

Analysis of Translational Correspondence in view of Sub-sentential Alignment

Lieve Macken

LT³ Language and Translation Technology Team
School of Translation Studies, Hogeschool Gent
Belgium

Lieve.Macken@hogent.be

Abstract

This paper reports on the first results of an empirical study of translational correspondence in different text types for the English-Dutch language pair. A Gold Standard was created, which can be used as a standard data set for evaluating sub-sentential alignment. The manually indicated translational correspondences were analyzed in view of different heuristics used in existing sub-sentential alignment modules.

1 Introduction

The distinction between literal and free translation is age-old. Literal and free translations can be seen as the opposite ends of a continuum, and most translations are situated in between the two extremes. The most appropriate translation method for a given text is determined by a.o. the type of the source text, the purpose of the source text, the target audience and the purpose of the translation.

For certain text types, the translation typically follows more closely the forms of the source texts, while for other text types, the translation is more oriented towards the target language, and aims at using more natural phrasings and idioms of the target language.

Therefore, in order to study the problem of translational correspondence, an English-Dutch corpus was compiled, consisting of three different text types: user manuals, press releases and proceedings of plenary debates. We assume that for each of the three text types another translation style was adopted, with the user manuals

being the most literal and the proceedings being the most free translations.

This paper reports on the first results of an empirical study of translational correspondences in different text types. The purpose of this paper is to examine the problem of translational correspondence in view of different heuristics used in existing sub-sentential alignment modules.

Sub-sentential alignment is used a.o. in statistical phrase-based machine translation and example-based machine translation. As sub-sentential alignment is mostly part of the MT-system (e.g. as phrase translation table), its performance is measured indirectly, i.e. in terms of the performance of the MT system.

A stand-alone sub-sentential alignment module however, is also useful for human translators if incorporated in CAT-tools, e.g. sophisticated concordance programs, or in sub-sentential translation memory systems (Gotti et al., 2005).

We assume that, for literal translations, the problem of translational correspondence is more a problem of combining word-by-word correspondences, while for freer translations, the problem is more complex. Hence, depending on the text type, the task of a sub-sentential alignment module will be relatively straightforward or rather complicated.

The remainder of this paper is organized as follows: Section 2 describes the corpus used. In section 3 we give more details on the creation of the gold standard and the annotation guidelines and we describe how inter-annotator reliability was assessed. In section 4, the translational correspondences of the manual reference corpus are analyzed. Section 5 concludes the paper and section 6 discusses future work.

2 Corpus

To examine the problem of translational correspondence in different text types, we compiled an English-Dutch corpus of three different text types. Table 1 summarizes the formal characteristics of the corpus.

2.1 User Manuals

The user manuals come from three different companies: SAP¹, IBM and Philips. These technical documents are typically written for an end user. The texts are characterized by a high percentage of technical terms. The sentences are short, with an average of 13 words per sentence. Mathematically, there is a perfect 1:1-correspondence between source and target sentences.

2.2 Press Releases

The press releases were provided by Barco, a Belgian high-tech company in the visualization industry. The texts were selected from the company's archive of press releases based on content: all texts describe the use of the company's technology on worldwide events, or announce new partnerships. Due to the nature of the company, the press releases contained a high percentage of technical terms. The texts were written for investors.

The average sentence length is 24.6 words for the English source texts and 23.7 words for the Dutch target texts. The ratio source/target sentences is 0.97, which means that only a few source sentences are translated by two or more target sentences.

2.3 Proceedings of plenary debates

The Europarl corpus² (Koehn, 2005) contains the proceedings of the European Parliament plenary debates in each of the eleven official languages of the European Union. The target audience of the proceedings are all European citizens. The texts are a written reproduction of spoken language.

Every speech is identified by a language tag and a speaker tag. Based on the language tag, it was possible to extract all *direct* translations from English to Dutch. As we preferred English source sentences produced by native English

speakers, the speaker identity tag was matched against a compiled list of British and Irish Members of Parliament. The result was an English-Dutch subcorpus of 6 million words. From this subcorpus, the texts were selected from the debates in the days following the 9/11-attacks of 2001.

The English source sentences are relatively long, with an average sentence length of 24.8 words. The ratio source/target sentences is 0.78, which means that a considerable number of source sentences are translated by two or more target sentences. This is also reflected in the lower average sentence length of the Dutch target texts (21.1 words).

| Text type | Words | Sentence length (source) | Sentence length (target) | Ratio S/T sentences |
|----------------|--------|--------------------------|--------------------------|---------------------|
| User Manuals | 9,558 | 13.1 | 13.1 | 1.00 |
| Press releases | 13,871 | 24.6 | 23.7 | 0.97 |
| Proceedings EP | 9,463 | 24.8 | 21.1 | 0.78 |

Table 1: Corpus characteristics

3 Gold standard

In order to create an a-priori reference alignment for a set of English-Dutch parallel texts, translational correspondence was indicated manually for all texts of the corpus described above. To that end an annotation style guide was created.

3.1 Annotation Style guide

The annotation style guide was to a large extent based on the annotation guidelines of related alignment projects (Melamed, 2001a; Merkel, 1999; Och and Ney, 2000; Véronis, 1998). The annotators were asked to indicate the *minimal* language unit in the source text that corresponds to an equivalent in the target text, and vice versa. To determine this minimal language unit, two major rules were formulated (Merkel, 1999; Véronis, 1998):

- Select *as many words as necessary* in the source and in the target sentence to ensure a two-way equivalence.
- Select *as few words as possible* in the source and in the target sentence, while preserving two-way equivalence.

The annotators had to link all the words of the source and the target text. To facilitate the anno-

¹ The SAP files are part of the training kits of the EColoRe project, available online at <http://ecolore.leeds.ac.uk>

² Available online at <http://www.statmt.org/europarl/>

tation process, a graphical annotation tool, Hand-Align³, was used.

They could link different units (e.g. word, word groups, paraphrased sections, punctuation). The corresponding units were not necessarily contiguous.

As translations are characterized by both correspondences and changes, two different types of alignments were introduced: regular links were used to connect straightforward correspondences; fuzzy links for translation-specific shifts of various kinds (syntactic shifts, e.g. active-passive transformations, paraphrases, etc.). An example of an annotated sentence is given in Figure 1. In the example, a fuzzy link is used to connect *The weight of the TV ... is* with *de TV weegt*.

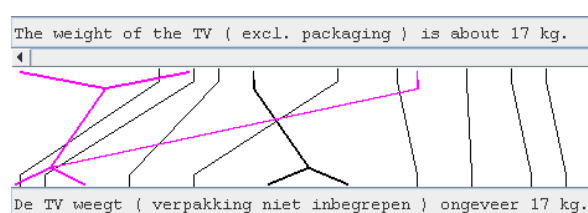


Figure 1: Graphical representation of the links

A multi-level annotation was used in case of fuzzy links. If words or word groups within a fuzzy link clearly corresponded, these were marked with a regular link. In the fuzzy link example of Figure 1, a regular link is used to connect English *the* and *TV* to Dutch *De* and *TV*.

Null links were used for source text units that had not been translated or target text units that had been added. Null links were also used for segments that are paraphrased in such a way that no correspondence could be indicated for those segments.

For further processing, the output of the graphical annotation tool was converted into a table (see Figure 2). In the table representation, the first column contains the source text segment, the second column the target text segment, and the third column the type of link (1 is fuzzy, 2 is regular).

| | | |
|-----------------------------|-----------------|---|
| The+weight+of+the+TV+...+is | De+TV+weegt | 1 |
| the | De | 2 |
| TU | TU | 2 |
| (| (| 2 |
| excl. | niet+inbegrepen | 2 |
| packaging | verpakking | 2 |
|) |) | 2 |
| about | ongeveer | 2 |
| 17 | 17 | 2 |
| kg. | kg. | 2 |

Figure 2: Table representation of alignments

3.2 Inter-annotator reliability

To assess inter-annotator reliability, three volunteers of the English department of the School of Translation Studies of Ghent manually annotated nine texts of the corpus of press releases. The alignments of the volunteers were compared with the author's alignments for those nine texts (pairwise comparisons). Two different statistical metrics were used to assess inter-annotator reliability.

A widespread measure for evaluating inter-annotator agreement for tagging tasks in the field of computational linguistics is the Kappa statistic (Carletta, 1996; Di Eugenio and Glass, 2004). The Cohen's Kappa Statistic measures pairwise agreement among coders making category judgments. For a similar task, Daumé III and Marcu (2005) used the Kappa-statistic to compute inter-annotator agreement for word-to-word and phrase-to-phrase alignments between existing abstract-document pairs for automatic document summarization.

To satisfy the needs of the annotation scheme presented above, the procedure of Daumé III and Marcu was slightly adapted. After the conversion of all phrase-to-phrase alignments into word-to-word alignments by linking each word of the source phrase to each word of the target phrase (all-pairs heuristic), each possible word combination of a given source and target sentence was placed into a specific category, depending on the type of connection between the source and target word.

One-to-one alignments were categorized as "direct links", whereas words connected within phrase alignments were categorized as "indirect links". To account for null links, one extra virtual "null-word" was added in each source and target sentence, and null links were treated as one-to-null or as null-to-one links. The distinction between regular links and fuzzy links was retained, but regular links within fuzzy links were ignored. This resulted in six different categories: not linked, direct regular links, indirect regular links, direct fuzzy link, indirect fuzzy links and null links.

Kappa was computed over these categories and results between 0.7 and 0.9 were obtained.

According to Carletta (1996), a kappa score over 0.8 reflects good agreement, and kappa values between 0.67 and 0.8 allow tentative conclusions to be drawn.

³ Available online at <http://www.cs.utah.edu/~hal/HandAlign/>

To be able to compare the obtained inter-annotator results with other alignment projects, the *Word Alignment Agreement* score (Davis, 2002) was calculated.

As for Kappa, phrasal alignments were converted into word-to-word alignments using the all-pairs heuristic. Inter-annotator agreement was measured in terms of similarity between sets of corresponding words. To normalize the inter-linked word-to-word links from the phrasal alignments, a weight was assigned to each word-to-word link.

The WAA-score is a symmetric measure and gives a number between zero and one, with zero being no agreement and one being perfect agreement. For the inter-annotator experiment WAA-scores between 0.84 and 0.94 were obtained. These results are similar to the scores reported by Melamed⁴ (2001b).

4 Analysis

As we expected, a different degree of “freeness” was observed in the different text types, which is reflected in the percentage of fuzzy links and null links. The proceedings of the debates contain the highest number of fuzzy links (10%) and the highest number of null links (8%). The user manuals contain the lowest percentage of fuzzy links (5%), and very few null links (2.7%).

In the following subsections the translational correspondences are analyzed in more detail.

| Text type | Regular | Fuzzy | Null |
|----------------|---------|-------|------|
| User Manuals | 92.0 | 5.3 | 2.7 |
| Press releases | 89.3 | 6.0 | 4.7 |
| Proceedings EP | 81.6 | 10.3 | 8.1 |

Table 2: Percentage of regular, fuzzy and null links

4.1 Syntactically motivated chunks

Some EBMT alignment modules only incorporate syntactically motivated chunks in their example database (Gough and Way, 2004; Stroppa et al., 2006). Within that context, it was interesting to verify on our corpus to what extent the alignments could be mapped onto syntactic chunks.

We used the English memory-based shallow parser of the CNTS/ILK-group (Daelemans et

al., 1999), to generate base chunks for the English source texts (NP, VP, PP, ADVP, ADJP).

We then compared the syntactic chunks with the reference alignments. A chunk was considered compliant to the set of alignments if no alignment of the set crossed the chunk boundaries. For example, if a chunk consists of three words, the chunk was considered compliant if those three words were not part of an alignment spanning a larger segment than those three words.

For the user manuals and the press releases, 82-83% of the chunks were compliant with the alignments. For the proceedings of the debates, 70% were compliant.

From the figures we can conclude that if the alignment module is restricted to syntactically motivated chunks, 20-30% of the chunks are missed. This is in line with the findings of Koehn et al. (2003). They demonstrated that the performance of their phrase-based SMT-system decreased when the phrase translation table was restricted to only syntactically motivated chunks, as too many useful phrases were eliminated.

While one can argue that some fixed expressions can be coded easily as “word with spaces” with one part-of-speech tag in the tagger or parser, other cases are more difficult to solve.

If we examine the alignments that cross phrase boundaries, we come across the whole range of multiword expressions, going from fixed or semi-fixed phrases and multi-word prepositions, over phrasal verbs and compounds to idioms and paraphrases:

- fixed phrases
*My party and group are committed to giving our full support, **in particular** to the British Government...*
*Mijn partij en mijn fractie geven onze volledige steun aan de Britse regering **in het bijzonder**...*
- multiword prepositions
*May I thank Mr Watson **on behalf of** our committee for bringing this report to plenary.*
*Staat u mij toe om de heer Watson **namens** onze commissie te bedanken voor de presentatie van dit verslag in de plenaire vergadering.*
- complex nouns
*...state-of-the-art system that will take Portugal to the **head of the field** in maritime traffic surveillance*
...dit geavanceerde system dat Portugal

⁴ The WAA-score is a further refinement of the metrics used by Melamed.

naar de **kopgroep** van het maritieme verkeerstoezicht zal brengen

- phrasal verbs
*The U2 Elevation Tour 2001, the band's first tour since 1992, **kicked off** on March 24 in Miami.*
*De U2 Elevation Tour 2001 is de eerste arena tournee van de groep sinds 1992. De tournee **startte** op zaterdag 24 maart 2001 in Miami.*
- light verb constructions
*The best memorial for those innocent people who died so tragically last week is **to take action** to prevent its repetition.*
*Het beste gedenkteken voor de onschuldige mensen die vorige week zo tragisch om het leven zijn gekomen, is **het nemen van maatregelen** om een herhaling te voorkomen.*
- idioms
*I ask whether now is not the time to **con-sign to the rubble of history** the clumsy, outdated second and third pillars of the EU cooperation...*
*Is het geen tijd om de logge, achterhaalde tweede en derde pijler van de communautaire samenwerking **definitief overboord te zetten** ...*
- paraphrases
*If Acrobat Reader is not detected on your system, **you are given the opportunity to** install it at this time.*
*Als Acrobat Reader niet op uw systeem wordt aangetroffen, **kunt u** dit programma direct installeren.*

Presenters include Carlos Santana, Toni Braxton, ...

Deze happening zal worden gepresenteerd door Carlos Santana, Toni Braxton...

Developing a system that can automatically align these different types of multiword expressions is quite a challenge. While some types of multiword expressions occur frequently in a corpus, and can be captured by statistical techniques (e.g. multiword prepositions), other types (e.g. idioms) are more difficult to deal with.

4.2 Non-contiguous correspondences

While the examples given in section 4.1 only contain contiguous source and target segments, some of the multiword expressions are flexible in the sense that the parts of the multiword are not always contiguous.

Some examples of flexible multiword expressions:

- *XYZ **focuses solely on** providing the most advanced video display products. XYZ **richt zich uitsluitend op** het leveren van de meest geavanceerde video display producten.*
- *our aim remains **to drive the industry forward** het blijft onze doelstelling de markt **een duwtje in de rug te geven***
- *at the **top left corner** ...in de linker **bovenhoek***

Apart from the flexible multiword expressions, another category of non-contiguous segments is formed by determiner-noun and auxiliary-verb pairs. This is a direct result of some specific rules of the annotation guidelines. The annotators were asked to link extra determiners and extra auxiliaries in source or target text together with the noun or main verb to the noun's or verb's translation:

- *...we **were soon sitting** in front of television sets*
*Al snel **zaten** wij voor de televisie...*

Between 2.5 and 5 percent of the aligned target language segments and less than one percent of aligned source language segments are not contiguous (see Table 3).

| Text type | Source | Target |
|----------------|--------|--------|
| User Manuals | 0.8 | 3.2 |
| Press releases | 0.7 | 2.5 |
| Proceeding EP | 0.9 | 4.9 |

Table 3: Percentage of non-contiguous segments in source and target texts

The difference in percentage of non-contiguous segments in source and target language is due to the high percentage of verbs with separable prefixes in Dutch and the less strict word order in Dutch, where verbal groups are more often separated than in English.

4.3 Length of alignments

Koehn et al. (2003) investigated the impact of phrase length on statistical phrase-based machine translation. They concluded that a phrase length of maximally three words was sufficient. Learning longer phrases did not improve the overall system. On the other hand, when the phrase length was limited to two words, performance decreased.

To verify their findings on our corpus, we computed the segment lengths of the manual alignments. Table 4 gives an overview of the percentage of links with different segment length. Column 1 contains the 1:1 links, column 2 the 1:2, 2:2 and 2:1 links, and so forth.

When you look at the global figures, you can see that limiting the segment length to three words on both source and target side, covers 98% of the alignments of the texts of the user manuals, 96% of the alignments of the press releases and 93% of the alignments of the proceedings of the plenary debates.

| Text type | 1 | 2 | 3 | >3 |
|-------------------|----|----|---|----|
| User Manuals | 79 | 15 | 4 | 2 |
| Press releases | 78 | 13 | 5 | 4 |
| Proceedings EP | 75 | 13 | 5 | 7 |

Table 4: Percentage of links with different segment length

If only regular links are taken into account, the percentage is even higher: 99% of the alignments are covered by segments of maximally three words (see Table 5). Eighty-three percent of the regular links are word-to-word links.

| Text type | 1 | 2 | 3 | >3 |
|-------------------|----|----|---|----|
| User Manuals | 83 | 14 | 2 | 1 |
| Press releases | 83 | 12 | 4 | 1 |
| Proceedings EP | 83 | 13 | 3 | 1 |

Table 5: Percentage of regular links with different segment length

However, when we look at the segment length of the fuzzy links, we see a rather different picture. The overall segment length of fuzzy links is much higher. By limiting the segment length to three words, depending on the text type, 30-45% of the fuzzy alignments are not covered (see Table 6).

| Text type | 1 | 2 | 3 | >3 |
|-------------------|----|----|----|----|
| User Manuals | 6 | 31 | 30 | 33 |
| Press releases | 16 | 30 | 24 | 30 |
| Proceedings EP | 17 | 19 | 19 | 45 |

Table 6: Percentage of fuzzy links with different segment length

By way of illustration, we give some examples of translational correspondence of segments with a segment length of more than 3 words:

- *Last week's attacks on innocent civilians in New York and Washington ... De aanvallen op onschuldige burgers die vorige week hebben plaatsgevonden in New York en Washington ...*
- *As the Americans seek the arrest of people on European soil who may be responsible, they must **shake their heads in disbelief** at the multitude of bilateral extradition agreements, which they will need to employ. De Amerikanen die op Europees grondgebied proberen mogelijk verantwoordelijke personen te arresteren, moeten wel **verbijsterd zijn** over de enorme hoeveelheid bilaterale uitleveringsovereenkomsten waarmee zij rekening dienen te houden.*

4.4 Null links

Null links were used for source text units that had not been translated or target text units that had been added. Null links were also used for segments that are paraphrased in such a way that no correspondence could be indicated for those segments.

Especially in the freer translations, a high percentage of null links can be observed: 8.1% of the Europarl links, versus 4.7% of the press releases links and only 2.7% of the user manuals links are null links (see Table 2).

For the manual reference corpus, also punctuation marks were to be linked. Around 30-40% of the null links were punctuation marks for which no correspondence could be found in either source or target text (see Table 7).

In the texts of the press releases and the proceedings of the debates more multi-word units could not be linked (30%), which could be an indication of a higher degree of rewriting.

| Text type | One word | More words | Punctuation |
|----------------|----------|------------|-------------|
| User Manuals | 37 | 22 | 41 |
| Press releases | 30 | 30 | 40 |
| Proceedings EP | 37 | 30 | 33 |

Table 7: Percentage of null links divided per type

In the example below, one English sentence corresponds to four Dutch sentences. The segments of the target sentences for which no correspondence could be found in the source sentence are indicated in bold:

Measures are needed too in foreign and security policy: greater cooperation in the UN, in the G8, in the Council of Europe, urging Member States to sign the UN Convention on the financing of terrorism; export controls on sensitive products and tighter association agreements with third countries.

*Ook in het buitenlands en veiligheidsbeleid zijn maatregelen nodig. **Ik denk daarbij aan** meer samenwerking binnen de VN, de G8 en de Raad van Europa. **Er moet bij** lidstaten op worden aangedrongen dat zij het Internationaal Verdrag ter bestrijding van de financiering van terrorisme ondertekenen. **Andere voorbeelden zijn** exportcontroles op gevoelige producten en nauwe associatieovereenkomsten met derde landen.*

4.5 Multi-level annotations

Recall that a multi-level annotation was used in case of fuzzy links. If words of word groups within a fuzzy link clearly corresponded, these were marked with a regular link. In the context of building an automatic module that can align fuzzy links, it is interesting to investigate to what extent lexical clues (word-to-word correspondences) can be used to identify fuzzy correspondences.

For each fuzzy link containing at least 2 source words or 2 target words, we verified whether the annotators indicated at least one regular link inside the fuzzy link. Table 8 shows that in 60% of the multiword fuzzy links, a lexical clue could be found.

| Text type | Lexical clue | No lexical clue |
|----------------|--------------|-----------------|
| User Manuals | 61 | 39 |
| Press releases | 60 | 40 |
| Proceedings EP | 60 | 40 |

Table 8: Percentage of fuzzy links > 2 words with and without lexical clues

In the following example, the lexical clues within the fuzzy link are underlined. The words connected with a fuzzy link are indicated in bold:

- *It is incumbent on this House to make clear to the Commission and the Council that **aid needs to be provided to the European airline industry and to the aviation industry in general...***
*Het is de taak van dit Parlement om de Commissie en de Raad duidelijk te maken dat de Europese luchtvaartmaatschappijen en de luchtvaartindustrie in het algemeen **steun nodig hebben.***
- *...**their objectives are political***
*...**zij hebben allen bepaalde politieke doelstellingen***

No lexical clues can be indicated in the following translated sentence:

- *I challenge any minister who may resist these proposals **in the secrecy of the Council chamber...***
*Ik daag iedere minister die deze voorstellen **achter de gesloten deuren van de Raadskamer verwerpt uit...***

5 Conclusion

In this paper, a deeper insight is given in the problem of translational correspondence in different text types. A manual reference corpus was created, in which three different types of links were used: regular links for straightforward correspondences, fuzzy links for translation-specific shifts of various kinds, and null links for words for which no correspondence could be indicated.

We showed that a different translation style was adopted in different text types. A freer translation style was characterized by a high degree of fuzzy links and null link. Texts with a high degree of regular links follow more closely the forms of the source texts, and are easier to translate and hence more suited for machine translation.

We examined whether alignments could be restricted to syntactic chunks. We demonstrated that translational correspondence is often established at the level of multiword expressions, which do not necessarily coincide with syntactic chunks. As some types of multiword expressions are flexible in the sense that the parts of the multiword are not always contiguous, the problem of multiword expressions and the problem of non-contiguous segments are interwoven.

The percentage of null links is another indication of the degree of “freeness” of a translation. Null links were used for source text units that had not been translated or target text units that had been added. In most texts, these are local phenomena involving just a few words.

But null links were also used for segments that are paraphrased in such a way that no correspondence could be indicated for such segments. We have also demonstrated that for 40% of the multiword fuzzy links, no lexical clue could be found. Hence, it will be very difficult for an alignment module to distinguish between fuzzy links without lexical clues and heavily paraphrased segments.

It depends on the text type, whether it is sufficient to focus on straightforward correspondences (regular links). If only straightforward correspondences need to be extracted, an acceptable performance can be achieved by focussing only on contiguous translation units of maximally three words.

However, if more complex correspondences need to be extracted, the alignment module faces a more complex task of matching larger and non-contiguous units. Especially for a human translator those less straightforward correspondences are extremely useful, as it is not trivial to find direct lexical equivalents for those correspondences, and literal translations are not always acceptable.

6 Future work

In the short term, we intend to add syntactic information to the Dutch target texts. This will allow us to quantify the correspondences on an abstract (linguistic) level, and make an inventory of the most frequent structural translation shifts.

Next, we want to extend our Gold Standard by adding direct Dutch-English translations extracted from the Europarl corpus to investigate the influence of the translation direction.

In the long run, our objective is to build a robust alignment module that can cope with some

of the more difficult translational correspondences.

References

- Jean Carletta. 1996. Assessing agreement on classification tasks: the Kappa statistic. *Computational Linguistics*, 22(1), 249-254.
- Walter Daelemans, Sabine Buchholz, and Jorn Veenstra. 1999. *Memory-based shallow parsing*. Proceedings of the CoNLL-99, Bergen, Norway.
- Hal Daumé III, and Daniel Marcu. 2005. Induction of word and phrase alignments for automatic document summarization. *Computational Linguistics*, 31(4), 505-530.
- Paul C. Davis. 2002. *Stone Soup Translation: The Linked Automata Model*. Unpublished Ph.D., Ohio State University, Columbus, Ohio.
- Barbara Di Eugenio, and Michael Glass. 2004. The Kappa statistic: a second look. *Computational Linguistics*, 30(1), 95-101.
- Fabrizio Gotti, Philippe Langlais, Elliott Macklovitch, Didier Bourigault, Benoit Robichaud, and Claude Coulombe. 2005. *3GTM: a third-generation translation memory*. Proceedings of the 3rd Computational Linguistics in the North-East (CLiNE) Workshop, Gatineau, Québec.
- Nano Gough, and Andy Way. 2004. *Robust large-scale EBMT with marker based segmentation*. Proceedings of the Tenth Conference on Theoretical and Methodological Issues in Machine Translation (TMI-04) Baltimore, MD.
- Philipp Koehn. 2005. *Europarl: a parallel corpus for statistical machine translation*. Proceedings of the Tenth Machine Translation Summit, Phuket, Thailand.
- Philipp Koehn, Franz Josef Och, and Daniel Marcu. 2003. *Statistical Phrase-based translation*. Proceedings of the HLT-NAACL 2003: conference combining Human Language Technology conference series and the North American Chapter of the Association for Computational Linguistics conference series, Edmonton, Canada.
- Dan I. Melamed. 2001a. Annotation style guide for the Blinker Project. In D. I. Melamed (Ed.), *Empirical methods for exploiting parallel texts* (pp. 169-182). Cambridge, Massachusetts: MIT Press.
- Dan I. Melamed. 2001b. Manual annotation of translational equivalence. In D. I. Melamed (Ed.), *Empirical methods for exploiting parallel texts* (pp. 65-77). Cambridge, Massachusetts: MIT Press.
- Magnus Merkel. 1999. Annotation Style Guide for the PLUG Link Annotator. Retrieved 08-06-2005, from http://www.ida.liu.se/~magma/publications/plugin_kannot.pdf
- Franz Josef Och, and Hermann Ney. 2000. *A comparison of alignment models for statistical machine translation*. Proceedings of the 18th International

Conference on Computational Linguistics (COLING 2000), Saarbrücken, Germany.

Nicolas Stroppa, Declan Groves, Andy Way, and Kepa Sarasola. 2006. *Example-Based Machine Translation of the Basque Language*. Proceedings of the 7th Conference of the Association for Machine Translation in the Americas (AMTA), Boston, MA.

Jean Véronis. 1998. Arcade. Tagging guidelines for word alignment. Version 1.0. Retrieved 15-11-2005, from <http://www.up.univ-mrs.fr/veronis/arcade/arcade1/2nd/word/guide/index.html>