INTERSPEECH 2008
9th Annual Conference of the International Speech Communication Association

Brisbane, Australia September 22-26, 2008

Czech-to-Slovak Adapted Broadcast News Transcription System

Jan Nouza, Jan Silovsky, Jindrich Zdansky, Petr Cerva, Martin Kroul, Josef Chaloupka

Technical University of Liberec, Czech Republic

The first broadcast news (BN) transcription system for Slovak is introduced. It employs the same modules as the system we developed earlier for Czech. We utilize similarity between the two languages in efficient lexicon building, in mapping Slovak specific (rarely occurring) phonemes onto Czech ones and in low-resource cross-lingual adaptation of acoustic model. The system uses 166K-word lexicon and on the Slovak part of European COST278 BN database achieves 23.6% WER (which is only 5% less than the original, long-term optimized Czech system). Similar results were achieved also on recently recorded data from four Slovak stations.

Full Paper

Bibliographic reference.   Nouza, Jan / Silovsky, Jan / Zdansky, Jindrich / Cerva, Petr / Kroul, Martin / Chaloupka, Josef (2008): "Czech-to-slovak adapted broadcast news transcription system", In *INTERSPEECH-2008*, 2683-2686.