Title: Statistical parametric evaluation on new corpus design for malay speech articulation disorder early diagnosis

Author/Authors: Mohd. Nizam Mazenan, Tantian Swee, Tanhui Ru, Azran Azhima

Abstract: Speech-to-Text or always been known as speech recognition plays an important role nowadays especially in medical area specifically in speech impairment. In this study, a Malay language speech-to-Text system was been designed by using Hidden Markov Model (HMM) as a statistical engine with emphasizing the way of Malay speech corpus design specifically for Malay articulation speech disorder. This study also describes and tests the correct number of state to analyze the changes in the performance of current Malay speech recognition in term of recognition accuracy. Statistical parametric representation method was utilized in this study and the Malay corpus database was constructed to be balanced with all the phonetic placed and manner of articulation sample appeared in Malay speech articulation therapy. The results were achieved by conducting few experiments by collecting sample from 80 patient speakers (child and adult) and contain for almost 30,720 sample training data.