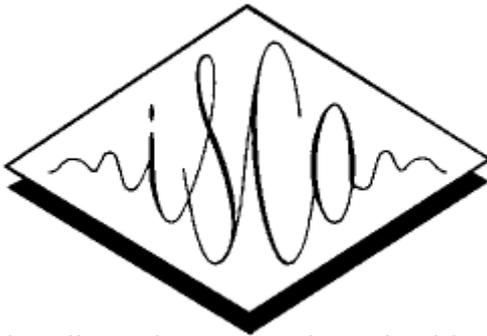**INTERSPEECH
2007
8th Annual
Conference of
the
International
Speech
Communication
Association**

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

The VoiceTRAN Machine Translation System

Jerneja Žganec Gros, Stanislav Gruden

Alpineon Research, Slovenia

Freely available tools and language resources were used to build the VoiceTRAN statistical machine translation (SMT) system. Various configuration variations of the system are presented and evaluated. The VoiceTRAN SMT system outperformed the baseline conventional rule-based MT system in both English-Slovenian in-domain test setups. To further increase the generalization capability of the translation model for lower-coverage out-of-domain test sentences, an "MSD-recombination" approach was proposed. This approach not only allows a better exploitation of conventional translation models, but also performs well in the more demanding translation direction; that is, into a highly inflectional language. Using this approach in the out-of-domain setup of the English-Slovenian JRC-ACQUIS task, we have achieved significant improvements in translation quality.

[Full Paper](#)

Bibliographic reference. Gros, Jerneja Žganec / Gruden, Stanislav (2007): "The voiceTRAN machine translation system", In *INTERSPEECH-2007*, 1521-1524.