A Bilingual Grammar for Translation of English-Swedish Verb Frame Divergences

Sara Stymne and Lars Ahrenberg

Department of Computer and Information Science
Linköpings universitet, SE-58183 Linköping (Sweden)

{sarst | lah}@ida.liu.se

Abstract

We describe a bilingual grammar used for translation of verb frame divergences between Swedish and English. The grammar is used both for analysis and generation with Minimal Recursion Semantics as interlingua. Our grammar is based on the DELPH-IN resources for which semantic transfer is proposed for MT. We show that an interlingua strategy based on a bilingual grammar can handle many cases of verb frame divergences minimising the need of transfer.

1 Introduction

Translation via semantic representations of source language input is a common approach in research, although less frequent in commercial systems, where semantic distinctions tend to be localised at the word level and motivated mostly by practical necessity. With the advent of grammars that express relations between surface strings and semantic representations for a large part of the constructions of a language, such as the ERG (Flickinger, 2000) and the JACY grammar (Siegel, 2000), the idea of performing practical translation on a coupling of general parsers and generators for large formal language descriptions using a common semantic framework seems less esoteric. A recent example of this approach is the LOGON project (Oepen et al., 2004).

Just as in LOGON, Minimal Recursion Semantics representations (MRS; Copestake, Flickinger, Sag, & Pollard, 2003) are used as interface structures. However, unlike the LOGON architecture, which uses different grammars based on different formalisms and linguistic theories, we see it as an advantage to use the same grammatical framework for both source and target languages, as this means that the same parser and generator can be used throughout. We use HPSG-like Typed Feature Structure Grammars as our framework because of the availability of the LKB workbench (Copestake, 2001) and the store of type definitions known as the Matrix (Bender, Flickinger, & Oepen, 2002), made available by the DELPH-IN collaboration (Bond, Oepen, Siegel, Copestake, & Flickinger, 2005). In addition, we see it as desirable that grammars are designed on similar principles, so that solutions to translation problems can be coordinated between languages and implemented using a shared inventory of types. This makes the system more homogenous and facilitates the addition of new languages.

We also believe that the development of practical applications benefits from the existence of a core system that provides a library of solutions to the translation problems that are likely to be encountered in application domains. The work reported here should be seen as a step towards such a core system that supports the development of application systems using English and Swedish.

Another desideratum is that semantic transfer rules should be restricted to the cases where they are absolutely necessary. Everything else being equal, we prefer MRS structures to pass unchanged from parser output to generator input. When this is

the case, we may actually view all strings, whether belonging to the source language or the target language, to be part of the same, bilingual grammar.

The specific goal of this work has been to investigate the possibility of handling a large number of verb frame divergences (VFDs) between Swedish and English in a bilingual HPSG grammar. We have considered a large number of divergences where verbs from the two languages differ syntactically and/or lexically, but where their semantics can be considered to be the same, i.e. where a common interlingual relation can be assumed.

We have found that a number of cases of VFDs can actually be treated in a bilingual grammar. Another result of this research is a taxonomy of VFDs with English-Swedish instances and an implemented bilingual grammar based on the Matrix. The results have equal application to other Scandinavian languages, and with modifications, to other Germanic language pairs as well.

In the following section we will give some examples of identified English-Swedish VFDs. In section 3 we will review related work and describe the semantic transfer approach that has previously been used with the DELPH-IN resources. Section 4 describes BiTSE, the bilingual grammar that is the core of our MT system. In section 5 we explain our treatment of several types of divergences. Section 6 contains a discussion on the merits and limits of our approach and section 7 contains the conclusion.

2 Verb Frame Divergences

As part of this study we have investigated verb frame divergences (VFDs). A verb frame consists of a verb and its arguments. A verb frame divergence is when two verb frames with the same meaning have different structures. Some examples of this, based on the categories suggested by Dorr (1994), will be presented here. All examples in this article are taken from the Europarl corpus (Koehn, 2005). Some of the examples shown contain more than one type of divergence.

(1) In that case, the matter turns out to be a national problem after all.

Till sist kommer ändock ärendet att visa sig vara ett nationellt problem.

(2) This appears to be the case with the events which Mr Lomas reports in his question.

Det tycks nämligen vara fallet med de fakta som herr Lomas för på tal i sin fråga.

(3) But that is precisely why we first need a clear strategy.

Just därför är till att börja med en klar strategi nödvändig här.

In (1) "turns out to be" corresponds to "visa sig vara" ("show itself be") which contains two structural divergences, when two logical constituents have different structure. In English a phrasal verb with the particle "out" is used and Swedish has a reflexive verb with the fake reflexive "sig". The verbal complement has an infinitive marker in English, but not in Swedish.

(2) contains a conflational divergence, the main verb "reports" in English corresponds to "för på tal" ("brings on speech") where the concept "speech" is conflated in English but explicit in Swedish.

An example of a categorial divergence can be seen in (3). Categorial divergences occur when semantically equivalent constituents have different syntactic categories. Here the English main verb "need" semantically corresponds to the Swedish adjective "nödvändig" ("necessary").

3 Related Work

Interlingual approaches to machine translation have been tried at least since the beginning of the sixties with much discussion and debate about the nature of interlinguas and the merits and drawbacks of interlingual approaches as compared to transfer approaches (e.g. Boitet, 1988; Nirenburg & Goldman, 1990). Basically a MRS relation is a place-holder for a concept with known argument structure which is associated with one or more linguistic expressions in a lexicon. Semantic relations such as hyponomy and antonomy, and even some semantic decomposition, could be added, but is not part of the current setup, though hyponymy

could be dealt with within the type system. Domain knowledge, as used in knowledge-based interlingual MT such as the KANT system (Mitamura, Nyberg, & Carbonell, 1991), is also not handled.

Similarities between two (or more) languages can be encoded in formal grammars in different ways. The Rosetta project (Rosetta, 1994) explored the idea of isomorphic grammars. Our framework does not require grammars to be isomorphic; the important thing is that they produce a common MRS for sentences that are translations of one another. In addition, the grammars are actually implemented as one bilingual (or multilingual) grammar, allowing types to be shared between languages.

3.1 MT using DELPH-IN resources

There have been previous suggestions for MT using the DELPH-IN resources, most of them using a semantic transfer strategy, but also an experimental multilingual grammar used as the core of a small MT system.

Copestake, Flickinger, Malouf, Riehemann, & Sag (1995) describe how MRS can be used for translation. They suggest a design that is based on semantic transfer using MRS. The transfer component works on MRS to produce output that the target grammar can accept. It is possible that the transfer component can output more than one form, some of which may be unacceptable by the generator. When several forms are output they will be ordered by a control mechanism that is distinct from both the transfer component and the generator.

The transfer component suggested by Copestake et al. (1995) is based on setting up symmetric and bidirectional transfer equivalences between each pair of languages. Their suggestion also allows interlingual predicates that are common for all languages such as negation.

A large-scale project where semantic transfer with MRS is used is LOGON, which focus on translation between Norwegian and English (Oepen et al., 2004). The main architecture is: analysis of Norwegian to MRS using the Norwegian LFG grammar NorGram (Dyvik, 1999), MRS transfer as described above, and generation to English using the ERG.

The LOGON system is unidirectional. It only translates from Norwegian to English, due to the design with different grammars for analysis and generation. However, Bond et al. (2005) notes that in the general MT design the HPSG grammar for each language is reversible and can be used both for parsing and generation. The transfer rules are also reversible, except for context and filter information in some cases.

Bond et al. (2005) discusses the open source resources for MT made available by the DELPH-IN collaboration and the general strategies used, including the basic ideas presented in this section. They also raise some proposals for future work, including "How much of the semantic representation can be shared between languages (and thus require little or no transfer)?" (p. 20).

A different MT architecture using Matrix-based grammars, based on a multilingual grammar has been suggested by Søgaard & Haugereid (2005). This design was an inspiration for our approach.

4 BiTSE - a bilingual grammar as core of MT

The core of our MT system is BiTSE, the Bilingual grammar for Translation between Swedish and English. Figure 1 shows the basic design of the system. The transfer module in Figure 1 is not currently part of our system, but is a possible extension for possible non-interlingual part of MRS structures.

BiTSE was developed using the Linguistic Knowledge Builder (LKB; Copestake, 2001) and the parser and generator of LKB are used when running BiTSE for MT. The coverage of BiTSE is currently the core of the languages and some VFDs, including basic verb and noun phrases, some adjectival and prepositional modifiers, phrasal verbs, fake reflexives, polar question and main and subordinate clause word order. The lexicon is small and basically includes one representative lexical item for each type of verb considered.

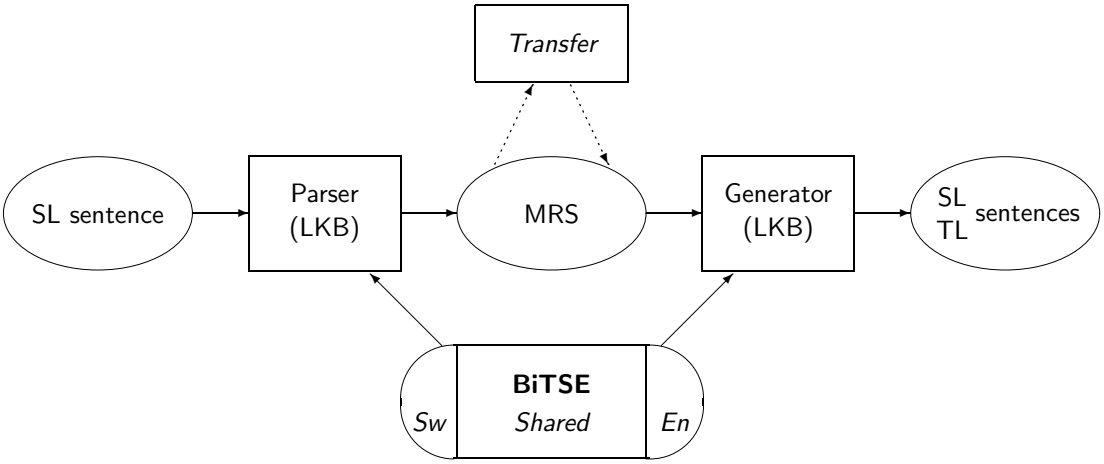

Figure 1: The design of the MT system with BiTSE as its knowledge source. The transfer module is not currently part of the system but could be added and transfer the non-interlingual part of the MRS structure.

BiTSE is based on the LinGO Grammar Matrix (Bender et al., 2002), a cross-linguistic starter-kit for HPSG grammars providing a store of grammatical and lexical types with MRS as the semantic representation. As illustrated in (4) for the sentence "The big dog sleeps", a MRS is a tuple containing a top handle (h1), an instance or event variable (e2), a bag of elementary predications and a bag of handle constraints (qeq).

(4)
```
<h1,e2,
 {h3:def_q(x4,h5,h6),
  h7:big(e8,x4),
  h7:dog(x4),
  h9:sleep(e2,x4),
  h1:proposition(h10)},
 {h5 qeq h7, h10 qeq h9}>
```

A MRS structure can be scope-resolved in one or several ways by equating all handle arguments and the top handle with a handle from a relation respecting all handle constraints, forming a tree. See Copestake et al. (2003) for a more detailed description of MRS.

4.1 Constructions for language

To include more than one language in a grammar, a feature that constrains the language of signs had to be added in addition to the basics of the Matrix. Following Søgaard & Haugereid (2005) a feature for language was added to SIGNs. Constraints were then added on all rules to make them work on a single language at the time, and for language-specific rules to work on only one of the languages. Figure 2 shows two of these types.

```
binary-lang-agree-phrase :=
 binary-headed-phrase &
  [ LANG #lang,
    HEAD-DTR.LANG #lang,
    NON-HEAD-DTR.LANG #lang ].

swedish-only-rule := headed-phrase &
  [ LANG #sw,
    HEAD-DTR.LANG #sw & sw ].
```

Figure 2: Types for language handling

As for Søgaard & Haugereid (2005) LANGUAGE is not a semantic feature, which makes the semantic representation language-independent, resulting in generation giving all equivalent sentences in both languages. Thus, the MRS in (4) will generate both the English "The big dog sleeps" and the Swedish "Den stora hunden sover". In order for the MRS to be language-independent all equivalent relations must have the same names. To achieve this all relations have English names, such that each English word generally has a relation with the same name as the word, and each Swedish word has a relation with the corresponding English name.

4.2 Sharing types between languages

An advantage of this grammar design is that it allows the parts of the grammar that are the same to be shared. This avoids the redundancy of entering the same information twice in a two grammar MT system. The shared information can make up a considerable part of the grammar, at least for related languages like Swedish and English. As Table 1 shows, the shared part of BiTSE contains more than half of the grammar. The fact that the Swedish part have nearly double the amount of types compared to English is largely due to 27 types for declination and conjugation of nouns and verbs in Swedish, which are not needed in English. The coverage for phrasal verbs is also larger for Swedish.

Table 1: Size of the different parts of BiTSE

	No. of types
Shared	188
Swedish	76
English	32

The number of language-specific verb lexemes is higher than for other word classes, mostly because verbs are the current focus of BiTSE, and thus are more specialised than other word classes. It could be expected that if BiTSE were to grow, the percentage of types that are shared might decrease.

This grammar design can be seen as an extension of the Matrix for two languages in this case, but possibly for a larger group of languages. We have also tried this principle out by adopting the Norwegian grammar NorSource (Hellan & Haugereid, 2003) to Swedish, which showed that only small modifications were needed.

5 Treatment of verb frame divergences

In this section we describe how some verb frame divergences are handled in BiTSE.

5.1 Structural divergences

Structural divergences are very common between Swedish and English. They occur when constituents that are logically equivalent in two languages have different structures. We have found the following four types:

- prep. complement vs. NP object:
  talking about quality – diskuterar kvaliteten
- refl. verb vs. plain verb:
  uttalat sig – spoke
- phrasal verb vs. plain verb
  gått ut – expired
- infinitive + marker vs. plain infinitive
  needs to undergo – behöver underkastas

Combinations of these divergences are also common. In all these cases there are one more word in one language than the other. The general solution is to treat one of these words as empty, i.e. carrying no semantics, and let the other carry all semantic information. Verbs then specify which empty constituents it needs as complements. Sigurd (1995) suggests a solution for particles, reflexives and prepositions in the Swetra Referent Grammar MT system based on the same principle.

The Matrix does not provide good support for empty complements, so a number of basic types for this were incorporated into BiTSE. The existing Matrix types for words with complements do not give correct semantics to empty complements, which should receive no semantic bindings at all. Swedish verbs can have up to two empty complements, as in (5), which shows all four types of structural divergences: empty reflexive and particle, and a prepositional complement with an empty preposition and an empty infinitive marker "att" in the prepositional complement.

(5) know how you can escape into Europe

veta hur man bär sig åt
know how one carries oneself PART
för att fly till Europa
for to flee to Europe

Besides empty complements verbs can of course also have complements that carry semantics, which we call *contentive* constituents. Thus new types for different combinations of empty and contentive complements in different order were needed. As an example Figure 3 shows the BiTSE base type for a verb with one empty complement, like "uttala sig" ("express oneself"). Only the first argument, the subject, is mapped as an argument of the verbal relation. The second argument, the empty one, gives no semantic contribution.

```
intrans-empty2ndarg-lex-item :=
  basic-two-arg &
   [ ARG-ST < [ LOCAL.CONT.HOOK.INDEX
                  ref-ind & #ind ],
              synsem >,
     SYNSEM.LKEYS.KEYREL [ ARG1 #ind ]].
```

Figure 3: Type for intransitive verb with an empty complement

Empty prepositions and infinitive markers are handled as empty syntactic heads of the phrase that is later chosen by verbs as complements.

As an example we will show in some more detail how fake reflexive pronouns are handled.

5.1.1 Fake reflexive pronouns

Verbs can occur with reflexive pronouns of two types: *fake reflexives*, which obligatory occur with reflexive verb*s*, as in "I perjure myself", and *ordinary reflexives*, which occur as objects to ordinary transitive verbs, as in "I shave myself".

The category that takes part in structural divergences is the fake reflexives. They are analysed as semantically empty, with the reflexive verb carrying the relation for the meaning of the verb plus the fake reflexive and selecting the correct reflexive pronoun. Thus they get the same MRS as a non-reflexive verb with the equivalent meaning.

The type for intransitive reflexive verbs like "perjure oneself", shown in Figure 4, inherits from two BiTSE types, one for verbs in general, and one for the type for intransitive verbs with an empty complement shown in Figure 3. The *png* (person, number and gender) value of the reflexive is co-indexed with that of the subject, to ensure agreement.

```
intrans-refl-comp-verb-lex := ord-verb-lex &
             intrans-empty2ndarg-lex-item &
 [ SYNSEM.LOCAL [ CAT.VAL
       [ SPR <  >,
         SUBJ < #subj  >,
         COMPS < #refl & [ OPT - ] >,
         SPEC < > ] ],
   ARG-ST < #subj & [ LOCAL
             [ CONT.HOOK.INDEX [ PNG #p ],
               CAT np & [ HEAD.CASE nom ]]],
            #refl & [ LOCAL
               refl-0-local &
               [ CONT.HOOK.INDEX.PNG #p ]]>].
```

Figure 4: Type for intransitive reflexive verbs

This treatment ensures that the same interlingual relation can be used for a plain verb as for a reflexive verb.

5.2 Conflational divergences

Conflational divergences occurs when an argument that is explicit in one language is implicit, or conflated, in the other language, such as "report"/"föra på tal" ("bring on speech"). Sometimes arguments can be optionally conflated in one language, as

(6) I shave [myself]

Jag rakar mig

which is implicitly reflexive if the object is left out in English. In Swedish it is not possible to leave the object out for this type of verb, which causes a divergence.

There are two features on SYNSEMs to handle optionality: a boolean feature OPT, which is used to mark SYNSEMs as optional and OPTTYPE which describes which relation should be inserted in place of the removed optional argument. The default value for OPTTYPE is *unspec*, which means that removing the optional complement should result in it being left unspecified, as for "I eat". For this type of optionality there is a rule that simply removes the object from the verb's valence list if it is not present.

For verbs like "shave" where removal of the object should result in a reflexive re-

lation being added, the object has OPTTYPE *refl-opt*. There is also a unary phrasal rule that adds a relation for a reflexive pronoun when removing an optional complement. This rule is constrained to work only for OPTTYPE *refl-opt*. It further assures that the added reflexive pronoun relation agrees with the subject on person, number and gender.

5.3 Head-inversion divergences

Head-inversion occurs when a main verb in one language corresponds to another constituent, usually an adverb, in the other language. An example of this is the Swedish raising verb "brukar" which corresponds to the English scopal adverb "usually". The standard HPSG analysis for these two types of constituents based on Matrix types assign them similar semantics:

(7) "Bob brukar sova"

```
<h1,e2,
 {h3:named(x4, ‘‘Bob’’),
  h5:def_q(x4,h6,h7),
  h8:brukar(e2,h9),
  h9:sleep(e10,x4),
  h1:proposition(h11)},
 {h6 qeq h3, h11 qeq h8}>
```

(8) "Bob usually sleeps"

```
<h1,e2,
 {h3:named(x4, ‘‘Bob’’,)
  h5:def_q(x4,h6,h7),
  h8:usually(e2,h9),
  h10:sleep(e11,x4),
  h1:proposition(h12)},
 {h6 qeq h3, h9 qeq h10,
  h12 qeq h8}>
```

The only difference between these two MRS structures is that "brukar" has "sleep" directly as an argument and "usually" has it via a qeq-relation.

These structures are both underspecified. The one with the scopal adverb have two scope-resolved versions: (9), which is equivalent to the one scope-resolved version of the raising verb structure, and (10).

(9) (proposition(def_q(x4, named(x4, Bob), usually(e2, sleep(e11, x4)))))

(10) (proposition(usually(e2, def_q(x4, named(x4, Bob), sleep(e11, x4)))))

Even though both these readings can be considered semantically correct, we believe that it is not necessary to underspecify scope in such a precise way in a grammar for Swedish–English MT. Swedish and English are very similar with regard to scope underspecification, and thus we believe it suffices to choose one of the two possible readings in cases like above. It is then possible to give scopal adverbs like "usually" the same semantics as raising verbs like "brukar", resulting in equivalent MRS structures. Even though we have not been able to identify any cases where this type of underspecification makes a difference to MT we do not rule out that there might be some rare cases where it does.

5.4 Syntactic divergences

Syntactic divergences occurs when synonymous verbs have different argument frames. One example of this is divergences that occur because of dative alternations. In both languages the dative object can be either a noun phrase or a prepositional object but the distribution is different.

(11) I tell him a story

I tell a story to him

(12) *Jag berättar honom en historia

Jag berättar en historia för honom

(11) and (12) shows the possible alternations for the verb "tell"/"berätta", which has two possible patterns in English and only one in Swedish. This translation divergence is actually present within one language as well, since the two English sentences in (11) are equivalent and should have the same MRS in English.

In our approach no other treatment of this divergence is needed than that which is anyway needed within one language. In this case the strategy to handle dative alternations can also be shared between English and Swedish, which further eliminates redundant representations.

6 Discussion

The mechanisms used to solve the problems of translating VFDs we have illustrated above, such as an independent level of semantic representation, semantically empty words, and constraints on subcategorization, were to a large extent already available, or potentially available, in the monolingual framework. And, of course, a major reason for choosing an interlingual, "deep grammar" approach to translation has always been that translation in such a framework comes for free.

However, there are several problems of VFD-translation that remain to be treated. Some of them, such as categorial divergencies are discussed in Stymne (in press). A more general problem is given by translations that do not use synonyms. For example, the English verb "put" is generally translated by Swedish verbs with a more specific meaning, such as "ställa" (cause to stand somewhere), "sätta" (cause to sit somewhere) and "lägga" (cause to lie somewhere). For such cases we note that the translation relation must not be taken as transitive, i.e., Swedish sentences such as

(13) Hon ställde vasen i lådan

(14) Hon lade vasen i lådan

must not be treated as equivalent, although an English sentence such as "She put the vase in the box" may be used to translate both of them. For this to be possible we must distinguish the translation relation from the synonymy relation and allow semantic relations between concepts of the interlingua to be defined and utilized in mappings of MRS representations. This goes well beyond what the Matrix framework currently allows.

7 Conclusion

We have shown that translation of VFDs that are handled by semantic transfer in the general DELPH-IN MT design could instead naturally be handled by an interlingual design in many cases, minimising the need of transfer.

The work has also produced BiTSE, a bilingual grammar of Swedish and English, covering basic phrase constructions and a number of VFDs. In this grammar more than half the types are common for the two languages, which shows that a bilingual grammar design reduces the redundancy that occurs in two separate grammars.

References

Bender, E. M., Flickinger, D., & Oepen, S. (2002). The Grammar Matrix: An open source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th Conference on Computational Linguistics* (pp. 8–14). Taipei, Taiwan.

Boitet, C. (1988). Pros and cons of the pivot and transfer approaches in multilingual machine translation. In S. Nirenburg, H. Somers, & Y. Wilks (Eds.), *Readings in machine translation* (pp. 273–279). Cambridge, MA: MIT Press. (Reprinted from D. Maxwell, K. Schubert, T. Witkam (Eds.), 1988, *Recent Developments in Machine Translation*, Dordrecht, The Netherlands: Foris)

Bond, F., Oepen, S., Siegel, M., Copestake, A., & Flickinger, D. (2005). Open source machine translation with DELPH-IN. In *Proceedings of the Open-Source Machine Translation Workshop at MT Summit X* (pp. 15–22). Phuket, Thailand.

Copestake, A. (2001). *Implementing typed feature structure grammars.* Stanford, CA: CSLI Publications.

Copestake, A., Flickinger, D., Malouf, R., Riehemann, S., & Sag, I. (1995). Translation using Minimal Recursion Semantics. In *Proceedings of the Sixth International Conference on Theoretical and Methodological Issues in Machine translation, TMI-95.* Leuven, Belgium.

Copestake, A., Flickinger, D., Sag, I., & Pollard, C. (2003). Minimal Recursion Semantics: An introduction. *Language and Computation*, *1*(3), 1–47.

Dorr, B. J. (1994). Machine translation divergences: A formal description and pro-

posed solution. *Computational Linguistics*, *20*(4), 597–633.

Dyvik, H. (1999). The universality of f-structure: Discovery or stipulation? The case of modals. In *Proceedings of the 4th International Lexical Functional Grammar Conference.* Manchester, UK.

Flickinger, D. (2000). On building a more effcient grammar by exploiting types. *Natural Language Engineering (Special Issue on Efficient Processing with HPSG)*, *6*(1), 15–28.

Hellan, L., & Haugereid, P. (2003). Norsource - an excercise in the matrix grammar building design. In E. M. Bender, D. Flickinger, F. Fouvry, & M. Siegel (Eds.), *Proceedings of the Workshop on Ideas and Strategies for Multilingual Grammar Development, ESSLLI 2003* (pp. 41–48). Vienna, Austria.

Koehn, P. (2005). Europarl: A parallel corpus for statistical machine translation. In *Proceedings of MT Summit X* (pp. 79–86). Phuket, Thailand.

Mitamura, T., Nyberg, E. H., & Carbonell, J. G. (1991). An efficient interlingua translation system for multi-lingual document production. In *Proceedings of the Third Machine Translation Summit.* Washington, DC.

Nirenburg, S., & Goldman, K. (1990). Treatment of meaning in MT systems. In S. Nirenburg, H. Somers, & Y. Wilks (Eds.), *Readings in machine translation* (pp. 281–293). Cambridge, MA: MIT Press. (Reprinted from *Proceedings of the Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages* (pp. 15–22). Austin, TX)

Nirenburg, S., Somers, H., & Wilks, Y. (Eds.). (2003). *Readings in machine translation.* Cambridge, MA: MIT Press.

Oepen, S., Dyvik, H., Lønning, J. T., Velldal, E., Beermann, D., Carroll, J., Flickinger, D., Hellan, L., Johannessen, J. B., Meurer, P., Nordgård, T., & Rosén, V. (2004). Som å kapp-ete med trollet? Towards MRS-based Norwegian–English machine translation. In *Proceedings of the 10th International Conference on Theoretical and Methodological issues in Machine Translation* (pp. 11–20). Baltimore, MD.

Rosetta, M. T. (1994). *Compositional translation.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Siegel, M. (2000). HPSG analysis of Japanese. In W. Wahlster (Ed.), *Verbmobil: Foundations of Speech-to-Speech Translation* (pp. 264–279). Berlin, Germany: Springer.

Sigurd, B. (1995). Analysis of particle verbs for automatic translation. *Nordic Journal of Linguistics*, *18*, 55–65.

Søgaard, A., & Haugereid, P. (2005). *The noun phrase in mainland Scandinavian.* Presented at the 3rd meeting of the Scandinavian Network of Grammar Engineering and Machine Translation, Gothenburg, Sweden. (Retrieved April 28, 2006, from `http://www.cst.dk/anders/publ/gothenburg04.pdf`)

Stymne, S. (in press). *Swedish-English verb frame divergences in a bilingual Head-driven Phrase Structure Grammar for machine translation.* Master's thesis, Linköpings universitet, Linköping, Sweden.