145 ways to insert punctuation for speech translation

Vincent Vandeghinste (Centre for Computational Linguistics), Lyan Verwimp (ESAT-PSI), Joris Pelemans (ESAT-PSI) and Patrick Wambacq (ESAT-PSI)

firstname.lastname@kuleuven.be

In speech translation the first step often consists of automatic speech recognition (ASR), which outputs an unsegmented stream of words. Translating this stream of words, using off-the-shelf machine translation (MT) results in a lower translation quality compared to translating punctuated input.
We compare different strategies and techniques to deal with this problem.

We compare different punctuation strategies: *Preprocessing* inserts punctuation in the source text and allows to use an MT system trained on a normal (i.e. punctuated) parallel corpus. *Implicit* insertion inserts punctuating during the translation, and requires a dedicated MT system. *Postprocessing* inserts punctuation in the output of the MT system, which is trained on unpunctuated data.

We predict punctuation using language modeling techniques, such as *n*-grams and long short-term memories (LSTM), sequence labeling LSTMs (unidirectional and bidirectional), and monolingual phrase-based, hierarchical and neural MT, and intrinsically evaluate punctuation prediction accuracy.

For actual translation phrase-based, hierarchical and neural MT are investigated.

We set up an experiment in which we combine all these strategies, punctuation prediction methods and machine translation methods, resulting in 145 experimental conditions.
We observe that for punctuation prediction, phrase-based statistical MT and neural MT reach similar results, and are best used as a preprocessing step which is followed by neural MT to perform the actual translation. Implicit punctuation insertion by a dedicated neural MT system, trained on unpunctuated source and punctuated target, yields similar results.