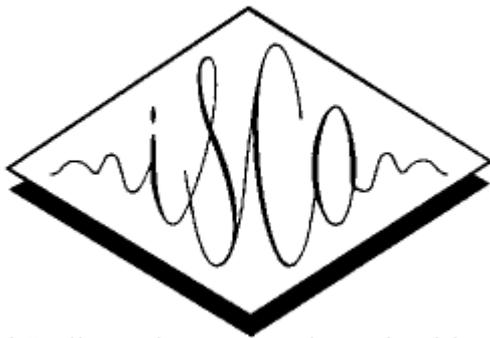


ISCA Archive

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

**INTER_SPEECH
2004 - ICSLP
8th International
Conference on
Spoken Language
Processing**

**Jeju Island, Korea
October 4-8, 2004**



Multilingual Corpora for Speech-to-Speech Translation Research

Genichiro Kikui, Toshiyuki Takezawa, Seiichi Yamamoto

ATR, Japan

Multilingual spoken language corpora are indispensable for developing new speech-to-speech machine translation (S2SMT) technologies. This paper first discusses characteristics that corpora for S2SMT should have, then surveys existing corpora. Finally, it compares these corpora focusing on relations between collected data and collection scheme including instructions given to speakers.

[Full Paper](#)

Bibliographic reference. Kikui, Genichiro / Takezawa, Toshiyuki / Yamamoto, Seiichi (2004): "Multilingual corpora for speech-to-speech translation research", In *INTER_SPEECH-2004*, 357-360.