



The Prague Bulletin of Mathematical Linguistics

NUMBER 102 OCTOBER 2014 93-104

Multilingual Dependency Parsing: Using Machine Translated Texts instead of Parallel Corpora

Loganathan Ramasamy, David Mareček, Zdeněk Žabokrtský

Charles University in Prague, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics

Abstract

This paper revisits the projection-based approach to dependency grammar induction task. Traditional cross-lingual dependency induction tasks one way or the other, depend on the existence of bitexts or target language tools such as part-of-speech (POS) taggers to obtain reasonable parsing accuracy. In this paper, we transfer dependency parsers using only approximate resources, i.e., machine translated bitexts instead of manually created bitexts. We do this by obtaining the the source side of the text from a machine translation (MT) system and then apply transfer approaches to induce parser for the target languages. We further reduce the need for the availability of labeled target language resources by using unsupervised target tagger. We show that our approach consistently outperforms unsupervised parsers by a bigger margin (8.2% absolute), and results in similar performance when compared with delexicalized transfer parsers.

1. Introduction

Inducing dependency structures has been an important topic of research within the parsing community for many years. Dependency parsers that can produce dependency structures for novel sentences often rely on manually constructed treebanks for training the parsers. Unlike other annotation tasks such as POS tagging, treebank annotation is much more complex and expensive. Zeman et al. (2012) identified the availability of treebanks for 30 languages. However, still majority of languages do not have treebanks. Thus, inducing treebanks for languages that have small or no training data is definitely a challenging task. Though fully unsupervised dependency parsing approaches (Mareček and Straka, 2013; Spitkovsky et al., 2013, Blunsom and



Figure 1. Schematic depiction of how the target parse trees are obtained

Cohn, 2010) are quite attractive for they don't require any hand annotated data for training, their quality is still lower than other class of approaches mainly known as *cross-lingual syntactic transfer* techniques. This is a useful alternative in transferring syntactic knowledge from one or more languages.

Hwa et al. (2005) used parallel corpus and word alignments to project English parse trees to Spanish and Chinese. They have also used small number of language specific transformation rules to reduce projection errors due to different annotation choices in the treebanks. Most of the earlier transfer based approaches (Ganchev et al., 2009; Kuhn, 2004) heavily rely on bitexts or some other target¹ language resources (such as POS taggers).

Transfer based techniques such as Zeman and Resnik (2008) and McDonald et al. (2011b) decouple this target language resource requirement by directly parsing target sentences via delexicalized source parser (trained with source POS tag sequence). Delexicalized parsing depends only on target POS tagger which uses the same POS tagset as the source language tagger. Augmenting delexicalized parsers by cross-lingual clusters (Täckström et al., 2012), bilingual lexicon (Durrett et al., 2012) and target adaptation techniques (Täckström et al., 2013) further improved the delexicalized parsers. Most of the recent works on transfer parsers sought to reduce POS annotation differences between source and target languages by mapping to a common coarse-grained tagset (Petrov et al., 2012). Addressing annotation differences at the structural level in transfer parsers is still an open problem, though there are some early attempts such as Smith and Eisner (2009) and more recently Zeman et al. (2012); McDonald et al. (2013) through treebank harmonization and by common annotation standards.

It has been well established from previous works that the availability of bitexts or target POS taggers (or both) is very crucial for transferring dependency parsers from one or multiple source languages. Imagine a situation where we don't have direct access to bitexts or a target POS tagger but only to a translation system from a resource-poor (RP) language to a resource-rich (RR) language. This presents an interesting scenario for the existing transfer based approaches. In this paper, we propose to combine bitexts obtained from machine translation and target POS obtained from

¹*target* refers to language(s) for which we would be interested in inducing dependency parser. *source* in other words refers to language(s) from which we will transfer the dependencies. It is also assumed that target languages are resource-poor whereas source languages are resource-rich.

unsupervised clusters to obtain transfer parsers. We use MT system to translate target language texts to resource-rich source language texts (for which parsers and taggers are available). Our overall approach is similar to Hwa et al. (2005) and McDonald et al. (2011b), but the main difference lies in the nature of bitexts we use for transferring the parsers.

Later in the results section, we show that this approach outperforms state-of-the-art unsupervised approaches even though we use only approximate bitexts for our transfers.

2. Dependency Transfer With Machine Translated Texts

The heart of our approach lies in how we obtain bitexts and target POS taggers (for resource-poor languages), which are crucial for transfer parsers. Unlabeled data is available in plenty even for resource-poor languages. We obtain both the resources from unlabeled target texts only. For word aligned bitexts, we first translate target texts into English via an MT system, in our case Google Translate API.² Our system is single source, i.e., all our experiments are carried out with English as a source language against a variety of target languages. For word alignments, we use alignment links provided by the MT system. If word alignments are not provided by the MT system, then any automatic word aligners can be used to obtain the word alignments. We parse translated English source texts using the parsing model trained on the English treebank from CoNLL shared task (Nivre et al., 2007). Our overall approach is depicted in Figure 1.

We obtain fully connected target language parse trees by projecting English parse trees onto target sentences via word alignment links. Before projection, we initialize the target tree by connecting all the target tokens to the default root node of the tree. The projection algorithm then starts from the source root and visits all the source nodes in a pre-order fashion while making adjustments to parents in the target tree.

In the case of 1-M alignments, we first determine the head of the chunk on the target side and connect the remaining members of that chunk to the chunk head. We make a simplistic assumption about chunk head: i.e., we consider the last member (its absolute position should also be higher than other members) to be the chunk head. Unlike Hwa et al. (2005), we do not add empty nodes on the target in the case of determining target parents for 1-M and unaligned source nodes. We do that so for simplifying the evaluation of target trees.

We use the projected target trees to train various parsing models. Previous works have mostly relied on using target POS taggers for training the parsers. McDonald et al. (2011b) demonstrated that POS information alone carries much of the syntactic information, thus making use of target POS taggers considerably improved the accuracy of the target parse trees. To make our work applicable in realistic scenarios, we

²Google Translate API – <https://developers.google.com/translate/>

#	ar	bg	ca	cs	da	de
Sentences	3.0K	13.2K	14.9K	25.6K	5.5K	38.0K
Tokens	116.8K	196.2K	443.3K	437.0K	100.2K	680.7K
Train/Test(%)	96/4	97/3	88/12	99/1	94/6	95/5
#	el	es	et	fi	hi	hu
Sentences	2.9K	16.0K	1.3K	4.3K	13.3K	6.4K
Tokens	70.2K	477.8K	09.5K	58.6K	294.5K	139.1K
Train/Test(%)	93/7	90/10	90/10	90/10	91/9	94/6
#	it	nl	pt	sl	sv	tr
Sentences	3.4K	13.7K	9.4K	1.9K	11.4K	5.9K
Tokens	76.3K	200.7K	212.5K	35.1K	197.1K	69.7K
Train/Test(%)	93/7	97/3	97/3	79/21	97/3	95/5

Table 1. Target language treebank texts that are translated to English using MT system

induce target POS information using unsupervised techniques. Unsupervised POS tagging is arguably less complex than inducing tree structures and previous works on unsupervised POS techniques (Blunsom and Cohn, 2011; Clark, 2000) have proven to be effective even in practical applications. In this work, we use unsupervised target POS tags instead of supervised or universal POS tags, but for the sake of comparison, we also provide results with supervised and universal POS tags. We experiment with various tagset size and show results for the tagset size that gives the best average accuracy on target languages. Our approach can be used within the transfer framework for languages that lack even POS taggers, thus making the approach suitable for languages that do not have any labeled target language resources.

3. Experiments

We use 18 treebanks (see Table 1) for most of our experiments. In certain experiments, we show results for a subset of those languages for comparison with other works. For training/testing, we use the same data split as described in Zeman et al. (2012). The target language treebanks we use mostly come from past CoNLL shared tasks (2006, 2007 and 2009). For Hindi, we have used the latest version (ICON 2012) of the treebank instead of the version mentioned in Zeman et al. (2012). All our results show only *unlabeled attachment score (UAS)* accuracies – similar to other works in the field.

3.1. Projection

The schema of the projection procedure is depicted in Figure 1. It consists of four steps:

Lang.	Baseline		UDP	Projection (+ reparsing)					sup
	left	right		unsup40	dir proj	univ	sup	gold	
ar	5.2	58.8	34.1	51.8	40.3	56.1	58.3	60.0	74.2
bg	17.9	38.8	56.7	46.4	41.4	55.3	53.6	56.3	81.5
ca	24.7	28.8	24.6	56.9	46.0	-	59.0	60.2	87.6
cs	24.1	28.9	55.0	45.4	51.4	54.7	58.4	62.3	74.9
da	13.2	47.9	42.0	41.6	36.9	41.6	42.1	43.1	79.8
de	24.2	18.9	43.5	46.4	43.0	47.8	48.7	50.1	84.2
el	32.0	18.5	30.9	49.2	52.0	59.2	65.4	65.2	78.5
es	24.7	29.0	36.3	56.9	47.0	-	58.2	59.0	88.1
et	34.1	17.4	63.8	54.8	58.0	-	58.3	66.0	72.9
fi	39.3	13.6	36.7	37.1	43.1	-	39.5	46.2	60.7
hi	24.4	27.3	15.6	24.9	24.6	-	28.0	28.3	75.1
hu	42.8	5.3	34.7	43.4	41.1	48.2	51.2	53.2	75.2
it	23.0	37.4	49.7	55.6	53.1	53.7	59.3	62.0	80.5
nl	27.9	24.7	27.6	60.9	53.2	-	59.3	61.4	76.7
pt	25.8	31.1	39.8	61.8	56.4	62.2	62.9	66.5	84.2
sl	24.4	26.6	45.9	44.4	50.4	45.7	53.1	56.6	76.7
sv	25.9	27.8	49.1	53.9	48.2	56.8	54.3	55.8	82.7
tr	65.1	2.0	42.3	46.0	51.4	-	57.2	57.0	75.5
avg	27.7	26.8	40.5	48.7	46.5	52.8	53.7	56.1	78.3

Table 2. UAS for baselines, supervised/unsupervised parsers and various projected parsing models. Resource requirements for projection under various settings: un-sup40 requires bitext, source parser and unsupervised target tagger; dir proj requires only bitext; univ requires bitext, source parser, and universal source/target taggers; sup (column 8) requires bitext, source parser, and target tagger; the resource requirements for gold is similar to sup but requires gold POS during testing. sup (column 10) is a regular supervised parser results. UDP shows results for unsupervised dependency parsing. Numbers in **bold** indicate the best accuracy for each row (excluding columns 9 and 10).

1. The target language corpus is translated to English using Google Translate API v1. The API provides also alignment links, which will be used for projection. The translated English sentences are then tokenized and the original alignment links are adjusted.³ For the translation task, we confined ourselves to translating only the treebank data, but much larger texts can be used in the future.

³For instance, when a punctuation is separated from a form on the English side, we link the separated punctuation to the corresponding punctuation on the treebank data.

2. English sentences are tagged by the Morce tagger (Spoustová et al., 2007) and parsed by the MST parser (McDonald et al., 2005). For parsing English, we used the parser model trained on the version of Penn treebank supplied during the CoNLL 2007 shared task (Nivre et al., 2007).
3. English dependency trees are projected to the target language sentences (only the training part of the treebank) using the alignment links.
4. Three target parser models are trained (using MST parser with 2nd order and non-projective setting) on the projected target corpus with different POS annotations (next subsection).

3.2. Training with different POS tags

To tag target test data, we train two supervised taggers and one unsupervised tagger on the training section of the target treebanks.

- **Supervised POS:** We train Stanford tagger (Toutanova and Manning, 2000) on the training section of the treebanks.
- **Universal POS:** We first convert the annotated training data to universal POS tags (Petrov et al., 2012) and train the tagger on it.
- **Unsupervised POS tagger:** We use unsupervised hidden Markov model (HMM) POS tagger by Blunsom and Cohn (2011). Not knowing which tagset size is suitable for the projected treebanks, we obtain POS tags for different arbitrary tagset size: 20, 40, 80 and 160. Besides the training and testing parts of the treebanks⁴, we used additional monolingual texts from W2C Wikipedia corpus (Majliš and Žabokrtský, 2012) to enlarge the size of the data to one million words for each language.

3.3. Direct transfer of delexicalized parser

We train delexicalized English parsers under two settings. In the first setting, we convert the POS tags of CoNLL 2007 English data into universal POS tags using the mapping provided by Petrov et al. (2012), strip all the word forms and train the parser. In the second setting, we first tag the English translations from Google Translate using the Morce tagger (Spoustová et al., 2007), convert them to universal POS tags and after stripping all the word forms, we train delexicalized parser on them. We obtain target universal POS tags using the universal POS tagger trained from the training part of the target treebanks. So, the main difference between McDonald et al. (2011b) and our approach is that they obtained target POS tags by POS projection from English, whereas we used POS information from target language treebanks and trained uni-

⁴The unsupervised POS tagger by Blunsom and Cohn (2011) does not produce a trained POS model. Given an unlabeled data and tagset size, the tagger produces unsupervised tags for the unlabeled data. Our unlabeled data included the training/testing part of the treebanks and some additional monolingual corpus.

versal POS taggers on them. At the moment, our experiments on unsupervised POS taggers deal with varying tagset size, and it would also be interesting in the future to make a comparison with different unsupervised approaches such as unsupervised POS projection (Das and Petrov, 2011).

3.4. Unsupervised parsing

To compare the projection results with a completely unsupervised approach, we used the software for dependency grammar induction by Mareček and Straka (2013).⁵ We run the experiments in the same manner as they described for all our testing languages.

4. Results

Our major results are presented in Table 2 and 3. The left and right baselines in Table 2 indicate that some languages have a strong preference for either left or right branching.

- *sup* presents UAS scores from a supervised parser trained on the training portion of the target treebanks.
- *UDP* presents UAS scores achieved in unsupervised dependency parsing (Mareček and Straka, 2013) on the test portion of the target treebank data.
- Projected results under different settings
 - *unsup40* presents UAS scores for parsing the test data with unsupervised POS tags (tagset size 40). We also experimented with different tagset size: 20, 80 and 160. We chose the tagset size that gave the best average accuracy.
 - *dir proj* shows UAS scores from directly projecting translated English test data onto target test sentences.
 - *univ*: test data is tagged by universal POS tagger before parsing.
 - *gold*: test data is tagged with gold POS tags before parsing.
 - *sup*: test data is tagged by supervised POS tagger before parsing.

For the sake of comparison, we reproduce delexicalized parser results on our data with two settings: (i) delexicalized parser trained on the POS tags of CoNLL 2007 (Nivre et al., 2007) English data and (ii) delexicalized parser trained on the POS tags of English translations obtained from the MT system. The results are shown in Table 3. For both settings, we obtain target POS tags in a supervised manner. We also provide delexicalized parser results from McDonald et al. (2011a). One intriguing aspect of these results is that delexicalized parsers obtained from the machine translated texts perform better than the delexicalized parser obtained from the CoNLL 2007 data (48.4 vs. 47.8). *McD 2011* has better overall results compared to our delexicalized parsers. We attribute this to difference in parser training parameters as well as the usage of

⁵<http://ufal.mff.cuni.cz/udp/>

Lang.	Delex CoNLL	Delex GT	McD 2011
ar	29.1	31.9	-
bg	52.8	51.4	-
cs	35.6	35.4	-
da	44.6	39.7	45.5
de	47.5	48.1	47.5
el	59.0	60.5	65.2
es	-	-	52.4
hu	45.0	46.2	-
it	52.9	57.2	56.3
nl	-	-	66.5
pt	65.4	63.5	67.7
sl	36.5	44.2	-
sv	57.7	54.5	59.7
avg	47.8	48.4	57.6

Table 3. Delexicalized direct transfer parsers comparison (unlabeled attachment score - UAS). Delex CoNLL - delexicalized parser trained on the CoNLL 2007 (Nivre et al., 2007) data. Delex GT - delexicalized parser trained on the POS tags of English translations obtained using Google Translate API. McD 2011 - results from McDonald et al. (2011a).

POS projection instead of a supervised POS tagger. When we make an overall comparison (Tables 2 & 3), the advantages of training the supervised taggers (both original and universal POS) are clearly visible. However, *unsup40* outperforms *UDP* by 8.2% absolute UAS score, and also gives slightly better results with respect to *delex CoNLL* and *delex GT*. Remember, both *delex CoNLL* and *delex GT* use supervised target taggers, that means, *unsup40* does not use any labeled target resources, but still performs better than other strategies. This suggests that, the syntactic projection which crucially relies on bitexts can still be beneficial even in the absence of high quality bitexts.

5. Discussion

From the results given in the previous section we can see that the average UAS of parsing models trained on the projected trees outperforms left and right baselines, while it is well below the supervised parser UAS, which was expected. In most cases it is above the unsupervised parser UAS, which was not guaranteed before performing our experiments. The performance difference between using hand-designed POS tags of the original treebanks and unsupervised tags (induced from large unannotated data) is surprisingly small.

Many techniques have been proposed in the past to address variety of resource poor scenarios, to which we would like to add one more scenario, the availability of an

MT system from a resource-poor language to a resource-rich language. To summarize, if one wants to parse a text in a language, the following procedure typically leads to the best performance:

1. If there is a treebank for the language, use it for training a supervised parser.
2. If there is no treebank, but a POS tagger exists, then develop a conversion to the universal tagset and use the delexicalized parser.
3. If the tagger is not available, but an MT system from the language to English exists, then use our approach.
4. If none of the previous steps is applicable, then use unsupervised parsing.

There are published treebanks for around 40 languages, and there are more than 70 languages supported by Google Translate. So one can expect around 30 languages for which either step 2 or 3 leads to state-of-the-art results. Given the quick growth of the number of languages covered by Google Translate, we believe that our approach will be viable for more and more languages.

In addition, our approach might be helpful for languages whose treebanks and taggers are available only under very restrictive licenses.

6. Conclusion

In this paper, we have considered the dependency induction task for a specific resource-poor scenario in which only MT system from a resource-poor language to a resource-rich language is available. We have used machine translated bitexts as a substitute for high quality bitexts, and used source language MT outputs and target language texts as a basis for projection-based transfer approach. The experimental results show that, in realistic scenarios, the projection-based transfer can be combined with unsupervised target POS tagger to achieve better parsing performance than unsupervised parser and similar performance as delexicalized transfer parsers. In the future, it would be interesting to compare transfer parsers induced from human translated bitexts and machine translated bitexts and ascertain whether MT outputs can be used as a substitute for tasks which require bitexts.

Acknowledgements

This research has been supported by the grant no. GPP406/14/06548P of the Grant Agency of the Czech Republic. This work has been using language resources developed and/or stored and/or distributed by the LINDAT/CLARIN project of the Ministry of Education of the Czech Republic (project LM2010013).

Bibliography

Blunsom, Phil and Trevor Cohn. Unsupervised induction of tree substitution grammars for dependency parsing. In *Proceedings of the 2010 Conference on Empirical Methods in Natural*

- Language Processing*, EMNLP '10, pages 1204–1213, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics.
- Blunsom, Phil and Trevor Cohn. A hierarchical Pitman-Yor process HMM for unsupervised part of speech induction. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 865–874, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.
- Clark, Alexander. Inducing syntactic categories by context distribution clustering. In *Proceedings of the 2Nd Workshop on Learning Language in Logic and the 4th Conference on Computational Natural Language Learning - Volume 7*, CoNLL '00, pages 91–94, Stroudsburg, PA, USA, 2000. Association for Computational Linguistics.
- Das, Dipanjan and Slav Petrov. Unsupervised part-of-speech tagging with bilingual graph-based projections. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*, HLT '11, pages 600–609, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics. ISBN 978-1-932432-87-9. URL <http://dl.acm.org/citation.cfm?id=2002472.2002549>.
- Durrett, Greg, Adam Pauls, and Dan Klein. Syntactic transfer using a bilingual lexicon. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 1–11, Jeju Island, Korea, July 2012. Association for Computational Linguistics. URL <http://www.aclweb.org/anthology/D12-1001>.
- Ganchev, Kuzman, Jennifer Gillenwater, and Ben Taskar. Dependency grammar induction via bitext projection constraints. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1 - Volume 1*, ACL '09, pages 369–377, Stroudsburg, PA, USA, 2009. Association for Computational Linguistics. ISBN 978-1-932432-45-9. URL <http://dl.acm.org/citation.cfm?id=1687878.1687931>.
- Hwa, Rebecca, Philip Resnik, Amy Weinberg, Clara Cabezas, and Okan Kolak. Bootstrapping parsers via syntactic projection across parallel texts. *Natural Language Engineering*, 11:311–325, 9 2005. ISSN 1469-8110. doi: 10.1017/S1351324905003840. URL http://journals.cambridge.org/article_S1351324905003840.
- Kuhn, Jonas. Experiments in parallel-text based grammar induction. In *Proceedings of the 42Nd Annual Meeting on Association for Computational Linguistics*, ACL '04, Stroudsburg, PA, USA, 2004. Association for Computational Linguistics. doi: 10.3115/1218955.1219015. URL <http://dx.doi.org/10.3115/1218955.1219015>.
- Majliš, Martin and Zdeněk Žabokrtský. Language richness of the web. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC 2012)*, Istanbul, Turkey, May 2012. European Language Resources Association (ELRA). ISBN 978-2-9517408-7-7.
- Mareček, David and Milan Straka. Stop-probability estimates computed on a large corpus improve Unsupervised Dependency Parsing. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 281–290, Sofia, Bulgaria, August 2013. Association for Computational Linguistics.
- McDonald, Ryan, Fernando Pereira, Kiril Ribarov, and Jan Hajič. Non-Projective Dependency Parsing using Spanning Tree Algorithms. In *Proceedings of Human Language Technology Con-*

- ference and Conference on Empirical Methods in Natural Language Processing (HTL/EMNLP), pages 523–530, Vancouver, BC, Canada, 2005.
- McDonald, Ryan, Slav Petrov, and Keith Hall. Multi-source transfer of delexicalized dependency parsers. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '11, pages 62–72, Stroudsburg, PA, USA, 2011a. Association for Computational Linguistics. ISBN 978-1-937284-11-4. URL <http://dl.acm.org/citation.cfm?id=2145432.2145440>.
- McDonald, Ryan, Slav Petrov, and Keith Hall. Multi-source transfer of delexicalized dependency parsers. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 62–72, Edinburgh, Scotland, UK., July 2011b. Association for Computational Linguistics. URL <http://www.aclweb.org/anthology/D11-1006>.
- McDonald, Ryan, Joakim Nivre, Yvonne Quirnbach-Brundage, Yoav Goldberg, Dipanjan Das, Kuzman Ganchev, Keith Hall, Slav Petrov, Hao Zhang, Oscar Täckström, Claudia Bedini, Núria Bertomeu Castelló, and Jungmee Lee. Universal dependency annotation for multilingual parsing. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 92–97, Sofia, Bulgaria, August 2013. Association for Computational Linguistics. URL <http://www.aclweb.org/anthology/P13-2017>.
- Nivre, Joakim, Johan Hall, Sandra Kübler, Ryan McDonald, Jens Nilsson, Sebastian Riedel, and Deniz Yuret. The CoNLL 2007 Shared Task on Dependency Parsing. In *Proceedings of the CoNLL Shared Task Session of EMNLP-CoNLL 2007*, pages 915–932, Prague, Czech Republic, June 2007. Association for Computational Linguistics.
- Petrov, Slav, Dipanjan Das, and Ryan McDonald. A universal part-of-speech tagset. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey, may 2012. European Language Resources Association (ELRA). ISBN 978-2-9517408-7-7.
- Smith, David A. and Jason Eisner. Parser adaptation and projection with quasi-synchronous grammar features. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 2 - Volume 2*, EMNLP '09, pages 822–831, Stroudsburg, PA, USA, 2009. Association for Computational Linguistics. ISBN 978-1-932432-62-6. URL <http://dl.acm.org/citation.cfm?id=1699571.1699620>.
- Spitkovsky, Valentin I., Hiyan Alshawi, and Daniel Jurafsky. Breaking out of local optima with count transforms and model recombination: A study in grammar induction. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1983–1995, Seattle, Washington, USA, October 2013. Association for Computational Linguistics.
- Spoustová, Drahomíra, Jan Hajič, Jan Votrubec, Pavel Krbeč, and Pavel Květoň. The Best of Two Worlds: Cooperation of Statistical and Rule-Based Taggers for Czech. In *ACL '07: Proceedings of the Workshop on Balto-Slavonic Natural Language Processing*, pages 67–74, Morristown, NJ, USA, 2007. Association for Computational Linguistics.
- Täckström, Oscar, Ryan McDonald, and Jakob Uszkoreit. Cross-lingual word clusters for direct transfer of linguistic structure. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 477–487, Montréal, Canada, June 2012. Association for Computational Linguistics. URL <http://www.aclweb.org/anthology/N12-1052>.

- Täckström, Oscar, Ryan McDonald, and Joakim Nivre. Target language adaptation of discriminative transfer parsers. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1061–1071, Atlanta, Georgia, June 2013. Association for Computational Linguistics. URL <http://www.aclweb.org/anthology/N13-1126>.
- Toutanova, Kristina and Christopher D. Manning. Enriching the Knowledge Sources Used in a Maximum Entropy Part-of-Speech Tagger. In *Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora (EMNLP/VLC-2000)*, pages 63–70, Hong Kong, October 2000.
- Zeman, Daniel and Philip Resnik. Cross-language parser adaptation between related languages. In *IJCNLP 2008 Workshop on NLP for Less Privileged Languages*, pages 35–42, Hyderabad, India, 2008. Asian Federation of Natural Language Processing, International Institute of Information Technology.
- Zeman, Daniel, David Mareček, Martin Popel, Loganathan Ramasamy, Jan Štěpánek, Zdeněk Žabokrtský, and Jan Hajič. HamleDT: To Parse or Not to Parse? In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey, may 2012. European Language Resources Association (ELRA). ISBN 978-2-9517408-7-7.

Address for correspondence:

Loganathan Ramasamy
ramasamy@ufal.mff.cuni.cz
Institute of Formal and Applied Linguistics
Faculty of Mathematics and Physics,
Charles University in Prague
Malostranské náměstí 25
118 00 Praha 1, Czech Republic