



# Machine Translation Optimization using Hybrid Architectures

Neeha Ashraf

Research Scholar computer sciences  
University of Kashmir  
Jammu and Kashmir, India

Manzoor Ahmad

Research Supervisor computer sciences  
University of Kashmir  
Jammu and Kashmir, India

## ABSTRACT

Humans learn their culture through language. Its inquisitive nature of human and passion to travel across the world, warrants different cultures interact with each other, the means to achieve this is through human language, often interacting cultures communicate through different languages. To add to this, off-late, the increased use of social networking and the flood of information in foreign languages through web, the use of machine translation technology became inevitably significant. Researchers have predominantly employed two approaches in machine translations, based either on rule-based or statistical approaches. Both approaches have their strengths and weaknesses. Researchers working on both Rule-based and Statistical approaches have shown a keen interest in the area of hybrid machine translation. These systems try to profit from other respective approaches, combining data-driven and knowledge-driven elements. The interest is to investigate different hybrid arrangements required for optimizing the machine translations which will contribute to an overall increase in MT quality.

## Keywords

Machine translation; natural language processing; hybrid; architecture; optimization; RMT; SMT; hytra; SMT Extensions; RMT Extensions ; Multi-engine hybridization; Parallel Coupling; Serial Coupling

## 1. INTRODUCTION

Humans learn their culture through language. Its inquisitive nature of human and passion to travel across the world, warrants different cultures interact with each other, the means to achieve this is through human language, often interacting cultures communicate through different languages. To add to this, off-late, the increased use of social networking and the flood of information in foreign languages through web, the use of machine translation technology [1] became inevitably significant. Researchers have predominantly employed two approaches in machine translations, based either on rule-based[2] or statistical approaches[3]. Both approaches have their strengths and weaknesses. Researcher working on both Rule-based and Statistical approaches have shown a keen interest in the area of hybrid machine translation[4]. These systems try to profit from other respective approaches, combining data-driven and knowledge-driven elements. The interest is to investigate various hybrid translation arrangements of machine translations, which will contribute to an overall increase in MT quality.

In recent years, the increased use of social networking and the flood of information in foreign languages through web, the use of machine translation technology became inevitably significant. Research has been carried out in Machine

Translation especially in the area of data-driven (SMT) as opposed to knowledge-driven (RMT) MT. Recent evaluations (Callison-Burch et al, 2009)[5] resulted that both types of machine translation have comparable translation quality, but the level of output acceptance is about 50%. The success of MT is limited by its limited acceptability by human readers. Since the current state of art in MT is performing well but for specific scenarios and languages. The current approaches in machine translation are based either on rule-based or statistical approaches. Both approaches have their strengths and weaknesses.

Rule-based systems contain a wealth of linguistic knowledge about the language involved and can deal with long distance dependencies[6]. They translate more accurately by “trying to represent each and every bit of input. RMT system lacks in lexical selection in transfer and robustness in case of analysis failures sentences. It requires more expensive resources if the linguistic analysis goes deeper. Lacking monolingual and multi-lingual resources leads to resource poverty.

SMT system relies on the presence of parallel corpora and does not involve any linguistic analysis in deeper sense. These systems are more robust and always produce output. These systems read more fluent due to Language Models and perform better in lexical selection, as they have the ability to grasp implicit knowledge contained in co-occurrence statistics. These system face problems to cope with phenomenon which requires linguistic knowledge like morphology, syntactic functions and word order. They have problems with long distance reordering and also lose adequacy due to missing or spurious translation. The size of monolingual and bilingual corpora for the language pair in question defines the degree of resource poverty for such systems.

During the last few years, researchers working on both Rule-based and Statistical approaches have shown a keen interest in the area of hybrid machine translation. These systems try to profit from other respective approaches, combining data-driven and knowledge-driven elements. The purpose of this paper is to discuss different architectures of hybrid system proposed recently to optimize the existing machine translation techniques. The interest is to study how the hybrid systems can improve the quality of machine translation in terms of fluency and adequacy, how they can overcome the restrictions of respective other approaches, how these systems can overcome the resource poverty of a source or target language like Urdu, Kashmiri/Kashur etc., and how they can be adopted to new domains.

The evaluation of MT quality is a difficult task because there may exist many possible ways to translate a given source



sentence. Moreover, the usability of a given translation depends on numerous factors such as the intended use of the translation, the characteristics of the MT software, and the nature of the translation process. Human assessment of four MT systems: AnglaMT, Google, Microsoft Translator and Anuvadaksh, each employing different translation methodology, has been carried out by relatively ranking multiple translations of the same input to each other. Anuvadaksh, hybrid machine translation systems yielded better results, addressing both quality and time-to-market limitations. Hybrid machine translation approach seems to address the issue of MT resource poverty in source-poor/target-rich or target-rich language/ source-poor language pairs by exploiting available symbolic and statistical target-language (TL) resources [7]. As hybrid machine translation suited some languages better than others, especially low resource languages, it turned to be our research endeavor to examine different arrangements of hybrid machine system, by means of studying their architectures.

## 2. HYBRID SYSTEMS

Customarily, MT frameworks have been worked inside of one of the two noteworthy architectures: rule-based machine translation (RBMT) and statistical machine translation (SMT). The RBMT systems explicitly represent and process linguistic knowledge about languages (grammar, lexicon) and about translation equivalents between source and target (translation dictionaries, corresponding grammatical structures). These frameworks have higher precision of linguistic analysis (e.g., they can effectively handle multiword linguistic constructions and re-arrangements of the word order, long-distance dependencies between words or overall syntactic structure of sentences), they have smaller size and use less computational power. Notwithstanding, they have a slower advancement cycle and need manually built dictionaries, grammars and processing algorithms, which seriously limits the number of supported languages.

SMT systems, on the other hand, are built using substantial accumulations of beforehand translated texts (parallel text corpora), which are consequently adjusted on the sentence and word levels and which are stored as large databases of phrases that are translations of each other; translations for new text is generated by intelligent search algorithms that recombine the segments from the database into a faithful and natural translation. SMT systems have faster development cycle, are more precise in determining ambiguities, however have more issues in taking care of sentence-level linguistic phenomena. Moreover, they require more storage room and utilize more computational power: commonly they run as web services on capable servers or computer clusters, which restrain their utilization for off-line mobile applications. All the more as of late scientists endeavored to join SMT and SMT approaches (supposed Hybrid MT), more often than not including some phonetic principles, elements and sentence structure representations on top of SMT frameworks.

The main aim of the machine translation system based on hybridization is to take advantage of the strengths of both rule-based techniques and statistical techniques. Hybrid systems can be categorized into two main groups:

- Single-Engine Hybridization
- Multi-Engine Hybridization

## 3. SINGLE-ENGINE HYBRIDIZATION-(ARCHITECTURE EXTENSIONS)

The architecture of this scheme is based on Rule-based or Statistical paradigm but is modified by using respective other approach. Thurmair (2009)[8] refers this type of hybrid configuration as “Architecture Extensions”. The modifications can occur at various levels depending on the targeted problem.

For example:

- The system data are pre-processed
- Phrase tables are extended by other approach
- Dictionaries are enlarged by other approach.

The first modification is referred to as pre-editing and latter two modifications are referred to as core-modification. Furthermore, a Statistical approach can be enhanced by reordering the source side of parallel corpus to handle the word order issues (Jawaid and Zeinan, 2011)[9].

The Single-Engine Hybridization in two categories based on the main approach used

- Rule-based machine translation modified by statistical techniques(RMT-Extension)
- Statistical Machine translation modified by linguistic information(SMT-Extension)

### 3.1 RMT Extensions:

This approach improves rule-based systems with data driven techniques. It's implemented in two ways:

- Pre-editing : It is done both on the dictionary side by running Term-Extraction tools[10] and enriching dictionary with phrases from an SMT system, and on the grammar side by automatically extracting grammar rules from corpora
- Core-modification: Modification of the system core is done by adding probabilistic information to the analysis and the parsing process. Manipulations in the transfer selection process[11] have also been attempted.

#### 3.1.1 Pre-Editing

This approach implies modifications in language resources of rule-based systems such as dictionaries and grammar rules using data driven techniques.

- I. Pre-editing using dictionary entries: Dictionaries are one of the major component of Rule-based systems. Data-driven techniques are applied for terminology extraction from corpora, either on a monolingual basis or from bilingual corpora. Monolingual corpora helps to find missing entries in the system's dictionaries whereas bilingual corpora helps in finding translation candidates. These approaches are already in use in RMT systems, but still there are few challenges to deal with such as
  - Recognition of the multiword terms: Most of the semantically meaningful words are multiword terms having an internal linguistic structure.
  - Linguistic annotating of the recognized terms: - Terms must be brought into



correct citation form (lemmatized) and annotated with POS etc.

Data-driven techniques have their limitations. Though MT quality improves moderately depending on the amount of reductions of the out-of-vocabulary words, which depends on the size and coverage of the already existing dictionary. This approach helps filling the dictionary gaps. However, as a side-effect, the problem of lexical selection gets aggravated, as amount of translations among which to select increases. This problem becomes much more difficult to solve than the problem of dictionary gaps.

- II. Pre-editing using rules in RMT: The approach of using grammar rules by data-driven techniques does not seem to improve MT output quality. These approaches suffer due to extraction of large quantity of rule candidates from even small corpora. Because of the noise produced by the grammar extraction techniques[12], selection of a grammar rule from a large set of possible candidates lead to the problems of combinatorics and unexpected side-effects, such as parse failure.

### 3.1.2 RMT Core System Modification

Another approach using data-driven techniques is the modification of the core engine of the Rule-based system. This approach has been tried in several respects e.g. using probabilistic information in parsing .The focus of current hybrid methods is towards translation selection in the transfer phase.

Traditional approach to RMT transfer selection rely on two techniques:

- I. Technique 1: Assignment of subject area codes to translations, that is, if a text belongs to a given subject area the respective translation is activated. In case of specific domains general readings of terms in question are also found, so this method is not reliable.
- II. Technique 2: Tests and actions on certain contextual/structure properties such as presence of direct object, certain prepositions, passive voice. These properties trigger a specific translation .However such conditions cannot be reliably stated for lexical selection most often, especially if the number of alternative translation grows. Moreover such tests rely on correct parses of the input sentence which is not guaranteed

Thus need for additional and robust means for lexical selection arises:

**M1:** One way of doing it is to use the more frequently used translation of a given term as default. But this way is not sensitive to the specific context in which a term must be translated and mostly returns the default.

**M2:** Another way is to utilize the contextual information for disambiguating the lexical selection process. The process is based on contextual disambiguation[13]. For a given candidate translation relevant clusters of contextual terms are built at training time from a corpus, at the run-time these contexts are matched against the context of the text to be translated and the best translation is selected .This technique

requires broadening the analysis scope of the system from sentence-based to paragraph-based contexts. It achieves good disambiguation results for the terms it was built for improvement in accuracy are also reported. They follow a probabilistic approach. (Lateral semantic approach)[14] and use a smaller contextual window.

RMT core modifications can improve the transfer selection process but are less successful in case of robustness/ parse failure

## 3.2 SMT Extensions

SMT systems can also be extended to improve translation quality. It is implemented in two ways:

- Pre-editing: It is done to prepare the data, the steps involved are morphology, POS information/syntax information and word reordering
- System core modifications are done by adding RMT information to the phrase tables and by using factored translation

### 3.2.1 Pre-editing

Pre-editing in the SMT extension encompasses following:

- I. Morphology: Morphology has been researched in depth, mainly for the language with rich morphological schemas. Lemmatization and part-of-speech (POS) tagging were used on the source side and on the target side. The goal is to reduce data sparseness using lemma-based language models[15] instead of textform-based. It tries to improve results for smaller corpora. Since surface information has shown to be beneficial, therefore both lemma based and textform analysis should be done. Another important factor is compounding/ decompounding to parallelize alignment. Furthermore, pre-processing is required in languages with agglutinative behaviors e.g. Arabic to split complex word strings into meaningful parts.
- II. Syntax: Recently many statistical MT approaches have included syntactic information as part of the preprocessing phase or decoding phase. With the help of syntax based transformations problems like word order differences[16] or the distance based reordering can be sorted. The basic idea is to apply transformation rules to the source language parse tree to make the order of the source sentence closer to the target sentence and only let syntactically well-formed phrases enter the phrase table. Both source and target side corpora are parsed, matching subtrees are identified and aligned in the phrase table. There may be few parsed phrases on the phrase table but still can help improving the MT output.
- III. Reordering: Reordering is a critical factor for SMT systems. It is because the languages have different word and constituent order[17] and also because the constituent order is



meaningful. Standard phrase-based models are able to handle local reorderings to some extent, but longer distance reordering requires different means.

Many proposals have been put forward to extend the input word sequence into a lattice containing different reordering of the input words. It can be based on Part-of Speech (POS) information. Distortion rules may be set up manually or automatically for continuous and discontinuous POS sequences, by matching them on source and target side of the training corpora. The input lattice contains the distorted strings respectively, with weights on the probability of the distortion. An alternative approach is proposed. In this approach position is not used at all, instead they try a global alignment in a kind of sentence-based bag-of-words strategy. While decoding all possible permutations are created that are allowed by the language model. The problems like multiple occurrences of words in the target need to be checked. In addition, if the source languages information is missing results are not too promising.

(Birch et al-2009)[18] claims that reordering problems depend on the selection of translation models. Long term reordering is handled by hierarchical models, whereas short and medium reordering are handled by phrase-based models. It is still to be researched.

### 3.2.2 SMT Core modifications

The following approaches are identified to incorporate RMT resources into an SMT architecture

- I. Extension of the phrase table by importing RMT resources: This approach was proposed to run RMT systems together with SMT systems and to add terms and phrases produced by RMT systems in the SMT phrase tables. This approach uses knowledge coded in the bilingual dictionaries of the RMT systems. The system can be extended to include text from different domains. Since SMT decoder runs last, it effects the output, which can be less grammatical than one of the original RMT. The system reacts on the data sparseness problem of the SMT training, where as it does not react on the output grammatically problem.
- II. Using Grammars: Recently syntax is being used in decoding .It is a major topic of research.(Charniak et al 2003)[19] showed improvements by increasing the no. of grammatical translation in test by 45%. Many proposals where suggested to learn grammar and transfer rules from bilingual corpora. (Lavoie et al 2002, Hannemann et al. 2008) identified structural contexts for translation selection from bitexts. Hierarchical translation (Chiang 2007)[20] uses synchronous context-free grammar in decoding. Melane 2004 adapts parsing to allow multiple input strings (multi-trees).
- III. Factored Translation: This approach (Koehn et

al 2007b)[21] proposed an integrated approach to incorporate linguistic information in statistical MT. it enriches system by providing more information at word level. It treats words as vectors of features such as lemma, the POS, morphology etc. In other words, phrases are represented as a sequence containing multiple levels of information. According to factored model a word is vector of k factors:

$$w = f_i^1, f_i^2, \dots, f_i^k$$

The factors can be things like morphological classes, stems, roots or it can be data-driven word classes or semantic features useful for sparsely inflected languages. A two-factor language model is generated by standard class-based language models (Brown et al 1992) where one factor is the word class and other is word itself. A factor based model is a model over factors

$$P(f_{1:k}^t / f_{1:t-n}^{1:k})$$

That can be factored as a product of probabilities of the form

$$P(f / f_1, f_2, \dots, f_n)$$

In this approach two tasks are dealt with, first is to find appropriate set of factors. Second is to include an appropriate statistical model over those factors. (Bilmes & Kirchoof, 2004; Dep, 2008). This model follows the statistical modeling methods used in phrase-based models. Each step is modeled by a feature function. This function is learned from the training data, resulting in translation tables and generation tables. The approach decomposes phrase translations into a sequence of mapping steps, with translation steps operating on phrase level and generation steps on word level. Models are combined in bi-linear fashion. The approach uses structural information for both source to target mapping as well as target decoding. This improves quality for systems combining these factors. They use both knowledge-driven and data-driven information. But they rely on availability of bilingual corpora. This fact may reduce their applicability.

## 4. MULTI-ENGINE HYBRIDIZATION

Multi-engine architecture is also refereed as coupling by Thurmair. It involves combination of two or more existing systems to produce improved MT output. Coupling can be either done in a serial way or the parallel way. In a serial way, the most researched approach is statistical post-editing of rule-based system. In a parallel way the best translation is produced from the output of several systems.

### 4.1 Parallel Coupling

This type of coupling uses several MT systems in parallel, and produces the best output from the result set. It can be divided into two paradigms

- I. The first and simple approach is to select the best

output/translation from a list of n-best outputs/translations (Hildebr and Vogel 2008)[22]. They search all output hypothesis available for the best n-grams and then select the best hypothesis from the candidate list. As the resulting text can integrate sentences from different MT system outputs, they report an improvement of 2-3 BLEU compared to the best single system.

- II. The second approach is more sophisticated one. It recombines the best parts from multiple hypothesis into a new utterance that can be better than the best of the given candidate. This approach doesn't work on whole sentences but on smaller segments like phrases, words. It uses confusion networks, and generates an output sentence on the basis of the available MT outputs. One of its variants is phrase-level approach. It extracts sentence-specific phrase translation tables from system outputs with alignments to source and runs a phrasal decoder with this new translation table. A Skelton is selected as a base, for each position of the Skelton the best translation alternative is identified and composed to the overall output sentence. Skeletons can be selected on sentence as well as phrase level. The choice of the best Skelton is critical because it determine the structure/word order of the target sentence.

## 4.2 Serial Coupling

Systems which offer serial coupling consists of modifying RMT O/P by means of a SMT post editing component. The SMT component uses a bilingual training component. System using small domains ( Simal et al. 2007, data from Canadian Job Bank)[23] showed that even using relatively small training data, significant improvements of the MT O/P can be achieved. The combination of RMT and SPE component outperforms pure SMT systems in case where only limited data are available

Experiments on a broader database show the following result

- I. Combination of RMT and SPE systems are highly competitive in MT quality. The O/P seems to be grammatical and main result of the combination is an increase in lexical selection quality.
- II. Errors due to SMT postprocessor need to be avoided. These errors include
  - Syntactic structure of the output can be confused by PSE component
  - Accuracy drops due to omission of some parts of translation.
- III. To overcome lexical selection problem of the RMT systems, use RMT system's syntactic structure and only try local alternatives.

### 4.2.1 Multi Engine MT driven by s SMT decoder

These systems transform the information contained in the MT O/P into a form that is suitable as input for an existing SMT decoder. The material extracted from the rule-based. MT results is combined resources used in standard phrase-based SMT. The task of the combination is to find good relative weights for the various phrase table entries.

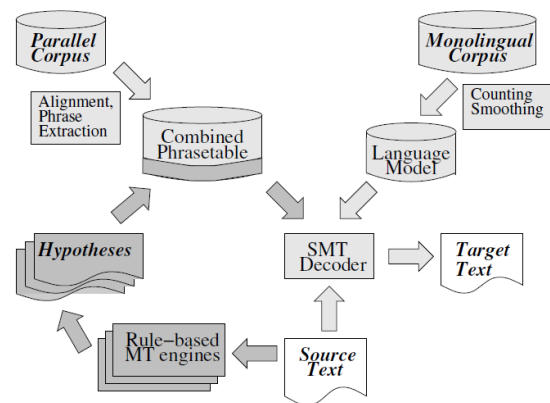


Fig 1. Architecture for Multi-engine MT Driven by a SMT Decoder

The above mentioned architecture can be seen as a variant of SMT and places a strong emphasis on statistical models. Here lexical information from rule-based engines is used to increase lexical coverage. Since rule-based MT engine suffer from missing lexical coverage but an advantage of SMT is that lexical entries can be automatically induced from existing translations. Thus the control of whole arrangement is under rule-based engine, which has an advantage that well-formed syntactic structures generated using linguistic rules cannot be broken apart by the SMT components. Rule-based systems lack mechanism for ruling out implausible results and cannot cope with errors that creep into the lexicon due misalignments. Therefore entries taken from statistical alignments need to be carefully filtered. Also information derived from word alignment of a given translation lack linguistic information that is required by a rule-based system. The corresponding expressions in a parallel corpus are formed as inflected full forms, the entries in bilingual dictionary contain normalized forms and also morphological classification that defines all possible inflectional forms of the given entry.

## 5. GENUINE HYBRID ARCHITECTURE

These hybrid architecture combine entire system components of the respective approaches to form a new system. They use following basic components:

- Identification of SLanguage chunks (words, phrases etc.)
- Transformation of these chunks into target language using bilingual resource
- Generation of target language sentence.

### 5.1 Hybrid System using Rule-based Analysis

This approach has been investigated using METIS Projects (Vandeghinste et al. 2006)[24]. Analysis is done using available Natural language tools such as lemmatisers, taggers, chunkers. Transfer is based on existing dictionaries consisting of lemma and POS in source and target language including single and multiword terms and generation uses a language model based on a tokenized and tagged English corpus (BCN).



## 5.2 Hybrid System using Data Driven Analysis

This approach has been proposed by ( Carbonell et al.2006). The resources required are

- A bilingual dictionary
- n-gram indexed target language corpus.

In analysis an n-gram window is moved over the sentence and all words in the window are translated using the bilingual dictionary. Based on these translations the target language corpus is searched for the closest n-gram containing all words of the source. The result is a lattice of n-gram translations. Among these the segments with the strongest left and right overlaps and highest density of terms are selected by the decoder.

## 6. CONCLUSION

For optimizing the machine translation, hybrid architecture produce suitable and quality results, however the results may vary from architecture to architecture. Following factors are responsible for selection of appropriate architecture of MT system

- I. Translation domain
- II. Translation quality
- III. Availability of resource and data, both on monolingual and bilingual level.

Most of the research combines sources of information (rules and data), but there are also projects combining various corpus-based approaches. It is difficult to assess which is the most relevant or promising hybrid type of architecture, but it would seem reasonable to use the best-performing system as a guide, and the others for additional information. The good results produced by hybridization have led to a corresponding spread of MT applications such as speech translation, cross-language information retrieval, computer-aided and post-edited MT systems. Work with hybrid strategies in both in MT and its applications bring significant improvement because they allow the simultaneous exploitation of a variety of systems. For more complex languages (Hindi, Urdu, Kashmiri/Kashur) it is worthwhile to invest in hybrid systems.

## 7. ACKNOWLEDGMENT

This research paper is made conceivable through the assistance and backing from everybody, counting: folks, educators, family, companions, and generally, all conscious creatures.

## 8. REFERENCES

- [1] Sneha Tripathi, Juran ksrihna Srakhel – Approach to Machine Translation. Annals of Library and Information Studies Vol. 57, December 2010, PP. 388-393
- [2] Kaji H, An efficient method for rule-based machine translation (1988). Available at: <http://www.aclweb.org/anthology/C/C88/C88-2167.pdf>. (Accessed on 3rd January, 2016)
- [3] Macherey,L, Bender O and Ney H, Applications of Statistical machine translation approached to spoken language understanding, IEEE transactions on Audio Speech and Language Processing, 17(4)(2009) PP.803-818
- [4] Marta R. Costa-Jussa, Jose A.R. Fonollosa, Latest trends in hybrid machine translation and its applications, Computer Speech and Language 32 (2015) PP.3-10
- [5] Callison-Burch, Chr., Koehn, Ph.,Monz,Ch., Schroeder, J., 2009:Finding of the 2009 workshop on Statistical Machine Translation. Proc. 4th Workshop on SMT, Athens.
- [6] Raffaella Bernardi, Computational Linguistics: Long-Distance Dependencies, KRDB, Free University of Bozen-Bolzano.P.zza Domenicani, Room:228,email:BERNARDI@INF.UNIB2.IT
- [7] Neeha Ashraf, Manzoor Ahamd, Experimental Framework using Web-based Tools for Evaluation of Machine Translation Techniques, International Journal of Advanced Research in Computer Science and Software Engineering, V6I4, 2016 PP.223-228.
- [8] Available at <http://www.mt-archive.info/MTS-2009-Thurmair.pdf>. (Accessed on 17th January, 2016)
- [9] Available at: <https://ufal.mff.cuni.cz/pbml/95/art-jawaid-zeman.pdf> (Accessed on 19th January, 2016)
- [10] Available at: <http://www.termcoord.eu/discover/free-term-extractors/term-extraction-tool/> (Accessed on 4th February, 2016)
- [11] Available at: <https://www.aclweb.org/antologu/P15-2040> (Accessed on 10th February, 2016)
- [12] Fei Xia, Martha Palmer, Aravind Joshi, A Uniform Method of Grammar Extraction and its Applications. Available at: <http://www.aclweb.org/anthology/W00-1307> (Accessed on 14th February, 2016 )
- [13] Eric Foxley, Godwin M. Gwei, Synonymy and Contextual Disambiguation of Words, Int J lexicography (1989) 2920:11-134.
- [14] Available at <https://arxiv.lateral.io/document/bac54fo741e92f6185bee68d7f2f9dd5> ( Accessed on 15th February, 2016
- [15] Tanel Alumae, Toomas Kirt; LSA-based Language Model Adaptation for Highly Inflected Languages. Available at: <https://phon.ioc.ee/dokuwiki/lib/exe/fetch.php?media=people:tanel:eurospeech07.pdf> (Accessed on 19th February, 2016)
- [16] Sabine Mohr, Universitat Stottgart Word order difference and their implications, March 22nd, 2004.
- [17] Available at: [http://courses.washington.edu/asian401/lectures/asian401\\_wk6\\_movement\\_agreement.pdf](http://courses.washington.edu/asian401/lectures/asian401_wk6_movement_agreement.pdf) (Accessed on 01 March, 2016).
- [18] Birch, A., Blunsom, Ph., Osborne, M., 2009: A Quantitative Analysis of Reordering Phenomena. Proc. 4th Workshop on SMT, Athens .
- [19] Charniak, E., Knight, K., Yamada, K., 2003, Syntax based Language Models for Statistical Machine Translation. Proc. MT Summit IX, New Orleans .
- [20] Chiang. D., 2007: Hierarchical Phrase-Based Translation in: Computational Linguistics, 33, 2.
- [21] Koehn, Ph., Hoang, H., 2007: Factored Translation



Models. Proc. ACL, Prague.

- [22] Hildebrand, A.S., Vogel St., 2008, Combination of Machine Translation System via Hypothesis Selction from Combined N-Best Lists. Proc. AMTA, Hawaii.
- [23] Simard, M., Ueffing, N., Isabelle, P., Kuhn, R., 2007: Rule-based Translation with Statistical Phrase-based
- Post-editing. AL 2007 second workshop on Statistical Machine Translation. Prague.
- [24] Vandeghinste, V., Schuurman, I., Carl, M., Markantonatous, St., Badia, T., 2006, METIS-II: Machine Translation for Low Resource Languages. Proc. LREC Genoa