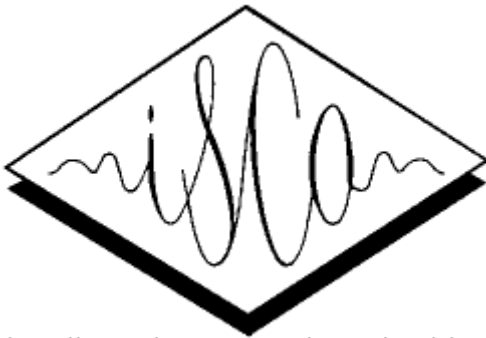


ISCA Archive

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

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



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

Improvements in Machine Translation for English/Iraqi Speech Translation

**S. Saleem, K. Subramanian, R. Prasad, David Stallard, Chia-Lin Kao, P.
Natarajan, R. Suleiman**

BBN Technologies, USA

In this paper, we describe techniques for improving machine translation quality in the context of speech-to-speech translation for significantly different language pairs. Specifically, we explore three broad approaches for improving translation from English to Iraqi and vice versa. First, we investigate normalization techniques which address the differences in spoken and written forms of both languages. Second, we incorporate additional knowledge sources into the translation process such as a bilingual lexicon and named entity detection. Third, we exploit the rich morphological structure of Iraqi Arabic using two different approaches. The first approach decomposes words in Iraqi Arabic whereas the second approach, a novel one inflects English by combining key phrases into words using the minimum descriptive length criterion. Significant gains in accuracy are observed, while translating from text as well as speech recognition output.

[Full Paper](#)

Bibliographic reference. Saleem, S. / Subramanian, K. / Prasad, R. / Stallard, David / Kao, Chia-Lin / Natarajan, P. / Suleiman, R. (2007): "Improvements in machine translation for English/Iraqi speech translation", In *INTERSPEECH-2007*, 2445-2448.