A Hybrid System for Patent Translation

Ramona Enache **Cristina España-Bonet** **Aarne Ranta** **Lluís Màrquez**
Dept. of Computer Science and Engineering TALP Research Center
Chalmers University of Technology LSI Department
University of Gothenburg Universitat Politècnica de Catalunya
{enache,aarne}@chalmers.se {cristinae,lluism}@lsi.upc.edu

Abstract

This work presents a HMT system for patent translation. The system exploits the high coverage of SMT and the high precision of an RBMT system based on GF to deal with specific issues of the language. The translator is specifically developed to translate patents and it is evaluated in the English-French language pair. Although the number of issues tackled by the grammar are not extremely numerous yet, both manual and automatic evaluations consistently show their preference for the hybrid system in front of the two individual translators.

1 Introduction

The predominant core of machine translation (MT) systems has been changing through the years. From the very beginnings in the 50s where only dictionary-based MT systems existed, the technology evolved towards rule-based systems (RBMT). Later in the 90s the everyday more powerful computers allowed to develop empirical translation systems. Recently a type of empirical system, the statistical one (SMT), has become a widely used standard for translation. At this point the two main paradigms, RBMT and SMT, coexist with their strengths and weaknesses. Luckily these strengths and weaknesses are complementary and current efforts are being made to hybridise both of them and develop new technologies. A classification and description of hybrid translation can be found in (Thurmair, 2009).

In general RBMT provides high precision, due to an analysis of the text, but has limited coverage and a considerable amount of effort and linguistic knowledge is required in order to build such a system. On the other hand, SMT can achieve a huge coverage and is good at lexical selection and fluency but has problems in building structurally and grammatically correct translations.

Hybrid MT (HMT) is an emerging and challenging area of machine translation, which aims at combining the known techniques into systems that retain the best features of their components, and reduce the disadvantages displayed by each of the methods when used individually.

This work presents a hybrid translation system specifically designed to deal with the translation of patents. The language of patents follows a formal style adequate to be analysed with a grammar, but at the same time uses a rich and particular vocabulary adequate to be gathered statistically. We focus on the English-French language pair so that the effects of translating into a morphologically rich language can be studied.

With respect to the engine, a grammar-based translator is developed to assure grammatically correct translations. We extend GF (Grammatical Framework, Ranta (2011)) and write a new grammar for patent translation. The SMT system that complements the RBMT is based on Moses (Koehn et al., 2007). This system works on two different levels. First, it is used to build the parallel lexicon of the GF translator on the fly. Second, it is the top level decoder that takes the final decision about which phrases should be used.

In the following Section 2 describes recent work both in patent translation and hybrid systems. Section 3 explains our hybrid system and Section 4 evaluates its performance. Finally, Section 5 summarises the work and outlines possible lines to follow.

2 Related work

This work tackles two topics which are lately attracting the attention of researchers, patent translation and hybrid translators.

The high number of patents being registered and the necessity for these patents to be translated into several languages are the reason so that important efforts are being made in the last years to automate its translation between various language pairs. Different methods have been used for this task, ranging from SMT (Ceausu et al., 2011; España-Bonet et al., 2011a) to hybrid systems (Ehara, 2007; Ehara, 2010). Besides full systems, various components associated to patent translation are being studied separately (Sheremetyeva, 2003; Sheremetyeva, 2005; Sheremetyeva, 2009).

Part of this work is being done within the framework of two European projects, PLuTO (Patent Language Translations Online[1]) and MOLTO (Multilingual Online Translation[2]). PLuTO aims at making a substantial contribution to patent translation by using a number of techniques that include hybrid systems combining example-based and hierarchical techniques. On the other hand, one of MOLTO's use cases aims at extending a grammar-based translator with an SMT to gain robustness in the translation of patents. This paper is carried out within MOLTO.

HMT is not only useful in this context but is being applied in different domains and language pairs. Besides system combination strategies, hybrid models are designed so that there is one leading translation system assisted or complemented by other kinds of engines. This way the final translator benefits from the features of all the approaches. A family of models are based on SMT systems enriched with lexical information from RBMT (Eisele et al., 2008; Chen and Eisele, 2010). On the other side there are the models that start from the RBMT analysis and use SMT to complement it (Habash et al., 2009; Federmann et al., 2010; España-Bonet et al., 2011b).

Our work can be classified in the two families. On the one hand, SMT helps on the construction of the RBMT translator but, on the other hand, there is the final decoding step to integrate translations and complete those phrases untranslated by RBMT. We use GF as rule-based system.

GF is a type-theoretical grammar formalism, mainly used for multilingual natural language applications. Grammars in GF are represented as a pair of an *abstract syntax* –an interlingua that captures the semantics of the grammar on a language-independent level, and a number of *concrete syntaxes* –representing target languages. There are also two main operations defined, *parsing* text to an abstract syntax tree and *linearising* trees into raw text. In this way one can translate between two target languages of the same multilingual grammar, by combining parsing and linearization.

The GF resource library (Ranta, 2009) is the most comprehensive grammar for dealing with natural languages, as it features an abstract syntax which implements the basic syntactic operations such as predication and complementation, and 20 concrete syntax grammars corresponding to natural languages. This layered representation makes it possible to regard multilingual GF grammars as a RBMT system, where translation is possible between any pair of languages for which a concrete syntax exists. However, the translation system thus defined is first limited by the fixed lexicon defined in the grammar, and secondly by the syntactic constructions that it covers. For this reason, GF grammars have a difficult task in parsing free text. There is some recent work on parsing the Penn Treebank with the GF resource grammar for English (Angelov, 2011), whereas the current work on patent translation is the first attempt to use GF for parsing un-annotated free text.

3 HMT system

The patent translator is a hybridisation between rule-based and statistical techniques. So, the final system is not only a combination of two different engines but the subsystems also mix different components. We have developed a GF translator for the specific domain that uses an in-domain SMT system to build the lexicon; an SMT system is on top of it to translate those phrases not covered by the grammar. In the following we describe the individual translators and the data used for their development.

3.1 Corpus

A parallel corpus in English and French has been gathered from the corpus of patents given for the CLEF-IP track in the CLEF 2010 Conference[3]. These data are an extract of the MAREC corpus,

[1] http://www.pluto-patenttranslation.eu/

[2] http://www.molto-project.eu/

[3] http://clef2010.org/

containing over 2.6 million patent documents pertaining to 1.3 million patents from the European Patent Office[4] (EPO). Our parallel corpus is a subset with those patents with translated claims and abstracts into the two languages. From this first subset we selected those patents that deal with the biomedical domain.

The final corpus built this way covers 56,000 patents out of the 1.3 million. That corresponds to 279,282 aligned parallel fragments extracted from the claims. A fragment is the minimum aligned segment in the two languages, so, it is shorter than a claim and, consequently, shorter than a sentence. The length of the fragments is variable and depends on the aligned units that can be extracted from the xml mark-up within the patent such as paragraph tags for example. Two small sets for development and test purposes have also been selected with the same restrictions: 993 fragments for development and 1008 for test.

3.2 In-domain SMT system

The first component is a standard state-of-the-art phrase-based SMT system trained on the biomedical domain with the corpus described in Section 3.1. Its development has been done using standard freely available software. A 5-gram language model is estimated using interpolated Kneser-Ney discounting with `SRILM` (Stolcke, 2002). Word alignment is done with `GIZA++` (Och and Ney, 2003) and both phrase extraction and decoding are done with the `Moses` package (Koehn et al., 2006; Koehn et al., 2007). Our model considers the language model, direct and inverse phrase probabilities, direct and inverse lexical probabilities, phrase and word penalties, and a non-lexicalised reordering. The optimisation of the weights of the model is trained with MERT (Och, 2003) against the BLEU (Papineni et al., 2002) evaluation metric.

A wider explanation of this system, the preprocess applied to the corpus before training the system and a deep evaluation of the translations can be found in España-Bonet et al. (2011a).

3.3 GF system

As explained in Section 2, the extension of GF to a new domain implies the construction of a specialised grammar that expands the general resource grammar. Since in our case of application we are far from a close and limited domain, some probabilistic components are also necessary. The general architecture is illustrated by Figure 1. A GF grammar-based system alone cannot parse most patent sentences. Consequently, the current translation system aims at using GF for translating patent chunks, and assemble the results in a later phase.

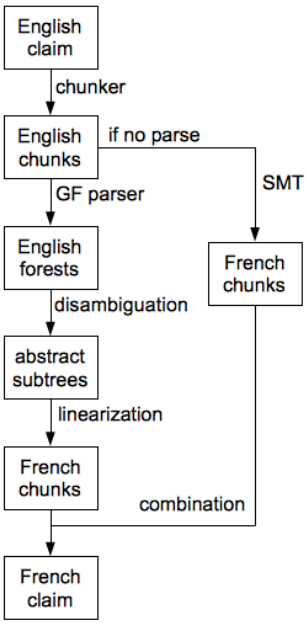

Figure 1: Architecture of the GF translation system.

As a pre-process, claims are tagged with part-of-speech (PoS) with Genia (Tsuruoka et al., 2005), a PoS tagger trained on the biomedical domain. From the PoS-tagged words only the ones labelled as nouns, adjectives, verbs and adverbs are kept, since the GF library already has an extensive list of functional parts of speech such as prepositions and conjunctions. We use the extensive GF English lexicon[5] as a lemmatiser for the PoS-tagged words, so that one can build their correspondent abstract syntax entry. Moreover, all the inflection forms of a given word are obtained from the same resource.

This process is made online. For every sentence to translate, the lexicon is enlarged with the corresponding vocabulary. The French version of the lexicon is built by translating the individual entries from the English lexicon (all inflection forms) with the SMT individual system trained on the patent

[4] http://www.epo.org/

[5] The GF English lexicon is based on the Oxford Advanced Learner's Dictionary, and contains around 50,000 English words.

corpus. The French translations are lemmatised with an extensive GF French lexicon, based on the large morphological lexicon Morphalou (Romary et al., 2004) in order to get their inflection table. The part-of-speech is assumed to be the same as in the English counterpart.

When this procedure is applied on the test set, the part-of-speech tagger is able to find 2,013 lexicon entries. However, due to part-of-speech mismatching or to the fact that a given word was not found in the SMT lexical table, 43.81% of the entries could not be translated to French.

In order to increase the coverage of the final GF translation, the grammar is adapted to deal with chunks instead of with full sentences. So, the source text is chunked into noun phrases (NP), adjective phrases (AP), adverbial and prepositional phrases (PP), relative pronouns (RP) and verb phrases (VP). Other kinds are ignored.

Some technical details have to be taken into account in order to build the patents grammar for chunks. Whereas NPs can be translated directly, a VP, RP or AP needs to have an NP to agree with, otherwise the GF grammar cannot know which linearisation form to choose. For NP and PP which can be translated independently, a mapping into corresponding GF categories is defined, whereas for VP, RP and AP, their GF mapping requires an NP in order to build their correspondent linearisation. If the required NP is not found, the chunk is sent to the SMT. Also, the VP category from the English and French GF resource grammars is implemented as a discontinuous category, so that it can handle discontinuous constituents in English and clitics in French. The patent grammar uses a category built on top of VP, which represents the flattened version of a VP, with all the constituents combined.

Because the syntactical structure of chunks is important in this case, a post-processing step is needed. This is meant to ensure that the PoS-tagging is consistent and that certain aspects captured in the grammar can be properly reflected in the claims. One can see the importance of this step with an example.

Ex1 *The use of claim 1 , wherein said use is intramuscular .*

In the previous example, "said", a frequent used word in patent claims, acts as a definite article, whereas Genia tags it as a verb and therefore is

Word	**PoS Genia**	**Chunk Genia**	**PoS Final**	**Chunk Final**
the	DT	B-NP	DT	B-NP
use	NN	I-NP	NN	I-NP
of	IN	B-PP	IN	I-NP
claim	NN	B-NP	NN	I-NP
1	CD	I-NP	CD	I-NP
,	,	O	,	O
wherein	IN	B-PP	RP	B-RP
said	V	B-VP	DT	B-NP
use	NN	B-NP	NN	I-NP
is	VBZ	B-VP	VBZ	B-VP
intramuscular	JJ	B-ADJP	JJ	I-VP
.	.	O	.	O

Table 1: Chunk detection for the example sentence Ex1.

it not merged with the following noun into a noun phrase. Moreover, the relative pronoun "wherein" is labelled as an adverb or noun phrase. The post-processing process updates the tags of certain entries and the tag of the following word, when needed.

Table 1 shows how the original tagging from Genia is converted into the correct GF parse chunks: *the use* (NP), *of claim 1* (PP), *wherein* (RP), *said use* (NP), *is intramuscular* (VP). As one can notice, chunks are merged when needed, like for the PP *of claim 1*, where the preposition was merged with the NP into a single chunk. The same goes for the VP chunk, as it is aimed to combine two-placed verbs or copulas with their objects before parsing.

GF parses the corresponding English chunks to obtain a forest of abstract syntax trees. In order to disambiguate among the possible options, all of them are linearised, looked up in the French corpus and the most frequent linearisation is kept as the best translation.

The translation sequence is done from left to right, so that the last-occurring NP is retained, and is used to make the agreement with VP, RP or AP. If no such NP can be found, or if the GF grammar is not capable to parse the one indicated by the chunker, the current chunk is passed to the SMT. In the working example, this in not necessary, and GF grammar alone obtains a translation for the full sentence:

1. *the use* → "l' utilisation" (NP)
2. *of claim 1* → "selon la revendication 1" (PP)
3. *wherein* → "dans laquelle" (RP agreeing with *"l' utilisation"*)

4. *said use* → "ladite utilisation" (NP)

5. *is intramuscular* → "est intramusculaire" (VP agreeing with *"ladite utilisation"*)

Finally, chunks are combined together with the punctuation marks, other non-included elements and untranslated chunks in the same order as in the source language.

3.4 Top SMT layer

The grammar-based translator already makes use of the SMT system trained on patents to translate the GF English lexicon. This way, the vocabulary is disambiguated towards the biomedical domain, but still there are non-parseable chunks with unknown vocabulary in the lexicon that cannot be translated using the grammar.

To gain robustness in the final system, the output of the GF translator is used as *a priori* information for a higher level SMT system. The SMT baseline is fed with phrases which are integrated in two different ways. In both cases SMT leads the translation since it is the system that chooses the final reordering of the translation, GF constraints parts of the translation.

Hard Integration (HI): Phrases with GF translation are forced to be translated this way. The system can reorder the chunks and translates the untranslated chunks, but there is no interaction between GF and pure SMT phrases.

Soft Integration (SI): Phrases with GF translation are included in the translation table with a certain probability so that the phrases coming from the two systems interact. Probabilities in the SMT system are estimated from frequency counts in the usual way; the probabilities in the GF system are a fixed value in the interval [0, 1] for all the phrases. This probability is given to the chunk translation pair as a whole, so when competing with SMT translations that have four translation probabilities (phrase-to-phrase and word-to-word in the two directions) the probability mass is divided among them to combine the systems in the translation table. Notice that a probability of one for a phrase does not imply a sure translation not only because of this, but also because at the end, the language model chooses the translation.

4 Results and discussion

The complete hybrid system and the individual components introduced in Section 3 are evaluated on the patents test set both automatic and manually.

	GF	**SMT**
NP	2,366 (14.9%)	2,199 (13.8%)
VP	275 (1.7%)	1,302 (8.2%)
AP	1,960 (12.3%)	1,935 (12.2%)
RP	648 (4.1%)	86 (0.5%)
Other	–	5,099 (32.0%)
Total	*5,301 (33.3%)*	*10,621 (66.7%)*

Table 2: Number and percentage of individual chunks translated by the HI system.

After the pre-process, the test set is divided in 15,922 chunks. From these chunks 33.3% can be translated using the GF patents grammar, and the remaining 66.7% must be passed to the SMT system. Table 2 shows the concrete percentages for every kind of chunk. Notice that GF only is designed to deal with the four most frequent types of chunks, and punctuation and conjunctions for example are ignored by GF. For these majority categories, GF can handle half of NP and AP, almost all RP but only 17.4% of VP.

There are several reasons why GF cannot translate the chunks. In 18.3% of the cases the chunks could not be parsed by the GF English grammar. When parsed, 15.5% of the chunks could not be translated due to missing words in the bilingual lexicon and to a lesser extent 1.1% could not be translated because of the missing information about agreement. 31.3% of the chunks are labelled as *Other* (punctuation marks, item markers, etc.) and ignored by GF.

Splitting the sentences in chunks proved to be crucial for the final translation. 84.7% of the fragments to be translated contained at least one chunk that could not be parsed by the English grammar, and even more, 93.1% of the fragments contained at least one chunk that could not be translated. So, the coverage of a GF translation at sentence level would be of only 6.9%. At chunk level the coverage increases up to 33.3%.

Still this limited coverage cannot compete with that of a statistical system. Table 3 reports an automatic evaluation using several lexical metrics for both GF and SMT individual systems (top rows). This set of metrics is a subset of the metrics available in the Asiya evaluation package (Giménez and Màrquez, 2010). For all the metrics the SMT sys-

	WER	PER	TER	BLEU	NIST	GTM-2	MTR-pa	RG-S*	ULC
GF	60.96	50.08	58.90	26.56	5.57	22.74	38.76	29.00	16.17
SMT	27.03	17.50	25.32	63.18	9.99	44.58	71.64	72.65	67.14
HI	33.56	21.95	31.24	55.88	9.24	38.81	67.30	67.80	58.84
SI1.0	26.76	17.39	25.10	63.56	10.02	**44.86**	**71.96**	72.89	67.56
SI0.5	**26.63**	**17.32**	**25.02**	**63.60**	**10.03**	44.84	71.94	**72.93**	**67.60**
SI0.0	27.08	17.48	25.36	63.15	9.99	44.54	71.60	72.66	67.11

Table 3: Automatic evaluation of the baselines and hybrid systems.

	SMT	Tied	SI0.5
Tester1	4	9	10
Tester2	3	13	7
Tester3	2	17	4
Tester4	6	5	12
Total	15	44	33

Table 4: Manual evaluation of the 23 different sentences from a random subset of 100 sentences.

tem beats the GF one in a significant way. This is mainly due to the coverage, SMT is able to translate the whole sentence which is not the case of GF. However, GF is able to deal with some grammatical issues that cannot be recovered statistically. The most evident example is agreement in gender and number. Contrary to English, French adjectives and nouns agree in gender and number and relative pronouns agree with their relative. This is taken into account by construction in GF so that mistaken SMT translations such as "le médicament séparée" is correctly translated as "le médicament séparé" (*the separate medicament*) or "composition pharmaceutique selon la revendication 1, dans lequel" is correctly translated as "composition pharmaceutique selon la revendication 1, dans lequelle" (*the pharmaceutical composition of claim 1, wherein*).

These are minor details from the point of view of the lexical evaluation metrics however, they make a difference to the reader. Although in few occasions the understanding of the sentence is compromised because of the lack of agreement, the fluency of the output is not harmed.

Therefore we incorporate these well-formed translations into the SMT system. A hard integration of the translations does not allow them to interact. GF translations are always used and the statistical decoder reorders them and completes the translation with its own phrase table. This system is named HI in Table 3. Results are below those of the SMT system because the system is being forced to use the high quality translations together with translations of elements not considered. Just to give an example, GF will highly benefit from incorporating a grammar to deal with compounds and numbers. Currently these elements typical of the domain are not specifically approached.

A softer integration of the translations is done by the family of systems denoted by SI in Table 3. In this case, GF translations are given a probability which ranges from null to one with the same value given to all the phrases. Several experiments have been carried out for different values in the interval. We show in the bottom rows of Table 3 just three of them: 0, 0.5 and 1. Relative probabilities between the systems result not to be as important as the fact of allowing the interaction.

The combination of all the phrases improves the translations according to all the lexical metrics considered. There is an increment of 0.42 points of BLEU, 0.30 of TER and 0.46 of ULC, an uniform linear combination of 13 variants of the metrics considered. Improvements are moderate because of two reasons. First, SMT translations are already good for a start. Second, the amount of issues that GF handles are limited to be reflected on automatic metrics.

We have conducted a manual evaluation of the translations. To do this, 100 sentences have been randomly selected and four evaluators have been asked to indicate the grammatically most syntactically correct translation between two options: the SMT translation and the SI0.5 hybrid translation. The main aspects that we evaluated were correct agreement and properly inflected words.

For the whole testing corpus, 78.47% of the sentences were identically translated by the SMT and HMT. For our manually tested corpus, we only in-

spected the 23 sentences where the systems had a different output. The results can be seen in Table 4. The hybrid system is better than the SMT one according to the four evaluators, and the improvements come from discrepancies in gender, number and agreement. The SMT translations were preferred in the cases where the hybrid translation failed to translate certain words, so that the final claim has a visible hole –which makes it syntactically incorrect.

Figure 2 shows an example sentence where these features are observed. GF is doing the gender agreement between noun and adjective correctly ("séparée" vs. "séparé") but is not able to translate the full sentence ("at the same time as"). The two hybrid systems in this case are able to construct the correct translation which coincides with the reference.

5 Conclusions and future work

This work presents a HMT system for patent translation. The system exploits the high coverage of statistical translators and the high precision of GF to deal with specific issues of the language.

At this moment the grammar tackles agreement in gender, number and between chunks, and reordering within the chunks. Although the cases where these problems apply are not extremely numerous both manual and automatic evaluations consistently show their preference for the hybrid system in front of the two individual translators.

The coverage of the grammar can be extended in order to deal with more typical structures present in patent documents. The coverage of VP is particularly low because of the missing verbs from the French lexicon and the syntactically complex verb phrases –such as cascades of nested verbs, which are not handled by the patents grammar yet. Also, a grammar to translate compounds will be included as they are a significant part of the biomedical documents. Moreover, the grammar component can be extended to handle the ordering at sentence level besides of the reordering within the chunks. This is specially interesting to deal with languages like German where the structure of the sentence is different from the structure in English for example.

The previous improvements will increase the number of chunks that can be parsed by the grammar; in order to increase the percentage of translations it is also necessary to improve the lexicon building procedure. An obvious improvement would be a bilingual dictionary of idioms, so that the translation would not just map word-to-word, but also phrase-to-phrase.

Finally, we plan to implement another version of the hybrid system where GF grammars are applied at an later stage –after the English chunks are translated into French by the SMT system. The GF grammars will be used to to restore the agreement for chunks like VP, RP and AP, like before. The main difference is that due to an earlier use of SMT, one can capture idiomatic constructions better, and use GF just in the end for improving syntactic correctness.

Acknowledgements

This work has been partially funded by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 247914 (MOLTO project, FP7-ICT-2009-4-247914).

References

Angelov, Krasimir. 2011. *The Mechanics of the Grammatical Framework*. Ph.D. thesis, Chalmers University of Technology, Gothenburg, Sweden.

Ceausu, A., J. Tinsley, A. Way, J. Zhang, and P. Sheridan. 2011. Experiments on Domain Adaptation for Patent Machine Translation in the PLuTO project. *Proceedings of the 15th Annual Conference of the European Association for Machine Translation (EAMT 2011)*.

Chen, Y. and A. Eisele. 2010. Hierarchical hybrid translation between english and german. In Hansen, Viggo and Francois Yvon, editors, *Proceedings of the 14th Annual Conference of the European Association for Machine Translation*, pages 90–97. EAMT, EAMT, 5.

Ehara, Terumasa. 2007. Rule based machine translation combined with statistical post editor for japanese to english patent translation. *MT Summit XI Workshop on patent translation, 11 September 2007, Copenhagen, Denmark*, pages 13–18.

Ehara, Terumasa. 2010. Statistical Post-Editing of a Rule-Based Machine Translation System. *Proceedings of NTCIR-8 Workshop Meeting, June 1518, 2010*, pages 217–220.

Eisele, A., C. Federmann, H. Saint-Amand, M. Jellinghaus, T. Herrmann, and Y. Chen. 2008. Using moses to integrate multiple rule-based machine translation engines into a hybrid system. In *Proceedings of the Third Workshop on Statistical Machine Translation*, StatMT '08, pages 179–182.

GF	Une utilisation selon la revendication 3, dans laquelle le médicament séparé est administré at the same time as...
SMT	Utilisation selon la revendication 3, dans laquelle le médicament séparée est administré en même temps que...
HI	Une utilisation selon la revendication 3, dans laquelle le médicament séparé est administré en même temps que...
SI0.5	Utilisation selon la revendication 3, dans laquelle le médicament séparé est administré en même temps que...
Ref.	Utilisation selon la revendication 3, dans laquelle le médicament **séparé** est administré **en même temps que**...

Figure 2: Example where GF translates with the correct gender of the adjective and the SMT completes the untraslated words.

España-Bonet, C., R. Enache, A. Slaski, A. Ranta, L. Màrquez, and M. Gonzàlez. 2011a. Patent translation within the molto project. In *Proceedings of the 4th Workshop on Patent Translation, MT Summit XIII*, pages 70–78, Xiamen, China, sep.

España-Bonet, C., G. Labaka, A. Díaz de Ilarraza, L. Màrquez, and K. Sarasola. 2011b. Hybrid machine translation guided by a rule-based system. In *Proceedings of the 13th Machine Translation Summit*, pages 554–561, Xiamen, China, sep.

Federmann, C., A. Eisele, Y. Chen, S. Hunsicker, J. Xu, and H. Uszkoreit. 2010. Further experiments with shallow hybrid mt systems. In *Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR*, pages 77–81, July.

Giménez, Jesús and Lluís Màrquez. 2010. Asiya: An Open Toolkit for Automatic Machine Translation (Meta-)Evaluation. *The Prague Bulletin of Mathematical Linguistics*, (94):77–86.

Habash, N., B. Dorr, and C. Monz. 2009. Symbolic-to-statistical hybridization: extending generation-heavy machine translation. *Machine Translation*, 23:23–63.

Koehn, Philipp, Wade Shen, Marcello Federico, Nicola Bertoldi, Chris Callison-Burch, Brooke Cowan, Chris Dyer, Hieu Hoang, Ondrej Bojar, Richard Zens, Alexandra Constantin, Evan Herbst, and Christine Moran. 2006. Open Source Toolkit for Statistical Machine Translation. Technical report, Johns Hopkins University Summer Workshop. http://www.statmt.org/jhuws/.

Koehn, Philipp, Hieu Hoang, Alexandra Birch Mayne, Christopher Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondrej Bojar, Alexandra Constantin, and Evan Herbst. 2007. Moses: Open source toolkit for statistical machine translation. In *Annual Meeting of the Association for Computation Linguistics (ACL), Demonstration Session*, pages 177–180, Jun.

Och, Franz Josef and Hermann Ney. 2003. A systematic comparison of various statistical alignment models. *Computational Linguistics*, 29(1):19–51.

Och, Franz Josef. 2003. Minimum error rate training in statistical machine translation. In *Proc. of the Association for Computational Linguistics*, Sapporo, Japan, July 6-7.

Papineni, Kishore, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the Association of Computational Linguistics*, pages 311–318.

Ranta, Aarne. 2009. The GF resource grammar library. *Linguistic Issues in Language Technology*, 2(1).

Ranta, Aarne. 2011. *Grammatical Framework: Programming with Multilingual Grammars*. CSLI Publications, Stanford. ISBN-10: 1-57586-626-9 (Paper), 1-57586-627-7 (Cloth).

Romary, Laurent, Susanne Salmon-Alt, and Gil Francopoulo. 2004. Standards going concrete: from lmf to morphalou. In *Proceedings of the Workshop on Enhancing and Using Electronic Dictionaries*, ElectricDict '04, pages 22–28, Stroudsburg, PA, USA. Association for Computational Linguistics.

Sheremetyeva, Svetlana. 2003. Natural language analysis of patent claims. In *Proceedings of the ACL-2003 workshop on Patent corpus processing - Volume 20*, PATENT '03, pages 66–73, Stroudsburg, PA, USA. Association for Computational Linguistics.

Sheremetyeva, Svetlana. 2005. Less, Easier and Quicker in Language Acquisition for Patent MT. *MT Summit X, Phuket, Thailand, September 16, 2005, Proceedings of Workshop on Patent Translation*, pages 35–42.

Sheremetyeva, Svetlana. 2009. On Extracting Multiword NP Terminology for MT. *EAMT-2009: Proceedings of the 13th Annual Conference of the European Association for Machine Translation, ed. Lluís Màrquez and Harold Somers, 14-15 May 2009, Universitat Politcnica de Catalunya, Barcelona, Spain*, pages 205–212.

Stolcke, A. 2002. SRILM – An extensible language modeling toolkit. In *Proc. Intl. Conf. on Spoken Language Processing*.

Thurmair, G. 2009. Comparing different architectures of hybrid machine translation systems. In *Proc MT Summit XII*.

Tsuruoka, Y., Y. Tateishi, J.D. Kim, T. Ohta, J. McNaught, S. Ananiadou, and J. Tsujii. 2005. Developing a robust part-of-speech tagger for biomedical text. In Bozanis, P. and eds. Houstis, E.N., editors, *Advances in Informatics.*, volume 3746, page 382392. Springer Berlin Heidelberg.