Effects of Cognitive Exercises on Speaking Characteristics of Dementia Patients

by

Seyedeh Ghazaleh Zakipour

A thesis submitted to the Faculty of Graduate Studies of

the University of Manitoba

In partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Graduate Program of Biomedical Engineering

University of Manitoba

Winnipeg

Copyright © 2022 by Seyedeh Ghazaleh Zakipour

Table of Contents

Abstract ..... *ix*

Keywords: ix

Chapter 1: Introduction *1*

1.1 Objectives 2

1.2 Thesis Organization 3

Chapter 2: Background *4*

2.1 Brain Anatomy 4

2.2 Dementia 5

2.2.1 Alzheimer's Disease (AD) 5

2.2.2 Mild Cognitive Impairment (MCI) 6

2.2.3 Frontotemporal Dementia (FTD) 7

2.3 Speech Analysis in Dementia Population 8

2.4 Cognitive Exercises for Neurorehabilitation in Older Adults 23

2.5 Cognitive Exercises in Dementia Populations 26

2.6 Summary 28

Chapter 3: Methodology *30*

3.1 Study Design 30

3.2 Participants 32

3.3 Data Analysis and Feature Extraction 33

3.4 Feature selection from Pitt Corpus 36

3.5 Classifier Training with Pitt Corpus 37

3.6 Classifier Testing with Exr Group 37

3.7 WMS-IV Correlation Analysis for Exr Group 37

3.8 Feature Categorization Using Pitt Corpus 38

3.9 Post-Intervention and Follow-Up Analyses for Exr Group 38

3.10 Summary 38

Chapter 4: Results *40*

4.1 Feature Selection and Categorization Using Pitt Corpus 40

4.2 Classifier Training Using Pitt Corpus 44

4.3 Classifier Testing Using Exr Group 45

4.4 WMS-IV Correlation Analysis for Exr Group 46

4.5 Post-Intervention and Follow-Up Analyses for Exr Group 48

4.6 Summary 53

Chapter 5: Discussion *54*

Chapter 6: Future Work and Conclusion *58*

References: ..... *LIX*

Appendix A: ..... *LXXVI*

List of Tables

Table 1-Brain lobes and their functions 4
Table 2-Summary of the dementia speech analysis literature review 11
Table 3-Story scripts in the logical memory sections of WMS-IV assessment 32
Table 4-Participants' demographic 33
Table 5-Summary of analyzed linguistic features 34
Table 6-Summary of analyzed acoustic features 34
Table 7- Acoustic features with significant differences between AD and control participants 40
Table 8- Linguistic features with significant differences between AD and control participants 43
Table 9-Classification results for machine learning algorithms 45
Table 10-Feature correlations to the change of LM I score of the WMS-IV assessment from baseline to post-intervention (extracted from LM I section) 46
Table 11-Feature correlations to change of WMS-IV total score of the WMS-IV assessment from baseline to post-intervention (extracted from LM I section) 46
Table 12- Feature correlations to the change of LM II score of the WMS-IV assessment from baseline to post-intervention (extracted from LM II section) 47
Table 13- Feature correlations to change of WMS-IV total score of the WMS-IV assessment from baseline to post-intervention (extracted from LM II section) 47
Table 14-Changes in mean values of LM I, LM II, and WMS-IV total scores 48

List of Figures

Figure 1-Brain lobes 5
Figure 2-Study procedure flowchart 39
Figure 3-LM I score over time 49
Figure 4-LM II score over time 49
Figure 5-WMS-IV total score over time 49
Figure 6-Mean of acoustic features during LM I (group H) 51
Figure 7-Mean of acoustic features during LM I (group D) 51
Figure 8-Mean of linguistic features during LM I (group H) 51
Figure 9-Mean of linguistic features during LM I (group D) 51
Figure 10-Mean of acoustic features during LM II (group H) 52
Figure 11-Mean of acoustic features during LM II (group D) 52
Figure 12-Mean of linguistic features during LM II (group H) 52
Figure 13-Mean of linguistic features during LM II (group D) 52
Figure 14-The Cookie Theft picture LXXVI

Acknowledgments

First, I would like to thank my advisor, Dr. Zahra Moussavi, for her guidance and for giving me the opportunity to work with her. Getting accepted into her team and coming to Canada marked a new, exciting era in my life, which I will always appreciate. I believe working with her has helped me gain a lot of new technical knowledge, and she has given me great advice on life, which I will value forever.

I would like to thank Dr. Ahmed Ashraf and Dr. Ji Hyun Ko for accepting to be my committee members. Their suggestions during my thesis proposal guided me in the right way and helped me complete my project.

Also, I would like to thank Dr. Natasha Jacobson for helping me edit this thesis and providing me with valuable suggestions on conducting it.

And lastly, I would like to thank all my friends and team members for supporting and motivating me. Because of them, I will never forget my time in Winnipeg.

List of Abbreviations

AD	Alzheimer's Disease
FTD	Frontotemporal Dementia
MCI	Mild Cognitive Impairment
PET	Positron-Emission tomography
MRI	Magnetic Resonance Imaging
$APOE_4$	apolipoprotein E4
WMS-III, -IV	Wechsler Memory Scale (III or IV)
LM I	Logical Memory I
LM II	Logical Memory II
MMSE	Mini Mental State Examination
f0	Fundamental Frequency
HNR	Harmonic to Noise Ratio
MFCC	Mel-frequency Cepstral Coefficient
MoCA	Montreal Cognitive Assessment
CoV	Coefficient of Variation
std	Standard Deviation
SVM	Support Vector Machine
RF	Random Forest
DT	Decision Tree
LR	Logistic Regression
KNN	K Nearest Neighbor
LDA	Linear Discriminant Analysis
ANOVA	Analysis of Variance
AUC-ROC	Area Under the Receiver Operating Characteristic Curve

Dedication

To my father who always believed in me and supported me even when I did not believe in myself.

To my mother who always encouraged me to chase my dreams and not give up.

To my grandparents who helped me become a better person.

And to my best friends who always listened to me when no one else would.

Abstract

Dementia is primarily associated with memory problems in older adults, though it can also affect speaking abilities. Speech aphasia in affected patients is characterized by long pauses, unfinished sentences, and word-finding difficulties. Recently, cognitive exercises have been shown to enhance cognition in older adults and adults with dementia; though not all dementia subjects suffer from speech aphasia. Thus, it was the effort of the present research to evaluate whether cognitive exercises similarly improve speech in the dementia population in general. In this pilot study, 8 participants (4 females, age 77.13 ± 8.32) previously diagnosed with dementia were recruited from an ongoing neuro-cognitive rehabilitation clinical trial (ClinicalTrials.gov NCT03587012). The primary outcome measure was the Wechsler Memory Scale IV (WMS-IV) assessment. In two sections of the assessment, participants must recall two short stories. These spoken recalls were stored in audio files for speech analysis. As such, acoustic and linguistic features were extracted from the audio files and their transcriptions. Due to the limited sample size of the pilot study, an online dataset (DementiaBank (Pitt corpus): 166 Alzheimer's disease, 64 healthy controls) was utilized to evaluate the extracted acoustic and linguistic features for statistically significant differences between participants with probable Alzheimer's disease and age-matched healthy controls. In turn, significant acoustic and linguistic features were analyzed in the pilot study dataset. The results of the analysis of the collected data indicated that the linguistic abilities of dementia patients improved from baseline to post-intervention after the 4-week period of memory exercises. Moreover, linguistic features revealed more interpretable distinctions between controls and dementia patients using the online databank. In addition, the acoustic features: fundamental frequency, noise to harmonic ratio, Mel-frequency cepstral coefficients, voiced duration, harmonic differences, bandwidth of the first formant, shimmer, spectral entropy, fractal dimensions, and number of zero-crossings showed high correlations with WMS-IV scores. This pilot study indicates that cognitive trainings have benefits for the speaking abilities of dementia patients, but to confirm the results, a much larger dataset is required.

Keywords:

Dementia, Alzheimer's Disease, Cognitive Exercise, Speech Analysis, Machine Learning

Chapter 1: Introduction

Dementia is a general term for problems with memory, speech, decision making, problem-solving, and other cognitive abilities that negatively influence a person's daily life [1]. The World Health Organization has reported that worldwide, more than 55 million people endure dementia, and around 10 million new cases are diagnosed annually [2]. Without proper intervention, the global number of dementia cases is estimated to reach to 78 million by 2030 and 139 million no later than 2050 [2].

Age is a significant risk factor for dementia with young-onset dementia (diagnosis prior to age 65) comprising only 3% of all cases [3]. Other potential risk factors for dementia include reduced cognitive activity, low education levels, social isolation, depression, and insufficient sleep [1]. However, risk of dementia may be reduced with a physically and socially active lifestyle, refraining from smoking, avoiding excessive alcohol use, and maintaining a healthy diet [3].

Alzheimer's disease (AD) is the most widespread kind of dementia [1]. Typically, the first noticeable sign of AD is problems with episodic memory, such as forgetting a recent event [4]. AD comprises 60-80% of the total dementia cases and is the 7th leading cause of death, globally [2], [3]. In Canada alone, over 500,000 adults have dementia, and without appropriate prevention, this number may rise to 912,000 by 2030 [5]. The emotional and financial effects of AD can be stressful and overwhelming for patients and their families. Economically, dementia accounted for $1.3 trillion in global health spending in 2019 and is expected to increase to $2.8 trillion by 2030 [2].

Other symptoms of AD include word-finding difficulties, struggling to navigate a familiar neighborhood, and more complex behavioral changes [4]. In general, 10 warning signs of dementia include memory problems, difficulties in conducting daily tasks, issues with language and speech, lack of awareness of time and place, difficulties in judgment, inability to keep track of events, misplacing items, mood and behavioral changes, difficulties in understanding visual and spatial cues, and withdrawal from social activities and work [3]. Usually, these symptoms are first noticed by patients' family members and/or close friends. Of interest in the present thesis is speech deficits as a part of AD development.

Studies have shown that some speech factors are significantly different between AD patients

and healthy controls [6]–[23]. Difficulty in speaking and understanding speech (aphasia) is detectable at early stages of the disease [24]. Speech in AD is characterized by longer hesitations with shorter speech segments while speaking [22]. Hesitations can include unfilled pauses, filled pauses (e.g., “ah”, “mhm”, “um”), repetition, laughter, false starts, vowel-lengthening, and throat-clearing [22]. Also, AD patients tend to speak more slowly and spend more time for finding the correct word [22]. Furthermore, AD patients have difficulties in recognizing and naming objects (anomia) [13], [25]. As AD progresses, the aphasia worsens, and the vocabulary of patients becomes very limited, to the extent that they cannot finish their sentences [13]. It should be noted that similar language difficulties and decay happen in other forms of dementia as well, including mild cognitive impairment (MCI) [26] and frontotemporal dementia (FTD) [27], [28].

Even in normal aging, cognitive abilities, such as memory, conceptual reasoning, and speed of processing decline to some extent [29]. However, the rate of decline increases in dementia patients [30]. Given that neuroscience research has proven that neuroplasticity is still possible even at old ages [31], cognitive exercises have been suggested to delay the onset of cognitive decline and/or to improve cognitive functioning in older adults [32].

1.1 Objectives

Given the prevalence of dementia on a global scale, and limited understanding of speech as a symptom of the disease, it was the objective of this thesis to **investigate the effects of cognitive exercises on speaking characteristics of dementia patients**. To succeed in this global objective, a series of sub-objectives were defined, as follows:

1. Collate existing acoustic and linguistic features from literature and extract said features from collected audio files.
    a) With oral and transcribed participant responses, acoustic and linguistic features may be evaluated for further analysis.
2. Detect significantly different acoustic and linguistic parameters between participants with probable AD and healthy controls.
    a) With data acquired from the Pitt corpus [33] (from DementiaBank databank), acoustic and linguistic parameters are compared between cognitively healthy and unhealthy participants to identify significantly different features for future diagnostic consideration.

3. Exploit significantly different acoustic and linguistic parameters to predict cognitive diagnosis.
    a) Train machine learning methods on the Pitt corpus [33] data to investigate if tested algorithms can correctly label study participants as cognitively healthy or unhealthy.
4. Identify correlations between acoustic and linguistic parameters with standardized memory assessments.
    a) Using previously collected acoustic and linguistic features, speech may be directly compared to participant Wechsler Memory Scale-IV (WMS-IV) assessments [34].
5. Analyze treatment methods for speech changes in dementia patients.
    a) Explore changes in the linguistic and acoustic features of study participants pre- and post-intervention to determine the effects of treatment on speech.

1.2 Thesis Organization

This thesis contains six chapters. The current chapter provided a brief explanation of dementia and the present study's objectives. Chapter 2 expands on this foundation to further understanding of two types of dementia and its precursor. Moreover, Chapter 2 considers speech problems in dementia patients and speech features that are different between dementia patients and healthy controls. Furthermore, the effects of cognitive training on older adults and older adults with dementia are elaborated upon. Chapter 3 elaborates on the study design, the selected features, and the analysis methods. The results of this study are demonstrated in Chapter 4 and discussed in Chapter 5. Finally, Chapter 6 completes the work with recommendations for future work.

Chapter 2: Background

2.1 Brain Anatomy

According to Johns Hopkins medicine, the brain consists of three main parts [35]: cerebrum, cerebellum, and brain stem. The cerebrum has four main lobes. Each lobe and its main attributed functions are presented in Table 1. In addition, Figure 1 displays the locations of the lobes.

Table 1-Brain lobes and their functions

Lobe	**Function(s)**
Frontal	Executive function [36], [37] Behavior and personality characteristics [38], [39] Motor control [39] Cognition [39] Attention [39] Working memory [39], [40] Long-term memory [39] Mood [39] Speaking ability [39] Awareness of other's feelings and mental states (theory of mind) [39], [41] Smell recognition [42]
Parietal	Motor control [43] Space perception [44] Speech [45] Writing [46] Interpretation of somatosensory information [47]
Temporal	Episodic memory [48] Speech [49], [50] Smell recognition [42]
Occipital	Vision [51]

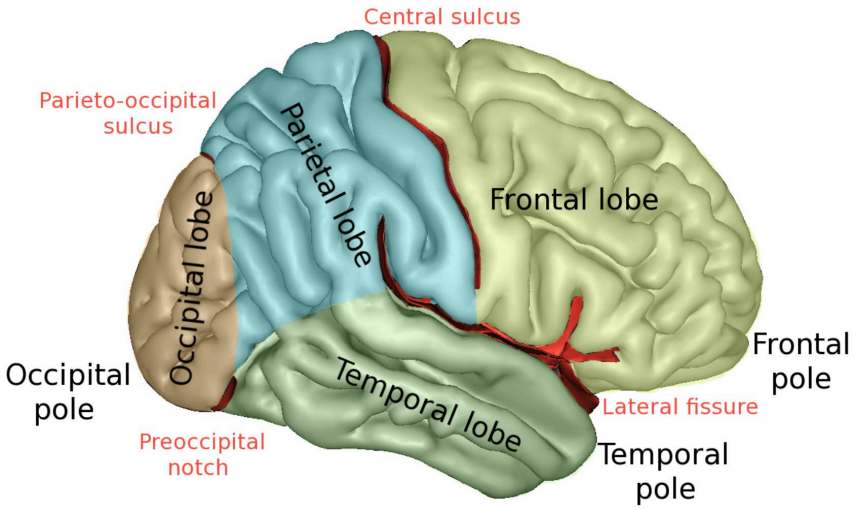


Figure 1-Brain lobes

(figure from Creative Commons [52])

2.2 Dementia

Dementia can be broken into 5 leading types: AD, vascular dementia, dementia with Lewy bodies, FTD, and mixed dementia [3]. In addition, MCI is usually considered as the stage before neurodegenerative dementia [53]. Since the participants of this study were diagnosed with AD, MCI, or FTD, these types of dementia are explored further in the following paragraphs.

2.2.1 Alzheimer's Disease (AD)

AD is commonly identified by the deterioration and death of neurons [4]. This neuronal damage has been attributed to aggregations of amyloid-beta (Aß) peptides and tau protein [54]. However, these biomarkers may manifest years before symptoms arise [4]. Aß peptides form plaques outside of neurons, while tau proteins shape tangles inside the neurons [1]. Plaques can disrupt communication among neurons, resulting in the deterioration and death of the nerve cells [1]. In addition, tau tangles block the transport of nutrients and other essential chemicals that neurons need to survive [1]. As a result, communication between neurons that support memory, learning, and other cognitive functions, is lost [4].

Currently, the routine practice for diagnosing AD is a combination of blood and cognitive testing, cerebrospinal fluid analysis, and neuroimaging methods (positron emission tomography (PET) scans and magnetic resonance imaging (MRI)) [3]. Blood testing is performed to eliminate other diseases that can cause cognitive deterioration/change [3]. In addition, cognitive tests are utilized to assess patients' cognitive levels. Cognitive tests measure cognitive abilities such as

memory, attention, visuospatial, language, and executive function and include evaluations such as Mini Mental State Examination (MMSE) and Montreal Cognitive Assessment (MoCA) [3]. The existing cognitive tests may be insensitive to early cognitive and language deficits in AD or MCI [55], [56], because in early dementia, patients might conceal their language deficits by using fabricated words or simplified speech [55], [57]. Therefore, due to the invasive nature of physical tests and limitations of cognitive assessments, researchers are working to develop alternative techniques for accurate diagnosis of AD.

Genetics may also be a contributing factor for AD. For instance, the apolipoprotein E4 ($APOE_4$) gene can increase the risk of AD in younger patients [3]. However, though the $APOE_4$ gene exists in about 15% of people, it does not guarantee an AD diagnosis [3]. It should be noted that no existing medication can cure AD, and the effectiveness of these drugs to alleviate the symptoms of AD varies from one person to another [3]. A review study on donepezil [58], one of the most common medications for AD, concluded that donepezil elicited a positive response in 26% to 63% of patients.

As AD progresses, the brain encounters drastic changes. For instance, AD patients typically show signs of inflammation and atrophy in the brain [1]. Brain inflammation is the result of a disparity between immune system response and speed of detriment [1]. Brain atrophy occurs due to the loss of neurons [1] and may happen in various brain regions including sulcal spaces in the frontal and temporal cortices [59], frontal and temporal horns of the lateral ventricles [60], posterior cortical areas [61], and medial temporal regions such as the amygdala and the hippocampus [62], [63]. Thus, the overall volume of the brain decreases in AD patients [59]. The hippocampus plays a significant role in learning and memory; therefore, its volume loss may explain the memory problems associated with AD [64].

2.2.2 Mild Cognitive Impairment (MCI)

MCI patients have a level of cognitive decline that does not highly interfere with daily life but can be detectable while performing more complex tasks [54]. MCI has 2 main sub-types [65]: amnestic and non-amnestic. In the amnestic type, patients experience memory loss. Older adults may misplace items or have difficulty remembering names, but memory problems become more concerning when patients forget appointments or events that would normally interest them [65]. Other cognitive skills such as executive function, visuospatial skills, and language remain relatively intact. In the non-amnestic type, cognitive functions including visuospatial skills,

language, and attention decline, but memory problems are not reported. Some people with MCI remain stable or even return to normal cognitive function, while some others develop AD within 5 years [66].

Annually, 5-15% of MCI patients develop AD [67]. Studies have suggested that MCI patients who have the $APOE_4$ gene [68], hippocampal atrophy and larger ventricles [69], [70], hypometabolism in the temporal and parietal lobes (as indicated in fluorodeoxyglucose positron-emission tomography scans [71]), low levels of β-amyloid peptide 42 (Aβ42), high levels of tau protein in cerebrospinal fluid [72], or evident amyloid on PET scans [73] are at higher risk of progressing into AD. There is no medication to treat MCI, but some studies have suggested that patients who take medications that are antihypertensive, cholesterol-lowering, antioxidants, or anti-inflammatories, in combination with oestrogen therapy have a reduced risk of developing AD [74].

2.2.3 Frontotemporal Dementia (FTD)

FTD is the most prevalent type of frontotemporal lobar degeneration [27]. FTD is caused by atrophy and loss of neurons in the frontal and temporal lobes, anterior cingulate cortex, and insular cortex [28], [75]. PET scans have suggested that ventromedial frontal cortex is one of the earliest affected areas of the brain during the course of FTD [76]. In a review paper by Neary et al. [27], it was discussed that patients with FTD demonstrate behavioral, cognitive, and language changes. For instance, FTD patients showed signs of distraction, mental inflexibility, changes in eating habits, utilisation and repetitive behaviours, issues with planning, problem-solving difficulties, judgement errors, and repeated or cluttered speech. In addition, patients with FTD have difficulty with empathy [77], and may show absence of emotions and social emotions [27]. However, spatial skills, elementary visual perception, and memory are retained in FTD patients, while they are significantly impacted in AD patients [27].

FTD has 3 main variants [28]: 1) behavioral variant frontotemporal dementia, which causes changes in personality and behavior and executive function; 2) primary progressive aphasia, which leads to deterioration of speech, grammar, object naming, and word comprehension; and 3) semantic variant primary progressive aphasia, which negatively affects semantic knowledge, naming and word finding skills. FTD patients may also show physical symptoms of primitive reflexes, rigidity, and tremor [78]. Therefore, a supplementary physical exam is required to conclude a correct diagnosis.

Men and women are at equal risk of developing FTD [79]. FTD usually develops between the ages of 45-65 [79], [80], but has also been observed at younger and older ages [81], [82]. A family history of dementia increases the risk of developing FTD (in approximately 40% of cases) [83]. In almost all cases of FTD, accumulation of tau protein, TAR DNA-binding protein with molecular weight 43, or the fused-in-sarcoma protein are detectable [28]. No medication can cure FTD, though available treatments attempt to ameliorate symptoms [28].

2.3 Speech Analysis in Dementia Population

Recently, speech analysis has been proposed as an inexpensive and non-invasive option to classify between dementia patients and healthy controls [6]–[23], [84], [85]. Speech analysis can be divided into two subgroups: 1) acoustic analysis and 2) linguistic analysis. In acoustic analysis, features are extracted from voice signals with the help of voice analysis programs, such as Praat [86] or openSMILE [87]. In contrast, in linguistic analysis, features are extracted from the transcriptions of voice signals. Natural Language Toolkit [88] is a common software for exploiting the linguistic parameters. Transcribing patient interviews can be performed manually or with automatic voice recognition systems. Researchers usually develop their own automatic voice recognition systems using deep neural networks as available automatic systems are more error-prone to older voices [89]. To collect voice samples, participants are recorded while performing an oral task, such as reading a paragraph, telling a story, or describing a picture. The choice of the tasks differs among studies; however, it is essential to choose the task such that the extracted parameters are sensitive to acoustic and linguistic changes in dementia patients.

The main patterns of speech in AD can be categorized into 3 groups [25]: 1) dysfluency (broken speech), 2) dysprosody (related to fundamental frequency and intensity), and 3) agrammatism (can be studied using natural language processing techniques). Furthermore, speech in AD is characterized by deficits in semantic knowledge, single-word comprehension, articulatory and phonological levels [25]. In a study performed by Beltrami et al. [90], most of the syntactic features were significantly different between patients with dementia and healthy controls while talking about a working day and a recent dream. Another study [19] used Swedish and English datasets collected while participants were describing the Cookie Theft picture, a standardized assessment tool from the Boston Diagnostic Aphasia Examination [91]. The Cookie Theft picture (Appendix A:) shows a mother and her two children in a kitchen and must be orally described by the participant. The more actions a person can describe for the picture, the greater

their score becomes. During said study, it was found that MCI patients in both Swedish and English datasets regressed in identifying information units of the Cookie Theft picture within a given timeframe. However, no significant difference was identified in idea density, efficiency, or total speech duration between MCI and healthy participants, despite the slower speech pace of MCI patients. But overall, speech in MCI patients was closer to that in healthy control participants than AD patients. Moreover, MCI patients show impairments in language tests, such as category fluency, object knowledge, and object naming [92]. Lastly, speech in FTD may be characterized by aspontaneity or press of speech, mutism and preservation of speech, echolalia (repeating speech of others), and stereotypy of speech [78]. Also, one study demonstrated that language scores declined faster for FTD patients than AD patients [93]. Considering the language impairments observed in dementia patients, several linguistic and acoustic features are different among them and healthy older adults.

Linguistic features may consist of rates and ratios of parts of speech tags [6], [9], [19], [20], [94], number of words [20], [94], Honoré's statistic [6], [9], Brunet's index [6], [9], type token ratio [6], [9], [94], verbal rate [55], transformed phonation rate [55], and rate of grammatical errors [95]. Bucks et al. [6] argued that AD participants had higher means of pronoun rate, adjective rate, and verb rate, but lower noun rate (which showed their anomia). Also, Honoré's statistic, Brunet's index, and type token ratio showed richer speech in most of the healthy controls. In accordance, Khodabakhsh et al. [11] indicated that pronoun frequency and pronoun-to-noun ratio were significantly different between AD and control groups regardless of education level. In addition, AD patients used less nouns than pronouns. In another study [7], AD and healthy participants significantly differed in semantic and phonological verbal fluency, object naming, the ability to remember items of shopping and word lists, and description of a person. These tests were part of Cuetos-Vega test, a cognitive test introduced in a 2007 study [96]. Besides, Beltrami et al. [90] observed that some parts of speech rates and lexical richness measures were statistically different between dementia groups and healthy controls. Furthermore, Cho et al. [94] noticed that patients with frontotemporal lobar degeneration spoke less words, had lower lexical diversity, and used more nouns than the control group. Additionally, the patient and control groups were significantly different in number of unique nouns, pronouns, wh-words, tense-influenced verbs, verbs, speech errors, adverbs, total words, adjectives, and prepositions. Given the support of previous works, it

is proposed that linguistic features are different between dementia patients and age-matched healthy controls.

Acoustic features may comprise of duration [97]–[99], fundamental frequency (F0) [14], [17], [22], [23], [84], [97], [98], [100], [101], formants [17], [22], [23], jitter [14], [22], [97], [98], [101], shimmer [8], [14], [22], [97], [98], [101], harmonic to noise ratio (HNR) [22], [97], [101], Mel-frequency cepstral coefficients (MFCCs) [9], [22], [101], fractal dimensions [22], energy [13], [14], [22], zero crossing rate [22], and features related to pause [10], [20], [55], [97], [99], and voiced/unvoiced segments [8], [10], [12]–[14], [22], [97]. It has been observed that AD patients have deeper voices, slower speech rate, and a slower rate of glottal pulses [8]. They also had high voice breaks and shimmer, implying that their voices had more tremble, less intensity, and less airflow control [8]. Additionally, AD patients had more unvoiced segments and less voiced segments in their speech [14], leading to longer pause durations [55]. Moreover, by using a regression model, Martinez-Sanchez et al. [7] showed that speech rate and articulation rate with age could account for 45.5% of the variance in MMSE scores and 70.05% of the variance in Cuetos-Vega test scores. Also, one study conducted by Meilan et al. [97] investigated acoustic features in AD speech and concluded that the increase in the percentage of voiceless sections described more than 34% of the variance in the language and memory test scores used. Hoffman et al. [95] added to this by stating that hesitation ratio was significantly higher for AD patients while speech tempo was considerably higher for control participants. Besides, features that describe the non-linearity of voice have shown to be useful for AD/control classification. For instance, one study utilized fractal dimension in their feature set [102]. Fractal dimension can explain the complexity and dynamics of a system and detect small changes [102]. Adding fractal dimension to the feature set improved the classification accuracy [102] and decreased the classification error rate [13]. It can be concluded that acoustic features are distinguishable among dementia patients and healthy controls.

A summary of the reviewed papers on speech analysis in dementia is shown in Table 2. It should be noted that the “Extracted Speech Features” column shows the features that yielded the highest accuracy, though a more extensive set may exist in the source manuscript. Researchers have used the features described above or features analogous to them for classifying AD/MCI patients from healthy controls. Researchers have employed machine learning or deep learning methods, such as linear discriminant analysis (LDA), K nearest neighbors (KNN), decision trees

(DT), random forests (RF), logistic regression (LR), support vector machines (SVM), and deep neural networks for classification. The achieved accuracies varied from 75% to 96% [6], [8]–[11], [13]–[15], [17], [19]–[21], [85], [103], while the sample size in most studies was limited ($10 < n < 270$). Only one study included more than 7000 participants [23]. Nevertheless, the extracted acoustic and linguistic parameters were able to classify dementia patients and healthy controls with reasonable accuracies.

Table 2-Summary of the dementia speech analysis literature review

Study	**Sample Diagnosis and Male/Female Ratio**	**Dataset Language(s)**	**Task**	**Extracted Speech Features**	**Classification Algorithm**	**Outcome**
Bucks et al., 2000 [6]	8 AD (M: 4, F: 4) 16 HC (M: 7, F: 9)	English	Semi-structured interviews with open-ended questions: participants talked about themselves and their experiences.	Noun rate Pronoun rate Brunet's index	LDA	ACC: 87.5%
Singh et al., 2001 [55]	8 AD (M: 4, F: 4) 8 HC (M: 2, F: 6)	English	Semi-structured interviews: participants talked about their lives and their experiences.	Verbal rate Transformed phonation rate Mean duration of pauses Standardized phonation time Standardized pause rate	-	-
Hoffman et al., 2010 [95]	30 AD (M: 9, F: 21) 15 HC (M: 7, F: 8)	Hungarian	Semi-structured interviews: participants answered why they were at the clinic, remembered some important events of their lives, and described a daily hobby or activity.	Articulation rate Speech tempo Hesitation ratio Rate of grammatical errors	-	-
Pakhomov et al., 2010 [93]	19 bvFTD 12 progressive non-fluent aphasia	English	Cookie Theft [91] picture description	Out-of-vocabulary rate: the rate of words that were only used	-	-

Study	Sample Diagnosis and Male/Female Ratio	Dataset Language(s)	Task	Extracted Speech Features	Classification Algorithm	Outcome
	6 progressive logopenic aphasia 11 semantic dementia 23 younger adults and 9 older adults as HCs Overall, for patients: M: 25, F: 23			by the patients and not the healthy adults. Perplexity index: degree of the difference of word patterns used by the patients and healthy participants.		
Pakhomov et al., 2010 [99]	38 patients diagnosed with bvFTD, progressive non-fluent aphasia, and semantic dementia.	English	Cookie Theft [91] picture description	Pause-to-word ratio Variance of F0 Pronoun to noun ratio Word count Word-level and part-of-speech perplexity Total duration of speech Mean prosodic phrase length Information units Normalized filled, silent, and long (> 400 ms) pause counts Normalized repetition count Normalized false start count Normalized dysfluent event count	-	-

Study	Sample Diagnosis and Male/Female Ratio	Dataset Language(s)	Task	Extracted Speech Features	Classification Algorithm	Outcome
Horley et al., 2010 [100]	20 AD 20 HC	English	Two production tasks (repeating and reading sentences and express emotions) and two perceptual tasks (identification of emotions and neutral content). Stimuli for each task was based on 2 positive emotions (happiness and surprise) and 2 negative emotions (anger and sadness).	Mean and std of F0 Rate of speech	-	-
Meilán et al., 2012 [97]	21 AD	Spanish	Participants read a series of sentences from the screen and filled out some missing information about themselves.	Duration F0 Jitter Shimmer HNR NHR Percentage of pauses Voiceless segments Voice interruptions Phonation	-	-
Martínez-Sánchez et al., 2013 [7]	35 AD (M: 27, F: 8) 35 HC (M: 28, F: 7)	Spanish	Reading a paragraph	Speech rate	ROC Curve	Area under the curve: 80%
Meilán et al., 2014 [8]	30 AD (M: 10, F: 20) 36 HC (M: 7, F: 29)	Spanish	Reading a paragraph	Percentage of voice breaks Number of periods of voice Number of voice breaks Shimmer NHR	LDA	ACC: 84.8%

Study	Sample Diagnosis and Male/Female Ratio	Dataset Language(s)	Task	Extracted Speech Features	Classification Algorithm	Outcome
Rentoumi et al., 2014 [103]	18 Pure AD (M: 12, F: 6) 18 Mixed AD (M: 6, F: 12)	English	Cookie Theft [91] picture description	Frequency unigrams	Naïve Bayes	ACC: 75%
Fraser et al., 2015 [9]	DementiaBank [33] 167 AD 97 HC	English	Cookie Theft [91] picture description	Part of speech tags Syntactic complexity: Mean length of T-units, clauses, and sentences. Grammatical constituents: Counting the number of noun phrases. Psycholinguistics: Measuring the frequency of highly familiar words. Vocabulary richness: Type-token ratio, Brunet's index, Honoré's statistics. Information content: Measuring the number of lexical items that indicate information units in the picture. Repetitiveness: Using a bag-of-words and cosine distance between sentences. Acoustic measures: MFCC and its statistics.	LR	ACC: 81.92%
Khodabakhsh et al., 2015 [11]	28 AD (M: 18, F: 10) 51 HC (M: 31, F: 20)	Turkish	Unstructured interviews	Silence ratio with and without phoneme rate	SVM	ACC: 83.5%

Study	Sample Diagnosis and Male/Female Ratio	Dataset Language(s)	Task	Extracted Speech Features	Classification Algorithm	Outcome
Khodabakhsh et al., 2015 [10]	27 AD (M: 17, F: 10) 27 HC (M: 12, F: 15)	Turkish	Conversations	Logarithm of the response time Continuous speech and pause duration Pauses per second Silence/speech ratio Voicing ratio Voicing per second Variance of phoneme frequency Average absolute delta energy Delta pitch Delta formants Phonemes per second	SVM	ACC: 96.3%
König et al., 2015 [12]	26 AD (M:13, F: 13) 23 MCI (M: 11, F: 12) 15 HC (M: 6, F: 9)	French	Count down Picture description Sentence repeating Categorical verbal fluency	Voiced/unvoiced segments Periodic/aperiodic segments Vocal reaction time Number of deletions, insertions, and the distances in time of second to ninth detected word position from the first word position	SVM	Equal Error Rate: AD-HC: 13% MCI-HC: 21% AD-MCI: 20%
Lopez-de-Ipiña et al., 2015 [14]	AZTITXIKI 5 AD 5 HC	English, French, Spanish, Catalan, Basque, Chinese, Arabic	Telling pleasant stories Remembering pleasant feelings	Voiced/unvoiced segments Short-time energy Spectral centroid Emotional features: F0, intensity, jitter, and shimmer, as well as emotional temperature	SVM	ACC: 93.79%
Lopez-de-Ipiña et al., 2015 [13]	AZTIAHORE 20 AD (M: 8, F: 12) 20 HC (M:10, F: 10)	English, French, Spanish, Catalan, Basque,	Spontaneous speech	Voiced/unvoiced segments Short-time energy Spectral centroid	Multi-Layer Perceptron	ACC: 92.43%

Study	Sample Diagnosis and Male/Female Ratio	Dataset Language(s)	Task	Extracted Speech Features	Classification Algorithm	Outcome
		Chinese, Arabic		Automatic selection of max, min, variance, std, median and mode average for the whole signal and its voiced parts		
Nasrolahzadeh et al., 2016 [104]	30 AD (M: 16, F: 15) 30 HC (M: 15, F: 15)	Persian	Participants were asked to tell stories, express their feelings, and have friendly conversation.	Bi-spectrum Bi-coherence estimation	-	-
Martinez de Lizarduy et al., 2017 [15]	Categorical Verbal Fluency [105]: 38 MCI, 187 HC Spanish Picture Description [105]: 6 AD, 12 HC Spanish Spontaneous Speech: AZTIAHORE 20 AD (M: 8, F: 12) 20 HC (M: 10, F: 10)	English, French, Spanish, Catalan, Basque, Chinese, Arabic	Categorical verbal fluency Picture description Spontaneous speech	Classical features Perceptual features Advanced features Non-linear features	CNN	ACC: Spontaneous speech: 95% Picture Description: 94% Categorical verbal fluency: 82%
Martínez-Sánchez et al., 2017 [16]	45 AD (M: 16, F: 29) 82 HC (M: 30, F: 52)	Spanish	Reading a paragraph	std of the duration of syllabic intervals	ROC Curve	AUC: 87%
Beltrami et al., 2018 [90]	16 Amnestic MCI (M: 8, F: 8) 16 Multiple Domain MCI (M: 8, F: 8) 16 Early Dementia (M: 8, F: 8)	Italian	Picture description Describing a typical workday Remembering latest dream	Acoustic features Rhythmic features Lexical features Syntactic features	-	-

Study	Sample Diagnosis and Male/Female Ratio	Dataset Language(s)	Task	Extracted Speech Features	Classification Algorithm	Outcome
	44 HC (M: 24, F: 24) I					
Themistocleous et al., 2018 [17]	Guthenburg dataset 25 MCI (M: 12, F: 13) 30 HC (M: 11, F: 19)	Swedish	Reading a paragraph	The first five formants at 15%, 25%, 75% of vowel's duration Logarithmic values of the first three formants Vowel duration F0 (min, max, mean)	Deep neural network	ACC: 83%
Al-Hameed et al., 2019 [101]	15 Functional Memory Disorder (M: 6, F: 9) 15 Neurodegenerative Disorder (M: 7, F: 8)	English	Conversations	9 selected features from acoustic parameters. Speech and silent features F0 HNR NHR Shimmer Jitter Number and degree of voice breaks MFCC Spectral centroid Filter bank energy coefficient	SVM	ACC: 97%
Chien et al., 2019 [18]	Mandarian_Lu dataset and NTU dataset 30 AD, 30 HC for fluency tasks 15 AD, 15 HC for picture description task 10 AD (M: 6, F: 4), 10 HC (M: 4, F: 6) (NTUH dataset)	Mandarin Taiwanese	Fruit fluency Location fluency Picture description	Sequence of tokens. Each token represented one element of the speech.	GRU	AUC: 83.8%

Study	Sample Diagnosis and Male/Female Ratio	Dataset Language(s)	Task	Extracted Speech Features	Classification Algorithm	Outcome
Fraser et al., 2019 [19]	GothenBurg [106]: 31 MCI 26 HC Karolinska [107]: 96 HC DementiaBank [33] : 19 MCI-19 HC 166 AD-97 HC	Spoken and written Swedish Spoken English	Cookie Theft [91] picture description	1-Cluster features (such as the average cosine distance between the centroid and all the words within a cluster), 2-34 features selected from cluster features, summary features (such as the number of clusters and the number of words assigned to each cluster), baseline features (such as the total number of nouns and verbs divided by the total number of words). 3-73 features selected from all feature sets.	SVM	ACC: 1-AD-HC in DementiaBank: 82% 2-MCI-HC (English): 63% 3-MCI-HC (Swedish): 72%
Gosztolya et al., 2019 [20]	25 Mild AD 25 MCI 25 HC	Hungarian	Immediate and delayed recall of a short movie Talking about the previous day	Articulation rate Speech tempo Length of utterance Duration and number of silent and filled pauses (hesitation) Hesitation rate Phoneme features	SVM	ACC: Mild AD-HC: 86% MCI-HC: 80% MCI-Mild AD: 80%

Study	Sample Diagnosis and Male/Female Ratio	Dataset Language(s)	Task	Extracted Speech Features	Classification Algorithm	Outcome
				Linguistic features: Number of words and punctuation marks, number and rate of part of speech tags, number of pauses, number of uncertain and unknown words, number of negation words and content words.		
Liu et al., 2020 [21]	VBSD 23 AD (M: 10, F: 13) 13 HC (M: 5, F: 8)	Chinese	Picture description immediately and from memory Repeating a set of sentences Repeating the pronunciation "pa-ka-ka"	Spectrogram features	LR	ACC: 86.1%
Lopez-de-Ipiña et al., 2020 [22]	PGA-OREKA [105] 38 MCI 62 HC	Spanish	Categorical verbal fluency	Top 25 features from classical features. Jitter F0 Shimmer NHR HNR Harmonicity Spectrum centroid Amplitude perturbation quotient Formants and their variants (min, max, mean, median, mode, std) Voice breaks Voiced/unvoiced segments Zero-crossing rate and its variants (min, max, mean, median, mode, std)	SVM	Classification Error Rate: 5%

Study	Sample Diagnosis and Male/Female Ratio	Dataset Language(s)	Task	Extracted Speech Features	Classification Algorithm	Outcome
				Energy, short-time energy, intensity, and spectrum centroid and their first- and second-order regression coefficients. Perceptual features including MFCC and its variants (min, max, mean, median, mode, std) and their first- and second-order regression coefficients. Advanced features including perceptual linear predictive coefficients, modulation spectra coefficients, inferior colliculus coefficients, adaptive component weighted coefficients, linear predictive cosine transform coefficients, linear predictive cepstral coefficients, and their variants (min, max, mean, median, mode, std).		

Study	Sample Diagnosis and Male/Female Ratio	Dataset Language(s)	Task	Extracted Speech Features	Classification Algorithm	Outcome
				Non-linear features including fractal dimension and its variants (min, max, mean, median, mode, std), Shannon entropy, multiscale permutation entropy and their variants (min, max, mean, median, mode, std).		
Meilán et al., 2020 [98]	73 Non-degenerative MCI (M: 20, F: 53) 13 Pre-clinical AD (M: 4, F: 9)	Spanish	Reading a paragraph	Duration Parameters related to speech fluency and rhythm F0 and spectral analysis Intensity Amplitude Jitter Shimmer	-	-
Nagumo et al., 2020 [23]	1601 MCI 367 GCI 368 MCI + GCI 6343 HC	Japanese	Vowel utterance Tongue twister Diadochokinesis Reading 2 short sentences 3 times	Mean and std of F0 and the first two formants, triangular vowel space area, vowel articulation index, utterance duration, number and length of pauses, number of uttered syllables, the mean and std of syllable duration, the mean and std of the max and min power, mean utterance time ratio of tongue twister and short sentences tasks.	LR	Area under the receiver operating characteristics curve: MCI-HC: 61% GCI-HC: 67% MCI+GCI-HC: 77%
Cho et al., 2021	74 bvFTD (M: 48, F: 26)	English	Cookie Theft [91] picture description	Counts of part of speech tags	-	-

Study	Sample Diagnosis and Male/Female Ratio	Dataset Language(s)	Task	Extracted Speech Features	Classification Algorithm	Outcome
[94]	22 non-fluent/agrammatic variant primary progressive aphasia (M: 11, F: 11) 42 semantic variant primary progressive aphasia (M: 19, F: 23) 37 HC (M: 13, F: 24)			Number of tense influenced verbs Number of unique nouns Number of wh-words Total number of words Rate of abstractness of nouns Semantic ambiguity Word frequency Word familiarity Type-token ratio Cross-entropy estimation Speech errors.		
Mahajan et al., 2021 [84]	ADReSS dataset [108] 82 AD 82 HC	English	Cookie Theft [91] picture description	1. Token-level psycholinguistic features, token-level sentiment features and demographic features. 2. Token-level psycholinguistic features, token-level sentiment features and demographic features. Acoustic features such as MFCC, line spectral pairs, and F0. Targeted features relating to lengths of speech segments, disfluency, interventional features, and idea density.	1. CNN-LSTM 2. CNN-LSTM and GRU	ACC: 1-72.92% 2-72.92%

Study	Sample Diagnosis and Male/Female Ratio	Dataset Language(s)	Task	Extracted Speech Features	Classification Algorithm	Outcome
Roshanzamir et al., 2021 [85]	DementiaBank [33] 170 AD 99 HC	English	Cookie Theft [91] picture description	Large bidirectional encoder representations from transformers (BERT) [109] embeddings.	LR	ACC: 88.08%

std: Standard Deviation
min: Minimum
max: Maximum
AD: Alzheimer's Disease
MCI: Mild Cognitive Impairment
GCI: Global Cognitive Impairment
HC: Healthy Control
NHR: Noise to Harmonic Ratio
HNR: Harmonic to Noise Ratio
F0: Fundamental Frequency
bvFTD: behavioral variant frontotemporal dementia
ACC: Accuracy
ROC: Receiver Operating Characteristic
LDA: Linear Discriminant Analysis
LR: Logistic Regression
SVM: Support Vector Machine
CNN: Convolutional Neural Network
GRU: Gated Recurrent Unit
LSTM: Long Short-Term Memory

2.4 Cognitive Exercises for Neurorehabilitation in Older Adults

Neuroplasticity is the ability of the brain to change its structure or function with training, experience, or external stimulus, and compensate for lost neurons [31]. One example of neuroplasticity is recovery in stroke patients. After many hours of therapy, new parts of the brain perform functions previously executed by the damaged parts of the brain [31]. Thus, the brain is forced to use available plasticity to restore normal brain function [31]. Inspired by this phenomenon is the theory that cognitive training for processing speed, memory, and reasoning may activate neuroplasticity of the brain [110].

There are two other theories related to aging and cognition as well: the cognitive reserve model [111] and the scaffolding theory of aging and cognition [112]. The cognitive reserve model indicates that certain experiences, including education and active lifestyle, increase the supply of neural resources [113], [114]. As a result, these resources offer protection against age-related cognitive decline [31]. The scaffolding theory of aging and cognition explains that compensatory scaffoldings develop in the brain due to age-related damages [31] such as decreased hippocampal activity [115], volumetric shrinkage in different brain regions [116], and excessive amyloid

distributions [117]. The scaffolding theory of aging and cognition also suggests that protection against cognitive decline can happen with proper cognitive training or engagement in a new task [31].

Generally, there are three groups of cognitive trainings [118]: 1) strategy-based training that leads to durable improvements in the training task but not so much to other tasks (near-effect); 2) multidomain training that targets several cognitive abilities but has limited transfer effects (far-effect); 3) process-based training that focuses on general processing abilities such as executive function, the speed of processing, and working memory. Process-based trainings, especially those focused on executive function and working memory, might be more effective when compared to other forms of training as demonstrated by small- to medium-sized transfer effects (even in older adults) [118]. The far-effect of a training intervention is assessed by how well the participants improve in dissimilar tasks, while its near-effect is evaluated by the participants' performance on a task that focuses on the same idea as the training [118].

The outcomes of cognitive exercise interventions on older adults have been controversial. Some researchers claim only near-effect benefits of training [118], [119], some claim both near- and far-effects [120], while other studies claim there is no benefit to such interventions [121]. Thus, the reported efficacy of cognitive training interventions differs between studies as well as among younger and older adults. Controversies may stem from variations in frequencies and durations of interventions. Alternatively, the mode of delivery has demonstrated importance, such as in-person, or self-administered [120]. For instance, in a review conducted by Karbach et al. on 13 working memory studies [118], the mean number of sessions, mean duration of each session (in hours), mean number of training hours, and mean time between pre- and post-intervention (in days) were 16.66 (std = 11.16), 0.67 (std = 0.25), 10.69 (std = 11.05), and 32.25 (std = 21.28) respectively. In addition, it must be considered that studies used protocols targeting different brain regions [121]. Consequently, it is essential to critically examine the purpose of a study and determine training frequency and target areas accordingly.

In one small, randomized trial with older adults [122], the authors hypothesized that cognitive training might result in significant and gradual cognitive improvements over the course of training. In this instance, cognitive training consisted of word-image matching and animal-shape matching. Wechsler Memory Scale-III (WMS-III) was administered as an outcome measure pre- and post-trial. The participants (13 active, 11 control) performed cognitive exercises for 30 minutes

per day, 3 days per week, for 8 weeks. Considering the improvements in the WMS-III scores and correct responses to the two tasks, the authors concluded that memory performance was enhanced post-intervention. In contrast, a meta-analysis review of 52 studies [123] derived that cognitive exercise for older adults were ineffective for executive function, working memory, and verbal memory, but effective for non-verbal memory, processing speed, and visuospatial outcomes. In another study [121], older adults were trained on 2 tasks (token search and dual n-back test) and tested on spatial span and digit span. Participants in the training group (n = 47) significantly improved in the training tasks, but their performance on the test tasks did not improve and were almost identical to the passive control group (n = 31). As a result, the authors inferred that brain trainings did not have a transfer effect, even on similar tasks [121]. Conversely, Schmiedek et al. [119] tested working memory exercises on 19 younger (mean age = 25.3 years) and 21 older (mean age = 74.5 years) adults in the active group, and 27 younger (mean age = 26.4 years) and 20 older (mean age = 73.3 years) adults in the control group. In post-intervention, after 45 days of 15-min sessions, considerable improvements were observed in the training tasks of both younger and older adults. The interventions resulted in positive near-effect transfers and the effect was similar among younger and older adults. At a 3-month follow-up, both groups maintained their practice gains and near-effect gains, though gains were larger in younger adults. Also, no far-effect transfer was detected. Thus, in general, near-effect transfers are more widely reported in older age groups, supporting their potential in future interventions.

Despite research disproving far-effect transfers of cognitive training in older populations [123], [121], [119], both near-effect and far-effect transfers have been observed in younger adults [124]–[126]. One such instance is that of Caeyenberghs et al. [126] who recruited participants between the ages of 19 and 40. Caeyenberghs's research group developed computerized training including spatial span tasks under several conditions which focused on working memory. The conclusion was that working memory and executive function tasks led to near- and far-effect transfers on untrained working memory tasks and tasks of reasoning and inhibition. In a similar study administered by Dodge et al. [127], 11 participants (young adults) trained for 25 minutes per day for 2 months. The training tasks consisted of a visuospatial working memory task, operation n-back test, and dual n-back test. Participant performance significantly improved on the trained tasks in the last three training sessions. In addition, the authors found that working memory training improved the connections of white matter in parietal lobes and the anterior section of corpus

callosum. Thus, findings from Dodge et al. suggest cognitive exercises may lead to improvements in brain structure [127].

The choice of control groups can also affect the results on the studies. After reviewing 87 publications, Melby-Lervag et al. [128] stated weak confirmation of near-effect transfers and no far-effect transfer in studies using treated control groups. In addition, the working memory trainings rarely transferred to daily cognitive abilities. The authors added that, without treated control groups, the training effects could be due to participant familiarity with the tasks, or additional assistance from tutors. Differences in the motivations of the studies should also be considered [128]. As a result, study design plays an important role in the final effect of trainings.

In all, despite the potential of neuroplasticity in the brain, near- and far-effects of cognitive training remain controversial. Efforts to improve far-effect transfer of cognitive training are essential for improvement in working memory and executive function, though, variation in study designs have yielded conflicting results in training viability. As such, it is recommended to standardize study protocols to consider age-related impacts, control group effects, and bias from study motivation to offer a more critical and robust analysis of cognitive training in older adults.

2.5 Cognitive Exercises in Dementia Populations

Individuals with cognitive impairments may benefit more from working memory training because they have more capacity for improvement [128]. In a review on the effects of cognitive exercises in people with dementia [129], it was concluded that people with mild to moderate dementia benefit from cognitive exercises. The benefits were evident from cognitive test scores (Alzheimer's Disease Assessment Scale-Cognitive Subscale and MMSE). However, behaviour, mood, and performance in daily activities were not significantly enhanced by cognitive exercises. That said, three of the reviewed studies reported that the positive effects on cognitive function lasted for 1 to 3 months post-intervention [130]–[132], with one study reporting no improvement at a 10 month follow-up [133]. Further, Hwang et al. [134] claimed that cognitive training can delay memory deterioration in AD patients. They indicated that the participants improved in listing words and recognizing and listing objects after 4 weeks of training. Also, participant scores significantly increased in orientation, registration, and recall sections of MMSE. In a 2019 study [135], 80 patients with mild AD participated. They were separated into sham and experiment groups. The experiment group used a software for training, whereas the sham group had free access to the internet to read articles, play games, or solve puzzles. Both groups had the same frequency

of interventions. Patients in the experiment group gained more positive effects in areas such as working memory, executive operations, and language comprehension. However, the effects started to decline to baseline measures after 12 months post-intervention, though remained stable for the first 6 months. In another study [136], AD patients in a younger age group (61-69 years old) showed more cognitive improvements than the participants in the older age group (70-81 years old). This suggests that the cognitive progress potential is age-dependant. In all, findings have supported cognitive training in dementia patients.

It has been suggested that web-based cognitive exercises with various cognitive functions may work better than traditional methods for dementia patients [137]. In a randomized study [138], 60 patients with MCI and AD were divided into two groups: the first group (n = 37) participated in computerized cognitive training combined with traditional cognitive training (using pen and paper) for 3 months, while the second group (n = 23) participated in traditional cognitive training, only. Computerized training was effective for reducing anxiety and promoting better decision-making at a 12-month follow-up, but no significant difference was reported on the scores of basic cognitive functions. That said, higher effect sizes were seen with computerized cognitive training than in just traditional cognitive training, alone. The authors hypothesized that this could be the result of a ceiling effect. Another study [139] conducted on patients with mild to moderate dementia (n = 15) and healthy older adults (n = 32) indicated that participants who received a combination of traditional and computerized cognitive training significantly improved in Dementia Rating Scale scores and logical memory tasks in WMS-III assessment, with effects lasting for 8 weeks. Alternatively, Hagovska et al. [140] compared 2 groups of MCI patients (n = 30, each). One group performed computerized cognitive exercises while the other completed traditional cognitive exercises. After training, the first group had significantly higher scores for quality of life, attention, and a language subscale. The transfer effect to functional activities was identical between the two groups. In a different study [141], 21 healthy older adults and 14 MCI patients received computerized working memory trainings with adaptive difficulty levels. After intervention, healthy participants had lower prefrontal activation (higher processing efficiency) while doing tasks with high working memory load. On the other hand, no changes in the prefrontal activation were observed for MCI patients. They did, however, improve in behavioral performance in working memory tasks with low load. Finally, in a more recent study [120], it was observed that 9 dementia patients improved in WMS-IV scores after playing cognitive games on an iPad. Thus,

switching from traditional cognitive trainings to computerized trainings, or a combination of the two, is recommended for its reported benefits.

Combining help or therapy with cognitive exercises has led to interesting results as well. Hyer et al. [142] tested tutored working memory cognitive trainings on older adults with MCI. Participants were divided into active and sham (n = 34, each) groups. Both groups performed the same exercises, though the level of difficulty remained constant in the sham group. Active and sham participants improved, but those in the active group maintained their improvement at a 3-month follow-up, whereas the sham group declined. In another study [143], 348 participants comprised of healthy older adults, mild AD, and MCI patients received 3 months of computerized cognitive training with reminiscence therapy. All 3 groups improved in verbal memory. Healthy adults also showed a small improvement in executive function tasks which lasted up to the 3-month follow-up. In addition, memory improvements for MCI and healthy participants were maintained at follow-up. Though AD and MCI patients had enhanced MMSE scores, the positive effects did not remain until follow-up. More research is needed on the confounding effects of therapy and cognitive exercises to determine the overall benefits to patients.

In sum, cognitive training has some benefits for AD/MCI patients, however, no evidence was found for computerized cognitive training for patients with behavioral variant of frontotemporal dementia [144]. Small sample sizes and overlap of early symptoms with AD can be the reason for undersupply of studies for this group [144], [145]. As for the AD/MCI groups, the positive effects may start to decline after several months. Hence, it is proposed to perform cognitive exercises regularly and continuously, preferably with the help of a tutor [120].

2.6 Summary

Aging strikes at the health of neural structures of the brain. However, for dementia patients, the rate of neural deterioration is more advanced than cognitively healthy older adults. For FTD, AD, and MCI patients, most of the neural loss happens in the frontal and temporal lobes, most prominently affecting their behavior and memory. Moreover, speaking abilities decline in patients with dementia. Acoustically, dementia patients present with more pauses, voiceless segments, and shimmer in their speech. Linguistically, patients use more pronouns, verbs, and adjectives than nouns. Also, dementia patients often encounter difficulties with word finding and object naming. Cognitive training has been suggested as a method for alleviating some aspects of cognitive decline for dementia patients, such as working memory. Studies have suggested that training offers

positive effects to the patients, however, the effects start to fade after several months. Therefore, it is advised to conduct the trainings frequently and consistently to support long-term cognitive benefits.

Chapter 3: Methodology

The current research is part of an on-going study (ClinicalTrials.gov NCT03587012) in which participants' voice samples have been collected from May 2020 to July 2021. The following paragraphs elaborate on the study design, the selected cognitive exercises, the participants, and the data collection method. Further, detailed explanations are given for feature extraction and analysis procedures.

3.1 Study Design

Participants practiced a series cognitive exercises with the help of a tutor for 40 minutes per day, 5 days per week, and for 4 weeks. Due to the Covid-19 pandemic, all trainings were performed online, using either Zoom or Microsoft Teams platforms. Cognitive training was conducted using the app, MindTriggers [146], on an iPad. MindTriggers consists of 7 serious games of varying difficulties for targeting spatial orientation, associative memory, left-right brain connectivity, and short-term memory. Each of the games are briefly explained in the following:

- **Associative Memory:** A 2 by 2 grid is presented with a shape containing an animal in each cell. Three shapes are the same with changing orientations in the x-y plane, and the fourth shape is different. Participants must select the different shape in a series of training trials. After the training trials, players must associate the appropriate animal with its commonly assigned shape.
- **Word-Image Association:** A number of randomly chosen images (n) are displayed for a given time. Next, n + 1 boxes appear with a word on top, and the player must remember whether the picture associated with each word was presented before.
- **Memorize and Re-write:** A sentence to be memorized is presented to the player. Following, the words making up the sentence are scrambled, and the player must rewrite the original phrase.
- **Visual Memory:** Players are presented with n images across the screen and asked to memorize them. The objective of this game is to identify which images they were originally shown amongst an increased number (3n) of assorted figures.
- **Make a Sentence:** Several words are scattered across the screen, from which the player is required to form as many sentences as they can within a given time limit.

- **Categorization:** The purpose of this game is to categorize a series of images into subgroups. For example, a number of animals and vehicles may be presented, and the player must move the images into their appropriate categories (animals versus modes of transportation).
- **Spatial Orientation:** In this game, the player sees a symmetric hexagonal room with three landmarks and a central avatar. The room has colored walls and a grid tile floor. The player must move their avatar to find a Target Area: a highlighted tile in the room whose location throughout the game does not change. The aim of the game is to relocate the Target AreA continuously. To increase difficulty, landmarks are removed successively, the room is rotated, and the grid is removed from the floor.

The primary outcome measure of the study was WMS-IV assessment, which again, was managed online at baseline, post-intervention, and 1-month follow-up. To note, participant no. 4 did not have a follow-up assessment. WMS-IV is a neuropsychological examination, which can evaluate different memory functions. The six primary sections of WMS-IV are auditory memory (maximum score: 142), delayed memory (maximum score: 92), immediate memory (maximum score: 136), recognition memory (maximum score: 80), visual memory (maximum score: 86), and visual working memory (maximum score: 50) [120], [147]. Twelve sub-sections of the WMS-IV assessment were used in the current study to assess the target brain functions. These sub-sections included brief cognitive status exam, visual reproduction I and II, logical memory I and II (LM I and LM II), verbal paired associates I and II, symbol span, logical memory II recognition, verbal paired associates II recognition, verbal paired associates II recall, and visual reproduction II recognition [120]. The sum of the selected sub-sections resulted in a total maximum score of 358. The Kruskal-Wallis test was utilized to examine the effects of gender, education, age, and nativity in English with the WMS-IV outcomes. For feature extraction, LM I and LM II sections (maximum scores of 53 and 39, respectively) were chosen for their continuous speech segments. In these segments, participants were asked to recall 2 short stories once immediately (LM I) and once after a 30–45-minute gap (LM II). In LM I, the first story was repeated twice, but only the second repetition was considered in this research. The contents of the 2 stories are displayed in Table 3.

As mentioned, the current study was part of bigger project on testing brain exercises on healthy older adults and dementia patients. The current study was authorized by the Biomedical Research Ethics Board of the University of Manitoba (ethics number: HS21624 (B2018:029)). All

participants provided written consent. After ethics approval, participants were recorded during WMS-IV assessments. The audio recordings were trimmed to only keep LM I and LM II sections. Next, the audio files were manually transcribed. Transcriptions were reviewed for accuracy by a second research assistant.

Table 3-Story scripts in the logical memory sections of WMS-IV assessment

Story 1	Story 2
Ruth and Paul have been friends for 30 years. They meet every Tuesday at Alma's diner for breakfast and then they go for a walk in Mason Park.	Anna Thompson of South Boston employed as a cook in a school cafeteria reported at the police station that she had been held up on State Street the night before and robbed of 56 dollars. She had 4 small children. The rent was due, and they had not eaten for 2 days. The police touched by the woman's story took up a collection for her.

3.2 Participants

The main study had more than 35 participants; the outcomes of the study on 29 of them were published by Moussavi et al. [120]. As this current study for speech analysis was added later and immediately after the start, it was hit by the pandemic, and only 8 participants were enrolled during the duration of this study who also had speech data recorded. The 8 participants' demographics are detailed in Table 4. All volunteers were assessed by MoCA test after they signed the consent forms. To note, participant no. 5 did not present with AD or MCI, though they were a probable case of FTD. From this point forward, the participants of this study will be called the Exr group. The inclusion criteria for the Exr group were the following [120]:

- Older than 60 years old
- MoCA score between 7 and 25
- Clinically diagnosed with MCI or mild/moderate stage of a dementia sub-type.
- Other than dementia, have no diagnosis of other neurological disorders, such as Huntington disease, amyotrophic lateral sclerosis, Parkinson's disease, multiple sclerosis, major depression, bipolar disorder, and schizophrenia.

- Fluent in reading and writing English.
- Be able to enroll in the clinical trial for 5 days per week for 4 weeks.

Table 4-Participants' demographic

Participant	Gender	Age	Native in English?	Education	MoCA	Diagnosis
1	F	91	Yes	PhD	23	AD
2	F	76	Yes	Diploma	23	MCI
3	M	79	Yes	Masters	21	AD
4	F	65	Yes	Bachelors	9	AD
5	M	70	No	PhD	23	FTD
6	M	86	Yes	Masters	14	AD
7	F	74	Yes	Grade 11	26	MCI
8	M	76	Yes	Grade 11	23	AD
Mean ± std		77.13 ± 8.32			20.25 ± 5.73	

AD: Alzheimer's Disease
MCI: Mild Cognitive Impairment
FTD: Frontotemporal Dementia
MoCA: Montreal Cognitive Assessment
std: Standard Deviation

3.3 Data Analysis and Feature Extraction

As indicated in **Chapter 2: Background**, many acoustic and linguistic features have been introduced in literature to distinguish between dementia patients and healthy adults. To evaluate the Exr group, 28 linguistic and 100 acoustic features were selected based on previous literature. Linguistic features were extracted with the help of Natural Language Processing Toolkit [88] in Python. Participants' transcriptions were pre-processed and cleaned. The pause words ("ah", "hmm", "mhm", "oh", "uh") were excluded after counting and abbreviations were replaced with their complete forms (i.e. haven't became have not).

Before computing the acoustic features, audio files were manually trimmed to only keep the patients' voice clips. In addition, audio files were denoised using a noise gate in Audacity [148]. Eighty-eight of the acoustic features were chosen from the Geneva Minimalistic Acoustic Parameter Set (GeMAPS) [149] and were extracted using the "OpenSMILE" library in Python. The non-linear parameters such as fractals, entropy, and zero crossing rate were calculated with the "antropy" library in Python. The rest of the acoustic features which are related to voice activity detection were computed with the "webrtcvad" library in Python. Table 5 and Table 6 are presented

to summarize evaluated linguistic and acoustic features, respectively. The descriptions of each feature were taken from the referred papers.

Table 5-Summary of analyzed linguistic features

Feature	Description
Number of sentences [9]	-
Number of words [20]	-
Number of pause words [20]	Number of words that indicate hesitation including "ah", "hmm", "mhm", "oh", "uh".
Number of unique words [6], [94]	Number of words that are only repeated once (after stemming or only keeping the roots of the words).
Text length [6]	Number of characters in the entire transcription file.
Number of parts of speech tags [20]	Number of nouns, pronouns, adjectives, adverbs, determiners, and coordinating conjunctions.
Rate of parts of speech tags [20]	Number of parts of speech tags divided by the total number of words.
Parts of speech ratios [6]	Number of pronouns/number of nouns, number of adjectives/number of nouns, number of verbs/number of nouns.
Type token ratio [6]	Number of unique words/text length.
Honoré's statistic [6], [150]	$\frac{100 * \log(\text{text_length})}{(1 - \frac{\text{number of words that are only repeated once}}{\text{number of words}})}$
Brunet's index [6], [151]	$\text{text_length}^{\text{number_of_words}^{-0.165}}$
Verbal rate [55]	Number of words/total locution time
Standardized phonation time [55]	Number of words/total locution time excluding the pauses

Table 6-Summary of analyzed acoustic features

Feature	Description
Pitch (F0) (mean and CoV, 20^{th}, 50^{th}, 80^{th} percentile, the range of the 20^{th} to 80^{th} percentile, the mean and std of the rising and falling slopes) [149]	F0 (calculated on a semitone frequency scale, starting at 27.5 Hz).

Jitter (mean and CoV) [149]	Frequency perturbation in consecutive F0 periods divided by the average length of pitch period.
Formants 1, 2, 3 frequencies and band widths (mean and CoV) [149] (Hz)	-
Shimmer (mean and CoV) [149] (dB)	Peak amplitude perturbation in consecutive F0 periods divided by the average amplitude.
Loudness (mean and CoV, 20^{th}, 50^{th}, 80^{th} percentile, the range of the 20^{th} to 80^{th} percentile, the mean and std of the rising and falling slopes) [149]	Estimation of signal's intensity.
Harmonics to noise ratio (mean and CoV) [149] (dB)	Logarithm (energy of the harmonic (periodic) segments of the signal/energy of the noisy (aperiodic) segments of the signal).
Alpha ratio (mean and CoV) [149] (dB)	Logarithm (Energy in 50-1000 Hz/energy in 1-5 kHz).
Hammarberg index (mean and CoV) [149] (dB)	Logarithm (Strongest energy peak in 0-2 kHz/strongest energy peak in 2-5 kHz).
Spectral slope 0-500 Hz (mean and CoV) [149]	Slope of the power spectrum in 0-500 Hz (logarithmic) calculated using linear regression.
Spectral slope 500-1500 Hz (mean and CoV) [149]	Slope of the power spectrum in 500-1500 Hz (logarithmic) calculated using linear regression.
Relative energy of formants 1, 2, 3 (mean and CoV) [149] (dB)	Logarithm (energy of the spectral peak at formant's center frequency/energy of spectral peak at F0).
H1-H2 harmonic difference (mean and CoV) [149] (dB)	Logarithm (energy of the first F0 harmonic (H1)/energy of the second F0 harmonic (H2)).
H1-A3 harmonic difference (mean and CoV) [149] (dB)	Logarithm (energy of the first F0 harmonic (H1)/energy of the strongest harmonic in the third formant range).
Mel frequency coefficients (MFCC) [1-4] (mean and CoV) [149]	The first 4 Mel frequency coefficients.
Spectral flux (mean and CoV) [149]	Difference in spectra of 2 successive frames.

Rate of loudness peaks [149]	Number of loudness peaks in each second.
Voiced duration (F0 > 0) (mean and std) [149]	Length of voiced sections of the signal per second.
Unvoiced duration (F0 = 0) (mean and std) [149]	Length of unvoiced sections of the signal per second.
Pseudo syllable rate [149]	Number of uninterrupted voiced segments per second.
Equivalent sound level [149]	Logarithm of average root mean square energies in each frame.
Total duration [97] (s)	Total duration of the signal including voiced and unvoiced.
Voiced/unvoiced ratio [12]–[14], [22]	Voiced duration/total duration, unvoiced duration/total duration, unvoiced duration/voiced duration.
Spectral entropy [22]	Based on Shannon entropy, measures disorganization of the signal.
Fractal dimensions [22]	Higuchi, Katz, Petrosian fractal dimensions.
Zero crossings [22]	Number of zero crossings.
CoV: Coefficient of variation, which is standard deviation normalized by the arithmetic mean [149]. std: Standard Deviation	

3.4 Feature selection from Pitt Corpus

To ensure only features that were significantly different among dementia patients and control participants were being analyzed, a supplementary database called DementiaBank was utilized. DementiaBank consists of several datasets in different languages, such as English, German, Mandarin, Spanish, and Taiwanese. Each dataset contains audio or video recordings of patients with different types of dementia and control groups while performing oral tasks. The Pitt corpus [33] from DementiaBank is comprised of English audio recordings of patients with probable AD and cognitively healthy volunteers. In addition, CHAT protocol was used to transcribe the audio files, for which transcriptions were made available. The data was collected by the Alzheimer and Related Dementias Study at the University of Pittsburgh School of Medicine from 1983 to 1988. The tasks in the Pitt study were story recall, picture description, sentence construction, and verbal fluency. The picture description section of this dataset has a larger sample

size compared to the other parts and has been particularly popular in previous works for AD versus control classification [9], [19], [84], [85]. The picture selected for this task was the Cookie Theft [91] picture. Some volunteers were assessed more than one time and on a yearly basis, but for the purpose of the present work, only the first assessments were analyzed. For the first assessment, 171 and 73 recordings were available for participants with and without probable AD. However, because of poor audio quality, only 166 (109 females, age 77.19 ± 8.76) and 64 (38 females, age 63.84 ± 8.82) files from probable AD and control recordings were selected, respectively. These files were analyzed identically to the description set forth for Exr group. Finally, to identify statistically significant features between AD and control participants, Wilcoxon rank-sum test was applied.

3.5 Classifier Training with Pitt Corpus

To check how well the statistically significant features can separate patients with probable AD and control participants in Pitt corpus [33], some classic machine learning classification algorithms such as SVM, RF, DT, LR, KNN were tested using Scikit Learn library [152] in Python (80% for training and 20% for testing, data was randomly shuffled before splitting). The training data was scaled to have 0 mean and 1 standard deviation (std). Further, using analysis of variance (ANOVA) F-values [152], 'K' most important features were selected as inputs to the algorithms. Since the dataset was unbalanced, the minority class was over-sampled, and the majority class was under-sampled [153]. Also, a 10-fold cross validation method was utilized to select the best model parameters. As classification metrics, accuracy, precision, sensitivity, specificity, F-score, and area under the receiver operating characteristic curve (AUC-ROC) were considered.

3.6 Classifier Testing with Exr Group

The machine learning models with the selected features were tested on the Exr group to check if they could correctly classify the participants as dementia patients.

3.7 WMS-IV Correlation Analysis for Exr Group

To compute correlations between the features and assessment results (LM I, LM II, and WMS-IV total scores), Kendall correlation coefficient was employed. The correlations were calculated between the changes in features from baseline to post-intervention ($\frac{\text{Post-Intervention-Baseline}}{\text{Baseline}}$). The objective was to identify the acoustic and linguistic features that had high or very high correlations [154] to changes in WMS-IV scores.

3.8 Feature Categorization Using Pitt Corpus

The means of the significantly different features in the Pitt corpus [33] were compared to see which group (AD or control) had a higher mean for each feature. The features with higher means in the control groups were put into group 'H' and the other features were put in group 'D'.

3.9 Post-Intervention and Follow-Up Analyses for Exr Group

One of the primary objectives of the current study was to analyze how cognitive exercises affect the speech of dementia patients. Therefore, the acoustic and linguistic features needed to be analyzed pre- and post-intervention. For Exr group, linguistic and acoustic features were divided into groups 'H' or 'D'. First off, missing values were replaced by the medians of the features of other participants. Next, for each participant, means of normalized (each feature was divided by its maximum value, so all features had values between 0 to 1) acoustic and linguistic features at each point in time (baseline, post-intervention, and follow-up) were calculated separately. Subsequently, the acoustic and linguistic means of all participants were calculated and plotted overtime to demonstrate the changes. An improvement was defined as when the linguistic or acoustic overall mean increased for the H features and decreased for the D features over time. This part of the analysis was conducted in R software [155]. Furthermore, the changes in LM I, LM II, and WMS-IV total scores were plotted.

3.10 Summary

To summarize, using both open source and locally derived datasets, acoustic and linguistic features for the classification of dementia were analyzed. Figure 2 illustrates the study procedure and data analysis process.

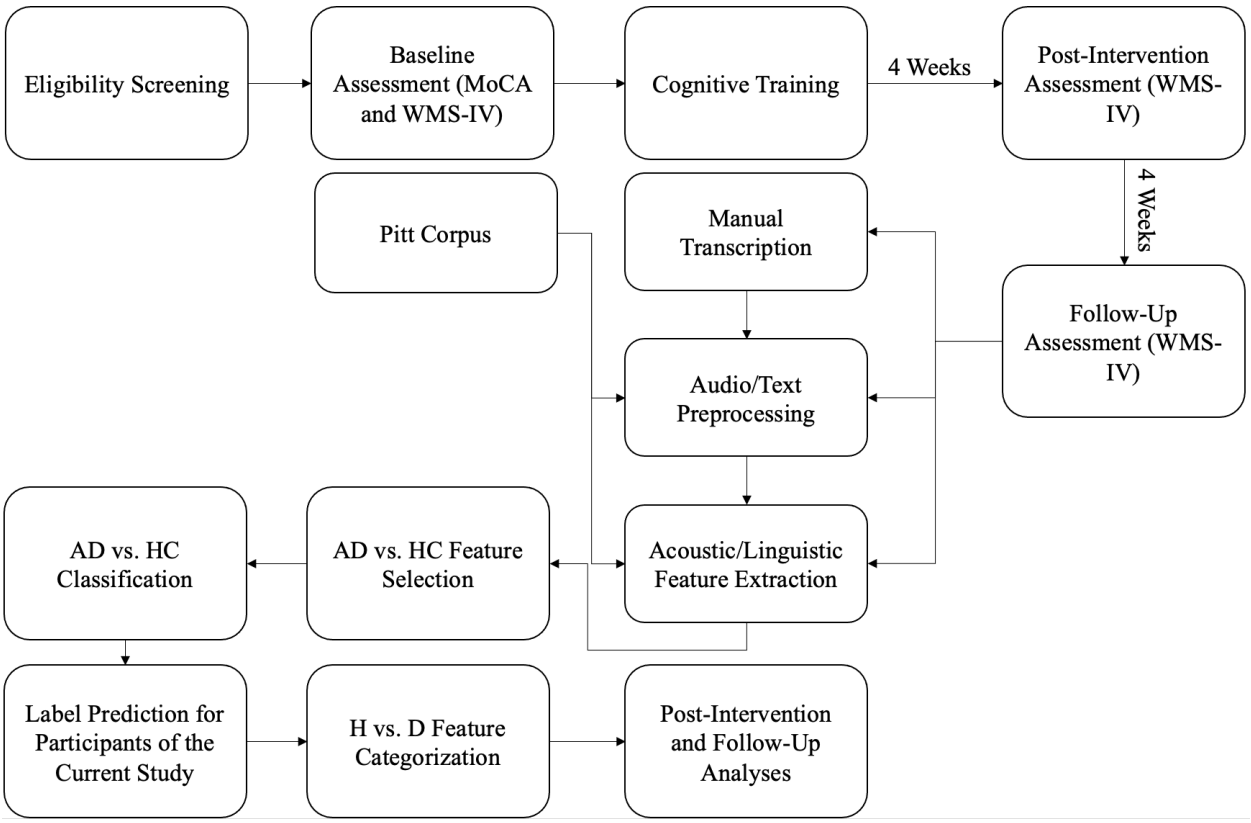


Figure 2-Study procedure flowchart

Chapter 4: Results

Using the Pitt corpus [33], features that were significantly different amongst dementia patients and healthy controls were detected. Next, the features were categorized into two groups ('D' or 'H') based on their respective means among the two groups. Afterwards, five classifiers were trained on the Pitt data with top 20 features that were selected by ANOVA F-values.

Following, the participants of this study (Exr group) were used as a testing dataset to check if the models could identify the participants as dementia patients. In parallel, for the Exr group, correlations of all features to LM I, LM II, and WMS-IV total scores were investigated, and the changes in the trends of the means of acoustic and linguistic features were analyzed.

After performing the Kruskal-Wallis test, overall, it was observed that age, gender, level of education, English as a first language, and diagnosis of the participants in the Exr group did not affect the average total WMS-IV score calculated from baseline, post-intervention, and follow-up at 0.05 significance level. Furthermore, the above-mentioned factors presented no effect on the average LM I and LM II scores. The following sections elaborate on the results.

4.1 Feature Selection and Categorization Using Pitt Corpus

Using the Pitt corpus [33] dataset, acoustic and linguistic features were compared between control and AD participants by applying a non-paired Wilcoxon rank-sum test at 0.05 significance level. The groupings of the features were based on which group (dementia or healthy control) had the higher mean. The results are presented in Table 7 and Table 8.

Table 7- Acoustic features with significant differences between AD and control participants

Feature	Mean of Dementia Group (Pitt Corpus)	Mean of Control Group (Pitt Corpus)	P-Value (Wilcoxon Rank-Sum Test)	Group
H1-A3 harmonic difference (mean) (dB)	23.78	15.73	0.00005	D
H1-H2 harmonic difference (mean) (dB)	-0.31	-2.26	0.011	D

Hammarberg index (voiced, mean) (dB)	29.32	25.43	0.004	D
Hammarberg index (voiced, CoV) (dB)	0.57	-1.94	0.016	D
HNR (mean) (dB)	3.77	2.63	0.003	D
Jitter local (CoV)	1.68	1.49	0.0001	D
Lengths of voiced segments per second (mean)	0.25	0.20	0.001	D
Lengths of voiced segments per second (CoV)	0.22	0.17	0.001	D
Loudness (mean of falling slope)	4.74	3.57	0.015	D
Loudness (mean of rising slope)	6.52	4.85	0.017	D
Loudness (CoV)	1.28	1.06	0.000004	D
Loudness (std of falling slope)	3.59	2.70	0.011	D
Loudness (std of rising slope)	4.68	3.43	0.009	D
MFCC 1 (mean)	22.49	20.15	0.018	D
MFCC 1 (voiced, mean)	34.99	29.50	0.00006	D
MFCC 4 (mean)	-3.18	-7.12	0.0002	D
MFCC 4 (voiced, mean)	-17.43	-20.92	0.023	D

Equivalent sound level (dB)	-39.21	-44.01	0.0007	D
Shimmer Local (CoV) (dB)	0.68	0.61	0.00005	D
Spectral flux (CoV)	1.89	1.63	0.002	D
Spectral flux (voiced, mean)	0.42	0.31	0.014	D
Spectral slope 500-1500 Hz (unvoiced, mean)	-0.004	-0.006	0.017	D
Lengths of unvoiced segments per second (mean)	0.66185	0.66173	0.02	D
Alpha Ratio (voiced, mean) (dB)	-16.64	-12.82	0.001	H
F0 (range of 20^{th} to 80^{th} percentile)	9.09	12.45	0.0006	H
F0 (CoV)	0.25	0.29	0.014	H
Formant 1 frequency (CoV) (Hz)	0.34	0.38	0.003	H
Formant 2 frequency (CoV) (Hz)	0.19	0.20	0.001	H
Formant 3 bandwidth (CoV) (Hz)	0.48	0.52	0.0008	H

Higuchi fractal dimension	1.32	1.38	0.020	H
Katz fractal dimension	2.61	2.68	0.038	H
Pseudo syllable rate	1.29	1.51	0.0007	H
Loudness (20^{th} percentile)	0.03	0.04	0.003	H
MFCC 1 (voiced, CoV)	0.51	0.65	0.011	H
MFCC 4 (voiced, CoV)	-10.90	-0.86	0.015	H
Rate of loudness peaks	1.57	1.78	0.020	H
Shimmer Local (mean) (dB)	1.59	1.72	0.003	H
Spectral slope 500-1500 Hz (voiced, mean)	-0.029	-0.025	0.019	H

D: Dementia
H: Healthy
CoV: Coefficient of variation, which is standard deviation normalized by the arithmetic mean [149].

Table 8- Linguistic features with significant differences between AD and control participants

Feature	**Mean of Dementia Group (Pitt Corpus)**	**Mean of Control Group (Pitt Corpus)**	**P-Value (Wilcoxon Rank-Sum Test)**	**Group**
Adjective noun ratio	0.27	0.19	0.0002	D

Number of adjectives	5.62	4.55	0.042	D
Rate of adjectives	0.06	0.04	0.0001	D
Rate of adverbs	0.05	0.04	0.033	D
Rate of pronouns	0.10	0.08	0.036	D
Brunet's index	17.15	17.63	3.57e-11	H
Honoré's statistic	1042.86	1103.57	0.082	H
Number of determiners	13.61	15.82	0.003	H
Number of unique words	35.06	41.57	0.0001	H
Number of verbs	21.21	26.31	0.00001	H
Rate of verbs	0.23	0.26	0.0000003	H
Standardized phonation time	4.05	9.59	0.0013	H
Text length	423.21	484.72	0.007	H
Verbal rate	1.96	2.44	0.0000028	H
Verb to noun ratio	1.05	1.20	0.0008	H

D: Dementia
H: Healthy

Considering the statistically different features introduced in Table 7 and Table 8, the 38 acoustic features and 15 linguistic features were used to classify AD and control participants.

4.2 Classifier Training Using Pitt Corpus

Age and gender were added to the feature set as demographics information. The mean and std of evaluation metrics for different classifiers are presented in Table 9 after 10-fold cross validation. For all classifiers the best results were yielded using the 20 most important features consisting of F0 (the range of the 20th to 80th percentile), loudness (coefficient of variation (CoV)), MFCC 4 (mean), MFCC 4 (CoV), shimmer (mean), shimmer (CoV), H1-H2 harmonic difference (mean), A1-H3 harmonic difference (mean), first formant frequency (CoV), spectral flux (voiced, mean), pseudo syllable rate, number of verbs, rate of verbs, rate of adjectives, adjective to noun ratio, number of unique words, Brunet's index, verbal rate, standardized phonation time, and age. As can be seen the LR and RF classifiers outperformed the others.

Table 9-Classification results for machine learning algorithms

Classifier	Accuracy	Precision	Sensitivity	Specificity	F-score	AUC-ROC
SVM	0.80 ± 0.07	0.86 ± 0.07	0.85 ± 0.07	0.63 ± 0.13	0.85 ± 0.05	0.83 ± 0.07
KNN	0.71 ± 0.07	0.87 ± 0.08	0.70 ± 0.07	0.74 ± 0.13	0.77 ± 0.06	0.77 ± 0.08
LR	0.83 ± 0.06	0.92 ± 0.08	0.83 ± 0.08	0.83 ± 0.16	0.87 ± 0.05	0.87 ± 0.07
DT	0.73 ± 0.10	0.87 ± 0.08	0.74 ± 0.17	0.70 ± 0.17	0.78 ± 0.10	0.73 ± 0.09
RF	0.85 ± 0.05	0.89 ± 0.08	0.91 ± 0.07	0.71 ± 0.21	0.89 ± 0.03	0.86 ± 0.08

SVM: Support Vector Machine
KNN: K Nearest Neighbors
LR: Logistic Regression
DT: Decision Tree
RF: Random Forest
AUC-ROC: Area Under the Receiver Operating Characteristic Curve

4.3 Classifier Testing Using Exr Group

All classifiers were tested on how accurately they could predict dementia labels for participants of this study. RF and LR models with top 10 features led to the best results. The top 10 features included loudness (CoV), shimmer (mean), frequency of the first formant (CoV), rate of verbs, rate of adjectives, adjective to noun ratio, Brunet's index, verbal rate, standardized phonation time, and age. With RF classifier (entropy criterion[1], complexity parameter[2] = 0.01) and during LM I, only 1 AD patient was wrongly categorized as control. With LM II, all participants were correctly grouped as dementia patients. With LR classifier (l2 penalty[3], inverse regularization strength[4] = 1), during LM I, 2 AD patients were misclassified as controls. During LM II, only the FTD patient was not correctly classified.

[1] Entropy works based on information gain and it determines the quality of a tree split.
[2] This parameter is used for pruning the trees. The tree with the largest cost smaller than the complexity parameter will be chosen for next rounds.
[3] L2 penalty adds a regularization term to the loss function that is equal to the square of the magnitude of coefficients.
[4] Smaller values of inverse regularization strength indicate stronger regularization.
The above footnotes were extracted from https://scikit-learn.org/stable/.

4.4 WMS-IV Correlation Analysis for Exr Group

The correlations of the changes in acoustic and linguistic features to the change in LM I score of the WMS-IV assessment from baseline to post-intervention were evaluated, with results summarized in Table 10. Note that the higher LM I or LM II scores, the better cognitive status.

Table 10-Feature correlations to the change of LM I score of the WMS-IV assessment from baseline to post-intervention (extracted from LM I section)

Feature	Correlation Coefficient
Shimmer (mean)	0.79**
HNR (mean)	-0.79**
HNR (CoV)	0.79**
Number of adverbs	0.72*
* $p < 0.05$ ** $p < 0.01$ CoV: coefficient of variation, which is standard deviation normalized by the arithmetic mean [149].	

Also, the change in F0 (CoV) extracted from LM I section had the highest correlation to the change in WMS-IV total score. The next features were MFCC 1 (mean and CoV) and voiced duration. The correlation coefficients and the levels of significance can be observed in Table 11.

Table 11-Feature correlations to change of WMS-IV total score of the WMS-IV assessment from baseline to post-intervention (extracted from LM I section)

Feature	Correlation Coefficient
F0 (CoV)	0.86**
MFCC 1 (mean)	-0.79**
MFCC 1 (CoV)	0.79**
Voiced duration	0.79**
Katz fractal dimension	0.71*
Number of zero-crossings	0.71*
Number of pronouns	0.71*
* $p < 0.05$ ** $p < 0.01$	

CoV: coefficient of variation, which is standard deviation normalized by the arithmetic mean [149].

Changes in F0 (50^{th} percentile), spectral entropy, and the number of words extracted from LM II had the highest correlations with the change in LM II score from baseline to post-intervention.

Table 12- Feature correlations to the change of LM II score of the WMS-IV assessment from baseline to post-intervention (extracted from LM II section)

Feature	Correlation Coefficient
F0 (50^{th} percentile)	0.79**
Spectral entropy	0.79**
Voiced duration	0.71*
Rate of verbs	0.71*
Number of determiners	0.71*
* $p < 0.05$ ** $p < 0.01$ CoV: coefficient of variation, which is standard deviation normalized by the arithmetic mean [149].	

At last, the change in H1-A3 harmonic difference (CoV) extracted from LM II section had the highest correlation with the change in WMS-IV total score. Following that, H1-H2 harmonic difference (mean) and the bandwidth of the first formant (CoV) had the highest correlation coefficients.

Table 13- Feature correlations to change of WMS-IV total score of the WMS-IV assessment from baseline to post-intervention (extracted from LM II section)

Feature	Correlation Coefficient
H1-A3 harmonic difference (CoV)	0.86**
First formant bandwidth (CoV)	0.79**
H1-H2 harmonic difference (mean)	0.79**
F0 (CoV)	0.71*
F0 (percentile range)	0.71*
* $p < 0.05$ ** $p < 0.01$	

CoV: coefficient of variation, which is standard deviation normalized by the arithmetic mean [149].

4.5 Post-Intervention and Follow-Up Analyses for Exr Group

Table 14 summarizes the changes in mean WMS-IV scores from baseline to follow-up sessions. All scores except the LM I score improved with time from baseline to post-intervention. Moreover, all scores increased from baseline to follow-up. Since participant no. 4 did not have any follow-up scores, their scores were replaced with the medians of the scores of other participants (for the purpose of averaging).

Table 14-Changes in mean values of LM I, LM II, and WMS-IV total scores

	Baseline	Post-Intervention	Follow-Up
LM I Score	20.63 ± 7.84	20.5 ± 8.77	24.25 ± 10.51
LM II Score	6.5 ± 5.76	7.88 ± 8.27	10.25 ± 8.03
WMS-IV Total Score	123.58 ± 48.33	141.30 ± 63.12	149.93 ± 60.78

Figure 3, Figure 4, and Figure 5 illustrate the trends of LM I, LM II, and WMS-IV total scores over time. For the 2 MCI patients, both LM I and LM II scores improved at one-month follow-up with respect to baseline. For participant no. 2 the total WMS-IV score progressed as well, but for participant no. 7, the increase was only observed at post-intervention.

For the FTD patient, LM I and LM II scores increased at post-intervention and follow-up sessions. Nevertheless, the total WMS-IV score was approximately 3 points lower in post-intervention than in follow-up. Although, they were still higher than baseline values. It is worth mentioning that the total WMS-IV scores for this participant were interpolated to justify some missing sections of the assessments.

Among AD patients, participant no.4 did not have a follow-up assessment. In this instance, the LM I score remained unchanged, LM II score decreased, and WMS-IV total score increased at post-intervention. Of the remaining 4 participants (no. 1, no. 3, no. 6, and no. 8), only participant no. 1 progressed in LM I score in follow-up. However, LM II and WMS-IV total scores improved for participants 1, 3, and 6 at post-intervention. The improvements in WMS-IV total scores only persisted for participant 1 at follow-up.

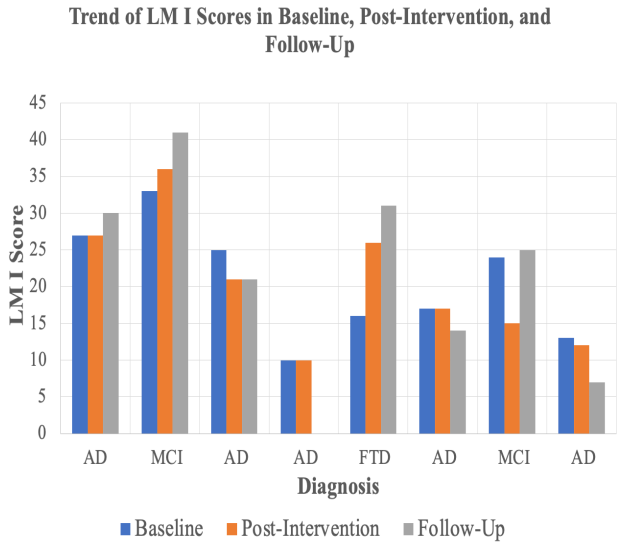


Figure 3-LM I score over time

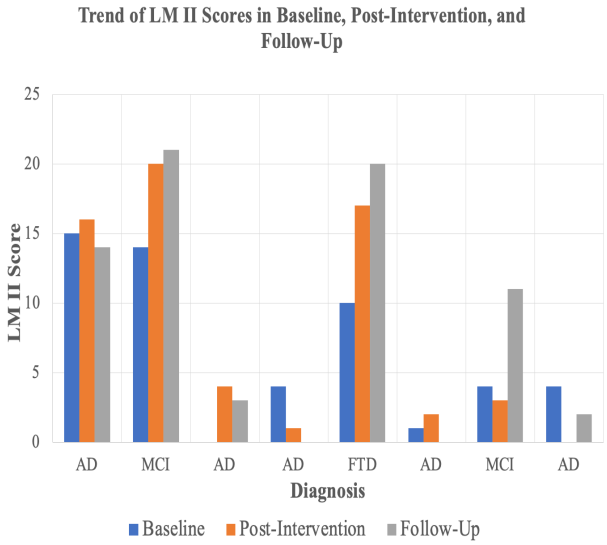


Figure 4-LM II score over time

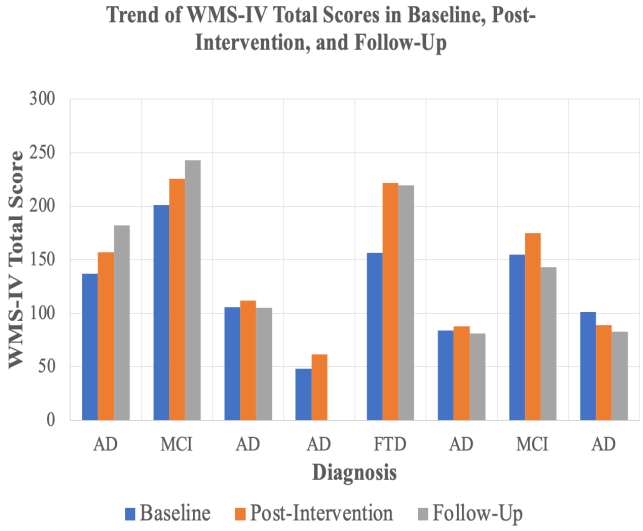


Figure 5-WMS-IV total score over time

Figures 6-13 demonstrate the means of H and D acoustic and linguistic features extracted from LM I and LM II sections. Figure 6 and Figure 7 show that during LM I, means of H and D acoustic features increased from post-intervention to follow-up and from baseline to follow-up. In addition, during LM I, the mean of H acoustic features declined from baseline to post-intervention, while the mean of D acoustic features improved. Figure 8 and Figure 9 indicate that during LM I and from baseline to post-intervention, the mean of the H linguistic features improved, while the mean of D linguistic features was reduced. The opposite trend was observed for post-intervention to follow-up and baseline to follow-up sessions. Figure 10 and Figure 11 show that during LM II, mean of H acoustic features increased from baseline to follow-up and from post-intervention to follow-up sessions, while the mean of D acoustic features only increased from baseline to post-intervention. As for the linguistic features during LM II, Figure 12 and Figure 13 display that both H and D parameters increased from baseline to post-intervention and follow-up sessions. It must be noted that based on Wilcoxon rank-sum test, none of these changes were found to be significant.

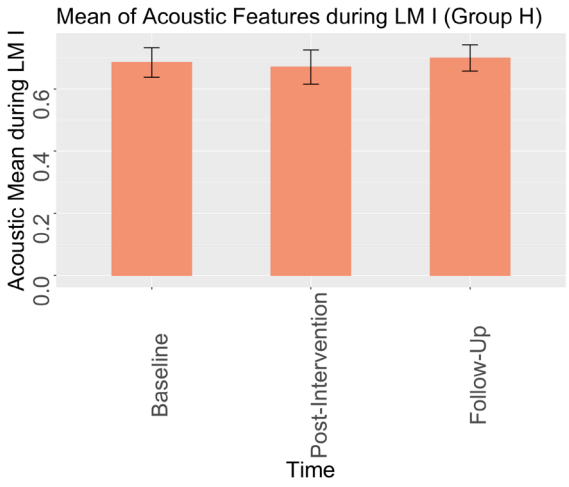


Figure 6-Mean of acoustic features during LM I (group H)

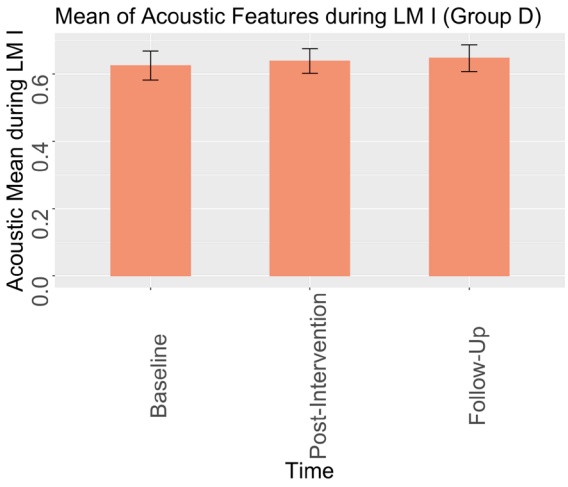


Figure 7-Mean of acoustic features during LM I (group D)

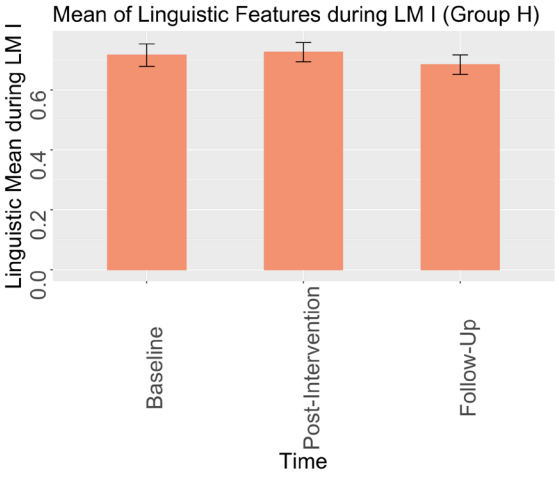


Figure 8-Mean of linguistic features during LM I (group H)

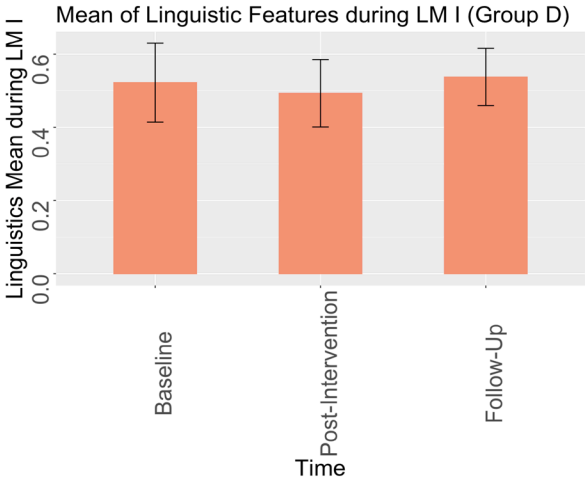


Figure 9-Mean of linguistic features during LM I (group D)

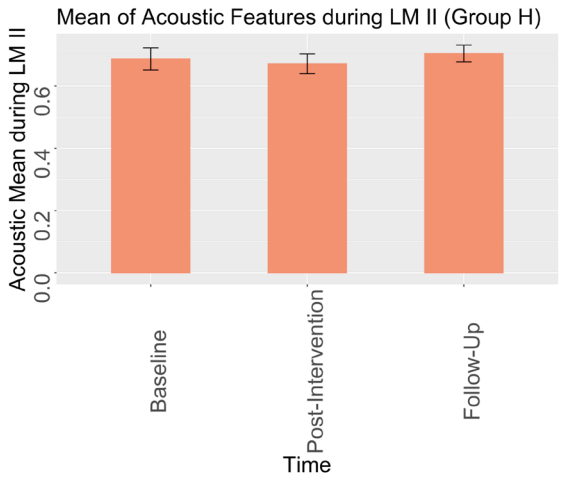


Figure 10-Mean of acoustic features during LM II (group H)

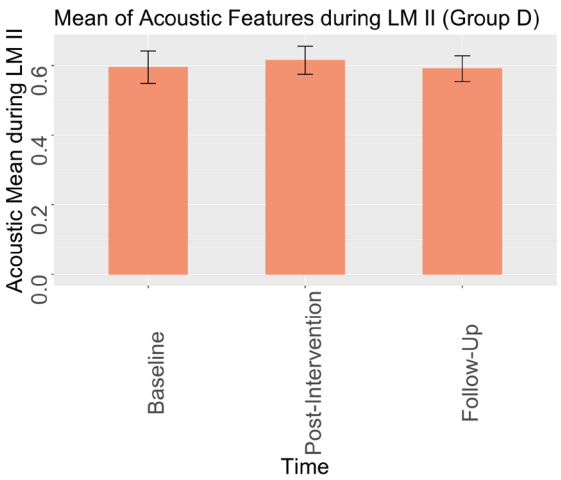


Figure 11-Mean of acoustic features during LM II (group D)

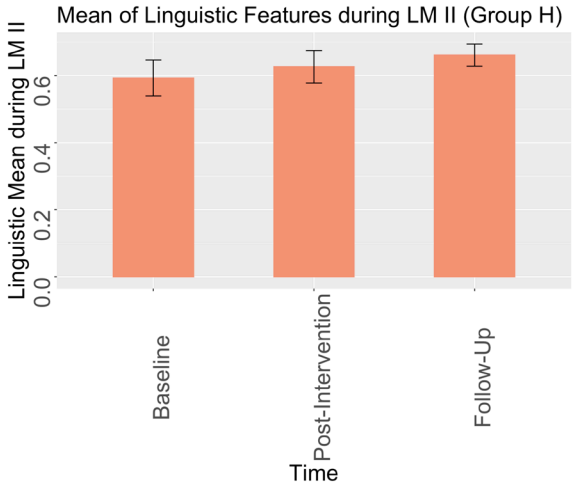


Figure 12-Mean of linguistic features during LM II (group H)

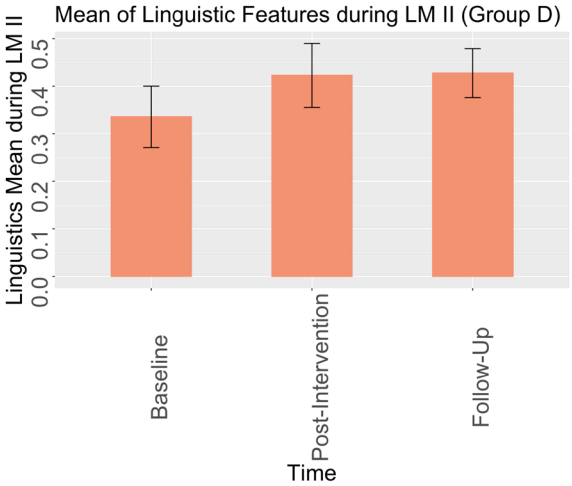


Figure 13-Mean of linguistic features during LM II (group D)

4.6 Summary

The Wilcoxon rank-sum test on the Pitt data [33] indicated that acoustic features relating to F0, formants, harmonic differences, Hammarberg index, jitter, shimmer, loudness, MFCC, pseudo syllable rate, spectral flux, spectral slope, lengths of voiced/unvoiced segments, and fractal dimensions significantly differed between patients with probable AD and the control group. Alternatively, linguistic features such as adjective to noun ratio, verb to noun ratio, number of adjectives, determiners, verbs, and unique words, rate of adjectives, adverbs, pronouns, and verbs, text length, verbal rate, standardized phonation time, Brunet's index, and Honoré's statistic significantly differed between patients with probable AD and the control group. In addition, mean of various features including harmonic differences, Hammarberg index, HNR, jitter, lengths of voiced segments per second, loudness, MFCC 1 and 4, equivalent sound level, shimmer, spectral flux, spectral slope, adjective to noun ratio, number of adjectives, rate of adjectives, adverbs, and pronouns were higher for the (probable) AD participants. The best validation results were achieved using a RF model with 85% accuracy, 89% precision, 91% sensitivity, 71% specificity, 89% F-score, and 86% AUC-ROC. Using the features from LM I section, only 1 AD patient from the Exr group was wrongly classified. However, with features from LM II section all the participants were correctly classified as dementia patients. Further, it was observed that the highest correlation to changes in WMS-IV scores was 0.86, which occurred between the changes in WMS-IV total score and F0 (CoV) and the changes in WMS-IV total score and H1-A3 harmonic difference (CoV) extracted from LM I and LM II sections, respectively. Moreover, feature improvement analysis indicated that mean of H linguistic parameters (extracted from LM I section) increased from baseline to post-intervention, while the mean of D linguistic parameters decreased.

Chapter 5: Discussion

In this study, acoustic and linguistic features were extracted from patients' interviews and an available online dataset called the Pitt corpus [33]. Next, the correlations of changes in the features (from baseline to post-intervention) with changes in WMS-IV scores were evaluated. Following, with the help of the Pitt dataset [33], features that were significantly different between dementia patients and control participants were identified. Afterwards, machine learning algorithms were trained on the Pitt data [33] and tested on the participants of this study. This was performed to check if the trained models with the selected features could correctly label the participants as dementia patients. Finally, the changes in means of the acoustic and linguistic features for the participants of this study were investigated to examine how the changes improved after cognitive training. To this end, this research has investigated the effects of cognitive exercises on a small dataset of dementia patients and suggests linguistic and acoustic features for monitoring.

Using the Pitt dataset [33], the mean of linguistic parameters such as text length, number of verbs, rate of verbs, number of determiners, verb to noun ratio, Honoré's statistic, number of unique words, Brunet's index, verbal rate, standardized phonation time were shown to be higher for the control group. This was expected because control participants have smoother speech with more vocabulary than AD patients [6]. In accordance with literature [6], [11], AD patients used more pronouns, adjectives, and adverbs probably because word-finding difficulties prevented them from remembering the actual words or phrases [4], [22]. As such, it is recommended to utilize linguistic parameters for future studies since they seem to be capable of distinguishing dementia patients from healthy controls.

Two acoustic findings from the Pitt data [33] contradicted previous studies on acoustic changes in dementia patients. Both lengths of voiced segments per second (mean) and lengths of voiced segments per second (std) had higher values for AD patients. These observations conflicted with literature suggesting that dementia patients typically talk less and use more pauses than controls [7], [14], [55], [95]. One reason for this observation may be that AD patients may spend more time describing a word that they cannot remember, thus adding to the total voiced time ("empty speech" [156]). Also, this disagreement may indicate that, though some acoustic features are distinguishable between AD and control groups such as F0, formants, fractal dimensions, lengths of voiced or unvoiced segments per second, equivalent sound level, harmonic differences,

jitter, shimmer, HNR, Hammarberg index, loudness, MFCC, spectral flux, spectral slope, and alpha ratio, they do not necessarily lead to meaningful interpretations.

AD and depression have a bi-directional relationship, with almost half of AD patients also presenting with depression [157] [158]. However, rates of depression are not always detectable by psychological tests, thus, additional diagnostic features are considered [159]. One such feature for consideration in depression diagnosis is voice quality [159]. In a recent study, for instance, Silva et al. [160] recruited 54 patients with depression and 90 control participants (aged between 18 and 60). Participants were recorded while answering some assessment questions. The conclusion was that the patients with depression had higher values for std of F0, jitter, and shimmer and lower glottal to noise excitation ratio and cepstral peak prominence-smoothed than control participants. The authors also found jitter and shimmer were significantly associated with the Beck Depression Inventory-Second Edition score [161]. In another study [162], authors extracted a combination of features related to formant frequencies and delta-Mel cepstral coefficients from 292 depressed people. The features had 80% correlation with the Beck scores. Also, parameters such as loudness, zero-crossing rate, energy, spectral flux, entropy, MFCC 1-16, variance, skewness, kurtosis, harmonicity, flatness, psychoacoustic sharpness, F0, probability of voicing, jitter, and logarithmic HNR have been shown to be able to predict Beck scores with low errors [163]. In the present study, depression was evaluated with Montgomery-Asberg Depression Rating Scale [164], and 3 participants had mild depression at different sessions. Thus, it can be expected that comorbid depression may have altered the acoustic features extracted in this study such as F0, jitter, shimmer, loudness, zero-crossing rate, energy, spectral flux, entropy, MFCC, and HNR.

The results from the machine learning algorithms indicated that the models learned the speaking characteristics of dementia patients to an acceptable extent. However, it must be noted that the dataset in this study has a very small sample size. Thus, it is hard to derive conclusions from it. The participants in the Pitt dataset [33] were recorded while doing a picture description task while the participants of the present study were asked to recall 2 short stories. To recall the stories, access to short-term memory and activation of brain regions responsible for speaking and thinking are required. Maybe if both datasets had the same task, the results would have been more reliable. In addition, the outcomes were different with features extracted from LM I and LM II sections, which again shows the choice of the tasks matter. Moreover, since features such as F0 (the range of the 20^{th} to 80^{th} percentile), loudness (CoV), MFCC 4 (mean), MFCC 4 (CoV),

shimmer (mean), shimmer (CoV), H1-H2 harmonic difference (mean), A1-H3 harmonic difference (mean), first formant frequency (CoV), spectral flux (voiced, mean), pseudo syllable rate, number of verbs, rate of verbs, rate of adjectives, adjective to noun ratio, number of unique words, Brunet's index, verbal rate, and standardized phonation time, and age were recognized as more important in the feature selection algorithm, more attention should be paid to them in future studies.

The correlation coefficients indicated that changes in some acoustic and linguistic parameters had high correlations with changes in WMS-IV scores when extracted from LM I and LM II sections. For instance, the change in F0 (CoV) had a high correlation with the change in WMS-IV total score when extracted from LM I and LM II sections. Also, the change in voiced duration had high correlations with the changes in WMS-IV total score and LM II score when extracted from LM I and LM II sections, respectively. However, the rest of the highly correlated features were not uniformly repeated in all scenarios. This could be because of the small sample size and the fact that LM II also has a memory. With a larger population, participants should be recorded while performing other short speaking tasks to check whether the features are task-dependant. If the features continue having high and significant correlations with WMS-IV scores, there will not be a need for lengthy assessments, and patients' cognitive status can be estimated from their speaking characteristics.

It was interesting to observe that during LM I section of the WMS-IV assessment, mean of H linguistic features increased while the mean of D linguistic features decreased. The improvement in H linguistic features cannot be due to practice effect because the follow-up mean was lower than baseline and post-intervention means. Also, a pilot study [122] has shown that there is no significant practice effect for performing the WMS assessment in older adults. Thus, the observed improvement in H linguistic features can be an indicator that the linguistic knowledge of the participants in the Exr group became closer to cognitively healthy older adults and improved. That said, this trend was not observed during LM II. One reason could be that LM II requires more memory demand since it is a delayed recall. As a result, it may not be an appropriate task for investigating speaking abilities. Moreover, this trend was not observed for acoustic features. It must be considered that physiological changes in the vocal tract during aging affect the quality of voice [165]. As such, the reason for not observing improvements in the acoustic features may be the fact that cognitive trainings cannot reverse the biological effects of aging.

Though some patients in the Exr group showed improvements in WMS-IV scores after intervention, the effects often did exhibit long-lasting results for most of the AD participants. MCI and FTD patients progressed more because the rates of cognitive decline were lower for them. Perhaps the solution for AD patients is to practice the exercises from early stages of the disease and more frequently and for longer durations. It should be investigated whether longer durations of interventions can lead to more improvements without becoming tedious for the participants.

The main drawback of this study is the small sample size. Because of the small number of participants, the outcomes are not reliable. Besides, some degree of error might have occurred in manual transcriptions of the interviews. Furthermore, it must be noted that English was not the native language for one of the participants (no. 5) and this may have affected the linguistic measures. Also, as the assessments were conducted online, inherent errors in WMS-IV scores may be attributed to technical difficulties and internet interruptions. In a review of 11 studies, Pang et al. [166] concluded that online memory exercises improved memory measures with moderate effect sizes. Despite this finding, another study [167] indicated that MCI patients were less likely to participate and complete online cognitive trainings, suggesting that they might need more support. Even though online trainings are more convenient and cost-effective, usability concerns may prevent the patients from participating [168].

Chapter 6: Future Work and Conclusion

For future studies, a larger sample size and in-person assessments and training are encouraged. Also, it is recommended to recruit participants with only one sub-type of dementia (such as AD) to achieve more uniform results. It would be advantageous to have a healthy control group as well. This way, with enough data, there would be no need for an auxiliary dataset to train the machine learning algorithms on and there would be no differences between the tasks in the training and testing datasets. For classification, more advanced deep learning algorithms such as deep neural networks and recurrent neural networks with different hidden layers can be tested (with a larger sample size) to achieve higher accuracies. Furthermore, the dependency of the features to different oral tasks should be further investigated. Therefore, dementia and control participants should be recorded while doing different oral tasks. Afterwards, it can be inspected if the features are still significantly different between the two groups. Moreover, it is important to pay attention to the quality of recoded samples. Noisy signals would need more denoising, which could destruct the raw signal and acoustic features can be altered by this. Besides, it would be helpful to have access to patients' MRI scans before and after intervention to see how the exercises affect the brain structures. Additionally, it would be beneficial to combine cognitive exercises with other methods that ameliorate AD symptoms, such as Transcranial Alternating Current Stimulation. In fact, recent work has supported complimentary treatment programs, with one such study observing long-term cognitive improvement [120].

This study once again supports the fact that cognitive trainings have some benefits for dementia patients. This research focused on speaking abilities and LM I, LM II, and WMS-IV total scores as outcome measures. The results suggest that on average, linguistic abilities of dementia patients improve with cognitive training and become closer to linguistic abilities of cognitively healthy older adults. Additionally, changes in some features such as F0, MFCC 1, Katz fractal dimension, number of zero-crossings, voiced duration, spectral entropy, harmonic differences, first formant bandwidth, HNR, Brunet's index, text length, rate of conjunctions, and number of words, nouns, determiners, and unique words showed to have high correlations with changes in WMS-IV scores. It is expected that the features introduced in this study can be beneficial for future research on speech analysis in dementia and early detection of dementia.

References:

[1] "2021 Alzheimer's disease facts and figures," *Alzheimer's and Dementia*, vol. 17, no. 3, pp. 327–406, Mar. 2021, doi: 10.1002/alz.12328.

[2] World Health Organization (WHO), "Dementia," Sep. 02, 2021.

[3] S. Gauthier, P. Rosa-Neto, J. A. Morais, and C. Webster, "World Alzheimer Report 2021: Journey through the diagnosis of dementia."

[4] L. Mucke, "Neuroscience Alzheimer's disease," *Nature (London)*, vol. 461, no. 7266, pp. 895–897, 2009, doi: 10.1038/461895a.

[5] L. W. Chambers, C. Bancej, and I. Mcdowell, "Report summary Prevalence and monetary costs of dementia in Canada (2016): a report by the Alzheimer Society of Canada," *Health promotion and chronic disease prevention in Canada*, vol. 36, no. 10, pp. 231–232, 2016, doi: 10.24095/hpcdp.36.10.04.

[6] R. S. Bucks, S. Singh, J. M. Cuerden, and G. K. Wilcock, "Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance," *Aphasiology*, vol. 14, no. 1, pp. 71–91, 2000, doi: 10.1080/026870300401603.

[7] F. Martínez-Sánchez, J. J. G. Meilán, J. García-Sevilla, J. Carro, and J. M. Arana, "Oral reading fluency analysis in patients with Alzheimer disease and asymptomatic control subjects," *Neurología (English Edition)*, vol. 28, no. 6, pp. 325–331, 2013, doi: 10.1016/j.nrleng.2012.07.017.

[8] J. J. G. Meilán, F. Martínez-Sánchez, J. Carro, D. E. López, L. Millian-Morell, and J. M. Arana, "Speech in alzheimer's disease: Can temporal and acoustic parameters discriminate dementia?," *Dementia and Geriatric Cognitive Disorders*, vol. 37, no. 5–6, pp. 327–334, 2014, doi: 10.1159/000356726.

[9] K. C. Fraser, J. A. Meltzer, and F. Rudzicz, "Linguistic features identify Alzheimer's disease in narrative speech," *Journal of Alzheimer's Disease*, vol. 49, no. 2, pp. 407–422, 2015, doi: 10.3233/JAD-150520.

[10] A. Khodabakhsh and C. Demiroglu, "Analysis of speech-based measures for detecting and monitoring Alzheimer's disease," *Methods in Molecular Biology*, vol. 1246, pp. 159–173, 2015, doi: 10.1007/978-1-4939-1985-7_11.

[11] A. Khodabakhsh, F. Yesil, E. Guner, and C. Demiroglu, “Evaluation of linguistic and prosodic features for detection of Alzheimer’s disease in Turkish conversational speech,” *Eurasip Journal on Audio, Speech, and Music Processing*, vol. 2015, no. 1, Dec. 2015, doi: 10.1186/s13636-015-0052-y.

[12] A. König *et al.*, “Automatic speech analysis for the assessment of patients with predementia and Alzheimer’s disease,” *Alzheimer’s and Dementia: Diagnosis, Assessment and Disease Monitoring*, vol. 1, no. 1, pp. 112–124, 2015, doi: 10.1016/j.dadm.2014.11.012.

[13] K. López-De-Ipiña *et al.*, “Feature selection for spontaneous speech analysis to aid in Alzheimer’s disease diagnosis: A fractal dimension approach,” *Computer Speech and Language*, vol. 30, no. 1, pp. 43–60, 2015, doi: 10.1016/j.csl.2014.08.002.

[14] K. López-de-Ipiña *et al.*, “On Automatic Diagnosis of Alzheimer’s Disease Based on Spontaneous Speech Analysis and Emotional Temperature,” *Cognitive Computation*, vol. 7, no. 1, pp. 44–55, Feb. 2015, doi: 10.1007/s12559-013-9229-9.

[15] U. Martinez de Lizarduy, P. Calvo Salomón, P. Gómez Vilda, M. Ecay Torres, and K. López de Ipiña, “ALZUMERIC: A decision support system for diagnosis and monitoring of cognitive impairment,” *Loquens*, vol. 4, no. 1, p. 037, Dec. 2017, doi: 10.3989/loquens.2017.037.

[16] F. Martínez-Sánchez *et al.*, “Speech rhythm alterations in Spanish-speaking individuals with Alzheimer’s disease,” *Aging, Neuropsychology, and Cognition*, vol. 24, no. 4, pp. 418–434, 2017, doi: 10.1080/13825585.2016.1220487.

[17] C. Themistocleous, M. Eckerström, and D. Kokkinakis, “Identification of Mild Cognitive Impairment From Speech in Swedish Using Deep Sequential Neural Networks,” *Frontiers in Neurology*, vol. 9, pp. 1–10, 2018, doi: 10.3389/fneur.2018.00975.

[18] Y. W. Chien, S. Y. Hong, W. T. Cheah, L. H. Yao, Y. L. Chang, and L. C. Fu, “An Automatic Assessment System for Alzheimer’s Disease Based on Speech Using Feature Sequence Generator and Recurrent Neural Network,” *Scientific Reports*, vol. 9, no. 1, Dec. 2019, doi: 10.1038/s41598-019-56020-x.

[19] K. C. Fraser, K. Lundholm Fors, and D. Kokkinakis, “Multilingual word embeddings for the assessment of narrative speech in mild cognitive impairment,” *Computer Speech and Language*, vol. 53, pp. 121–139, 2019, doi: 10.1016/j.csl.2018.07.005.

[20] G. Gosztolya, V. Vincze, L. Tóth, M. Pákáski, J. Kálmán, and I. Hoffmann, "Identifying Mild Cognitive Impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features," *Computer Speech and Language*, vol. 53, pp. 181–197, Jan. 2019, doi: 10.1016/j.csl.2018.07.007.

[21] L. Liu, S. Zhao, H. Chen, and A. Wang, "A new machine learning method for identifying Alzheimer's disease," *Simulation Modelling Practice and Theory*, vol. 99, Feb. 2020, doi: 10.1016/j.simpat.2019.102023.

[22] K. López-de-Ipiña *et al.*, "On the analysis of speech and disfluencies for automatic detection of Mild Cognitive Impairment," *Neural Computing and Applications*, vol. 32, no. 20, pp. 15761–15769, Oct. 2020, doi: 10.1007/s00521-018-3494-1.

[23] R. Nagumo *et al.*, "Automatic Detection of Cognitive Impairments through Acoustic Analysis of Speech," *Current Alzheimer Research*, vol. 17, no. 1, pp. 60–68, 2020, doi: 10.2174/1567205017666200213094513.

[24] K. López-de-Ipiña *et al.*, "On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis," *Sensors (Switzerland)*, vol. 13, no. 5, pp. 6730–6745, 2013, doi: 10.3390/s130506730.

[25] P. Gómez-Vilda *et al.*, "Phonation biomechanic analysis of Alzheimer's disease cases," *Neurocomputing (Amsterdam)*, vol. 167, pp. 83–93, 2015, doi: 10.1016/j.neucom.2015.03.087.

[26] L. Toth *et al.*, "A Speech Recognition-based Solution for the Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech," *Current Alzheimer Research*, vol. 15, no. 2, pp. 130–138, Nov. 2017, doi: 10.2174/1567205014666171121114930.

[27] D. Neary, J. Snowden, and D. Mann, "Frontotemporal Dementia," *Lancet neurology*, vol. 4, no. 11, pp. 771–780, 2005, doi: 10.1016/S1474-4422(05)70223-4.

[28] J. Bang, S. Spina, and B. L. Miller, "Frontotemporal dementia," *The Lancet (British edition)*, vol. 386, no. 10004, pp. 1672–1682, 2015, doi: 10.1016/S0140-6736(15)00461-4.

[29] C. N. Harada, M. C. Natelson Love, and K. L. Triebel, "Normal cognitive aging," *Clinics in Geriatric Medicine*, vol. 29, no. 4, pp. 737–752, Nov. 2013, doi: 10.1016/J.CGER.2013.07.002.

[30] M. S. Albert, "Changes in cognition," *Neurobiology of Aging*, vol. 32, no. 1, pp. S58–S63, Dec. 2011, doi: 10.1016/j.neurobiolaging.2011.09.010.

[31] D. C. Park and G. N. Bischof, “The aging mind: neuroplasticity in response to cognitive training,” *Dialogues Clin Neurosci*, vol. 15. pp. 109–119, 2013. [Online]. Available: www.dialogues-cns.org13_AG_1004_BA_INTERIEUR.qxd:DCNS#551/03/13

[32] C. Basak, W. R. Boot, M. W. Voss, and A. F. Kramer, “Can Training in a Real-Time Strategy Video Game Attenuate Cognitive Decline in Older Adults?,” *Psychology and Aging*, vol. 23, no. 4, pp. 765–777, Dec. 2008, doi: 10.1037/a0013494.

[33] J. T. Becker, F. Boller, O. L. Lopez, J. Saxton, and K. L. McGonigle, “The natural history of Alzheimer’s disease: description of study cohort and accuracy of diagnosis.,” *Archives of neurology (Chicago)*, vol. 51, no. 6, pp. 585–594, 1994, doi: 10.1001/archneur.1994.00540180063015.

[34] D. Wechsler, “WMS-IV: Wechsler memory scale-fourth edition.” Tex: Pearson, San Antonio, 2009.

[35] “Brain Anatomy and How the Brain Works,” *Johns Hopkins Medicine*. https://www.hopkinsmedicine.org/health/conditions-and-diseases/anatomy-of-the-brain#:~:text=The%20largest%20part%20of%20the,hearing%2C%20touch%20and%20other%20senses. (accessed Mar. 23, 2022).

[36] R. D. Rogers and S. Monsell, “Costs of a Predictable Switch Between Simple Cognitive Tasks,” *Journal of Experimental Psychology: General*, vol. 124, no. 2, pp. 207–231, 1995, doi: 10.1037/0096-3445.124.2.207.

[37] E. Koechlin and C. Summerfield, “An information theoretical approach to prefrontal executive function,” *Trends in Cognitive Sciences*, vol. 11, no. 6, pp. 229–235, Jun. 2007, doi: 10.1016/j.tics.2007.04.005.

[38] A. L. Benton, “Differential behavioral effects in frontal lobe disease,” *Neuropsychologia*, vol. 6, no. 1, pp. 53–60, 1968, doi: 10.1016/0028-3932(68)90038-9.

[39] C. Chayer and M. Freedman, “Frontal Lobe Functions,” *Current Neurology and Neuroscience Reports*, vol. 1, no. 6, pp. 547–552, 2001, doi: 10.1007/s11910-001-0060-4.

[40] M. D’Esposito, B. R. Postle, and B. Rypma, “Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies,” *Experimental brain research*, vol. 133, no. 1, pp. 3–11, 2000, doi: 10.1007/s002210000395.

[41] D. T. Stuss, G. G. Gallup, and M. P. Alexander, “The frontal lobes are necessary for ‘theory of mind,’” *Brain*, vol. 124, no. 2, pp. 279–286, 2001, doi: 10.1093/brain/124.2.279.

[42] J. A. Gottfried, “Smell: central nervous processing,” *Advances in Otorhinolaryngology*, vol. 63, pp. 44–69, 2006, doi: 10.1159/000093750.

[43] P. H. Weiss, C. Dohle, F. Binkofski, A. Schnitzler, H.-J. Freund, and H. Hefter, “Motor impairment in patients with parietal lesions: disturbances of meaningless arm movement sequences,” *Neuropsychologia*, vol. 39, no. 4, pp. 397–405, 2001, doi: 10.1016/S0028-3932(00)00129-9.

[44] L. Fogassi and G. Luppino, “Motor functions of the parietal lobe,” *Current Opinion in Neurobiology*, vol. 15, no. 6. pp. 626–631, Dec. 2005. doi: 10.1016/j.conb.2005.10.015.

[45] N. S. Dhanjal, L. Handunnetthi, M. C. Patel, and R. J. S. Wise, “Perceptual Systems Controlling Speech Production,” *The Journal of neuroscience*, vol. 28, no. 40, pp. 9969–9975, 2008, doi: 10.1523/JNEUROSCI.2607-08.2008.

[46] S. L. E. Brownsett and R. J. S. Wise, “The contribution of the parietal lobes to speaking and writing,” *Cerebral Cortex*, vol. 20, no. 3, pp. 517–523, Mar. 2010, doi: 10.1093/cercor/bhp120.

[47] C. M. Klingner and O. W. Witte, “Somatosensory deficits,” in *Handbook of Clinical Neurology*, vol. 151, Elsevier, 2018, pp. 185–206. doi: https://doi.org/10.1016/B978-0-444-63622-5.00009-7.

[48] B. C. Dickerson and H. Eichenbaum, “The episodic memory system: Neurocircuitry and disorders,” *Neuropsychopharmacology*, vol. 35, no. 1. pp. 86–104, Jan. 2010. doi: 10.1038/npp.2009.126.

[49] L. Jäncke, F. Liem, and S. Merillat, “Are language skills related to structural features in Broca’s and Wernicke’s area?,” *European Journal of Neuroscience*, vol. 53, no. 4, pp. 1124–1135, Feb. 2021, doi: 10.1111/ejn.15038.

[50] J. Reilly, J. Troche, and M. Grossman, “Language Processing in Dementia,” in *The Handbook of Alzheimer’s Disease and Other Dementias*, 2011, pp. 336–368. doi: 10.1002/9781444344110.ch12.

[51] D. Pins and D. ffytche, “The Neural Correlates of Conscious Vision,” *Cerebral cortex (New York, N.Y. 1991)*, vol. 13, no. 5, pp. 461–474, 2003, doi: 10.1093/cercor/13.5.461.

[52] Sebastian023, “‘Lateral Lobes,’” *Creative Commons-Licensed under CC BY-SA 3.0.*.

[53] R. C. Petersen, “clinical practice Mild Cognitive Impairment,” *The New England journal of medicine*, vol. 364, no. 23, pp. 2227–2261, 2011, doi: 10.1056/NEJMcp0910237.

[54] C. R. Jack *et al.*, "NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease," *Alzheimer's and Dementia*, vol. 14, no. 4, pp. 535–562, Apr. 2018, doi: 10.1016/j.jalz.2018.02.018.

[55] S. Singh, R. S. Bucks, and J. M. Cuerden, "Evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech," *Aphasiology*, vol. 15, no. 6, pp. 571–583, 2001, doi: 10.1080/02687040143000041.

[56] J. K. Brian Roark, Margaret Mitchell, John-Paul Hosom, Kristy Hollingshead, "Spoken Language Derived Measures for Detecting Mild Cognitive Impairment," in *IEEE Trans Audio Speech Lang Processing*, 2011, vol. 19, no. 7, pp. 2081–2090. doi: 10.1109/TASL. 2011.2112351.

[57] C. Crockford and R. Lesser, "Assessing functional communication in aphasia: Clinical utility and time demands of three methods," *International Journal of Language & Communication Disorders*, vol. 29, no. 2, pp. 165–182, 1994, doi: 10.3109/13682829409041490.

[58] A. Burns *et al.*, "Defining Treatment Response to Donepezil in Alzheimer's Disease: Responder Analysis of Patient-Level Data from Randomized, Placebo-Controlled Studies," *Drugs & Aging*, vol. 25, no. 8, pp. 707–714, 2008, doi: 10.2165/00002512-200825080-00007.

[59] D. P. Perl, "Neuropathology of Alzheimer's disease," *Mount Sinai Journal of Medicine*, vol. 77, no. 1. pp. 32–42, Jan. 2010. doi: 10.1002/msj.20157.

[60] M. A. Deture and D. W. Dickson, "The neuropathological diagnosis of Alzheimer's disease," *Molecular Neurodegeneration*, vol. 14, no. 1. BioMed Central Ltd., Aug. 02, 2019. doi: 10.1186/s13024-019-0333-5.

[61] L. Rami *et al.*, "Distinct functional activity of the precuneus and posterior cingulate cortex during encoding in the preclinical stage of Alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 31, no. 3, pp. 517–526, 2012, doi: 10.3233/JAD-2012-120223.

[62] A. Serrano-Pozo, M. P. Frosch, E. Masliah, and B. T. Hyman, "Neuropathological alterations in Alzheimer disease," *Cold Spring Harbor Perspectives in Medicine*, vol. 1, no. 1, Sep. 2011, doi: 10.1101/cshperspect.a006189.

[63] L. G. Apostolova *et al.*, “Hippocampal atrophy and ventricular enlargement in normal aging, mild cognitive impairment (MCI), and Alzheimer Disease,” *Alzheimer Dis Assoc Disord*, vol. 26, no. 1, pp. 17–27, 2012, doi: 10.1097/WAD.0b013e3182163b62.

[64] B. Opitz, “Memory function and the hippocampus,” *The Hippocampus in Clinical Neuroscience*, vol. 34, pp. 51–59, Apr. 2014, doi: 10.1159/000356422.

[65] R. C. Petersen, “Mild cognitive impairment as a diagnostic entity,” *J Intern Med*, vol. 256, no. 3, pp. 183–194, 2004, doi: 10.1111/j.1365-2796.2004.01388.x.

[66] S. Gauthier *et al.*, “Mild cognitive impairment,” *The Lancet*, vol. 367, no. 9518, Apr. 2006, doi: 10.1016/S0140-6736(06)68542-5.

[67] R. A. Dunne *et al.*, “Mild Cognitive Impairment: The Manchester consensus,” *Age and Ageing*, vol. 50, no. 1. Oxford University Press, pp. 72–80, Jan. 01, 2021. doi: 10.1093/ageing/afaa228.

[68] R. C. Petersen *et al.*, “Vitamin E and Donepezil for the Treatment of Mild Cognitive Impairment,” *N Engl J Med*, vol. 352, no. 23, pp. 2379–2388, 2005, doi: 10.1056/NEJMoa050151.

[69] L. K. McEvoy *et al.*, “Alzheimer disease: Quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment,” *Radiology*, vol. 251, no. 1, pp. 195–205, Apr. 2009, doi: 10.1148/radiol.2511080924.

[70] C. R. Jack *et al.*, “Brain beta-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to Alzheimer’s disease,” *Brain*, vol. 133, no. 11, pp. 3336–3348, Nov. 2010, doi: 10.1093/brain/awq277.

[71] B. C. Dickerson, R. A. Sperling, B. T. Hyman, M. S. Albert, and D. Blacker, “Clinical Prediction of Alzheimer Disease Dementia Across the Spectrum of Mild Cognitive Impairment,” *Arch Gen Psychiatry*, vol. 64, no. 12, pp. 1443–1450, 2007, doi: 10.1001/archpsyc.64.12.1443.

[72] O. Hansson, H. Zetterberg, P. Buchhave, E. Londos, K. Blennow, and L. Minthon, “Association between CSF biomarkers and incipient Alzheimer’s disease in patients with mild cognitive impairment: a follow-up study.,” *Lancet neurology*, vol. 5, no. 3, pp. 228–234, 2006, doi: 10.1016/S1474-4422(06)70355-6.

[73] D. A. Wolk *et al.*, “Amyloid imaging in mild cognitive impairment subtypes,” *Annals of Neurology*, vol. 65, no. 5, pp. 557–568, May 2009, doi: 10.1002/ana.21598.

[74] B. Winblad *et al.*, “Mild cognitive impairment-beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment,” *J Intern Med*, vol. 256, no. 3, pp. 240–246, 2004, doi: 10.1111/j.1365-2796.2004.01380.x.

[75] A. R. Varma *et al.*, “Diagnostic patterns of regional atrophy on MRI and regional cerebral,” *Acta Neurol Scand*, vol. 105, no. 4, pp. 261–269, 2002, doi: 10.1034/j.1600-0404.2002.1o148.x.

[76] E. Salmon *et al.*, “Predominant ventromedial frontopolar metabolic impairment in frontotemporal dementia,” *Neuroimage*, vol. 20, no. 1, pp. 435–440, 2003, doi: 10.1016/S1053-8119(03)00346-X.

[77] S. Lough, C. Gregory, and J. R. Hodges, “Dissociation of social cognition and executive function in frontal variant frontotemporal dementia,” *Neurocase*, vol. 7, no. 2. Oxford University Press, pp. 123–130, 2001. doi: 10.1093/neucas/7.2.123.

[78] D. Neary *et al.*, “Frontotemporal lobar degeneration: A consensus on clinical diagnostic criteria,” *Neurology*, vol. 51, no. 6, pp. 1546–1554, 1998, doi: 10.1212/WNL.51.6.1546.

[79] S. M. Rosso *et al.*, “Frontotemporal dementia in The Netherlands: Patient characteristics and prevalence estimates from a population-based study,” *Brain*, vol. 126, no. 9, pp. 2016–2022, Sep. 2003, doi: 10.1093/brain/awg204.

[80] E. Ratnavalli, C. Brayne, K. Dawson, and J. R. Hodges, “The prevalence of frontotemporal dementia,” *Neurology*, vol. 58, no. 11, pp. 1615–1621, 2002, doi: 10.1212/WNL.58.11.1615.

[81] T. B. Gislason, M. Sjögren, L. Larsson, and I. Skoog, “The prevalence of frontal variant frontotemporal dementia and the frontal lobe syndrome in a population based sample of 85 year olds,” *Journal of Neurology Neurosurgery and Psychiatry*, vol. 74, no. 7, pp. 867–871, Jul. 2003, doi: 10.1136/jnnp.74.7.867.

[82] J. S. Snowden, D. Neary, and D. M. A. Mann, “Autopsy proven sporadic frontotemporal dementia due to microvacuolar-type histology, with onset at 21 years of age,” *Journal of Neurology, Neurosurgery and Psychiatry*, vol. 75, no. 9, pp. 1337–1339, 2004, doi: 10.1136/jnnp.2003.028498.

[83] J. D. Rohrer *et al.*, “The heritability and genetics of frontotemporal lobar degeneration,” *Neurology*, vol. 73, no. 18, pp. 1451–1456, 2009, doi: 10.1212/WNL.0b013e3181bf997a.

[84] P. Mahajan and V. Baths, “Acoustic and Language Based Deep Learning Approaches for Alzheimer’s Dementia Detection From Spontaneous Speech,” *Frontiers in Aging Neuroscience*, vol. 13, no. February, pp. 1–11, 2021, doi: 10.3389/fnagi.2021.623607.

[85] A. Roshanzamir, H. Aghajan, and M. Soleymani Baghshah, “Transformer-based deep neural network language models for Alzheimer’s disease risk assessment from targeted speech,” *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, pp. 1–14, 2021, doi: 10.1186/s12911-021-01456-3.

[86] P. Boersma, “Praat, a system for doing phonetics by computer.,” *Glot International*, vol. 5, no. 9/10, pp. 341–345, 2001.

[87] F. Eyben, M. Wöllmer, and B. Schuller, “OpenSMILE - The Munich versatile and fast open-source audio feature extractor,” in *MM’10 - Proceedings of the ACM Multimedia 2010 International Conference*, 2010, pp. 1459–1462. doi: 10.1145/1873951.1874246.

[88] S. Bird, E. Loper, and E. Klein, “Natural Language Processing with Python.” O’Reilly Media Inc, 2009.

[89] K. Fraser, F. Rudzicz, N. Graham, and E. Rochon, “Automatic speech recognition in the diagnosis of primary progressive aphasia,” 2013.

[90] D. Beltrami, G. Gagliardi, R. R. Favretti, E. Ghidoni, F. Tamburini, and L. Calzà, “Speech analysis by natural language processing techniques: A possible tool for very early detection of cognitive decline?,” *Frontiers in Aging Neuroscience*, vol. 10, 2018, doi: 10.3389/fnagi.2018.00369.

[91] H. Goodglass and E. Kaplan, “The Boston Diagnostic Aphasia Examination.” 1983.

[92] A. L. R. Adlam, S. Bozeat, R. Arnold, P. Watson, and J. R. Hodges, “Semantic knowledge in mild cognitive impairment and mild Alzheimer’s disease,” *Cortex*, vol. 42, no. 5, pp. 675–684, 2006, doi: 10.1016/S0010-9452(08)70404-0.

[93] S. V. S. Pakhomov *et al.*, “A computerized technique to assess language use patterns in patients with frontotemporal dementia,” *Journal of Neurolinguistics*, vol. 23, no. 2, pp. 127–144, Mar. 2010, doi: 10.1016/j.jneuroling.2009.12.001.

[94] S. Cho *et al.*, “Automated analysis of lexical features in frontotemporal degeneration,” *Cortex*, vol. 137, pp. 215–231, Apr. 2021, doi: 10.1016/j.cortex.2021.01.012.

[95] I. Hoffmann, D. Nemeth, C. D. Dye, M. Pákáski, T. Irinyi, and J. Kálmán, “Temporal parameters of spontaneous speech in Alzheimer’s disease,” *International Journal of*

Speech-Language Pathology, vol. 12, no. 1, pp. 29–34, 2010, doi: 10.3109/17549500903137256.

[96] Cuetos-Vega F., Menéndez-González M., and Calatayud-Noguera T., "Descripción de un nuevo test para la detección precoz de la enfermedad de Alzheimer," *Rev Neurol*, vol. 44, pp. 469–474, 2007, doi: https://doi.org/10.33588/rn.4408.2006403.

[97] J. J. G. Meilán, F. Martínez-Sánchez, J. Carro, J. A. Sánchez, and E. Pérez, "Acoustic Markers Associated with Impairment in Language Processing in Alzheimer's Disease," *Span J Psychol*, vol. 15, no. 2, pp. 487–494, Jul. 2012, doi: 10.5209/rev_sjop.2012.v15.n2.38859.

[98] J. J. G. Meilán, F. Martínez-Sánchez, I. Martínez-Nicolás, T. E. Llorente, and J. Carro, "Changes in the Rhythm of Speech Difference between People with Nondegenerative Mild Cognitive Impairment and with Preclinical Dementia," *Behavioural Neurology*, vol. 2020, 2020, doi: 10.1155/2020/4683573.

[99] S. V. S. Pakhomov *et al.*, "Computerized analysis of speech and language to identify psycholinguistic correlates of frontotemporal lobar degeneration," *Cognitive and Behavioral Neurology*, vol. 23, no. 3, pp. 165–177, Sep. 2010, doi: 10.1097/WNN.0b013e3181c5dde3.

[100] K. Horley, A. Reid, and D. Burnham, "Emotional prosody perception and production in Dementia of the Alzheimer's type," *Journal of Speech, Language, and Hearing Research*, vol. 53, no. 5, pp. 1132–1146, 2010, doi: 10.1044/1092-4388(2010/09-0030).

[101] S. Al-Hameed, M. Benaissa, H. Christensen, B. Mirheidari, D. Blackburn, and M. Reuber, "A new diagnostic approach for the identification of patients with neurodegenerative cognitive complaints," *PLoS ONE*, vol. 14, no. 5, May 2019, doi: 10.1371/journal.pone.0217388.

[102] K. Lopez-de-Ipiña *et al.*, "Spontaneous speech and emotional response modeling based on one-class classifier oriented to Alzheimer disease diagnosis," *IFMBE Proceedings*, vol. 41, pp. 571–574, 2014, doi: 10.1007/978-3-319-00846-2_141.

[103] V. Rentoumi, L. Raoufian, S. Ahmed, C. A. de Jager, and P. Garrard, "Features and machine learning classification of connected speech samples from patients with autopsy proven Alzheimer's disease with and without additional vascular pathology," *Journal of Alzheimer's Disease*, vol. 42, pp. S3–S17, 2014, doi: 10.3233/JAD-140555.

[104] M. Nasrolahzadeh, Z. Mohammadpoory, and J. Haddadnia, “A novel method for early diagnosis of Alzheimer’s disease based on higher-order spectral estimation of spontaneous speech signals,” *Cognitive Neurodynamics*, vol. 10, no. 6, pp. 495–503, Dec. 2016, doi: 10.1007/s11571-016-9406-0.

[105] “CITA-Alzheimer Foundation, PGA project.” http://www.cita-alzheimer.org/investigacion/proyectos (accessed Apr. 30, 2022).

[106] A. Wallin *et al.*, “The Gothenburg MCI study: Design and distribution of Alzheimer’s disease and subcortical vascular disease diagnoses from baseline to 6-year follow-up,” *Journal of Cerebral Blood Flow and Metabolism*, vol. 36, no. 1. Nature Publishing Group, pp. 114–131, Jan. 01, 2016. doi: 10.1038/jcbfm.2015.147.

[107] K. Cromnow and T. Landberg, “Skriftliga beskrivningar av bilden Kakstölden.” 2009.

[108] S. Luz, F. Haider, S. de la Fuente, D. Fromm, and B. MacWhinney, “Alzheimer’s Dementia Recognition through Spontaneous Speech: The ADReSS Challenge,” Apr. 2020, [Online]. Available: http://arxiv.org/abs/2004.06833

[109] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.” [Online]. Available: https://github.com/tensorflow/tensor2tensor

[110] D. C. Park, A. H. Gutchess, M. L. Meade, and E. A. L. Stine-Morrow, “Improving cognitive function in older adults: Nontraditional approaches,” *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, vol. 62, no. SPEC. ISSUE 1, pp. 45–52, 2007, doi: 10.1093/geronb/62.special_issue_1.45.

[111] Y. Stern, “What is cognitive reserve? Theory and research application of the reserve concept,” *Journal of the International Neuropsychological Society*, vol. 8, no. 3, pp. 448–460, 2002, doi: 10.1017/S1355617702813248.

[112] D. C. Park and P. Reuter-Lorenz, “The adaptive brain: Aging and neurocognitive scaffolding,” *Annual Review of Psychology*, vol. 60. pp. 173–196, Jan. 2009. doi: 10.1146/annurev.psych.59.103006.093656.

[113] X. Meng and C. D’Arcy, “Education and dementia in the context of the cognitive reserve hypothesis: A systematic review with meta-analyses and qualitative analyses,” *PLoS ONE*, vol. 7, no. 6, Jun. 2012, doi: 10.1371/journal.pone.0038268.

[114] A. M. Tucker and Y. Stern, “Cognitive reserve in aging,” *Curr Alzheimer Res*, vol. 8, no. 4, pp. 354–360, 2011, doi: 10.2174/156720511795745320.

[115] A. H. Gutchess *et al.*, “Aging and the Neural Correlates of Successful Picture Encoding: Frontal Activations Compensate for Decreased Medial-Temporal Activity,” *Journal of cognitive neuroscience*, vol. 17, no. 1. pp. 84–96, 2005. [Online]. Available: http://mitprc.silverchair.com/jocn/article-pdf/17/1/84/1757034/0898929052880048.pdf

[116] N. Raz *et al.*, “Regional brain changes in aging healthy adults: General trends, individual differences and modifiers,” *Cerebral Cortex*, vol. 15, no. 11, pp. 1676–1689, Nov. 2005, doi: 10.1093/cercor/bhi044.

[117] K. Rodrigue *et al.*, “Amyloid burden in healthy aging Regional distribution and cognitive consequences,” *Neurology*, vol. 78, no. 6. pp. 387–395, 2012. [Online]. Available: www.neurology.org

[118] J. Karbach and P. Verhaeghen, “Making Working Memory Work: A Meta-Analysis of Executive-Control and Working Memory Training in Older Adults,” *Psychological Science*, vol. 25, no. 11, pp. 2027–2037, 2014, doi: 10.2307/26661347.

[119] S. C. Li, F. Schmiedek, O. Huxhold, C. Röcke, J. Smith, and U. Lindenberger, “Working Memory Plasticity in Old Age: Practice Gain, Transfer, and Maintenance,” *Psychology and Aging*, vol. 23, no. 4, pp. 731–742, Dec. 2008, doi: 10.1037/a0014343.

[120] Z. Moussavi, K. Kimura, L. Kehler, C. de Oliveira Francisco, and B. Lithgow, “A Novel Program to Improve Cognitive Function in Individuals With Dementia Using Transcranial Alternating Current Stimulation (tACS) and Tutored Cognitive Exercises,” *Frontiers in Aging*, vol. 2, Mar. 2021, doi: 10.3389/fragi.2021.632545.

[121] B. Stojanoski, K. M. Lyons, A. A. A. Pearce, and A. M. Owen, “Targeted training: Converging evidence against the transferable benefits of online brain training on cognitive function,” *Neuropsychologia*, vol. 117, no. July, pp. 541–550, 2018, doi: 10.1016/j.neuropsychologia.2018.07.013.

[122] M. T. Garcia-Campuzano, J. Virues-Ortega, S. Smith, and Z. Moussavi, “Effect of Cognitive Training Targeting Associative Memory in the Elderly: A Small Randomized Trial and a Longitudinal Evaluation,” *Journal of the American Geriatrics Society (JAGS)*, vol. 61, no. 12, pp. 2252–2254, 2013, doi: 10.1111/jgs.12574.

[123] A. Lampit, H. Hallock, and M. Valenzuela, "Computerized Cognitive Training in Cognitively Healthy Older Adults: A Systematic Review and Meta-Analysis of Effect Modifiers," *PLoS Medicine*, vol. 11, no. 11, 2014, doi: 10.1371/journal.pmed.1001756.

[124] J. M. Chein and A. B. Morrison, "Expanding the mind's workspace: Training and transfer effects with a complex working memory span task," *Psychonomic Bulletin and Review*, vol. 17, no. 2, pp. 193–199, Apr. 2010, doi: 10.3758/PBR.17.2.193.

[125] E. Dahlin, A. Stigsdotter Neely, A. Larsson, L. Bäckman, and L. Nyberg, "Transfer of Learning After Updating Training Mediated by the Striatum," *Science (1979)*, vol. 320, no. 5882, pp. 1510–1512, Jun. 2008, doi: 10.1126/science.1159051.

[126] K. Caeyenberghs, C. Metzler-Baddeley, S. Foley, and D. K. Jones, "Dynamics of the human structural connectome underlying working memory training," *Journal of Neuroscience*, vol. 36, no. 14, pp. 4056–4066, 2016, doi: 10.1523/JNEUROSCI.1973-15.2016.

[127] H. H. Dodge *et al.*, "Web-enabled conversational interactions as a method to improve cognitive functions: Results of a 6-week randomized controlled trial," *Alzheimer's and Dementia: Translational Research and Clinical Interventions*, vol. 1, no. 1, pp. 1–12, 2015, doi: 10.1016/j.trci.2015.01.001.

[128] M. Melby-Lervåg, T. S. Redick, and C. Hulme, "Working Memory Training Does Not Improve Performance on Measures of Intelligence or Other Measures of 'Far Transfer': Evidence From a Meta-Analytic Review," *Perspectives on Psychological Science*, vol. 11, no. 4, pp. 512–534, 2016, doi: 10.1177/1745691616635612.

[129] B. Woods, E. Aguirre, A. E. Spector, and M. Orrell, "Cognitive stimulation to improve cognitive functioning in people with dementia," *Cochrane Database of Systematic Reviews*, vol. 2012, no. 2, pp. CD005562–CD005562, Feb. 2012, doi: 10.1002/14651858.CD005562.pub2.

[130] G. Wallis, M. Baldwin, and P. Higginbotham, "Reality orientation therapy--a controlled trial.," *British Journal of Medical Psychology*, vol. 56, no. 3, pp. 271–277, 1983, doi: 10.1111/j.2044-8341.1983.tb01556.x.

[131] M. v Baldelli, A. Pirani, M. Motta, E. Abati, E. Mariani, and V. Manzi, "Effects of reality orientation therapy on elderly patients in the community," *Archives of Gerontology and Geriatrics*, vol. 17, no. 3, pp. 211–218, 1993, doi: 10.1016/0167-4943(93)90052-j.

[132] S. Baines, P. Saxby, and K. Ehlert, “Reality Orientation and Reminiscence Therapy A Controlled Cross-over Study of Elderly Confused People,” *The British Journal of Psychiatry*, vol. 151, no. 2, pp. 222–231, 1987, doi: 10.1192/bjp.151.2.222.

[133] S. B. Chapman, M. F. Weiner, A. Rackley, L. S. Hynan, and J. Zientz, “Effects of cognitive-communication stimulation for Alzheimer’s disease patients treated with donepezil,” *Journal of speech, language, and hearing research*, vol. 47, no. 5, pp. 1149–1163, 2004, doi: 10.1044/1092-4388(2004/085).

[134] J.-H. Hwang, H.-G. Cha, Y.-S. Cho, T.-S. Kim, and H.-S. Cho, “The effects of computer-assisted cognitive rehabilitation on Alzheimer’s dementia patients memories,” *Journal of Physical Therapy Science*, vol. 27, no. 9, pp. 2921–2923, 2015, doi: 10.1589/jpts.27.2921.

[135] M. Cavallo and C. Angilletta, “Long-Lasting Neuropsychological Effects of a Computerized Cognitive Training in Patients Affected by Early Stage Alzheimer’s Disease: Are They Stable Over Time?,” *Journal of Applied Gerontology*, vol. 38, no. 7, pp. 1035–1044, Jul. 2019, doi: 10.1177/0733464817750276.

[136] N. Mendoza Laiz, S. del Valle Díaz, N. Rioja Collado, J. Gomez-Pilar, and R. Hornero, “Potential benefits of a cognitive training program in mild cognitive impairment (MCI),” *Restorative Neurology and Neuroscience*, vol. 36, no. 2, pp. 207–213, 2018, doi: 10.3233/RNN-170754.

[137] E. Irazoki, L. M. Contreras-Somoza, J. M. Toribio-Guzmán, C. Jenaro-Río, H. van der Roest, and M. A. Franco-Martín, “Technologies for cognitive training and cognitive rehabilitation for people with mild cognitive impairment and dementia. A systematic review,” *Frontiers in Psychology*, vol. 11. Frontiers Media S.A., Apr. 09, 2020. doi: 10.3389/fpsyg.2020.00648.

[138] A. Gaitán, M. Garolera, N. Cerulla, G. Chico, M. Rodriguez-Querol, and J. Canela-Soler, “Efficacy of an adjunctive computer-based cognitive training program in amnestic mild cognitive impairment and Alzheimer’s disease: A single-blind, randomized clinical trial,” *International Journal of Geriatric Psychiatry*, vol. 28, no. 1, pp. 91–99, Jan. 2013, doi: 10.1002/gps.3794.

[139] M. Eckroth-Bucher and J. Siberski, “Preserving cognition through an integrated cognitive stimulation and training program,” *American Journal of Alzheimer’s Disease and other Dementias*, vol. 24, no. 3, pp. 234–245, 2009, doi: 10.1177/1533317509332624.

[140] M. Hagovská, O. Dzvoník, and Z. Olekszyová, "Comparison of two cognitive training programs with effects on functional activities and quality of life," *Research in Gerontological Nursing*, vol. 10, no. 4, pp. 172–180, Jul. 2017, doi: 10.3928/19404921-20170524-01.

[141] A. Vermeij, R. P. C. Kessels, L. Heskamp, E. M. F. Simons, P. L. J. Dautzenberg, and J. A. H. R. Claassen, "Prefrontal activation may predict working-memory training gain in normal aging and mild cognitive impairment," *Brain Imaging and Behavior*, vol. 11, no. 1, pp. 141–154, Feb. 2017, doi: 10.1007/s11682-016-9508-7.

[142] L. Hyer *et al.*, "Cognitive Training Program to Improve Working Memory in Older Adults with MCI," *Clinical Gerontologist*, vol. 39, no. 5, pp. 410–427, Oct. 2016, doi: 10.1080/07317115.2015.1120257.

[143] F. Barban *et al.*, "Protecting cognition from aging and Alzheimer's disease: A computerized cognitive training combined with reminiscence therapy," *International Journal of Geriatric Psychiatry*, vol. 31, no. 4, pp. 340–348, Apr. 2016, doi: 10.1002/gps.4328.

[144] S. Lasaponara, F. Marson, F. Doricchi, and M. Cavallo, "A scoping review of cognitive training in neurodegenerative diseases via computerized and virtual reality tools: What we know so far," *Brain Sciences*, vol. 11, no. 5, May 2021, doi: 10.3390/BRAINSCI11050528.

[145] S. Shinagawa *et al.*, "Non-pharmacological management for patients with frontotemporal dementia: a systematic review," *Journal of Alzheimer's disease*, vol. 45, no. 1, pp. 283–293, 2015, doi: 10.3233/JAD-142109.

[146] Z. Moussavi, "MindTriggers." Available at: https://apps.apple.com/us/app/mind-triggers/id1508008769., 2020.

[147] C. M. Cullum and G. J. Larrabee, "Chapeter six-WAIS-IV use in neuropsychological assessment," in *WAIS-IV Clinical Use and Interpretation*, San Diego: Academic Press, 2010, pp. 167–187.

[148] "Audacity," *Available at: https://www.audacityteam.org/*.

[149] F. Eyben *et al.*, "The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing," *IEEE Transactions on Affective Computing*, vol. 7, no. 2, pp. 190–202, Apr. 2016, doi: 10.1109/TAFFC.2015.2457417.

[150] A. Honore, "Some Simple Measures of Richness of Vocabulary," *Association of Literary and Linguistic Computing Bulletin*, vol. 7, pp. 172–177, 1979.

[151] E. Brunet, *Le Vocabulaire de Jean Giraudoux : structure et évolution : statistique et informatique appliquées à l'étude des textes à partir des données du Trésor de la langue française*. Genève : Slatkine, 1978.

[152] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011, [Online]. Available: http://scikit-learn.sourceforge.net.

[153] N. v Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.

[154] M. M. Mukaka, "Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research," *Malawi Medical Journal*, vol. 24, no. 3, pp. 69–71, 2012, [Online]. Available: www.mmj.medcol.mw

[155] "R Core Team (2021). R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

[156] S. Ahmed, A. M. F. Haigh, C. A. de Jager, and P. Garrard, "Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease.," *Brain*, vol. 136, no. Pt 12, pp. 3727–3737, 2013, doi: 10.1093/brain/awt269.

[157] F. di Iulio *et al.*, "Occurrence of neuropsychiatric symptoms and psychiatric disorders in mild Alzheimer's disease and mild cognitive impairment subtypes," *Int Psychogeriatr*, vol. 22, no. 4, pp. 629–640, 2016, doi: 10.1017/S1041610210000281.

[158] N. Steck, C. Cooper, and V. Orgeta, "Investigation of possible risk factors for depression in Alzheimer's disease: A systematic review of the evidence," *Journal of Affective Disorders*, vol. 236, pp. 149–156, Aug. 2018, doi: 10.1016/j.jad.2018.04.034.

[159] C. Solomon, M. F. Valstar, R. K. Morriss, J. Crowe, and A. Vinciarelli, "Objective methods for reliable detection of concealed depression," *Frontiers in ICT*, vol. 2, 2015, doi: 10.3389/fict.2015.00005.

[160] W. J. Silva, L. Lopes, M. K. C. Galdino, and A. A. Almeida, "Voice Acoustic Parameters as Predictors of Depression," *Journal of Voice*, 2021, doi: 10.1016/j.jvoice.2021.06.018.

[161] A. T. Beck, R. A. Steer, R. Ball, and W. F. Ranieri, "Comparison of Beck Depression Inventories-IA and-II in Psychiatric Outpatients," *J Pers Assess*, vol. 67, no. 3, pp. 588–597, 1996, doi: 10.1207/s15327752jpa6703_13.

[162] J. Williamson, T. Quatieri, B. Helfer, R. Horwitz, B. Yu, and D. Mehta, "Vocal biomarkers of depression based on motor incoordination," in *Proceedings of the 3rd ACM international workshop on audio/visual emotion challenge*, 2013, pp. 41–48. doi: 10.1145/2512530.2512531.

[163] L. He and C. Cao, "Automated depression analysis using convolutional neural networks from speech," *Journal of Biomedical Informatics*, vol. 83, pp. 103–111, Jul. 2018, doi: 10.1016/j.jbi.2018.05.007.

[164] S. A. Montgomery and M. Asberg, "A New Depression Scale Designed to be Sensitive to Change," *British journal of psychiatry*, vol. 134, no. 4, pp. 382–389, 1979, doi: 10.1192/bjp.134.4.382.

[165] R. T. Sataloff, D. C. Rosen, M. Hawkshaw, and J. R. Spiegel, "The Aging Adult Voice," 1997.

[166] S. H. Pang, S. F. Lim, and C. J. Siah, "Online memory training intervention for early-stage dementia: A systematic review and meta-analysis," *Journal of Advanced Nursing*, vol. 77, no. 3. Blackwell Publishing Ltd, pp. 1141–1154, Mar. 01, 2021. doi: 10.1111/jan.14664.

[167] K. L. Gigle, K. Blomeke, E. Shatil, S. Weintraub, and P. J. Reber, "Preliminary evidence for the feasibility of at home online cognitive training with older adults," *Gerontechnology*, vol. 12, no. 1, pp. 26–35, 2013, doi: 10.4017/gt.2013.12.1.007.00.

[168] S. Teles, C. Paúl, P. Lima, R. Chilro, and A. Ferreira, "User feedback and usability testing of an online training and support program for dementia carers," *Internet Interventions*, vol. 25, pp. 100412–100412, Sep. 2021, doi: 10.1016/j.invent.2021.100412.

Appendix A:

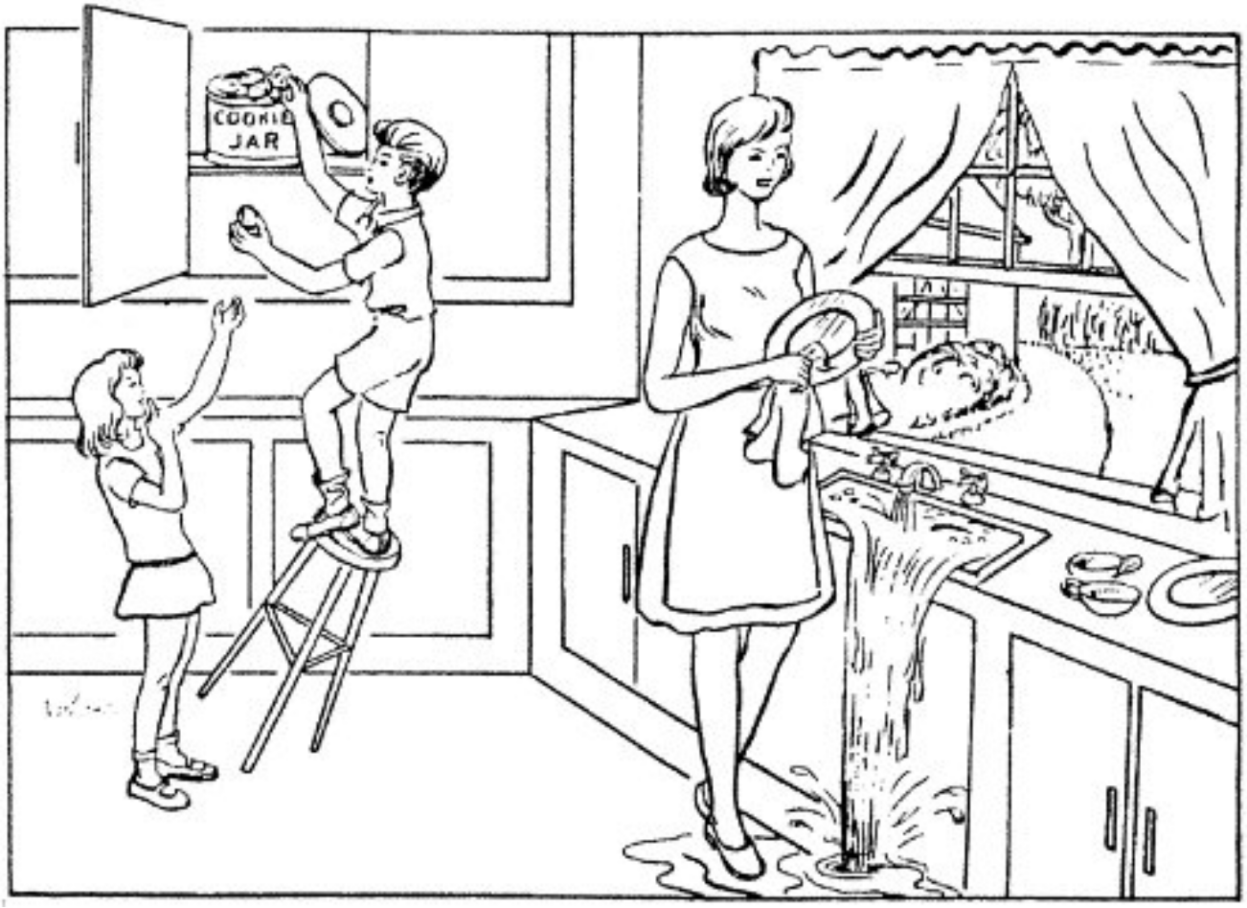


Figure 14-The Cookie Theft picture