An Open Source Desktop Post-Editing Tool

Lane Schwartz   lanes@illinois.edu

Department of Linguistics, University of Illinois at Urbana-Champaign,
Foreign Languages Building, 707 S Mathews Ave, Urbana, IL 61801, USA

We present a simple user interface for post-editing that presents the user with the source sentence, machine translation, and word alignments for each sentence in a test document (Figure 1). This software is open source, written in Java, and has no external dependencies; it can be run on Linux, Mac OS X, and Windows.

This software was originally designed for monolingual post-editors, but should be equally usable by bilingual post-editors. While it may seem counter-intuitive to present monolingual post-editors with the source sentence, we found that the presence of alignment links between source words and target words can in fact aid a monolingual post-editor, especially with regard to correcting word order. For example, in our experiments using this interface (Schwartz et al., 2014), post-editors encountered some sentences where a word or phrase was enclosed within bracketing punctuation marks (such as quotation marks, commas, or parentheses) in the source sentence, and the machine translation system incorrectly reordered the word or phrase outside the enclosing punctuation; by examining the alignment links the post-editors were able to correct such reordering mistakes.
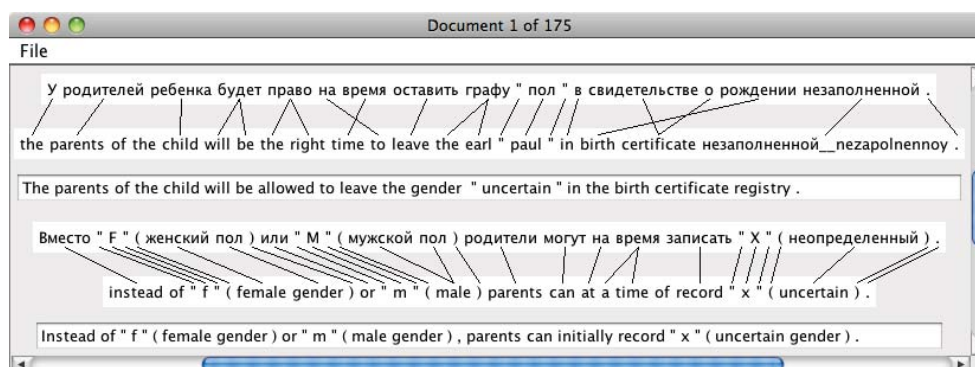

Figure 1: Post-editor user interface

References

Schwartz, L., Anderson, T., Gwinnup, J., and Young, K. M. (2014). Machine translation and monolingual postediting: The AFRL WMT-14 system. In *Proceedings of the Ninth Workshop on Statistical Machine Translation*, pages 186–194, Baltimore, Maryland. Association for Computational Linguistics.