

Translation of Multi-word Expressions into Under-resourced Language: Case of English-Latvian Statistical Machine Translation

- Two series of experiments with MWE extraction and integration in a phrase-based SMT system:
 - experiments with linguistically motivated MWEs
 - experiments with MWE candidates extracted using association measures
- Manual inspection of obtained results showed some improvement of fluency and adequacy

Method	BLEU
Baseline	62.40
Baseline + MWE training data	62.44
Two translation tables	62.55
Additional feature	62.27

English: the agency shall draw up analytical accounts of its revenue and expenditure.

Human: aģentūra **veic** analītisku ieņēmumu un izdevumu uzskaiti.

Baseline: aģentūra **izstrādā** analītisko uzskaiti par tā ieņēmumiem un izdevumiem.

Improved SMT: aģentūra **veic** analītisku ieņēmumu un izdevumu uzskaiti.