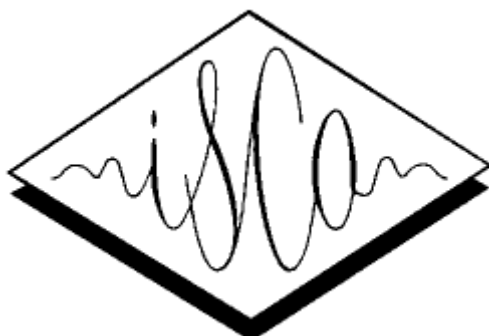


ISCA Archive

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

**EUROSPEECH
2003 -
INTER_SPEECH
2003
8th European
Conference on
Speech
Communication
and Technology**

**Geneva, Switzerland
September 1-4, 2003**



Example-Based Bi-Directional Chinese-English Machine Translation with Semi-Automatically Induced Grammars

K.C. Siu, Helen M. Meng, C.C. Wong

Chinese University of Hong Kong, China

We have previously developed a framework for bi-directional English-to-Chinese/Chinese-to-English machine translation using semi-automatically induced grammars from unannotated corpora. The framework adopts an example-based machine translation (EBMT) approach. This work reports on three extensions to the framework. First, we investigate the comparative merits of three distance metrics (Kullback-Leibler, Manhattan-Norm and Gini Index) for agglomerative clustering in grammar induction. Second, we seek an automatic evaluation method that can also consider multiple translation outputs generated for a single input sentence based on the BLEU metric. Third, our previous investigation shows that Chinese-to-English translation has lower performance due to incorrect use of English inflectional forms - a consequence of random selection among translation alternatives. We present an improved selection strategy that leverages information from the example parse trees in our EBMT paradigm.

[Full Paper](#)

Bibliographic reference. Siu, K.C. / Meng, Helen M. / Wong, C.C. (2003): "Example-based bi-directional Chinese-English machine translation with semi-automatically induced grammars", In *EUROSPEECH-2003*, 2801-2804.