COMPLEX VERB TRANSFER PHENOMENA IN THE SLT SYSTEM

Björn Gambäck Ivan Bretan[1]
gamback@sics.se ivan@sics.se
Natural Language Processing Group
Swedish Institute of Computer Science
Box 1263, S - 164 28 Kista, Stockholm, Sweden

Abstract

The paper discusses several complex transfer problems and their prospective solutions within an English-to-Swedish spoken language translation system. The emphasis in the text is on transfer problems which are not lexically triggered, concentrating mainly on the translation of differences in mood and tense. Laying the groundworks for the translation part, the treatment of verb-phrase syntax and semantics is described in detail. The paper also shortly discusses some lexically triggered complex transfer problems.

1 Introduction

When developing a system aimed at translating between two natural languages, several problems soon emerge. Some of these are specific to the translation methodology chosen or the language-pair at hand, while others can be generalized and tend to occur in any translation system. Many of the problems of the general kind are rooted in the verb-phrases; phrases that not only show a rich cross-linguistic variety, but also are of vital importance for the interpretation of utterances in almost any language. In this paper we will discuss some of these problematic transfer phenomena, mainly concentrating on those that arise as an effect of tense and mood differences.

The problems concerning complex transfer may in general be split into two quite different categories, those that are and those that are not lexically triggered. The first group contains such phenomena as argument- and head switching, object raising, and changes from one syntactic category to another. The second group includes problems with changes in aspect, tense, mood, and determination.

The question of lexically triggered complex transfer problems is not the main topic of this paper — a nice overview of such problems as having surfaced in several other systems can be found in (Pulman (ed.) 1991) — although some such problems will be discussed in Section 4. The framework for our discussion there, as well as in the entire paper as such, is the transfer component of a spoken-language English-to-Swedish translation system called SLT, a short overview of which is the topic of Section 2. The transfer formalism used in the SLT system is the general point of discussion in Section 3.

The main problem addressed in the paper is that of translating verb-phrases. The syntactic and semantic treatment of these in our grammars is thus the topic of Section 5. The discussion in that section is mainly centered around the treatment of tense and mood; our solution to translating these phenomena is then described in Section 6. Finally, Section 7 sums up the previous discussion and points to some areas of future research.

'Thanks to Jussi Karlgren, Manny Rayner, Christer Samuelsson and Mats Wirén for valuable comments.

2 The SLT system

The Spoken Language Translator (SLT) is a system prototype which can translate queries from spoken English to spoken Swedish in the domain of air travel planning (ATIS). The system was developed as a joint effort by the Swedish Institute of Computer Science, SRI International (Menlo Park, US and Cambridge, UK), and Telia Research AB (Haninge, Sweden). Most of it was constructed from previously existing pieces of software, which were adapted for use in the speech translation task with as few changes as possible. The overall SLT system is described shortly in this section, for a complete description see (Rayner *et al* 1993) or (Agnäs *et al* 1994).

The main components of the SLT system are connected together in a pipelined sequence as follows. The input signal is processed by SRI Menlo Park's DECIPHER(TM), a speaker-independent continuous speech recognition system based on Hidden Markov Model technology. It produces a set of speech hypotheses which is passed to the English-language processor, the SRI Cambridge Core Language Engine, CLE (Alshawi (ed.) 1992), a general natural-language processing system which is based completely on unification and has a reversible phrase-structure type grammar.

The CLE grammar associates each speech hypothesis with a set of possible quasi-logical forms, QLFs, typically producing 5 to 50 QLFs per hypothesis. A preference component is then used to give each of them a numerical score reflecting its linguistic plausibility. When the preference component has made its choice, the highest-scoring logical form is passed to the transfer component, which uses a set of simple non-deterministic recursive pattern-matching rules (described in detail below) to rewrite it into a set of possible corresponding Swedish representations.

The preference component is now invoked again, to select the most plausible transferred logical form. The result is fed to a second copy of the CLE, which uses a Swedish-language grammar and lexicon developed at SICS (Gambäck & Rayner 1992) to convert the form into a Swedish string and an associated syntax tree. Finally, the string and tree are passed to the Telia Prophon speech synthesizer, which utilizes polyphone synthesis to produce the spoken Swedish utterance.

The SLT system's current performance figures measured on previously unseen data (the 1001-utterance December 1993 ATIS corpus) are: 78.8% of all utterances are such that the top-scoring speech hypothesis is an acceptable one. If the speech hypothesis is correct, then an acceptable translation is produced in 68.3% of the cases and the overall performance of the system is 53.8%. Limiting the test corpus to sentences of 10 words or less (688 utterances), these figures move up to 83.9% for speech recognition and 74.2% for language processing, with a 62.2% overall performance. For about 10% of the correctly recognized utterances, an unacceptable translation is produced. Nearly all of these are incorrect due to their containing errors in grammar or naturalness of expression, with errors due to divergence in meaning between the source and target sentences accounting for less than 1% of all translations. SLT performance is discussed at length in (Rayner *et al* 1994).

3 QLF-based transfer

In this section we describe the transfer component of the SLT system, starting out by describing the transfer formalism used and then discussing the formalism's adequacy for the task with respect to its ability to deal with complex transfer phenomena. The two copies of the CLE had prior to the SLT project been used together to form a machine translation system called the Bilingual Conversation Interpreter, BCI (Alshawi *et al* 1991). Several transfer problems ensued in that project. Some of these will be also be discussed in this section.

The QLF transfer framework has basically been adopted unchanged from the BCI project and will thus only be described briefly here. Unification-based QLF transfer is based on the notion of compositionally translating a QLF of the source language to a QLF of the target language, through matching QLF fragments against QLF pair patterns. By means of these patterns, transfer rules are specified declaratively using the following format:

trule(<Comment>, <QLF pattern 1> <0perator> <QLF pattern 2>).

The left hand side of a rule (QLF pattern 1) matches a fragment of the source language QLF and the right hand side the corresponding target QLF. The patterns can contain (possibly constrained) variables that match QLF fragments of arbitrary size, but no additional conditions can be associated with the rules. The resulting formalism is therefore kept very simple, the main idea that only cross-linguistic data should be specified in QLF transfer. The particulars of how to form a QLF which corresponds to a grammatical sentence in a language is monolingual knowledge, and best left to the grammars, statistical preferences and lexica.

This fairly simple transfer rule formalism allows for succinct formulation of rules that deal with mappings between phrases that vary significantly in surface syntactic realization. The reason for this is the high level of abstraction in the QLF representation with respect to features essential for translation such as predicate-argument structure, mood, tense, and aspect, while enough structural information is kept to enable generation of a surface structure faithful to the original formulation.

4 Lexically triggered complex transfer

Given the fact that we have decided on a particular intermediate representation which will abstract over certain grammatical features, it is important to assess what differences between source and target language expressions the formalism requires non-trivial (non-atomic) transfer rules to handle. We should expect a need for such rules where the grammatical structures of the two expressions are fundamentally different, and where this difference is localized to specific words or phrases. A typical example of such a difference is variation in argument structure. Such types of discrepancies between source and target language are well-known in the machine translation field, and besides difference due to purely idiomatic expressions, include the phenomena exemplified in Table 1.

Table 1: Examples of complex transfer phenomena

Complex transfer type	English-to-German/Swedish example	
Argument switching	John **likes** Mary	Mary **gefällt** John
Head switching	John **likes** swimming	John **schwimmt** gem
Object raising	John **wants** Mary **to** go	John **vill att** Mary skall åka
Passive to active	Insurance **is included**	Forsakring **ingar**
Verb to adjective	John **owes** Mary $20	John **ar skyldig** Mary $20

As discussed in (Gambäck *et al* 1991), our transfer formalism handles most of these, some others, and most importantly, the *interaction* between different complex transfer phenomena; however, for two of the ones given in Table 1 the formalism at present seems to fail to provide enough expressive power. Neither object raising nor head switching can currently be solved without adding rules very specific to the translation of particular sentences. This will be further discussed below (Section 7).

5 Verb-phrases in the grammar

We will now describe the way verb-phrases are treated in the grammars of SLT. The most important differences between English and Swedish syntax are to be found within in the verb-phrases; however, in this section we will see that this difference of main verb syntax can be factored away in the QLF-based semantics, while some auxiliary verb cases still cause some problems.

The treatment of verb-phrase semantics in the CLE follows the strategy suggested by (Alshawi & Crouch 1992): information about tense, modality, etc., is packaged declaratively in the compositional semantics. This information could then be unpacked later on to determine the implicit points in time, etc., not shown in the surface form of the sentence. For a translation system like the SLT, this is not necessarily important: the tense/aspect information can in many cases be translated as it is to the other language (see Section 6.1). Part of the treatment here of verb semantics is based on the fact that it indeed is going to be used in a translation system, so even though auxiliaries behave quite differently in English and Swedish, the semantic representations should be as close to each other as possible.

Looking at the main verb case first, we note that the syntax of verb-phrases of almost any kind in both English and Swedish can be treated in a simple and uniform way by a rule common to unification-based grammar approaches as the one in the CLE: we can remove the information regarding subcategorization schemes (i.e., the number and type of verbal complements, such as objects, particles, etc.) from the grammar by a rule schema such as

$$VP\langle\hat{V}\rangle \rightarrow V\langle\hat{V}\rangle\ (\textbf{subcat} = \textit{Complements})\quad \textit{Complements} \qquad (1)$$

where the notation is to be interpreted so that subcat is a feature of the verb having the value *Complements,* which in turn is unified with the rest of the verb-phrase. The value of subcat is specified for a particular verb in its lexical entry and can of course be empty (e.g., for intransitives). The semantic information is the one within angle brackets, so $\hat{V}$ by unification percolates the verb's semantic interpretation up to become the interpretation of the entire verb-phrase.

The semantic interpretations of the complements are simply unified into the verb's semantics in the lexicon. The lexical entry for e.g. a transitive particle verb can schematically be viewed as

$$V\langle\lambda\hat{x}.\hat{E}(\hat{x},\widehat{NP})\rangle\ (\textbf{subcat} = P\ NP\langle\widehat{NP}\rangle) \qquad (2)$$

where $\hat{E}$ represents the interpretation of the event as such, $\widehat{NP}$ the semantics of the object, and the lambda-abstraction is over the semantics of the (sought) subject.

5.1 Tense and related phenomena

Now we will discuss a range of problematic phenomena that in principle just have a few properties in common, namely that they are (to a certain extent) inherent parts of the verb sequence surface syntax or morphology, but in general have to be interpreted at a deeper semantic level to provide for a transfer representation at an interesting level of generalization, since as mentioned earlier, these phenomena are subject to significant surface syntactic cross-linguistic variation. Apart from "tense" and "mood", we will also include "aspect" and "voice" as belonging to the same broad type, while we (admittedly rather arbitrarily) will exclude for example "negation".

Auxiliaries that change the tense of the verb-phrase (e.g., to past as *hade,* "had", or future as *ska,* "shall") can be treated on a syntactic level with a rule just like the one above (rule 1), but must in the semantics be treated separately from the main-verb case. For main verbs, the tense information of the verb-phrase is the same as the one of the daughter verb and is simply unified up together with the other semantic information. For the auxiliary case, the semantic interpretation of the mother verb-phrase should still be the one of the daughter verb-phrase, but the tense should be taken from the auxiliary. Introducing t-subscripts for the tense information, we get:

$$VP_1(\hat{V}_{t_1}) \rightarrow V(\hat{\emptyset}_{t_1})\,(\mathbf{subcat} = VP_2) \quad VP_2(\hat{V}_{t_0}) \;. \tag{3}$$

with the $\hat{\emptyset}$ indicating that the auxiliary perform no other semantic function besides carrying the tense information (in t_1). The value t_0 for the tense of the daughter verb-phrase indicates that its tense will not influence that of the mother; however, it may, but this should be treated in the lexical entry for the auxiliary.

Modal auxiliaries can in English be treated by the rules already introduced, but the semantics of Swedish ones complicate the picture somewhat: we need to treat two cases, one for finite and one for non-finite (i.e., infinite plus supine) verb forms, the difference being that the former (unlike in English) can modify other modals as in a sentence like

Jag skulle vilja kunna flyga. "I would like to be able to fly"
(lit. "I should want could fly")

In examples like this one, finite modals behave quite a bit like tense auxiliaries; they do not affect the semantic content as such, but rather a modal operator *m* that parallels the tense operator *t* introduced in rule 3. How this is actually done is described in detail in (Gambäck 1993).

For the implementation of the system sketched out here, we represent the information about tense, mood, etc., in the compositional semantics as a functor

```
verb(Tense,Aspect,Action,Mood,Voice)
```

The arguments of which are in order: Tense past, present, or future; Aspect perfective or imperfective aspect; Action progressive or non-progressive action; Mood the speaker's view on the event as expressed with a modal, imperative, etc; Voice active or passive voice.

6 Verb-phrase transfer

Given the treatment of the verb-phrases in the grammars, this section goes into detail on how they are translated: the current version of the SLT transfer component has 144 atomic lexical transfer rules for verb constants, including constants representing multi-word phrases, such as verbs taking complex complements involving partitives or reflexives. An example is the mapping between *"ta om"* (lit. take over) and *"repeat":*

```
trule(lex(simple), repeat_SayAgain =< 'ta_Om_Något').
```

The verbs *"be"* and *"have"* are also translated as lexical constants. As semantically vague verbs, these two words can of course translate into a multitude of target words. In order to avoid combinatorial explosions, these contextually sensitive translations are deferred to structural rules, which

decreases compositionality, but increases efficiency. An example of this is the translation of *"be"* into *"finnas"* (to exist), which is governed by a number of structural rules. There are only two non-atomic lexical rules for verbs, namely *"arrive"* and *"depart",* which when used with a direct object, as in *"arrive Boston",* need to be translated into a VP with post-modifying PP, *"anlända till Boston":*

```
trule(semi_lex(complex, arrive_TurnUpAt_place-anlända_till_ställe),
  [[arrive_TurnUpAt,D,tr(subj),tr(place)]] >=
  [form(prep(till),_,F^[F,v(D),tr(place)]),[anlända_2p,D,tr(subj)]]).
```

There are 11 lexically triggered complex transfer rules that deal with specific modal verbs. Eight of them deal with uninteresting design differences between the treatment of modal verbs in Swedish and English. One handles the translation of the non-modal *"want",* as in *"I want a ticket",* into *"vill ha",* which is a modal followed by *"have"* as main verb.

```
trule(semi_lex(modal_intro, want_WishFor~'vill_ha_Något'),
  form(verb(X,Y,Z,no,W),A,B^[B,[want_WishFor,A,tr(term1),tr(term2)]],_) ==
  form(verb(X,Y,no,vill1,W),A,B^[B,['ha_Något',A,tr(term1),tr(term2)]],_)).
```

Another translates *"I would like a ticket"* (modal + main) into *"Jag skulle vilja ha en biljett"* (modal + modal + main). The last of these specific rules deals with translating *"I want to"* + VP into *"Jag vill"* + VP', which would not be a difficult translation were it not for the fact that *"want"* is classified as non-modal and *"vill"* as modal. This means that the subject of *"want"* needs to be "moved" to the main verb, i.e., a head switching operation as described in Section 4. Since this type of complex transfer cannot be specified generally, a number of specific rules were coded to cover the cases that occurred in ATIS.

6.1 Non-lexically triggered complex transfer

Now we turn our attention to some non-lexically triggered complex transfer phenomena. In the literature, these are not discussed as frequently as the lexically triggered ones (see Section 4). One reason for this is that interlingua-based MT system can avoid complex, restructure! transfer when treating things like tense, aspect and determination; however, the *representation* of such phenomena still is a difficult task which must be addressed in any MT system (or at least in any NLP system aiming at a deeper level of utterance interpretation).

Some difficult non-lexical problems like number differences in coordinate structures and the translation of anaphoric relationships are discussed in (Pulman (ed.) 1991), while (Gawrońska 1992) discusses the translation of aspect; here, we will concentrate on the translation of tense and mood.

In the SLT system, there are two atomic rules translating logical constants used to signal mood feature values:

```
trule(lex(simple),sai_do >= no).
trule(lex(simple),emphatically_do >= no).
```

The first eliminates subject-aux inversion, which is not marked in Swedish. The fact that a QLF represents a yes/no-question will be sufficient to generate the corresponding Swedish sentence through inversion of subject and verb. Emphatic *"do",* as in *"Do fly to Boston!"* does not

have a corresponding auxiliary in Swedish, and needs to be expressed through prosodic features, which are currently not supported by the QLF format. A structural rule deals with transferring a *to*-infinitive into a bare infinitive, which is needed when translating a non-modal into a modal.

Another structural rule suppresses the progressive marker for finite verb phrases in order to provide for translating e.g, *"I am flying to Boston"* into *"Jag flyger till Boston"* (I fly to Boston). A more complex structural rule maps between `NP + progressive-VP` and `NP + relative-clause` as in *"flights going to Boston" => "flygningar som gar till Boston"* (flights that go to Boston). The rule is necessary since Swedish lacks the possibility of marking the progressive aspect through the present participle in the same way as in English.

```
trule(struct(np_progressive_vp-np_relative_clause),
  [and,tr(head),form(verb(no,A,yes,M,D),V,tr(restr))] >=
  [and,tr(head),[island,form(verb(pres,A,no,M,D),V,tr(restr))]]).
```

No special rules were needed to account for tense or the active/passive distinction. However, it should be noted that future as expressed using *"will",* as in *"I will fly tomorrow",* should preferably be translated through the temporal verb *"komma att",* as in *"Jag kommer att flyga imorgon",* rather than using the direct translation of *"will", "skola",* which implies commitment.

7 Conclusions and future work

We have discussed some complex transfer problems relating mainly to differences of mood and tense. In the text we have outlined the treatment of these problems within the SLT system, an English-to-Swedish spoken language translation system. Such a discussion would be rather meaningless without a detailed knowledge of how the syntax and semantics of verb-phrases is implemented in the system at hand, thus Section 5 contained a description of the treatment of VPs within the SLT system grammars. The main emphasis in the paper has been on transfer problems which are not lexically triggered; however, we have also shortly discussed some lexically triggered complex transfer problems such as object raising and head switching.

As noted in Section 4, the current transfer rule formalism does not yet provide enough expressive power to handle two of the phenomena listed in Table 1: object raising and head switching. Object raising works fine in the English-to-Swedish direction, but not the opposite. It is possible to unify the object *"John"* with the subject of *"fly"* in *"Mary wants John to fly"* when translating that sentence into *"Mary vill att John ska flyga"* (lit. Mary wants that John should fly). However, translating from the Swedish sentence into English would require the reversal of a (λ-bound) variable substitution. In addition, it seems as though the transfer framework also needs an extension to be able to cope with head switching. Thus, the arguments of the source verb need to be "moved" to the proper, arbitrarily embedded position in the source QLF.

A possible solution to these problems is to extend the formalism with a mechanism to express structural changes; this would require a way of manipulating different significant parts of a QLF verb `form` without spelling out its structure in detail. When translating the head switching example of Table 1, the mechanism should allow for specifying that, in order to generate the German *"schwimmen"* VP, we need to translate the English verb predication for *"swim"* under the premise that the John `term` actually fills its subject slot, in practice "moving" the subject from *"like"* to *"swim"* (and this change of course needs to be reversible). Similar mechanisms have been used within comparable machine translation frameworks, e.g., in (Kinoshita *et al* 1992).

References

M-S. Agnäs, H. Alshawi, I. Bretan, D. Carter, K. Ceder, M. Collins, R. Crouch, V. Digalakis, B.Ekholm, B. Gambäck, J. Kaja, J. Karlgren, B. Lyberg, P. Price, S. Pulman, M. Rayner, C. Samuelsson, and Tomas Svensson. "Spoken Language Translator: First-Year Report". Joint Research Report R94:03 and CRC-043, SICS and SRI International, Stockholm, Sweden and Cambridge, England, January 1994.

H. Alshawi, editor. *The Core Language Engine.* The MIT Press, Cambridge, Massachusetts, 1992.

H. Alshawi and R. Crouch. "Monotonic Semantic Interpretation". In *Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics,* pp. 32-39, Newark, Delaware, 1992.

H. Alshawi, D. Carter, B. Gambäck, and M. Rayner. "Translation by Quasi Logical Form Transfer". In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics,* pp. 161-168, University of California, Berkeley, California, 1991.

B. Gambäck. "Towards a Uniform Treatment of Swedish Verb Syntax and Semantics". In *Proceedings of the 16th Scandinavian Conference of Linguistics and the 8th Conference of Nordic and General Linguistics,* University of Gothenburg, Gothenburg, Sweden, 1993.

B. Gambäck and M. Rayner. "The Swedish Core Language Engine". In *Papers from the 3rd Nordic Conference on Text Comprehension in Man and Machine,* pp. 71-85, Linköping University, Linköping, Sweden, 1992. Also available as SICS Research Report, R92013, Stockholm, Sweden.

B. Gambäck, H. Alshawi, D. Carter, and M. Rayner. "Measuring Compositionality in Transfer-Based Machine Translation Systems". In *Natural Language Processing Systems Evaluation Workshop,* pp. 141-145, University of California, Berkeley, California, 1991.

B. Gawrońska. "Aspect — A Problem for MT". In *Proceedings the 14th International Conference on Computational Linguistics,* volume 2, pp. 652-657, Nantes, France, 1992.

S. Kinoshita, J. Phillips, and J. Tsujii. "Interaction between Structural Changes in Machine Translation". In *Proceedings the 14th International Conference on Computational Linguistics,* volume 2, pp. 679-685, Nantes, France, 1992.

S. Pulman, editor. *Eurotra ET6/1: Rule Formalism and Virtual Machine Design Study.* Commission of the European Communities, 1991.

M. Rayner, H. Alshawi, I. Bretan, D. Carter, V. Digalakis, B. Gambäck, J. Kaja, J. Karlgren, B. Lyberg, S. Pulman, P. Price, and C. Samuelsson. "A Speech to Speech Translation System Built from Standard Components". In *Proceedings of the Workshop on Human Language Technology,* Princeton, New Jersey, 1993. ARPA, Morgan Kaufmann.

M. Rayner, D. Carter, P. Price, and B. Lyberg. "Estimating Performance of Pipelined Spoken Language Translation Systems". In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing,* Kyoto, Japan, 1994.