8.14: Annual Public Report

S. Busemann, O. Bojar, C. Callison-Burch, M. Cettolo, R. Garabik, J. van Genabith, P. Koehn, H. Schwenk, K. Simov, P. Wolf

Distribution: Public

EuroMatrixPlus
Bringing Machine Translation
for European Languages to the User

ICT 231720 Deliverable 8.14

Project funded by the European Community
under the Seventh Framework Programme for
Research and Technological Development.

Project ref no.	ICT-231720
Project acronym	EuroMatrixPlus
Project full title	Bringing Machine Translation for European Languages to the User
Instrument	STREP
Thematic Priority	ICT-2007.2.2 Cognitive systems, interaction, robotics
Start date / duration	01 March 2009 / 38 Months

Distribution	Public
Contractual date of delivery	November 15, 2011
Actual date of delivery	December 05, 2011
Date of last update	December 5, 2011
Deliverable number	8.14
Deliverable title	Annual Public Report
Type	Report
Status & version	Final
Number of pages	13
Contributing WP(s)	all
WP / Task responsible	DFKI
Other contributors	All Participants
Author(s)	S. Busemann, O. Bojar, C. Callison-Burch, M. Cettolo, R. Garabik, J. van Genabith, P. Koehn, H. Schwenk, K. Simov, P. Wolf
EC project officer	Michel Brochard
Keywords	

The partners in EuroMatrixPlus are:
DFKI GmbH, Saarbrücken (DFKI)
University of Edinburgh (UEDIN)
Charles University (CUNI-MFF)
Johns Hopkins University (JHU)
Fondazione Bruno Kessler (FBK)
Université du Maine, Le Mans (LeMans)
Dublin City University (DCU)
Lucy Software and Services GmbH (Lucy)
Central and Eastern European Translation, Prague (CEET)
Ludovit Stur Institute of Linguistics,
Slovak Academy of Sciences (LSIL)
Institute of Information and Communication Technologies,
Bulgarian Academy of Sciences (IICT-BAS)

For copies of reports, updates on project activities and other EuroMatrixPlus-related information, contact:

The EuroMatrixPlus Project Co-ordinator
Prof. Dr. Hans Uszkoreit, DFKI GmbH
Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany
uszkoreit@dfki.de
Phone +49 (681) 85775-5282 - Fax +49 (681) 85775-5338

Copies of reports and other material can also be accessed via the project's homepage: http://www.euromatrixplus.net/

Chapter 1

General Remarks

This document describes the highlights of EuroMatrixPlus achieved during its third project year.

The Moses toolkit remains and grows as the leading platform for open source machine translation. We observe informally a growing use of the toolkit in industry, especially the language service provider space, where it emerges as a disruptive technology with wide impact. The integration of novel methods into the open source Moses toolkit is an ongoing effort – mainly done in parallel with the development of these methods. The main advances were in the area of machine learning methods and efficient model estimation and inference, especially for tree-based models.

EuroMatrixPlus continued to gain broad visibility within academia and industry by again organizing several workshops and documenting their results. A key event is the shared task evaluation series that was continued with the WMT event. The Third Joint EM+/CNGL Workshop "Bringing MT to the User: Research Meets Translators" was highly successful, focussing on the use of MT in translation workflows and specifically the use of MT and other translation technologies by professional translators. The tradition to annually organize one-week MT marathons started in the predecessor project EuroMatrix was continued further in 2011. The Sixth MT marathon was held in September in Trento (Italy). It included lectures, talks and lab sessions, addressing a broad audience.

The next chapter overviews the significant research results achieved in the individual work packages during the reporting period. The output of MT systems and 44 refereed publications demonstrates the impact of EuroMatrixPlus on machine translation.

Chapter 2

Significant Results

WP1: Rich Tree-Based Statistical Translation

The annotation of the Prague Czech-English Dependency Treebank 2.0 has been accomplished: 49,000 pairs sentence are now equipped with manual tectogrammatical trees in both Czech and English. Further data checks are under way, documentation is being written and the CD packaging is soon to be polished.

The implementation Treex platform (Žabokrtský, 2011), previously called TectoMT, was significantly restructured for better modularity and thoroughly tested during WMT evaluation campaign (Popel et al., 2011; Mareček et al., 2011). A preliminary version of Treex has been released to CPAN with further improvements and unit tests to come in the following months.

A hierarchical lexical reordering was implemented in Moses and employed for developing the system used in WMT 2011 evaluation campaign. Moreover, the investigation on linguistically informed reordering models has continued.

Publications related to this WP are (Hardmeier et al., 2011; Bisazza et al., 2011).

WP2: Hybrid Machine Translation

The main effort during this period of the project was spent on improving the hybrid modules developed previously.

As the terminology extraction was deemed most useful for expanding the lexicon of the rule-based engine, the tool was evaluated and refined. Importing malformed terms, for example incorrectly inflected terms, proved to be most harmful to the translation process. As such, the term lists were inspected for incorrect entries and the both the filtering as well as the post-processing was improved to yield terms of better quality.

Furthermore, we enquired into additional features for the tree selection process. LUCY adapted the available output to provide more information.

Results of this work package was published in (Wolf et al., 2011; Federmann and Hunsicker, 2011a; Federmann and Hunsicker, 2011b).

WP3: Advanced Learning Methods for Machine Translation

During this period of the project, all tasks of WP3 have progressed towards the respective goals. In short, the SMT systems for the translation from Arabic to Italian via English setup in a previous stage of the project have been refined. Procedures for the extraction of parallel data from comparable corpora have been made more efficient, and our unsupervised training algorithms have been improved. A method improving target-side vocabulary coverage for morphologically rich languages using target-side monolingual data was designed and tested. Moreover, techniques developed within the project for data selection and for handling of pronominal anaphora have been employed in the WMT 2011 evaluation campaign. We also worked on making more

effective the tuning of SMT systems, on topic adaptation via bilingual latent semantic models and on a fill-up technique for adapting translation models. The continuous space language model toolkit has been improved too. The algorithms resulting from the research in the workpackage were very successfully employed in the systems developed by the partners for the IWSLT 2011 evaluation campaign.

Publications related to this WP are (Bisazza et al., 2011; Bojar and Tamchyna, 2011; Cettolo et al., 2011a; Cettolo et al., 2011b; Hardmeier et al., 2011; Lambert et al., 2011; Macháček and Bojar, 2011; Rousseau et al., 2011; Ruiz and Federico, 2011; Schwenk et al., 2011)

WP4: Open Source Tools and Data

The Moses toolkit remains and grows as the leading platform for open source machine translation. We observe informally a growing use of the toolkit in industry, especially the language service provider space, where it emerges as a disruptive technology with wide impact.

The integration of novel methods into the open source Moses toolkit is an ongoing effort — mainly done in parallel with the development of these methods. The main advances were in the area of machine learning methods and efficient model estimation and inference, especially for tree-based models.

Further minor improvements of the Moses toolkit include eppex, a faster approximate alternative to phrase extraction (Przywara and Bojar, 2011), and the support of new metrics during MERT parameter optimisation (Servan and Schwenk, 2011).

We organized the 6th MT Marathon, hosted by FBK on September 5-10, 2011[1].

We continued our efforts to acquire parallel training data, including a focused web crawl effort for the language pairs used in the evaluation campaign. We plan to release this data for the last evaluation campaign at the end of the EuroMatrixPlus project.

The parallel corpus CzEng is now being updated and a new release 1.0 (larger in size but also more thoroughly filtered to avoid wrong sentence pairs) will be available for the WMT12 evaluation campaign.

WP5: "WikiTrans" Community-Based Translation Environments

During this period of the project, all tasks of WP5 have progressed towards the respective goals. In short, We have been investigating methods for optimising systems for specific domains, which may be (a) predefined, (b) automatically obtained by clustering, (c) user-specific domains. Domain ranges in granularity from genres to single documents. This work is ongoing through the end of the project.

We have continued our ongoing development of computer aided translation tools. We developed word-level confidence measures with the aim of highlighting error-prone parts of the machine translation in a post-editing scenario to such tools.

We have been making progress on the platform for non-expert translators to contribute translations of Wikipedia articles. We have investigated how a large number of non-expert translators can collaborate to produce translations that approach the level of professional translations (inspired by the Wikipedia collaborative model). We expect to extend the platform to collect translations of Wikipedia articles for low-resource languages. To investigate syntactic models of translation we will target verb-final languages such as those spoken on the Indian subcontinent or in Southeast Asia.

Publications related to this WP are (Walia, 2011; Zaidan and Callison-Burch, 2011; Ahn et al., 2011)

[1] http://mtmarathon2011.fbk.eu/

WP6: Integrated Localisation Workflow

The Third Joint EuroMatrixPlus–CNGL Workshop (JEC 2011) on "Bringing MT to the User: Research Meets Translators"[2] was a noted success. The workshop was co-organised by the EuroMatrixPlus Project, the Centre for Next Generation Localisation (CNGL)[3], the Directorate-General for Translation (DGT, European Commission)[4] and Autodesk[5], and hosted by the DGT in Luxembourg on October 14th, 2011. The objective of the workshop series is to bring together translators, users, academic and industrial MT researchers and developers to discuss issues that are most important in real world industrial settings and applications involving MT, but currently under-represented in research circles. The focus of the 2011 workshop was on the use of MT in translation workflows and specifically the use of MT and other translation technologies by professional translators.[6]

Yanjun Ma, Yifan He and Josef van Genabith presented a tutorial "From MT Confidence Estimation to the Integration of MT and TM Technologies" at the MT-Summit XIII in Xiamen, China, 19–23 September 2011.[7]

Work has progressed on *tight* TM/MT (Translation Memory/Machine Translation) integration: while some of our previous work (He et al., 2010a; He et al., 2010b) concentrated *loose* coupling between TM and MT (recommendation and ranking), (Ma et al., 2011a; Ma et al., 2011b) show that deep linguistic information (including syntactic dependencies) and sophisticated machine learning approaches (in particular tree kernels) improve TM/MT integration: to give an example, we use weighted tree kernels to identify TM segments which contain useful matching chunks to constrain the MT system (in a tight coupling approach), based on the intuition that divergences higher up in fuzzy dependency matches are more "severe" than divergences lower down the dependency tree.

Work has progressed on statistical post-editing (SPE): (Bechara et al., 2011) show that SPE using PB-SMT as both first and second-stage systems can improve translation quality in an approach that uses source context information and alignment-based thresholding.

Publications related to this WP include (Ma et al., 2011a; Ma et al., 2011b; Bechara et al., 2011).

WP7: Evaluation Campaign

In the first evaluation workshop, we performed a large-scale manual evaluation of 148 machine translation systems and 41 system combination entries. We used the ranking of these systems to measure how strongly automatic metrics correlate with human judgments of translation quality for 21 metrics. The proceedings of the workshop resulted in 18 publications on various scientific topics related to machine translation, along with 47 short papers describing the machine translation systems and automatic evaluation metrics that were submitted to the shared task.

The manual evaluation of WMT shared tasks was thoroughly examined and several suggestions for next years were proposed (Bojar et al., 2011a). We were also experimenting with other manual and automatic techniques of MT evaluation (Bojar, 2011; Berka et al., 2011).

The internal evaluation for the second year of the project was completed and successful. The internal evaluation for the third year is scheduled for the beginning of February 2012 to coincide with the public evaluation campaign. The results are summarized in the following table:

[2] http://web.me.com/emcnglworkshop/JEC2011/Home.html
[3] http://cngl.ie
[4] http://ec.europa.eu/dgs/translation
[5] http://www.autodesk.ch
[6] Proceedings are available at http://web.me.com/emcnglworkshop/JEC2011/Program.html.
[7] http://mt.xmu.edu.cn/mtsummit/

	Best System Score Year 1	Best System Score Year 2	Percent improvement
Czech → English	58%	60%	+2%
English → Czech	64%	65%	+1%
German → English	63%	60%	-3%
English → German	56%	60%	+4%
French → English	61%	64%	+3%
English → French	64%	64%	+0%

Our annual goal is to satisfy the following two criteria:

1. The best performing system should be better than the previous year's best system for every language pair.

2. The best system should be at least 5% better for at least half of the language pairs.

The best performing system is designated as the system where the translations are judged to be better than some other system the greatest number of times. A 5% improvement indicates that the system will be judged to be better than the previous best system by 5% of the sentences.

We satisfied criteria 1 except for the case of German → English, where we failed to see improvement (the drop in score is not statistically significant). All other language pairs showed modest gains, with cumulative average of 2.0% across all pairs. We thus did not satisfy criterion 2 this year. This may be due in part to the fact that the gains between years 0 and 1 were quite large, increasing the difficulty of achieving further gains in the current year.

Publications directly funded in this WP are (Callison-Burch et al., 2011b; Bojar et al., 2011a; Callison-Burch et al., 2011a; Williams and Koehn, 2011; Haddow et al., 2011; Macháček and Bojar, 2011; Bojar and Tamchyna, 2011; Mareček et al., 2011; Zaidan, 2011; Levenberg et al., 2011; Hardmeier et al., 2011; Schwenk et al., 2011; Weese et al., 2011; Xu et al., 2011; Lewis et al., 2011). All publications in WMT 2011 (Callison-Burch et al., 2011b) are related to this WP in using the data we created.

WP8: Project Management and Dissemination

Dissemination activities were carried out by organizing

- Regular evaluation campaigns with the exchange of all the translation results and the results of human evaluations help to see and understand the differences and relative advantages of the different approaches and motivate all groups to engage in collaborations towards better, integrated approaches.

- Machine translation marathons, which form are a very effective instrument that helps to increase the mutual understanding of teams that address the MT problem from different backgrounds and work together on shared projects intensively for a week with subsequent collaborations for much longer time.

In the reporting period the project has co-organized several major events with participation from academia and industry:

- Sixth Workshop on Statistical Machine Translation, July 2011, Edinburgh (UK)[8]

- Sixth MT Marathon, September 2011, Trento (Italy)[9]

- Third Joint EM+/CNGL Workshop: "Bringing MT to the User: Research meets Translators", October 2011, Luxembourg[10]

[8] http://www.statmt.org/wmt11/
[9] http://mtmarathon2011.fbk.eu/
[10] http://web.me.com/emcnglworkshop/JEC2011/Home.html

The Seventh Workshop on Statistical Machine Translation is co-located with NAACL in Montreal, June 2012.

In its third year, the project has generated 44 scientific publications in refereed conference proceedings, journals or books. They are listed at the end of this report.

More information about EuroMatrixPlus can be found on the project website.[11]

WP9: Integrating Slovak Language Resources into the EUROMATRIXPLUS Framework

The parallel Enlish-Slovak corpus of fiction has been successfully compiled (1.5 million sentence pairs); together with the parallel Czech-Slovak corpus (720 thousand sentence pairs) the corpora search interface is open to public on the webpage of the Ľ. Štúr Institute of Linguistics.

The corpora have been used to train the MOSES MT system (both language pairs) and to evaluate the translations. Two other MT systems for translation between closely related languages (Czech and Slovak) are being evaluated as well – Česílko and Česílko 2. Several automatic scores (BLEU, METEOR, NIST, TER and TERp) and three different data sets were used for the evaluation. Two different types of manual evaluation have been used – various types of errors were manually counted and the outputs of the MT systems were manually ranked by three annotators.

Publications related to this WP are (Galuščáková and Bojar, 2011; Bojar et al., 2011b).

WP10: HPSG-based Statistical Translation

In WP10 we have concentrated on the error analysis from the first experiments with the system Moses, trained on Bulgarian-English aligned data. We have identified the following shortcomings: (1) low coverage on lexical level because of the rich Bulgarian morphology; (2) unrecognized named entities; (3) word order errors; and (4) lexical meaning mismatches between the original and the translation.

In order to handle some of these errors, we focused on the extension of the processing tools for Bulgarian to cover the whole SETIMES corpus. This includes: (1) processing of most frequent missing words and named enties; (2) tuning the lemmatizer and the MRS analyser to the new lexical material. We have also started the processing of the Bulgarian part of SETIMES corpus. This will result in alignments on sentence level of Bulgarian processed text and English text. We will perform experiments for training of statistical machine translation on these parallel data.

The processing of the English part of SETIMES corpus was discussed with Dan Flickinger from Saarbrücken during his visit in Sofia.

Publications related to this WP are (Simov and Osenova, 2011), (Simov et al., 2011), (Osenova, 2011), and (Savkov et al., 2011). Additionally, we have an accepted paper for the Workshop on Treebanks and Linguistic Theories 2012, Heidelberg, Germany — (Simov and Osenova, 2012).

[11] http://www.euromatrixplus.net

Publications

Byung Gyu Ahn, Benjamin Van Durme, and Chris Callison-Burch. 2011. WikiTopics: What is popular on Wikipedia and why. In *Proceedings of the Workshop on Automatic Summarization for Different Genres, Media, and Languages*, pages 33–40, Portland, Oregon, June. Association for Computational Linguistics.

Hanna Bechara, Yanjun Ma, and Josef van Genabith. 2011. Post-editing for a statistical MT system. In *Machine Translation Summit XIII*, Xiamen, China.

Jan Berka, Martin Černý, and Ondřej Bojar. 2011. Quiz-based evaluation of machine translation. *The Prague Bulletin of Mathematical Linguistics*, 95:77–86.

Arianna Bisazza, Daniele Pighin, and Marcello Federico. 2011. Chunk-lattices for verb reordering in Arabic-English statistical machine translation. *Machine Translation*, Accepted for publication.

Ondřej Bojar and Aleš Tamchyna. 2011. Improving translation model by monolingual data. In Chris Callison-Burch, Philipp Koehn, Christof Monz, and Omar Zaidan, editors, *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 330–336, Edinburgh, UK. Association for Computational Linguistics.

Ondřej Bojar, Miloš Ercegovčević, Martin Popel, and Omar Zaidan. 2011a. A grain of salt for the WMT manual evaluation. In Chris Callison-Burch, Philipp Koehn, Christof Monz, and Omar Zaidan, editors, *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 1–11, Edinburgh, UK. Association for Computational Linguistics.

Ondřej Bojar, Petra Galuščáková, and Miroslav Týnovský. 2011b. Evaluating quality of machine translation from Czech to Slovak. In Markéta Lopatková, editor, *Information Technologies – Applications and Theory*, pages 3–9.

Ondřej Bojar. 2011. Analyzing error types in English-Czech machine translation. *The Prague Bulletin of Mathematical Linguistics*, 95:63–76.

Chris Callison-Burch, Philipp Koehn, Christof Monz, and Omar Zaidan. 2011a. Findings of the 2011 Workshop on Statistical Machine Translation. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 22–64, Edinburgh, Scotland, July. Association for Computational Linguistics.

Chris Callison-Burch, Philipp Koehn, Christof Monz, and Omar F. Zaidan, editors. 2011b. *Proceedings of the Sixth Workshop on Statistical Machine Translation*. Association for Computational Linguistics, Edinburgh, Scotland, July.

Mauro Cettolo, Nicola Bertoldi, and Marcello Federico. 2011a. Bootstrapping Arabic-Italian SMT through comparable texts and pivot translation. In *Proceedings of EAMT*, Leuven, Belgium.

Mauro Cettolo, Nicola Bertoldi, and Marcello Federico. 2011b. Methods for smoothing the optimizer instability in SMT. In *Proceedings of MT-Summit*, Xiamen, China.

Christian Federmann and Sabine Hunsicker. 2011a. Improving an existing RBMT system by stochastic analysis. In *Hamburg Working Paper in Multilingualism*, Hamburg Working Paper in Multilingualism.

Christian Federmann and Sabine Hunsicker. 2011b. Stochastic parse tree selection for an existing RBMT system. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 351–357, Edinburgh, Scotland, July. Association for Computational Linguistics.

Petra Galuščáková and Ondřej Bojar. 2011. Czech-Slovak parallel corpora. In *Proceedings of*

Slovko 2011, Natural Language Processing, Multilinguality, pages 65–71, Bratislava, Slovakia. Slovenská akadémia vied.

Barry Haddow, Abhishek Arun, and Philipp Koehn. 2011. Samplerank training for phrase-based machine translation. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 261–271, Edinburgh, Scotland, July. Association for Computational Linguistics.

Christian Hardmeier, Jörg Tiedemann, Markus Saers, Marcello Federico, and Mathur Prashant. 2011. The Uppsala-FBK systems at WMT 2011. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 372–378, Edinburgh, Scotland, July. Association for Computational Linguistics.

Yifan He, Yanjun Ma, Josef van Genabith, and Andy Way. 2010a. Bridging SMT and TM with translation recommendation. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 622–630, Uppsala, Sweden.

Yifan He, Yanjun Ma, Andy Way, and Josef van Genabith. 2010b. Integrating N-best SMT outputs into a TM system. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING-2010:Posters)*, pages 374–382, Beijing, China.

Patrik Lambert, Holger Schwenk, Christophe Servan, and Sadaf Abdul-Rauf. 2011. Investigations on translation model adaptation using monolingual data. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 284–293, Edinburgh, Scotland, July. Association for Computational Linguistics.

Abby Levenberg, Miles Osborne, and David Matthews. 2011. Multiple-stream language models for statistical machine translation. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 177–186, Edinburgh, Scotland, July. Association for Computational Linguistics.

William Lewis, Robert Munro, and Stephan Vogel. 2011. Crisis MT: Developing a cookbook for MT in crisis situations. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 501–511, Edinburgh, Scotland, July. Association for Computational Linguistics.

Yanjun Ma, Yifan He, Andy Way, and Josef van Genabith. 2011a. Consistent translation using discriminative learning: A translation memory-inspired approach. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 1239–1248, Portland, Oregon.

Yanjun Ma, Yifan He, Andy Way, and Josef van Genabith. 2011b. Features for translation memory-inspired consistent translation. In *Machine Translation Summit XIII*, Xiamen, China.

Matouš Macháček and Ondřej Bojar. 2011. Approximating a deep-syntactic metric for MT evaluation and tuning. In Chris Callison-Burch, Philipp Koehn, Christof Monz, and Omar Zaidan, editors, *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 92–98, Edinburgh, UK. Association for Computational Linguistics.

David Mareček, Rudolf Rosa, Petra Galuščáková, and Ondřej Bojar. 2011. Two-step translation with grammatical post-processing. In Chris Callison-Burch, Philipp Koehn, Christof Monz, and Omar Zaidan, editors, *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 426–432, Edinburgh, UK. Association for Computational Linguistics.

Petya Osenova. 2011. Localizing a core HPSG-based grammar for Bulgarian. In *Proceedings of GSCL 2011: Multilingual Resources and Multilingual Applications*, pages 175–180, Hamburg, Germany, September.

Martin Popel, David Mareček, Nathan Green, and Zdeněk Žabokrtský. 2011. Influence of parser choice on dependency-based MT. In Chris Callison-Burch, Philipp Koehn, Christof Monz, and Omar Zaidan, editors, *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 433–439, Edinburgh, UK. Association for Computational Linguistics.

Česlav Przywara and Ondřej Bojar. 2011. eppex: Epochal phrase table extraction for statistical machine translation. *The Prague Bulletin of Mathematical Linguistics*, 96:89–98.

Anthony Rousseau, Fethi Bougares, Paul Deléglise, Holger Schwenk, and Yannick Estève. 2011.

LIUM's systems for the IWSLT 2011 speech translation task. In *International Workshop on Spoken Language Translation (IWSLT)*.

Nick Ruiz and Marcello Federico. 2011. Topic adaptation for lecture translation through bilingual latent semantic models. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 294–302, Edinburgh, Scotland, July. Association for Computational Linguistics.

Aleksandar Savkov, Laska Laskova, Petya Osenova, Kiril Simov, and Stanislava Kancheva. 2011. A web-based morphological tagger for Bulgarian. In *Proceedings of Slovko 2011. Sixth International Conference. NLP, Multilinguality*, Modra/Bratislava, Slovakia, October.

Lane Schwartz, Chris Callison-Burch, William Schuler, and Stephen Wu. 2011. Incremental syntactic language models for phrase-based translation. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 620–631, Portland, Oregon, USA, June. Association for Computational Linguistics.

Holger Schwenk, Patrik Lambert, Loïc Barrault, Christophe Servan, Sadaf Abdul-Rauf, Haithem Afli, and Kashif Shah. 2011. LIUM's SMT machine translation systems for WMT 2011. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 464–469, Edinburgh, Scotland, July. Association for Computational Linguistics.

Christophe Servan and Holger Schwenk. 2011. Optimising multiple metrics with MERT. *The Prague Bulletin of Mathematical Linguistics (PBML)*.

Kiril Simov and Petya Osenova. 2011. Towards Minimal Recursion Semantics over Bulgarian dependency parsing. In *Proceedings of RANLP*, Hissar, Bulgaria, September.

Kiril Simov and Petya Osenova. 2012. Bulgarian-english treebank: Design and implementation. In *Proceedings of the International Workshop on Treebanks and Linguistic Theories 2012. Special Issue of the Linguistic Issues in Language Technology (LiLT) Online Journal, volume: 7, issue: 14*, Heidelberg, Germany, January.

Kiril Simov, Petya Osenova, Laska Laskova, Aleksandar Savkov, and Stanislava Kancheva. 2011. Bulgarian-English parallel treebank: Word and semantic level alignment. In *Proceedings of Second Workshop on Annotation and Exploitation of Parallel Corpora*, Hissar, Bulgaria, September.

Anmol Walia. 2011. Word level confidence estimation for statistical machine translation. Technical report, August. MSc Thesis, University of Edinburgh, School of Informatics.

Jonathan Weese, Juri Ganitkevitch, Chris Callison-Burch, Matt Post, and Adam Lopez. 2011. Joshua 3.0: Syntax-based machine translation with the thrax grammar extractor. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 478–484, Edinburgh, Scotland, July. Association for Computational Linguistics.

Philip Williams and Philipp Koehn. 2011. Agreement constraints for statistical machine translation into german. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 217–226, Edinburgh, Scotland, July. Association for Computational Linguistics.

Petra Wolf, Ulrike Bernardi, Christian Federmann, and Sabine Hunsicker. 2011. From statistical term extraction to hybrid machine translation. In *Proceedings of the 15th Annual Conference of the European Association for Machine Translation*. EAMT, EAMT, 5.

Jia Xu, Hans Uszkoreit, Casey Kennington, David Vilar, and Xiaojun Zhang. 2011. DFKI hybrid machine translation system for WMT 2011 – on the integration of SMT and RBMT. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 485–489, Edinburgh, Scotland, July. Association for Computational Linguistics.

Zdeněk Žabokrtský. 2011. Treex – an open-source framework for natural language processing. In Markéta Lopatková, editor, *Information Technologies – Applications and Theory*, volume 788, pages 7–14.

Omar F. Zaidan and Chris Callison-Burch. 2011. Crowdsourcing translation: Professional quality from non-professionals. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 1220–1229, Portland, Oregon, USA, June. Association for Computational Linguistics.

Omar Zaidan. 2011. MAISE: A flexible, configurable, extensible open source package for mass

AI system evaluation. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, pages 130–134, Edinburgh, Scotland, July. Association for Computational Linguistics.