

Reddy, A., Rose, R., Safadi, H., Larkin, S. and Boulianne, G. "Incorporating Knowledge of Source Language Text in a System for Dictation of Document Translations" In *Machine Translation Summit XII.* , Ottawa, Ontario, Canada, August 26 - September 30, 2009.

Abstract

A method is presented which integrates target language automatic speech recognition (ASR) models with source language statistical machine translation (SMT) and named entity recognition (NER) information at the phonetic level. Information extracted from a source language document including translation model probabilities and translated named entities are combined with acoustic-phonetic information obtained from phone lattices produced by the ASR system. Phone-level integration allows the combined MAHT system to correctly decode words that are either not in the ASR vocabulary or would have been incorrectly decoded by the ASR system. It is shown that the combined MAHT system results in a decrease in word error rate on the dictated translations of up to 32% relative to a stand alone baseline ASR system.