Spanish-to-Basque MultiEngine Machine Translation for a Restricted Domain

**Iñaki Alegria (1), Arantza Casillas (2), Arantza Diaz de Ilarraza (1), Jon Igartua (1),
Gorka Labaka (1), Mikel Lersundi (1), Aingeru Mayor (1), Kepa Sarasola (1)**

(1) Informatika Fakultatea. IXA group. (2) Zientzia eta Teknologia Fakultatea.
University of the Basque Country. UPV-EHU

`i.alegria@ehu.es`

Abstract

We present our initial strategy for Spanish-to-Basque MultiEngine Machine Translation, a language pair with very different structure and word order and with no huge parallel corpus available. This hybrid proposal is based on the combination of three different MT paradigms: Example-Based MT, Statistical MT and Rule-Based MT. We have evaluated the system, reporting automatic evaluation metrics for a corpus in a test domain. The first results obtained are encouraging.

1 Introduction

Machine translation for Basque is both a real need and a testing ground for our strategy to develop language tools. The first development was Matxin, a Rule-Based MT system (Mayor, 2007). Later on a Data-Driven Machine Translation system was built and both systems compared (Labaka et al., 2007). As both approaches have their limits, and each deals with a different kind of knowledge, it was decided to try combining them to improve their results. On the one hand, after improvements in 2007 (Labaka et al., 2007) the Spanish-to-Basque RBMT system Matxin proved useful for assimilation, but is still not suitable for unrestricted use in text dissemination. On the other hand, data-driven MT systems base their knowledge on aligned bilingual corpora, and the accuracy of their output depends heavily on the quality and the size of these corpora. When the pair of languages used in translation, such as Spanish and Basque, has very different structures and word orders, the corpus obviously needs to be

bigger. However, since Basque is a lesser-used language, large and reliable bilingual corpora are unavailable. At present, domain-specific translation memories for Basque are no bigger than two or three million words, much smaller than corpora used for other languages; for example, Europarl corpus (Koehn, 2005), a standard resource, has 30 million words. So, although domain-restricted corpus-based MT for Basque shows promising results, it is still not ready for general use.

Therefore, it is clear that we should combine the basic techniques for MT (rule-based and corpus-based) in order to build a hybrid system with better performance. Due to the pressing need for translation in public administration and taking into account that huge parallel corpora for Basque are not available, we have tested a first strategy by building a MT engine for a restricted domain related to public administration for which translation memories were available.

The rest of the paper is organized as follows. Section 2 presents some related work. Section 3 describes the corpus we have compiled to carry out the experiments. Section 4 explains the single engines built up for Basque MT and how we have combined them. Section 5 reports our experiments. Finally, we draw conclusions and refer to future work.

2 Related Work

(van Zaanen and Somers, 2005), (Matusov et al., 2006) and (Macherey and Och, 2007) review a set of references about MEMT (Multi-Engine MT) including the first attempt by (Frederking and Nirenburg, 1994). All the papers on MEMT reach the same

conclusion: combining the outputs results in a better translation. Most of the approaches generate a new consensus translation combining different SMT systems using different language models and in some cases combining also with RBMT systems. Some of the approaches require confidence scores for each of the outputs. The improvement in translation quality is always lower than 18% relative increasing in BLEU score.

(Chen et al., 2007) reports 18% relative increment for in-domain evaluation and 8% for out-domain, by incorporating phrases (extracted from alignments from one or more RBMT systems with the source texts) into the phrase table of the SMT system and use the open-source decoder Moses to find good combinations of phrases from SMT training data with the phrases derived from RBMT.

(Matusov et al., 2006) reports 15% relative increment in BLEU score using consensus translation computed by voting on a confusion network. Pairwise word alignments of the original translation hypotheses were estimated for an enhanced statistical alignment model in order to explicitly capture reordering.

(Macherey and Och, 2007) presented an empirical study on how different selections of translation outputs affect translation quality in system combination. Composite translations were computed using (i) a candidate selection method based on inter-system BLEU score matrices, (ii) a ROVER-like combination scheme, and (iii) a novel two-pass search algorithm which determines and re-orders bags of words that build the constituents of the final consensus hypothesis. All methods gave statistically significant relative improvements of up to 10% BLEU score. They combine large numbers of different research systems.

(Mellebeek et al., 2006) reports improvements of up to 9% BLEU score. Their experiment is based in the recursive decomposition of the input sentence into smaller chunks, and a selection procedure based on majority voting that finds the best translation hypothesis for each input chunk using a language model score and a confidence score assigned to each MT engine.

(Huang and Papineni, 2007) and (Rosti et al., 2007) combines multiple MT systems output at word-, phrase- and sentence-levels. They report im-

provements of up to 10% BLEU score.

3 The Corpus

Our aim was to improve the precision of the existing Spanish-to-Basque MT system by trying to translate texts in a restricted domain, because reliable Spanish-Basque corpora are not sufficiently available for a general domain. Also, we were interested in a kind of domain where a formal language would be used and in which many public organizations and private companies would be interested.

The Basque Institute of Public Administration (IVAP[1]) collaborated with us in this selection by examining some possible domains, available parallel corpora, and translation needs. We selected the domain related to labor agreements. Then, we built the Labor Agreements Corpus using a bilingual parallel corpus with 585,785 words in Basque and 839,003 in Spanish.

To build the test corpus, we randomly chose the full text of several labor agreements. We chose full texts because we wanted to ensure that several significant but short elements, such as headers and footers, would be represented, and also because it is important to measure the coverage and precision we get when translating the whole text in one document and not only some parts of it. First, we automatically aligned the corpus at sentence level, and then we performed manual revision. We did not allow system developers to see the test corpus.

As we have said, our goal was to combine different MT approaches: Rule-Based (RBMT), Example-Based (EBMT) and Statistical (SMT). Once we had the corpus, we split it into three parts for SMT (training, development and test corpus) and into two parts for EBMT (development and test corpus). In SMT we used the training corpus to learn the models (translation and language model), the development corpus to tune the parameters, and the test corpus to evaluate the system. In RBMT and EBMT there are no parameters to optimize, and so we considered only two corpora: one for development (combining the training and development parts used in SMT) and one for the test.

Table 1 shows the size, number of documents, sentences and words in the training, development,

---

[1] *http://www.ivap.euskadi.net*

Subset	Lang.	Doc.	Senten.	Words
Train	Basque	81	51,740	839,393
	Spanish	81		585,361
Develop	Basque	5	2,366	41,408
	Spanish	5		28,189
Test	Basque	5	1,945	39,350
	Spanish	5		27,214

Table 1: Labor Agreements Corpus

and test subsets of each language.

4 The MultiEngine MT system

In the next subsections we explain the three single MT strategies we have developed: Example-Based Approach, Statistical Machine Translation Approach and Rule-Based Machine Translation Approach. Finally, we explain how we have combined these three approaches.

4.1 Example Based Approach

In this subsection we explain how we automatically extract translation patterns from the bilingual parallel corpus and how we exploit them.

Translation patterns are generalizations of sentences that are translations of each other, replacing various sequences of one or more words by variables (McTait, 1999).

Starting from the aligned corpus we carry out two steps to automatically extract translation patterns. First, we detect some concrete units (mainly entities) in the aligned sentences and then we replace these units by variables. To detect the units, due to the morphosyntactic differences between Spanish and Basque, we need to execute particular algorithms for each language. We have developed algorithms to determine the boundaries of dates, numbers, named entities, abbreviations and enumerations.

After detecting the units, they must be aligned, relating the Spanish and Basque units of the same type that have the same meaning. For numbers, abbreviations, and enumerations, the alignment is almost trivial; however, the alignment algorithm for named entities is more complex. It is explained in more detail in (Martínez et al., 1998). Finally, to align the dates, we use their canonical form. Table 2 shows an example of how a translation pattern is extracted.

ES-EU Sentences	Sentences with generalized units	Translation Pattern
En Vitoria-Gasteiz, a 22 de Diciembre de 2003	En<rs type=loc> Vitoria -Gasteiz</rs> , a<date date=22/12/2003> 22 de Diciembre de 2003</date>	En<rs1> , a<date1>.
Vitoria Gasteiz, 2003ko Abenduaren 22.	<rs type=loc> Vitoria -Gasteiz </rs>,<date date=22/12/2003> 2003ko 22 Abenduaren </date>	<rs1> <date1>

Table 2: Example of Translation Pattern extraction

Once we have automatically extracted all the possible translation patterns from the training set, we store them in a hash table for use in the translation process.

When we want to translate a source sentence, we check if that sentence matches any pattern in the hash table. If the source sentence matches a sentence in the hash table with no variable, the translation process will immediately return its translation. A Word Error Rate (WER) metric was used to compare the two sentences. Otherwise, if the source sentence does not match anything in the hash table, the translation process will try to generalize that sentence and will check the hash table again for a generalized template. To generalize the source sentence, the translation process will apply the same detection algorithms used in the extraction process.

In a preliminary experiment using a training corpus of 54,106 sentence pairs we automatically extracted 7,599 translation patterns at the sentence level. These translation patterns covered 35,450 sentence pairs of the training corpus. We also consider an aligned pair of sentences as a translation pattern if it does not have any generalized unit but appears at least twice in the training set.

As this example-based system has very high precision but very low coverage, it is interesting to com-

bine it with the other MT engines, especially in this kind of domain where a formal and quite sublanguage is used.

4.2 Statistical Machine Translation Approaches

Two different approaches have been implemented: a conventional SMT system and a morpheme-based system. These corpus-based approaches have been carried out in collaboration with the National Center for Language Technology in Dublin. The system exploits SMT technology to extract a dataset of aligned chunks. Based on a training corpus, we conducted Spanish-to-Basque translation experiments (Labaka et al., 2007).

We used freely available tools to develop the SMT systems:

- GIZA++ toolkit (Och, 2003) for training the word/morpheme alignment.

- SRILM toolkit (Stolcke, 2002) for building the language model.

- Moses Decoder (Koehn et al., 2007) for translating the sentences.

Due to the morphological richness of Basque, some Spanish words, like prepositions or articles, correspond to one or more suffixes in Basque. In order to deal with this problem, we built a morpheme-based SMT system.

Adapting the SMT system to work at the morpheme level consists of training the basic SMT on the segmented text. The translation system trained on this data will generate a sequence of morphemes as output. In order to obtain the final Basque text, words have to be generated from those morphemes.

To obtain the segmented text, we analyzed Basque texts using Eustagger (Aduriz et al., 2003). This process replaces each word with the corresponding lemma followed by a list of morphological tags. The segmentation is based on the strategy proposed in (Agirre et al., 2006).

We optimized both systems (the conventional SMT and the morpheme-based) by decoding parameters using Minimum Error Rate Training. The metric used to carry out the optimization is BLEU. Table 3 shows: the conventional SMT system reported 9.51 for BLEU accuracy measure and 3.73

	BLEU	NIST	WER	PER
SMT	9.51	3.73	83.94	66.09
Morpheme based SMT	8.98	3.87	80.18	63.88

Table 3: Evaluation for SMT Systems

for NIST; the morpheme-based SMT system reported 8.98 BLEU and 3.87 NIST accuracy measures.

4.3 Rule-Based Machine Translation Approach

In this subsection we present the main architecture of an open-source RBMT engine named Matxin (Alegria et al., 2007), the first implementation of which translates from Spanish to Basque using traditional transfer, based on shallow and dependency parsing.

The design and the programs of the Matxin system are independent from this pair of languages, so the software can be used for other projects in MT. Depending on the languages included in the adaptation, it will be necessary to add, reorder and change some modules, but this will not be difficult because a unique XML format is used for communication among all the modules.

The project has been integrated in the Open-Trad[2] initiative, a government-funded project shared among different universities and small companies, which includes MT engines for translation among the main languages in Spain. The main objective of this initiative is the construction of an open, reusable and interoperable framework.

In the OpenTrad project, two different but coordinated architectures have been developed:

- A shallow-transfer-based MT engine for similar languages (Spanish, Catalan and Galician).

- A deeper-transfer-based MT engine for the Spanish-Basque and English-Basque pair. It is named Matxin, and stored it in *matxin.sourceforge.net*. It is an extension of previous work by the IXA group.

For the second engine, following the strategy of reusing resources, another open-source engine,

---

[2]*http://www.opentrad.org*

FreeLing (Carreras et al., 2004), was integrated for parsing Spanish sentences.

The transfer module is divided into three phases which match the three main objects in the translation process: words or nodes, chunks or phrases, and sentences.

1. First, it carries out lexical transfer using a bilingual dictionary compiled into a finite-state transducer.

2. Then, it applies structural transfer at sentence level, transferring information from some chunks to others, and making some chunks disappear. For example, in the Spanish-Basque transfer, person and number information for the object is imported from other chunks to the verbal chunk. As in Basque the verb also agrees in person and number with the object, later on the generation of the verb in Basque will require this information.

3. Finally, the module carries out the structural transfer at chunk level. This process can be quite simple (e.g. noun chains between Spanish and Basque) or more complex (e.g. verb chains between these same languages).

Then the XML file coming from the transfer module is passed on to the generation module.

- In the first step, this module performs syntactic generation in order to decide the order of chunks in the sentence and the order of words in the chunks. It uses several grammars for this purpose.

- The last step is morphological generation. In generating Basque, the main inflection is added to the last word in the phrase (the declension case, the article and other features are added to the whole noun phrase at the end of the last word), but in verb chains other words need morphological generation. We adapted a previous morphological analyzer/generator for Basque (Alegria et al., 1996) and transformed it according to the format used in *Apertium*.

The results for the Spanish-Basque system using FreeLing and Matxin are promising. The quantitative evaluation uses the open-source evaluation tool

IQMT and we give figures using BLEU and NIST measures (Giménez et al., 2005). We also carried out an additional user-based evaluation, using Translation Error Rate (Snover et al., 2006). (Mayor, 2007) shows the results of the RBMT system's evaluation: 9.30, using the BLEU accuracy measure. In interpreting the results, we need to keep in mind that the development of this RBMT system was based on texts of newspapers.

We adapted this RBMT system to the domain of Labor Agreements in three main ways:

1. Terminology. Semiautomatic extraction of terminology using Elexbi, a bilingual terminology extractor for noun phrases (Alegria et al., 2006). Additionally, we carried out an automatic format conversion to the monolingual and bilingual lexicons for the selected terms. We extracted more than 1,600 terms from the development corpus, examined them manually, and selected nearly 807 to be include in the domain-adapted lexicon.

2. Lexical selection. Matxin does not address the lexical selection problem for lexical units (only for the preposition-suffix translation); it always selects the first translation in the dictionary (other possible lexical translations are stored for the post-edition process). For the domain adaptation, we calculated a new order for the possible translations based on the parallel corpus using GIZA++.

3. Resolution of format and typographical variants found frequently in the administrative domain.

After these improvements, the RMBT engine was ready to process sentences from this domain.

4.4 Approaches Combination

We experimented with a simple mixing alternative approach up to now used only for languages with huge corpus resources: selecting the best output in a multi-engine system (MEMT, Multi-engine MT). In our case, we combined RBMT, EBMT, and SMT approaches. In our design we took into account the following points:

1. Combination of MT paradigms: RBMT and data-driven MT.

2. Absence of large and reliable Spanish-Basque corpora.

3. Reusability of previous resources, such as translation memories, lexical resources, morphology of Basque and others.

4. Standardization and collaboration: using a more general framework in collaboration with other groups working in NLP.

5. Open-source: this means that anyone having the necessary computational and linguistic skills will be able to adapt or enhance it to produce a new MT system,

For this first attempt, we combined the three approaches in a very simple hierarchical way, processing each sentence with the three engines (RBMT, EBMT and SMT) and then trying to choose the best translation among them. First, we divided the text into sentences, then processed each sentence using each engine (parallel processing when possible). Finally, we selected one of the translations, dealing with the following facts:

- Precision of the EBMT approach is very high, but its coverage is low.

- The SMT engine gives a confidence score.

- RBMT translations are more adequate for human postedition than those of the SMT engine, but SMT gets better scores when BLEU and NIST are used with only one reference (Labaka et al., 2007). Table 4[3] summarizes the results of the automatic evaluation (BLEU) with one reference and those of the user-driven evaluation (HTER). Those evaluations where perfomed with two more general corpora related to news in the Basque Public Radio-Television (EiTB) and to articles in a magazine for consumers (Consumer).

With these results for the single approaches we decided to apply the following combinatory strategy:

1. If the EBMT engine covers the sentence, we chose its translation.

---

[3]*The Consumer corpus used for evaluation is the one referenced in Table 3 but before a cleaning process.*

Corpus	BLEU RBMT	BLEU SMT	HTER RBMT	HTER SMT
EiTB corpus	9.30	9.02	40.41	71.87
Consumer	6.31	8.03	43.60	57.97

Table 4: Evaluation using BLEU and HTER for single SMT and RBMT systems

	Coverage	BLEU	NIST
EBMT	EBMT 100%	29.02	4.70
RBMT	RBMT 100%	7.97	3.21
SMT	SMT 100%	14.37	4.43
EBMT +RBMT	EBMT 46.42% RBMT 53.58%	35.57	6.19
EBMT +SMT	EBMT 46.42% SMT 53.58%	38.31	6.82
EBMT +SMT +RBMT	EBMT 46.42% SMT 31.22% RBMT 22.36%	37.84	6.68

Table 5: Evaluation for MEMT systems using the development corpus

2. We chose the translation from the SMT engine if its confidence score was higher than a given threshold.

3. Otherwise, we chose the output from the RBMT engine.

5 Evaluation

In order to assess the quality of the resulting translation, we used automatic evaluation metrics. We report the following accuracy measures: BLEU (Papineni et al. , 2002) and NIST (Doddington, 2002).

The results using the development corpus for this second approach appear in Table 5.

Table 6 shows the results using the test corpus.

The best results, evaluated by using automatic metrics with only one reference, came from combining the two Data-Driven approaches: EBMT and SMT. Taking into account the single approaches, the best results are returned with EBMT strategy.

The results of the initial automatic evaluation showed very significant improvements. For example, a 193% relative increase for BLEU when comparing the EBMT+SMT+RBMT combination

	Coverage	BLEU	NIST
EBMT	EBMT 100%	32.42	5.76
RBMT	RBMT 100%	5.16	3.08
SMT	SMT 100%	12.71	4.69
EBMT +RBMT	EBMT 64.92% RBMT 35.08%	36.10	6.84
EBMT +SMT	EBMT 64.92% SMT 35.08%	37.31	7.20
EBMT +SMT +RBMT	EBMT 64.92% SMT 23.40% RBMT 11.68%	37.24	7.17

Table 6: Evaluation for MEMT systems using the test corpus

to the SMT system alone. Furthermore, we realized a 193.55% relative increase for BLEU when comparing the EBMT+SMT combination with the SMT system alone and 15.08% relative increase when comparing EBMT+SMT combination with the EBMT single strategy.

The consequence of the inclusion of a final RBMT engine (to translate just the sentences not covered by EBMT and with low confidence score for SMT) is a small negative contribution of 1% relative decrease for BLEU. Of course, bearing in mind our previous evaluation trials with human translators (Table 4), we think that a deeper evaluation using user-driven evaluation is necessary to confirm similar improvements for the MEMT combination including a final RBMT engine.

For example in the translation of the next sentence in Spanish (it is taken from the development corpus) "La Empresa concederá préstamos a sus Empleados para la adquisición de vehículos y viviendas, en las siguientes condiciones" the RBMT system generates "Enpresak maileguak emango dizkio haren Empleados-i ibilgailuen erosketarentzat eta etxebizitzak, hurrengo baldintzetan" and the SMT system "Enpresak mailegu ibilgailuak bertako langileei emango, eta etxebizitza erosteko baldintzak". The figures using BLEU and NIST are higher for the SMT translation, but only the RBMT translation can be understood.

The results of the MEMT systems are very similar in the development and test corpora. Although the percentage of coverage of the EBMT single system is lower for the development corpus, its precission is higher.

Most of the references about Multi-Engine MT do not use EBMT strategy, SMT+RBMT is the most used combination in the bibliography. One of our main contributions is the inclusion of EBMT strategy in our Multi-Engine proposal; our methodology is straightforward, but useful.

6 Conclusions and Future Work

We applied Spanish-to-Basque MultiEngine Machine Translation to a specific domain to select the best output from three single MT engines we have developed. Because of previous results, we decided to apply a hierarchical strategy: first, application of EBMT (translation patterns), then SMT (if its confidence score is higher than a given threshold), and then RBMT.

It has carried out an important improvement in translation quality for BLEU in connection with the improvements obtained by other systems. We obtain 193.55% relative increase for BLEU when comparing the EBMT+SMT combination with the SMT system alone, and 15.08% relative increase when comparing EBMT+SMT combination with the EBMT single strategy.

Those improvements would be dificult to get for single engine systems. RBMT contribution seems to be very small with automatic evaluation, but we expect that HTER evaluation will show better results.

In spite of trying the strategy for a domain, we think that our translation system is a major advance in the field of language tools for Basque. However the restriction in using a corpus in a domain is given by the absence of large and reliable Spanish-Basque corpora.

For the near future, we plan to carry out new experiments using a combination of the outputs based on a language model. We also plan to define confidence scores for the RBMT engine (including penalties when suspicious or very complex syntactic structures are present in the analysis; penalties for high proportion of ignored word senses; and promoting translations that recognize multiword lexical units). Furthermore, we are planning to detect other types of translation patterns, especially at the phrase or chunk level.

Acknowledgements

References

Itziar Aduriz and Arantza Díaz de Ilarraza. 2003. *Morphosyntactic disambiguation ands shallow parsing in Computational Processing of Basque.* Inquiries into the lexicon-syntax relations in Basque. Bernarrd Oyharabal (Ed.).

Eneko Agirre, Arantza Díaz de Ilarraza, Gorka Labaka, Kepa Sarasola. 2006. *Uso de información morfológica en el alineamiento Español-Euskara.* XXII Congreso de la SEPLN.

Iñaki Alegria, Xabier Artola, Kepa Sarasola. 1996. *Automatic morphological analysis of Basque.* Literary & Linguistic Computing Vol. 11, No. 4, 193-203. Oxford University Press.

Iñaki Alegria, Antton Gurrutxaga, Xabier Saralegi, Sahats Ugartetxea. 2006. *ELeXBi, A Basic Tool for Bilingual Term Extraction from Spanish-Basque Parallel Corpora.* Proceedings of the 12th EURALEX International Congress. pp 159-165

Iñaki Alegria, Antantza Díaz de Ilarraza, Gorka Labaka, Mikel Lersundi, Aingeru Mayor, Kepa Sarasola. 2007. *Transfer-based MT from Spanish into Basque: reusability, standardization and open source.* LNCS 4394. pp. 374-384.

Xavier Carreras, Isaac Chao, Lluis Padró, Muntsa Padrò. 2004. *FreeLing: An open source Suite of Language Analyzers.* Proceedings of the 4th International Conference on Language Resources and Evaluation.

Yu Chen, Andreas Eisele, Christian Federmann, Eva Hasler, Michael Jellinghaus, Silke Theison. 2007. *Multi-engine machine translation with an open-source decoder for statistical machine translation.* proceedings of the Second Workshop on Statistical Machine Translation, pp. 193-196.

George Doddington. 2002. *Automatic evaluation of Machine Translation quality using n-fram cooccurrence statics* Proceedings of HLT, pp. 128-132.

Robert Frederking and Sergei Nirenburg. 1994. *Three heads are better than one.* Proceedings of the fourth ANLP

Jesús Giménez, Enrique Amigó, Chiori Hori. 2005. *Machine Translation Evaluation Inside QARLA.* Proceedings of the International Workshop on Spoken Language Technology.

Fei Huang and Kishore Papineni. 2007 *Hierarchical system combination for machine translation.* Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, pp. 277-286.

Philipp Koehn. 2005. *Europarl: A parallel corpus for statistical machine translation.* MT Summit X.

Philipp Koehn, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcelo Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondrej Bojar, Alexandra Constantin, Evan Herbst. 2007. *Moses: Open Source Toolkit for Statistical Machine Translation.* Annual Meeting of the ACL.

Gorka Labaka, Nicolas Stroppa, Andy Way, Kepa Sarasola. 2007. *Comparing Rule-Based and Data-Driven Approaches to Spanish-to-Basque Machine Translation* Proceedings of MT-Summit XI.

Wolfgang Macherey and Franz J. Och. 2007. *An Empirical Study on Computing Consensus Translations from Multiple Machine Translation Systems.* Proceedings of the EMNLP and CONLL 2007.

Raquel Martínez, Joseba Abaitua, Arantza Casillas. 1998. *Alingning Tagged Bitext.* Proceedings of the Sixth Workshop on Very Large Corpora.

Evgeny Matusov, Nicola Ueffing, Hermann Ney. 2006. *Computing Consensus Translation from Multiple Machine Translation Systems Using Enhanced Hypotheses Alignment.* Proceedings of EACL 2006.

Aingeru Mayor. 2007. *Matxin: erregeletan oinarritutako itzulpen automatikoko sistema.* Ph. Thesis. Euskal Herriko Unibertsitatea.

Kevin M. McTait. 1999. *A Language-Neutral Sparse-Data.* Algorithm for Extracting Translation Patterns. Proceedings of 8th International Conference on Theorical and Mathodological Issues in Machine Translation.

Bart Mellebeek, Karolina Owczarzak, Josef Van Genabith, Andy Way. 2006. *Multi-engine machine translation by recursive sentence decomposition.* Proceedings of the 7th Conference of the Association for Machine Translation in the Americas, Visions for the Future of Machine Translation, pp.110-118.

Franz J. Och and Hermann Ney. 2003. *A Systematic Comparison of Various Statistical Alignment Models.* Computational Linguistics, 29(1): 19-51.

Kishore Papineni, Salim Roukos, Tod Ward, Wei-Jing Zhu. 2002. *BLEU: a method for automatic evaluation of machine translation.* Proceedings of 40th ACL, pp. 311-318.

Antti-Veikko I. Rosti, Necip Fazil Ayan, Bing Xiang, Spyros Matsoukas, Richard Schwartz, Bonnie J.Dorr. 2007. *Combining outputs from multiple machine translation systems.* NAACL-HLT-2007 Human Language Technology: the conference of the North American Chapter of the Association for Computational Linguistics, pp.228-235.

K. Sim, W. Byrne, M. Gales, H. Sahbi, P. Woodland. 2007. *Consensus network decoding for statistical machine translation system combination.* Proceedings of ICASSP, 2007.

Matthew Snover, Bonnie Dorr, Richard Schwartz, Linnea Micciulla, John Makhoul. 2006. *A study of translation edit rate with targeted human annotation.* Proceedings of AMTA'2006.

Andreas Stolcke. 2002. *SRILM - An Extensible Language Modeling Toolkit.* Proceedings Intl. Conerence. Spoken Language Processing.

Nicolas Stroppa, Declan Groves, Andy Way, Kepa Sarasola. 2006. *Example-Based Machine Translation of the Basque Language.* Proceeding of the 7th Conference of the AMTA.

Menno van Zaanen and Harold Somers. 2005. *DEMOCRAT: Deciding between Multiple Outputs Created by Automatic Translation.* MT Summit X.

Andy Way and Nano Gough. 2005. *Comparing Example-Based and Statistical Machine Translation.* Natural Language Engineering, 11(3):295-309.