PhraseFix: Statistical Post-Editing of TectoMT
Petra Galuščáková, Martin Popel, and Ondřej Bojar
Charles University in Prague, Faculty of Mathematics and Physics
Institute of Formal and Applied Linguistics
{galuscakova,popel,bojar}@ufal.mff.cuni.cz

Conclusion
● SPE works: PhraseFix better than TectoMT.
● SPE is further improved by marking identical and reliable phrases using a special feature.
● However, SMT still outperforms SPE.
● Using TectoMT output as additional data improves the SMT baseline.

Variants
● Corpus Combination (CComb) – output of TectoMT appended to the baseline parallel corpus
● Phrase Table Combination (PTComb) – output of TectoMT provides a secondary phrase table, MERT selects weights

Variant Results
Baseline: CzEng
CzEng + tmt(newstest2013)
18.2
18.4
18.6
18.8
19

This research is supported by the grants GAUK 9209/2013, FP7-ICT-2011-7-288487 (MosesCore) of the European Union and SVV project number 267 314.