5.86)
   ```
   tau_lex(andere,other,[tau(#NP)],[],adj_syn(#NP)) => sem_t.
   ```

3. *Rigid adjectives.* The rigid adjectives are transfered with two lexical TAU rules (ex.(2.86a-b)). The first argument of each of them refers to the relevant type of rigid predicate – *card*, or *ord*. A reduction of these two rules into one is technically possible if the type of rigid predicate will be substituted with a predicate *ad_rigid* (see Martin Emele 94, tau_ad.stuf), summarizing the common features of the lexical predicates *card* and *ord*.

   (5.87)
   ```
   (a) tau_lex(card,ad_rigid,#SemArgs,#Args,#SourceSign) =>
           tau_ad_rigid(card,#SemArgs,#Args,#SourceSign).
   (b) tau_lex(ord, ad_rigid,#SemArgs,#Args,#SourceSign) =>
           tau_ad_rigid(ord, #SemArgs,#Args,#SourceSign).
   ```

4. *Comparision degrees.* The comparision degrees of adjectives do not have special transfer rules. The adjective in positive degree is translated by means

of a lexical TAU rule (ex.(2.87)), and the semantic structure of the relevant comparision degree (see ex.(2.80),(2.81)) is parallelilly built in the Transfer module.

(5.88) `tau_lex(gut,good,[tau(#NP,#Args)],#Args,adj_syn(#NP)) => sem_t.`

The German PredNames of the translated adjective are substituted with the corresponding English ones (see ex.(2.88) and compare with ex.(2.81)).

(5.89) We will arrange the BEST appointment.

$$\begin{bmatrix} \mathit{DRS} \\ \begin{bmatrix} \mathit{drs\_t} \\ \mathit{conds}: \\ \begin{bmatrix} \mathit{basic\_condition} \\ \mathit{pred}: \mathit{appointment} \\ \mathit{inst}: \boxed{1} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \mathit{dimen\_condition} \\ \mathit{dimen\_inst}: \boxed{1} \\ \mathit{dime\_value}: \boxed{3} \\ \mathit{dimen\_pred}: \mathit{good} \\ \mathit{dimension}: \mathit{quality} \end{bmatrix} \begin{bmatrix} \mathit{neg\_condition} \\ \mathit{neg\_arg}: \\ \begin{bmatrix} \mathit{drs\_t} \\ \mathit{conds}: \\ \begin{bmatrix} \mathit{dimen\_condition} \\ \mathit{dime\_value}: \boxed{5} \\ \mathit{dimen\_pred}: \mathit{good} \\ \mathit{dimension}: \mathit{quality} \end{bmatrix} \begin{bmatrix} \mathit{comp\_condition} \\ \mathit{comp\_arg1}: \boxed{5} \\ \mathit{comp\_arg2}: \boxed{3} \\ \mathit{comp\_inst}: \boxed{11} \\ \mathit{comp\_op}: \mathit{greaterthan} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

5. *Negative adjective.* A special lexical TAU rule translates a negative predicate into a negative predicate, making explicit the syntactic constituent of the negation - ad_syn(#Arg) (see ex.(2.89)).

   (5.90) `tau_lex(neg_ad, neg_ad, [tau(#Arg)], [], ad_syn(#Arg)) => sem_t.`

5.19.3 Structural Transfer

The German sentence in ex.(2.90a) cannot be literally translated into English, as shown in ex.(2.90b).

(5.91) (a) DE008/VM009: das paßt echt schlecht bei mir.
(b) *It does not suit me at all.*

Among the other peculiarities of this sentence with respect to the semantic representation and transfer[29], the German positive sentence containing a modified (or intensified) modifier (the dimensional adjective *schlecht*) – *echt schlecht* is translated into a negative sentence and a negation intensifier[30] – *at all*, classified as focus adverb. Thus, two different semantic representations were required for the German sentence (see ex.(2.91), a repeated ex.(2.71)) and for the derived English sentence (ex.(2.92)).

(5.92) das paßt echt schlecht bei mir.

$$\begin{bmatrix} \mathit{DRS} \\ \begin{bmatrix} \mathit{drs\_t} & \\ \mathit{conds}: & \begin{bmatrix} \mathit{basic\_condition} \\ \mathit{pred}: \mathit{passen} \\ \mathit{inst}: \boxed{1} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \mathit{drs\_t} & \\ \mathit{conds}: & \begin{bmatrix} \mathit{dimen\_condition} \\ \mathit{dimen\_inst}: \boxed{1} \\ \mathit{dime\_intensity}: \boxed{1} \\ \mathit{dimen\_pred}: \mathit{schlecht} \\ \mathit{dimension}: \mathit{quality} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

[29]The chalanges these two sentences offer to the Semantic Construction and the Transfer modules have been brightly discussed in the paper "Semantik-orientierter rekursiver Transfer in HPSG am Beispiel des Referenzdialogs" Dorna et al. (1994).

[30]This term intuitively corresponds to the semantic role of *at all* in the scope of negation.

(5.93) It does not suit me at all.

$$\begin{bmatrix} DRS \\ \begin{bmatrix} neg\_condition \\ neg\_arg : \begin{bmatrix} drs\_t \\ conds : \\ \begin{bmatrix} phi\_cindition \\ phi\_arg : \begin{bmatrix} drs\_t \\ \begin{bmatrix} basic\_condition \\ pred : suit \\ inst : \boxed{1} \end{bmatrix} \end{bmatrix} \\ phi\_op : grad \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

A special transfer rule was created[31] to make this structural transfer possible. It is shown in ex.(2.93).

(5.94)
```
tau_lex(schlecht,negation,[focus_adv_sem1(at_all,grad,[tau(#VP,#Args)])],
        [at_all|#Args],adv_syn(#VP&lex_pred(passen;passen1;gehen3;gehen2)))
        => sem_t.
```

5.19.4 Transfer PredNames of Adjectives

The transfer of adjectives in MDS was basically performed by means of lexical TAU rules. Thus, the determination of English PredNames, corresponding to the German PredNames of the word list, was pursued after the principles desribed in section 2.9.3., and 2.9.4. The transfer PredNames of the two groups of dimensional and relational adjectives are given in table 5.11, and in table 5.12 respectively. The other groups of adjectives we discussed so far did not need proper transfer PredNames, as they were transfered by means of structural representations, as discussed in Section 2.9.18.3.

5.19.5 Future Work

A fool word form morphologic dictionary was used instead of a morphologic lemmatizer. Thus, there was a semantic lexical entry for each inflected adjectival form (ex.(2.94a-d)).

(5.95)
```
(a) lex(frueh) => ad_dimen_sem(frueh,timeloc).
(b) lex(fruehe) => ad_dimen_sem(frueh,timeloc).
(c) lex(fruehen) => ad_dimen_sem(frueh,timeloc).
(d) lex(frueher) => ad_dimen_sem(frueh,timeloc).
```

It would be good to think of efficient integrating of the morphologic analyser into the system. Furthermore, some lexical entries occur twice in the semantic lexicon, as they are classified in different semantic subclasses. Consequent matching of entries show:

[31] The analysis of the German input sentences and the English output sentence, the structural conception of the transfer rule and the creation of this rule are due to Michael Schiehlen.

SourcePred	→	TargetPred
andere	→	other
darauffolgende	→	following
folgende	→	following
kommende	→	next
letzte	→	last
naechste	→	following
naechste	→	next
uebernaechste	→	after_next
vorhergehende	→	preceding
vorherige	→	previous
vorige	→	last

Table 5.11: Transfer PredNames of *relational adjectives*

1. the dimensional adjectives with the standard adverbs (ex.(2.95a-b)).

   (5.96) (a) `lex(offensichtlich) => ad_dimen_sem(offensichtlich1,modality).`
   (b) `lex(offensichtlich) => ad_sem(offensichtlich2).`

2. the comparision degrees of dimensional adjectives with dimension value *time-loc* with focus adverbs with focus operator *grad* or *temp* (ex.(2.96a-b)).

   (5.97) (a) `lex(spaet) => ad_dimen_sem(spaet,timeloc).`
   (b) `lex(spaetestens) => focus_adv_sem(spaetestens,temp).`

3. comparison degrees of dimensional adjectives with intensifiers (ex.(2.97a-b)).

   (5.98) (a) `lex(eher) => comp_ad_sem(bald,timeloc,lessthan).`
   (b) `lex(eher) => intens_sem(eher).`

A consistent classification and semantic representation of these classes would contribute for the efficient transfer in one future system.

SourcePred	→	TargetPred	SourcePred	→	TargetPred
ausgeschlossen	→	out_of_question	laut1	→	loud
ausgezeichnet	→	perfectly	lieb1	→	dear
belegt	→	taken	lieb2	→	rather2
belegt	→	booked	metaphysisch	→	metaphysical
betrueblich	→	sad	moeglich	→	possible
botanisch	→	botanic	nett	→	nice
ehrlich	→	honest	neu	→	new
ehrlich	→	honestly	offensichtlich1	→	obviously
einfach	→	easy	okay1	→	okay
einfach	→	just2	prima1	→	fine
einfach	→	simple	recht	→	right2
einfach	→	simply	recht	→	all_right
einverstanden	→	agree	richtig	→	correctly
entschieden	→	definitely	richtig	→	right2
erledigt	→	settled	ruhig	→	null_pred
ernst	→	serious	ruhig	→	feel_free_to_do
ernst	→	seriously	schlecht	→	negation
fest1	→	fixed	schlecht	→	not_well
frei	→	free	schlecht	→	bad
frueh	→	early	schlecht	→	difficult
ganz1	→	whole	schlecht	→	inconvenient
gebraucht	→	used	schoen	→	nice
genau	→	exactly	spaet	→	late
gern	→	with_pleasure	ungern	→	unwillingly
geschickt	→	suit	ungeschickt	→	inconvenient
gewiss	→	certain	unguenstig	→	inconvenient
gut	→	good	unguenstig	→	unfavourably
gut	→	well	unmoeglich	→	impossible
hervorragend	→	great	viel1	→	much
in_ordnung	→	all_right	voll	→	booked
knapp	→	tight	voll	→	fully_booked
kurz	→	short	wunderbar	→	great
kurzfristig	→	short_term	langfristig	→	long_term
lang	→	long	angenehm	→	all_right

Table 5.12: Transfer PredNames of *dimensional adjectives*.

Chapter 6

The Transfer of Particles and Idiomatic Expressions

6.1 Greetings

A kind of pragmatic controlled transfer is implemented for greetings, goodbyes, and thank. These are conventional dialogue acts thus they are mapped onto a concept and only that concept is transferred to Generation.

6.2 Idioms, Exclamations, Particles

6.3 Idioms

There is no common treatment for idioms nor an agreed on list of expressions classified as idiomatic. For the demonstrator some expressions like *außer hause, ich weißnicht so recht, wie war das nochmal, etc* aretreated by syntax as one phrase which are directly translated into corresponding English phrases.

6.4 Exclamations

In the demonstrator expressions like *ja, also, prima etc.* occurring exclusively at the beginning of an utterance are classified as exclamations. They have no semantics and there are no transfer rules for them. A list with possible translations was given to the Generation module which was then responsible to decide whether to verbalize them or not.

6.5 Particles

Also for particles there was no common general treatment in the demonstrator. Discourse particles at the beginning of an utterance were treated as exclamations. Intensifiers like *ja, ruhig* occuring in the middle of an utterance are simply deleted, because there is no lexical equivalent for them in English. The intended meaning has to be expressed by other means like intonation. But the formalism used in the MDS does not allow to formulate constraints about prosodic information. Some discourse particles were classified in the semantic group of adverbs and were treated as adverbs in the transfer module (cf. 5).

Chapter 7

The Transfer of Prepositions

Prepositions are highly ambiguous. The meanings one might distinguish within the source language are often not specific enough in order to determine their appropriate translation into the target language. Considering prepositional meanings and their lexicalization in different languages, we are faced with extreme interferences. Hence, their translation represents one of the difficulties of MT, cf. Hutchins and Somers (1992).

7.1 General Remarks

The translation of prepositions differs significantly with respect to the semantic structure they form part of, cf. (7.1) - (7.3):

(7.1) Die Sekretärin wartet *auf* Herrn Brown.
The secretary is waiting *for* Mr. Brown.
(7.2) Die Sekretärin wartet *auf* dem Platz.
The secretary is waiting *on* the square.
(7.3) Die Sekretärin ist *auf* dem Platz.
The secretary is *on* the square.

If they are used as the head of an argument, prepositions often do not have a meaning of their own. This influences their translation. It depends above all on the translation of their head predicate. In (7.1), for example, *auf* is subcategorized by the verb *warten*. Its translation equivalent *wait* selects the TL preposition *for* that does not stand in a regular translation correspondence to *auf*, cf. section 4.2. On the other hand, prepositions that show up in modifiers are meaningful. Their translation is mostly predictable. It can be determined by the kind of entity designated by their internal argument. In (7.2), where the *auf*-PP modifies the waiting situation, and in (7.3), where it localizes the secretary, the TL preposition *on* corresponds systematically to *auf*, describing a spatial relation.

The sketched behavior has implications for the treatment of prepositions in the transfer component. Prepositions occurring in argument PPs have to be transferred together with their head predicate, cf. section 4.2. Those occurring in modifiers can be translated separately.

The semantic representation of prepositions that is passed to the transfer module supports this distinction. Based on the davidsonian analysis, the λ-DRT formalism interprets PPs which contain prepositions that are used idiosyncratically with verbs, nouns, or adjectives as arguments. PPs occurring as predicatives, adjuncts and directionals are semantically analysed as modifiers. Thus, prepositions with no meaning, whose translation is rather idiosyncratic, are distinguished from meaningful ones, which are translated systematically. In the following, we concentrate on the latter that can be approached by an independent mapping device.[1]

7.2 Motivation for the Concept-Based Approach

For the translation of prepositions that occur in prepositional modifiers we put forward a concept-based method. Let us motivate the chosen approach by some examples:

(7.4) *Im* Mai bin ich *im* Urlaub.
I'll be *on* vacation *in* May.
→ *unspec_temporal_inclusion_rel* and *temporal_spatial_inclusion_rel*

(7.5) *Am* Montag bin ich *an* der Universität.
On Monday I'll be *at* the university.
→ *unspec_temporal_inclusion_rel* and *institutional_inclusion_rel*

(7.6) *Am* Ende der Woche bin ich *auf* einer Tagung.
At the end of the week I'll be *at* a conference.
→ *unspec_temporal_inclusion_rel* and *temporal_spatial_inclusion_rel*

(7.7) *Am* Abend bin ich *in* einer Vorlesung.
In the evening I'll be *at* a lecture.
→ *unspec_temporal_inclusion_rel* and *temporal_spatial_inclusion_rel*

(7.8) *In* dieser Zeit war ich *auf* der Universität.
At this time I've been *at* the university.
→ *unspec_temporal_inclusion_rel* and *institutional_inclusion_rel*

In (7.4) - (7.8) the German prepositions *in*, *an* and *auf* exhibit different meanings. They are used to express a temporal (*unspec_temporal_inclusion_rel*) and a temporal-spatial localization (*temporal_spatial_inclusion_rel*) of an event as well as its localization with respect to an institution (*institutional_inclusion_rel*). The translation of these prepositions differs with regard to the various interpretations, but it is also distinct within one interpretation. Figure 1 illustrates the distribution of the German and English prepositions with respect to the expressed meaning.[2]

[1]For detailed discussion of the treatment of prepositions in the VERBMOBIL demonstrator, see Buschbeck-Wolf and Nübel (1995).

[2]For the sake of consistency we use the relation names that are implemented in the MDS: *inclusion_rel* means that the individual modified by the PP is localized or included with respect to the space, the time, etc., the internal argument refers to.

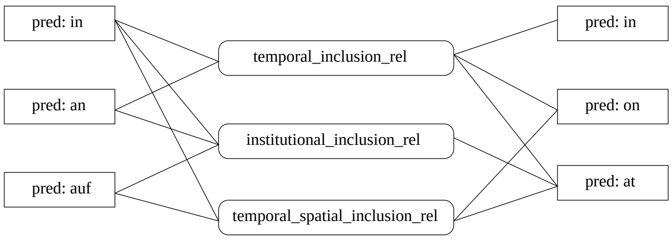


Figure 1: Bilingual conceptual relations and their language specific lexicalization

Each of the German prepositions *in*, *an* and *auf* expresses different relations, i.e. they are polysemous. On the other hand, they share some meanings, i.e. they are partially synonymous. This holds for their English correspondences *in*, *on* and *at* as well. The various interpretations displayed by these prepositions within one language affect their translation, too:

1. If a SL preposition is interpreted differently its target language correspondences might also be distinct:

   (7.9a) Mr. Brown kommt *im* Mai.
   Mr. Brown is coming *in* May.

   (7.9b) Mr. Brown ist *im* Urlaub.
   Mr. Brown is *on* vacation.

   (7.9c) Mr. Brown ist *in* der Universität.
   Mr. Brown is *at* the university.

2. The distinct meanings captured by a SL preposition might be covered by a TL correspondence that exhibits the same kind of ambiguity:

   (7.10a) Mr. Brown ist *in* der Vorlesung.
   Mr. Brown is *at* the lecture.

   (7.10b) Mr. Brown ist *in* der Schule.
   Mr. Brown is *at* school.

   (7.10c) Mr. Brown kam *in* dieser Zeit .
   Mr. Brown came *at* this time.

3. Different SL prepositions share one meaning that might be captured by one TL preposition, since the TL does not display the same distribution of synonymical relations:

   (7.11a) Mr. Brown war *an* der Universität.
   Mr. Brown was *at* the university.

   (7.11b) Mr. Brown war *in* der Universität.
   Mr. Brown was *at* the university.

   (7.11c) Mr. Brown war *auf* der Universität.
   Mr. Brown was *at* the university.

These observations lead us to approach the translation of prepositions by a concept-based method in favour of direct transfer mappings that would turn out to be highly specific for the prepositions considered. The introduction of concepts that cover the denoted meanings allow to link SL and TL prepositions systematically. This approach accounts for the meaning distribution over the prepositions in the languages involved.

There are some further advantages of this approach:

1. By the assignment of concepts, the contextually relevant meaning of a preposition is conserved and thus can be used for further disambiguation. In contrast, if transfer rules were applied the identified interpretation would not be explicitly anchored.

2. The mapping to concepts reflects various meaning relations. The mapping of one predicate to different concepts makes its ambiguity explicit and the mapping of different predicates to one concept accounts for their synonymy.

3. The identification of conceptual relations that are shared by two languages gives theoretically interesting insights in the specification and hierarchical organization of bilingual concepts.

The concept-based approach we adopt is based on a two-step mapping. As a first step, SL prepositional predicates are mapped onto bilingual concepts by the application of *refinement rules*. As a second step, the appropriate TL prepositional predicates are derived by *lexicalization rules*. The mapping to the denoted relation requires the disambiguation of the SL preposition that is achieved by several kinds of selectional restrictions on the preposition's arguments, see below. The derivation of the appropriate TL correspondence has to obey particular lexicalization constraints.

7.3 Refinement and Lexicalization Rules

Refinement rules are predicate-to-concept mapping rules. They specify semantically underspecified prepositional relations, such as *mit_rel* or *durch_rel*, with respect to their contextually relevant interpretations, i.e. they assign the appropriate bilingual concepts. For prepositions that exhibit case alternation, directional relations are distinguished from non-directional ones already in the semantic analysis. The underspecified relation *an_rel*, for example, is intersected with *goal_rel* if the case of the preposition's internal argument is accusative, otherwise it is intersected with the negation of *goal_rel*.

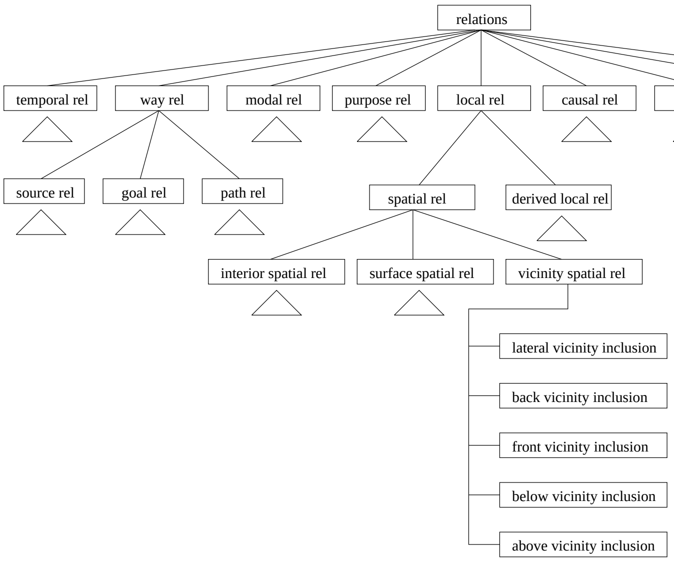


Figure 2: Parts of the relations' hierarchy

All bilingual relations are hierarchically organized in the relational part of the sort hierarchy. Figure 2 displays a part of it.
On the one hand, the high number of distinctions between the conceptual relations is motivated by the bilingual situation. For a straightforward lexicalization, a strong partition between the transfer relevant readings is necessary. On the other hand, it can be explained by theoretical considerations. We have considered most of the systematically corresponding pairs of prepositional relations (e.g. pairs of static/directional interpretations) that are subtypes of different supertypes and introduced conceptual relations also for unambiguous prepositions that have only one TL correspondence[3]

The range of the conceptual interpretations an underspecified prepositional relation might have is defined disjunctively, cf. (7.12).[4]

(7.12)
```
mit_rel = general_comitative_rel;instrumental_rel;
          general_concomitant_rel;modality_mood_rel.
```

Underspecified prepositional relations are mainly specialized by the use of sortal restrictions on their internal and external arguments.[5] The sorts that are used for the disambiguation of prepositions are anchored in the entity part of the STUF hierarchy which defines the inheritance and disjointness between transfer relevant sorts. It is described in more detail in section 8.2.1. The general format of refinement rules is declared in (7.13).

[3]Possibly, some of the readings could be grouped together if such considerations would be ignored.

[4]The code follows the syntax of STUF, cf. Momma et al. (1994).

[5]For other constraints that are introduced in order to identify the interpretation that is shared by a TL, see Buschbeck-Wolf and Nübel (1995).

(7.13) `@tau_prep(pred_name,relation_rel,entitaet_c,entitaet_c) => sign.`

Refinement rules are applied by unification. If a rule of this type unifies with the prepositions' predicate name (*pred_name*) and with the contextually given sorts of the external and internal arguments of the preposition (both of the type *entitaet_c*), it instantiates a bilingual relation of the type *relation_rel*.[6]

In the next step, lexicalization rules are applied. (7.14) shows their general definition. They instantiate the appropriate TL predicates (subtypes of *top*) for the determined bilingual conceptual relation (of the type *relation_rel*) if sortal constraints on the internal argument (type *entitaet_c*) are fulfilled. Other restrictions, e.g. referential properties of the internal argument, are fixed on the rules' RHS. Concept-to-predicate mapping rules are also applied by unification.

(7.14) `@tau_pred(relation_rel,top,entitaet_c) => sem_t.`

The lexicalization rules, thus formulate TL specific constraints for the generation of the appropriate TL preposition from a conceptual relation.

7.4 A Worked Example

In the following, we illustrate the transfer of the preposition *in*. Consider the following examples:

(7.15a) das Treffen *im* Januar ↔ the meeting *in* January
(7.15b) die Vorlesung *in* dieser Woche ↔ the lecture this week

(7.16a) die Eingangshalle *im* Hotel ↔ the entrance hall *in* the hotel
(7.16b) das Büro *im* Erdgeschoss ↔ the office *on* the ground floor

(7.17a) die Studenten *in* der Vorlesung ↔ the students *at* the lecture
(7.17b) *im* Urlaub sein ↔ to be *on* holidays

(7.18) das Treffen *in* der Universität ↔ the meeting *at* the university

(7.19) *in* Schwierigkeiten sein ↔ to be *in* trouble

Examples (7.15) - (7.19) show some of the static interpretations of the German preposition *in*. These are the temporal localization (7.15), the localization in the interior of a location or an object (7.16), the temporal-spatial localization (7.17), the localization with respect to an institution (7.18), and the modal interpretation (7.19). In (7.20) - (7.24) we illustrate how these particular meanings are identified by refinement rules and lexicalized by the corresponding concept-to-predicate mapping rules.

(7.20)

[6] Note that *relation_rel* might be already instantiated with respect to directional and nondirectional interpretations.

```
tau_prep(in,unspec_temporal_inclusion_rel,
            temporal_c,
            time_period_c;saison_c;month_c) => sign.

tau_pred(unspec_temporal_inclusion_rel,in,
         daytime_c;month_c;time_period_c;saison_c)
=> ref_type(non_deictic).

tau_pred(unspec_temporal_inclusion_rel,null_prep,
         time_obj_c) => ref_type(deictic).
```

The temporal interpretation *unspec_temporal_inclusion_rel* is assigned if the external argument denotes a temporally localizable entity and the internal argument is either a time period, a saison, or a month. This bilingual concept that is also assigned to other prepositions, such as *an* or *zu*, is either lexicalized by English *in* (7.15a) or is omitted if the temporal expression is deictically referred to (7.15b).

(7.21)

```
tau_prep(in,spatial_inclusion_rel,
            entity_c,
            three_dim_c; substance_c) => sign.

tau_pred(spatial_inclusion_rel,in,
         three_dim_c & ~floor_c;substance_c) => sem_t.

tau_pred(spatial_inclusion_rel,on,
         floor_c) => sem_t.
```

The underspecified relation *in_rel* is refined to *spatial_inclusion_rel* - the localization in an object's interior -, cf. (7.16), if the internal argument is a three-dimensional object or a substance. This meaning is expressed by English *in*, if the English reference object is also viewed three-dimensionally (7.16a), or by *on* if it is conceptualized two-dimensionally (7.16b).[7]

(7.22)

```
tau_prep(in,temporal_spatial_inclusion_rel,
            situation_c;human_c,
            event_c) => sign.

tau_pred(temporal_spatial_inclusion_rel,at,
         ~holiday_c) => sem_t.

tau_pred(temporal_spatial_inclusion_rel,on,
         holiday_c) => sem_t.
```

[7]Since we wanted to keep the sorts language independent, we circumscribed the dimensionality information by the sort *floor_c* which covers the objects that are conceptualized differently in the domain.

If a situation or a human being is localized with respect to an event, the concept *temporal_spatial_inclusion_rel*, cf. (7.18), is assigned. This relation is lexicalized idiosyncratically by *on* if the English *holiday* or *vacation* show up as internal argument. Otherwise, *at* is the appropriate TL preposition.

(7.23)

```
tau_prep(in,institutional_inclusion_rel,
            human_c;funkt_sit_c,
            institution_c) => sign.

tau_pred(institutional_inclusion_rel,at,
         entity_c) => sem_t.
```

If a human being or a functional situation is localized with respect to an institution the relation *in_rel* is refined to *institutional_inclusion_rel*. From this bilingual concept, only *at* can be derived.

(7.24)

```
tau_prep(in,modality_mood_rel,
            human_c; situation_c,
            mental_property_c) => sign.

tau_pred(modality_mood_rel,in,
         entity_c) => sem_t.
```

If *in* is used to characterize a human being or a situation with respect to a property, the relation *modality_mood* is assigned and lexicalized straightforwardly by English *in*.

The directional interpretations of *in* stand in systematic correspondence to its static relations. Since they can be analyzed using the same sortal restrictions as the latter, we will not outline their translation in detail.[8]

[8]For detailed information about the treatment of other prepositions we refer to the rule documentation available at the IMS of Stuttgart University.

Chapter 8

Semantic Sort Hierarchy in the Verbmobil Demonstrator

8.1 General remarks

To be efficient any NLP system must integrate disambiguation devices. Sortal information on lexical items can be used to reduce the number of syntactic output structures by formulating semantic coocurrence restrictions which eliminate semantic inconsistent syntactic structures. It supports lexical disambiguation, the resolution of structural and referential ambiguities as well as contextual representation and inferencing. Especially MT systems like the speech-to-speech-translation system VERBMOBIL requires sortal information in order to resolve translational ambiguities, which arise when a single source language word can be potentially translated into a number of different target language words or expressions. Let us consider an example for illustration:

(8.1) einen Termin *verlegen* → to *postpone* an appointment
(8.2) ein Buch *verlegen* → to *publish* a book
(8.3) ein Kabel *verlegen* → to *lay* a wire
(8.4) eine Firma *verlegen* → to *tran sfer* a company
(8.5) einen Notizzettel *verlegen* → to *misplace* a note

The translation of *verlegen* differs with respect to the different readings in (8.1) - (8.5), which can be captured by sortal restrictions on its theme argument.[1] The translation of prepositions, cf. section 7, is a further example where sortal constraints are used on a large scale in order to solve transfer ambiguities.[2]

The STUF sort hierarchy we are going to describe here is also part of the FLEX domain model. In contrast to the much broader FLEX domain model, cf. Quantz et al. (1994), the STUF hierarchy represents only sortal and relational information which is frequently used in the semantic and transfer components and thus should

[1] However, the ambiguity between the *misplace*-reading of *verlegen* on the one hand, and the *publish*- and *lay*-readings on the other, requires a deeper analysis, since the difference cannot be fixed by sortal restrictions alone.

[2] For illustration see also Buschbeck-Wolf and Nübel (1995).

not be treated via interfaces in an independent module.[3] Hence, the functionality of the STUF sort hierarchy is rather restricted. It is mainly used to verify selectional restrictions for disambiguation in the semantic and transfer modules. Information necessary for speech event recognition, calender modelling, context interpretation and inferencing, cf. Quantz et al. (1994), is not considered here.
The application of the STUF formalism for disambiguation imposes monotonicity. Selectional restrictions are formulated as hard constraints. In contrast, FLEX allows for soft constraints and default specifications. Nevertheless, at a first step of semantic interpretation we should figure out to which degree we can do without defaults and where the boundaries of monotonic interpretation are.
On the other hand, the STUF implementation of the sort hierarchy turned out to be more appropriate for seveal reasons, cf. Momma et al. (1994):

- the formalism supports the modelling of type hierarchies with multiple inheritance
- the formalism allows type-based inferencing
- subsumption as the main test for selectional restrictions is a built-in device of the unification process, so that
  - semantic selectional restrictions are processed *during* the construction
  - transfer constraints are applied *during* the transfer recursion
- semantic construction, transfer and those parts of semantic evaluation which can be processed monotonically are performed in *one integrated* STUF process.

The STUF type hierarchy represents conceptual information associated with lexical items. It contains both sortal and relational specifications, where sortal information corresponds to concepts of things and situations, and relational information to conceptual relations.[4] It has to be emphasized that a type hierarchy used for disambiguation in a MT system cannot be regarded as a general epistemic ontology, known from the domain of knowledge engeneering. Especially the lower parts of the conceptual categorization differ significantly from traditional ontologies, since they are grouped and fine-tuned with respect to the solution of translational ambiguities. However, taking a closer look at other ontologies, e.g. at the naive semantics ontology in Dahlgren (1988), the LILOG ontology designed for a knowledge based understanding system in Klose et al. (1992) or the KBMT ontology used as interlingua in MT, cf. Nirenburg et al. (1987), it becomes obvious that even the upper-level conceptual categorizations do not coincide. This leads to the assumption that the concrete design always depends on the requirements of the specific application.

[3]In our opinion the semantic and transfer components should take the responsibility for a sort hierarchy which supports the verification of semantic selectional restrictions (on upper-level concepts) and transfer relevant disambiguation (on lower-level concepts), since it is one of their primary sources. For consistency the sortal specification provided by these components has also to be integrated in the general domain model, which has a much broader application.

[4]Regarding the borderlines between linguistic, conceptual and world knowlegde on the one hand, and the content and degree of specifity of conceptual information on the other, we refer to the detailed discussion in Quantz et al. (1994), and share the position that the conceptual specification should be first of all application-oriented.

8.2 The structure of the sort hierarchy

In this subsection we describe the main parts of the hierarchy in more detail and give some explanatory remarks on the assumed sortal and relational distinctions. For the sake of simplicity we name the used sorts in English. The sort hierarchy covers only the domain-relevant meanings of lexical items which occur in the word list of the VERBMOBIL demonstrator. It consists of two main parts:

1. the sortal part with the supertype `entity_c`
2. the relational part with the supertype `relation_rel`

Since the functionality of these two part is slightly different we describe them separately.

8.2.1 The hierarchy of entities

This part of the hierarchy comprises sorts that correspond to nouns and to the situation argument of verbs. Its upper structure is illustrated in Figure 1.

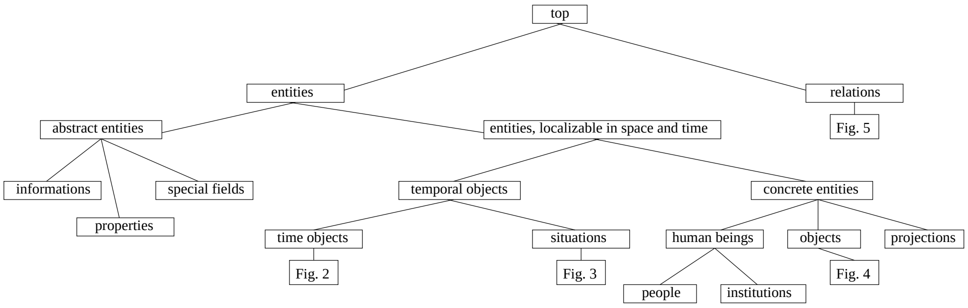


Figure 1: Upper structure of the sort hierarchy

Entities break down into abstract things and entities which are localizable in time and space. The latter are temporal objects, grouped into time objects (Figure 2) and situations (Figure 3), and concrete things which are localizable only in space. Concrete things are separated into objects (Figure 4), human beings and projections.

The entity part of the hierarchy is not designed as an ontology for interlingual MT, as e.g. in the KBMT system Nirenburg et al. (1987), which would require a deep modelling and a strong partition between concepts in order to allow a straightforward lexicalization. It is used to verify selectional restrictions in the semantic and transfer procedures.[5] This task guides the depth of sortal specification. We decided not to go down to the word level. Although one might find a concept for each meaningful lexical item, it turned out to be unnecessary for disambiguation. In most cases the sortal restrictions for the choice of a particular TL correspondence

[5]Examples can be found in Chapter 2

can be formulated on a higher level.[6] The assignment of sortal information to lexical items is provided by the semantic lexicon, cf. section 8.3.

Let us take a closer look at the lower-level parts of the entities' hierarchy. Figure 2 shows the categorization of time objects which plays an important role in the VERBMOBIL domain because of the high frequency of time expressions. The choosen differentiation is first of all motivated by the translation of temporal prepositions.

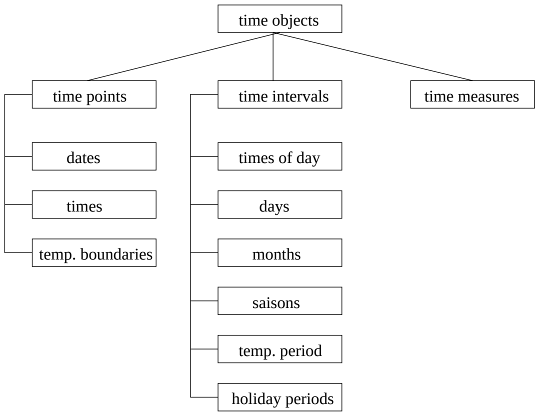


Figure 2: The structure of time objects

Figure 3 shows parts of the cross-classification of situations. On the one hand, situations are characterized with respect to the fields of their denotation, e.g. communication situations, movements, attitude-expressing or appointment-scheduling situations, and on the other hand, with regard to the thematic relations they usually involve. The second kind of information is used as a heuristic for the interpretation of prepositions, especially to discover way relations, i.e. sources, goals or paths, which are semantically analysed as modifiers so that the distinction between the stative and directional interpretation of prepositions, which influences their translation, gets neutralized.

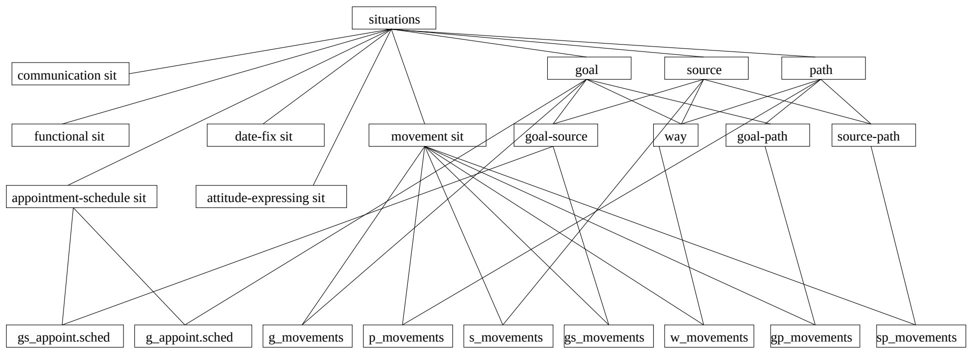


Figure 3: Parts of the situations' classification

[6]It is a particularly rare case that the translation of a lexical item depends on one special word in its context, as e.g. '*im* Urlaub/ sein' - 'be *on* holidays'.

Figure 4 depicts the cross-classification of concrete things. They are characterized by the features of artificiality[7], dimensionality and boundedness. The latter information is required for the transfer of spatial expressions. The entities which inherit values of these types are named and might be further subdivided into sorts which are motivated by other criteria, e.g. concerning instruments or buildings - their different functionality.

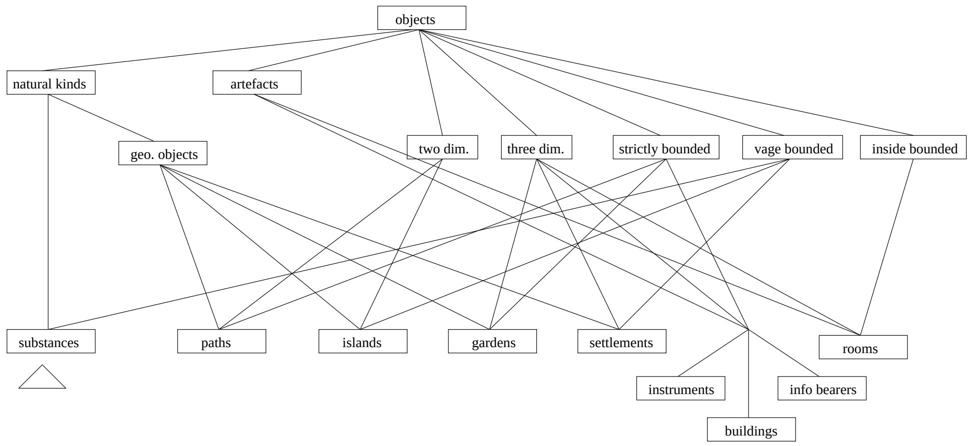


Figure 4: The object's part of the hierarchy

In order to express the sortal ambiguity of systematically polysemous nouns, cf. Bierwisch (1983) and Nunberg (1979), we declare disjunctive types which capture the interpretations occurring in the VERBMOBIL domain. Disjunctive specifications in the lexicon can thus be avoided.

(8.6) `inst_loc_c    = institution_c; building_c.`

(8.7) `info_medium_c = information_c; info_bearer_c.`

(8.8) `meal_food_c   = event_c; food_c.`

Nouns denoting institutions, such as *Universität*, are often used to refer to the building which houses the institution and to the group of people which is associated with it. These interpretations are anchored in the type `inst_loc_c` in(8.6). Other possible meanings are not considered here. Similary, nouns like *Buch* are regarded as being an abstract information with respect to their content, and an information bearer considering their material manifestation, cf.(8.7). Other nouns, such as *Frühstück*, cover the meal as well as the food reading, cf. (8.8). With respect to the hierarchy introduced so far, the different sortal interpretations of polysemous nouns are often disjoint at the upper level. Therefore, they are defined as subtypes of `entity_c`.

[7]Although they belong to the category of natural kinds, human beings are excluded from this part. With respect to sectional restriction they behave quite differently.

8.2.2 The hierarchy of relations

The relational part of the hierarchy represents conceptual relations which are associated with nouns, adverbs and prepositions. They are defined as relations between two entities. At least at the upper level, they correspond to thematic relations. Considering the lower parts, these relations are highly specialized. The reason for this is their special functionality. While upper level relations may be used for the specification of selectional restrictions, the fine-grained lower-level relations represent interlingual concepts which are used for the translation of temporal, spatial, modal and other expressions, cf. 7.[8] For the sake of modularity one may argue to keep that sorts which are used to fix selectional restrictions separate from sorts, which are used to lexicalize target language expressions separate. On the other hand, there is reason to include the fine-grained relations into the hierarchy, since they present a refinement of general upper-level relations. Figure 5 introduces a small part of the hierarchy of relations. The refinement of upper-level relations is shown here with the example of spatial relations, i.e. the localization in an object's vicinity. The fine-grained relations differ with respect to the concrete part of the vicinity in which something is included. All these relations are lexicalized by different prepositions, e.g. *by*, *behind*, *in front of*, *below* or *above*.

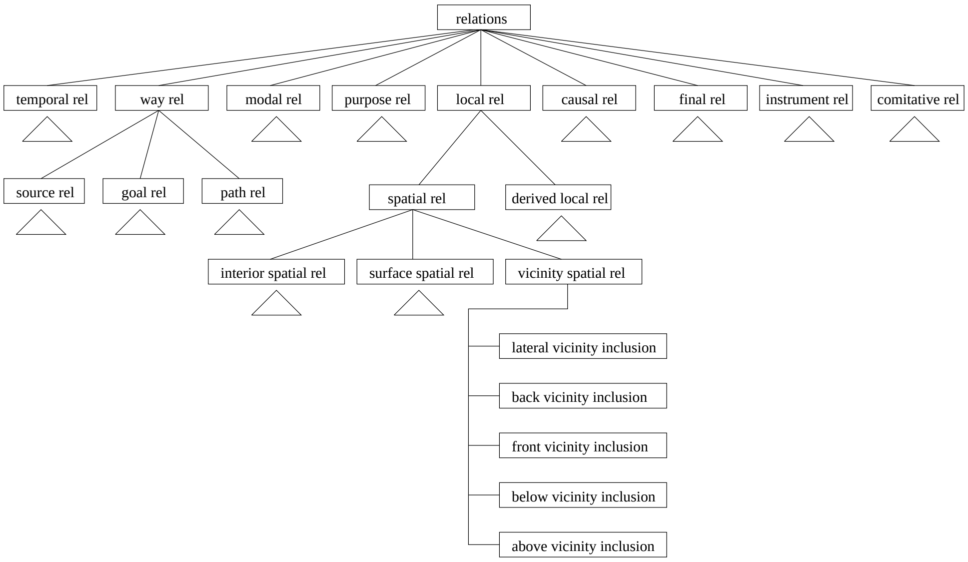


Figure 5: Parts of the relations' hierarchy

[8] The transfer module has a hybrid architecture. It combines direct transfer mappings with interlingual specifications, cf.1.

8.3 The assignment of sorts and relations in the lexicon

In this subsection we briefly sketch the assignment of sortal and relational information in the semantic lexicon.

8.3.1 The sortal specification

In the semantic lexicon, the sortal feature of a discourse marker, is assigned the fine-grained sort of a lexical item. It inherits both upper-level sortal information relevant for the semantic construction and lower-level sortal information for the resolution of transfer ambiguities. We exemplify the sort assignment with some entries of the semantic lexicon. (8.9) shows the assignment of unambiguous sorts to semantically different types of nouns.
(8.9)

```
lex(karlsruhe) => propername_sem(karlsruhe,settlement_c).
lex(konferenz) => common_noun_sem(konferenz,event_c).
```

The polysemous nouns in (8.10) are provided with disjunctive sortal specifications which cover their various interpretations, cf. (8.6) -(8.8).
(8.10)

```
lex(universitaet) => common_noun_sem(universitaet,inst_loc_c).
lex(fruehstueck)  => common_noun_sem(fruehstueck,meal_foot_c).
```

Verbs are described according to the dynamic/static distinction and with respect to their situation type, which also includes information about their thematic structure. Their entry contains the thematic relations of their arguments including a sortal restriction on their range, cf. (8.11).
(8.11)

```
lex(kommen)  => intransitive_verb_sem(kommen,dynamisch_c & movement_wi_c,
                                       agent_rel,person_c).
lex(glauben) => transitive_verb_sem(glauben,statisch_c & mental_sit_c,
                                     agent_rel,person_c,
                                     theme_rel,entity_c).
```

The fine-grained sortal specification thus appears already in the semantics, althought it is not used in this module with this specificity. However, semantic construction and transfer work as an integrated STUF process and the assignment is done only once. For the sake of modularity we prefer a more modular sort assignment in the next phase of the project. The sortal information part of the lexicon should be structured in a way that it serves semantic construction and transfer separately. Operating on the same type hierarchy a structured lexicon could provide the semantic construction with upper-level sorts and the transfer with lower-level sorts. This way, changes in the lower structure of the sortal hierarchy do not affect the basic upper sorts used in the semantic construction.

8.3.2 The relational specification

Conceptual relations of the type `relation_rel` are assigned to the role feature of arguments. These are, on the one hand, thematic relations, like agent, theme,

experiencer, whose range is constrained by sortal restrictions, e.g. (8.12)
(8.12)

```
lex(schlagen) => ditransitive_prefix_verb_sem(vor,vorschlagen1,
                                 dynamisch_c & communicat_sit_c,
                                 agent_rel,human_c,
                                 experiencer_rel,human_c,
                                 theme_rel,entity_c.
```

On the other hand, there are relations usually expressed by modifiers, such as temporal, local, modal, causal, etc. relations and various subtypes of them. They are, for example, assigned to adverbs:

(8.13)

```
lex(danach)    => pronoun_prep_adv_sem(loc_far_pred,nach1,~human_c,demonstr,
                                    temporal_span_posterior_rel).
lex(mittags)   => temporal_adv_sem(mittag,quant(temploc(pexact(time_of_day_c,
                                temporal_inclusion_rel)))).
lex(irgendwie) => pronoun_adv_sem(irgendwie,entity_c,std,modal_rel).
```

In contrast to most adverbial modifiers, the specific contextual interpretation of prepositional modifiers is much broader, especially with respect to their translation. The description of their meaning would require a disjunctive specification of all their admissible interpretations. In order to avoid massive disjunctions in the lexicon, prepositions are assigned relations, cf.(8.14)-(8.16), which get their disjunctive definition in the type hierarchy.

(8.14) `lex(mit)  => prep_sem(mit,mit_rel) & case_insensitive(dat).`
(8.15) `lex(an)   => prep_sem(an,an_rel & #R) & case_sensitive(#R).`
(8.16) `lex(ans)  => det_prep_sem(an,an_rel & goal_rel).`

The underspecified relation `mit_rel` in(8.14) restricts the range of interpretations possible for the preposition *mit*. Its concrete contextual meaning is determined by refinement rules, cf. section 7. However, for prepositions which exhibit case alternation, the set of interpretations can be restricted already in the lexicon. The underspecified `an_rel` in(8.15), for example, is intersected with `goal_rel` if the internal argument of the preposition is accusative. Similary, the scope of interpretation for clitical prepositions, as e.g. *ans* in(8.16), can be restricted already in the lexicon.

Chapter 9

Tense

As part of the semantic construction a tense condition is instantiated within the DRS. The tense morphology of the verbs provides the crucial clues towards constructing the relevant Reichenbachian temporal relations between E (event time), R (reference time) and S (speech time). Furthermore, the presence of temporal adverbials, negation, and quantification is registered and represented within the tense condition.

The evaluation and simultaneous disambiguation of the information collected in the tense condition results in the instantiation of a *sur(face)_tense* feature within the tense condition. This is undertaken as the last step of the semantically-based recursive transfer (see section 2).

This surface tense feature provides a clue to the generation component. In priniciple, however, the generation component is not restricted to the information provided in the surface tense feature, but can also work with the interlingua representation contained in the tense condition. Thus, the generation component is given an interesting degree of freedom: for example, if another verb were chosen as being subtly more appropriate within the given context than one produced by the transfer component, and if that verb changed the Aktionsart of the expression, then the generation component would still have access to the interlingua representation, and be able to generate a more appropriate surface tense.

9.1 Basic Approach

9.1.1 Temporal Relations

Ehrich (1992) presents the attractively simple, and yet sufficiently powerful schema in (7) as the basis for an analysis of the German tense system. The distinction she makes between *contextually* and *intrinsically* determined relations is also sometimes viewed as the difference between tense (relation between R and S) and aspect (relation between E and R) (e.g., Apello 1986, Allegranza et al. 1991).

(7)

		Contextually Determined	
		S, R	R < S
Intrinsic	E, R	Present	Past
Relations	E < R	Perfect	Past Perfect
	E > R	—	—

The notation "S,R" signifies that these times stand in some sort of relation to one another, though whether this relation is one of overlap or temporal precedence is underspecified and is further determined by the context (temporal adverbials or discourse context).[1]

Within VERBMOBIL, morphological and syntactic temporal information are mapped to interlingua representations according to the correlations presented in Table (8).

(8)

Present Tense Verbs	$E \bigcirc R$,	$R > S$ or $S \subseteq R$
Past Tense Verbs	$E \bigcirc R$	$R < S$
Infinitives/Participles	E,R	R,S
Future Auxiliary		$R > S$
Present Perfect Auxiliary	$E < R$	$R \bigcirc S$
Past Perfect Auxiliary	$E < R$	$R < S$

For a lengthier discussion of the theoretical assumptions which the described implementation is based on, see Butt (1994, 1995).

9.1.2 Tense Representation within the Semantic Formalism

Concretely, the above information is integrated into the Semantic Formalism in form of a tense condition. The tense condition is a complex condition within the *conds* slot of a DRS.here.

```
tense_condition   ::    e_rel_r                 : tense_rel,
                        r_rel_s                 : tense_rel,
                        tense_inst              : marker,
                        e_time                  : marker,
                        r_time                  : marker,
                        s_time                  : marker,
                        tloc                    : tloc_type,
                        sur_tense               : etense_val.
```

The types e_time, r_time, and s_time and the encoding of the relations between them is as described above. The s_time is coindexed with a (contextual) time anchor, whose value is "now". The e_time is coindexed with the instantiation of the verb. Both are thus bound by existential closure.[2]

As it is as yet not possible to evaluate anaphoric relations, or to evaluate temporal adverbials with respect to a calendar model, it is not possible to instantiate an

[1]The future tense is not included in Table (7) as Ehrich follows Vater (1975) in treating the German future auxiliary *werden* as a modal.

[2]The label e_time applies to both states and dynamic events – no crucial distinctions are lost with regard to this simplification of terminology since the precise nature of the eventuality is encoded both in terms of sortal information, and in terms of Aktionsarten.

evaluation procedure which would take a temporal adverbial, place it in relation to the speech time, and determine the reference time for the event. The temporal relations as they stand are thus not *inferred* from calendar and contextual information, but are specified lexically. A hierarchical modeling of the set of temporal relations, which are based on the proposals in Allen (1983), further ensures that the various lexical specifications can be combined compositionally and yet monotonically.

The feature *tense_inst* serves to identify the tense condition uniquely. In most cases, the value of the tense_inst is exactly that of the e_time, but under quantification and negation the tense_inst is flagged with sortal information in order to be able to identify the tense condition as having been introduced by negation or quantification. The details of the treatment and implementation of quantification with regard to tense are described in the section on quantification.

Finally, the *sur_tense* is instantiated as a last step of transfer and contains the English surface tense that the expression should be generated with.

9.1.3 Temporal Adverbials

In the *tloc* slot information about the presence and particular nature of temporal adverbials is gathered. On the basis of the detailed analysis of temporal adverbs in terms of the hierarchy defined by K. Eberle, macros within the semantic construction fill the *tloc* slot with the appropriate information. The four differing kinds of temporal adverbs shown in Table (9) are distinguished.

(9)

Type		**Examples**
st_dist	(not coincident with S)	*morgen*, *im April*, *montags*
st_equ	(coincident with S)	*heute*, *jetzt*
st_prec	(precedes S)	*gestern*, *vorhin*
st_perf	(perfectivizing)	*seitdem*, *bisher*

Temporal adverbials of the type st_dist are by far the most common and serve to disambiguate the German present tense to a future reading. Temporal adverbials of the type st_equ restrict the German present tense to a present reading. The type st_perf was introduced to allow for the perfectivizing effect of *seitdem* (see discussion in the section on perfects). Finally, the st_prec serves to prevent utterances like (10) from being realized in the future. Instead, a translation in the present or present progressive is rendered, depending on the type of verb (stative or not).

(10) #Ich komme gestern.

The *tloc* thus in effect situates the event, and is analogous to the notion of reference time (r_time). However, the temporal adverbial which introduces the *tloc* specifications is not explicitly identified with the r_time. This is because they are not always identical. In the case of quantification, for example, there may be a temporal adverb (and, hence, a *tloc*), but when it is contained within the scope of the quantiifier, it may not serve as the r_time for the expression. Furthermore, when there is no explicit temporal adverb in an expression, a reference time which situates the event must still be assumed. This is represented by the r_time, underspecified though it may be. See Butt (1995a,b) for further discussion of the notion of reference time.

9.2 The German Present Tense

The German present tense is analysed as fundamentally ambiguous: it can refer to past, present, or future occurences (e.g., Ehrich 1992). Within the scope of VERBMOBIL, we ignore the historical present, and model the present as being ambiguous between a present and a future reading. The present tense must be disambiguated by means of temporal adverbials or context. As contextual information is not yet available for the Demonstrator, in the absence of an overt temporal adverbial, the German present tense is realized as either English present or present progressive, depending on whether the verb describes a stative or a dynamic event.

Representative examples from various VERBMOBIL dialogs in the simple present are shown in (11)–(13).[3]

Dynamic predicates (events) usually occur with a temporal adverb. Examples like (12), in which a dynamic predicate occurs on its own are extremely rare in the dialogs. As (11) shows, when an event predicate occurs with a temporal adverb, then it must be realized in the future in English. When an event predicate occurs on its own, it must be realized with the present progressive.

Event in combination with a Temporal Adverb

(11) DRF:28: genau, wir **treffen** uns **dann** in der Eingangshalle des Czerczinsky mit den Unterlagen
ERF:29: we **will meet** in the lobby of the Czerczinsky with the papers.

Unmodified Event

(12) wir **treffen** uns in der eingangshalle.
we **are meeting** in the lobby

Statives, on the other hand, are always realized in the present tense in English: whether or not a temporal adverb is present is irrelevant.

Statives

(13) a. D7:08 Dienstag **ist** etwas ungünstig.
E7:08 Tuesday **is** a bit inconvenient.

b. D7:14 **Meinen** Sie das **reicht** uns?
E7:14 Do you **think** that **is** enough for us?

c. D7:07 Da **habe** ich noch Zeit.
E7:07 I still **have** time then.

[3] The labeling of the examples is as follows: DRF and ERF refer to German and English sentences, respectively, from the *Referenzdialog.* D7, E7, etc., refer to utterances from the Blaubeuren Dialogs, which are considered to be the core corpora for the Demonstrator. The numbering here is according to a standardized version agreed upon by the syntax, semantics, transfer, and generation partners. Examples labeled "Bonn", etc., stem from further VERBMOBIL dialogs which are not included in the core corpora, but which have been tagged, translated, and investigated in Tübingen. Examples not labeled at all have been made up.

d. BonnD1:89 und **am siebenundzwangisten**, da **ist** leider der Kongress in Berlin wieder
BonnE1:89 and **on the twenty-seventh** there **is** the conference in Berlin again

The tense condition intially introduced by a verb with present tense morphology is as follows.

The German Present Tense

e_rel_r:	assoc	(underspecified)
r_rel_s:	follow_eq	(ambiguous)
r_time :	sort: temporal_c	(underspecified)
s_time :		(specified as "now" in anchors)
e_time :		(coindexed with the verb)
tense_inst:		(same as the e_time)

The rules which disambiguate underspecified temporal relations and instantiate the English surface tense follow the general schema shown below. The information contained in a tense condition which is relevant for a calculation of the surface tense is checked: the type of the *tense_inst* (temporal vs. nontemporal for quantification), the type of the temporal adverb (*tloc*), whether the event is dynamic or stative (the sort of the e_time), and the nature of the specification for the temporal relations between E and R, and R and S, which are also simultaneously specified further.

General Format for the Instantiation of Surface Tense

```
saturate_tense =>  (e_rel_r:#ER & r_rel_s:#RS & tense_inst:sort:#TIS &
     e_time:sort:#ES & tloc:#TL) &
     saturate_tense1(#ER,#RS,#TIS,#ES,#TL).
```

The *saturate_tense* rules are instantiated through *surface_tense* in the recursive transfer statement below. The definition of *trs* allows the recursive traversal of a sign, and the translation of lexical items and passing along of interlingua representations through *tau* (see section 2).

Recursive Transfer and Surface Tense

```
trs([#F&prag:#P|#R]) =>
    [sem:(saturate_sem(tau(#F,[]))&surface_tense)&prag:#P|trs(#R)].
```

At each stage, *surface_tense* looks for a tense condition, checks whether an English surface tense value has been provided yet, and instantiates an appropriate value for the English surface tense where necessary through the specific *saturate_tense* rules (e.g., 1–4 below).[4]

[4]The integration of the tense module into the overall transfer component was undertaken by the Transfer group at IMS-Stuttgart (in particular, K. Eberle, C.J. Rupp, M. Dorna, and M. Emele). The relatively complicated sign-traversal macros needed for the identification of tense conditions and satisfaction of the *sur_tense* value, as well as general technical support at all stages were also provided by the IMS group.

As was illustrated in the above source and target language semantic representations, the *tloc* slot is filled by the introduction of a temporal adverb into the semantic representation. For a treatment of the simple German present tense, two factors need to be taken into account.

- If the verb is dynamic (an event), and there is a temporal adverb (st_dist), then the English tense is future (see 1).[5]
- If the verb is dynamic, and there is no temporal adverb, then the English tense is present progressive. (see 2)
- If the verb is stative, the presence of a temporal adverb has no influence: the English tense is always present (see 3 and 4).

In 1–4, the *saturate_tense* statements needed for a treatment of simple German present tense sentences are shown.

1. Present Disambiguated to Future by Temporal Adverbial → Future

```
saturate_tense1(~precede,follow_eq,temporal_c,dynamisch_c,st_dist) =>
      e_rel_r: overlap_rel &
      r_rel_s: follow &
      sur_tense: fut.
```

2. Present of Dynamic Verbs → Present Progressive

```
saturate_tense1(~precede,follow_eq,temporal_c,dynamisch_c,~st_dist &
                         ~st_perf&~st_prec) =>
     e_rel_r: overlap_rel &
     r_rel_s: equal &
     sur_tense: presprog.
```

3. Present of Stative Verbs → Present

```
saturate_tense1(~precede,follow_eq,temporal_c,statisch_c,
                         ~st_dist & ~st_perf & ~st_prec) =>
     e_rel_r: overlap_rel  &
     r_rel_s: equal &
     sur_tense: present.
```

[5]Various kinds of temporal adverbs are defined in terms of what intervals around the speech time they describe – for a disambiguation of the German present tense, only the type *st_dist* is relevant.

4. Present of Stative Verbs with Temporal Adverb → Present

```
saturate_tense1(~precede,follow_eq,temporal_c,statisch_c,st_dist) =>
      e_rel_r: overlap_rel &
      r_rel_s: follow &
      sur_tense: present.
```

The above rules are part of the transfer component, but, as already mentioned, also simultaneously disambiguate or further specify the information about the relations between E, R and S.[6]

For further discussion on the English futurate and interactions of the present tense with performatives and imperatives, see Butt (1995).

9.3 The German Past Tense

Occurences of the German past tense within the VERBMOBIL dialogs are rare and fairly uninteresting. They always correspond to the English past or past progressive, depending on whether or not the predicate in question is a state or an event. Typical examples are shown in (14) and (15).

(14) D1:19 nein halt das **war** jetzt Mai
E1:19 no wait, that **was** May now

(15) D1:07 oh, ich **dachte** eigentlich an den fünfzehnten April
E1:07 oh, I **was** actually **thinking** of the fifteenth of April, a week later

The tense condition as initially introduced by a verb with past tense morphology is as follows.

Past Tense

e_rel_r:	assoc	(underspecified)
r_rel_s:	precede	(unambiguously past)
r_time :	sort: temporal_c	(underspecified)
s_time :		(specified as “now” in anchors)
e_time :		(coindexed with the verb)
tense_inst:		(same as the e_time)

The transfer rules for the English surface tense are shown in 5 and 6.

5. German Past of Dynamic Verbs → Past Progressive

```
saturate_tense1(~precede,precede,##,dynamisch_c,##) =>
     e_rel_r: overlap_rel &
     sur_tense: pastprog.
```

[6] For ease of presentation, we have abstracted away from implementational details here — in a non-procedural language like STUF further conditions need to be added to ensure that not more than one solution per expression is instantiated. For the rules as they actually appear see Appendix A.

6. German Past of Stative Verbs → Simple Past.

```
saturate_tense1(~precede,precede,##,statisch_c,##) =>
     e_rel_r: overlap_rel &
     sur_tense:spast.
```

9.4 The Periphrastic Tenses

9.4.1 The Future

Utterances in the future also occur very rarely in the VERBMOBIL dialogs. When they do occur, they can always be translated straightforwardly with the English future. Some typical examples are shown in (16) and (17).

(16) Karls5:12 und dann **werden** wir uns am Mittwoch **sehen**
Karls4:12 and then we**'ll meet** on Wednesday

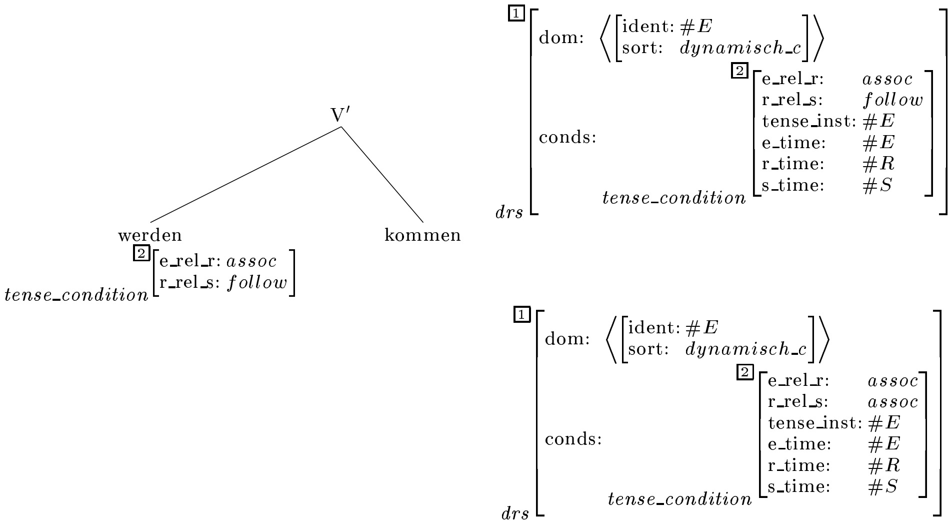
(17) Karls5:16 auf jeden Fall **werde** ich am Donnerstag **kommen**
Karls5:16 in any case, I**'ll come** on Thursday

The tense information here must be determined compositionally from the future auxiliary *werden*, and the main verb of the sentence. As can be seen in the following example, where the compositional construction of the tense condition is illustrated,[7] the infinitive main verb introduces completely underspecified values with regard to the temporal relations that hold between E, R and S. The auxiliary *werden* then serves to provide more specific values for the relations: the reference time must be located after (follow) the speech time.

(18) Wir werden kommen.

[7] For ease of presentation, only the tense condition is shown.

FIGURE 3: Source Language (German) Representation



The transfer rule for the simple German future is given in 7.[8]

7. German Future → English Future

```
saturate_tense1(~precede,follow,temporal_c,##,##) =>
      e_rel_r: equal &
      sur_tense:fut.
```

9.4.2 The Perfect Tenses

Current Treatment

Occurences of the perfect are also very rare in the dialogs, and always correspond to the English simple past. Since our primary aim was to implement a broad treatment of the German present tense for the Demonstrator, we have concentrated on not more than the occurences in the VERBMOBIL dialogs, and have translated all German present perfects as English simple past. An example is shown in (19).

(19) D3:01 der Termin den wir neulich **abgesprochen haben** ...
E3:01 the appointment that we **decided on** the other day ...

Utterances in the German past perfect are even rarer, and always correspond to the English past perfect.

(20) BonnD1:02 das **hatten** wir ja **vorgesehen**
BonneE1:02 we **had planned** on that

The tense condition for the perfect and the past perfect is constructed compositionally, in analogy to the future.[9] The perfect auxiliaries *sein* and *haben* introduce

[8]Future perfects have not as yet occured in the VERBMOBIL dialogs. However, the model presented here can be readily extended to include a treatment of these constructions as well.

[9]The realization of this approach is due to C.J. Rupp, who coordinated the specification of the *mtv* (mood-tense-voice) access predicate in close cooperation with the syntax (S. Schachtl).

the condition that the event time must lie before the reference time. The relation between the speech time and the reference time is determined by the morphological tense of the auxiliary: if the auxiliary carries present tense morphology, the relation between R and S will be one of coincidence. If the auxiliary carries past tense morphology, the reference time will precede the speech time.

The Perfect

e_rel_r:	precede	(introduced by auxiliary)
r_rel_s:	equal	(present tense morphology)
r_time :	sort: temporal_c	(underspecified)
s_time :		(specified as "now" in anchors)
e_time :		(coindexed with the verb)
tense_inst:		(same as the e_time)

The Past Perfect

e_rel_r:	precede	(introduced by auxiliary)
r_rel_s:	precede	(past tense morphology)
r_time :	sort: temporal_c	(underspecified)
s_time :		(specified as "now" in anchors)
e_time :		(coindexed with the verb)
tense_inst:		(same as the e_time)

The transfer rules for the English surface tense are shown in 8 and 9.

8. German Present Perfect → English Simple Past

```
saturate_tense1(precede,equal,##,##,##) =>
      sur_tense:spast.
```

9. German past perfect → English past perfect

```
saturate_tense1(precede,precede,##,##,##) =>
      sur_tense:pastperf.
```

9.5 Progressives

No special treatment of progressives in terms of an operator was necessary. For a detailed discussion see Butt (1995).

9.6 Quantification

The interaction between quantification and tense is represented in terms of two tense conditions: one introduced by the event in the scope of the quantifier, and one introduced by the quantifier.

The quantifier does not itself denote an event, so the tense condition of the quantifier is flagged as being nontemporal (*entitaet_c*). The information with regard to the temporal relations that hold between E, R, and S are "copied" up from the embedded tense condition. The connection between the two tense conditions is guaranteed by the fact that they both refer to the same event (e_time).

The details of the copying/raising approach were suggested by K. Eberle to allow an easier integration of the tense module into the transfer component: direct access to temporal information at the top level is thus ensured and tense is given wide scope. For a more detailed discussion of quantification and the interaction with tense see Butt (1995).

The fact that the tense condition of a quantifier is flagged as nontemporal allows a triggering of special rules needed for the present tense. As we saw above, German present tense sentence are generally ambiguous between a present and future reading. Furthermore, depending on the Aktionsart of the verb, either a simple present, or a present progressive is appropriate in English. Quantified present tense sentences, however, can only be realized in the simple habitual present tense.

(21) Ich mache immer montags Termine aus.
I always make appointments on Mondays.
*I am always making appointments on Mondays.
*I will always make appointments on Mondays.

The flagging of the tense condition for the quantifier as nontemporal instantiates the special rules in 13 and 14. The rule in 13 ensures that (21) will be realized in the English present tense.

13. Quantification without Temporal Adverb → Present

```
saturate_tense1(assoc,follow_eq,entitaet_c,##,~st_dist) =>
      r_rel_s : equal &
      sur_tense:present.
```

14. Quantification with Temporal Adverb → Future

```
saturate_tense1(assoc,follow_eq,entitaet_c,##,st_dist) =>
       r_rel_s : follow &
       sur_tense:fut.
```

The rule in 14 allows for cases like (22), in which the temporal adverb situates the set of appointment-making events in the next year.

(22) Nächstes Jahr mache ich immer montags Termine aus.
Next year I'll make appointments on Mondays.

For a more in depth discussion of issues yet to be resolved, especially with regard to possible differing scoping relations between quantifiers and temporal adverbs, again see Butt (1995).

9.7 Modals and Conditionals

A tense condition each is constructed for the modal and the embedded main verb. The speech time (s_time) in each tense condition is the speech time specified contextually (in the *anchors* feature). The tense condition of the embedded verb remains underspecified and is assigned a value of *nonfin*(inite) for the English surface tense. For the modal a finite surface tense value is instantiated according to the tense transfer rules described above. The presence of temporal adverbs has no effect on tense realization because modals are consistently analysed as states.

Though modals are quite straightforward with respect to tense and transfer, several complications arose in the implementation. One of the more interesting ones, and one which has not as yet received a satisfactory solution, is the problem posed by conditionals (subjunctives) like *könnte*, *sollte*, and *ständen*. For one, all of these are analyzed as past tense in the syntax because of their morphological form. The rules which map from the syntactic information in the *mtv* (mood, tense, voice) access predicate to temporal relations therefore initially assigned *precede* as the relation between R and S and situated all conditionals squarely in the past. And while it could be argued that at least *sollte* in principle has a past tense reading, in actual fact, none of the conditionals in question are ever used in the past tense in the VERBMOBIL dialogs. Examples are shown in (23).

(23) a. D7:05b dann **sollten** wir unseren Termin davor ausmachen
E7:05b then we **should** arrange for our appointment before then

b. D7:03a Anfang Juli **hätte** ich noch Zeit
E7:03a I **would** still **have** time at the beginning of July

This problem is circumvented easily for most of the conditionals by taking mood into account in addition to the morphological tense in the mapping from syntax to temporal relations.

9.8 Summary

The approach to the transfer of temporal phenomena presented here allows a broad coverage of the VERBMOBIL corpora. A tense condition within a DRS is constructed compositionally through a lexical specification of Reichenbachian temporal relations in the semantic lexicon, and information about the morphological tense of a predicate that is provided by the syntax. The compositional construction of tense conditions is realized in parallel with the Semantic Construction. The subsequent evaluation and instantiation of the English surface tense takes place within the Transfer module. The instantiation of target language tense also takes into account the interaction of tense with temporal adverbs, and quantification. In particular, temporal adverbs serve to disambiguate the German present tense.

Since our primary focus for the Demonstrator was the German present tense, the system naturally needs be extended to allow a more complete treatment of tense and aspect in the second phase of the project. The extension is planned in terms of a greater reliance on Aktionsart, the use of a calendar model and world knowledge (modeled in BACK/FLEX) to allow a more precise evaluation of temporal relations and a modeling of temporal anaphora.

Chapter 10

Example Output

The output of the transfer module for the sentence in (1) is shown in (2):
(1) *dann schlage ich vor donnerstag achter juli um halb vier*
(2)

```
%%% used time (msec): 9750
%%% result:

[(sign &
  sem:(sem_t &
       lambda:[] &
       drs:(drs_t &
            dom:[] &
            conds:[(phi_condition &
                    phi_arg:(drs_t &
                             dom:[(A &
                                   marker &
                                   sort:dynamisch_c)] &
                             conds:[(alfa_condition &
                                     alfa_arg:(B &
                                               marker &
                                               snumber:individual &
                                               sort:mensch_c) &
                                     alfa_type:std &
                                     alfa_restr:(drs_t &
                                                 dom:[B] &
                                                 conds:[])),
                                    (alfa_condition &
                                     alfa_arg:(C &
                                               marker &
                                               sort:uhrzeit_c) &
                                     alfa_type:std &
                                     alfa_restr:(drs_t &
                                                 dom:[C] &
                                                 conds:[(rigid_condition &
                                                         rigid_pred:tofd &
                                                         rigid_inst:C &
                                                         designator:330)])),
                                    (basic_condition &
                                     pred:at &
                                     inst:(D &
                                           marker &
                                           sort:(entitaet_c & ~situation_c)) &
```

```
                         args:[(arg_role &
                                arg:C &
                                role:unspec_temporal_inclusion_rel)]),
                        (alfa_condition &
                         alfa_arg:(E &
                                   marker &
                                   sort:tage_c) &
                         alfa_type:std &
                         alfa_restr:(drs_t &
                                     dom:[E] &
                                     conds:[(rigid_condition &
                                             rigid_pred:dofw &
                                             rigid_inst:E &
                                             designator:5),
                                            (basic_condition &
                                             pred:timeloc &
                                             inst:E &
                                             args:[(arg_role &
                                                    arg:(F &
                                                         marker &
                                                         sort:monate_c) &
                                                    role:relation_rel)]),
                                            (rigid_condition &
                                             rigid_pred:mofy &
                                             rigid_inst:F &
                                             designator:7),
                                            (rigid_condition &
                                             rigid_pred:ord &
                                             rigid_inst:F &
                                             designator:8)])),
                        (basic_condition &
                         pred:timeloc &
                         inst:D &
                         args:[(arg_role &
                                arg:E)]),
                        (G &
                         basic_condition &
                         pred:suggest &
                         inst:A &
                         args:[(arg_role &
                                arg:B &
                                role:agent_rel),
                               (arg_role &
                                arg:D &
                                role:theme_rel)]),
                        (tense_condition &
                         e_rel_r:overlap_rel &
                         e_time:A &
                         r_rel_s:equal &
                         r_time:(marker &
                                 sort:temporal_c) &
                         s_time:(H &
                                 marker &
                                 sort:temporal_c) &
                         sur_tense:presprog &
                         tense_inst:A &
                         tloc:(tloc_type & ~st_dist & ~st_perf & ~st_prec))]) &
             phi_op:phi_op_val &
             phi_pred:then1)]) &
quants:[] &
```

```
      anchors:[(disc_anchor &
                 param:B &
                 discourse_role:speaker),
                (time_anchor &
                 time_param:H &
                 time_role:now)] &
      ip:(ip_t &
          cond:G &
          idx:A) &
      persp:(persp_t &
             p_inst:A)) &
prag:(prag_t &
       imp:no &
       key_wds:[dann] &
       prosody:prosody_val &
       verb_position:verb2 & whq:no))]
```

Chapter 11

Coverage

An overview about the tested coverage can be found in table 11.1:

4	dialogues
108	sentences
889	words
107	sentences through parser
103	sentences through semantic construction
100	sentences through transfer

Table 11.1: Tested Coverage

Table 11.2 shows the figures of the potential coverage acccording to the Verbmobil demosntrator word list:

	defined	used in dialogues
VERBMOBIL wordlist	1292	279
German Semantic Lexicon (full form)	1606	279
German Semantic Lexicon (lemmata)	619	207
Transfer Rules	742	255
English Semantic Lexicon (entries)	553	

Table 11.2: Potential Coverage

The following semantic phenomena were treated within the MDS semantic construction and transfer component:

- idioms
- isolated words and phrases
- anaphoric and elliptical expressions
- date and time expressions

- wh-questions
- reflexive and prefix verbs
- verbs with sentential and prepositional complements
- determinerless noun phrases
- nouns functioning as modifiers, etc.

Bibliography

V. Allegranza, P. Bennett, J. Durand, F. van Eynde, L. Humphreys, P. Schmidt, and E. Steiner. Linguistics for machine translation: The Eurotra linguistic specifications. In *Studies in Machine Translation and Natural Language Processing*. Office for Official Publications of the Commission of the European Community, Luxembourg, 1991.

J. Allen. Maintaining knowledge about temporal intervals. *Comm.ACM*, 26:832–843, 1983.

L. Appelo. A compositional approach to the translation of temporal expressions in the Rosetta system. In *Proceedings from the Eleventh International Conference on Computational Linguistics (COLING)*, pages 313–318, Bonn, 1986.

N. Asher. Belief in discourse representation theory. *Journal of Philosophical Logic*, 15:127–189, 1986.

M. Bierwisch. Semantische und konzeptuelle Repräsentation lexikalischer Einheiten. In R. Ruzicka and W. Motsch, editors, *Untersuchungen zur Semantik*, volume XXIX of *studia grammatica*, pages 61–99. Akademie-Verlag, 1983.

J. Bos. Presupposition as Anaphora in the Verbmobil Semantic Formalism. Verbmobil-Report 25, Universität Saarlandes, 1994.

J. Bos. Focusing Particles and Ellipsis Resolution . Verbmobil-Report 61, Universität Saarlandes, 1995.

B. Buschbeck-Wolf and R. Nübel. Die Behandlung übersetzungsambiger Präpositionen im Transfer des VERBMOBIL-Demonstrators. Verbmobil-Report 86, IBM/IAI, 1995.

M. Butt. An abstract representation of reference times. In *Proceedings of the DGfS/CL, Düsseldorf*, 1995a. To Appear.

M. Butt. Transfer I: Tense and Aspect. Verbmobil-Report 55, Universitaet Tübingen, 1995b.

K. Dahlgren. *Naive Semantics for Natural Language Understanding*. Kluwer Academic Publishers, Boston, 1988.

M. Dorna, K. Eberle, M. Emele, and C. J. Rupp. Semantik-orientierter rekursiver Transfer in HPSG am Beispiel des Referenzdialogs. VM-Report 39, IMS, Universität Stuttgart, 1994.

J. Dörre and M. Dorna. CUF - A Formalism for Linguistic Knowledge Representation. DYANA-2 Deliverable R.1.2A, IMS, Universität Stuttgart, 1993.

J. Dörre, M. Dorna, and J. Junger. *CUF User's Manual.* IMS, Universität Stuttgart, 1995.

K. Eberle. Ereignisse: Ihre logik und ontologie aus textsemantischer sicht. IWBS Report 192, IBM Deutschland, WT LILOG, Stuttgart, 1991.

K. Eberle. Zu einer Semantik für Dialogverstehen und Übersetzung. VM-Report 83, IMS, Universität Stuttgart, 1995.

K. Eberle and W. Kasper. Tenses as anaphora. In *Proceedings of E-ACL*, Manchester, 1989.

K. Eberle and W. Kasper. French past tenses and temporal structure. In R. Thieroff and J. Ballweg, editors, *Tense Systems in European Languages*, pages 149–171. Niemeyer, Tübingen, 1994.

M. Egg. *Aktionsart und Kompositionalität: zur kompositionellen Ableitung der Aktionsart komplexer Kategorien.* Akademie-Verlag, Berlin, 1994.

V. Ehrich. *Hier und jetzt: Studien zur lokalen und temporalen Deixis im Deutschen.* Max Niemeyer Verlag, Tübingen, 1992.

B. Hamp. Temporal Anaphorische Adverben im Deutschen: *da* und *dann.* Memo 67, Verbmobil, Universität Tübingen, 1995.

W. J. Hutchins and H. L. Somers. *An Introduction to Machine Translation.* Academic Press, Cambridge, 1992.

H. Kamp and C. Rohrer. Tense in texts. In R. Bäuerle, R. Schwarze, and A. von Stechow, editors, *Meaning, Use and Interpretation of Language.* de Gruyter, Berlin, 1983.

H. Kamp and C. Rohrer. Temporal reference in french. (ms.), IMS, Universität Stuttgart, 1985.

G. Klose, E. Lang, and T. Pirlein. *Ontologie und Axiomatik der Wissensbasis von LILOG.* Springer, Berlin, 1992.

E. König. *The Meaning of Focus Particles. A Comparative Perspective.* Routeledge, London and New York, 1991.

T. Lexandowski. *Linguistisches Wörterbuch.* Quelle & Meyer, 1994.

S. McGlashan. Lexical Semantic Macros. VerbMobil-Memo 75, Universität Saarlandes, 1994.

S. Momma, A. Opalka, I. Raasch, J. Schneider, and R. Seiffert. *STUF-III User Manual.* IBM, 1994.

S. Nirenburg, V. Raskin, and A. Tucker. The stucture of interlingua in TRANSLATOR. In S. Nirenburg, editor, *Machine Translation*, pages 90–114. Cambridge University Press, 1987.

G. Nunberg. The Non-uniqueness of Semantic Solutions: Polysemy. *Linguistics and Philosophy*, 3:143–184, 1979.

B. Partee. Some structural analogies between tenses and pronouns in english. *Journal of Philosophy*, 70:601–609, 1973.

B. Partee. Nominal and temporal anaphora. *Linguistics and Philosophy*, 7(3): 243–287, 1984.

J. Quantz, M. Gehrke, U. Küssner, and B. Schmitz. The VERBMOBIL Domain Model Version 1.0. Verbmobil-Report 29, Technische Universität, Berlin/Siemens AG, München, 1994.

H. Reichenbach. *Elements of Symbolic Logic.* MacMillan, New York, London, 1947.

S. Reinhard. Transfer of Verbal Predicates. Verbmobil-Report 98, SfS, Universität Tübingen, 1995.

B. Schmitz and K. Fischer. Pragmatisches Beschreibungsinventar für Diskurspartikeln und Routineformeln anhand der Demonstratorswortliste. VerbMobil-Memo 83, Technische Universität Berlin, Universität Bielefeld, 1995.

R. Steinberger. Lexikoneinträge für deutsche Adverbien. *Tagungsband, KONVENS'94. Verarbeitung natürlicher Sprache*, pages 320–330, 1994.

P. Terell, R. Breitsprecher, V. Calderwood-Schnorr, and W. V. Morris. *PONS. Komparativwörterbuch. English-Deutsch.* Klett, Verlag für Wissen und Bildung, Stuttgart, 1991.

R. van der Sandt. Presupposition projection as anaphora resolution. *Journal of Semantics*, 9(4), 1992.

H. Vater. *Werden* als Modalverb. In Calbert and H. Vater, editors, *Aspekte der Modalität*, pages 71–148. Narr, Tübingen, 1975.

H. Winhart. Semantische Auswertung und Transfer von FVG in VERBMOBIL. Verbmobil-Report 78, SfS, Universität Tübingen, 1995.