EUROSPEECH 2003 - INTERSPEECH 2003
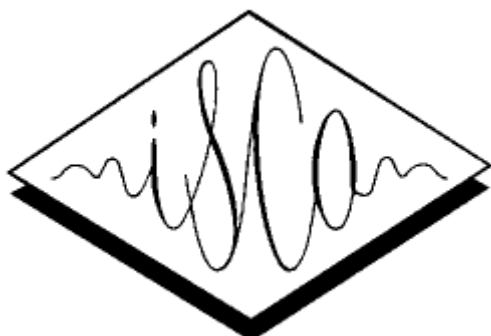## 8th European Conference on Speech Communication and Technology

Geneva, Switzerland
September 1-4, 2003

Coupling vs. Unifying: Modeling Techniques for Speech-to-Speech Translation

Yuqing Gao

IBM T.J. Watson Research Center, USA

As a part of our effort to develop a unified computational framework for speech-to-speech translation, so that sub-optimizations or local optimizations can be avoided, we are developing direct models for speech recognition. In direct model, the focus is on the creation of one single integrated model p(text| acoustics), rather than a complex series of artifices, therefore various factors such as linguistics and language features, speaker or speaking rate differences, different acoustic conditions, can be applied to the joint optimization. In this paper we discuss how linguistic and semantic constraints are used in phoneme recognition.

Full Paper

Bibliographic reference. Gao, Yuqing (2003): "Coupling vs. unifying: modeling techniques for speech-to-speech translation", In *EUROSPEECH-2003*, 365-368.