PAPER # 6

- **TITLE:**

Optimal Entropy to enhance the structure of the Wavelet-Packets-Best-Tree for Automatic Speech Recognition.

- **YEAR OF PUBLICATION:** September 2021.
- **JOURNAL:** Egyptian Journal of Language Engineering
- **ABSTRACT:**

Best Tree Encoding (BTE) is a promising feature extraction technique based on wavelet packet decomposition that is utilized in Automatic Speech Recognition (ASR). This research introduces an enhancement of Wavelet Packet Best Tree (WPBT) Calculations. The standard features BTE encodes the tree structure using a mathematical model into a features vector of 4 components. The best tree structure has been calculated using the entropy function. In the standard version of BTE, Shannon entropy has been chosen as the entropy function. In this research, Shannon Entropy (SE), Renyi Entropy (RE), and Tsallis Entropy (TE) are used to construct the Best Tree. The encoding of the Best Tree has been done using the same mathematical model approach in the standard 4-Point BTE. The proposed model is tested and Verified against the most widely used feature Mel Frequency Cepstral Coefficient (MFCC) plus delta and delta-delta coefficients (39 parameters) to evaluate its performance. The TIMIT database is used in this research. All phones are divided into five classes: Vowels, Fricatives, Silences, Nasals, and Plosives. The acoustical model

has been implemented using Hidden Markov Model (HMM). No language model has been applied. The HMM Tool Kit (HTK) software is used for model implementation. The experiments show that BTE using Tsallis entropy yields the highest overall success rate of 75.85% which is better than MFCC's overall success rate of 71.76%. Comparing the vector of 4 components of BTE to the 39 components vector of MFCC makes it a very promising feature vector to be considered for research and development.