Introduction

Dementia is a neurodegenerative disease that results in a gradual decline in cognition and memory. Recent reports on dementia suggest 23 in 100 people aged 60-84 currently possess the disease, whereby the prevalence is expected to double every five years. Dementia is known to affect cognition and behavior in the elderly and is comprised of both behavioral and language impairments. These language impairments heavily impact communication, with many individuals experiencing declines in expressive and receptive language abilities. However, it is important to note that individuals with dementia often have difficulties with the transmission of words in written and spoken communication rather than a deficit in word knowledge and word meanings. For individuals with Mild Cognitive Impairment (MCI), a syndrome that lies between normal cognitive aging and dementia, there is a high risk of developing dementia in the following months and years. Language variants between those with MCI and dementia have not been heavily studied computationally. For this reason, there is a critical need to develop algorithms and tools that are able to assess speech in ways that would lead to earlier diagnoses and treatment. The present study aims to elucidate the psycholinguistic properties of dementia and MCI patients, with particular focus on speech complexity through natural language processing techniques.

Methods

DementiaBank’s Cookie Theft Test to determine lexical variance

Metric of focus: type to token ratio (unigram to 11-gram models)

In cases where the type to token is smaller, it is predicted that individuals will experience a token-type ratio of 5:6 (where a normally aging individual may experience a token-type ratio of 6:10, whereby a normally aging individual may experience a token-type ratio of 6:10, whereby a normally aging individual may experience a token-type ratio of 6:10, whereby a normally aging individual may experience a token-type ratio of 6:10, whereby a normally aging individual may experience a token-type ratio of 6:10, whereby a normally aging individual may experience a token-type ratio of 6:10, whereby a normally aging individual may experience a token-type ratio of 6:10, whereby a normally aging individual may experience a token-type ratio of 6:10, whereby a normally aging individual may experience a token-type ratio of 6:10).

1. Transcription: Participants were asked to describe the Cookie Theft picture: descriptions were transcribed into Dementia Bank’s Transcript Database

2. Pre-parsing: Transcriptions were parsed and removed of extraneous data, including part of speech tagging of words, emotion tagging (“laughs”, “coughs”, etc.), and removal of all investigator questions and responses.

3. Main Algorithm: N-grams of size 1-11 were created and analyzed for uniqueness.

4. Type:token ratios were calculated and outputted.

Acknowledgements

I would like to thank Dr. Serguei Pakhomov Dr. Graeme Hirst for their undying support and guidance.