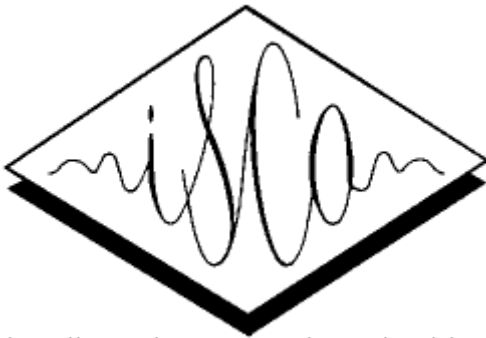


ISCA Archive

<http://www.isca-speech.org/archive>

**INTERSPEECH
2007
8th Annual
Conference of
the
International
Speech
Communication
Association**



**Antwerp, Belgium
August 27-31, 2007**

Bilingual LSA-Based Translation Lexicon Adaptation for Spoken Language Translation

Yik-Cheung Tam, Tanja Schultz

Carnegie Mellon University, USA

We present a bilingual LSA (bLSA) framework for translation lexicon adaptation. The idea is to apply marginal adaptation on a translation lexicon so that the lexicon marginals match to in-domain marginals. In the framework of speech translation, the bLSA method transfers topic distributions from the source to the target side, such that the translation lexicon can be adapted before translation based on the source document. We evaluated the proposed approach on our Mandarin RT04 spoken language translation system. Results showed that the conditional likelihood on the test sentence pairs is improved significantly using an adapted translation lexicon compared to an unadapted baseline. The proposed approach showed improvement on BLEU-score in SMT. When both the target-side LM and the translation lexicon were adapted and applied simultaneously for SMT decoding, the gain on BLEU-score was more than additive compared to the scenarios when the adapted models were individually applied.

[Full Paper](#)

Bibliographic reference. Tam, Yik-Cheung / Schultz, Tanja (2007): "Bilingual LSA-based translation lexicon adaptation for spoken language translation", In *INTERSPEECH-2007*, 2461-2464.