**Literature Survey**

People with neurological diseases like PD also suffer from Cognitive Impairment(CI). The cognitive impairment causes speech disfluencies and errors. The author of this paper [3] analyses the changes in read patterns caused by errors and disfluencies as a function of CI. They start by analysing errors and disfluencies in manual transcripts with annotations specifying where and which errors and disfluencies occur. Then, they detect the same errors and disfluencies from manual transcripts without annotations and couple that detection with Automatic Speech Recognition(ASR) to generate transcripts and detect errors and disfluencies. They found that these features can be used in a regression framework to predict MoCA Scores. Thus, they hypothesize that these features will be sensitive to cognitive decline for people without PD, but the specific error and disfluency patterns may differ.

The articulatory and phonatory aspects of speech can be used to assess neurological diseases better [1]. They conclude that although many studies use the acoustic features such as jitter and shimmer, the usefulness of phonatory approaches in the diagnosis of neurological disorders is not apparent yet. So, these features cannot help in distinguishing between patients with neurodegenerative diseases and controls. Some new studies indicate that the phonatory features like noise can be used to differentiate between the two but are more suitable for automatic detection or assessment approaches rather than being employed as single biomarkers. Regarding the articulatory approaches, VSA features family has also shown useful differentiation between patients and controls as it basically characterizes a reduction of the articulatory movements’ extension while producing vowels in connected speech. But we have to keep in mind that all these patients are trying to speak in English.

This study [5] uses two classifier architectures: traditional pipeline approach and end-to-end approach. In the pipeline approach, they use glottal features to extract the articulatory and phonatory features. A similar approach was used in the paper [1]. Here [5], they use the iterative adaptive inverse filtering (IAIF) and quasi-closed phase (QCP) glottal inverse filtering methods. The end-to-end approach uses deep learning models which were trained using both raw speech waveforms and raw voice source waveforms. They used the PC-GITA speech database. Apart from the speech waveform, the usage of the glottal source signal as a raw waveform in end-to-end systems is justified because the voice pathologies affect the vocal folds. The accuracy of the SVM-based pipeline method was improved from 65% to 67% when the baseline features were merged. They extracted the baseline features using the NeuroSpeech toolkit. They also concluded that their model could be extended to predict the neurological state of PD patients as well.

The author of this paper [2] studied the performance of neurological disease detection systems. The performance of such a system degrades when there is an acoustic mismatch between training and operating conditions caused mainly by degradation in test signals. They try to regulate the quality of the voice signals(recordings) using two different approaches by identifying the presence and type of short-term and long-term degradations and protocol violations. They use the mPower Mobile PD data set under different degradation conditions to test their approaches. Basically, they want to increase/control the quality of the voice signals in order to optimize the vocal characteristics and increase the performance of these detection systems.