IJCLA VOL. 4, NO. 2, JUL-DEC 2013, PP. 63–78
RECEIVED 07/12/12 ACCEPTED 04/03/13 FINAL 14/06/13

Syntax-Based Reordering in Phrase-Based English–Hungarian Statistical Machine Translation

LÁSZLÓ J. LAKI, ATTILA NOVÁK, AND BORBÁLA SIKLÓSI

Pázmány Péter Catholic University, Hungary

ABSTRACT

Phrase-based statistical machine translation systems can generate quite high quality translations in the case of language pairs with similar structure and word order. However if the languages are more distant from a grammatical point of view, the quality of translations is much behind the expectations, since the baseline translation system cannot cope with long distance reordering of words and morphological synchronization. In our paper, we present a method that tries to overcome these problems in the case of English-to-Hungarian translation. We describe how we defined some reordering rules on the English sentences in order to approximate the syntax of a hypothesized Hungarian translation prior to the actual process of translation. Due to the limited training corpus and data sparseness, and problems caused by the agglutinating characteristics of Hungarian, we applied a morpheme-based translation system. We show that although automatic evaluation cannot reliably reflect the improvement, human evaluation of the systems shows that readability and grammatical correctness of the translations were improved.

KEYWORDS: *Statistical machine translation, morphology, reordering.*

1 Introduction

Currently, the most widespread method for machine translation is to train statistical machine translation (SMT) systems without much explicit specific knowledge of the actual language pair, instead of creating sophisticated language dependent rule-based systems. For syntactically similar and morphologically simple language pairs, methods of phrase-based SMT perform quite well. However, in the case of more distant languages (such as English and Hungarian), there are less promising results. Studies have also shown that increasing the size of the training corpus still does not provide significant increase in the quality of translation [1]. Due to free word order and rich variability of word forms in Hungarian, even big corpora represent grammatical phenomena very sparsely. It implies that SMT systems applied for the English-Hungarian language pair are compromised by data sparseness problems. Our goal was to create a hybrid translation system that, while exploiting the advantages of statistical methods, tries to decrease the above mentioned difficulties.

2 Machine Translation from English to Hungarian

2.1 *Characteristics of Hungarian*

Hungarian is an agglutinating and compounding language with a practically unlimited number of different word forms. This, combined with free word order of main grammatical constituents and systematically different word order in NP's and PP's, results in a poor performance of simple phrase-based English to Hungarian translation systems. The great number of mismatches in word order and word count, the frequent need of long distance word movement and the low representing power of unanalyzed corpora for an agglutinating language like Hungarian, are all factors that make English-to-Hungarian machine translation difficult. The following comparison of language-specific corpus characteristics illustrates the latter problem. While the number of different word tokens in a 10 million word English corpus is generally below 100,000, it is well above 800,000 in the case of a Hungarian corpus of the same size. However, the 1:8 ratio does not correspond to the ratio of the number of possible word forms between the two languages: while there are no more than about 4–5 different inflected forms for an English word, there are about a 1000 for a Hungarian word, which indicates that a corpus of the same size is much less representative for Hungarian than it is for English [2].

2.2 *SMT and Word Order Differences*

If we evaluate the performance of phrase-based machine translation systems between English and various other European languages, we find that these systems perform much worse for languages which differ significantly from English in terms of word order. This indicates that the generic reordering algorithms implemented in phrase-based SMT systems cannot handle long distance word order mismatches effectively. In this paper, we describe a system that uses language-pair-dependent movement rules to handle word order differences, which were implemented as pre- and postprocessing steps around the core of a phrase-based SMT system.

3 Applying Reordering Rules

In order to reduce the complexity of the translation task, our system applies reordering rules prior to training the statistical models. The transformations applied to the source sentences make them more similar to the structure of the corresponding target sentences. In order to perform the required word movements, the rules rely on constituent structure and typed dependency relations in the English source sentences. To process raw sentences, the Stanford parser [3] is used as described in Section 4.2. This enrichment of the grammatical description of the sentence provides enough information for defining rules that can transform the source sentence structures to others that correspond to those occurring in the corresponding hypothesized Hungarian sentence. Since the SMT system is based on data extracted from aligned phrases in the parallel training corpus, the quality of the alignment phase is of crucial importance [4]. Thus one of our goals for the reordering rules was to create a better source for the alignment module. We expected that training the system on such a set of transformed English–Hungarian parallel sentences, more representative statistics can be built than in the case of the baseline model.

Approximating the structure of the source and target languages to each other can on the one hand decrease word alignment errors that result from differences in the organization of morphemes to surface word forms. On the other hand, results published on the research of other language pairs (such as English–German or English–Turkish) have shown that by applying reordering rules to the source sentence, the number of words left without translation during decoding can be decreased [5–7].

We created rules only for those word order differences which are systematically present between the two grammars: e.g. prepositions vs. case endings/ postpositions, possessive determiners vs. possessive suffixes etc. We did not intend to handle free word order variations of Hungarian, where the same meaning can be expressed with several different orderings, since in Hungarian, the actual word order in a sentence is not only determined by syntactic, but also by pragmatic factors.

Reordering rules rely both on phrase structure and dependency relations in the English input sentences. Once having these relations extracted, transformations are carried out along the relevant relations. A simple example is a phrase like *in my house*, which is transformed to the form *house_my_in* corresponding to the single word *házamban* in Hungarian. The morphological segmentation of this word is *ház[N] + am[PxS1] + ban[Ine]*, with the Hungarian morphemes corresponding to 'house[Noun] + my[Possessor:1Sg] + in[Case:Inessive]'.

Defining and applying the rules for such short phrases is not particularly difficult. However, related words in longer sentences can be much further separated from each other and they may be involved in more than one relation which often results in an interaction of word order constraints. In a similar manner, some rules insert morphological elements corresponding to those that are present in the Hungarian sentence, but not explicitly expressed in English, such as the accusative case suffix. These morphemes are important for the accuracy and fluency of the translation.

Our reordering rules fall into three categories:

3.1 *Rules Affecting Word Order and Morpheme Division/Unification*

Once having the dependency relations extracted from the sentence, these rules are responsible for moving each word to its reordered position and at the same time performing unification of English function words in order to make English sentence structures more similar to Hungarian. Besides typed dependencies, these transformations also rely on the constituent parsing of the sentences. Some examples of these rules are the ones transforming passives, positioning auxiliaries, prepositions and transforming possessive phrases. The order of performing these rules is important, especially when longer sequences are affected. In the following sentence in Table 1, we perform two transformations.

While heavy participle phrases in English generally follow the NP they modify, this is never the case in Hungarian where modifiers containing participles strictly precede the noun just like ordinary adjectival modifiers. Moreover, any arguments or adjuncts of the participle must precede it (unlike in the corresponding English structure where they fol-

Table 1. An example of reordering and word form restructuring

Original sentence:	The/DT sons/NNS of/IN the/DT many/JJ merchants/NNS living/VBG in/IN the/DT city/NN ./.
Reordered sentence:	the/DT city/NN_in/IN living/VBG many/JJ merchants/NNS sons/NNS_of/IN ./.

low it). This is an example of a systematic word order difference between the two languages. Correspondingly, the prepositional phrase *living in the city* is transformed along the relations PARTMOD(merchant, living)[1], PREP(living, in)[1] and POBJ(in, city)[1]. First the preposition is attached to the child of the POBJ relation (the head of the dependent NP), then this unified word is moved before the participle and the whole participial modifier phrase before the head noun. Thus the resulting word forms and their order is corresponding to the Hungarian translation: *a város_ban élő* ('the city_in living'). The other phrase (*the sons of the merchants*) is transformed similarly to the resulting *merchants sons_of* order, which corresponds to the order of morphemes in the Hungarian translation of the phrase: *kereskedők fi_ai*.

Table 2. Examples of reordering and morpheme insertion

Original sentence:	That/DT is/VBZ the/DT account/NN at/IN the/DT largest/JJS bank/NN in/IN Bern/NNP ./. "/"
Reordered sentence:	That/DT is/VBZ the/DT Bern/NNP_in/IN xxx/xxx largest/JJS bank/NN_at/IN xxx/xxx account/NN ./. "/"
Original sentence:	Buckets/NNS containing/VBG milk/NN must/MD be/VB covered/VBN
Reordered sentence:	Milk/NN_acc/ACC containing/VBG Buckets/NNS must/MD covered/VBN_MD_they/P3

Although in most cases the English sentence has more words than the corresponding Hungarian sentence since English grammatical words usually correspond to bound morphemes in Hungarian, there are situations where some words are missing and have to be inserted in order to get the Hungarian sentence structure. One construction where this hap-

[1] PARTMOD=participal modifier, PREP=prepositional modifier, POBJ=object of preposition. The full list of dependency relations can be found in `http://nlp.stanford.edu/software/dependencies_manual.pdf`

pens is the case of postnominal modifiers not containing a participle (e.g. *the largest bank in Bern*) which are transformed into prenominal modifiers in Hungarian that do contain one. Since the participle to be inserted depends on the context, we insert only an abstract character string representing the participle, the actual realization of which is determined by the SMT system during translation based on similar transformed examples in the training corpus. One such example is the sentence in Table 2 containing the string *xxx/xxx* that is translated to Hungarian as *levő* 'being'. The other example in Table 2 shows insertion of the accusative ending in addition to movement and reordering of the participle modifier that contains it.

3.2 *Rules Affecting Only Morphological Structure, Not Word Order*

English sentences contain several types of implicit structural information that are represented as explicit suffixes in Hungarian. E.g., while objects are identified by their position in English, the same dependency relation is explicitly marked by the accusative case suffix *-t* in Hungarian. Since dependency parsing identifies the object relation in English, it can be transferred as an additional morpheme to the reordered sentence. For example, the original sentence *She/PRP shot/VBD herself/PRP ./.* is transformed into the sentence *shoot/VB_Past_she/PRP herself/PRP_acc/ACC ./.*

There are cases when English represents some morphemes as separate words, while these are only suffixes in Hungarian. To avoid the aligner connecting these morphemes to some other words on the Hungarian side, these words are attached to their corresponding position. For example, if the sentence contains a possessive determiner and the object of the possession, then these are connected. Thus the phrase *"my/PRP$ own/JJ mother/NN"* is transformed to the form *"own/JJ mother/NN_my/PRP$"*, which corresponds to the Hungarian phrase *"saját anyá_m"*.

3.3 *Minor Adjustment Rules*

Rules in this group make some adjustments necessary to make the results of previous transformations well-formed. E.g., the transformations produce two consecutive definite articles if the possessor and the possessed are both definite in a possessive construction or if a definite noun has a modifier that contains another definite dependent. E.g., the phrase *the house standing in the forest* would be transformed to *the ∗the forest_in standing house*. Only one definite article is present in Hungarian in constructions of this kind: the extra articles are deleted by a minor adjustment rule. We also classified some simple movement rules as minor adjustment rules, as these do not interact with others in a complicated manner. One example is the attachment of the genitive *'s* (see Table 3) or the transposition of currency symbols after the sum they belong to.

Table 3. An example of possessive reordering

Original sentence:	John's cat
Dependency relations:	poss(cat, John)
	possessive(John, 's)
Reordered sentence:	John/NNP cat/NN_'s/PoS
Hungarian sentence:	John macská_ja

4 TOOLS AND RESOURCES

4.1 *Corpora*

The available English–Hungarian corpora are usually not suitable for training a general purpose SMT system, since they contain the terminology of a certain specific domain. That is why we used the largest and thematically most general corpus, called Hunglish[8], created by BME MOKK[1] and the Research Institute for Linguistics of the Hungarian Academy of Sciences. This corpus contains parallel texts from the following domains: literature and magazines, law and movie subtitles. There is a great degree of variation in the quality of different parts of the corpus. We automatically eliminated sentence pairs from the corpus that caused technical problems, but overall translation quality was not checked. Finally, the number of sentence pairs we used for training the system was 1,202,205 parallel sentences, which is 12,396,277 words on the English side and 12,316,157 on the Hungarian side.

[1] MOKK Centre for Media Research and Education at the Department of Sociology and Communication, Budapest University of Technology and Economics

4.2 *Constituent and Dependency Parsing*

For the first step of preprocessing, the English sentences were parsed, and dependency relations were extracted. To perform a morpheme-based translation, a part-of-speech tagger was also necessary for Hungarian.

To annotate the Hungarian side of the corpus, we used the PurePos automated morphological annotation system [9]. We parsed the Hungarian side of the corpus using this tool decomposing morphologically complex words in order to have a denser representation of the corpus than the unanalyzed version containing only word forms.

Since the original surface word forms can be reconstructed from the lemma and the morphological tags, the statistics for word alignment and translation can be improved by considering only the lemmas, as they occur more frequently in the corpus than any of the inflected forms. By applying this methodology, the translations generated by the SMT system also contain sequences of lemmas and morphosyntactic tags, thus in order to generate the final form of the translated sentence, the surface form of the words have to be regenerated. We did this by applying the word form generator module of Humor morphological analyzer to the output of the decoder [10, 11].

For parsing English, we used the state-of-the-art Stanford parser [3]. Since the quality of syntactic analysis is a crucial factor for reordering, we used the slower, but better lexicalized version of the parser. This results in a bit more accurate parses than the baseline unlexicalized parser, but it still very frequently generates parses which are often agrammatical with agreement errors and odd PoS sequences like the ones in Table 4.

Table 4. Examples of low level errors affecting reordering

POS-tagged sentence	-/: 100/CD million/CD **sound/NN** good/JJ to/TO me/PRP ./.
Reordered sentence	-/: me/PRP_to/TO xxx/xxx 100/CD million/CD sound/NN good/JJ ./.
POS-tagged sentence	For/IN airline/NN personnel/NNS ,/, we/PRP **cash/NN** personal/JJ **checks/VBZ** up/RP to/TO $/$ 100/CD ./.
Reordered sentence	airline/NN personnel/NNS_For/IN ,/, cash/NN personal/JJ up/RP_checks/VBZ_we/PRP 100/CD_$/$_to/TO ./.

Due to the sequentially pipelined construction of the system, errors are propagated from the very first PoS tagging step through the whole transformation and translation process. Each component of the pipeline assumes correct input, which they do not try to correct. Rather, they try their best to accommodate to whatever input they receive, often resulting in an absurd output. Word and phrase misplacements due to these wrong analyses yield a critical source of errors in the whole system, since the reordering rules are executed on erroneous input. It means that if we reorder an erroneously parsed sentence, then it is likely that the reorderings worsen the final result of the translation rather than improving it. The first such source of error is wrong PoS tag assignment. The most typical error is confusing nouns, adjectives and verbs, which is usually of fatal consequences regarding the translation of the sentence. Since both constituency and dependency parsing are based on such misleading information, the error propagates resulting in mistakes such as the ones displayed in Table 4.

5 The Moses Toolkit

In our present work, we used the phrase-based Moses SMT toolkit [12] to perform our translation experiments. Moses is the most widely used SMT tool. It is a practical solution for the tasks of both training and decoding. It depends on several external tools for the creation of the language models and the evaluation of the system.

The Moses system is suitable for implementing a so-called factored translation system. Instead of relying on just the surface form of the words, further annotations, such as morphological analysis, can be used in the process of a factored translation. Translation factors might be the surface form of each word, its lemma, its main PoS tag, its morphosyntactic features. During factored translation, there is an opportunity to use multiple translation models, generation models or contextual language models. Since the system has the possibility to use any combination of these, in theory, it is able to generate better translations using sparse linguistic data than a word-based baseline system. This feature is vital in cases where some abstraction is necessary, because some words in the sentence to be translated or generated are missing form the training set.

We investigated both factored and morpheme-based translation as possibilities to cope with data sparseness problems when translating from English to Hungarian. However, we found that traditional factored training and decoding is not suitable to handle the massive data sparseness issues encountered when translating to aggutinating languages like Hungarian or Finnish (see e.g. [13] for similar conclusions for the applicability of factored models to translation to Finnish). Nevertheless, factored models may be applicable to the solution of certain problems and are subject of our further investigation. The baseline system that we used for comparison is trained on the raw corpus without any preprocessing.

5.1 *Morpheme-based Translation*

In the morpheme-based implementation, morphological analysis, parsing and the reordering rules were applied to the corpus before training and translation, but at the end, no generation of word forms were performed within the Moses framework: the output of the decoder is a sequence of morphemes. We performed an automatic evaluation of this morpheme-based translation output using the BLEU metric. In contrast to the traditional surface-word-form-based BLEU score (w-BLEU), this score, which we term mm-BLEU, is based on counts of identical abstract morpheme sequences in the generated and the reference translations instead of identical word sequences. Note that this also differs from m-BLEU as used e.g. in [13], which is BLEU applied to (pseudo-)morphs generated by an unsupervised segmenter. mm-BLEU represents the ability of the system to generate the correct morphemes in the translations. After having these morphemes translated, a morphological generator was applied to the output of the Moses decoder in order to acquire the final word forms. As shown in Table 5, this resulted in lower w-BLEU scores than that of the baseline system. Nevertheless, manual investigation of the translation outputs revealed that the morpheme-based system is better at capturing grammatical relations in the original text and rendering them in the translation by generating the appropriate inflected forms. Although it is not reflected by the w-BLEU scores, it generates better translations from the perspective of human readability than the baseline system.

6 RESULTS

Since human evaluation is slow and expensive, machine translation systems are usually evaluated by automated metrics. However, it has been shown that system rankings based on single-reference BLEU scores often do not correspond to how humans evaluate the translations, for this reason, automatic evaluation has for a long time not been used to officially rank systems at Workshops on Statistical Machine Translation

(WMT) [14]. In our work, we present results of automated evaluation using a single reference BLEU metrics, but we also investigated translations generated by each system using human evaluation applying the ranking scheme used at WMT workshops to officially rank systems.

Our experimental setting for automated evaluation consisted of three separate test sets of 1000 sentences each, which were separated from our corpus prior to training the system. Besides these, evaluation was performed on a test set of a different domain (news) that is not represented in the training set at all.

Table 5 contains the traditional word-based w-BLEU scores of the baseline, the morpheme-based mm-BLEU scores of the morpheme-based system with rule-based reordering and w-BLEU scores of the latter system with the target language surface word forms generated. The w-BLEU scores are lower compared to the baseline for all the test sets. However, as mentioned above, the decrease in these values does not necessarily correspond to worse translations.

It is also worth mentioning that morpheme-based mm-BLEU scores for the out of domain newswire test corpora is as high as for the in domain test sets, while the w-BLEU scores are significantly lower for the news test sets.

Table 5. BLEU scores of the word-based baseline and the reordered morpheme-based system

Name	Baseline	Reordered morph.-based	
	w-BLEU	mm-BLEU	w-BLEU
test1	15.82%	64.14%	12.61%
test2	14.60%	57.39%	13.95%
test3	15.04%	57.84%	12.98%
news2008	6.45%	59.73%	6.99%
news2009	7.36%	60.56%	7.26%

During the evaluation process, translations are compared to a single reference sentence. Thus if the machine translation result contains an absolutely wrong word or word form, the evaluation will be just as bad as if it contained a synonym of the correct word, or just a slightly different inflected form of it. The measurements clearly reflect however that translating a test set of different style and domain than the training set, results in much lower BLEU score.

6.1 *Human Evaluation*

We randomly selected a set of 50 sentences from test set 1 that underwent human evaluation as well. Four annotators evaluated translations generated by each of the three systems plus the reference translation in the corpus with regard to translation quality (considering both adequacy and fluency in a single quality ranking). The order of translations was randomized for each sentence. The systems were ranked based on a score that was defined as the number of times a system was found not worse than the other in pairwise comparisons divided by the number of pairwise comparisons. The aggregate results of human evaluation are listed in Table 6.

Table 6. Human evaluation including the reference translations

Name	Baseline	Morph.-based	Reference
test1	34.08%	**52.49**%	83.08%

The ranking produced by each annotator was identical. The rather low score (83.08%) for the reference translations indicates that there are quite serious quality problems with the corpus (mostly due to sentence alignment problems but also due to sloppy translations). The results also clearly indicate that the w-BLEU scores cited in the previous section clearly do not correspond to Human ranking. The morpheme-based reordered model having a lower BLEU score performed better than the baseline system.

6.2 *Error Analysis*

Besides the shortcomings of the evaluation metrics and the corpus itself, there are several real errors emerging during the translation process that can be compensated for in some future work.

1. Errors in parsing of the source-side English sentence can also cause problems in determination of the dependency relations, which will result in erroneous application of the reordering rules. In such cases words that were originally at their correct position will land at the wrong place.

2. Problems of the English PoS sequence: if a word has the wrong tag in the sentence that is to be translated, but it always occurred correctly tagged in the training set, then the system is not able to translate it, even if the word itself is not an unknown word. Likewise, if the translation model contains the same word with several possible PoS tags depending on the context, then if the word in the actual sentence gets the contextually wrong tag, its translation will be wrong (see e.g. *whisper* tagged as a verb (following a determiner!) and thus translated as a verb in Table 7). Tagging errors in the training corpus may result in wrong translation even if the actual parse is correct. Moreover, an incorrect PoS tag usually results in an erroneous syntactic analysis and wrong reordering.

Table 7. The effect of parsing errors

Original sentence:	For 50 years, barely a whisper.
Reordered sentence:	50/[CD] year/[NN] [PL] For/[IN] ,/[,] barely/[RB] a/[DT] **whisper/[VB]** ./[.]
Translated sequence:	50/[NUM_DIGIT] év/[N] [PL] [TER] ,/[PUNCT] alig/[ADV] egy/[DET] **suttog/[V] [S3]** ./[PUNCT]
Morpheme-based:	50 évekig, alig egy **suttog.**
Back-translation:	For 50 year, hardly a **he whispers.**
Baseline:	50 éve, alig egy suttogás.
Back-translation:	50 years ago, hardly a whisper.
Reference:	50 évig a sóhajtásukat sem hallottuk.
Back-translation:	For 50 years, we haven't heard a whisper from them.

3. The quality of the training and test sets has an immediate effect on the measured quality of the translation. The problem is not only that the translation model contains wrong translations learnt from the corpus, but the evaluation metrics compares the results to wrong reference translations. Although this affects translations generated by both the baseline and the morpheme-based system, this might play a role in BLEU score differences not corresponding to how humans rank the translations.
4. Since the smallest units of the translation are morphemes, some of them might be moved to a wrong position. It is often the case in longer sentences that instances of the same functional morpheme belong to more than one different word in the sentence. This causes indeterminacies in the alignment process (because the models imple-

mented in the Giza++ word aligner cannot be forced to assume locally monotonous alignment at the places where we in fact know that alignment should be monotonous) and this usually results in erroneous phrases being extracted from the training corpus. For example if there are two nouns in a sentence, one of them is plural, then the [PL] tag corresponding to this feature might land at another noun.

7 Conclusion

In this paper, we described a hybrid phrase-based translation system from English to Hungarian that is an extension of the baseline statistical methods by applying syntax and morphology-based preprocessing steps on the training corpus and morphological postprocessing during translation. The goal was to transform the source-side English sentences to a syntactic structure that is more similar to that of the target-side Hungarian sentences. We concentrated on syntactic structures that have systematically differing realizations in the two languages. We found that readability and accuracy of the translation are improved by the process of reordering the source sentences prior to translation, especially in the cases when the somewhat fragile PoS tagger–parser chain does not lead to wrongly reordered sentences, which has a deteriorating effect on translation quality. Although automatic evaluation assigned the morpheme-based system a significantly and consistently lower score than the baseline system, human evaluation found our systems better than the baseline. We found that several linguistic phenomena can be translated with a much better accuracy than using a traditional SMT system. We also described some problems that are to be solved in the future with the expectation of having an even stronger effect on translation quality.

Acknowledgement This work was partially supported by TÁMOP – 4.2.1.B – 11/2/KMR-2011-0002.

References

1. Lü, Y., Huang, J., Liu, Q.: Improving Statistical Machine Translation performance by training data selection and optimization (2007)
2. Oravecz, C., Dienes, P.: Efficient stochastic part-of-speech tagging for Hungarian. In: LREC, European Language Resources Association (2002)
3. Marneffe, M.C.D., Maccartney, B., Manning, C.D.: Generating typed dependency parses from phrase structure parses. In: In LREC 2006. (2006)

4. Och, F.J., Tillmann, C., Ney, H., Informatik, L.F.: Improved alignment models for Statistical Machine Translation. In: University of Maryland, College Park, MD. (1999) 20–28
5. Collins, M., Koehn, P., Kučerová, I.: Clause restructuring for statistical machine translation. In: Proceedings of the 43rd Annual Meeting of Association for Computational Linguistics. ACL '05, Stroudsburg, PA, USA, Association for Computational Linguistics (2005) 531–540
6. Gojun, A., Fraser, A.: Determining the placement of German verbs in English-to-German SMT. In Daelemans, W., Lapata, M., Màrquez, L., eds.: EACL, The Association for Computer Linguistics (2012) 726–735
7. Yeniterzi, R., Oflazer, K.: Syntax-to-morphology mapping in factored phrase-based statistical machine translation from English to Turkish. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. ACL '10, Stroudsburg, PA, USA, Association for Computational Linguistics (2010) 454–464
8. Halácsy, P., Kornai, A., Németh, L., Sass, B., Varga, D., Váradi, T., Vonyó, A.: A Hunglish korpusz és szótár [Hunglish corpus and dictionary]. In: III. Magyar Számítógépes Nyelvészeti Konferencia, Szeged, SZTE (2005) 134–142
9. Orosz, G., Novák, A.: PurePos – an open source morphological disambiguator. In: Proceedings of the 9th International Workshop on Natural Language Processing and Cognitive Science., Wroclaw, Poland (2012)
10. Prószéky, G., Kis, B.: A unification-based approach to morpho-syntactic parsing of agglutinative and other (highly) inflectional languages. In: Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics. ACL '99, Stroudsburg, PA, USA, Association for Computational Linguistics (1999) 261–268
11. Novák, A.: What is good Humor like? In: I. Magyar Számítógépes Nyelvészeti Konferencia, Szeged, SZTE (2003) 138–144
12. Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., Cowan, B., Shen, W., Moran, C., Zens, R., Dyer, C., Bojar, O., Constantin, A., Herbst, E.: Moses: Open Source Toolkit for Statistical Machine Translation. In: Proceedings of the ACL 2007 Demo and Poster Sessions, Prague, Association for Computational Linguistics (2007) 177–180
13. Clifton, A., Sarkar, A.: Combining morpheme-based machine translation with post-processing morpheme prediction. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1. HLT '11, Stroudsburg, PA, USA, Association for Computational Linguistics (2011) 32–42
14. Callison-Burch, C., Fordyce, C., Koehn, P., Monz, C., Schroeder, J.: (meta-) evaluation of machine translation. In: Proceedings of the Second Workshop on Statistical Machine Translation, Prague, Czech Republic, Association for Computational Linguistics (June 2007) 136–158

László J. Laki
MTA-PPKE Language Technology Research Group
and Faculty of Information Technology,
Pázmány Péter Catholic University,
50/a Práter street, 1083 Budapest, Hungary
E-mail: <laki.laszlo@itk.ppke.hu>

Attila Novák
MTA-PPKE Language Technology Research Group
and Faculty of Information Technology,
Pázmány Péter Catholic University,
50/a Práter street, 1083 Budapest, Hungary
E-mail: <novak.attila@itk.ppke.hu>

Borbála Siklósi
Faculty of Information Technology,
Pázmány Péter Catholic University,
50/a Práter street, 1083 Budapest, Hungary
E-mail: <siklosi.borbala@itk.ppke.hu>