# The Challenge of Simultaneous Speech Translation

**Anoop Sarkar**
Simon Fraser University
`anoop@sfu.ca`

## Abstract

Simultaneous speech translation attempts to produce high quality translations while at the same time minimizing the latency between production of words in the source language and translation into the target language. The variation in syntactic structure between the source and target language can make this task challenging: translating from a language where the verb is at the end increases latency when translating incrementally into a language where the verb appears after the subject.

In this talk I focus on a key prediction problem in simultaneous translation: when to start translating the input stream. I will talk about two new algorithms that together provide a solution to this problem. The first algorithm learns to find effective places to break the input stream. In order to balance the often conflicting demands of low latency and high translation quality, the algorithm exploits the notion of Pareto optimality. The second algorithm is a stream decoder that incrementally processes the input stream from left to right and produces output translations for segments of the input. These segments are found by consulting classifiers trained on data created by the first algorithm.

We compare our approach with previous work and present translation quality scores (BLEU scores) and the latency of generating translations (number of segments translated per second) on audio lecture data from the TED talks collection.

30th Pacific Asia Conference on Language, Information and Computation (PACLIC 30)
Seoul, Republic of Korea, October 28-30, 2016

7