Recent Advances in Spoken Language Translation

Marcello Federico

FBK, Trento, Italy

The talk is structured in three parts. The first part overviews problems and approaches to spoken language translation. The second part presents challenges and achievements of the European Project TC-STAR, that ended in 2007. The third part describes advances in the use of confusion networks as interface between automatic speech recognition and machine translation. In particular, I will discuss an efficient implementation of a decoding algorithm for confusion networks and describe its use to translate ASR word lattices and to enrich translations with punctuation marks. I will also report experimental results on the translation of speeches of the European Parliament, from Spanish to English and viceversa.