Human Evaluation of Online Machine Translation Services for English/Russian-Croatian

Sanja Seljan[1], Marko Tucaković[2], Ivan Dunđer[1]

[1] Faculty of Humanities and Social Sciences – University of Zagreb, Department of Information and Communication Sciences, Ivana Lucica 3, 10 000 Zagreb, Croatia
sanja.seljan@ffzg.hr, marko.tucakovic@gmail.com, ivandunder@gmail.com

Abstract. This paper presents results of human evaluation of machine translated texts for one non closely-related language pair, English-Croatian, and for one closely-related language pair, Russian-Croatian. 400 sentences from the domain of tourist guides were analysed, i.e. 100 sentences for each language pair and for two online machine translation services, Google Translate and Yandex.Translate. Human evaluation is made with regard to the criteria of fluency and adequacy. In order to measure internal consistency, Cronbach's alpha is calculated. Error analysis is made for several categories: untranslated/omitted words, surplus words, morphological errors/wrong word endings, lexical errors/wrong translations, syntactic errors/wrong word order and punctuation errors. At the end of this paper, conclusions and suggestions for further research are given.

Keywords: machine translation, human evaluation, Google Translate, Yandex.Translate, English, Russian, Croatian, adequacy, fluency, error analysis, inter-evaluator agreement

1 Introduction

Machine translation refers to the process of translating from one natural language to another by applying various computer programs. It is closely related to the development of computational linguistics, artificial intelligence, computer science and language processing methods. Natural languages are characterised by large variabilities of expressions, exceptions to grammatical rules and context dependent changes, which makes automatic machine translation a challenging and demanding task, especially if the languages are morphologically rich and syntactically complex [1]. Machine translation technology is becoming more accessible for everyday needs, but the output quality is not always satisfactory, especially for less-resourced languages. Still, online machine translation has been used for quick access to information and gisting purposes, for browsing or cross-lingual information retrieval.
……

References

1. Ni, Y., Saunders, C., Szedmak, S., Niranjan, M. Exploitation of Machine Learning Techniques in Modelling Phrase Movements for Machine Translation. In: Journal of Machine Learning Research 12, 1-3 (2011)
2. Koehn, P. Statistical Machine Translation. Cambridge University Press, New York, NY, USA, 1st ed. (2010)
3. Vasiljevs, A., Gornostay, T., Skadins, R. LetsMT! – Online Platform for Sharing Training Data and Building User Tailored Machine Translation. In: Proceedings of the Fourth International Conference Baltic HLT, pp. 133-140. Amsterdam, IOS Press, Netherlands (2010)
4. Latour, J. Evaluating Statistical Machine Translation from English to Dutch. In: Proceedings of the 1st Twente Student Conference on IT: Intelligent Interaction. Twente, Netherlands (2004)
5. Och, F.J. Minimum Error Rate Training in Statistical Machine Translation. In: Proceeding ACL '03 of the 41st Annual Meeting on Association for Computational Linguistics, vol. 1, pp. 160-167 (2003)
6. Koehn, P. Statistical Significance Tests for Machine Translation Evaluation, EMNLP (2004)
7. Specia, L., Hajlaoui, N., Hallett, C., Aziz, W. Predicting Machine Translation Adequacy. Machine Translation Summit XIII. Xiamen, China (2011)
8. Quirk, C. Training a Sentence-Level Machine Translation Confidence Measure. In: Proceedings of the 4th Conference on Language Resources and Evaluation, pp. 825–828. Lisbon, Portugal (2004)
9. Castilho Monteiro de Sousa, S., Aziz, W., Specia, L. Assessing the post-editing effort for automatic and semi-automatic translations of DVD subtitles. In: International Conference Recent Advances in Natural Language Processing, pp. 97–103. RANLP, Hissar, Bulgaria (2011)
10. Garcia-Santiago, L.; Olvera-Lobo, M. D. Automatic Web Translators as Part of a Multilingual Question-Answering (QA) System: Translation of Questions. In: Translation Journal. vol. 14, 1 (2010)
11. Hampshire, S.; Porta Salvia, C. Translation and the Internet : Evaluating the Quality of Free Online Machine Translators. In: Quaderns: revista de traducció, 17, pp. 197-209 (2010)
12. Kit, C., Wong, T. M. Comparative Evaluation of Online Machine Translation Systems with Legal Texts. In: Law Library Journal. vol. 100 (2), pp. 299–321 (2008)
13. Zervaki, T. Online Free Translation Services. In: Proceedings of the 24th International Conference on Translating and the Computer. London, Aslib (2002)
14. Seljan, S.; Brkić, M; Kučiš, V. Evaluation of Free Online Machine Translations for Croatian-English and English-Croatian Language Pairs. In: INFuture2011: The Future of Information Sciences – Information Sciences and e-Society, pp. 331-344. Zagreb (2011)
15. Seljan, S., Vičić, T., Brkić, M. BLEU Evaluation of Machine-Translated English- Croatian Legislation. In: Proceedings of the Eighth International Conference on Language Resources and Evaluation - LREC, pp. 2143-2148. ELRA (2012)
16. Ma, X., Cieri, C. Corpus Support for Machine Translation at LDC. In: Proceedings of the 5th edition of the International Conference on Language Resources and Evalution - LREC, pp. 859-864 (2006)
17. Soricut, E. TrustRank: Inducing Trust in Automatic Translations via Ranking. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, pp. 612–621. Uppsala, Sweden (2010)
18. He, Y., Ma, Y., van Genabith, J., Way, A. Bridging SMT and TM with translation recommendation. In: ACL 2010 - 48th Annual Meeting of the Association for Computational Linguistics. Uppsala, Sweden (2010)

19. Specia, L. Exploiting Objective Annotations for Measuring Translation Post-editing Effort.15th Annual Conference of the European Association for Machine Translation, pp. 73-80. Leuven, Belgium (2011)