# Using Hybrid Methods and Resources in Semantic-based Transfer*

**Bianka Buschbeck-Wolf** and **Michael Dorna**
Institut für Maschinelle Sprachverarbeitung (IMS)
Azenbergstr. 12
D–70174 Stuttgart
{bianka,michl}@ims.uni-stuttgart.de
http://www.ims.uni-stuttgart.de/projekte/verbmobil/

## Abstract

This paper presents ongoing work on the development of the semantic transfer component of the multi-lingual speech-to-speech MT system Verb*mobil*. It focuses on the use of symbolic and statistical methods for the acquisition of semantic transfer rules, the disambiguation of translational ambiguities and the selection of appropriate rule candidates at runtime.

## 1 Introduction

In this paper we describe how a combination of different methods and resources is used for the development of the transfer component of Verb*mobil* (Dorna & Emele 96b). Verb*mobil* (Wahlster 93) is a multi-lingual speech-to-speech MT system that is applied to the task of translating spoken language in the domain of appointment scheduling and travel planning. Currently, the system includes modules for German, English and Japanese.

Over the last decades, neither pure stochastic approaches to machine translation (MT), such as the statistical approach (Brown *et al.* 90) or example-based MT (Sato & Nagao 90; Sumita *et al.* 90), nor pure symbolic methods, as pursued in METAL (Slocum *et al.* 87), SYSTRAN (Wheeler 87) or LOGOS (Schmid & Gdaniec 96), turned out to be sufficient for high quality translation. A reasonable, task-specific combination of different techniques has proved to provide the best results (Carbonell *et al.* 92; Lehmann & Ott 92; Brown & Frederking 95).

With our semantic transfer approach, we present ongoing research on the integration of successful methods from different paradigms. We focus on the combination of stochastic and symbolic methods in both the *acquisition of bilingual semantic transfer lexica* and the *disambiguation of translational ambiguities*.

In large MT projects, such as Verb*mobil*, the linguistic resources, representations and tools for analysis and generation are developed in parallel with the translation component. Hence, one has to think of a strategy for isolating transfer from ongoing changes during the project by minimising the dependencies between different components. For this reason, we have developed *templates for transfer rules* that cope with classes of translation patterns in a systematic way and minimise necessary adaptations for actual representations.

Another well-known problem in MT is the selection of appropriate rule candidates at runtime if there are alternatives among applicable transfer rules. We introduced the *specificity principle* in (Dorna & Emele 96a) which we have extended for processing different translation alternatives in parallel.

This paper is organised as follows: in Section 2, we describe our semantic transfer approach together with the semantic representation language. Section 3 demonstrates the extraction of semantic transfer knowledge from a pool of bilingual resources. In Section 4, we show the exploitation of symbolic and statistical disambiguation techniques and illustrate the idea of disambiguation on demand with a series of examples. Finally, Section 5 outlines some technical details of the transfer rule compiler and runtime system.

## 2 Semantic-based Transfer

### 2.1 Our Approach

The semantic-based transfer we present is based on some central ideas of the MRS-based approach outlined in (Copestake 95) and the Shake-and-Bake approach to MT sketched in (Beaven 92; Whitelock 92). In order to preserve ambiguities that hold across languages and to reduce analysis efforts significantly, semantic transfer relates source language (SL) and target language (TL) semantic descriptions which are underspecified wrt. scope and attachment ambiguities (Bos *et al.* 96). To cope with multilinguality, the transfer includes a refinement component which introduces partially language-independent representations (see Section 2.3).

The semantic construction produces underspecified semantic representations (see Section 2.2) that form the input to the transfer. The transfer obtains additional information from a semantic evaluation component that keeps track of the dialogue history and provides discourse information, such as speech acts (Alexandersson *et al.* 97a). Besides this symbolic evaluation module, a statistical evaluation component allows the transfer to access information about TL co-occurrences (see Section 4). The transfer module reports its TL semantic representations to the generator which maps them to TL expressions (Kilger & Finkler 95; Copestake *et al.* 95). See Figure 1 for the Verb*mobil* architecture from the view point of transfer.

---

[*] This work was funded by the German Federal Ministry of Education, Science, Research and Technology (BMBF) in the framework of the Verb*mobil* project under grant 01 IV 701 N3. The work presented here is strongly influenced by discussions with our colleagues of the Verb*mobil* subproject TP 3 "Translation", namely A. Eisele, M. Emele and M. Schiehlen. We would also like to thank G. Carroll and the anonymous reviewers for helpful comments. The responsibility for the contents of this article lies with the authors.

## 2.2 Semantic Representation

The semantic representation is based on a variant of Underspecified Discourse Representation Structures (UDRS) (Reyle 93).

The semantic analysis result together with additional information, such as number, tense, aspect, prosodic accent, sorts, dialogue act, etc., is represented in a multi-dimensional data structure called Verb*mobil* Interface Term (VIT) (Dorna 96). This single information structure serves as interface representation for all components that operate on semantic structures, i.e. transfer, semantic evaluation and generation.

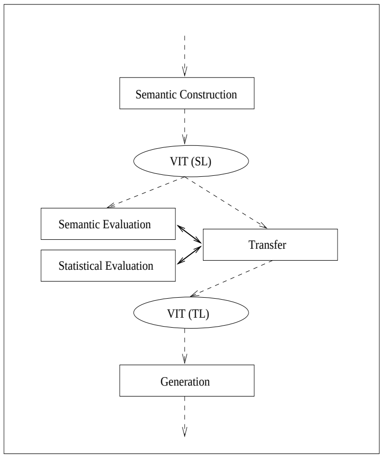

Figure 1: Interaction of the Transfer Component

We restrict the following presentation to that part of the VIT which contains a set of semantic entities. Each entity has a unique label `l` which is used as an address for linking information within and between multiple levels of a VIT. Besides their label, referential predicates introduce an instance `i`. Argument roles and modifier relations are represented in a Neo-Davidsonian way (Parsons 91). Semantic operators like quantifiers, modals or scopal adverbs take extra label arguments for referring to other elements which are in the relative scope of these operators. The semantic entities containing skolemised labels and instances encode a recursive representation in a flat set-oriented list structure. This data structure turned out to be convenient for the specification of transfer operations. Consider example (1a) and its favourable English translation (1b).

(1a) Wir wollten das Gespräch doch vorziehen?
(1b) We wanted to schedule the meeting earlier, didn't we?

This SL utterance is assigned the representation in (2). The relative scope of the sentence mood operator `ynq(l0,l1)` (yes-no question), the control verb `wollen(l1,i1)` and the pragmatic adverb `doch(l2,l3)` has been fully resolved by using the explicit labels of other predicates, these are `l1,l2` and `l3`.[1]

(2)
```
ynq(l0,l1),wollen(l1,i1),arg1(l1,i1,i2),
arg3(l1,i1,l2),doch(l2,l3),vorziehen(l3,i3),
arg1(l3,i3,i2),arg3(l3,i3,i4),
pron(l4,i2,speaker_hearer),def(l5,i4,l6),
gespraech(l6,i4)
```

## 2.3 Abstraction

Semantic transfer operates on a relatively abstract level of representation wrt. the Vauquois triangle (Vauquois 75). Here, morpho-syntactic realisations are abstracted away and a variety of language-independent categories are introduced, such as referentiality, tense, and mood, etc. Moreover, the semantic formalism used allows us to avoid resolving those ambiguities that hold across languages.

The transfer component includes several cascaded modules. Among them, there is a tense resolution module (Schiehlen 97) and a refinement module. The latter maps contextually synonymous predicates to bilingual abstractions, and decomposes complex predicates into language-independent semantic primitives (Buschbeck-Wolf & Tschernitschek 96). Additionally, particular ambiguous predicates are resolved, since it is often necessary to disambiguate before other transfer operations can start.[2]

The output of the transfer is a partial language-neutral representation that allows the generator to produce paraphrases. Generalisation and decomposition lead to a reduction of the number of transfer rules to the necessary minimum and lower the costs for introducing new languages. The advantages are similar to those of an interlingua approach to MT, see e.g. (Kay *et al.* 94).

## 2.4 Transfer Knowledge Bases

The primary knowledge bases of our transfer component are a set of monolingual refinement and restructuring rules (see Section 2.3) and the database of bilingual transfer equivalences. In the following, we consider only the latter.

The general form of a transfer rule is shown in (3). It establishes the equivalence between sets of SL semantic entities `SL_Sem` and sets of TL semantic entities `TL_Sem`. The operator `TauOp` indicates in which direction a rule is applied, i.e. bi-directional (↔) or uni-directional (→ or ←).[3]

(3) `SL_Sem # SL_Cond TauOp TL_Sem # TL_Cond.`

Below we present some simple transfer rules for the unambigous predicates in (2).[4]

---

[1] If the scope is underspecified, explicit subordination constraints are fixed in the scope slot of the VIT. The exact details of subordination are beyond the scope of this paper; see (Frank & Reyle 95) and (Bos *et al.* 96) for implementations.

[2] Since all restructuring and refinement operations are motivated by contrastive data, we assume this functionality to be part of the transfer process. This way, the modularity of the grammar components can be maintained.

[3] A rule application reduces the SL input by the set of semantic entities in `SL_Sem` if they form a matching subset of the input. On the other hand, the TL semantic enties `TL_Sem` are added to the TL output (see (Dorna & Emele 96a) for details).

[4] The transfer formalism provides a single metarule which can be used instead of mappings for identical pred-

```
(4) ynq(L,I,R) <-> ynq(L,I,R).
    def(L,I,L1) <-> def(L,I,L1).
    wollen(L,I) <-> want(L,I).
    pron(L,I,R) <-> pron(L,I,R).
```

The capitalised symbols `L`, `I`, `R`, etc. stand for logical variables which are bound to concrete values when applying a rule to a given input.

To resolve translational ambiguities, the rules are optionally provided with a condition part (`SL_Cond` and `TL_Cond`) which restricts their application to the relevant context. The condition part contains only tests (see also (Morimoto *et al.* 92)). The `#` sign separates the transfer mapping from the rule restriction.[5] Splitting the mapping from the condition part leads to smaller translation units. Thus, problems with the interaction of rules can be minimised.

(5) shows a mapping of the ambiguous predicate `vorziehen` to `prefer` under the condition that its theme argument `arg3` is of sort `time`.[6]

```
(5) vorziehen(L,I) # arg3(L,I,I1),sort(I1,time)
    <-> prefer(L,I).
```

### 2.5 Templates for Transfer Rules

For building the transfer rule bases we introduce a set of templates that can be called within or replace the actual transfer rules. Templates have a couple of advantages. They

- simplify the writing of transfer rules;
- allow the adaptation of changes in the semantic representations systematically;
- simplify the mapping of extracted translation patterns to transfer rules;
- capture generalisations of translation divergences.

Thus, a template mechanism ensures the adaptation and reusability of transfer rules independently of the concrete front and back end of a transfer component. Templates are defined for transfer equivalences as well as for frequently used conditions. (6) shows those that correspond to the rules in (4). The calls of *rule templates* are prefixed by the `@` operator.

```
(6) @mood(ynq, ynq).
    @quant(def, def).
    @verb(wollen, want).
    @pron(pron, pron).
```

(7) gives the definition of a template for the transfer of quantifiers. Its usage is restricted to the semantic type the predicate belongs to. In our case it is `qua` for quantifiers. The variables `SL` and `TL` are replaced by the incoming quantifiers when calling the template.

```
(7) quant(qua(SL), qua(TL)) :=
    SL(L,I,L1) <-> TL(L,I,L1).
```

While this is a very simple case, templates are more useful for capturing systematic changes in the semantic structure of the involved languages (Dorr 94).

```
(8) support_arg1verb(adx(SL), verb(TL)) :=
    support(L,I,L1), SL(L1,I1) <->
    TL(L,I), arg1(L,I,I1).
```

The definition of the `support_arg1verb` template in (8) covers the switch of a support (copula) into a verbal construction, as it is relevant for the translation of *einverstanden sein* into *agree*. The SL `support` predicate together with its predicative of the type `adx` (adjectival/adverbial modifier) is substituted by a TL verbal predicate. The instance of the SL predicative `I1` becomes the verb's `arg1`. Thus, the translation correspondence between *einverstanden sein* and *agree* can be established in a simple way (9).

```
(9) @support_arg1verb(einverstanden, agree).
```

*Condition templates* are applied in the rule's condition part in order to fix frequently occurring restrictions more efficiently. They are used to express, e.g., that a predicate is of a specific sort or semantic type, is modified, is quantified or is embedded in a particular way, etc. Condition template calls are prefixed with the `+` operator. As a result, the conditioned rule in (5) can be expressed by (10).[7]

```
(10) # +arg3_sort(I,time)
     @verb(vorziehen,prefer,I).
```

## 3 Transfer Rule Development

In this section, we show how we exploit existing resources for the acquisition of initial bilingual transfer lexica. These are domain-specific bilingual corpora of spoken material as well as on-line dictionaries, thesauri and off-line resources.

### 3.1 Extraction of Translation Candidates

MT is still most successful in restricted application domains because the number translational ambiguities can be kept small. Unfortunately, for such specific domains neither dictionaries nor parallel corpora of a sufficient size are available. This applies especially to spoken language resources. It is not yet clear whether or how knowledge extracted from large general corpora, e.g., mono- and bilingual co-occurrence frequencies (Kitamua & Matsumoto 96) or learned translation rules (Almuallim *et al.* 96), can be used in smaller domains. Moreover, domain specific readings and transfer mappings are not covered by models trained on the basis of unspecific resources.

---

icates on SL and TL side.

[5] The condition to the left of `TauOp` restricts the application direction →, and vice versa.

[6] The disambiguation of this predicate is explained in more detail in Section 4.

[7] The `verb` template is combined with a rule where only the condition on the SL side is specified. The complete notation would look like

```
true # +arg3_sort(I,time) <-> true
@verb(vorziehen,prefer,I).
```

where `true` stands for any predicate. The rule and the template are merged at compile time (see Section 5.1.1).

In Verb*mobil* we work in a restricted domain with a limited vocabulary constraint by the speech recognizer. Hence, we have no problems with unknown words, but the arbitrary word combinations cause different readings of the same words. Again, the usual on-line resources such as thesauri do not always cover domain specific readings and thus can hardly be used for automatic disambiguation of word senses.[8]

In this situation, it is more reasonable to stick to a symbolic approach concerning the overall architecture of an MT system and involve statistical methods wherever possible. So we still use mono- and bilingual annotation, alignment and extraction tools only for preparing the contrastive data for the rule writer. At the moment, this seems to be the only way to ensure translation quality in the domain of spoken language dialogues. See Section 3.3 for future directions.

### 3.2 Combining Existing Resources and Tools

In the following, we sketch the derivation of transfer rule skeletons using existing tools and resources at IMS. The process itself is mainly independent of the tools in use. However, the quality of the results reflects the precision of the tools.

Taggers (Schmid 94), lemmatisers and/or morphological analysers (for German (Schiller 95; Schulze 96)) are used to annotate independently the monolingual parts of parallel bilingual corpora. After sentence and word alignment (Eisele 97) the data is prepared for off-line and interactive corpus queries (Schulze & Christ 95). For each word to be covered by transfer rules these tools can produce parallel subcorpora containing all occurrences and translations found in the data.

Extracted translation correspondences are annotated with frequencies of their occurrence in monolingual corpora and in bilingual domain-specific corpora. This information is used to guide the manual rule definition and refinement to achieve a rapid growing coverage combined with a good quality of the transfer output.

### 3.3 Automatic Acquisition of Semantic Transfer Rules

Future research concerns the automation of transfer rule acquisition for symbolic transfer approaches. Robust parsers (Abney 97) already produce syntactic chunks which are used to construct partial semantic analyses (Light 96).

The quality of the chunks is getting better with the availability of larger resources that cover syntactic and semantic restrictions on argument bindings. Ongoing projects at IMS successfully develop corpora extraction techniques, e.g., for finding syntactic subcategorisation frames (Eckle & Heid 96) in combination with word sense clusters (Rooth 94) to produce huge on-line lexica including morphological, syntactic and semantic knowledge.

Given all these annotation and preparation steps for both languages in a parallel corpus, an alignment on different linguistic levels, such as words, phrases or even semantic fragments is possible by using similarity measures, see e.g., (Kitamua & Matsumoto 96). This alignment and co-occurrence information will be used to compute mappings to transfer templates (see Section 2.5). Subsequently, these templates are enriched by further contextual conditions that constrain the transfer mappings if there is more than one correspondence (see Section 4 below).

## 4 Disambiguation

Spoken language is highly ambiguous because of a large number of ambiguous word, anaphora, deictic expressions and ellipsis. To achieve an acceptable translation quality ambiguity and anaphora resolution as well as ellipsis reconstruction are of major importance in Verb*mobil*.

Since it is too expensive to resolve all ambiguities, anaphora, etc. we have developed the concept of *resolution on demand* (Buschbeck-Wolf 97). Besides the strategy of preserving scopal and attachment ambiguities by underspecification (see Section 2.1) this technique minimises the analysis efforts.

The need of particular resolution procedures can be best recognised in the transfer component. In case of alternative translation correspondences, we have to consider the contexts in which the one or the other TL expression is used. For the resolution of many translational ambiguities the local VIT context is sufficient. It allows us to formulate restrictions on the sort or the semantic type of a predicate, on its scopal embedding, mood, number, aktionsart, etc. (Buschbeck-Wolf & Tschernitschek 96) as well as on prosodic information (Lieske *et al.* 97). However, there are cases in which the transfer component needs more global information to choose a particular TL correspondence. This is, e.g., information about the antecedent of an anaphor or an ellipsis, domain-specific world knowledge, speech act and discourse stage information as well as information about the pragmatic function of discourse particles. This information can be obtained either from the semantic or the statistic evaluation component (see Figure 1).

By anchoring in the transfer rules specific requests to the semantic and statistical evaluation components, the transfer triggers the required resolution procedure (see Section 4.1 and 4.2). Some requests, such as the identification of an antecedent, are generated automatically. They are detected if required information, such as the sort of a referential predicate, is lacking in the semantic construction output.

### 4.1 Symbolic Disambiguation

Let us illustrate the disambiguation of translational ambiguities with sentence (1) introduced in Section 2.2.

(1a) Wir wollten das Gespräch doch vorziehen?
(1b) We wanted to schedule the meeting earlier, didn't we?

(1a) includes three ambiguous words. These are: the pragmatic particle *doch* which, among others, gets translated into *yes, after all* or into a question tag,

---

[8] The techniques for stochastic word sense disambiguation are getting better and better, see e.g., (Almuallim *et al.* 96; Dorr & Jones 96). But there is no hope to get word sense models without large corpora which are manually tagged with such senses (Ng 97).

but it can also be dropped in the English translation; *vorziehen* which means either *prefer* or *schedule earlier*; and the noun *Gespräch* which corresponds in our domain to *meeting*, *discussion* or *conversation*.

Let us consider some examples that show the different discourse functions of the particle *doch* (Stede & Schmitz 97).

(11a) *Doch*, ich komme am Montag.
(11b) *Yes*, I'll come on Monday.

(12a) Wir wollten uns *doch* am Montag treffen?
(12b) We wanted to meet on Monday, *didn't we?*

(13a) Dann würde das *DOCH* gehen.
(13b) Then, it would be possible, *after all*.

Sentence-intitial *doch* in (11) is used to deny the utterance of another speaker and to reaffirm one's own previously expressed opposite opinion. In (12), *doch* signals the return to a previously made arrangement. The speaker reminds the hearer of a scheduled meeting and expects his approving response. Similarly, in (1), the speaker reminds the hearer of their agreement to schedule the meeting earlier. With the prosodically marked *doch* in (13), the speaker refers to a previous dialogue stage. Something that was impossible before turned out to be feasible at the utterance time.

All three readings of *doch* in (11)-(13) need to be translated differently. While the reading in (11) can be recognised by its sentence initial occurrence (`sent_init(L)`), for (12) and (13) the transfer problem consists in the identification of the pragmatic function of *doch*. (Stede & Schmitz 97) developed a classification of discourse particles wrt. their discourse function. They consider the particle *doch* to be either a coherence marker which should not be translated or a pointer to something previously uttered (`particle_class(L,given)`). In case the latter function is identified, the translation of *doch* seems to differ wrt. the sentence mood. In yes-no questions (`mood(ynq)`) it is expressed by a question tag in English. In declarative or imperative sentences (`mood(decl;imp)`) it is mapped onto *after all* if it bears prosodic accent (`pros_accent(L)`).

The transfer rules in (15)-(16) integrate in these kinds of restriction in their condition parts.

(14)
```
# sent_init(L)
@particle(doch,yes,L).
```
(15)
```
# particle_class(L,given),mood(ynq)
@particle(doch,quest_tag,L).
```
(16)
```
# particle_class(L,given),mood(decl;imp)
  pros_accent(L)
@particle(doch,after_all,L).
```

Now let us consider the verb *vorziehen*. In (17)–(20), there are some contexts in which its mapping to one of its equivalents is quite obvious. (The contextual trigger is marked in bold face.)

(17a) Ich würde **den Dienstag** *vorziehen*.
(17b) I would *prefer* Tuesday.

(18a) Ich würde das Treffen **gerne** *vorziehen*.
(18b) I would like to *schedule* the meeting *earlier*.

(19a) Wir **sollten** das Treffen *vorziehen*.
(19b) We should *schedule* the meeting *earlier*.

(20a) Ich würde es *vorziehen*, **am Montag zu kommen**.
(20b) I would *prefer* to come on Monday.

If the theme argument is a time expression, *prefer* is the only correspondence (17). This is captured by the rule in (10) in Section 2.5. The translation with *schedule earlier* is not feasible then, because it would require a kind of movable object as `arg3` and times, such as Tuesday, are fixed.

(10)
```
# +arg3_sort(I,time)
@verb(vorziehen,prefer,L,I).
```

*Vorziehen* is an attitude verb. If it is modified by an adverb that expresses the speakers attitude, it cannot have an attitude reading itself (18). This excludes *prefer* as correspondence (21).[9]

(21)
```
# +attitude(I)
@verb(vorziehen,schedule,L,I)
@add_compadx(early,I).
```

The *prefer* reading of *vorziehen* is odd similar to those reasons discussed in the previous example when occurring within the scope of a modal, as in (19) and (1), where an attitude is also expressed. The corresponding rule is shown in (22).[10]

(22)
```
# +modal(L1,L2), L =< L2
@verb(vorziehen,schedule,L,I)
@add_compadx(early,I).
```

If *vorziehen* has a propositional `arg3` realisation (20), *prefer* is the appropriate equivalent (23).[11]

(23)
```
# +proposit_arg3(I)
@verb(vorziehen,prefer,L,I).
```

However, these rules are not sufficient to cope with the ambiguity of *vorziehen*. If its `arg3` belongs to the sort `situation` (24a) and it does not occur in one of the mentioned contexts (18)–(20), then the local semantic context does not allow to choose between the translations in (24b) and (24c).

---

[9] The template `attitude` stands for attitude adverbs. It is attached to the verbal predicate by coindexation of the instance `I`. The template `add_compadx` introduces a modifier and a comparative predicate for it.

[10] The template `modal` checks the semantic type of the incoming verbs. The equation says that the label of `vorziehen` is below the label of the modal, i.e. under its scope.

[11] The template `proposit_arg3` tests the propositional argument realisation.

| | |
|---|---|
| **sich in** ein *Gespräch* **einschalten** | **join** a *conversation* |
| **ins** *Gespräch* **kommen** | **get into** a *conversation* |
| ein *Gespräch* **fortsetzen** | **continue** a *conversation* |
| ein *Gespräch* **vorbereiten** | **prepare** a *meeting/discussion* |
| ein *Gespräch* **organisieren** | **organize** a *meeting/discussion* |
| ein *Gespräch* **planen** | **plan** a *meeting/discussion* |

Table 1: Co-ocurrences of *Gespräch* with verbal predicates

| | |
|---|---|
| ein **stockendes** *Gespräch* | a **faltering** *conversation* |
| ein **nettes** *Gespräch* | a **nice** *conversation* |
| ein *Gespräch* **unter vier Augen** | a **private** *conversation* |
| ein **angesetztes** *Gespräch* | a **scheduled** *meeting/discussion* |
| ein **vertagtes** *Gespräch* | a **postponed** *meeting/discussion* |
| ein **dringendes** *Gespräch* | an **urgent** *meeting/discussion* |

Table 2: Co-ocurrences of *Gespräch* with modifiers

(24a) Ich würde **das Treffen** am Montag *vorziehen*.
(24b) I would *prefer* the meeting on Monday.
(24c) I would *schedule* the Monday meeting *earlier*.

To resolve this ambiguity, the semantic evaluation component, which provides more information on the actual dialogue situation, is consulted by an `eval` call for resolving the particular reading of *vorziehen* (25) or (26).

```
(25) # +arg3_sort(I,sit), eval(L,vorziehen_pref)
     @verb(vorziehen,prefer,L,I).
(26) # +arg3_sort(I,sit), eval(L,vorziehen_move)
     @verb(vorziehen,schedule,L,I)
     @add_compadx(early,I).
```

Concerning our example sentence in (1a), so far the rules in (6), (15) and (22) have been applied in order to get the translation in (1b).

### 4.2 Statistical Disambiguation

Finally, the noun *Gespräch* has to be mapped onto its contextual appropriate correspondence. In contrast to verbal and modificational predicates, the disambiguation of nominal predicates is notoriously difficult, since it is impossible to fix manually all contexts in which the one or the other translation is preferred. For this task it is reasonable to rely on statistical information (Kameyama *et al.* 93).

Consider the German noun *Gespräch*. It corresponds to *meeting* or *discussion* when it is a kind of organised event, while the other equivalent *conversation* denotes a more casual event or refers to its course. There are several contexts which force the one or the other interpretation. Co-occurring with verbs, such as *organise, plan* or *prepare*, *meeting* or *discussion* seem to be appropriate translations, while in other contexts *conversation* is used (see Table 1). Similarly, modifiers can identify one of the alternatives (see Table 2).

To extract the correct translation (Gale *et al.* 92), we regard the TL contexts in which the TL correspondences of the ambiguous word occur. If they are close to the input context (Almuallim *et al.* 96) the statistical evaluation component validates the corresponding transfer rule (Eisele 97). The one that is relevant for our example sentence in (1) is shown in (27). In the context of *schedule*, the "official event" interpretation of *Gespräch* is appropriate, i.e. *meeting* or *discussion* are the translations. Concerning the choice between them, *meeting* is more likely to be used in the domain of meeting scheduling.

```
(27) # stat_eval(discussion;meeting,I)
     @nom(gespraech,meeting,I).
(28) # stat_eval(conversation,I)
     @nom(gespraech,conversation,I).
```

## 5 Processing of Transfer Rules

The IMS transfer system consists of a transfer rule compiler (`trc`) and a transfer runtime system (`trs`). The `trc` version described in (Dorna & Emele 96a) was extended by a template preprocessor. Furthermore, we have refined the specificity principle which handles the rule selections at runtime. This gives us a selection criterion when `trs` computes several transfer alternatives in parallel.

### 5.1 Compile time Processing

#### 5.1.1 Template Expansion

Rule and condition templates (see Section 2.5) are expanded before any further compilation takes place. All parts of a rule found during a template expansion are merged, respectively. I.e. the sets of semantic entities and rule conditions are united. Additionally, the direction of possible rule applications is determined. A bi-directional operator (`<->`) will be overridden by a uni-directional operator (`<-` or `->`) if a template definition or the rule itself contains this operator. If no overriding takes place, the application of a rule is always possible in both directions. The result of this preprocessing is a regular transfer rule (see Section 2.4).

#### 5.1.2 Rule Compilation

Transfer rules are always part of a module where each side of a rule belongs to a specific language. `trc` uses this information to check the compatibility of semantic entities wrt. language specific on-line lexica. The lexica are part of an ADT package for the VIT (Dorna 96). The semantic entities in transfer rules are sets of terms and `trc` partially orders these sets. The result are sequences (lists) which are used to collapse rules with the same prefix. Then `trc` builds an index over the prefixes for fast accessing applicable rules when matching rules against the input at runtime (Dorna & Emele 96a).

The compiled rules form a kind of transducer which takes a set of SL semantic entities as its input and produces a TL representation.

### 5.2 Runtime Processing

`trs` works incrementally on linguistically motivated segments of different sizes. The segment size may vary from words over constituents to sentences depending on the output of the recogniser and linguistic analysis components. Robustness is achieved by handling all sizes of segments. The quality of transfer output is improving with the length of segments.

At runtime, transfer solves a problem which can be reduced to a set covering problem. `trs` looks for the minimal number of subsets covering an input set. The subsets are defined by semantic entities found in SL matching parts (`SL_Sem`) of transfer rules. `trs` tries to find the most specific rules which cover the largest subsets. The *specificity principle* defined in (Dorna & Emele 96a) ensures a rule selection which is locally optimal. If at a particular processing state one such rule is found, all other candidates are blocked. This non-monotonic behaviour does not always give the optimal solution. Sometimes we find the most specific rule but not always the most specific sequence of rule applications (derivation).

Currently, we work on a parallel approach which looks for potential transfer results and selects those which are derived using a minimal number of rule applications. The minimal number of rule application is equivalent to the problem of finding the minimal number of subsets which was mentioned above. Therefore, this behaviour ensures a global optimum relating the input with the most appropriate rules designed for it.

The techniques for a possible realisation of parallel transfer range from chart processing (Amtrup 95), lemma table proof procedures (Johnson & Dörre 95) to transformation techniques for OR-parallel to AND-parallel programs using continuations (Ueda 86).

## 6 Summary

We have presented the semantic transfer approach of the speech-to-speech MT system Verb*mobil*. It is regarded to be a central component in Verb*mobil* that triggers inference processes in analysis and resolution components. The results are used for solving particular translation problems.

To cope with the multi-lingual scenario, we integrated a special refinement step that introduces language-independent elements into the language-specific semantic representation.

A template mechanism was developed to capture generalisations and to ensure the adaptation and reusability of transfer rules independently of the concrete input and output of the transfer component.

We exploited different bi- and monolingual resources for building the transfer rule base, i.e. for extracting translation correspondences. We have illustrated the strategy of resolution on demand by combining symbolic and statistical methods for choosing between translation alternatives.

Finally, the compilation of templates and transfer rules was sketched and the selection of appropriate transfer results in a parallel transfer approach was presented.

## References

(Abney 97) S. Abney. The SCOL Manual. Technical report, April 1997. URL: http://www.sfs.nphil.uni-tuebingen.de/~abney/scol.ps.gz.

(Alexandersson *et al.* 97a) J. Alexandersson, B. Buschbeck-Wolf, T. Fujinami, E. Maier, N. Reithinger, B. Schmitz, and M. Siegel. Dialogue Acts in Verbmobil-2. Verb*mobil* Report, forthcoming, 1997.

(Almuallim *et al.* 96) H. Almuallim, Y. Akiba, T. Yamazaki, A. Yokoo, and S. Kaneda. Two Methods for Learning ALT-J/E Translation Rules from Examples and a Semantic Hierarchy. In *Proceedings of the 15th COLING*, Kyoto, Japan, 1996.

(Amtrup 95) Jan W. Amtrup. Chart-based Incremental Transfer in Machine Translation. In *Proceedings of the 6th TMIMT*, Leuven, July 1995.

(Alshawi *et al.* 91) H. Alshawi, D. Carter, M. Rayner, and B. Gambäck. Translation by Quasi Logical Form Transfer. In *Proceedings of the 29th ACL*, Berkeley, CA, 1991.

(Beaven 92) J. L. Beaven. Shake-and-Bake Machine Translation. In *Proceedings of the 14th COLING*, Nantes, France, 1992.

(Bos 95) J. Bos. Predicate Logic Unplugged. In *Proceedings of the 10th Amsterdam Colloquium*, University of Amsterdam, Amsterdam, Holland, 1995.

(Bos *et al.* 96) J. Bos, B. Gambäck, C. Lieske, Y. Mori, M. Pinkal, and K. Worm. Compositional Semantics in Verbmobil. In *Proceedings of the 16th COLING*, 1996.

(Brown & Frederking 95) R. Brown and R. Frederking. Applying Statistical English Language Modeling to Symbolic Machine Translation. In *Proceedings of the 6th TMIMT*, Leuven, Belgium, 1995.

(Brown *et al.* 90) P. Brown, J. Cocke, S. Della Pietra, V. Della Pietra, J. Jelinek, F. anad Lafferty, R. Mercer, and P. Roossin. A Statistical Approach to Machine Translation. Vol. 16 of *Computational Linguistics*, 1990.

(Buschbeck-Wolf 97) Bianka Buschbeck-Wolf. Resolution on Demand. Verb*mobil* Report 196, IMS, Universität Stuttgart, 1997.

(Buschbeck-Wolf & Tschernitschek 96) B. Buschbeck-Wolf and C. Tschernitschek. What You Always Wanted to Know about Semantic Transfer. Verb*mobil* Report 114, IBM Heidelberg, Germany, 1996.

(Carbonell *et al.* 92) J. G. Carbonell, T. Mitamura, and E. H. Nyberg 3rd. The KANT Perspective: A Critique of Pure Transfer (and Pure Unterlingua, Pure Statistics, ... ). *Proceedings of the 4th TMIMT*, Montreal, 1992.

(Copestake 95) A. Copestake. Semantic Transfer in Verbmobil. Verb*mobil* Report 93, IMS, Universität Stuttgart, Germany and CSLI, Stanford University, 1995.

(Copestake *et al.* 95) A. Copestake, D. Flickinger, R. Malouf, S. Riehemann, and I. Sag. Translation using Minimal Recursion Semantics. In *Proceedings of the 6th TMIMT*, Leuven, Belgium, 1995.

(Dorna & Emele 96a) M. Dorna and M. Emele. Efficient Implementation of a Semantic-based Transfer Approach. In *Proceedings of the 12th ECAI*, 1996.

(Dorna & Emele 96b) M. Dorna and M. Emele. Semantic-based Transfer. In *Proceedings of the 16th COLING*, 1996.

(Dorna 96) M. Dorna. The ADT-Package for the Verb*mobil* Interface Term. Verb*mobil* Report 104, IMS, Universität Stuttgart, Germany, 1996.

(Dorr & Jones 96) B. J. Dorr and D. Jones. Acquisition of Semantic Lexicons: Using Word Sense Disambiguation to Improve Precision. In *Proceedings of SIGLEX-96, Santa Cruz, CA*, 1996.

(Dorr 94) B. J. Dorr. Machine Translation Divergences: a Formal Description and Proposed Solution. Vol. 20(4) of *Computational Linguistics*, 1994.

(Eckle & Heid 96) J. Eckle and U. Heid. Extracting raw material for a German subcategorisation lexicon from newspaper text. In *Proceedings of the 4th International Conference on Computational Lexicography*, Budapest, Hungary, 1996.

(Eisele 97) A. Eisele. IMS Alignment Tool Box. forthcoming, 1997.

(Frank & Reyle 95) A. Frank and U. Reyle. Principle Based Semantics for HPSG. In *Proceedings of the 7th EACL*, Dublin, Ireland, 1995.

(Gale *et al.* 92) W. A. Gale, K. W. Church, and D. Yarowsky. Using Bilingual Materials to Develop Word Sense Disambiguation Methods. *Proceedings of the 4th TMIMT*, Montreal, 1992.

(Johnson & Dörre 95) M. Johnson and J. Dörre. Memoization of Coroutined Constraints. *Proceedings of the 33rd ACL*, Cambridge, Mass., 1995

(Kameyama *et al.* 93) M. Kameyama, S. Peters, and H. Schütze. Combining Logic-Based and Corpus-Based Methods for Resolving Translation Mismatches. *Working Notes of the AAAI Spring Symposium Series*, Berkeley, 1993.

(Kay *et al.* 94) M. Kay, M. Gawron, and P. Norwig. *Verbmobil: a Translation System for Face-to-Face Dialogs*. CSLI Lecture Notes; No. 33. CSLI, Stanford University, 1994.

(Kilger & Finkler 95) A. Kilger and W. Finkler. Incremental Generation for Real–Time Applications. Research Report RR-95-11, DFKI GmbH, Saarbrücken, Germany, 1995.

(Kitamua & Matsumoto 96) M. Kitamua and Y. Matsumoto. Automatic Extraction of Word Sequence Correspondances in Parallel Corpora. In *Proceedings of the 4th Workshop on Very Large Corpora*, 1996.

(Lehmann & Ott 92) H. Lehmann and N. Ott. Translation Relations and the Combination of Analytical and Statistical Methods in Machine Translatiom. *Proceedings of the 4th TMIMT*, Montreal, 1992.

(Levin *et al.* 95) L. Levin, O. Glickman, Y. Qu, D. Gates, A. Lavie, C.P. Rosé C. Van Ess-Dykema, and A. Waibel. Using Context in Machine Translation of Spoken Language. *Proceedings of the 6th TMIMT*, Leuven, Belgium, 1995.

(Lieske *et al.* 97) C. Lieske, J. Bos, M. Emele, B. Gambäck, and C.J. Rupp. Giving Prosody a Meaning. In *Proceedings of Eurospeech'97*, 1997.

(Light 96) M. Light. CHUMP: Partial Parsing and Underspecified Representations. In *Proceedings of the ECAI-96 Workshop: Corpus-Oriented Semantic Analysis*, 1996.

(Morimoto *et al.* 92) T. Morimoto, M. Suzuki, T. Takezawa, G. Kikui, M. Nagata, and M. Tomokiyo. A Spoken Language Translation System: SL-TRANS2. In *Proceedings of the 15th COLING, Nantes, France*, 1992.

(Ng 97) Hwee Tou Ng. Getting serious about word sense disambiguation. In *Tagging Text with Lexical Semantics: Why, What, and How?* ANLP-97 Workshop, Washington, D.C., April 1997.

(Parsons 91) T. Parsons. *Events in the Semantics of English*. MIT Press, Cambridge, Mass., 1991.

(Qu *et al.* 96) Y. Qu, Rosé C.P., and B. Di Eugenio. Using Discourse Predictions for Ambiguity Resolution. In *Proceedings of the 16th COLING*, 1996.

(Reyle 93) U. Reyle. Dealing with Ambiguities by Underspecification: Construction, Representation and Deduction. Vol. 10(2) of *Journal of Semantics*, 1993.

(Rooth 94) Mats Rooth. Two-dimensional clusters in grammatical relations. Unpublished manuscript, 1994.

(Sato & Nagao 90) S. Sato and M. Nagao. Towards Memory-Based Translation. In *Proceedings of the 13th COLING*, Helsinki, Finland, 1990.

(Schiehlen 97) Michael Schiehlen. *Tense Resolution in* Verb*mobil*. Verb*mobil* Report, IMS, Universität Stuttgart, forthcoming.

(Schiller 95) Anne Schiller. DMOR: Entwicklerhandbuch und Benutzeranleitung. Unpublished manuscript, IMS, Universität Stuttgart, 1995.

(Schmid & Gdaniec 96) P. Schmid and C. Gdaniec. Evolution of the Logos Grammar: System Design and Development Methodology. In *Proceedings of the 2nd Conference of the Association for Machine Translation in the Americas*, Montreal, 1996.

(Schmid 94) H. Schmid. Probabilistic Part-of-Speech Tagging Using Decision Trees. In *International Conference on New Methods in Language Processing*, Manchester, UK, 1994.

(Schulze & Christ 95) B. M. Schulze and O. Christ. The XKWIC User's Manual. Technical report, IMS, Universität Stuttgart, 1995.

(Schulze 96) Bruno Maximilian Schulze. GERMLEX – ein Lemmatisierer für deutsche Textcorpora. Unpublished manuscript, IMS, Universität Stuttgart, 1996.

(Slocum *et al.* 87) J. Slocum, W. S. Bennet, J. Bear, M. Morgan, and R. Root. METAL: The LRC Machine Translation System. In S. Michaelson *et al.* (Eds.), *Machine Translation today: the state of the art*. Edinburgh University Press, 1987.

(Stede and Schmitz 97) M. Stede and B. Schmitz. Discourse Particles and Routine Formulas in Spoken Language Translation. In *Proceedings of the ACL/EACL'97 Workshop on Spoken Language Translation*. Madrid, 1997.

(Sumita *et al.* 90) E. Sumita, H. Iida, and H. Kohyama. Translating with Examples: A New Approach to Machine Translation. *Proceedings of the 3th TMIMT*, Austin, Texas, 1990.

(Ueda 86) K. Ueda. Making Exhaustive Search Programs Deterministic. In *Proceedings of the 3rd International Conference on Logic Programming*, London, 1986.

(Vauquois 75) B. Vauquois. *La Traduction Automatique á Grenoble*. Dunod, Paris, 1975.

(Wahlster 93) W. Wahlster. Verbmobil — Translation of Face-to-Face Dialogs. In *Proceedings of the 4th MT Summit*, Kobe, Japan, 1993.

(Wheeler 87) P.J. Wheeler. Systran. In M. King, editor, *Machine Translation Today: The State of the Art*. Edinburgh University Press, 1987.

(Whitelock 92) P. Whitelock. Shake-and-Bake Translation. In *Proceedings of the 14th COLING*, Nantes, France, 1992.