

## **Impact of acoustical voice activity detection on spontaneous filled pause classification**

### **ABSTRACT**

Filled pause detection is imperative for spontaneous speech recognition as it may degrade speech recognition rate. However, filled pause is commonly confused with elongation as they shared the same acoustical properties. Few attempts of classifying filled pause and elongation employed Hidden Markov model. Our proposed method of utilizing Neural Network as a classifier achieved 96% precision rate. We also proved that voice activity detection (VAD) affects the performance of speech recognition. Three acoustical-based VAD are compared and the best precision rate is achieved by incorporating volume and first-order difference features. Experiments are conducted using Malay language spontaneous speeches of Malaysia Parliamentary Debate sessions.

**Keyword:** Elongations; Filled pause; Multi-layer perceptron neural network; Voice activity detection