SPONTANEOUS SPEECH ANALYSIS FOR DETECTING MILD COGNITIVE IMPAIRMENT AND ALZHEIMER'S DISEASE IN THAI OLDER ADULTS

NATINEE NA CHIANGMAI

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR DOCTOR DEGREE OF PHILOSOPHY
IN RESEARCH AND STATISTICS IN COGNITIVE SCIENCE
COLLEGE OF RESEARCH METHODOLOGY AND COGNITIVE SCIENCE
BURAPHA UNIVERSITY
2023

-

ณัฐินี ณ เชียงใหม่

ดุษฎีนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา
วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
2566

SPONTANEOUS SPEECH ANALYSIS FOR DETECTING MILD COGNITIVE IMPAIRMENT AND ALZHEIMER’S DISEASE IN THAI OLDER ADULTS

NATINEE NA CHIANGMAI

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR DOCTOR DEGREE OF PHILOSOPHY
IN RESEARCH AND STATISTICS IN COGNITIVE SCIENCE
COLLEGE OF RESEARCH METHODOLOGY AND COGNITIVE SCIENCE
BURAPHA UNIVERSITY
2023

The Dissertation of Natinee Na Chiangmai has been approved by the examining committee to be partial fulfillment of the requirements for the Doctor Degree of Philosophy in Research and Statistics in Cognitive Science of Burapha University

Advisory Committee

Principal advisor

....

(Peera Wongupparaj)
Co-advisor

....
(Remo Job)

....
(Pattrawadee Makmee)

Examining Committee

.... Principal examiner

()

.... Acting Dean of the College of Research Methodology and Cognitive Science

()

....

This Dissertation has been approved by Graduate School Burapha University to be partial fulfillment of the requirements for the Doctor Degree of Philosophy in Research and Statistics in Cognitive Science of Burapha University

.... Dean of Graduate School
(Associate Professor Dr. Witawat Jangiam)

....

61810049: MAJOR: RESEARCH AND STATISTICS IN COGNITIVE SCIENCE; Ph.D. (RESEARCH AND STATISTICS IN COGNITIVE SCIENCE)
KEYWORDS: Neurocognitive assessment, Language-based tasks, Spontaneous speech analysis, Cognitive deficits, Dementia

NATINEE NA CHIANGMAI : SPONTANEOUS SPEECH ANALYSIS FOR DETECTING MILD COGNITIVE IMPAIRMENT AND ALZHEIMER’S DISEASE IN THAI OLDER ADULTS. ADVISORY COMMITTEE: PEERA WONGUPPARAJ, REMO JOB PATTRAWADEE MAKMEE 2023.

Memory deficits in Alzheimer’s disease (AD) and mild cognitive impairment (MCI) can be reflected in language-based tests, especially spontaneous speech tasks. Three spontaneous speech tests were developed in this study, including Thai Picture description (TPD), Thai Story Recall (TSR), and Semi-structured Interview for Thai (SIT). Ninety-eight Thai older adults underwent screening tests and three spontaneous speech tests. Then ,they were classified into three groups, including healthy control (HC), MCI, and AD. The verbal responses of the participant were extracted into content variables and acoustic features, and exanined thier discriminant ability and accuracy in differentiating HC, MCI, and AD with by Multivariate Discriminant Analysis (MDA), and analysis of ROC curve and AUC.

Two content variables showed significant differences among three groups of participants, i.e., correct information unit (CIU) of the TPD and delayed recall scores of the TSR. For acoustic features, ANOVAs revealed that three variables were significantly different among the three experimental groups, i.e., number of voice breaks in the TPD, number of voice breaks in the SIT, and total utterance time in delayed recall. The result of a stepwise estimation in MDA presented that the optimal combination of predictive model was CIU and backward digit span (BDS), in which provides 61.1% of correct classification. This discriminant function showed AUC of .81 in differentiating HC and MCI, AUC of .91 in distinguishing HC and AD, and AUC of .86 in detecting persons with cognitive impairments (MCI and AD) from HC.

In conclusion, the combination of CIU of TPD and BDS is suitable for differentiating AD and persons with cognitive impairments from HC. However, there is no appropriate predictor in distinguishing MCI and AD.

E

ACKNOWLEDGEMENTS

This work was supported by the Thailand Research Fund and the National Research Council of Thailand, through the Royal Golden Jubilee Ph.D. Scholarship Program; and Cognitive Science and Innovation Research Unit, Burapha University, Thailand. We thank the Bruno Kessler Institute and the Center for Neurocognitive Rehabilitation, University of Trento, all participants, their families, community healthcare volunteers, and technician staff for making this study possible.

Natinee Na chiangmai

TABLE OF CONTENTS

Page

List of tables

List of figures

CHAPTER 1
INTRODUCTION

Background

Neurodegenerative diseases are age-dependent disorders characterized by the ultimate death of neurons. These neuron deaths are caused by a series of abnormal protein syntheses inside and outside the neurons and affect different brain regions (Hussain et al., 2018). Dementia is one form of a neurodegenerative disease where this neuron damage causes significant cognitive impairments and behavior changes (Sachdev et al., 2014). Adults are suffering from disability and death due to dementia worldwide (Global Burden of Disease 2019 Ageing Collaborators, 2022; World Health Organization, 2023).

The global prevalence of dementia in people aged 60 and above is currently relatively small, ranging from 5%–7% of the population. Mild Cognitive Impairment (MCI) prevalence is slightly higher at 6%–12% of the population (Hampel & Lista, 2016). The World Health Organization (WHO) (2023) estimates that ten million cases of dementia will be found per year, approximately one every three seconds. According to the latest update of WHO in 2023, there are over 55 million dementia sufferers globally and almost 10 million new cases each year. In 2019, the estimated costs to society of dementia were US$ 1.3 trillion (World Health Organization, 2023).

Alzheimer's disease (AD) is the most common cause of dementia which is estimated at 60–70% of cases (World Health Organization, 2023). Of concern is that while trends in mortality rates between 2000 and 2013 showed decreases in stroke (23%), heart disease (14%), and prostate cancer (11%) respectively, death from Alzheimer's disease increased by 71% (Alzheimer's Association, 2016).

Across Thailand, the prevalence of dementia is 2.35% in people aged over 60 years of age was estimated to be almost 617,000 cases annually (Foundation of Thai Gerontology Research and Development Institute, 2016). Young onset dementia (where the symptoms occur before 65 years of age) accounts for 17% of all dementia patients, approximately a mean age of 58 years (Dharmasaroja et al., 2021). In the global study on dementia prevalence, the estimated number of Thai dementia cases of all ages was 670,047 people in 2019. According to the same study, the estimated percentage change in Thai cases between 2019–2050 could be 257% (Nichols et al., 2022). The etiology in the Thai cohort is Alzheimer's disease (50%), followed by vascular dementia (24%), dementia with Lewy bodies (6%), Parkinson's disease dementia (6%), frontotemporal dementia (2.6%), Progressive supranuclear palsy (2%), multiple system atrophy (0.8%), corticobasal syndrome (0.4%), other causes (4.4%), and unidentified causes 3% (Dharmasaroja et al., 2021). The prevalence of dementia is directly related to the increase in the aged population (an 'aging society'). It is also associated with socioeconomic status, particularly in low- and middle-income countries (Dharmasaroja et al., 2021; World Health Organization, 2023). By 2022, the number of individuals over the age of 60 in Thailand will reach 20% of the total population, meaning that the country will have become a 'complete aged society'. On current trends, Thailand will become a 'super-aged society' in 2023, with 28% of the population aged over 60 years (Foundation of Thai Gerontology Research and Development Institute, 2019). The key question is whether Thailand is ready for dementia.

In 2013, DSM-5 renamed dementia to 'major neurocognitive disorder' and further defined the earlier stages as 'mild neurocognitive disorder' (Sachdev et al., 2014). The diagnostic criteria were also modified. The six cognitive domains were treated as equally important evidence, including learning and memory, language, attention, executive function (EF), perceptual-motor function, and social cognition (Sachdev et al., 2014). In the case of MCI criteria attention was paid for clinical characteristics of memory impairment, resulting in two distinctive groups, namely an amnestic-MCI and a non-amnestic MCI (Petersen, 2016). Based on the aforementioned information, the early detection of cognitive decline with integrated tools and multimodal parameters may be considered and investigated accordingly.

The prodromal stage of dementia is arguably known as mild cognitive impairment (MCI), yet it can be evident for other diseases, that is, other neurodegenerative or psychiatric disorders (Filiou et al., 2019). MCI usually develops in middle adulthood while patients are still able to live a normal day-to-day life and engage in complex activities (Livingston et al., 2017). This silent and usually subtle decline is difficult to detect, as patients can either present with other memory problems or non-amnestic impairment, e.g., having poor decision-making on simple tasks, producing mistakes in conversation, or failing in visuospatial activities (Petersen, 2004).

Due to the fact that dementia is viewed as an incurable disease, early detection helps in planning interventions and delaying the progression of the disease (Kulkantrakorn, 2018). Early detection of dementia is particularly imperative, especially in detecting early onset symptoms in young adults, as 44% of patients diagnosed with dementia at a young age will develop Alzheimer’s disease (Dharmasaroja et al., 2021). The progression rate demonstrates the vital importance of early diagnosis as it will assist individuals, clinical practitioners, and policymakers in delaying progression, planning for appropriate treatments, and managing health and social care expenses (World Health Organization, 2017). An early diagnosis is further underscored by WHO, including the diagnostic rate in the Global Dementia Observatory indicator in 2018 (World Health Organization, 2018).

Neuropsychological, neurochemical, and neuroradiological techniques are suggested for early screening, differential diagnosis, and evaluating of treatment for dementia. These measures all offer promising sensitivity (Schmand et al., 2011). The neuropsychological tests used in Thailand have been almost universally adopted from international standardized tools. Today, approximately 30% of these tests are developed in Thailand (Na Chiangmai & Wongupparaj, 2020). Technology is also increasingly being utilized, providing an alternative measure that increases both accessibility and cost-effectiveness. The technology implementation, however, must be adequately refined for clinical practices (Koo & Vizer, 2019).

Previous studies relating to screening tests in Thailand have been largely focused on the neuropsychological tests which have been developed outside the country, e.g., Silpakit et al. (2007); Limpawattana et al. (2012); Tangwongchai et al. (2015); Julayanont et al. (2015); Charernboon (2019) (Na Chiangmai & Wongupparaj, 2020). Dementia screening tests are provided in both single and multiple domain(s) assessments. The standardized screening tests which are frequently used in Thai clinical settings are Mini Mental State Examination (MMSE, MMSE-Thai version 2002) and Montreal Cognitive Assessment (MoCA) in Thai version ((Neurological Institute of Thailand, 2014). Both tests are considered as the multiple-domain assessment in an interview-based form. When the translated tools were used, the original cut-off scores were applied to Thai cohort (Neurological Institute of Thailand, 2014). This implementation could cause a false positive because of the low education level in Thai elderly cohort as literacy status is related with both the performance score and the dementia classification (Charernboon, 2019; Silpakit et al., 2018). Although several studies tended to investigate for proper criteria for Thai older adults or in the different healthcare settings, the original scores are still implemented.

Another consideration is the cultural influences that impact an individual's thoughts and behaviors. Test developers and users should take into account these cultural influences, including values, beliefs, and behavioral styles (Ardila, 2005). Since dementia screening tools are not diagnosis criteria, tests with high sensitivity aim to not miss early cases. In addition, tests with high specificity should reduce mistaken diagnosis (Larner, 2017; Silpakit, 2013). The dementia screening test will be more useful in Thai clinical settings if these confounding factors are reduced or removed.

Cognitive abilities and declines can be captured examining several types of performances. Abilities and impairments can be elicited through neuropsychological testing, whose measures include response accuracy or response time. One method for dementia assessment is to analyze patient speech performance, e.g., König et al. (2015), Weiner et al. (2016), Weiner et al. (2017), Toth et al. (2018), and Gosztolya et al. (2019). As language is one of the concerning domains in neurodegenerative diseases, a patient's speech abilities carry not only linguistic information but also reflects memory and other cognitive abilities. Language disorders are also found in neurodegenerative patients, namely Primary Progressive Aphasia (PPA) (Pulido et al., 2020). Aphasia is known for the impairment of language due to brain injury, especially in the left hemisphere (Bayles et al., 2020, p. 71; Goldstein, 2019, p. 39). Speech problems are evident in PPA in various types of dementia and MCI. Speech problems caused by neurodegeneration can be observed as difficulties in naming objects, decreasing word comprehension, or non-fluency speech (Klimova & Kuca, 2016).

Language deficit is considered as a noticeable sign in the early stage of AD throughout communication of the patients (Szatloczki et al., 2015). However, language impairment is thought to be unnecessary or insufficient to seek for AD diagnosis (Smith & Bondi, 2013). Neurodegenerative diseases can lead to some form of aphasia, since normal language function relies on a wide range of brain regions. In AD case, the neural network damaged by AD defines consequently the symptoms of aphasia (Weekes, 2020). The atrophy pattern in proven cases of AD with progressive Aphasia reveals damages in the left posterior superior temporal lobe, inferior parietal area, medial temporal lobe, and posterior cingulate region. These atrophy areas are partially in Wernicke's area and Broca's area (Hickok, 2012; Seghier, 2013). Typically, language deficits in AD are naming difficulty, semantic paraphasia, language comprehension and fluent but meaningless articulation (Szatloczki et al., 2015; Weekes, 2020). Language measures such as verbal fluency, naming or recognition are sensitive to persons with AD (Arvanitakis et al., 2019; Budson et al., 2016; Muangpaisan et al., 2010). Growing of evidence suggests that spontaneous speech task, a task which resembles daily articulation and conversation, acquires a more comprehensive and ecological valid in assessing of both normal and patients with AD (Filiou et al., 2019; Martínez-Nicolás et al., 2021).

In a speech recording process, the instructions to pronounce, read and talk are given to participants in order to stimulate cognitive functions, and standardized cognitive screening tasks are administered, e.g., Wechsler Logical Memory I and II (WLM I&II). WLM is an immediate and delayed retelling of a story a person has been just orally presented with. Roark et al. (2011) found that only the delayed recall task differentiated between healthy and MCI participants. Verbal fluency, a conventional language task in which a person is asked to mention as many instances as possible of a given semantic (e.g. animals) and phonemic (e.g. words starting with the letter "f") categories; both verbal fluency tasks provided discrimination ability in different levels of cognitive declines, such as control group and MCI (Beltrami et al., 2018), control and Alzheimer's disease (AD) groups (Jokel et al., 2019) and MCI and AD (König et al., 2015).

Another form of testing activity is the 'free speech task', which takes the form of a picture description, dream recollection, and normal conversation. These tasks imitate usual daily life activities which a person utilizes to retrieve relevant information from memory and then present them in verbal expression. Linguistic impairment was widely accepted for the detection of MCI and Alzheimer's disease especially verbal responses derived from spontaneous speech, e.g., word finding, retrieval difficulties and reduction in discourse production (Gosztolya et al., 2019; Pulido et al., 2020; Satt et al., 2013).

The parameters in speech-based tasks can be quantified based on linguistic analyses taking into account several features. The distinctive features commonly considered in speech analysis studies are lexical, rhythmic, acoustic, or syntactic features (Beltrami et al., 2018). Currently, discriminability of linguistic features is focused on because each feature has different advantages and limitations, also inconsistent results. For example, syntactic features marked the difference between healthy control and participants with cognitive decline in Italian verbal response to open-ended questions (Beltrami et al., 2016. While the similar study with open-ended question elicited spontaneous speech in a Japanese study revealed the opposite result, participants with MCI showed larger vocabulary size than healthy group (Aramaki et al., 2016). They suggested that individuals with MCI may try to conceal their cognitive deterioration. Notably, acoustic feature profiles of patients with neurodegenerative disease were found to have significant differences from control groups in the studies with speech analysis (Al-Hameed et al., 2019; Toth et al., 2018; Weiner et al., 2016). Interestingly, the most significant differences between AD and MCI were shown in the temporal domain of speech. The physical characteristic of language regards acoustic features, i.e., the temporal property of speech can be analyzed across different languages (János et al., 2022; Ladefoged, 2006, p. 316). Therefore, acoustic variables are possibly beneficial in applying to Thai spoken response in order to explore the speech profile of people with neurodegenerative and healthy control in Thailand.

For over 30 years, clinical diagnosis studies have used speech analysis in Thailand, but none of these studies directly focused on dementia. Jack Gondour and his colleagues explored abnormal characteristics of voice in Thai aphasic patients and in patients with brain damage having adapted to Thai the original version of the Boston Diagnostic test ('A Thai adaptation of the Boston Diagnostic Aphasia

Examination' (TDAE) (Gandour et al., 1986; Gandour et al., 2000)). Speech processing was also of interest to the computer science and engineering disciplines that aimed at developing automatized speech recognition. Recently, speech characteristics have been highlighted again in medical literature and are purposive in classifying patients from cognitively intact people. Those studies were conducted with different groups of patients, including people with diabetes (Pinyopodjanard et al., 2019), depression (Yingthawornsuk, 2016), and neurodegenerative diseases (Manochiopinig et al., 2008; Nagarachinda et al., 2020; Suanpirintr et al., 2008).

Previous speech analysis studies showed that further studies with improved methodology, technology, and interpretation of the results were needed. The most common limitations of these prior studies were small sample sizes, non-homogeneous populations, and a lack of confidence in the generalizability of the results (Ambrosini et al., 2019; Fraser et al., 2019; Singh et al., 2001). In order to find reliable and systematic, predictive features, longitudinal studies have been planned (Beltrami et al., 2018; König et al., 2018; Roark et al., 2011).

One crucial aspect of these studies concerns the speech analysis methods that were adopted, specifically the utilization of automated or manual transcription. The manual option gave fewer errors but was very time-consuming (Tröger et al., 2018). The full automation of the process would be ideal but requires algorithm development (Asgari et al., 2017; Themistocleous et al., 2018). Another challenge in this area is the co-presence of several, possibly confounding factors, since spontaneous speech is related to multiples factors, including cognitive function, social and cultural effects, or physical limitations. The results of speech analysis, thus, should be carefully interpreted and exclusively targeted in dementia-related factors rather than other confounding factors (Filiou et al., 2019).

The aforementioned limitations can be improved in a Thai cohort study. Speech analysis studies in Thai older adults with neurodegenerative diseases were mostly interested in lexical and pragmatic features, such as Sangchocanonta et al. (2021) and Nagarachinda et al. (2020). They analyzed spontaneous speech of AD, MCI and healthy control. Regarding acoustic features, the study of Amonlaksananon et al. (2021) specifically examined speech disfluency of AD and MCI. One acoustic variable included in this study was silent pauses among other linguistic features such as syntactic error. However, they presented that silent pause was not able to differentiate the elderly with MCI from AD. This study would explore more other aspects and variables of acoustics features. Moreover, the usefulness of natural speech is a minimizing of the deviation on F0 (fundamental frequency); this feature is related to tonal articulation in Thai (Gandour et al., 1996; Suanpirintr et al., 2008). Thus, spontaneous speech is very appropriate for the speech analysis for dementia screening in older Thai adults. In addition, acoustic features in temporal and frequency-related domains can reveal the Thais speech profiles.

Objectives of This Project

1. To develop some dementia screening tasks based on spontaneous speech analysis for older Thai adults.
2. To compare the patterns of acoustic features profile in Thai older adults with MCI, AD, and cognitively intact persons.
3. To validate the speech analysis deriving from the developed tasks in classifying healthy older adults, MCI, and AD.

Conceptual framework

AD is the main cause of dementia and early detection in the prodromal stage or MCI allows health practitioners, patients, and families to develop strategies to attempt to delay the disease progression. Screening assessment with neuropsychological tests is suggested in the very first step of the diagnosis process. Several cognitive functions are evaluated through abilities and failures of participants' tests performance. AD has amnesia as a prominent symptom in relation to the certain subtype of MCI, which becomes AD in the later stage, called amnestic MCI. Memory function thus has been investigated in AD and MCI also through verbal responses to linguistic tasks. The speech characteristics of AD and MCI demonstrate similar patterns, such as a slow rate of utterance, low articulation, and speaking with hesitation; those speech deficits represent not only language impairment but memory and attention dysfunction.

Speech processing is entailed by several brain areas related to cognitive domains and physical functions. Speech production is initiated with simultaneous cognitive execution, including mentally forming the message with intention and previous knowledge, followed by the composition of linguistic elements, which are stored briefly in working memory (WM). Then, the message is briefly stored and manipulated in short-term memory as a capacity to chuck information (Baddeley, 2003; Cummins et al., 2015). Each cognitive function is explained in terms of the elements and processing in the following parts. Cognitive impairments of spontaneous speech in dementia relate to three main cognitive functions, i.e., attention, memory, and language. Each cognitive function is integrated with a specific and common area of the brain.

To begin with attention, a person grounds ideas or intentions in the conceptualization process. In the initiative process, a person needs to focus on the ideas or certain stimuli and not deliberately shift one's attention. The brain cannot equally process or allocate attention when confronted with more than one stimulus; concentration and neglect are needed a balancing control, according to James Williams (Goldstein, 2019, p. 95 & 102). The ability to engage with the stimulus is mediated in the parietal lobes. The slow disengagement process significantly correlates with parietal lobe damage, especially the lesion on the right. The parietal lobes strongly respond to attended stimulus rather than to disengaging (Kandel et al., 2013, p. 438). Among cognitively intact persons, selective attention is activated in the prefrontal cortex. This area controls the capacity to make shift-attention and divided attention with the function of inhibiting distraction effects. The prefrontal cortex therefore mediates working memory. The patients with the prefrontal damage consequently have a poor ability to disengage interference from their activities in mind, whether from the external or internal person (Lezak et al., 2012). The parietal lobes and prefrontal cortex have been linked to the regulation of visual attention, as evidenced by the individuals who experience attention problems after a stroke (Kandel et al., 2013, p. 438).

Memory loss is a common sign of Alzheimer's disease and MCI, particularly when this symptom obviously disrupts daily activities(Alzheimer's Association, 2020a). As mentioned in the previous section focusing on the attention the prefrontal cortex implicates in memory. The prefrontal cortex and medial temporal lobe mediate memory from the encoding to the retrieval of past knowledge (Kandel et al., 2013, p. 1449). The prefrontal cortex mediates episodic memory and working memory. The other prominent region of dementia is the medial temporal lobe and the hippocampus. The latter region is a structure in the limbic system, which lies on the medial temporal lobes. The limbic lobe consists of other structures apart from the hippocampus, which plays an important role in memory. The hippocampus is responsible for learning new experiences, such as a short story or information, along with timeframe, which is episodic memory. New information will be formed into long-term memories and consolidated by the mediating effect of the hippocampus. The inability to retrieve episodic memories is usually one of the earliest signs of AD (Budson et al., 2016, p. 9 & e12; Kandel et al., 2013, p. 409). The medial temporal lobe and hippocampus are affected early on in patients with Alzheimer’s disease, and therefor memory loss is an early observable symptom (Albert et al., 2011).

Speech processing is an expressive part of language; the other important part of language is the receptive component. Broca’s and Wernicke’s areas are two prominent parts of the brain implementing language. The former is in the left lateral frontal region; the latter is in the posterior superior temporal lobe. In the receptive process, auditory input is important, so Wernicke’s area lies near the primary auditory cortex, while Broca’s area is close to the motor cortex that regulates articulators (Kandel et al., 2013, p. 11 & 1360). Speech impairment or aphasia is caused by the lesion at language control, where are included the prefrontal cortex. The prefrontal cortex plays a crucial role in executive control and mediation of the working memory and attention processes. Visual encoding tasks present greater activation in the visual cortex, left prefrontal cortex, and medial temporal lobe. The Broca’s and Wernicke’s areas bidirectionally work together and connect with the prefrontal and premotor areas to convey speech. The lesion in the prefrontal cortex thus resulted in difficulties in word-finding, poor attention, and disinhibition. (Budson et al., 2016, p. 86; Kandel et al., 2013, p. 1363 & 1449).

The cognitive process of speech is at the intersection of attention, memory, and language. In the beginning, the conceptualization process portrays a meaning and purpose which needs attention to keep the idea in mind. Then cognition formulates those concepts or ideas to be ‘sound’ in mind or implicit memory. This stage plays a ‘bridge’ role between the speaker’s concepts and linguistic information in his/her memory. The formulation process is an encoding of the linguistic elements in order to produce the utterance by retrieval of semantic and episodic memory (Bürki, 2018; Postma, 2000). The phonological encoding process generates the sounds in the mental representation and comprises the sounds of the words in a particular spoken language. Although sentence production is extremely complex, this process is planned and executed effortlessly and unconsciously; the evidence is speech errors such as nonagreement of grammar in a sentence (Fernandéz & Cairns, 2010, p. 147 & 167). During articulation, the phonological syllables activate the articulatory gestures and their temporal relationships. In the final step, the syllabic gestures with all necessary details are in for the articulation apparatus controlled by motor activity (Ladefoged, 2006, pp. 6 & 310-312; Postma, 2000). The signals represent phonological encoding processes in speech analysis enclosed in the acoustic features, for example, pitch, loudness, and duration (Goswami, 2012).

However, speech impairment in patients with dementia is not only caused by the dysfunctional motor network. The aforementioned cognitive functions and brain areas are also involved. AD is found to show language production deficits such as difficulty in word finding or naming objects, word repetition, hesitation, long silence, or slow speech (Klimova & Kuca, 2016; Marczinski & Kertesz, 2006). The speech production information is vocal markers of MCI which present at a slighter level (Filiou et al., 2019). Several studies thus used acoustic features which represent the physical characteristic of speech for dementia detection, especially in naturally speaking or spontaneous speech (König et al., 2015; Ladefoged, 2006, p. 218 & 310; Pulido et al., 2020).

To the extent of replicating daily communication, the tasks will induce spontaneous speech with the phenomena of the phonological loop and common components of speech tasks. Three proposed tasks will initiate a person to retrieve previous information in long-term memory (LTM), both episodic and semantic memory during planning and execution to respond to the tasks. To assess short-term memory and working memory, the cognitive load is needed to test attentional control and the ability to retain information for short periods. The tasks thus include the process of immediate recall, delayed recall, and responding to the answer related to retained information. The characteristic of language stimuli will consist of both visual and audio modalities corresponding to the transfer of information between codes. The differences in familiarity, phonological similarity, and word length are composed in the proposed tasks. These components are embedded in three selected tasks, including picture description, story recalling, and semi-structured conversation.

The acoustic features are aimed to be classified parameters as a consequence of spontaneous speech. The previous studies of speech analysis for detecting MCI, AD, and cognitively intact older person temporal report different significant acoustic parameters. The acoustic parameters can be defined into four different categories, including prosodic, formant, source, and temporal features (König et al., 2019). The prosodic parameters relate to the stress and rhythm in speech; the sample feature is the fundamental frequency (F0). The formants are the indicative class of speech sound that carries acoustic resonance and vibration of the vocal cords. The frequently examined formants are F1-F3 and jitter. These parameters are an indicator of articulation coordination in speech motor control. The source of voice production presents a quality of voice. The sound features are examined by shimmer, harmonics-to-noise (HNR), utterance, and pause segments. The temporal features measure time-related parameters such as duration, rate, and proportion of sounding, pause, and hesitation (Al-Hameed et al., 2019; Ambrosini et al., 2019; Asgari et al., 2017; Beltrami et al., 2018; Gosztolya et al., 2019; Mueller, Koscik, et al., 2018; Roark et al., 2011; Satt et al., 2013; Themistocleous et al., 2018; Toth et al., 2018).

Speech performance is the representativeness of cognitive domains, not only of language but also of working memory and episodic memory. These two memories and the collaboration of related cognitions are reflected through abilities and impairments. To detect dementia, the speech performance of the patient can be evaluated by assessing the acoustic features. This study focuses on detecting MCI and AD because early detection of most causes of dementia will be a benefit for diagnosis, treatment planning, and preparation of relevant persons. The proposed tasks aim to elicit spontaneous speech from speakers by highly imitating daily activities. Voice recording will be analyzed by focusing on acoustic features, which are accepted in detecting cognitive impairment in dementia. At the same time, acoustic features give parameters of interval, frequency, and quality of voice. The application of spontaneous speech analysis is useful for developing telemedicine. The participants or patients can remotely perform the verbal response via telephone interview and assessment. Speech recorded over the telephone utilizes the screening process by reducing time, cost saving, and extending the opportunity to access dementia screening (Rapčan et al., 2009; Tröger et al., 2018).

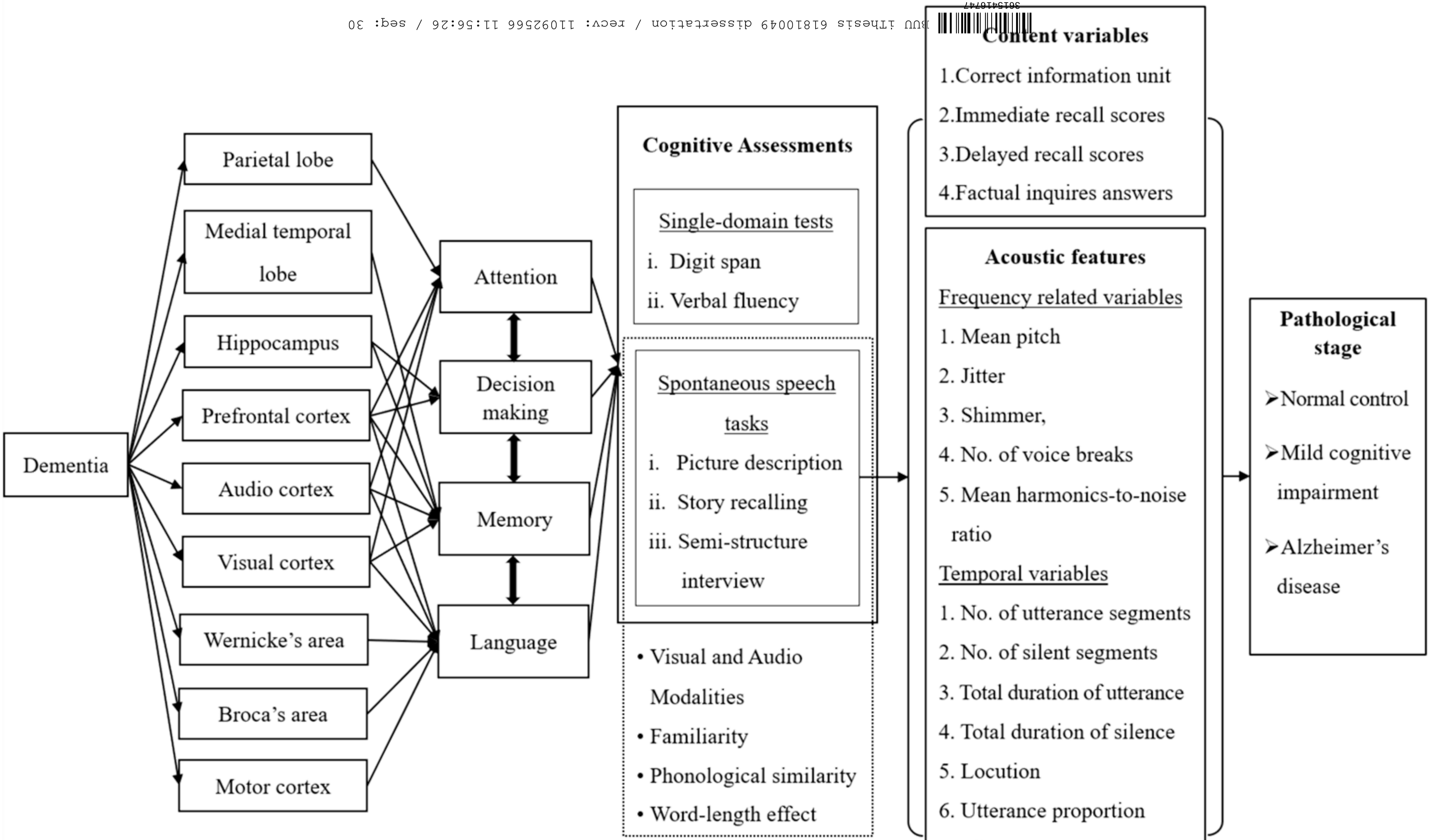


Figure 1 Conceptual framework

Hypotheses and Predictions

1. Participants in the healthy control group acquire the highest mean scores in the correct information units of the Thai Picture description test, followed by the participant with MCI and AD consecutively.

2. Participants in the healthy control group acquire the highest mean scores in immediate recall units, delayed recall unit and answer pertaining to the short story of the Thai Story Recall test, followed by the participant with MCI and AD consecutively.

3. The frequency-related variables are different among the three groups of participants, i.e., healthy control, MCI and AD.

4. Participants in the healthy control group acquired the highest mean values in temporal variables related to utterance followed by the participant with MCI and AD consecutively. The temporal variables related to utterance include utterance segments, total duration of utterance, total length of speech duration, and utterance proportion.

5. Participants in the AD group acquired the highest mean values in temporal variables related to silence followed by the participant with MCI and health control consecutively. The temporal variables related to silence include silence segments, total duration of utterance, silence proportion, pause rate, and silence-to-utterance ratio.

Definition of terms

Alzheimer's disease is a degenerative neuronal disease and the most common cause of dementia, as well as a major cause of death. According to the criteria, memory loss and a decline in thinking abilities are the earlier diagnosis criteria. The initial sign that is most frequently observed is a gradually declining of ability to remember new information. The common symptoms of Alzheimer's disease are related to memory and judgment, e.g., memory loss that disrupts daily life, difficulty completing familiar tasks, and new problems with words when speaking or writing.

Mild Cognitive Impairment (MCI) is a condition in which a person has subtle but measurable changes in cognitive abilities that are noticeable to the person experiencing them, to family members and friends but do not affect the individual's ability to carry out everyday activities.

Single-domain test is a test used in isolation to look at each cognitive domain separately. A single-domain test is purposely designed to assess a particular cognitive function or major cognitive domain. In this study, verbal fluency was utilized to assess language ability and digit span to assess working memory.

Spontaneous speech is self-generated discourse and connected language. It more closely approximates language production in everyday contexts and should be less intervened by examinations. Verbal responses derived by semi-structured and unstructured tasks are considered as spontaneous speech in this study.

Spontaneous speech task is the test that elicits spontaneous speech. The task involves continuous interactions across diverse cognitive processes, including semantic storage and retrieval, executive functions, and working memory. The task may only take a few minutes to complete and place a relatively low burden on a participant.

Acoustic features are related to intensity, duration, and frequency, which are physical characteristics of speech. Acoustic features can be extracted into four different aspects, including prosodics, formant, source, and temporal.

Utterance is a voiced segment that is identified by the Voice Activity Detection technique in Praat program. The duration of an utterance is at least 200 milliseconds (0.2 sec).

Silence is an unvoiced segment at least 1,000 milliseconds (1 sec) long. A long silence may manifest relatively greater difficulty in the retrieval of information or executing verbal responses (Singh et al., 2001). The silent segments are extracted with the same technique as the utterance.

CHAPTER 2
LITERATURE REVIEW

There has been a considerable and growing interest in investigating the early symptoms of older adults who are at risk of developing dementia. A range of tools has been developed across the globe and published in extant literature. This chapter briefly reviews (i) types of dementia, the underlying causes, clinical characteristics, and assessments; (ii) language and speech, including the neuroanatomy of language, the disorder of language and speech, and speech deficits in dementias; and (iii) speech analysis, measurements, and linguistic parameters. The chapter concludes with a discussion of the limitations of the presented studies and identifies areas that warrant further investigation.

2.1 Dementia

The symptoms of dementia are usually recognized when the cognitive decline in an individual is severe enough to affect activities in daily life. The patient presents an impairment in their behaviors and in their cognitive performance. Memory loss has traditionally been the most well-known and most easily recognized symptom in elderly people with dementia (Filiou et al., 2019). Other symptoms, however, can be portrayed as an inability to learn new things, the person continually repeating the same questions, forgetting known routes, agitation, or social withdrawal. Multiple cognitive functions can be affected, including memory, language, orientation, perception, attention, executive function, and social ability (Larner, 2017; McKhann et al., 2011). The different symptoms are caused by different etiology.

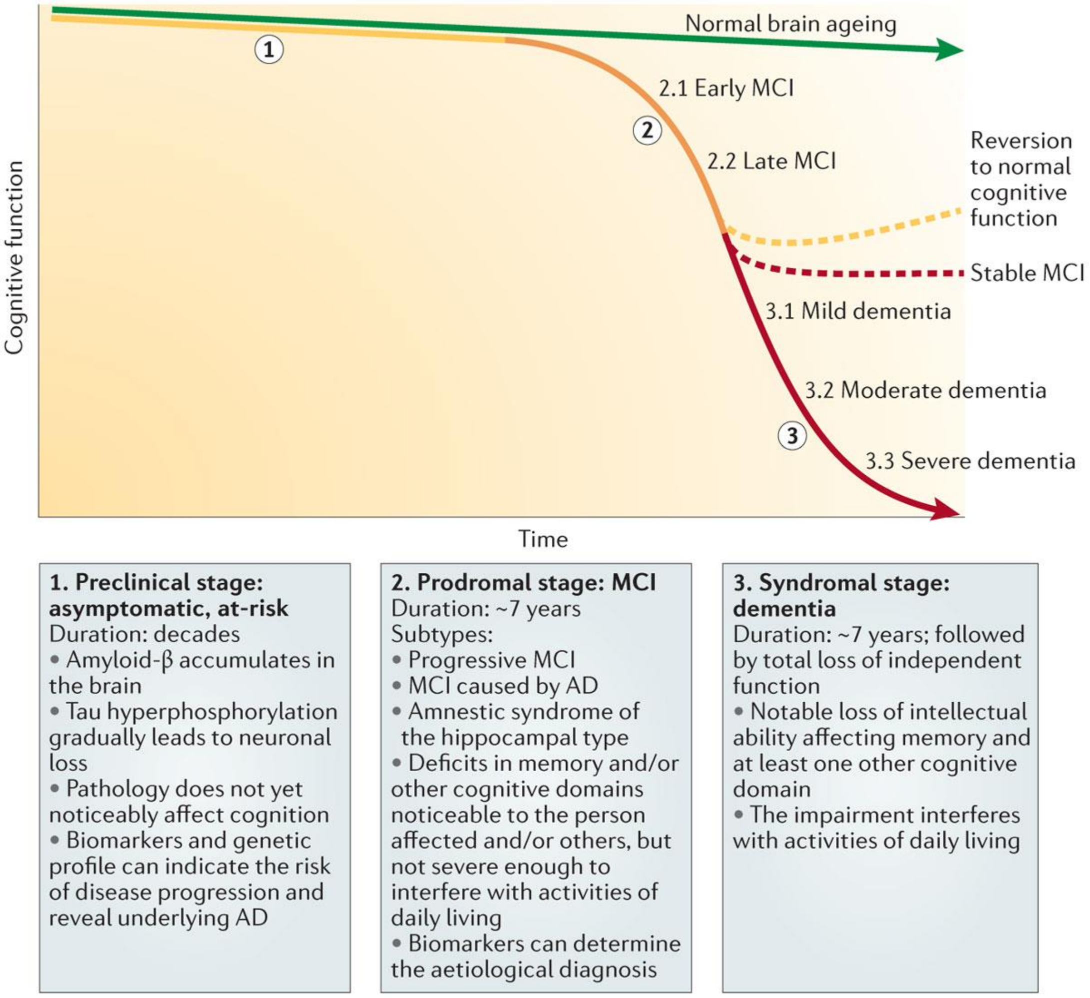
In 2013, when the American Psychiatric Association published the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), Dementia was renamed as a ‘major neurocognitive disorder’. DSM 5 also labeled the earlier stages of cognitive decline as ‘mild neurocognitive disorder’. The reclassification aimed to reduce the stigma associated with dementia and ensure that the diagnostic guidelines harmoniously correspond to clinical practices (Dementia Australia, 2018). The International Classification of Diseases 10th Revision (ICD-10) defines dementia through four indicators. Diagnosis of dementia by ICD-10 requires: (1) impairment in short- and long-term memory; (2) impairment in abstract thinking, judgment, higher cortical function, or personality change; (3) memory impairment and intellectual impairment, which cause significant social and occupational impairments; and (4) the occurrence of these traits when patients are not in a state of delirium (Grabowski & Damasio, 2004).

2.1.1 Causes/Etiology of dementia

The symptoms progression can identify the cause of dementia into two groups. The first group causes a progressive decline, while the other causes a progression that can be recovered. Although "reversible dementias" are described in some accounts, dementia actually refers to an irreversible deterioration in cognitive function brought on by biological processes that harm brain cells. Figure 2 presents the hypothetical staging model of dementia progression. Even though MCI is considered as the prodromal stage of dementia, individuals with MCI can remain stable, progress, or revert to normal cognitive function (Hampel & Lista, 2016). In the following sections, both progressive and reversible dementias are described with prevalence rates.

2.1.1.1 Progressive dementias/neurodegenerative diseases

Dementia is developed from a neurodegenerative disorder, causing progressive cognitive deficits (Lopez-De-Ipina et al., 2015). Different types of dementia are associated with distinctive brain abnormalities and relatively differentiable symptom patterns (Lezak et al., 2012). The decline of the nervous system can be found in all elderly people, but the difference in dementia patients is the reduced ability of damaged systems to recover. The nervous system can generate new neurons, axons, glia, and synapses in order to repair any damage, but dementia prevents this neuroplasticity. Abnormal protein synthesis is one of the obstructions in regeneration. The other dysfunctions cause different types of dementia (Hugo & Ganguli, 2014; Kulkantrakorn, 2018).



1. Preclinical stage: asymptomatic, at-risk
Duration: decades
- Amyloid-β accumulates in the brain
- Tau hyperphosphorylation gradually leads to neuronal loss
- Pathology does not yet noticeably affect cognition
- Biomarkers and genetic profile can indicate the risk of disease progression and reveal underlying AD

2. Prodromal stage: MCI
Duration: ~7 years
Subtypes:
- Progressive MCI
- MCI caused by AD
- Amnestic syndrome of the hippocampal type
- Deficits in memory and/or other cognitive domains noticeable to the person affected and/or others, but not severe enough to interfere with activities of daily living
- Biomarkers can determine the aetiological diagnosis

3. Syndromal stage: dementia
Duration: ~7 years; followed by total loss of independent function
- Notable loss of intellectual ability affecting memory and at least one other cognitive domain
- The impairment interferes with activities of daily living

Figure 2 Hypothetical staging model of dementia progression (Reprinted with permission from Hampel, H. & Lista, S. (2015) The rising global tide of cognitive impairment. Nat. Rev. Neurol. doi:10.1038/nrneurol.2015.250)

i. Alzheimer’s disease (AD)

Alzheimer’s disease is the single most common manifestation of neurodegenerative disease, accounting for between 60 and 70% of all patients with dementia (Hugo & Ganguli, 2014; World Health Organization, 2023). The predominant symptom is memory impairment; hence dementia was traditionally paired with memory loss. The decline of short-term memory prevents the patient from learning new things and gradually presents long-term memory deficits. The other characteristics of AD are visuospatial difficulties, word-finding struggles, and personality changes. Patients may not recognize the changes in themselves, but their caregivers and/or families will observe the problem behaviors which deviate from the individual's previously observed performance. (Budson et al., 2016; Kulkantrakorn, 2018)

These characteristics of a brain with AD are caused by the progressive loss of synapses and neurons, the aggregation of amyloid plaques, neurofibrillary tangles, and prominent cholinergic deficits (Hugo & Ganguli, 2014). Damaged areas are frequently found in the temporal, parietal, and frontal lobes, as well as the hippocampus. AD pathology consists of two main abnormal proteins, which are senile plaques and neurofibrillary tangles. A specific type of amyloid plaque can be found in AD, namely beta-amyloid or 'Aβ'. The extracellular plaques of b-amyloid aggregate the fibrillar peptide sheets (Alzheimer's Association, 2020b). This neuronal process causes inflammation and cell death. At the same time, the live cells have cytoskeleton abnormalities due to the accumulation of neurofibrillary tangles (Budson et al., 2016; Kandel et al., 2013).

ii. Vascular dementia VD

VD is believed to be the second most common cause of dementia, accounting for between 10 and 20 % of dementia. There are three types of vascular disease causing VD, namely cerebral infarction, intracerebral hemorrhage, and small vessel disease (Kulkantrakorn, 2018). The first two causes can contribute to an acute stepwise pattern or be rapid in its course of progression. While small vessel disease shows a more gradual pattern of neurocognitive decline, the patient thus shows slower executive function. The location of lesions is more relevant to the presenting symptoms than the volume of destruction (Kalaria, 2016). Patients with VD often have risk factors for heart disease or carotid artery stenosis/occlusion (Iadecola, 2013). The physical examination usually reveals a focal neurological deficit. Given that, the symptoms and time course vary with the locations and lesions.

The history of the effects of vascular disease, clinical assessment, and neuroimaging is an important diagnosis criterion of VD. A clear history of stroke or transient ischemic can always be found. The previous attacks due to those vascular diseases temporarily cause cognitive decline or neurological deficits. Cognitive assessment will objectively show the decline in the domains of complex attention and executive functions. The common symptoms of VD are gait disturbance, urinary symptoms, and personality changes or emotional instability. In the late-life period, the vascular neurocognitive disorder may be associated with depression, causing psychomotor and executive dysfunction, so-called vascular depression (Hugo & Ganguli, 2014; Kulkantrakorn, 2018).

iii. Dementia with Lewy bodies (DLB)

DLB accounts for almost 10-15 % of diagnosed dementia (Alzheimer’s Society, 2016). The growth of intracellular bodies (Lewy bodies) accumulates in the brain caused by a gene mutation. The underlying abnormal accumulation is primarily characterized by protein misfolding and aggregation within the pathognomonic Lewy bodies (Meeus et al., 2012). The common symptom of DLB is an atypical movement which is also found in Parkinson’s disease. DLB often occurs at the same time as Parkinson’s disease or within a year. The affected person’s cognition is prominently impaired in attention, visuospatial, and executive function. Markedly, the core characteristics of DLB are cognitive dysfunction, recurrent visual hallucinations, and parkinsonism. In the late stages of Parkinson’s disease, dementia symptoms can be found, so-called Parkinson’s disease dementia. It is thus believed these two diseases are common illnesses with different dominant symptoms in each onset stage (Gomperts, 2016).

Based on the temporal order of the movement dysfunction and the cognitive impairment, DLB and Parkinson's disease can be distinguished from one another. Cognitive decline occurs prior to parkinsonism in DLB, which is similar to the other cause of dementia, where cognitive impairment is the prominent symptom. In comparison, Parkinson’s disease reveals cognitive impairment later than a movement disorder Other symptoms that point to DLB include REM sleep behavior disorder, severe neuroleptic hypersensitivity, and poor dopamine transporter uptake in the basal ganglia, which can be shown using SPECT or PET imaging (Yamada et al., 2020). Various clinical signs that can be found include systematized delusions, sadness, frequent falls and syncope, brief, unexplained unconsciousness, severe autonomic dysfunction, hallucinations in various modalities, and severe autonomic dysfunction (Hanağası et al., 2016, pp. 400-402).

iv. Frontotemporal dementia (FTD)

FTD is rare compared to the other causes mentioned previously. as its prevalence is only 2.7% of all dementia. It, however, is the common cause of early-onset dementia. FTD generally develops in middle adulthood with the prominent characteristics of emotional disturbances. Cerebral atrophy occurs at the frontal and temporal lobes, with the predominant protein aggregation including hyperphosphorylated tau or ubiquitin protein. These two lobes dominate a range of performance and cognition, such as personality, emotions, behavior, thinking, and language. A patient with FTD may present with these different symptoms based on where the lobes are most affected(Ahmed et al., 2014; Kulkantrakorn, 2018).

The most prominent characteristics of FTD are behaviorally displayed in personality and behavior changes, a loss of interest in interpersonal activities and responsibilities, social withdrawal, a loss of personal hygiene, and an increase in socially dis-inhibited behavior. FTD early symptoms are different from that of AD, where memory loss is the primary evidence. FTD is more akin to psychiatric disease to frequent misdiagnosis as a major depressive or bipolar disorder. It also reacts with psychosis medicine and antidepressant drugs. Other evidence of FTD is perseverative or compulsive motor behaviors, hyperorality, and dietary changes. The neuropsychological profile of FTD is executive deficit (Miller & Yoon, 2016, pp. 392-394).

v. Mild cognitive impairment (MCI)

An individual with MCI will have problems with cognitive functioning, which will require diagnosis through clinical examination, not just from subjective complaints. In fact, approximately 7-25 % of individuals over the age of 60 have mild neurocognitive disorders (Jongsiriyanyong & Limpawattana, 2018). Unlike dementia, independence of functional activities and social engagement are preserved in MCI. MCI is considered the early onset of neurodegenerative disease and, therefore, can progress to dementia and other diseases, including Parkinson’s disease. Approximately 15% of people with MCI can progress onto Alzheimer’s disease within two years, and the progression rate is higher in later years; for instance, the rate can be 33% after five years (Alzheimer's Association, 2022). MCI prevalence is also associated with age; the probability of having MCI is 6.7% in people aged 60 – 64, 8.4% for 65 – 69, 10.1% for 70 – 74, almost 15 % for 75 – 79, and a quarter of 80 – 84 years-old (Petersen et al., 2018). While MCI may not always develop into dementia, , early detection is helpful for diagnosis, appropriate intervention, education, psychosocial support, and participation in shared decision-making about life planning, health care, research involvement, and financial matters (Dementia Australia, 2018; Muangpaisan et al., 2012; Petersen et al., 2018).

As MCI is a relatively mild dysfunction, patients and their family members usually have difficulty recognizing it until the symptoms have become a significant problem. In addition, most clinicians can misdiagnose these functional problems with those encountered by subjective memory complaints and also normal-aged individuals. The National Institute on Aging and the Alzheimer's Association (NIA/AA) stated that the points separating normal cognition and MCI, and MCI and dementia are blurred. Consequently, clinical decision-making is vital in determining these distinctions (Petersen & Morris, 2005; Sachdev et al., 2014). Any poor performance in one or more cognitive functions needs to be evaluated by comparing them to the patient's previous performance levels and also with performance of people of the same age and education level. The impairment is across a range of cognitions, including memory, executive function, attention, language, and visuospatial skills. Episodic memory is the most common cognitive decline in MCI patients who subsequently progress to a diagnosis of AD dementia. Individuals with MCI usually display a lower ability to learn and retain new information due to the decline of their episodic memory (Winblad et al., 2004). Most of all, memory complaint in MCI needs to be observed by an informant and through objective evidence (Petersen, 2004). The study of Petersen et al. (2005) operationally defined a memory deficit as performing 1.5 SD below an educated-adjusted norm on a 20-minute delayed recall of the logical memory subtest in the Wechsler Memory Scale. This criterion is now usual taken as an objective level for MCI, with values of either 1.0 SD or 1.5 SDs below the stated norms on memory tests (Chertkow et al., 2007). At the same time, the common criteria of Petersen and NIA/AA of daily activities being intact and lack of dementia (Albert et al., 2011) apply.

Petersen (2004) proposed memory impairment as the differential criterion for MCI subtypes on the bases of the discussion at the Key symposium in

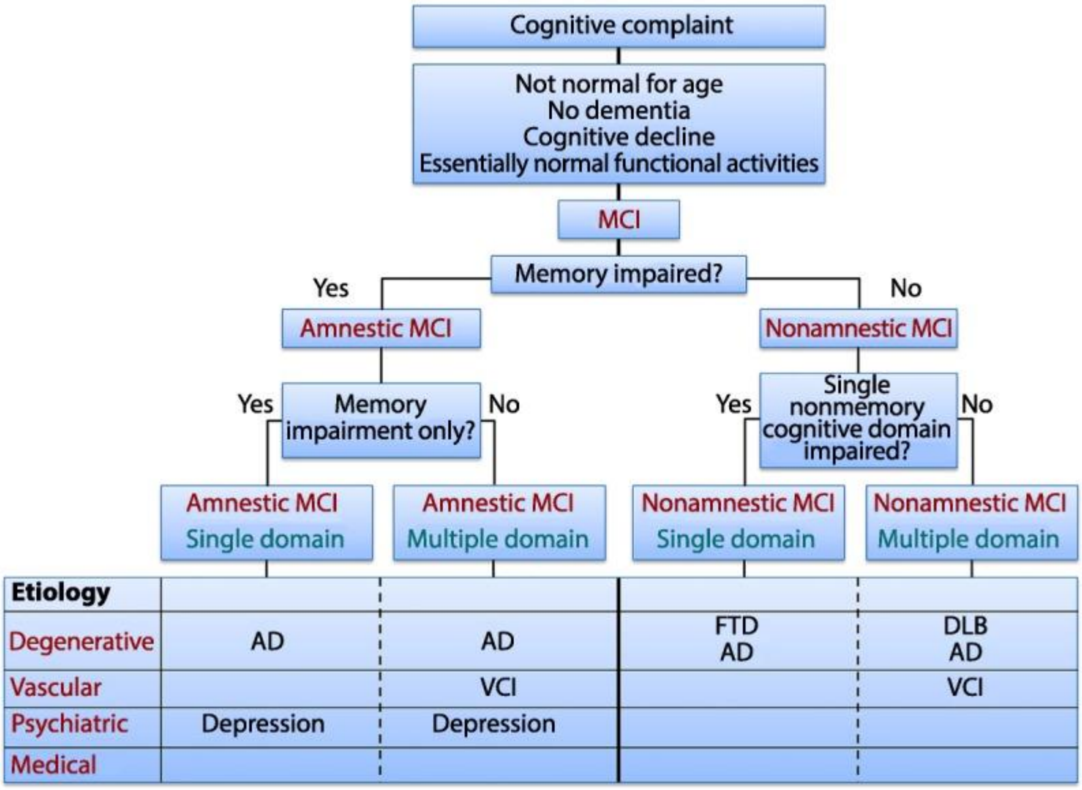
Stockholm, 2003. The subtypes are primarily considered with respect to possible dysfunctions of memory; two main subtypes consequently are called an amnestic MCI and non-amnestic MCI. Then the number of non-prominent cognitive domains are determined (Petersen, 2016). Figure 3 depicts the MCI subtypes and the possible progression of each subtype. Subsequently, the MCI characteristics regarding the Key symposium criteria can be categorized into four patterns of subgroups.

1) Amnestic MCI with a single domain, or amnestic single, is described as presenting only memory impairment with relative preservation of other cognitive domains (Chertkow et al., 2007). This subtype is more likely to convert to AD (Petersen, 2016). Therefore, it is commonly addressed as the pre-dementia stage of AD (Chetelat et al., 2005).

2) Amnestic MCI with multiple domains, or amnestic multiple, refers to the subtle additional impairments beyond memory deficit, e.g., naming, attention, and executive function (Chertkow et al., 2007). This phenotype possibly develops onto AD or vascular cognitive impairment (vascular MCI) (Dementia Australia, 2017; Petersen, 2016).

3) Non-amnestic MCI with a single domain is characterized by preserving memory ability but showing impairment in a single non-memory domain (Chertkow et al., 2007). The most common impairments in non-amnestic MCI are language and attention (Dementia Australia, 2017). Patients with this subtype can possibly develop onto AD or frontotemporal dementia, as well as another neurodegenerative disease such as the primary progressive aphasia. (Dementia Australia, 2017; Petersen, 2016).

4) Non-amnestic MCI with multiple domains is described as involving multiple cognitive dysfunctions other than memory domain (Chertkow et al., 2007; Dementia Australia, 2017). This type of MCI may progress onto AD, dementia with Lewy bodies, and vascular MCI (Dementia Australia, 2017; Petersen, 2016).



Etiology				
Degenerative	AD	AD	FTD AD	DLB AD
Vascular		VCI		VCI
Psychiatric	Depression	Depression		
Medical				

Figure 3 Key Symposium MCI subtypes and possible etiologies (Reprint with permission from Petersen R C, Mild Cognitive Impairment, Continuum (Minneap Minn), 22, 2. journals.lww.com/continuum/Abstract/2016/04000/Mild_Cognitive_Impairment.7.aspx © 2016, American Academy of Neurology.)

MCI is considered a prodromal stage of AD with common memory deficits. According to the DSM-5, mild neurocognitive disorder is termed MCI or prodromal dementia, while dementia is a major neurocognitive disorder (Albert et al., 2011; Petersen, 2016). The declination of MCI in Alzheimer’s disease is observed in episodic memory impairment and potentially in additional domains (Chetelat et al., 2005). The moderate level of impairment brings the problem to light whilst the insidious cognitive decline is gradually and steadily progressing. MCI in Alzheimer’s disease typically has amnestic signs, which cause impairments to memory and related executive functions in the earlier onset of the disease course. The other cognitive domains may occur later, such as language functions, perceptual- motor functions, visuoconstruction, and social cognition. However, non-amnestic presentations also do occur. Depression and apathy are not the prominent characteristics of Alzheimer's disease, and these may occur throughout the clinical spectrum (Hugo & Ganguli, 2014).

2.1.1.2 Potentially reversible dementia

Reversible dementia is caused by particular conditions which can be controlled. With an early diagnosis, appropriate treatment focused on the primary cause of the condition can cure dementia-like symptoms. Prevalence is highly variable; a range between 8% and 40% was reported in many studies, while it was found to be 7.50% in a Thai cohort study (Muangpaisan et al., 2012). A meta-analysis study reported an overall reversion rate of MCI to normal at approximately 24% (Malek-Ahmadi, 2016). The same study found that the reversion rate of Asian studies (12.24%) is lower than the studies in North American (22.38%) and Europe (26.59%). The causes of treatable dementia are depression, alcohol-induced cognitive impairment, normal pressure hydrocephalus (NPH), and vitamin B12 deficiency, those account for more than half the cases. The other causes are hypothyroidism, chronic subdural hematoma, anxiety, or the side-effects of neurological drugs, particularly Benzodiazepines. Accurate cause recognition is important in treatment plans, as reversible cases may be untreatable after delay and inaccurate diagnosis (Chari et al., 2015). As the cognitive deficits in reversible dementia are similar to those that characterize progressive dementia, including amnestic performance, attention, concentration, mental reaction, and occupational performance (Chari et al., 2015; Kulkantrakorn, 2018), it is important to distinguish the dementia type early in diagnosis (Neurological Institute of Thailand, 2014).

2.1.2 Clinical characteristics

The spectrum of cognitive impairment due to dementia is both wide and specific. The impairment of a person with dementia involves cognitive function, which impacts activities in daily life and causes distress. Social involvement is either the cause or effect of cognitive decline. Although the characteristic of dementia is similar to several diseases and aging features, one should be aware of the differentiation between dementia and delirium, depression psychosis, or reversible dementia (Neurological Institute of Thailand, 2014).

2.1.2.1 Cognitive functions

Amnesia is the most recognized and discussed feature of the early stages of dementia discussed by non-clinical professionals vis a vis other cognitive impairments (Arvanitakis et al., 2019). The most common presentations of amnesia are usually connected to recent events, e.g., struggling to recall telephone numbers or shopping list items, forgetting conversations or events of the day, and losing or misplacing items (NIH National Institute on Aging, 2020). Memory declination, over the course of time, will see an increase in the severity level of the amnesia with the involvement of long-term information forgetting (Goldstein & McNeil, 2012, pp. 163-165). This decline will manifest through actions such as the patient getting lost in their familiar surroundings or on their usual routes or struggling to recognize family members or to name ordinary items (Grabowski & Damasio, 2004). Memory impairment affects other cognitive domains by increasing the difficulty of retrieving information. For example, orientation can be progressively affected by the result of forgetting person, place, and time information. Language deficit can be involved when the patient has difficulty in naming objects, retrieving words, possessing a whispering voice or disconnected speech, or speaking using ungrammatical sentences (Klimova & Kuca, 2016). As personal behaviors result from a combination of cognitive processes, a single variant action may bring about both major and minor cognitive dysfunction.

Identifying the dominant cognitive impairment in different causes of dementia is vital for diagnosis. While the cognitive hallmark of AD is impaired long-term episodic memory (Gomez & White, 2006), the dominant neuropsychological profile of FTD is defective executive function. AD patients typically demonstrate anterograde amnesia and visuospatial disorientation, causing them to struggle in daily activities. The patient with FTD better preserves the ability to remember recent information and visuospatial orientation on objective testing. Language domains also provide differential indicators of dementia causes. DLB displays obvious speech problems due to atrophy in the motor areas of the brain. Patients with DLB have a loss of verbal fluency at the beginning of impairment until complete muteness at the late stages (Klimova & Kuca, 2016).

2.1.2.2 Activities of Daily Living

Impairment in activities of daily living (ADL) occurs when a person's personal care and occupational activities are worse in comparison with their previous levels. An elderly person may need greater assistance, but they are able to maintain some independence in some day-to-day activities. The performance of ADL is a predictive factor for dementia and can be measured in two types of activities, including activities with and without instruments. The impairment of instrumental ADL is significantly present in dementia patients, with respect to controls, in shopping, food preparation, or use of transport. The level of impairment in dementia is also higher than in MCI patients (Ouchi et al., 2016).

2.1.2.3 Emotion and behavior

The symptoms of dementia potentially cause stress to the patients; it may further exacerbate the variant behavior. When patients are well aware of their diagnosis, feelings of stress, confusion, uncertainty, and insecurity are common. The incomprehensibility and unpredictability of the disease, and its related changes, (Steeman et al., 2006) motivate behavior changes, and these can cause distress to family members or caregivers. The most painful behaviors may be incessant wandering, aggression, paranoia, sexual disinhibition, and depression. Behavior changes resulting from different causes of dementia, such as delusion, frequent falls, and a REM sleep behavior disorder, are often presented in DLB. While psychiatric symptoms such as hallucination and delusion are not only demonstrated in FTD, those may also present in the early stage of AD (Morris & Nagy, 2004). Behavior problems are frequently the precipitating factor of the need for institutional care.

2.1.2.4 Brain area related to dementia

Brain atrophy in dementia is present in the area where the abnormal proteins progressively accumulate; the dominant symptoms are a consequence of certain brain areas are interrupted or lost. In brain imaging studies, the initial site of the abnormal brain accumulation occurs along the hippocampal pathway (entorhinal cortex, hippocampus, and posterior cingulate cortex). These areas are related to early memory deficits. The Hippocampus and entorhinal cortex are part of medial temporal lobe (MTL) structures prominently associated with AD (Buckner et al., 2008). As the Posterior cingulate cortex (PCC) has connectivity with the hippocampus, a decrease of both PCC and hippocampus regions affects the default mode network in Alzheimer's patients, resulting in the disruption of attention processing (Greicius et al., 2004; van den Heuvel & Hulshoff Pol, 2010). For MCI caused by AD, language networks are found to be a potential discriminant indicator for MCI from NC better than EF networks. Pistono et al. (2021) proposed that the right inferior frontal gyrus, the right superior temporal gyrus, the left middle temporal gyrus, and the left middle temporal gyrus/angular gyrus comprise the language networks. In the later stages, neural loss in the temporal, parietal, and frontal regions of the cerebral cortex is associated with impairments in language, visuospatial, and behavior. (Frisoni et al., 2010).

2.1.3 Assessments

Techniques of neuroradiological, neurochemical, electroencephalogram neuropsychological are suggested for early screening, differential diagnosis, and evaluating of treatment for dementia. They are grouped in three categories and presented in the below sections.

2.1.3.1 Biomarkers

A biomarker is an objective indicator used to measure or assess normal or abnormal biological conditions or a response to drug treatment. Biomarkers are used to distinguish the causes of dementia and differentiate between degenerative states or other diseases and detect the pre-stage pathology, monitoring progressive decline and treatment. Ahmed et al. (2014)'s research suggested that biomarkers related to dementia are varied at different stages of the degenerative process. Different protein-misfolding conditions are used to define the etiology of dementia; it is thus an important biomarker of cognitive impairment disease. Neurodegenerative proteinopathies are the abnormal accumulation of proteins or peptides that cascade into complex molecules, causing progressive neurodegenerative diseases such as dementia. The major protein biomarkers in dementia are Aβ peptides and tau in AD, α-synuclein in LBD, and tau or TAR DNA-binding protein (TARDBP; alias TDP-43) in forms of FTD (Sonnen et al., 2008). While tau is typically higher in AD than in DLB, tau can be elevated in rapidly progressive cases of AD. Those biomarkers can be identified by biological specimens and brain imaging techniques. For example, dopaminergic loss in the basal ganglia as a suggestive feature of DLB can be revealed by Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT) measures. These techniques can be used to differentiate DLB from AD and some forms of FTD with reasonable diagnostic accuracy (Ahmed et al., 2014). Currently, biomarkers for Alzheimer's disease can be explored with several methods, none of which the literature suggests be used alone (McKhann et al., 2011). Several promising biomarkers can be investigated through neurofibrillary tangle lesions in brain imaging, tau in cerebrospinal fluid (CSF), blood and urine tests, and genetic risk profiling. Those techniques give different neurodegenerative markers and specific biomarkers (Cummings, 2012).

Although abnormal protein aggregations cause neuron apoptosis, there is an underlying process due to genetic markers and synaptic dysfunction. The Apolipoprotein E (APOE) gene produces apolipoproteins which carry cholesterol and fat in the bloodstream; the dysfunctional process increases protein aggregations. Especially APOE ε4 allele, this form of APOE associated with AD patients and early onset of AD, is called the risk-factor gene of dementia. While the APOE ε4 allele increases the risk of AD in late-onset, APOE ε2 was found as a protective factor that decreases the risk (Albert et al., 2011). Another biomarker associated with this genetic risk factor is the tau protein. However, the APOE ε4 allele is remarkedly related to Aβ; it is an unclear mediated process among tau concentration in cerebrospinal fluid and neurofibrillary tangle formation of AD (Tachibana et al., 2019; Zhou et al., 2016). Similarly, FTD relates to at least three genetic markers. The family history of autosomal dominant patterns plays an important role in 30% to 50% of FTD patients, especially tau protein accumulation. In DLB, tau levels are variably found in CFS. While the levels of CSF tau are very variable in DLB, typically being lower than in AD, although in rapidly progressive cases, they can be very elevated.

While biomarkers are accepted as promising indicators of dementia, laboratory methods may give different levels of stability and reliability (Alzheimer's Association, 2020b). The clinical techniques can be divided into imaging modalities, which consist of structural and functional brain imaging, cerebrospinal fluid (CSF) measures, and blood or urine tests. Brain imaging tools are also widely used in clinical and research settings. While both biological specimens are available for routine clinical assessment, blood and urine tests are less commonly used today. Although the developed techniques are more reliable, the laboratory service fees may limit their use. (Ahmed et al., 2014).

i. Structural brain imaging

UK, European, US, and Thailand guidelines recommend structural brain imaging for all patients being investigated for dementia. In Thailand, dementia clinical practice guidelines strongly recommend structural imaging for differentiating from reversible dementia and confirming etiology for treatment planning. These processes are not recommended in early diagnosis (Neurological Institute of Thailand, 2014). White matter signal alterations and vascular injury with a variety of reasons can be assessed using Magnetic Resonance Imaging (MRI) and Computerized Tomography (CT). The diagnostic criteria for stages of the condition include neuronal loss (atrophy) in certain parts of the brain as a positive predictor for various dementias (Petersen et al., 2000).

ii. Functional neuroimaging

Functional neuroimaging and related neuroimaging techniques are useful for dementia in displaying how the effect of neurotrophy on brain systems pertains to specific cognitive functions and behaviors. Functional neuroimaging allows medical practitioners and researchers to observe brain function over a period of time; these techniques are used to determine the severity of the effect and also changes due to rehabilitation. The functional neuroimaging tools measure an alteration of brain-process representing substances. Functional Magnetic Resonance Imaging (fMRI) uses as a representative of blood-oxygen-level-dependent (BOLD) signal captured by the magnetic properties of the cerebral venous blood in interested regions (Sonnen et al., 2008). Positron emission tomography (PET) using 18-F-fluorodeoxyglucose (FDG) and single photon emission tomography (SPECT) are two of the most commonly used brain imaging techniques in research because of their high reliability and requiring of professional practice (Crosson et al., 2010). Both techniques provide visualization and quantification of patterns of brain hypometabolism and hypoperfusion. Each type of dementia has a different characteristic pattern. Although functional imaging techniques offer accurate information, those tools require considerable expertise and a qualified set of devices to operate and analyze, which limits their widespread application as biomarkers (Brown et al., 2014; Sonnen et al., 2008).

iii. Cerebrospinal fluid (CSF)

In the context of dementia, the CSF biomarker is often used to distinguish an infection, malignancy, and inflammation in the nervous system. CSF examination is recommended in persons with cognitive impairment under 55 years of age, presenting with a rapid worsening progression or atypical dementia symptoms, or who are immunosuppressed. CSF analysis relies on a variety of immunochemical techniques that measure the range of neuron-specific or neuron-enriched proteins. The use of neuron-enhanced CSF markers β-amyloid and tau in the usual assessment of dementia varies greatly between countries and clinicians. Under Thai guidelines, CSF examination is only prescribed by a medical doctor with expertise. CSF is used in the differential diagnosis of specific dementias, e.g., AD is identified by an increasing of CSF hyperphosphorylated tau, decreasing of CSF βA decreased, and plasma Aβ (Neurological Institute of Thailand, 2014). In research, these biomarkers were increasingly combined with the new AD diagnostic criteria and were thoroughly considered for inclusion in further clinical trials (Budson et al., 2016; Sonnen et al., 2008).

iv. Blood and urine

The clear advantages of using blood or urine-based biomarkers for dementia include the profiles of vitamins, glucose, protein, or electrolytes (Neurological Institute of Thailand, 2014). Their profiles, however, give only indirect evidence of dementia. Protein molecules from the brain are much less concentrated in blood or urine than in CSF due to the activity of the blood-brain barrier and the large blood and urine volume in which these proteins are diluted. Moreover, the binding of many proteins of interest in veins increases the chance of being removed from the blood faster, causing a reduction of sensitivity. Dementia blood profile is, thus, largely used to explore coexisting dysfunction such as vitamin deprivation, age-associated immunosenescence, or inflammation (Chen et al., 2017; Sonnen et al., 2008).

2.1.3.2 Electroencephalography (EEG)

Even if the biomarkers yield high accuracy, they are expensive, less likely to be available in primary healthcare settings, and/or partially invasive, making them unsuitable for widespread AD risk assessment in large populations of elderly people. Electroencephalographic (EEG) indicators, in contrast, are more affordable, widely accessible, and entirely noninvasive, which are perfect features of regular clinical practices (Vecchio et al., 2013). Without requiring brain surgery to implant a sensor to record brain waves, the EEG approach offers the benefit of detecting the background electrical activities of the brain produced by neurons of the cerebral cortex effectively and non-invasively. (Amezquita-Sanchez et al., 2019; Poil et al., 2013; Vecchio et al., 2013). Additionally, using details on the positioning, sensitivity, number, and orientation of the sensors, tissue conductivity, and head geometry can result in the discovery and postulation of mechanisms of relationships between imaging and EEG data (Amezquita-Sanchez et al., 2019).

The study of quantitative EEG (qEEG) and event-related potentials (ERPs) as clinical indicators of the early stages of AD has received a lot of interest in recent years (Mazrooei Rad et al., 2021; Vecchio et al., 2013). ERPs are produced in conjunction with an event or when the brain experiences sensory stimulation (Mazrooei Rad et al., 2021). The amplitude and latency indices of the ERP components are affected by the excitation signal's amplitude, frequency, and excitation intervals. It has been demonstrated that in healthy people, the magnitude of the P300 component grows when the target stimulus frequency lowers or stimulus intervals widen. Additionally, this component's amplitude and delay vary with age, AD stages, and even the types of dementia. (Mazrooei Rad et al., 2021).

A fully standardized approach that can be executed quickly and simply in a clinical setting is the recording of resting state cortical EEG rhythms with the eyes closed (Poil et al., 2013). Unlike ERPs, the use of resting state EEG rhythms does not necessitate the use of stimulation devices or the recording of a participant's activity. It is also less prone to the tiredness and anxiety that are usually connected to task performance. This is ideal when EEG recordings are performed on elderly individuals (Vecchio et al., 2013). Mostly, frequency band characteristics are frequently employed as data processing tools in EEG. The alpha frequency band in patients with mild AD and dementia with Lewy bodies is reduced compared to healthy elderly, while the amplitude of patients in mild stage is increased in the delta frequency band (Mazrooei Rad et al., 2021). The use of EEG biomarkers in MCI and AD, like spectral measurements and synchronization between brain regions, has already been extensively documented in the literature. Lower cognitive performance and hippocampal atrophy are connected with decreased alpha power (Gaubert et al., 2019).

2.1.3.3 Neuropsychological tests

Neuropsychological assessment is of benefit for differential diagnoses between psychiatric and neurological symptoms and/or for identifying progression of neurological disorders in non-psychiatric groups. It also provides information for distinguishing between various neurological conditions, and behavioral data for lesion localization, or at least the specific hemisphere site of brain atrophy. Neuropsychological tests can be used as screening tools in the early diagnosis process (Lezak et al., 2012). The multiple cognitive domain assessment should be performed prior to single domain tests (Neurological Institute of Thailand, 2014).

Early assessment or screening test for cognitive impairment would assist patients and their families in receiving proper care and support at an earlier stage in the disease course, which could lead to both improved prognoses and extending patients' life (Lin et al., 2013). The screening tests are recommended in early diagnosis for both multiple and specific cognitive domains. The patient performance-related tests assign cognitive tasks to patients, and the test scores reflect the deteriorated and preserved cognitive abilities. Cognitive tasks in dementia screening tests mostly include memory, language, spatial ability, orientation, attention/calculation, and executive function, respectively (Na Chiangmai & Wongupparaj, 2020).

Although amnesia is a dominant characteristic of dementia, it is essential to assess additional domains apart from memory. These domains consist of attention (both simple and divided attention), language (including naming, fluency, expressive speech, and comprehension), executive functions such as reasoning, problem-solving, planning, and visuospatial skills (e.g., clock drawing, copying cubes, intersecting pentagons). It is suggested that several validated clinical neuropsychological measures are used to examine a single cognitive ability. For instance, the language domain can be tested by Boston Naming Test and letter and category fluency. Digit span forward is useful for attentional control, Trail Making Tests for executive function, and figure copying for spatial skills (Albert et al., 2011; Petersen, 2004)

As mentioned, assessment of multiple cognition functions should be performed prior to single domain testing in both clinical settings and research. Two well-known neuropsychological assessment tools used in dementia are the mini-mental state examination (MMSE) and the Montreal Cognitive Assessment (MoCA), both of which are available in the Thai language. While these tests are standardized and used across many countries, direct literal translations have been found to lead to different meanings and/or syllabic format from the original instrument (Shim et al., 2017). The MMSE is suitable for dementia screening, while MoCA is more sensitive to MCI (Julayanont et al., 2015; Nasreddine et al., 2005). In 1993, the Train the Brain Forum Committee developed an inclusive test known as the Thai Mental State Examination (TMSE). It incorporated Thai realities, including socioeconomic status, level of education, and culture of the examinees (Train The Brain Forum Commitee, 1993). Almost thirty years later, the cut-off scores of this test have not been adjusted in response to the changes resulting from the current aging of Thai society. The psychometric properties of the TMSE are thus changed. The TMSE provided low sensitivity and specificity in detecting early dementia (Silpakit, Silpakit, et al., 2017).

Dementia has at least four etiologies, and AD is the most common cause of dementia. The prodromal stage of dementia is MCI. This stage is difficult to detect because of the subtle presenting of cognitive impairment and preservation of daily activities. However, the prominent symptom of AD and amnesia-MCI is a memory deficit. The most common occurrences of amnesia are usually connected to recent events, e.g., struggling to recall telephone numbers or shopping list items, forgetting conversations or events of the day, and losing or misplacing items. Hence, early detection is important for intervention and life planning for patients and their families. Different dementia assessments present trade-off properties between accuracy, cost, accessibility, and invasive method. Three methods are mentioned, including biomarkers, EEG, and neuropsychological tests. This study chose the most easily accessible method, i.e., the neuropsychological test. Since neuropsychological tests can optimally use as dementia screening for Thailand nationwide. In addition, speech analysis received attention as a potential approach to improve the screening and allow early detection of symptoms. Speech disfunction reflects several dimensions of the cognitive impairment of a person, not only language ability but also memory and related cognitive functions. Speech analysis in detecting dementia. There were the studies which adapted paper-pencil form of neuropsychological assessments onto alternative platform such as telephone and video conference (Sánchez Cabaco et al., 2023; Tröger et al., 2018). The next section presents information and literature regarding speech pathology due to neurodegenerative disease.

2.2 Language & Speech

The cognitive process of speech is at the intersection of attention, decision-making, memory, and language. Speech processing is an expressive part of language; the other important part of language is the receptive component. The neuroanatomy of language revealed an importance of brain areas related to function and deficit in language and speech. The disorders of language and speech are separately presented in this section. Lastly, the speech deficits in dementia are stated to specifically characterize the symptoms found in the patients.

2.2.1 Language processing

Although the terms ‘speech’ and ‘language’ are frequently used interchangeably, each of these systems has distinct functions. Their operational systems depend on different sets of representations and processes; both share common purpose of communication (Zeki & Hillis, 2016, p. 123). Language is an intentionally and culturally driven system of voluntarily coupled signals and meanings (Fernandéz & Cairns, 2010). Two main aspects of language functions are comprised of receptive and expressive abilities, which grant comprehension and communication of information, respectively. For the purposes of understanding and communicating from sound to conversation, language processing involves several cognitive operations including retrieving vocabulary, concepts, or grammar; attentional control of processing abstract inferences, idioms, or verbal; problem-solving in working memory (Zeki & Hillis, 2016, p. 116 & 123).

Speech, additionally, consists of the further coordinated rapid motor function that accounts for the precise action of vocal expression for transmitting language. The basal ganglia, cerebellum, and cortical systems of the body, as well as the vagal, hypoglossal, and facial nuclei, play a role in the development of speech regulation. Eventually, phrenic nerves control and coordinate the muscles involved in the act of speaking, including the laryngeal, pharyngeal, palatal, lingual, oral, and respiratory muscles. Normally, two words per second spoken by a normal speaker are linguistically equivalent to 14 different sounds (phonemes), each phoneme relying on the contraction or relaxation of 100 muscle bundles (Zeki & Hillis, 2016, p. 123). When audible production is articulated, those sounds do not only convey information, but also linguistic features. This study was interested in acoustic features. Acoustic variables are related to intensity, duration, and frequency, which is a physical characteristic of speech (Boschi et al., 2017; Ladefoged & Johnson, 2015). A specific aspect of speech regards acoustic features combines of articulation (sounding), resonance (nasality), and pronunciation (voice controlling). The features are also often characterized by fluency and prosody (Zeki & Hillis, 2016).

Language proficiency depends on multiple cognitive domains, particularly perception, memory, and attentional allocation. It is also related to the mechanisms dominating the storage of concepts and structures in long-term memory and the mechanisms involved in retrieving and activating concepts and structures of working memory during a person’s ongoing language processing (Schmid, 2007). A variety of cognitive processes, which can be categorized into three fundamental basis, influence language use in an unconscious manner: (i) processes of social cognition, which concern the interaction between speaking participants; (ii) decision-making process in order to express thoughts; and (iii) memory-related processes that deal with storing, retrieving, and processing linguistic information (Diessel, 2019).

2.2.1.1 Attention & Social cognition

Language use is a specific type of social interaction that critically depends on the capacity to consider the information, intentions, and beliefs of other people. Joint attention is a fundamental social cognitive process. In order to communicate, the speaker and the listener should concentrate their attention on the common experience, which may involve an object or event in the surrounding situation or a concept that is evoked by the preceding discourse (Willems, 2015). Joint attention is a fundamental component of social contact, but communication requires more than just having everyone's attention on the same thing. As a result, communication requires that the interlocutors share a "common ground," which is defined as language users' awareness of their shared information. They also need to align their knowledge and views. The users are concerned not just with information about the physical speech situation surrounding the interlocutors but also with background information about the communicative partner and general world knowledge (Bung, 2016).

2.2.1.2 Decision-making process

Importantly, there are always several ways to state a specific communicative goal — or, more or less, the same thing. The range of linguistic tools that may be used to represent a certain communicative intention expands as language is used productively. Since there are frequently multiple ways to communicate essentially the same thing, speakers must make decisions. They must, in other words, "decide" how to convey a specific goal or meaning. Making decisions on the best linguistic strategy to use in a given scenario is a step in the development of language (Diessel, 2019; Goldstein, 2019).

2.2.1.3 Memory-related Processes

Language structure is driven by the semantic and pragmatic characteristics of speech and communication, according to functional and cognitive linguistics. However, recent literature has shifted their focus from communication and meaning to frequency and processing. Literature is interested in the features of linguistic information's representation and activation together with frequency and processing in memory (Diessel, 2019). In respect of verbal working memory, working memory is considered as a passive temporary storage of verbal information. Nevertheless, there is a disagreement about the previous interaction of language and memory. Working memory is reframed into both the maintenance and processing of verbal information. Since verbal information is always organized to serve goal-direct behavior, working memory is not only a passageway of information (Buchsbaum, 2016a; Schwering & MacDonald, 2020).

Memory is involved in speech production as a storage of long-term information and a control center. The phonetic and rhythmic representations are briefly stored in working memory as a temporal phonological store. A new model of working memory developed by Baddeley (2000) included speech production components that explain the cognitive functions of memory and language. Figure 4 illustrates the multi-component working memory model. In the model, the phonological loop, is a system where verbal and auditory information is held in the phonological store and repeated by the articulatory rehearsal process. Two functional components of the phonological loop are activated in listening, speaking, and subvocal speech (Baddeley, 2000; Goldstein, 2019, pp. 144-145). The message contains content and linguistic elements which are stored in LTM. The information is retrieved and sent into LTM via the episodic buffer. The buffer is an interface where crystallized cognitive systems accumulating in LTM and fluid capacities activating in WM are combined with a basis for conscious awareness (Baddeley, 2003).

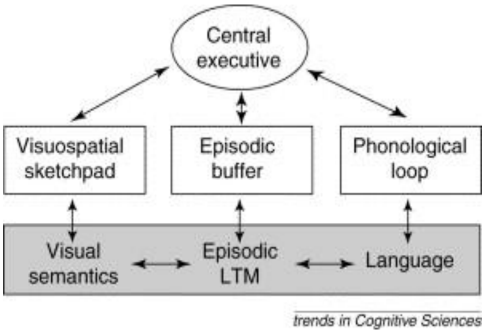


Figure 4 The current model of the multi-component working memory model. (Reprinted from Trends in Cognitive Sciences, 4, Alan Baddeley, The episodic buffer: a new component of working memory? , 7, Copyright (2000), with permission from Elsevier)

The central executive is responsible for binding information into coherent episodes. The components of WM then transform the linguistic representations into neuromuscular commands. The air from the lung, vocal folds, and glottis are the elements involved in producing the different phonemes. Meanwhile, a speaker monitors and controls his/her own speech via two feedback loops. To feel the movement and shape of muscles, the proprioceptive loop is in charge. The other feedback is the auditory loop, where the speaker's own speech is monitored (Cummins et al., 2015).

The phonological loop is a well-developed component of the working memory model since its function explains most behavior associated with verbal working memory (Buchsbaum, 2016b, p. 866). The fundamental role of the loop involves with the two components, namely the phonological store (the passive storage) and the articulatory rehearsal component (the active process) (Baddeley, 2000; Wongupparaj & Wongupparaj, 2012).

Four phenomena linked to phonology were explained by the revised Baddeley's working memory model: The phonological similarity effect; the word-length effect (Acheson & MacDonald, 2009; Buchsbaum, 2016b, p. 866); articulatory suppression; the transfer of information between codes.

1) The phonological similarity effect.

It was found that list of letters or words with similar sounds (e.g. rain, pain, point) are more difficult to remember than sound-dissimilar letters or words (e.g. rain, bug, desk) (Conrad & Hull, 1964). Since the overlapping representation increasingly interferes with the memory tracts of the shared phoneme or sound relative to the different words. This phenomenon relates to the passive buffer (Buchsbaum, 2016b, p. 866).

2) The word-length effect

Lists of longer words are more challenging to retain than list of shorter words. This phenomenon is assumed to arise because long words take more time to rehearse articulatory (Baddeley, 2000) and therefore, the overall elapsed time in the articulatory loops is greater in the set of polysyllable words (Baddeley et al., 1975). The longer rehearsal time for longer words results in less rehearsal of all words, and this may cause the decay of several items in the list. This time-based decay is related to both the rehearsal property and the capacity of storage (Baddeley et al., 1984; Buchsbaum, 2016b, p. 866).

3) Articulatory suppression

This phenomenon relates to the recall task. It happens when a person is engaged in covert or overt articulation (e.g. repeating the number “2” over and over) and this prevents the person performing an inner speech (subvocal) rehearsal during a delayed repetition task (Wongupparaj & Wongupparaj, 2012). The articulation can be an irrelevant sound. The articulatory suppression interferes with the ordinary mechanism of the articulatory rehearsal that allow to refresh the stored items (Baddeley et al., 1975), causing difficulty in recall performance (Buchsbaum, 2016b, p. 867).

4) Transfer of information between codes

The visual information (e.g., an orthographic input) must be re-coded into a verbal format before being committed to memory (Murray, 1968). The articulatory rehearsal reroutes an auditory code derived from visual code into phonological storage (Acheson & MacDonald, 2009). This phenomenon is supported by articulatory suppression. Articulation suppression effect happens when irrelevant information or sound is articulated during a person is trying to remembering an important information (Baddeley, 2000). The articulatory suppression blocks the articulatory rehearsal process. This effect prevents visual inputs from entering the phonological store (for recording from visual format into verbal format) and that information cannot be rehearsed. For auditory inputs, since verbal information can directly access into the phonological store, only rehearsal process is blocked (Baddeley et al., 1984; Henry, 2011, pp. 4-6).

2.2.1.4 Cognitive process involved in speech

Several models of speech attempt to describe cognitive representations and processes. Even though they do not agree on all aspects of speech production, they share some agreed-upon aspects on the following view of speech formation. Although several language production models are described different aspects of cognitive representations and process (e.g. Caramazza, 1997; Garrett, 1980; Levelt et al., 1999;), they share agreeable aspects in the following process (Bürki, 2018). Speech models are composed of three common processes: conceptualization, formulation, and articulation. The generation of utterance through three main processes is explained in the following parts.

The conceptualization and formulation processes are language processing, whereas articulation is observable product speech. The speaker establishes the ideas related to the message a person intends to convey during the conceptualization process. The grammatical, phonological, and phonetic encoding processes are the three encoding processes that constitute the formulation process. The formulation process cannot be clearly separated from the articulation process. During grammatical encoding, the cognitive process retrieves the syntax and semantic properties of the words (lemmas). These representations are combined with syntactic functions, resulting in the formulation of a syntactic framework (an ordered set of word and morpheme slots). The representations then create grammatical encoding and integrate with lexical–semantic representations, e.g., concepts and words form representations (Bürki, 2018).

The formulation process before the articulation of words begins with the phonological encoding process. Conceptual entities as input are connected with the representative words and their syntactic, morphological and phonological structure (Ramoo, 2021, p. 187). During phonological encoding, information regarding ‘frames’ and ‘fillers’ are early proposed. Given that word forms, two parts are represented including sub-lexical units (fillers) and related ‘metrical’ structure (frame). The frames and fillers are retrieved from the mental lexicon, forming a phonological word or phrase by inserting the fillers into the frames. Encoding of the phonetic unit is the last stage in the formulation process. Before utterance, the abstract phonological word or phrase is now mapped onto physical motor programs (Bürki, 2018). In psycholinguistic literature, little is known about the phonetic encoding process. The most well-known model of Levelt et al. (1999) portrayed speech production in which a person needs to access a repository of gestural scores for the frequently used syllables of the language or a syllabary. The novel or less frequent syllables are estimated when initially learned; a person thus inaccurately pronounces new words or phrases. The articulatory motions and their temporal correlations are described in these syllable scores.

The phonological syllables creating utterances of the word activates the articulatory gestures. Followingly, retrieval or computation of the phonetic syllables. Free parameters are set to specify the loudness, pitch, and duration of the curtain syllables, as well as the time of the articulatory activities. When the motor activity is completed, the production process ends (Bürki, 2018; Cummins et al., 2015; Postma, 2000). The articulators or mechanisms involved in producing speech are composed of an airstream, phonation, oral-nasal, and articulatory processes. Motor programs calculate prior to controlling the organs in each path of utterance mechanism, e.g., the lung, the muscle around the ribs, vocal cords, glottis, pharynx, or tongue. The air from the lung travels through the organs and is re-contoured until the word is out (Ladefoged, 2006).

2.2.2 Neuroanatomy of language (Localization)

The research in language processing and brain localization benefited from functional brain imaging, also on the basis of the classic lesion-based studies Language processing in the human brain is strongly left-hemispheric dominated, with two key regions grounding the network, as has been known since the second half of the 19th century (Kemmerer, 2015). Both regions, Broca’s area, and Wernicke’s area, are named after their founders. They are located on the inferior frontal and lateral temporal cortex, respectively. The two areas are notably crucial structures for language processes implemented by the normal brain (Andric & Small, 2015). Language localization was explored through studies of aphasia or other language disorders (Dronkers et al., 2017). The abnormal language symptoms were often associated with the incidence of a stroke, the occlusion or rupture of a blood vessel to a portion of a cerebral hemisphere. These conditions cause damage to brain tissue, and the damaging areas and aphasia relationship were thus reported (Goldstein, 2019, p. 39).

In order to communicate verbally, a person needs to comprehend the received matter and then convey his/her idea through spoken words. In 1861 Paul Broca noticed that some stroke patients who could not speak still understood language perfectly well. The patients could normally utter isolated words, whistle, and sing a melody but had difficulty in creating complete sentences, producing well-formed grammatically sentences, and expressing ideas in writing. When patients died, the brain was examined, and found a lesion in the inferior posterior region of the frontal lobe. This was hypothesized to be the expressive language model and labeled Broca’s area. The other important localization of language was reported in 1876 by Karl

Wernicke, who found the counterpart of Broca's aphasia. His patients were characterized by fluent speech but meaningless sentences. Notably, 'fluent' in this context solely means the flow of speech. Wernicke found that these patients had damage at the border of the temporal and parietal lobes; this area was named after him. Broca's patients could comprehend language but not speak, and Wernicke's patients could articulate fluently but failed to understand language. The two language patterns were hypothesized as receptive as opposed to expressive malfunction (Goldstein, 2019; Kandel et al., 2013).

Both locations of cortical language regions facilitate the cognitive function of speech. Wernicke's area is located at the intersection; most of the area is in the posterior temporal lobe, where the parietal and occipital lobes meet. This region lies near the primary auditory cortex and the angular gyrus, which is a cross-modal hub of information from other senses (Seghier, 2013). Therefore, Wernicke's aphasia causes a problem in understanding speech. Over time, Broca's area was widely recognized as an essential part of speech production, where it lies next to the motor cortex that controls the movement of articulatory organs. The roles of Broca's area are meditating on sensory representations of words in temporal lobes to the corresponding articulator in the motor cortex (Hickok, 2012). The coordination of information transformation (phonological word representation to articulatory code) is also controlled by Broca's area prior to words being spoken (Flinker et al., 2015). The two cortical language cortices collaborate through a bidirectional pathway, part of which is made up of the arcuate fasciculus. Receptive malfunction results in the patient speaking using irrelevant content (Kandel et al., 2013).

2.2.3 Disorders of Language

Aphasia or disorders in communication concern difficulties in understanding and/or expressing verbal and written language. Disorders of language in reading (alexia/dyslexia) and writing (agraphia/disgraphia) are typically associated with aphasia. The exploration of communicative disorder has been noticed in vascular cases, which can tell the relationship between brain atrophy and the disorders (Gindri & Fonseca, 2012). The aphasia model of such cases is profitably applied to the other causes of nervous disease, including neurodegenerative disorders (Zeki & Hillis, 2016).

2.2.3.1 Broca's Aphasia

Broca's aphasia is referred to as difficulty in expression, which is also called 'non-fluent,' 'motor,' or 'anterior' aphasia. The lesion or dysfunction in the posterior inferior frontal cortex is a precedent causing the symptoms, now known as Broca's aphasia. Speech dysfunction is characterized by troublesome speech, using simple phrasal verbs or nouns and leaving the grammar (telegraphic speech), or being interrupted by word-finding pauses. Both speech and writing are defined by agrammatic sentences and omissions or substitutions of function words.

The three hallmark symptoms of Broca's aphasia are preserved comprehension, agrammatism, and verbal apraxia. Dysfunction of the nervous motor system is known as apraxia. When motor programming of speech articulation is disturbed, it is called apraxia of speech. Patients with infarction solely in Broca's area typically have a brief deficit in motor speech. Their conditions present apraxia of speech in short period of course before quickly recovered. A person with this speech apraxia can be aware of their problem due to the attempt to correct the misarticulation by trial- and- error, repetitively yet still irregularly. These patients speak with irregular articulatory movements. This apraxia can be recovered very quickly. Although the syndrome can cause deficits in object naming, the concerning issue is deficits in action naming. (Zeki & Hillis, 2016).

2.2.3.2 Wernicke's Aphasia (receptive)

Wernicke's aphasia is a syndrome of the receptive language disorder. Sometimes termed 'posterior,' 'sensory,' or 'fluent' aphasia. The brain damage that causes Wernicke's aphasia can be found in Wernicke's area itself, which covers the posterior temporal lobe and inferior parietal lobe, or at the middle cerebral cortex where the artery supplies Wernicke's area. Lesions on the left hemisphere of the area which is the dominant language hemisphere (for most people), cause impairments resulting in meaningless articulation without disrupting fluency of both spontaneous speech and repetition. The main symptoms of this syndrome consist of impairment in language comprehension, failure in repetition, word-finding difficulty or anomia, semantically word substituting, i.e. the so-called semantic paraphasia, phonologically related word or nonword substituting such a phonemic paraphasia, and creating new words called neologisms (Zeki & Hillis, 2016).

2.2.3.3 Primary Progressive Aphasia (PPA)

Various neurodegenerative diseases cause progressive and selective language disorders; this set of syndromes is called Primary Progressive Aphasia (PPA). This syndrome is often associated with frontotemporal lobar degeneration (Boschi et al., 2017; Rohrer & Warren, 2016). It does not acutely happen after an accident or stroke, as the operational definition is a period of two years of progressive aphasia. Similar to MCI, activities of daily living and other functions rather than language are preserved. PPA also has common features with AD, such as difficulty in word finding. The language impairment in PPA gradually progresses from single dysfunction, such as anomia, to dissolution of language function or eventual mutism (Marczinski & Kertesz, 2006).

The subtypes of PPA are mixed with language impairment, from comprehension to expression. The current criteria define PPA subtypes into three groups, namely progressive non-fluent aphasia (PNFA), semantic dementia (SD), and logopenic/phonological aphasia (LPA). Patients with PNFA are also called a non-fluent variant and resemble Broca’s aphasia. There are many underlying deficiencies that can cause dysfluent speech, including a motor speech impairment or apraxia of speech and difficulties in computing proper grammatical sentences, which is called agrammatism. Consequently, these patients speak with shorter sentences or phrases, slow rate of speech, hesitation, and effortfulness. SD is described having semantic variant or fluent aphasia, similar to Wernicke’s aphasia. Patients with SD present with anomia and single-word comprehension deficits secondary to verbal semantic impairment. Non-fluent aphasia may occur with various primary language or speech dysfunction. The LPA subtype presents a slow speech rate, with long word-finding pauses, and occasional phonological paraphasias. Repetition of sentences is impaired. Grammar and articulation are usually preserved (Rohrer et al., 2012; Rohrer & Warren, 2016).

Although PPA subtypes are similar to Broca’s and Wernicke’s aphasia, the causes and disease progression are different. The latter two aphasia are caused by the acute aphasia syndromes of stroke, which result in changes in functional, structural, and neuroanatomical patterns of the language-related brain area. However, PPA closely corresponds with the frontotemporal dementia spectrum with clinical, genetic, and pathological overlap. PPA patients may especially present with a behavioral variant which is a prominent characteristic of FTD (Rohrer & Warren, 2016).

AD can lead to some form of aphasia, since normal language function relies on a wide range of brain regions which possibly be damaged by neurodegeneration (Weekes, 2020). To be clearer, the neural network damaged by AD defines the symptoms of aphasia, not AD itself. In autopsy-proven cases of AD presenting with progressive aphasia, the atrophy pattern included the left posterior superior temporal lobe, inferior parietal area, medial temporal lobe, and posterior cingulate region. These atrophy areas are partially in Wernicke's area and Broca's area (Hickok, 2012; Seghier, 2013). The medial temporal lobes are considered as the initial affected region in AD, episodic memory is thus observed before aphasia. Consequently, aphasia is reported to present in later stages of AD with anomia as an usual observed problem (Kirshner, 2012). However, PPA is mostly related to FTD rather than AD, there has been a hypothesis that PPA syndromes may also be associated with Alzheimer's disease. Recent evidence suggests that AD pathology plays an important role in LPA because of a high proportion of co-occurred cases (Kirshner, 2012). Despite this, some AD patients are reported that have PNFA and SD presenting characteristics not clearly belonging to a single category, so-called 'mixed' aphasia (Rohrer et al., 2012).

2.2.4 Disorders of Speech

Producing speech requires a highly coordinated regulation, from the control panel in the brain through to the motor organization of articulators. Damages in any one level of this track or multiple levels will eventually induce motor speech disorders. The underlying causes can be muscle weakness, paralysis, spasticity, and poor coordination. There are two categories of this movement disorder regarding phenotypic presentation, including dysarthria and apraxia of speech.

2.2.4.1 Dysarthria

The inability to organize and coordinate the speech-responsible muscles are a key feature of dysarthria. Unlike neurodegenerative diseases, the level of impairments appears at the cranial or spinal nerves, which innervate the muscles assisting speech. The impairments of those nerves cause a loss of control of speech-related muscles; accordingly, several aspects of speech are disrupted. Proper phonation and articulation require the use of more than 100 muscles and are crucial for speech that is loud, in time, accurate, steady, and of a certain tonal quality. In other words, losing control of a speech-responsible muscle can be described as a ‘drunken quality’ of speech. The common symptoms are tremors of the laryngeal and respiratory muscles as well as incomplete pronunciation. Patients with dysarthria are characterized by slurred, babbled, slowed, monotonous, poor pitch control, and irregular separation of the syllables in words. In addition, weakness of the face and tongue is observable. In clinical practice, objective criteria of speech function should be assessed, for instance, auditory abilities, perceptual characteristics, repetition rate, oral mechanisms, and intelligibility testing (Skodda et al., 2014; Zeki & Hillis, 2016)

2.2.4.2 Apraxia of Speech

Apraxia is a motor disorder where the syndrome mainly influences articulating ability; thus, it is called apraxia of speech. The anatomical lesion of this syndrome has been controversial; after all, it is usually affected by left hemisphere stroke. It is proposed that the two damaged areas for disease localization are the left superior middle cerebral cortex and the left inferior frontal cortex. The lesion can be due to a clot, tumor, abscess, or focal atrophy. This speech impairment can occur in the absence of aphasia or dysarthria. However, apraxia of speech can occur in the context of Broca’s aphasia. The two syndromes may be characterized in the following way: Patients with aphasia find difficulty in selecting the proper phonemes to utter, while apraxia affects the motor execution when pronouncing the selected phonemes (Jung et al., 2013; Rohrer et al., 2012).

Patients with apraxia of speech are aware of their problems but cannot normally control the rhythm of speaking. Dysarthria is differentiated from apraxia of speech by consistency and predictability of errors; the latter does not have patterns of abnormal utterances. They tend to articulate with trial and error. Characteristics of deficits may include difficulty initiating utterances, self-correction of errors, effortful repeating of the same utterance with inconsistent errors, and abnormal prosody, stress, and intonation. The features of apraxia of speech and agrammatic speech are accounted as possible characteristics of non-fluent agrammatic variants in primary progressive aphasia (Jung et al., 2013; Zeki & Hillis, 2016).

2.2.5 Speech deficits in dementias

Both the language level ('what is said') and the paralinguistic level ('how it is said') are impacted by dementia (König et al., 2018). Therefore, speech performances can reflect the deficits due to cognitive impairments in AD through cognitive tasks such as verbal fluency, word list recall, and naming objects (Beltrami et al., 2018; Gomez & White, 2006). Table 1 displayed change in language function in MCI and AD (Boschi et al., 2017; Ferris & Farlow, 2013; Szatloczki et al., 2015).

Difficulty in word finding is especially present in AD. Word retrieval problems are prominent and pervasive even in the early stages of dementia. Verbal fluency tests have been used to distinguish between AD patients and cognitively healthy people. AD had significantly lower scores in letter fluency in contrast with control groups. Therefore, fluency is acknowledged as a trustworthy indicator of later stages of dementia in senior people. Though they performed worse than healthy controls, prior research revealed that AD patients could produce more words in category and letter tasks than PPA patients (Marczinski & Kertesz, 2006). Word finding in the form of naming difficulties can be found in various forms of dementia, including AD, FTD, and VD (Paek et al., 2019).

Although the prominent impairment of AD is episodic memory, symptoms of 'typical' AD are also characterized by language deficits in the early stages of the disease. Otherwise, the language domain may be impaired with other cognitive domains due to the progression of the disease. Primary dysfunctions of language are based on abnormality in lexical semantic abilities, which demonstrates difficulty in word finding, unintended utterance or semantic paraphasia, a deficit in word comprehension, and verbal fluency impairment. At the same time, phonological (sounding) and syntactic (grammar) processing is relatively spared in the early stages but progressively impaired in the late stages. However, some studies found that AD patients in the early onset simplified their syntax and phonological structure. Eventually, language impairment becomes pervasive and severe due to the progression of the disease. Severe language deficits restrict speech abilities through echolalia and verbal stereotypes. Amnestic MCI patients also present language deficits, especially in the groups of a prodromal stage of AD. Those patients are reported as presenting characteristics closely similar to those of language impairment in the early stages of AD (Paek et al., 2019).

Table 1 Language characteristic changes in mild cognitive impairment and Alzheimer's disease

Language characteristic changes	**MCI**	**AD**	
		Early stage	Moderate-severe
Phonetics-phonology			
- Temporal changes in spontaneous speech (increasing hesitation number and time)	Impaired	Less fluent	Nonfluent, echolalic
Lexico-semantic			
- Word-finding and word-retrieval difficulties	Impaired	Mild impaired	Impaired
- Semantic knowledge	Impaired	Impaired for less frequently used words and objects	Impaired
Syntax-morphology			
- Syntactic comprehension	Intact	Intact	Impaired
- Reduced syntactic complexity	Intact	Impaired	Impaired
- Agrammatisms	Intact	Intact	Impaired in severe stage
Discourse-pragmatics			
- Reduction in productive and receptive discourse-level processing	Intact	Impaired	Impaired

Speech deficits further occur in other neurodegenerative diseases with different prominent characteristics. Parkinson’s Disease and LBD patients are described as having common errors in motor speech control, syntactic processing, verb inflection and generation, and sentence comprehension. According to motor disorders, both diseases are observed with slow speech and abnormality in speech sound. The motor speech deficits disrupt respiration, phonation, articulation, resonance, and prosody. Syntactic and pragmatic difficulties have been remarked on. The patients often lessen the difficulties in organizing narrative speech as well as production and comprehension syntactic complexity. Executive deficits have been linked to naming and verbal fluency deficiencies (Boschi et al., 2017).

Speech consists of linguistic components, cognitive processing, and motor function. Speaking can reflect cognitive impairment because language proficiency depends on multiple cognitive domains, particularly perception, memory, and attentional allocation. This study focused on memory-related processes. Memory is involved in speech production as a storage of long-term information and a control center. The phonetic and rhythmic representations are briefly stored in working memory as a temporal phonological store. To be noted that primary dysfunctions of language in AD are based on abnormality in lexical semantic abilities, which demonstrates difficulty in word finding, unintended utterance or semantic paraphasia, a deficit in word comprehension, and verbal fluency impairment. At the same time, phonological (sounding) and syntactic (grammar) processing is relatively spared in the early stages but progressively impaired in the late stages of AD. Therefore, a number of previous studies explored how to detect AD and MCI with speech analysis. They investigated several linguistic features, such as semantic fluency or discourse forming. One interesting parameter in detecting dementia is acoustic features. The next section elaborates more on speech tasks and linguistic parameters.

2.3 Speech analysis

Verbal responses are accepted to be a predictor of cognitive impairments and differential variables in dementia, especially spontaneous speech which closely imitates daily communication. The natural action of oral narration is called in several terms, e.g., spontaneous spoken, connected language, or natural language. Those are mainly referred to as spontaneous speech in this study. Spontaneous speech in the context of the research is characterized by self-generated discourse, close to language production in daily life, and less intervention by the examination (Mueller, Hermann, et al., 2018; Pulido et al., 2020). Stimulus or administered instruction is the key to eliciting the natural spoken language. Language-based measurements are categorized by the degree of free verbal responses.

Language assessments offer observable and unobservable information. Spoken responses provide the variable related to time and effort to complete the tasks, and also linguistic features, for instance, lexicon or phonological quality. Linguistic parameters can reflect preserved cognitive abilities of dementia patients from single to multiple components, such as semantic storage, retrieval ability, and execution of speech. The parameters can be measured with neuropsychological assessments, which can distinguish examiners with cognitive deficits based on performance or error scores. The verbal performance derived from neuropsychological assessments can further provide linguistic features as differential indicators. Meilán et al. (2014) further stated that the acoustic characteristics of the speech of AD patients could be generalized for discriminant analysis to languages other than English. Recently, verbal tasks have been directly used to bring articulation information aiming for differential diagnosis among variants of cognitive impairment between MCI, AD, Primary Progressive Aphasia, or other comorbid conditions such as depression (Mueller et al., 2016).

The following section explains the core components for developing spontaneous speech tasks, which extract cognitive deficits among MCI, AD, and cognitively intact persons.

2.3.1 Measurements

Speech provides inclusive linguistic aspects such as efficiency, fluency, quality and error of utterance, speech rhythms and intonation (prosody), grammar (syntax), and phonemic selection. Speech is recognized as a promising candidate as a source of information for new approaches to diagnosing dementia. However, a concerning issue in assessing patients by speech tasks is the extent of the spontaneity of responses. If patients may respond to neuropsychological tests differently from the way they speak at home or with friends in everyday activities, then the examiner would not be able to base his/her analyses on solid and generalizable data. The tasks and stimuli presented to the patients, therefore, should be created or chosen based on the level of familiarity and relevant personal factors such as first language or local dialect, if applicable (Weekes, 2020). The speech tasks used in cognitive assessment can actually be classified based on the ability to approximate language production and articulation in daily activities (Lezak et al., 2012).

2.3.1.1 Structured protocol

Word retrieval problems are prominent and pervasive even in the early stages of dementia. Deficiency of phonemic and semantic memory has also been documented. As a result, assessments that test phonemic and semantic fluency may also be useful for diagnosis of memory impairment (Gomez & White, 2006; Henry et al., 2004)

i. Verbal fluency

Verbal fluency tasks are recognized as language tests in the multiple cognitive domain assessments and in the specific cognitive domain test in neurodegenerative diseases (Charernboon, 2018). In neuropsychological tests, verbal fluency tasks appear in at least three standardized tasks in the Thai version, including the Seven Minute Screen, Rowland Universal Dementia Assessment Scale (RUDAS), and MoCA, which are suggested for MCI examination (Neurological Institute of Thailand, 2014). Examinees are asked to generate as many words as possible within a specified time limit on the basis of a cue provided by the examiner (e.g., 60 seconds). The two most used cues for verbal fluency are semantic (or category) (e.g., animals) and phonemic (or letter) (e.g., beginning with the letter ‘S’) (Na Chiangmai & Wongupparaj, 2020). Despite degenerative diseases such as dementia influencing the abilities of both category and letter fluency, the impairments rely on both some common and some distinct cognitive processes. Category fluency strongly depends on lexical representations. Letter fluency heavily relies on the central executive component of working memory. Both of these distinct deficits appear to depend on accessing stored knowledge rather than on verbal ability.

Responses in verbal fluency tasks offer info about several types of behavior and cognitive functions. Measures of the performance obtained from verbal fluency includes the raw number of words generated within a time limit; the possible semantic cluster(s) into which the produced words are organized; the word frequencies, based on the norms of the specific language, of the produced words; and the number of times a word is repeated within a trial. The word frequency allows to examine accessibility to lexical representations, with high-frequency words usually being produced more often than low-frequency words. Kim et al. (2019) was interested in alternative method in examining verbal fluency task, i.e., clustering (tendency to generate word chains that are grouped into semantic subcategories) and switching (changes from one category to another). Clustering and switching variables are significantly correlated with the number of words generated, with higher numbers of reported words allowing more possible clusters to be formed. Currently, technology provides an optional analysis for verbal fluency with clusters of semantically related items that are based on the corpus of words used on the internet (Kim et al., 2019). The number of repeating errors, with at least one intervening item between two repetitions, could indicate impaired working memory. The deficit of working memory decreases the ability to hold in memory the already generated words, so the examinee forgets a particular exemplar has already been produced and produces it again. When tested on verbal fluency, patients with AD typically underperform compared to healthy controls in terms of the number of words, the number of clusters, the number of switches, and the size of the clusters produced (Marczinski & Kertesz, 2006).

ii. Naming test

A naming test is included in the standardized neuropsychological tests for assessing language function, especially receptive ability. One of the most widely used is the Boston Naming Test (BNT). This test is often adopted as a single assessment of aphasia and also combined with other cognitive task tests, such as the Consortium to Establish a Registry for Alzheimer's Disease (CERAD). The original BNT contained 60 ink drawings across a range of high, medium, and low familiarity. To administer, examiners present the drawings depicted on stimulus cards one by one at a descending level of familiarity. Examinees are asked to provide the common name of the pictures within 20 seconds. A semantic cue may be offered when the examinees are unable to answer particular cards (Lezak et al., 2012; Mack et al., 1992). The administration time is approximately 15 to 20 minutes. In later times, the BNT has been shortened, with the stimulus pictures being condensed into shorter 15-

item and 30-item versions and translated into several languages, including Thai in the 10Thai-BNT (Aniwattanapong et al., 2018).

Both verbal fluency tests and naming tests require memory retrieval, but they differ for both procedures and stimuli. While the naming test uses pictures to assess semantic memory, the verbal fluency test requires only verbal instruction, which can elicit both semantic and phonemic representations. Executive function is also involved in controlling the demands by putting on effortful retrieval of both types of information. Patients with AD found significant difficulties in both semantic fluency and naming tests, which depend on semantic knowledge. The semantic memory was considered as obtaining the most differential sensitivity among naming test and letter fluency respectively. In MCI and dementia, cognitive abilities such episodic and semantic memory are more impaired than executive function. Notably, the memory impairments are mediated by the medial and lateral temporal lobes which are the main pathological area of dementia, rather than the frontal lobes (Gomez & White, 2006; Henry et al., 2004; Smith & Bondi, 2013, pp. 167-168)

iii. Reading

Reading tasks offer various beneficial variables such as comprehension ability, speech characteristics for reading aloud, and eye movement with eye-tracking technology. For diagnostic purposes, reading is used to evaluate the receptive and comprehensive abilities of visual material, especially in aphasic patients or individuals who are presumed to have left hemisphere atrophy (Lezak et al., 2012). To assess language ability, reading examinations can be designed to include reading-aloud tests and reading comprehension tests (Fraser et al., 2019). Reading aloud gives an acoustic perspective. It has been used in studies designed to detect both AD and MCI. The studies found that reading patterns for patients with AD were characterized by slower reading, shorter speech chunks, increased pauses, and speech dysfluency compared with patients with MCI and the control populations (Celine De et al., 2018). The other reading task is a comprehension test in which a set of questions are administered after a passage has been read. The questions can be true-false questions or multiple choice for evaluating comprehension accuracy. The study with reading task found that the statistical analysis revealed significantly higher correct answers and duration of the control than the MCI group. However, the score of the reading comprehension scores and duration negatively correlated with the scores and time needed to complete the visuo-perceptual test (Trail Making Test Part B) (Sofia et al., 2016). The receptive ability of reading in this study reflects opposite correlation with visuo-perceptual ability. Reading tasks have been argued to be promising indicators of AD and MCI since reading impairments correlate with late-stage and severe progression of the disease (Jokel et al., 2019).

2.3.1.2 Unstructured protocol (elicit spontaneous speech)

Test activities that highly resemble daily life communication need less restricting of the administering instruction and specific information conveyed by the tests. Consequently, individuals can generate their own spoken discourse more freely (Boschi et al., 2017; Lezak et al., 2012). Regarding the studies on spontaneous speech in MCI and AD, the tests base on single word processing such as verbal fluency or naming test are concerned having difficulties to elicit the extended language production (Boschi et al., 2017). Evidently, descriptive tasks such picture description were the most frequently found across the studies of connected speech in dementia, followed with narrative tasks, i.e., interview and conversation (Filiou et al., 2019).

i. Picture description/ Story narration

Picture description tools can be categorized into two set, i.e., single-story cards and booklets of sequential pictures (Mueller, Hermann, et al., 2018). The single-story cards usually depict a rather complex situation with some objects and people present, such as the “cookie-theft” (Goodglass and Kaplan, 1983); the booklet of sequential pictures usually depict scenes from a well-known story, such as Cinderella, but do not include any written words. To administer the test, the examiners provide the cards in sequential order to the participants, who are either asked to describe the story they see or to tell the story in their own words after the booklet has been removed (Treviso et al., 2018). Both tests are administered with minimal instructions or interruptions by the examiners. In general, the duration of the task is 2 -5 minutes for people without cognitive impairment. Participants attempt to spontaneously tell the story as if communicating in an everyday context. The test elicits a moderately unstructured speech output when compared to story recall test and interview. Since the speech responses of this kind of test can be quantified and scorable words can be predefined (Boschi et al., 2017). Both types of picture descriptions impose a predictable speech output; the description and tale provided by the participants should include important details, information, or semantic units (subjects, objects, activities, and locations) that are depicted in the photographs. As a result, the amount of accurate information units the patient correctly identified can be used to score the narration (Boschi et al., 2017; Mueller, Hermann, et al., 2018).

These tests are used to identify semantic deficits, word retrieval difficulties, and speech performance in neurodegenerative diseases. The participants need to select and retrieve appropriate vocabulary to match with those semantic units (Bradley et al., 2010). The ability to develop a narrative that is defined by a sequence of events or actions is also a requirement for picture narration. The narrative based on the assigned illustration requires understanding of the characters and events, including temporal and spatial alterations, as well as the goals and internal responses of the characters, eventually the tale is included with a structured and coherent framework. Visual perception is apparently examined, as well as several cognitive abilities involved in completing both tasks. Apparently, the coherence of a topic in discourse relies on attention control to prioritize major components and internalize frame representation.

The most used picture description tasks employed in the assessment of extended speech in neurodegenerative diseases are the Cookie Theft of the Boston Diagnostic Aphasia Examination (BDAE), the Picnic scene of Western Aphasia Battery (WAB) (Mueller, Hermann, et al., 2018), and the living room activities picture in Kentucky Aphasia Test (KAT) (Marshall & Wright, 2007; Nagarachinda et al., 2020). These pictures are in black and white. Apart from those standardized tests, the studies of connected speech analysis in MCI and AD further use Norman Rockwell prints such as ‘Easter Morning’ and ‘Cinderella’ as stimuli for storytelling tasks (Treviso et al., 2018). The illustrations of Norman Rockwell were frequently printed in color (Tomoeda et al., 1996). König et al. (2018) simply used a photo of an animal in its natural environment in their denomination picture description. Lately, Sangchocanonta et al. (2021) developed Thai picture description tasks with two pictures reflecting Thai cultural experiences, namely Thais-at-Home and Thai Temple Fair (See Figure 5). Those pictures were used to detect AD and MCI in Thai cohorts.

(a)

(b)

Figure 5 Picture set of Thai picture description tasks © [2021] IEEE. (Reprinted, with permission, from Sangchocanonta, et al., Development of Thai Picture Description Task for Alzheimer's Screening using Part-of-Speech Tagging, Annu Int Conf IEEE Eng Med Biol Soc, December 2021)

The pictures used in picture description tasks aim at activating episodic and semantic memory for familiar situations. They are characterized by rich context that contain several parts with a salient part in the foreground and less important elements in the background (Chapman et al., 1998; Giles et al., 1996). The assorted and complex aspects of the picture required fluency, judging and perception (Cummings, 2019). They should include persons and object, make explicit a place, and suggest a time-frame (Giles et al., 1996; Mueller, Hermann, et al., 2018). The objects and/or activities described should be among those early acquired in life and for which a vocabulary should be well-known (Chapman et al., 1998). The domestic scenario and common experience are likely familiar to all subjects (Tomoeda et al., 1996). These characteristics would clearly allow to detect whether frequent and familiar words are found to be a struggle for the patients (Chapman et al., 1998; Giles et al., 1996; Marshall & Wright, 2007).

The instruction for picture description administration is simple. After the picture is presented to the examinee, the examiner asks him/her to create the story from the given illustration using an instruction such as ‘Tell me everything you see going on in this picture.’ This sentence appears in at least two standardized tests, namely the Cookie Theft and the divided attention picture in the KAT (Giles et al., 1996; König et al., 2018; Marshall & Wright, 2007). Other instructions may be used, e.g. asking participants to generate a possible story (rather than describe the picture) (Chapman et al., 1998). While examinees are telling the story, verbal prompts on the part of the examiner should be avoided, but non-verbal encouragement is allowed (Giles et al., 1996). In several tests the examinees are allowed unlimited time to accomplish the task (Giles et al., 1996; Marshall & Wright, 2007) whereas recommended duration is three minutes in the Ester Morning picture (Tomoeda et al., 1996). A story is considered finished when the examinees explicitly indicate the end or pause for more than 15 seconds (Giles et al., 1996). Tomoeda et al. (1996) suggested encouraging the examinees by saying, ‘Is there anything else you can say about the picture?’ or ‘Tell me what else is happening.’, but do not ask for specific items or activities.

Discourse analysis widely differs in the scoring of storytelling. To illustrate, Tomoeda et al. (1996) selected eight measures to quantify the elicited speech from Easter Morning and the Runaway of Norman Rockwell (i.e. (1) total words, (2) information units, (3) conciseness, (4) circumlocutions, (5) frustrations, 6) aborted phrases, (7) revisions, (8) ideational repetitions. Interscorer agreement was also taken into consideration. While Chapman et al. (1998) used two indexes: (1) aspects of content are employed from frames of interpretation, proposition supporting frame, and propositions disrupting frame (2) aspects of form are included in the structure of information and reference. The discourse analysis used with the same tasks can be addressed with diverse variables.

The scoring criteria of picture descriptions can also be based on information units and duration of speech, as well as the combination of both parameters. The basis of scoring is informative, efficient, and concise storytelling. Giles et al. (1996) measured the storytelling from the Cookies Theft picture by three aspects, i.e., (1) duration of inclusive speech acts in seconds, (2) utterance within the speech duration, counting the number of syllables, and (3) information unit (IU), counting the pieces of information which are nonredundant meaningful fact or reasoning. The information units generated by healthy participants in KAT are used as references for unit counting (Marshall & Wright, 2007). According to the scoring rules of Giles and colleagues (1996), the same unit can be counted more than once if it is talked about in a new context on a second occasion. In contrast, repeated words are discounted if used for the same context. Only correct information is counted. A revision with repaired information is scored only by self-correction, not by an examiner's prompt. A further parameter deriving from the three aspects is conciseness which is the sum of information units divided by the total number of syllables. This variable is one of the discourse devices, and it was found to differentiate people with and without dementia (Bayles et al., 1999).

ii. Story recall (Immediate & Delayed)

The story recall can be considered similar to picture description in the aspect of natural speech samples and recently learned stimuli, but it differs for its reliance on features of audio stimuli. The use of additional questions may also be a differentiating aspect. Since story recalling is a natural form of communication, the elements forming a story are similar to a snapshot of activities in daily life (Mueller et al., 2020). Participants' performance is scored by story units, which are the important or outstanding elements of the presented story. Besides, working memory can be assessed by asking questions in relation to the presented story; these measures can give accuracy scores. The scores derived from this task can indicate cognitive impairments (Kent, 2013; Mueller et al., 2020). The higher the number of recalled elements and correct answers inferred, the higher the memory score (Kent, 2013; Roark et al., 2011). The recalling protocol, which requires the participant to learn and remember the stimuli and recall it immediately or after a time-interval, is considered to relying on episodic memory especially recent memory (Bradley et al., 2010).

Narrative retelling has been shown to discriminate between cognitively intact individuals and cognitively impaired people (Gomez & White, 2006; Prud'hommeaux et al., 2011; Roark et al., 2011). In neurodegenerative disease studies, recalling tasks are functional for detecting patients with MCI and AD because the task relies on episodic and working memory, which are impaired in both groups. Both MCI and AD groups exhibit working memory deficits on the sentence repetition task, as well as relatively lengthy pauses when they attempt to recall the story. Story recall requires both language and memory skills, and it is a task that can accurately detect memory deficits (Roark et al., 2007), including episodic memory and auditory memory (Bayles et al., 2020, p. 147; Holdnack & Drozdick, 2010, p. 242). The task can additionally examine executive functions such as planning, organizing, and monitoring data since language and memory interfaces need those organization skills (Roark et al., 2011; Treviso et al., 2018). It is considered a linguistically demanding test that involves multiple cognitive domains (Jokel et al., 2019).

Even though people cannot wholly remember a story word by word, they recall the concept or scheme of it and retell it in their own words with some parts corresponding to the actual presentation (Khan, 1986, p. 43). The degree of memorization in verbal memory tests can be assessed simply by counting the recalled elements (Treviso et al., 2018). The recalling or retelling tasks are typically executed twice, providing data for an immediate recall and a delayed recall of the story (Bayles et al., 2020, p. 147; Hodges, 2007). Story retelling tasks appear in at least two standardized memory test batteries, namely Wechsler Memory Scale-Revised (or other versions), specifically Wechsler Logical Memory I and II (WLM I & II)

(Holdnack & Drozdick, 2010, p. 240), and Arizona Battery for Communication Disorders of Dementia (ABCD-1 and 2) (Armstrong et al., 1996).

An important element of story retelling is the length of story, which should be short in order to avoid problems with natural forgetting. The standardized task shares a common structure with respect to stories' lenght. The number of sentences in each short story of LMS and ABCD is three (Hodges, 2007; Prud'hommeaux et al., 2011). The story in WLM is divided into 25 measuring units (Prud'hommeaux et al., 2011) that amounts to about four sentences an in the original Babcock story (Hodges, 2007) from which the WLM stories were adapted from (see Khan 1986, p. 44). Leal et al. 2021 tried to examine the emotional impact of short stories on memory by creating three emotionally stimulating stories modified from the WLM of the Wechsler Memory Scale III (WMS-III). Their stories contain 60 – 70 words per story, three sentences long and 25 memory units. In their study, the stories were separated into 7 thematic points or gist information.

American version of the short story in WLM is:

"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 fifty-six dollars. She had four small children, the rent was due, and they had not eaten for two days. The police, touched by the woman's story, took up a collection for her" (Roark et al., 2011).

The basis of story retelling is the immediate recall after listening to the story and reproducing the story again after a delay. Administration of this task needs the full attention of the examinees before starting and while the examiners verbally present the short story (National Institute on Aging, 2006). The examiner will ask the examinees to carefully listen to the story and will also prompt them with the retelling instruction (Hodges, 2007). The examiner reads the brief story with a slow pace, clear articulation, and normal inflections. After the story has been told, in its entirety, the participant is asked to retell the story immediately from the beginning, trying to remember as much as they can (Hodges, 2007; Khan, 1986; National Institute on Aging, 2006). This test requires the ability to encode and recall the story on the basis of auditory information (Lezak et al., 2012). The recall should be performed without interruption, conversation, or cues on the part of the examiner. After the immediate recall, and when the delayed onset is reached, participants are informed of the delayed recall (National Institute on Aging, 2006). The delayed recall may occur after a wide range of time intervals, such as 30-45 min. (Hodges, 2007), 30 min. (Lezak et al., 2012; Prud'hommeaux et al., 2011), 20 min. (Leal et al., 2021) and 10 min in the Babcock format (Lezak et al., 2012). The delayed phase is usually filled up with other neuropsychological tests. When the delayed period is finished, the examinees are asked to tell the story again. The second round of retelling assesses the delayed encoding and recall of the verbally presented story (Lezak et al., 2012).

The verbal presentation of the story was adjusted to aid participants from an overwhelming feeling and to obtain learning slope by presenting two stories and repeated reading. The Anna Thompson story (appeared firstly in the WMS-Revised in 1987) was considered to stimulate emotional bias, then some participants were distracted. In the WMS-III (lunched in 1997) and Babcock story recall test, the short story was secondly presented after an immediate (before delayed period) aiming to provide an opportunity to improve their learning. Given the read twice protocol, this administration had advantage for participants with a limited auditory span and brief attention (Lezak et al., 2012). However, a single reading is offered in the WMS-IV and several speech analysis studies in the dementia group. Prud'hommeaux et al. (2011) and Roark et al. (2011) used the story of Anna Thompson from WMS and presented the story one time in their studies of detecting MCI.

The score of the story retelling task is given by computing the number of the participant’s correctly recalled units. These can which can be thematic units or details depending on the criteria of the given test (Lezak et al., 2012). The story is divided into units (usually by graphically explicit slashes). Each unit may contain a single word or a few words (e.g., a phase). The scoring criteria should clearly define an important text of each unit which can be scorable. For example, in WLM, for the unit ‘in a school’ at least the mention of ‘school’ (either the response of a high school or a school also counts) is required for getting a score of 1 (Khan, 1986; National Institute on Aging, 2006, pp. 43-45). A half point is given as long as the basic idea-unit is preserved, although the responses are synonyms or substitutes of the original words, or there is the omissions of an adjective or a verb (Hodges, 2007; Khan, 1986, p. 43). Optional scoring criteria were created, such as thematic scoring or gist information which a point was given to correctly recalled number of predefined main ideas (Khan, 1986, pp. 43-45; Leal et al., 2021; Lezak et al., 2012). Accordingly, Leal et al. (2021) found that most older adults recalled a gist of information better than the details. Detailed (semantic) information seems to be frequently used. Older adults with high intelligence were expected to score more than ten recalled elements (out of X elements) in immediate recall with at least 60% retention in delayed recall (Hodges, 2007).

iii. Interview & Conversation

The most common form of daily communication is back-and-forth verbal interaction or dialog, and this format has also been used to investigate dementias (Mueller, Hermann, et al., 2018). The interview can be classified into three levels of structure: structured, semi-structured, and unstructured interview. The level of structure is related to the constraints provided for speech output:

a) In structured interviews, the form and order of questions are determined beforehand, and they are presented to participants in that specific format. This may result in quite low amounts of spontaneous speech (Boschi et al., 2017);

b) Semi-structured interviews can be administered by open-ended or mixed questioning with closed-ended questions and combines a set of pre-determined questions with the possibility for the interviewer to ask new questions, as a function of the responses given to previous questions, to further probe specific aspects. This middle level encourages interactive communication. Most standardized language or verbal memory assessments are therefore conducted using the systematic protocol of semi-structured interviews, as it is the case for the Western Aphasia Battery (WAB) and for the Autobiographical Memory Interview (AMI) (Boschi et al., 2017; Lezak et al., 2012);

c) Unstructured interviews, or informal conversations, do not have a systematic format to follow and have no prepared questions. They only focus on a selected theme in the conversation and develop according to the communicative context created by the participants. The turn-taking and narrative nature of the interviews require an elaboration of cognitive functions, with at least language comprehension, memory activation, attention control, and executive function. For example, telling the story of oneself allows the participants to be relaxed in the test environment and provide indexes for the evaluation of recent and/or remote episodic memory and personal semantic memory (Bradley et al., 2010; Mueller, Hermann, et al., 2018). Moreover, conversations provides an expressive variation of fluency which is superior to that of the tasks with a more structured protocol (Themistocleous et al., 2020). Generally, interviews about autobiographical memory aim at eliciting personal remote memory and at evaluating retrograde episodic memory (Hodges, 2007), while the story recall test or list learning tasks aim at assessing recent episodic memory.

In the context of dementias and MCI research, semi-structured interviews and conversations potentially offer useful results in differentiation of healthy control, MCI and AD (Asgari et al., 2017; Beltrami et al., 2016; Beltrami et al., 2018; Bung, 2016; Filiou et al., 2019). Although the semi-structured interview may be conducted and replicated easily, the unstructured interview has occasionally been used to evaluate language production in dementia (Singh et al., 2001). Speech responses of individuals reveal useful variables for distinguishing MCI, dementias, and cognitively intact processing (Filiou et al., 2019).

Conversations or interviews are intended to elicit samples of participants' natural speech. In speech analysis studies, the conversation usually starts with a very broad open-ended question, e.g., 'Tell me about your family/hobbies/career?', about negative or positive event in life, events happened yesterday, or the last dream the person remember (Beltrami et al., 2018; Boschi et al., 2017; König et al., 2018). The role of examiners is to start the conversation by introducing the topic and encouraging the examinees to speak (Boschi et al., 2017). By asking about familiar and generic topics, the participants are quite free to respond as they want.

Scoring criteria for an interview are found in most standardized tests, such as AMI. The first section of AMI tests the personal semantic facts in three epochs of life. The second section deals with the autobiographical incident schedule. Participants are asked to recall three specific incidents in three periods of their lives. Scoring for the autobiographic part is based on the descriptive richness and specificity of information. There are normative data and cut-offs for both sections of AMI (Hodges, 2007).

However, the impossibility of verifying the patients' personal history makes this procedure ineligible for the exact evaluation of the accuracy of the reported incidents. Even a family member or caregiver may not be knowledgeable about the facts or details that have happened in the past and the personal life of the participant. The normative data show that 90% of the responses was true, but the examiner should be cautious of confabulation (illogical information or a scenario that never actually happened) (Bradley et al., 2010). Scoring for accuracy or richness of the response seems not appropriate for the verbal production of conversation and interview.

Interviews and conversation can unfold in time as long as the experimenter allows it, and the typical duration of such activities can range from 5 – 20 minutes, depending on the type of interview. Without structured conversation, the duration varies according to the length of the responses of the examinees, and this may make it difficult to analyze and compare performance across groups. Furthermore, the analysis of speech output is very time-consuming, and it may be difficult to assign scores in the absence of predefined task constraints. Nevertheless, a computational analysis of physical features in speech is able to overcome scoring limitations (Boschi et al., 2017; Themistocleous et al., 2020).

2.3.2 Linguistic parameters

For spontaneous speech, the linguistic variables extracted from the participants' verbal production are categorized into four categories. These four components of linguistic features are mostly studied in the fields of neurodegenerative disorder and of older adults for detecting cognitive decline (Beltrami et al., 2018; Boschi et al., 2017; König et al., 2019). The four categories of linguistic variables are described in the following sections.

2.3.2.1 Lexical features (Semantics)

Lexico-semantic properties may allow to demonstrate impairments at word and content levels. At the lexical-grammatical level, part-of-speech is relevant for word-class categorization, e.g., noun, verb, preposition, etc. The average occurrence rate of each word type offers information on the lexical distribution of words delivered and the difficulties in accessing a specific word class. Part-of-speech can be used to identify place, time and person, which demonstrate the link between communicative context and speakers, including demonstratives (spatial deficits) and personal pronouns (person deficits); deictic word with no clear referents makes the discourse vague and ambiguous (Boschi et al., 2017). Lexical characteristics describe how informative a conversation is and/or give lexical richness. Word frequency may also be a sign of a discourse's level of information. For instance, using a lot of content words has been linked to speech that is less accurate (Fraser et al., 2014). (Fraser et al., 2014). Lexical and semantic errors can be distinguished at this level. Word repetitions, indeterminate phrases, modifications, and newly invented words are all examples of lexical mistakes. These mistakes affect how clearly speech is produced, as well as the grammatical structure and discourse organization. Semantic errors generally involve substituting a word with a semantically related word, using a superordinate label instead of a specific label (e.g. *animal* instead of *cat*) or producing coordinate terms (e.g. *dog* instead of *cat*). Lexical variables used in the dementias and MCI identification were, for example, part-of-speech rate, personal, spatial, and temporal deixis rate, type-token ratio, and propositional idea density. In particular, these variables tell the complexity of speech production, which marks the difference between participant groups (Beltrami et al., 2016).

2.3.2.2 Syntactic features (Grammar & Morphology)

Word inflection and agreement information, including tense, mood, aspect, person, number, and gender, are typically reported through morpho-syntactic characteristics. Because the allomorph selection is dependent on the phonological context, phonological processes are inextricably linked to inflectional processes (e.g., /*a cat*/ vs /*an uncle*/). The syntactic structure is also influenced by agreement and inflection. The usage of a nonexistent word form or the improper selection of an existing, incorrect morphological form of a word are both examples of morphological errors. The absence of function words, their improper use, or the incorrect use of verb tenses are all examples of morphological grammar errors. These mistakes also represent poor temporal coherence in conversation (Beltrami et al., 2018; Boschi et al., 2017).

General structural flaws and unfinished sentences are examples of syntactic errors. These variables indicate the type of created syntax, but they may also be used to identify impairments at various language levels. Sentences that are incomplete lack some essential component of the structure and are therefore not properly developed. Despite the fact that this feature represents a syntactic phenomenon, it may be caused by a variety of language impairments, including abnormalities at the lexico-semantic, syntactic, or discourse levels (Beltrami et al., 2018; Boschi et al., 2017; Ladefoged, 2006).

2.3.2.3 Pragmatic features (Discourse)

Pragmatic features refer to the ability to use language and speech appropriately in a given socio-cultural context taking into account the addresses, e.g., using informal language with friends and formal language in formal settings. In an individual's speech, the expressed content reveals pragmatic skills of speakers that can be continuity, cohesion, coherence, and correct usage of pronouns and conjunctions. Discourse product conveys information on how well the meaning is expressed and if the appropriate amount of information is provided. The connected speech needs cohesion to make connections within and between sentences. Cohesion can be classified into referential cohesion, temporal cohesion, and causal cohesion. Coherence also refers to relations of delivered contents but on a wider scale. This indicator of pragmatic features may be considered by local and global coherence. The former means the extent of congruence between the subsequent utterance and the preceding one. At the same time, global coherence tells the association of speech with the general topic and also links to the unassociated topics that counterparts to the preceding topic (Beltrami et al., 2018; Boschi et al., 2017; Nagarachinda et al., 2020).

2.3.2.4 Acoustic features (Phonetics)

Acoustic variables are related to intensity, duration, and frequency of the speech signal and, therefore, a physical characteristic of speech. Utterances can be quantified as a function of the execution time to produce sentences and/or pauses in the speech. The amount of time in dialogs is a function of both speech and pauses, hesitating pauses, etc. Acoustic features can be extracted and categorized into four groups, i.e., prosodics, formant, source, and temporal. The prosodic area relates to long-time variation in perceived stress and rhythm in spoken language, which may be unimportant in language with no stress. Fundamental frequency, or F0, is the variable that represents a prosodic feature, which relates to the vibration of the vocal fold. The formants represent spectrum components that carry information about the acoustic resonance of the vocal tract; they can also indicate articulation problems due to motor control deficits. Source variables provide information about the voice production when airflow passes through the glottal speech production, the glottal resistance, and voice quality. Lastly, temporal features measure speech characteristics based on a time scale, e.g., duration of sounding and pauses, total duration of speech, a ratio of sounding and silence, etc. (Boschi et al., 2017; Ladefoged, 2006). The longer duration of responses is comprised of pauses and hesitation as well as lower speech rate; those verbal products can be found in AD and MCI (Ambrosini et al., 2019; König et al., 2015; Slegers et al., 2018)

The current study explores the profile of acoustic features in Thai older adults in two aspects, i.e., the frequency-related domain and the temporal domain.

i. Frequency domain

1) Pitch is an audio property that refers to the relative highness or lowness of a tone and intonation as a result of the vibrations of the vocal cords (Gagliardi & Tamburini, 2021; Ladefoged & Johnson, 2015, p. 25). Generally, the pitch of females is higher than that of males (Xiu et al., 2022). This tone can be perceived by ears, whereas it is referred to as fundamental frequency (F0) in the physical attribute with hertz unit (Hz) (Ladefoged & Johnson, 2015, p. 25). Speech sound is a complex wave, and F0 is the lowest frequency component (Gagliardi & Tamburini, 2021). Meilán et al. (2014) mention that the AD group tends to have a lower pitch, which may be caused by the fewer cycles of vibrations per second. Consequently, the AD group presents with a deeper voice.

2) Jitter is a pitch instability or frequency perturbation. This parameter is a measure of period-to-period fluctuations in F0 (Mahon & Lachman, 2022; Xiu et al., 2022). In the present study relative jitter (local) is employed, which indicates the average absolute difference of time between consecutive periods divided by the average period, expressed as a percentage (Abhang et al., 2016, p. 64; Boersma & Weenink, 2023d). The unit is in percentages. A lower level of jitter is found in

healthy normal, and the threshold for pathology is above 1.04% (Asiaee et al., 2020; Boersma & Weenink, 2023d).

3) Shimmer is an amplitude variation from peak to peak (Meilán et al., 2014; Xiu et al., 2022). It is an amplitude perturbation due to the glottal signal during vowel formation (Abhang et al., 2016, p. 63). It can be referred to as instability of volume (intensity) (Mahon & Lachman, 2022). Relative shimmer (local), which is used in this study, is calculated as the average absolute difference of the amplitudes between consecutive periods divided by the average amplitude, expressed as a percentage (Abhang et al., 2016, p. 64; Boersma & Weenink, 2023e). The values of shimmer above 3.81% (local) are considered pathological voices (Asiaee et al., 2020; Boersma & Weenink, 2023e). In the AD group, voices tend to be tremulous with a lower intensity than in the normal group (Meilán et al., 2014).

4) Number of Voice breaks (NVB) is derived from the number of distances between consecutive pulses (burst of air when vocal cords open and close) that are classified as voice breaks (Ambrosini et al., 2019; Mahon & Lachman, 2022). The sudden change of pitch results in voice breaks and is considered a voice disorder (Meilán et al., 2014). A voice break is depicted as the time interval between consecutive pulses longer than 1.25 (constant number) divided by the pitch floor (Ambrosini et al., 2019; Asiaee et al., 2020; Boersma & Weenink, 2023c). Generally, the default value of pitch floor is at 75 Hz, then the time window between consecutive pulses longer than 16.67 milliseconds is considered as a voice break (Boersma & Weenink, 2023c). Since voice breaks are very short duration, this parameter is not perceived by human ears. This parameter tends to be lower in normal voices than in pathological voices (Asiaee et al., 2020). The speech of AD groups was found to present a higher NVB.

5) Harmonics-to-noise ratio (HNR) is a proportion of energy in the harmonics of the speech signal (periodic) to the noise energy (aperiodic), expressed in dB (Boersma & Weenink, 2023a; König et al., 2019). The signal-to-noise ratio is another term of HNR, in which signal is referred to the signal in the periodic part (Asiaee et al., 2020; Boersma & Weenink, 2023a). HNR is the degree of acoustic periodicity representing a voice quality (Boersma & Weenink, 2023a). This parameter tends to be higher in normal speech, while a hoarse speaker and pathological voices

have a lower HNR. The threshold for detecting pathological voice is below 7dB (Asiaee et al., 2020; Boersma & Weenink, 2023a). Lower HNR indexes higher noise in the voice signal associated to bubbles or tremors (Meilán et al., 2014).

ii. Temporal domain

Most of the variables in the temporal domain are produced by speech and pauses. Throughout the present project, 'utterance' refers to speech segment, phonation, or voiced part of verbal responses (Ambrosini et al., 2019; Beltrami et al., 2018; Roark et al., 2011) while 'silence' means unvoiced segment longer than 1,000 milliseconds (msec), long pauses, and hesitation (Roark et al., 2011; Satt et al., 2013; Singh et al., 2001; Weiner et al., 2016). Both utterance and silence are extracted from audio data of verbal responses by identifying boundaries of voiced and unvoiced segments (de Jong & Wempe, 2009; Khodabakhsh et al., 2015; Weiner et al., 2016). Acoustic features like duration are analyzed by reference to the number of utterance and silence segments produced. These parameters are the bases to calculate several relevant variables such as proportion, rate, and so on (Beltrami et al., 2018; Weiner et al., 2016).

Notably, slight changes in physiological and cognitive components sensitively affect speech which is noticeable in acoustic changes (König et al., 2018). Longer hesitation periods and slower speech speeds, in particular, seem to be connected to the vocal characteristics of dementia, especially in the early stages of the disease (Filiou et al., 2019; König et al., 2018; Toth et al., 2018). These variables are intrinsically related to difficulties in discourse planning, and difficulties in fluency and discourse planning are associated with hesitations or the time spent to correct spoken discourse, possibly due to some form of cognitive lapses, such as difficulty in retrieval of semantic or episodic memory and inability to retain information in working memory (Baddeley, 2000; Celine De et al., 2018). Hesitation manifests itself via fillers and pauses. Regarding their function in the discussion, fillers are taken into consideration. They may appear at the beginning or conclusion of a conversation to signal that something is difficult to comprehend, or that additional information is needed. They may also appear in the middle of a conversation to suggest that something needs to be clarified, retracted, or reworded. Pauses have a variety of interpretations and encompass other linguistic features, for instance, difficulty in utterance, deficits in accessing semantic storage, and impairments in syntax and discourse planning (Beltrami et al., 2018; Boschi et al., 2017). Those difficulties involved several cognitive domains, such as executive function (Singh et al., 2001).

Studies aiming to identify individuals with neurodegenerative progression have accepted acoustic features as one of the most discriminating features, especially the variables derived from spontaneous speech. A study of German native speakers found significant differences in acoustic parameters, such as F0 and loudness, between controls and patients with Parkinson's disease (Skodda et al., 2014). The German study used a reading task which was adequate to assess the performance of the two groups of participants and to try to differentiate healthy control and patients with Parkinson's disease on the basis of the acoustics parameters. In longitudinal studies, acoustic features from spontaneous speech showed significant change over a year (Robin et al., 2023).

In the machine learning classification of Al-Hameed et al. (2017), they obtained the best classification accuracies for healthy control and AD at 94.7%, healthy control and MCI at 95.0%, and AD and MCI at 95.0%. They suggested that accuracy increased by the time of the visit. Table 2 presents the sample cross-sectional studies in differentiating between AD, MCI and healthy control with the key variables in each test. The discriminative variables could be duration of speech, articulation, pauses and hesitation (Beltrami et al., 2018; König et al., 2018; Roark et al., 2011; Toth et al., 2018). Although each study named their variables different from other studies, they shared similar definitions and transformed formulation. Behind the three sample tests, the deficits in working memory and attentional-executive processing were speculated to be responsible for acoustic representation in both MCI and AD (Gosztolya et al., 2019; Hodges, 2007; Jokel et al., 2019)

Table 2 Sample cross-sectional studies in differentiating between AD, MCI and healthy control

Literature		Recall task	Picture description	Interview
Beltrami et al. (2018) compared among AD, MCI, & HC	Task	The last dream	Figure of a living room with some characters carrying out certain actions	Describe a typical working day
	Significant differences	- Silence segments duration - Speech segments duration - Transformed phonation rate - Standardized pause rate	- Speech segments duration - Transformed phonation rate - Standardized pause rate - Utterance length	- Silence segments duration - Speech segments duration - Transformed phonation rate - Standardized pause rate - Utterance length
König et al. (2018) compare between MCI & AD	Task		Photography of one animal in its natural environment	
	Significant differences		- Durations of silence segments - Durations of voice segments - Durations of unvoiced segments - Durations of voiced segments	
Toth et al. (2018) compared	Task	Black and white films		Describe the previous day
	Signifi	- Duration - Speech rate		- Duration

Literature	Recall task	Picture description	Interview
between HC & MCI	- Articulation rate - Total length of silent pauses - Length of pauses		- Number of silent pause - Number of filled pauses - Number of pauses, total length of silent pauses - Total length of filled pauses - Length of pauses - Lenthg of filled pause per duration

2.4 Previous studies

This section presents the exemplar literatures for the current study. All of the studies in this section investigated speech product of non-Thai participants. Speech-markers for detecting MCI and AD had received more attention yielded stronger significant evidence. Also, the techniques in speech processing and analysis had become more delicate according to modern technology.

The study of Hoffmann et al. (2010) They examined spontaneous speech of Hungarian native-speakers aged over 55 years with AD (classified into mild, moderate and sever level) and healthy control by asking them to talk about their visiting reason, critical life events and everyday activities. The recorded conversation was trimmed into 4-minute speech sample, then it was extracted into four variables, i.e., grammatical error ratio (total number of grammatical errors divided by total phonemes), speech tempo (total phonemes divided by locution or four minutes in this study), articulation rate (total phonemes divided by articulation periods exclusively), and hesitation ratio (total duration of hesitations divided by four minutes). All four temporal variables were found to show statistically significant differences among four groups of the participants. Their findings confirmed the effect of AD on temporal domain of acoustic features. Especially hesitation ratio, longer duration of hesitation explicated in relatively severe level. Besides, hesitation ratio was able to distinguish mild AD from healthy older adults. They further explained that lexical access and word finding difficulties might be accounted for poorer performance in AD (compared to healthy control), i.e., longer hesitation, shorter speech tempo and articulation rate, and higher grammatical error ratio.

The studies in speech-markers for detecting MCI and AD were developed with the changing technology trend. Automatic tools were introduced in each process of speech analysis and participants classifying, such in the study of Toth et al. (2018). The acoustic parameters were extracted from speech responses of healthy control and MCI. The tasks were immediate recall after presenting of the short film (the first story), delayed recall about the short film (the second story) after one-minute of delayed interval, and interview about the participants' previous day. The researchers compared the speech variables extracted by two different techniques, including manual analysis by the Praat software (Boersma & Weenink, 2022), and their developed automatic speech recognition (ASR). Several acoustic features were statistically significant different between two groups of the participants, e.g., speech tempo, hesitation ratio, length of utterance. They stated that speech tempo in the delayed recall task and the number of pauses in the interview showed the most significant differences between MCI and healthy control.

Acoustic features were investigated their ability to distinguish individuals with MCI from healthy control in several aspects rather than temporal domain. Themistocleous et al. (2020) determined the distinguished ability of two acoustic aspects, including voice quality and speech fluency. The recordings of the classical picture description task, the Cookie Theft, produced by MCI and healthy control were extracted and analyzed. Voice quality variables accounted for the adjusting of the sublaryngeal and laryngeal systems. The variables in these groups were calculated based on the amplitude, periodicity in the voice signal, and frequency. While the speech fluency variables were calculated based on the number of syllables, duration of articulation and locution. The speech parameters in both groups of acoustic features differed significantly between MCI and healthy control, such as shimmer, and articulation rate. For speech fluency, MCI presented slower articulation rate than healthy older adults which can be the results of slower cognitive process such word recall and grammar, and also an impairment of motor control in articulation. The fine control of articulatory organs was in line with the result of shimmer. Patients with MCI showed greater shimmer than healthy control which indicated greater instability of amplitude and less control of the sublaryngeal or pulmonary pressure.

The analysis of spontaneous speech in detecting AD and MCI is applied across different languages, and it presented promising results in several studies. The international pilot study of János et al. (2022) revealed that four temporal variables from the interview showed significant differences between MCI and healthy control in both Hungarian-speaking and English-speaking groups, i.e., silent pause duration rate, total pause duration rate, silent pause average duration, and total pause average duration. This study analyzed the speech products derived from one opened question asking about the participants' activities in the previous day. For the study in non-European countries, a pilot study in China presented a promising prospect in the early detection of AD with the speech analysis technique (Qiao et al., 2020). The participants with AD, MCI and healthy control groups underwent the Cookie-Theft picture description test. The results demonstrated significant differences of speech acoustic features between the three groups, e.g., number of speech segments, and ratio of hesitation/speech counts.

The present project focuses on the diagnostic potentiality of the acoustic features of speech to detect early sign of cognitive impairment and to differentiate between AD, MCI, and cognitively intact individuals in Thai. Since In Thailand, there are few studies on the aforementioned framework provided by the project. In considering the selection of cognitive tests, the assessment tools used to elicit natural speech need to be developed based on the specific purpose and the specific population (Thai). Since linguistic factor and education have an impact on sensitivity of the cognitive assessment, this study intended to minimize the confounding factors by creating the most suitable spontaneous speech tests for Thai older adults. It was found that 71% of Thai dementia assessments in the psychometric index studies were translated tests and the literacy bias was found in 52% of the validation studies (Na Chiangmai & Wongupparaj, 2020). Moreover, Thai language is uniquely different with English in linguistic and semantic aspects (Sangchocanonta et al., 2021). The current study aims to develop language stimuli that reduce potential confounding factors and enhance their capability to elicit spontaneous speech. Three spontaneous speech tests are developed, including story recall, picture description, and semi-structured interview. These three tests acquired different levels of restriction and freedom to response, they thus selected to be explored in this study. New stimuli are developed for the short story and picture suitable for Thai older adults.

So far as this study was conducted, there were no similar studies in Thai older adults that developed the spontaneous speech tests for speech analysis in detecting MCI and AD. There was a group of projects which had similar purposes but different methodology. Thai researchers developed the picture description task for AD screening, but they extracted part-of-speech, lexico-semantic features and some of acoustic-phonetic features, namely, duration of silence pauses (Amonlaksananon et al., 2021; Munthuli et al., 2021; Nagarachinda et al., 2020; Sangchocanonta et al., 2021). In this study, the samples of spontaneous speech collected from the participants are analyzed for their content variables and acoustic features (frequency and temporal domains) are expected to have discriminative ability in distinguishing between healthy control, MCI and AD. The final outcomes and products of the current study are a collection of Thai spontaneous speech tasks; a set of spontaneous speech samples collected from the participants; a fine-graded analysis of such rich database; and a set of significant variables which can distinguish MCI, AD, and cognitively intact elderly people.

CHAPTER 3

RESEARCH METHODOLOGY

The objectives of this study are to create dementia screening tools based on the analysis of spontaneous speech and to validate the developed tools with known groups of MCI, Alzheimer’s disease, and cognitively intact older adults. The study is cross-sectional quasi-experiment research with comparative design. All participants were examined with the same measurements for diagnosis and validation of the proposed tools. The main research procedure consists of two phases:

Phase 1 Developing speech tasks for Thai older adults

Phase 2 Collecting and analyzing data on the proposed tasks.

The details of these two phases are shown in the following figure.

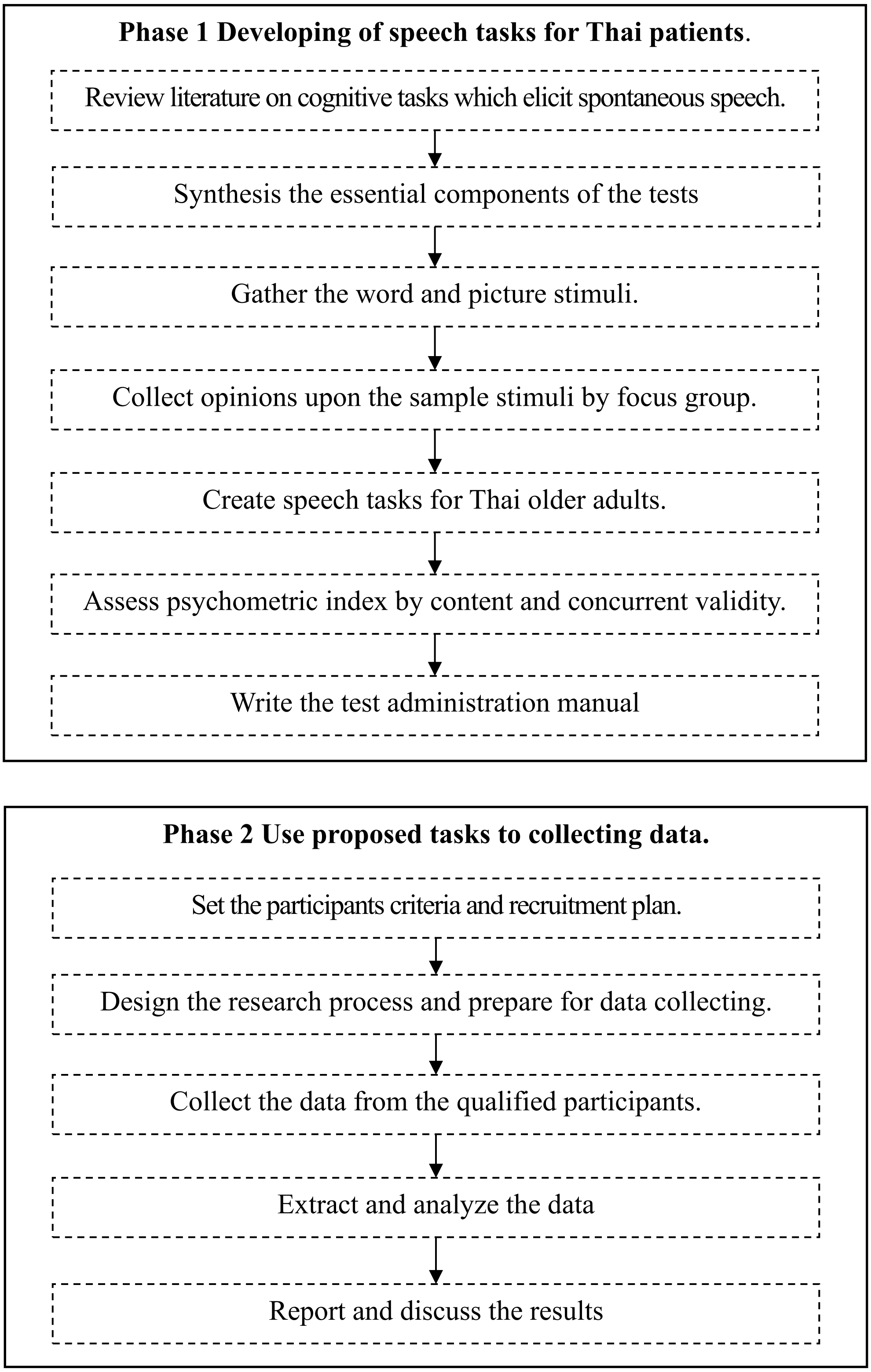


Figure 6 Research process

3.1 Phase 1 Developing the tree experimental tasks on spontaneous speech used in the study

The study process in this phase is presented in the following flowchart.

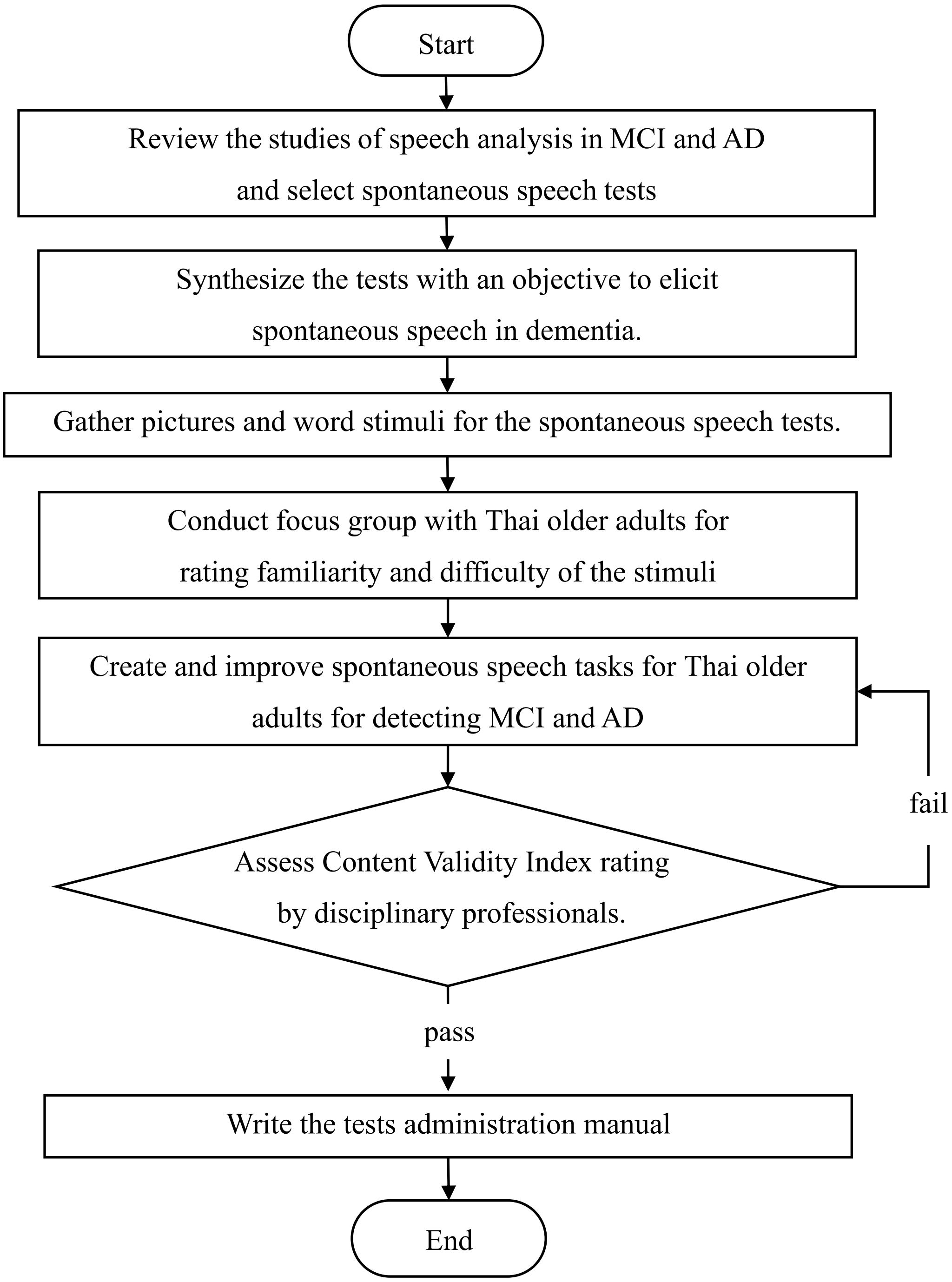


Figure 7 Development process of speech tasks for Thai patients

3.1.1 Identifying essential components of the cognitive assessment used in speech analysis.

There are several cognitive assessments used to elicit speech responses in dementia studies. This study selected three speech tasks to be further developed to assess the cognitive performance of older Thai adults: (1) picture description; (2) story retelling; and (3) semi-structured interview. The main reasons for the choice of the three tasks are that (a) according to previous research, picture description and story retelling are very productive for the analysis of diagnostic acoustic features of cognitive impairment; (b) picture description and story retelling rely on semantic memory that is also impaired in dementia. The episodic memory component of the story recall task is limited to a task-related recent memory; (c) the semi-structured interview involves long-term memories and is closely tied to everyday communication.

The three aforementioned tasks were derived from existing standardized and well-known tasks.

1) The picture description task is standardly used to assess language ability in tests such as Boston Diagnostic Aphasia Examination (BDAE: Goodglass & Kaplan, 1972) and Western Aphasia Battery (WAB: Kertesz, 1979) (Mueller, Hermann, et al., 2018). The components of picture description tasks, including stimulus illustration, instruction, and scoring, were derived from the Cookies Theft picture (Giles et al., 1996), Kentucky Aphasia Test (Marshall & Wright, 2007; Nagarachinda et al., 2020), and Norman Rockwell prints (Bayles et al., 1999; Chapman et al., 1998; Tomoeda et al., 1996). Specifically, the stimuli were selected from the Thai version of the short Boston Naming Test in the Consortium to Establish a Registry for Alzheimer’s Disease (CERAD) (Aniwattanapong et al., 2018; Tangwongchai et al., 2015) since this test takes into account Thai culture and allows participants to be familiar with the represented setting which is known to affect retrieval (Larner, 2017; Lim et al., 2018).

2) The story retelling task is also standardly used to assess language impairments (Hodges, 2007). The short story used in this project and the scoring criteria were derived from the sub-test “Logical Memory” in the Wechsler Memory Scale-IV (WMS-IV) (Drozdick et al., 2013, pp. 20-21; Wechsler, 2009), the Arizona

Battery for Communication Disorders of Dementia (Armstrong et al., 1996), and the Babcock story recall test (Khan, 1986; Lezak et al., 2012). Since familiarity modulates word-retrieval, stimuli with different ranges of familiarity this study were included in the short story.

3) The topic of the semi-structured interview, also used to assess language impairments, varies across studies but is generally focused on familiar, everyday events and routines. In preparing the task for this project, studies such as Beltrami et al. (2018); Gomez and White (2006); Tröger et al. (2018) were use as guideline.

The preparation of the tasks and material were also based on the outcomes of a focus group conducted with 13 cognitively intact adults who are at least 55 years old (Creswell & Creswell, 2018; Silpakit, Sukying, et al., 2017). The purpose of the focus group was for brainstorming to collect information about the to-be-selected stimuli and to obtain rating on the familiarity and difficulty of the word and picture stimuli. The proposed stimuli also integrated the cognitive phenomena which aimed to imitate everyday activities. The phenomena such the phonological loop were embedded in the picture stimulus and short story, including the phonological similarity effect, and the phenomenon of transferring information between codes (Baddeley, 2000).

After acquiring the essential components, the stimulus in picture and word forms was selected before developing the spontaneous speech tasks. The important stimuli in this study will be in audio and visual modalities. Single words and pictures were gathered from the standardized test in the previous step and added more objects regarding cognitive effects such as phonological similarity and familiarity effect. Then the pictures and words were used to create two types of stimuli which are a short paragraph of a narrative story and a black-and-white picture of one situation. The question in the semi-structured interview was selected from the responses of the focus group and based on autobiographical memory and simple events in daily life.

3.1.2 Psychometric indexes.

After the three tasks were constructed, they were assessed for their validity and adjusted afterward. A content validity test was conducted. The three proposed tasks were rated by professionals from different disciplines, such as experts in implementing of cognitive assessment, in developing of psychological measurement, and in dementia. They were asked to rate the content validity index (CVI) along five dimensions: (1) consistency with the basic theory and approaches, (2) appropriateness of the stimuli, (3) procedures of administration and construction, (4) appropriateness of the scoring criteria, and (5) the overall adequacy of the tasks.

3.2 Phase 2 Collecting data by the proposed tasks

The study process in this phase is presented in the following flowchart.

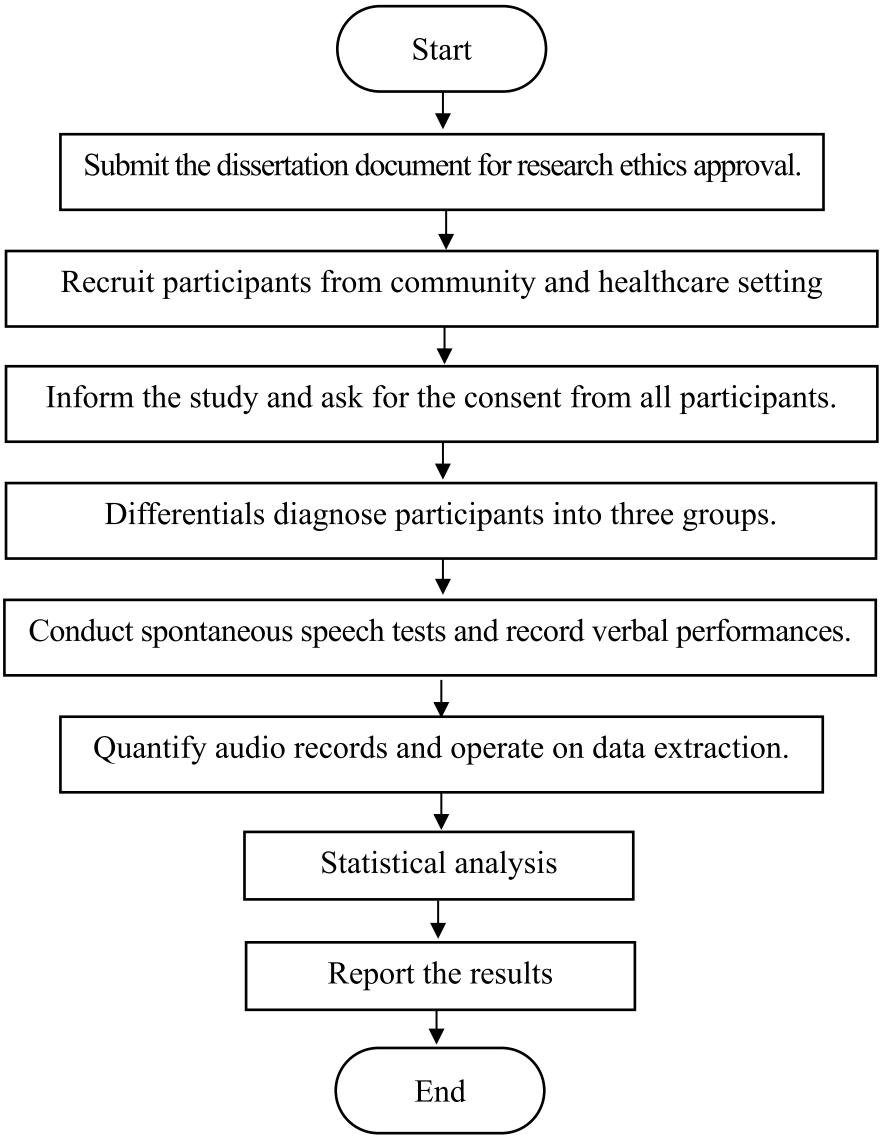


Figure 8 Collecting the data and analysis process.

3.2.1 Participants

Three groups of Thai older adult participated in the study, i.e. a group of people with Alzheimer’s disease (AD), a group of people with mild cognitive impairment (MCI), and a group of cognitively intact people, acting as a healthy control group (HC). Each group consisted of 25 individuals. The number was determined on the bases of current literature indication. A scoping review by Filiou et al. (2019) reported that the majority of the studies in spontaneous speech assessment included less than 15 participants (AD or MCI) in their studies. Mueller et al. (2016) conducted an AD risk assessment study by analyzing samples of connected speech of 39 individuals with Alzheimer and 39 healthy controls, for which a sensitivity analysis indicated a power of .80 to detect effect sizes of 0.40 - 0.50. According to McMillan and Schumacher (2014, p. 156), the sample size for a comparison study should be at least 15 participants and 8-10 participants for a highly controlled experiment. For this project, the sample size was also calculated by G*Power (v.3.1.9.2) (Faul et al., 2007). A three-group comparison by MANOVA with 2 – 6 variables, an effect size of .40, alpha of .05, and power of the test of .80 (Cohen et al., 2007; Kellar et al., 2013, p. 110) suggested to collect the data from between 18 - 30 participants. Thus, for this study the chosen number of participants is 25 people per group, for a total of 75 people.

Participants were recruited from hospitals, community healthcare units, and communities in Chonburi. Their ages ranged between 55 and 80 years. The researcher advertised and opened the application until each group reached 25 people. Participation was aided by word of mouth of previously enrolled participants, by the community health practitioners, or by the Village Health Volunteers (VHV).

On the data collection day, at the beginning of the session, all participants received information about the study and about the written informed consent to participate, as well as information about the possibility for the participant to withdraw from the study at any time. All participants were encouraged to bring a companion, such as a family member (if available), to the session. In case participants attended the test session alone; they were asked for the contact of their family member or close person at the beginning of the session. This practice is for an emergency that might happen during the experiment session.

All participants were screened with the same set of assessments shown in the instrument section (see Section 3.2.3.1). They should meet the inclusion criteria which are stated in the below section. Then they were classified by the standardized criteria suggested by the medical doctor, i.e., Neurological Institute of Thailand (2021); Petersen (2004)) and neuropsychological assessments (i.e., Morris (1993); Nasreddine et al. (2005)). The MoCA scores were additionally considered for the classification (Tangwongchai et al., 2009). Participants were classified into three groups, i.e., healthy control (HC), mild cognitive impairment (MCI), and Alzheimer's disease (AD).

Inclusion criteria

1) Voluntarily participate

2) Aged between 55 – 80 years

3) Native Thai speaker

4) Normal hearing and sight; wearing glasses is acceptable.

5) No known history of drug or alcohol abuse or history of neurological or major psychiatric illness (Gosztolya et al., 2019; Themistocleous et al., 2020).

Exclusion criteria

1) Participants need to withdraw from the study

2) Participant cannot complete the assessment process

3.2.2 Research design

The current study is a cross-sectional study and quasi-experiment research with a comparative design (McMillan & Schumacher, 2014). Three groups of participants were differentiated by a preexisting variable, i.e., dementia pathological stages. The analyzed behavior consisted of spontaneous speech, taking into account linguistic and acoustic features. Care was taken to equate the three groups by controlling possible confounding variables such as sex, age range, and educational level.

3.2.3 Tools and Instruments

Three sets of assessments were used in this study and two types of recording devices; screening tests; cognitive tests; spontaneous speech tests.

3.2.3.1 Screening assessments

Six questionnaires, self-report, and cognitive tasks were included.

i. Patient Health Questionnaire (PHQ-9)

The Thai version is a 9-item questionnaire, where the participants rate the frequency of certain feelings and behaviors on a Likert scale from 0 (not at all) to 3 (nearly every day). The total scores range between 0–27. It had satisfactory internal consistency with a Cronbach's alpha of .79, and an optimal cut-off score of ≥ 9 returned a sensitivity of 84% and a specificity of 77% (Lotrakul et al., 2008).

ii. Barthel's Activities of Daily Living Questionnaire (ADL)

The Thai version consists of 10 items evaluating the ability to perform certain tasks, e.g., using the toilet. Different scores are associated with the items, ranging from 0–1, 0–2, and 0–3. The score 0 indicates inability to do the activity on one' own and 3 the ability to do it independently. The maximum score is 20 points. In the study of Jantapo and Kusoom (2021), they report a high interrater reliability of .92. They also stated that the scores could be grouped into three ranges, i.e., 0–4 severe dependency, 5–11 moderate dependency, and 12–20 independency.

iii. Instrumental Activities of Daily Living Scale (i-ADL)

This assessment was adopted from the guideline of the Neurological Institute of Thailand (2014). This version is a 6-item questionnaire to evaluate older adults' independent levels of accomplishing the activities which require the use of instruments, e.g., the telephone and money. The response range is 0 (unable to use certain tools), 1 (able to use with assistance), and 2 (independently using certain tools). The total score was 12 points. Performance can be classified into four groups, namely, full dependency (0–2), severe dependency (3–5), moderate dependency (6–8), and independency (≥ 9) (Itprasert, 2017).

iv. Five-minute Hearing Test

This test was developed by the American Academy of Otolaryngology-Head Neck Surgery, aims at determining whether a person should be evaluated and treated for hearing loss. The test was translated into Thai with the original number of 15 items. The participants can rate their hearing on a scale of 0–3, where 0 is never, and 3 is almost always. The total score is 45 points, with 10 and higher points indicating hearing disabilities. This optimal cut-off score gave a sensitivity of 86% and a specificity of 69% among the Thai cohort (Yimtae et al., 2011).

v. Montreal Cognitive Assessment (MoCA)

The MoCA was developed by (Nasreddine et al., 2005) and translated into Thai by (Hemrungrojn et al., 2009) and it is accepted as a clinical evaluation for MCI (Vichitvejpaisal et al., 2015). For the Thai version, it reported a Cronbach's alpha of .91. The suggested optimal cut-off scores for MCI are 22 – 25 points, giving a sensitivity and specificity of 0.80 while suggested cut-off scores for AD screening is 21 and under, this cut-off had a sensitivity of 1.0 and a specificity of 0.98 (Tangwongchai et al., 2009).

vi. Clinical Dementia Rating Scale (CDR)

The CDR was adopted from Boonpeng (2014), which is a Thai version of Morris (1993) scale. The examiners interviewed the participants in six categories and rated them on the level of decline. This study opted for the global CDR scores because this approach primarily emphasizes the memory category, which is the hallmark of AD (Alzheimer's Association, 2022). The global CDR is calculated based on a standard algorithm proposed by Morris (1993) that weights memory as the primary category and all others as secondary. The algorithm calculation can be checked at the website of the National Alzheimer's Coordinating Center, https://naccdata.org/data-collection/tools-calculators/cdr. The severity of dementia is indicated by five ranges of the global CDR, i.e., no dementia (0), very mild dementia (0.5), mild dementia (1), moderate dementia (2), and severe dementia (3) (Neurological Institute of Thailand, 2021).

3.2.3.2 Cognitive tests

Two conventional cognitive tests were administered in this study.

i. Verbal fluency tests

Both Phonemic and Semantic Fluency tests were used to support the results of the speech tasks as they are commonly used as language assessments. The two versions of verbal fluency (VF) have similar protocols except for the stimulus instruction. This study asked the participants to generate words in the 'animal' category, and words begin with 'Kor' (ก) letter (Charernboon, 2018). In each fluency test, participants were asked to generate as many words as possible within 60 seconds. The instructions specified that repeated words and proper names were wrong responses. They also specified that words with similar prefixes or first words were considered as repeated words, e.g., ‘กิน’, ‘กินข้าว’, ‘กินน้ำ’. These words begin with the same first word but have changed only the second sequence. If participants began listing words with a letter or a category different from those specified in the instruction the examiner stopped the participant and explicated the correct letter or category again. In case the participants asked whether they gave the correct words, the examiners replied by saying, ‘Please go on’ without stopping.

Participant provided their responses in written form. For the scoring, only correct words were considered, and one point was allocated to each correct word beginning with the letter Kor or was an animal. Repeated words were only scored once in the correct list. However, any repetition contributed to the so-called VF Rep index. For example, when a participant said ‘elephant’ twice, a correct response was scored as well as a repeated score of one was computed. If a participants said it three times, repeated scores were two. All repeated points were summed for each VF; thus, there were two repeated scores, namely repeated words in category fluency (CF Rep) and repeated words in letter fluency (LF Rep).

Each fluency task was scored separately; consequently, there were two total scores which were called category fluency score (CF) and letter fluency score (LF).

ii. Digit span tests

In the study, both the forward and the backward version of the digit span test (DS) were assessed, using the digit lists of Monaco et al. (2013). Each span has lists with an increasing number of digits with different orders. The examiner clearly pronounced the list of digits at a rate of one digit per second while the participant listened carefully. At the end of the list, in the forward digit span (FDS) condition participants had to orally reported the digits in the exact order of mention. In the backward digit span (BDS) condition, participants had to orally report the digits in backward order, from the last to the first.

First, participants were presented with the shortest (3-item) list. If they successfully report the list, the next list, the 4-item, list is presented, and so on up to the 9-item (for FDS) and 8-item (for BDS) list. If they fail the 3-item list, another version of a 3-item list is presented. If they fail again to correctly report the digits the test stops. If they succeed, then the 4-item list is presented. The final score is the number of items in the longest list which the participants could correctly report. For example, if the participants can finish the sequence of seven digits but fail in both lists of eight digits, they receive seven points.

3.2.3.3 Spontaneous speech tests

Three tests were developed in the current study. They aimed to elicit spontaneous speech in older Thai adults to detect the dementia pathological stages.

i. Thai Picture Description Task (TPD)

Picture description began with presenting the pictorial stimulus (See Figure 18) and asking participants to tell a story pertaining to the seen objects, persons, and situations depicted in the picture. The instruction was, ‘Please tell me what is going on in this picture as much as you can’. If the participants started to mention each object, e.g., ‘This is a boy. This is a girl’ or ‘There are TV, curtain, broom, and coconut,’ the examiners asked them not to describe the picture but tell the story in general. This prompt aimed at helping a person who might have misunderstood the prior instruction and can only deliver once. While the participants verbally told the story, interruption, pointing at the illustration, or answering any question were avoided. Nonverbal encouragement could be given by nodding to encourage the participants to continually tell the story. Nodding was not used as an answer to inquiries asked by participants. Nonverbal encouragement should not be given too often. The examiners tried to naturally imitate everyday communication. The finished when the participants explicitly stated so. When participants were quiet for more than 10 seconds the examiners could ask whether they finished the story by saying, ‘Is there anything you want to say’ or ‘What else is happening.’ In the Thai Picture Description Task, the duration of telling the story is not time-limited in this initial study. Because this is the first data collection of the developed stimulus picture, this study aimed to obtain all possible range of responding time. Typically, the participants finished their storytelling within three minutes.

The amount of information generated was counted and scored. The spoken response needed to be transcribed into text form. Each word would be considered as a correct or incorrect content unit; only a content unit that is congruent with the picture and the story itself was given credit. The procedure was the following: First, each transcribed story was segmented into word units with the latest version of Thai tokenization (TLTK 1.5.7), which is available online (Aroonmanakun, 2002). Each word was reckoned as one information unit. Each word was inspected to determine whether it was a correct or incorrect information unit. Correct information units (CIU) were defined as relevant, informative, and accurate words. Only CIU could score one point. Briefly, words with the following characteristics were not counted as CIU; repetition, inaccurately portraying the picture, error sounding, incomplete words or phrases, unspecified pronoun or referent, redundant conjunctions or modifiers, filler words or phrases, and commentaries on the task, on performance, and on personal experience (see further details in Appendix A).

ii. Thai Story Recall Task (TSR)

The examiner told participants to listen carefully to the short story that was going to be presented, and to recall the whole story as accurately as possible at the end of presentation. Then the examiners read the story at a normal-to-slow pace, taking approximately 90 seconds. Figure 9 presents the story. Immediately after the story had ended, participants recalled the story for the first round (immediate recall). After some intervening tasks (see Figure 10), participants were asked to again recall the story (delayed recall). Recalling should not be interrupted or given a hint.

phong/[1] pen khon Chonburi/[2] yá:y ma: yù: thî: i:-sǎ:n/[3] daî sì: pi:/[4]
khǎo ɔ̀:k jà:k u-bon/[5] jà ao lú:k dɔ̀:k/[6] khɔ̌:ng lên/[7] pai haî nɔ́:ng bɔ:n/[8]
lǎ:n/[9] a:-yú sǎ:m khùa:p/[10] khà-nà thî: nâng rót-fai/[11] pai nán kə̀:t
phàe:n-din waǐ/[12] thî: muea:ng phon/[13] rót-fai khà-yàp daî chá: chá:/[14]
phà:n phu:-khǎo/[15] hǐn din-da:n pai/[16] phɔ́: rə̂:m khâo tua muea:ng/[17]
mɔ:ng hěn wong-wia:n láe:w/[18] jueng rú: wâ: klaî thǔeng sa-thǎ:-ni:/[19]
kwà: jà thǔeng u-dɔ:n/[20] kɔ̂: sǐa: we:-la: pai láe:w sɔ̌:ng chûa-mo:ng/[21]

Figure 9 Short story of TSR

There was no time limit for either the immediate or delayed recall. After the immediate recall, five factual inquiries were asked about the story. Each question was not related to the other. The questions were orally presented in the same order and one after another when the participants finished each question. The examiner should not give a hint to the participants. Whether the participants replied to all or only to some of them, the questions were never repeated.

At the end of the question-answering session, the verbal fluency tests and the semi-structured interview were conducted as fillers, aiming to last 7 – 10 minutes (dot-line box in Figure 10). Letter fluency and category fluency tests were always conducted in this order. The semi-structured interview was always the last task in this sequence. At the end of the interview, the participants were asked to retell the story again (delayed recall) without cue and interruption from the experimenter.

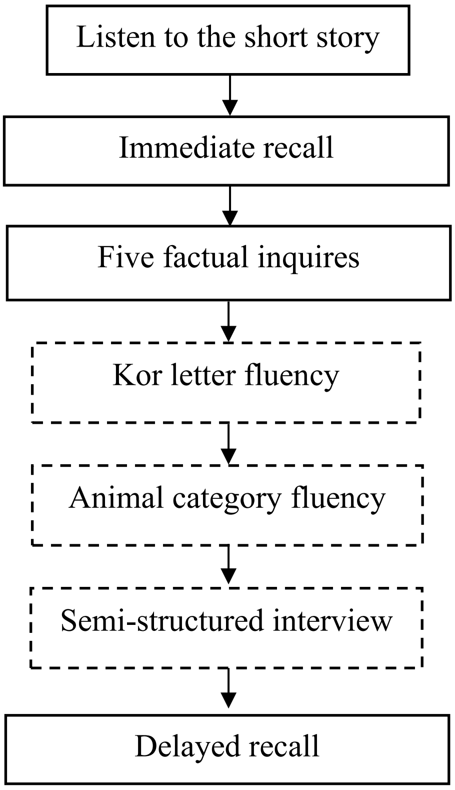


Figure 10 Tests sequence of TSR

The story was divided into 21 scoring segments with an underlying important word or phrase in each segment. Scores were given when the important items were recalled with similar formats and followed the storyline. For instance, there are three provinces in the story; the first province is a hometown, the second province is a departure, and the last province is a destination. A correct score was given to any correct name-location pairs but not when the name was associated to the wrong location. Words slightly deviating phonologically from the original word form were counted when their meaning could be ascertained, e.g., ‘Chonburi’ replaced by ‘Chon.’ The scoring method was similar in both immediate and delayed recall tests, that the total score was 21 points and followed the aforementioned rules. The correct recalled units were summed into total scores of the immediate recall (Imm) and delayed recall (Del) separately.

For the factual inquiries, one score was counted for the correct answer (Ans), counting as correct response phonological and/or semantically similar words which conveyed the meaning expressed in the story. For instance, the correct answer to the question ‘What kind of mountain?’ is “*shale*”, where the Thai word is /hǐn din-da:n/. The response can be /din-da:n/, which also means *shale*. The total score of factual inquiries was five points. (See further details in Appendix B)

iii. Semi-structured Interview for Thai (SIT)

The interview started with, ‘Please tell me about your favorite tourist attraction when you were an adolescent.’, during the participants’ verbal production the examiner asked questions from the list (Figure 11). Selecting the questions was consistent with the spoken information, e.g., when the participants talked about their company on the trip, the following question was ‘Can you name your company?’. An examiner was suggested to react naturally to the participants’ story with non-verbal encouragement such as nodding. When adequate, a short verbal prompt was given without disruption, such as ‘What next?’. The interview took approximately five minutes and a half. There is no score given to this task as responses in SIT were recorded for the acoustic features analyses.

Please tell me about your favorite tourist attraction when you were an adolescent.

- What makes you like it?
- How long did you stay/travel there?
- Can you name who was in the trip?
- What did you buy as a souvenir?

Figure 11 List of the questions for SIT

3.2.3.4 Recording devices

There are two types of recording devices in this study. The speech of almost 95% of the participants was recorded by the Zoom H1 Handy Recorder. This device generated an audio record in WAV, with a sampling rate of 44.1kHz, stereo sound, and 32-bit. The second device is a built-in microphone of the Huawei phone ELE-L29 model. It was used to record 5% of the participants when the first tool was not available. The Audio file of the second recorder was in M4A. These audio files were transformed into WAV. The properties of transformed files were 48kHz, stereo sound, and 32-bit. An independent t-test was performed to examine the differences in frequency variables between the two recorders. The mean values of pitch are not different between the two devices.

3.2.4 Procedure and data collecting

The data-collecting session was preceded by a short presentation of the project aims and procedures and the invitation to participants to ask for any details about the study and their participation. Participants were then asked to give their consent in written form. Following this, participants filled out a datasheet, which included demographic information and screening tests (see Figure 12). Research assistants helped the participants who needed assistance in filling out the sheet. The examiner did not det involve in this part to be blind to the actual condition of the participants.

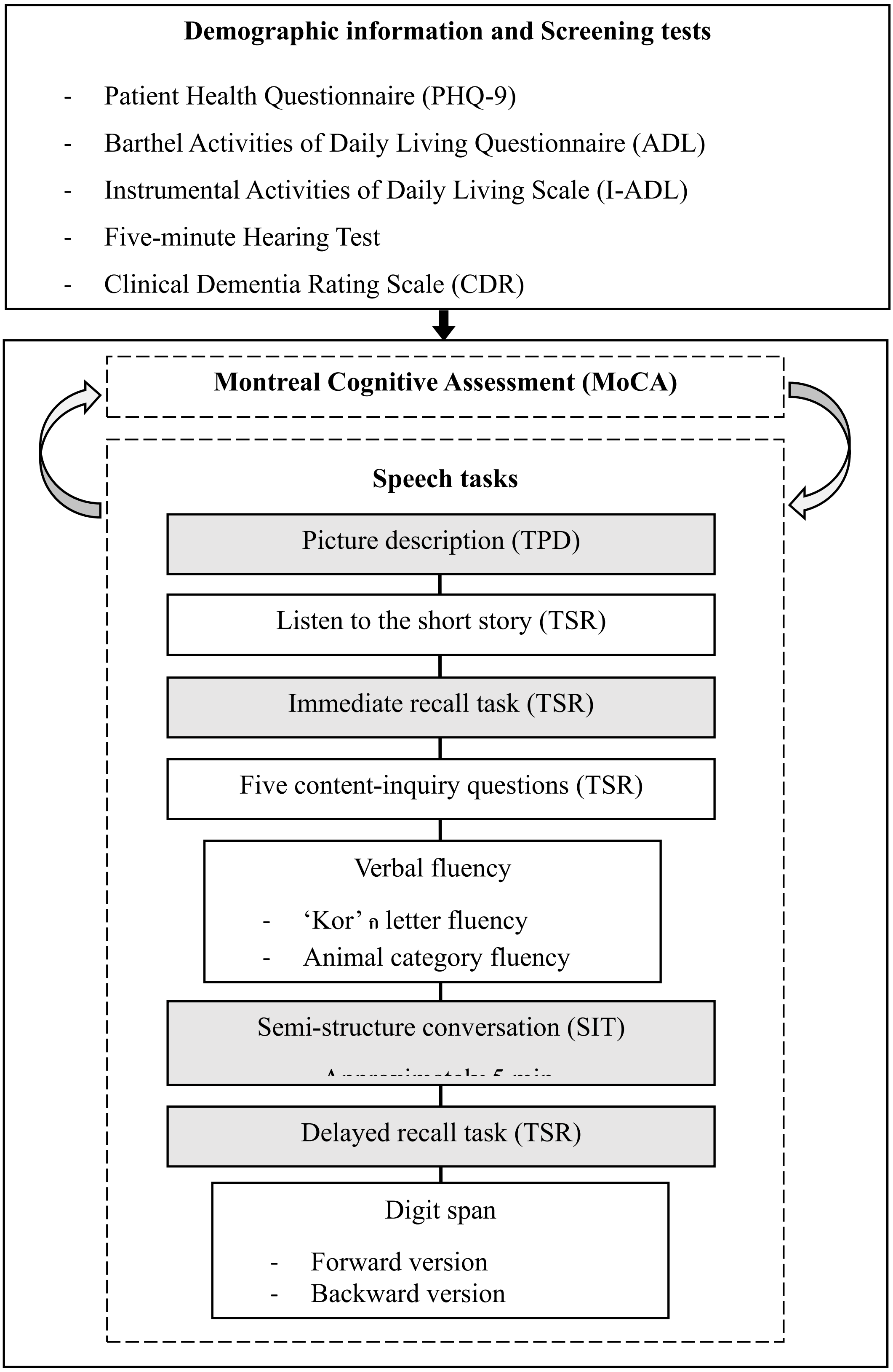


Figure 12 Experimental sequence in this study

The experimental session consisted of both neuropsychological assessments (i.e., MoCA) and speech tasks (i.e., VF, DS, TPD, TSR, and SIT). The order of the MoCA and the set of speech tasks was switched between participants. Since this study additionally conducted verbal fluency and digit span tests. Both tests were similar to the MoCA subtest, but different stimuli. For the letter fluency, the MoCA used ‘Aor’ (อ), but the set of speech tasks used ‘Kor’ (ก). The MoCA does not assess the category fluency, but this study tested for ‘animal’ category (see Section 3.2.3.2 i). For the digit span, this study used different sets of numeric spans and the participants were presented with the longer span until they fail, while the MoCA only presented one set of digits (see Section 3.2.3.2 ii). The three spontaneous speech tasks were completed during a single session and typically administered in the same order. The test sequence is shown in Figure 12. Before starting the experimental session, all participants were asked for permission to record their voice by saying, ‘While you respond to the tests presented in this session, may I record the voice? The voice will be replayed for transcribing and analysis processes. I will keep the voice file on my computer, and only me and my supervisor can access the voice file. You can ask me for more information or refuse to give permission’. In case they allowed the recording, the session started; in case they rejected it, the session was stopped. Speech responses were collected during the experiment sessions in the form of audio files. Only four of the spontaneous speech tasks were analyzed for their acoustic variables (gray box in Figure 12).

3.2.5 Speech extraction

Prior to the extraction step, speaker diarization was manually performed on the audio files for separating the voices of an examiner and participant. The boundaries of each speaker were marked and tagged on a textgrid in Praat version 6.2.10 (Boersma & Weenink, 2022). Then, the product of speaker segmentation was extracted by appropriate tools and techniques into temporal and frequency variables.

Since the audio recording was a conversation between the participants and the examiners, speaker identification was required. The audio files were visualized in Praat. The segmenting mechanism in this study was derived from the study of Khodabakhsh et al. (2015), as shown in Figure 13. The procedure consists in manually separating each recording into turns of the speakers. According to their work, a turn was considered as a segment in which the participants speak without interruption by the examiner and by long silences (silences longer than one second).

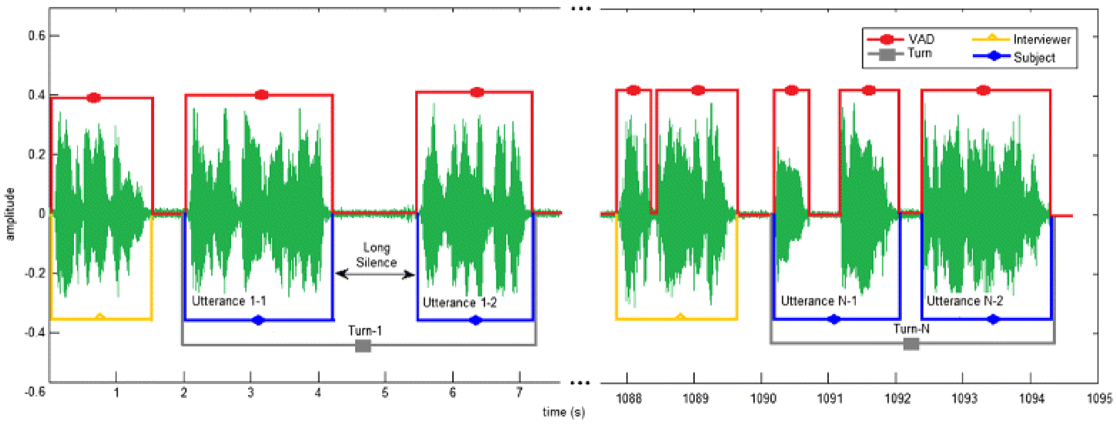


Figure 13 Segmentation of turns between a participant and an examiner (Reprint with permission from Khodabakhsh, A., Yesil, F., Guner, E. et al. (2015) Evaluation of linguistic and prosodic features for detection of Alzheimer’s disease in Turkish conversational speech. EURASIP Journal on Audio, Speech, and Music Processing. doi.org/10.1186/s13636-015-0052-y)

The boundaries of each speaker during a conversation were manually identified by listening to the audio track and examining the pitch and intensity. Visualization of pitch and intensity provided boundary information of each segment (a small part of each utterance of a speaker). The boundaries were marked on an interval textgrid. Figure 14 depicts the sound, spectrogram, pitch, intensity, and textgrid with speaker segments. The first two boxes from the top present a stereo sound with two channels. The lower dark box is a spectrogram that visualizes the pitch (blue line) and intensity (yellow line) information of this record. The bottom box is an interval textgrid that demarks boundaries of each speech segment. In this picture, ‘e’ is the symbol for the examiner, and ‘m’ is the symbol for participants’ response in the immediate recall task. Only the segments marked with the participants’ responses were processed for temporal and frequency extraction.

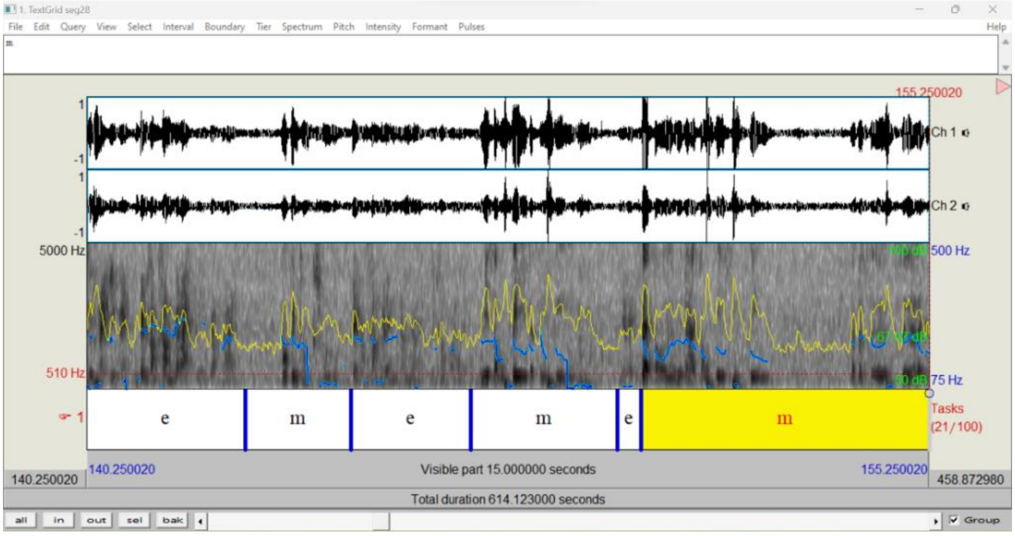
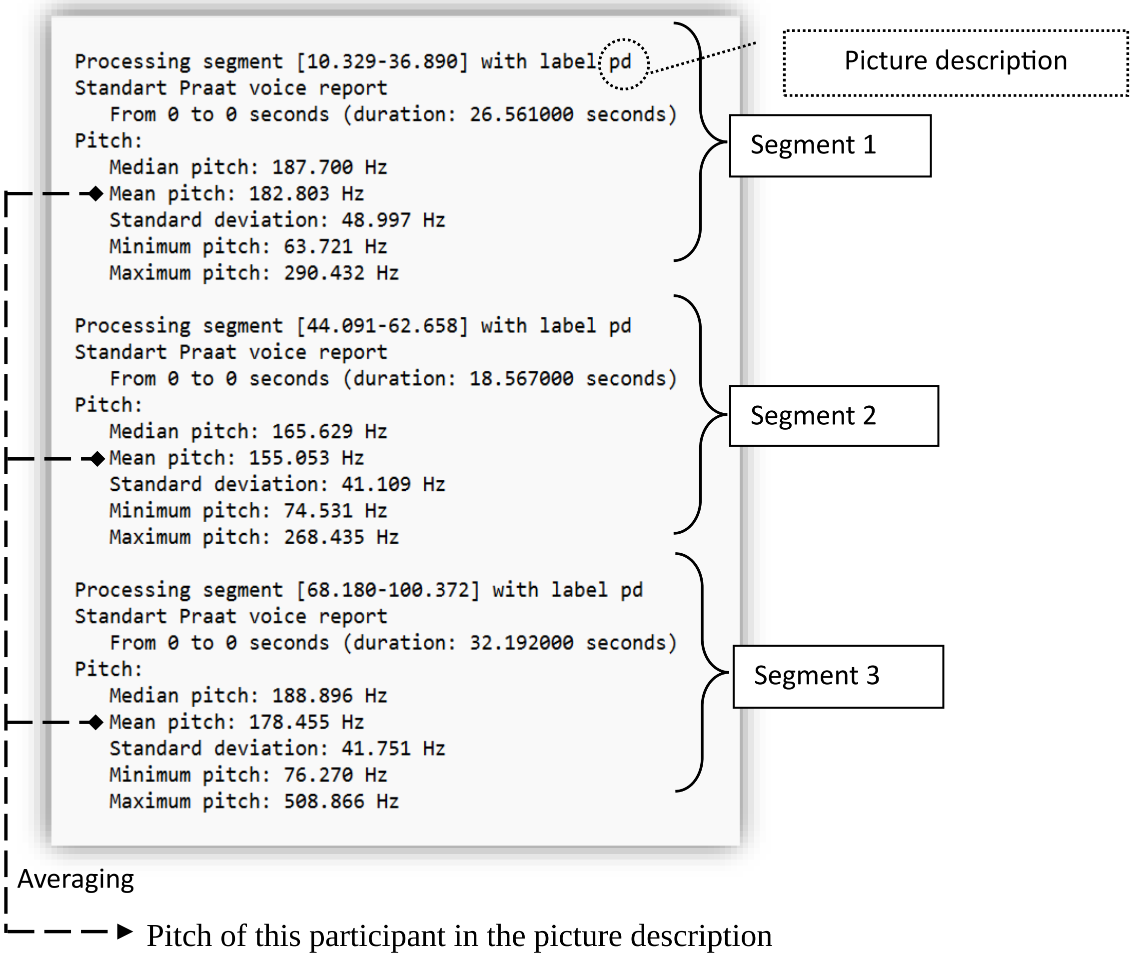


Figure 14 Speaker segmentation

3.2.5.1 Extraction of frequency domain

The frequency variables were extracted by the Python script (please see Appendix C) based on a Python library, namely Parselmouth. This library allows Python to work on the internal Praat code (Yanick Jadoul, 2022). The script was written to extract several variables within each task. However, only five variables were examined in this study, i.e., mean pitch (fundamental frequency, F_0), relative jitter, relative shimmer, number of voice breaks (NVB), and mean harmonics-to-noise ratio (HRN). The Python code can extract one task at a time. This means each the participants' recording was extracted four times pertaining to four tasks of three spontaneous speech tests, i.e., picture description, immediate recall, delayed recall, and semi-structured interview. The results provided values of each segment of a given participant in each task. An average value was used to acquire one value of certain variables per task. Figure 15 presents the results of frequency extraction from the Python script and the method of averaging the sample variable, which is the pitch in this sample, in a given task (picture description task).



```
Processing segment [10.329-36.890] with label pd
Standart Praat voice report
   From 0 to 0 seconds (duration: 26.561000 seconds)
Pitch:
   Median pitch: 187.700 Hz
   Mean pitch: 182.803 Hz
   Standard deviation: 48.997 Hz
   Minimum pitch: 63.721 Hz
   Maximum pitch: 290.432 Hz

Processing segment [44.091-62.658] with label pd
Standart Praat voice report
   From 0 to 0 seconds (duration: 18.567000 seconds)
Pitch:
   Median pitch: 165.629 Hz
   Mean pitch: 155.053 Hz
   Standard deviation: 41.109 Hz
   Minimum pitch: 74.531 Hz
   Maximum pitch: 268.435 Hz

Processing segment [68.180-100.372] with label pd
Standart Praat voice report
   From 0 to 0 seconds (duration: 32.192000 seconds)
Pitch:
   Median pitch: 188.896 Hz
   Mean pitch: 178.455 Hz
   Standard deviation: 41.751 Hz
   Minimum pitch: 76.270 Hz
   Maximum pitch: 508.866 Hz
```

Figure 15 Sample results of frequency extraction in picture description task and averaging method for the pitch parameter

3.2.5.2 Extraction of the temporal domain

Three steps were emplyed for deriving the temporal variables. The details of each step and the results are described in the following section.

1) Voice Activity Detection

Praat was performed to automatically identify silent and utterance intervals in each sound file. This study selected the Voice Activity Detection technique (VAD) in Praat, which is called *To TextGrid (silences)*. This function detects unvoiced pieces based on pitch and intensity threshold (Boersma & Weenink, 2023b). This study adopted the script for the *To TextGrid (silences)* function with the following parameter: The minimum pitch at 100 Hz, time step at 0, silence threshold -

30 dB, the minimum silent interval at 0.5 sec, and minimum sounding interval at 0.2 sec (see Appendix D) (Schweitzer, 2016). These parameters were determined by the default setting (i.e., pitch and time step), and proper value that minimized the error of incorrectly detecting silence in utterance. Figure 16 presents a sample result, where ‘P’ indicates a segment corresponding to a pause or unvoiced segment, and ‘U’ indicates an utterance or voiced segment.

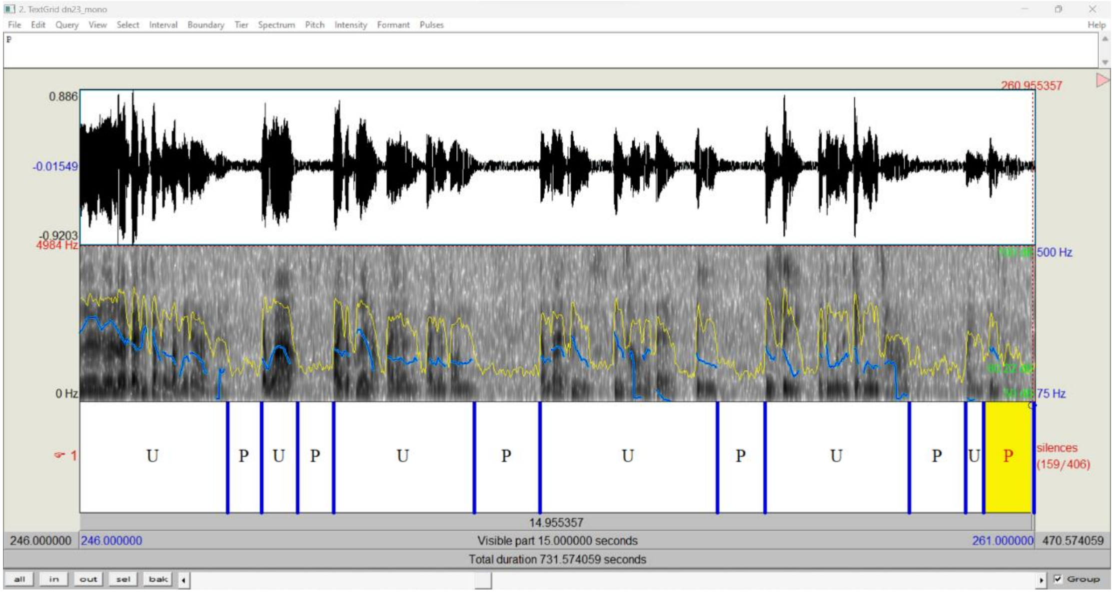


Figure 16 Utterance and silence segments resulting from the To TextGrid (silences) function.

2) The results from the primary process are segments of speakers, while the VAD was performed to separate voiced and unvoiced segments within the participants’ verbal response. Two textgrids are overlap by intersecting techniques. The Python script for intersection is a function developed by PraatIO (Mahrt, 2016). Operation of an intersection requires a textgrid file with two tiers together (speakers segments and voiced-unvoiced boundaries). The merge function in Praat allows two textgrid files to combine into one textgrid with two tiers. The results given from the intersection script were a textgrid with the third tier of intersected segments and a text file of time boundaries of the intersected segment (Appendix E). Figure 17 illustrates the textgrid product after the intersection processing. The three tiers in this picture

include voiced-unvoiced segments (silences tier), information about speakers (tasks tier), and intersected segments with information about both the speakers and voiced information (intersection tier). This picture presents the pieces of the immediate recall task (‘m’ refers to an immediate recall task).

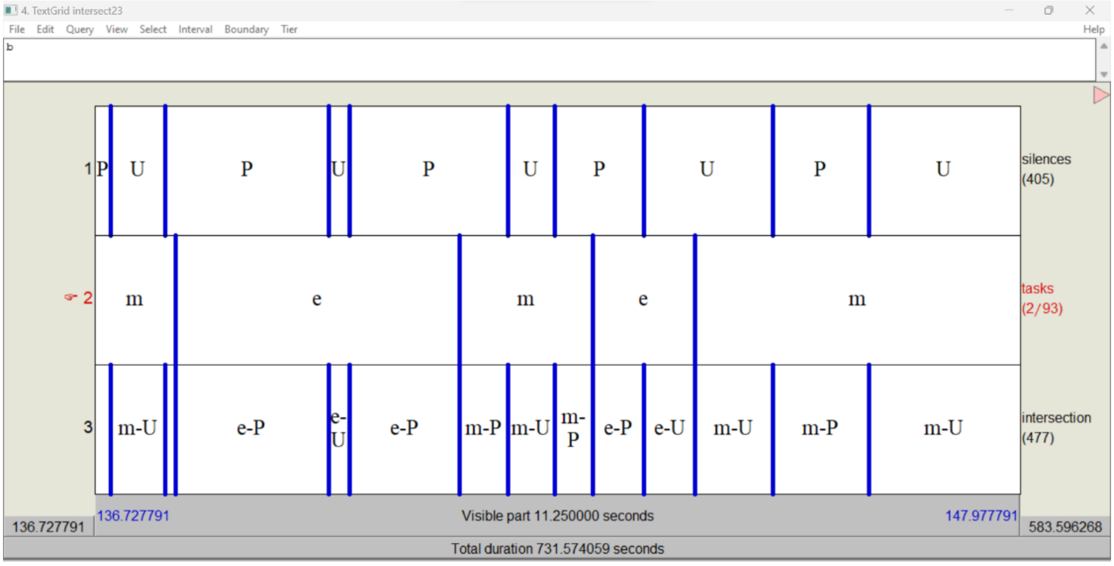


Figure 17 Intersecting segments of the textgrids of VAD and speaker segmentation

3) Text files from the intersection script would be transferred into an Excel file. The parameters were then prepared in Excel before transfer to SPSS. The first operation was re-labeling the silent events. Since the VAD labels pauses that were longer than 500 msec, this study defined silent segments greater than or equal to 1,000 msec. All silent segments which were equal to or longer than 1,000 msec were labeled as ‘P’. Silent segments shorter than 1,000 msec were included in the utterance time. Then two primary parameters were derived, which were the number of segments (utterance and silence separately) and total duration of segments (utterance and silence separately). The variables pertaining to the number of segments and total duration resulted in four variables per task (See Table 3).

Table 3 Primary parameters of temporal variables

Variables		TPD	Imm	Del	SIT
Number of segments (number)	1. Utterance	✓	✓	✓	✓
	2. Silence	✓	✓	✓	✓
Total duration (seconds)	3. Utterance	✓	✓	✓	✓
	4. Silence	✓	✓	✓	✓

Note: TPD = Thai Picture Description task; Imm = immediate recall task; Del = delayed recall task; SIT = Semi-structured Interview for Thai.

Table 4 Transformed parameters of temporal variables

Variables	Description
1. Locution, Total length of speech duration (seconds)	Total duration of utterance + Total duration of silence
2. Utterance proportion, Phonation rate (%)	$\frac{\text{Total duration of utterance}}{\text{Locution}} \text{ x } 100$
3. Silence proportion (%)	$\frac{\text{Total duration of silence}}{\text{Locution}} \text{ x } 100$
4. Pause rate (number per seconds)	$\frac{\text{Total number of silence segments}}{\text{Locution}}$
5. Hesitation rate, Silence-to-utterance ratio	$\frac{\text{Total number of silence segments}}{\text{Total number of utterance segments}}$

The primary sets of variables were then transformed into five temporal variables per task. Table 4 provides the formulas of each transformed variable. Five transformed variables were calculated based on the primary parameters of each task. Consequently, each task of the spontaneous speech tests consists of nine variables (four primary variable in Table 3 and five transformed variables in Table 4), that is, (1) number of utterance segments; (2) number of silent segments; (3) total duration of utterance; (4) total duration of silences; (5) locution (total length of speech duration); (6) utterance proportion (phonation rate); (7) silence proportion; (8) pause rate; and (9) hesitation rate (silence-to-utterance ratio).

3.2.6 Statistic analysis

The IBM SPSS version 26 for Windows was used to analyze all data. The list below is the set of analyses for this study.

3.2.6.1 Descriptive statistics, mean, median, standard deviation, and interquartile range.

3.2.6.2 Normality test by Shapiro-Wilk, the most powerful normality test (Razali & Wah, 2011), skewness, and kurtosis.

3.2.6.3 Pearson's correlation for continuous variables and Spearman's rho for categorical variables. Correlation coefficients were interpreted based on research in psychology areas. The Pearson's and Spearman's correlation coefficients (*r*) are classified in five levels, including perfect (1), strong (0.7 – 0.9), moderate (0.4 – 0.6), weak (0.1 – 0.3), and no relationship (0) (Akoglu, 2018).

3.2.6.4 Mean comparison.

- Kruskal-Wallis for non-normal distributed variables, i.e., age and years of education.
- Chi-square tests for categorical variables, i.e., sex, underlying diseases.
- One-way analysis of variance (ANOVA) for continuous variables, i.e., MoCA and cognitive tasks.
- Post-hoc analysis by the Turkey HSD test, which is less conservative but can exercise control over the type I errors (Lee & Lee, 2018). Cohen's *d* effect sizes were also reported with 0.20, 0.50, and 0.80 as small, medium, and large (Lakens, 2013), respectively.

3.2.6.5 Multivariate analysis of variance of the content variables (MANOVA) for examining the effect of dementia pathological stages on a group of selected variables.

3.2.6.6 Multivariate Discriminant Analysis (MDA) for determining the most functional variables for differentiating older adults in each pathological stage.

3.2.6.7 Receiver Operating Characteristic Curve (ROC curve) and Area Under Curve (AUC) for evaluating the discriminant ability. The AUC ranges from 0 to 1, and .50 means no discriminant ability. This study expected these three ranges of AUC, i.e., .70 – .80 (acceptable), .80 – .90 (excellent), and > .90 (outstanding) (Mandrekar, 2010). The optimal cut-off point was selected by maximizing sensitivity and specificity based on the Youden Index (YI), which was calculated from Sensitivity + Specificity – 1 (Hajian-Tilaki, 2013).

3.2.7 Protection of human subjects

The data collection procedures and protection of the human rights of the participants were approved by the Institutional Review Board [IRB] for Graduate Studies, Burapha University. Project number G-HS097/2564(C1) was approved on 3rd February 2022. After receiving the IRB approval, the researcher contacted the officer of Burapha University Hospital and Chonburi Hospital to discuss the purposes of the study and asked for permission to collect data in those settings. Participants were free to choose whether to participate, and data confidentiality was guaranteed. The researcher contacted those who indicated their willingness to take part in the study and secured their informed consent. After receiving their informed consent, participants were free to leave at any time. Individual replies were kept private during the presentation of the findings.

CHAPTER 4
RESULTS AND PREDICTIONS

This study aimed to differentiate people with dementia from cognitively intact people by means of the proposed spontaneous speech tasks administered to people with Alzheimer's disease (AD), to people with mild cognitive impairment (MCI), and to cognitively intact people (HC). The developed tasks were used to elicit participants' spontaneous speech and were employed as a neuropsychological assessment by means of the analyses of the acoustic feature of the produced oral texts. In the assessment session, the voice of participants was recorded with their consent. The voice was analyzed to represent the different patterns of speech among three groups of participants which are Two phases of this research were composed of the development of spontaneous speech tasks and the comparison of speech patterns.

4.1 Phase 1 Developing speech tasks for Thai patients

Three spontaneous speech tasks which were developed in this study were composed of the essential components from the standardized and accepted measurement. The consequence of developing tasks was reported in the following section.

4.1.1 Synthesized protocol of spontaneous speech tasks

Before the stimuli were created, the core components of each task were reviewed. The synthesis of the accepted and used measurement for spontaneous speech analysis for detecting dementia grounded the prototype of each task.

4.1.1.1 Picture description

This study selected a single composite picture rather than a sequential pictures protocol because the composite picture was mostly employed in speech analysis studies of neurodegenerative diseases (Boschi et al., 2017). According to the synthesis of the previous research, the picture description task demonstrates three common and essential components, which are in the following paragraphs.

1) The stimulus card depicts a situation with people, places, and actions. Elements in the picture should be highly familiar characteristics or daily life events in the context and culture of participants (Giles et al., 1996; Marshall &

Wright, 2007). Most situational pictures are in black and white (Mueller, Hermann, et al., 2018).

2) The administration should be simple and have no guidance. The examinee is asked to 'narrate' the picture rather than describe it (Bayles et al., 1999). The instruction can be, 'Tell me what is going on/happening in this picture?' (Giles et al., 1996; Marshall & Wright, 2007). There is no limit of time; the duration can be 2 – 5 minutes in the healthy control group (Boschi et al., 2017). The examiner may encourage or ask for more narration when the participant pauses or show a sign of ending the story (Tomoeda et al., 1996).

3) Scoring criteria can be defined by discourse units (pragmatic features), duration of narration, and information units (Giles et al., 1996; Nagarachinda et al., 2020; Tomoeda et al., 1996). This study focuses on acoustic features, which include interval parameters of speech. For language fluency or proficiency scoring, counting of information units was appropriated to represent cognitive abilities.

4.1.1.2 Story retelling

As per this study, the immediate recall administration adopted a protocol of a single short story without repeatedly reading. The administration, short story, and scoring of this task have at least four common characteristics, which were used to be the prototype of the proposed task in this study.

1) Recalling is needed to perform twice, which is immediately after listening to the story and the second round after the delayed period. The first round is called an immediate recall and a delayed recall for the latter round (National Institute on Aging, 2006; Wechsler, 2009). A delayed period is varied in a different test, e.g., four hours in the Story Memory Test of Heaton, Grant, and Matthews (Lezak et al., 2012), 30 min in the logical memory (Hodges, 2007), 20 minutes in the Learning and Memory Battery (LAMB) (Lezak et al., 2012), 15 min in the Mini-Mental State Examination 2nd edition, Expanded Version (Song et al., 2019), and 10 minutes in the Babcock protocol (Khan, 1986; Lezak et al., 2012).

2) After finishing an immediate recall, an examinee is prompted to memorize the story for the second round of retelling (Hodges, 2007; Leal et al., 2021; Wechsler, 2009).

3) The brief story is found to contain 60 – 70 words (Wechsler, 2009). The short stories in WMS and the Arizona Battery for Communication Disorders of Dementia (ABCD) are composed of three sentences, while the original Babcock story was four sentences (Hodges, 2007; Prud'hommeaux et al., 2011).

4) Scores are given to the correct recalled element. The scorable units are divided in the short story, which can be a single word or a group of words that contain an important text (Khan, 1986; National Institute on Aging, 2006). The total scores vary between 21 – 25 measuring units (Hodges, 2007; Prud'hommeaux et al., 2011).

4.1.1.3 Semi-structured interview

This task aims to imitate conversation in daily life. Spontaneous speech is the target of three developing tasks, especially this task. The stimuli of this task are open-ended questions related to the topics and inquiries in between the conversation.

1) Spontaneous speech is daily life communication that carries the flow of thought, not an expression of a single word, phrase, or even one sentence (Mueller, Hermann, et al., 2018; Pulido et al., 2020).

2) Speech impairment in people with MCI and Alzheimer’s disease reflects cognitive declination, especially in episodic memory and working memory. Thus, the task of imitating daily conversation could elicit speech impairment (Beltrami et al., 2018; Gomez & White, 2006; Tröger et al., 2018).

4.1.1.4 The phenomena in phonological loop embedded in the spontaneous speech tasks.

Spontaneous speech tasks rely on several components of memory. This study adopted the multi-component model of working memory developed by Baddeley (2000). This model is the prominent fundamental of a number of theoretical models of verbal working memory (Acheson & MacDonald, 2009). The model explains the process in LTM, STM, and WM, especially the functional specifications of language and verbal memory (Baddeley, 2003). The three proposed tasks and administration procedure integrated the phenomena in the phonological loop, which is evidence of verbal working memory.

The phenomena were embedded in the presentation of specific stimuli and the design of the assessment process. The first implementation was from the phonological similarity effect. Similar phonemic words in Thai were added in the short story of the story recall. As well as the Thai polysyllable words were included in the story to highlight the word-length effect (see Table 5). Articulatory suppression naturally occurs in the delay period of the story recall, where the participants were asked to perform filler tests in the form of verbal tasks. Regarding the phenomenon of transferring information between codes, the stimuli were presented in both visual (picture description) and auditory modalities (story recall).

4.1.2 Survey daily life events and assess familiarity with the stimuli.

This study conducted a focus group with older adults to gather information about common activities in their daily life. In the same session, the focus group participants were asked to rate the sample pictures for picture description. Participants voluntarily attended the focus group through snowball advertisement of the older adult who lived in Saen Suk sub-district, Muang district, Chonburi province. Enrollment criteria tried to replicate the characteristics of the participants in the task evaluation process (phase 2). The living area was one of the criteria. Apart from living in Saen Suk sub-district, they needed to be older than 55 years old and were able to attend a 2-hour session.

There were 13 participants (3 males and ten females) between the ages of 56 – 76. Before starting the session, they were asked to carefully read and provide informed consent in printed form. Focus group questions consisted of the main five sections with inquiries regarding the response of the participants. Questions in each section are shown below.

1) What is in your mind when speaking about a familiar place?
    - What activities did you do in that place?
    - Who do you normally go with?
    - How often did you go?
2) What were your daily activities when you were young?
    - What did you do?
    - Who was involved in those activities?
    - What instruments did you use?

3) What festival or cultural event which you like most?

- What activities did you do at that event?
- Where did the event take place?
- Who attended the event with you?
- What objects did specifically need in the events

4) The participants were given four pictures of different styles and stories and asked for their opinions on those pictures.

- Please consider the sketch line of the four pictures and tell me which picture you are familiar with.
- Which pictures are easy on your eyes?

5) Which is your favorite picture and why?

Questions number 1 – 3 were asked to receive individual answers, and then the most common two answers would be inquired for more information. The answers were selected and used to draft the picture stimulus. Question number 4 and 5 were related to their familiarity with the sample pictures and their opinions on the characteristics of the illusion, not the meaning of the pictures. The information was provided for a commercial artist with the details of stimuli which elicit spontaneous speech (Marshall & Wright, 2007). Features of the stimulus picture are described in the next section.

4.1.3 Created the stimuli and administration

Each task was developed based on essentially synthesized components and information of the focus group. The information from the focus group mostly appeared in the picture stimulus and the topics of the semi-structured interview but not in the story stimulus.

4.1.3.1 Thai Picture Description Task (TPD)

According to the most familiar place and activity of the focus group, the house was frequently mentioned with dated and antique details. For instance, cleaning equipment and actions at their age were a broom and damp towel, kneeling on the floor to wipe, and using a big plastic or tin tub to wash dishes. The researcher informed the artist of the details needed to be in the pictures, including the storyline and appearance of the picture. The storyline was designed to be a man and woman doing chores; the man hurried to wipe the floor inside the house while the woman was outside cleaning the dishes by using big tubs. The salient points are depicted inside and outside the house. The first point was the cat was trying to drink the water from the tub, so the woman was telling the cat to go away, and she could not notice that the water was overflowing from the tub. The second point was a strong wind blowing through the windows bringing dust and leaves into the house, where the man had already wiped the floor. The old-fashioned equipment was presented in the picture, such as a television in a box shape with four legs, a broom, and a wax pot with coconut coir. For the appearance of the stimulus, the components from the synthesized process were given, including black and white, clear and bright, and A4 size. Picture stimulus is shown in Figure 18.

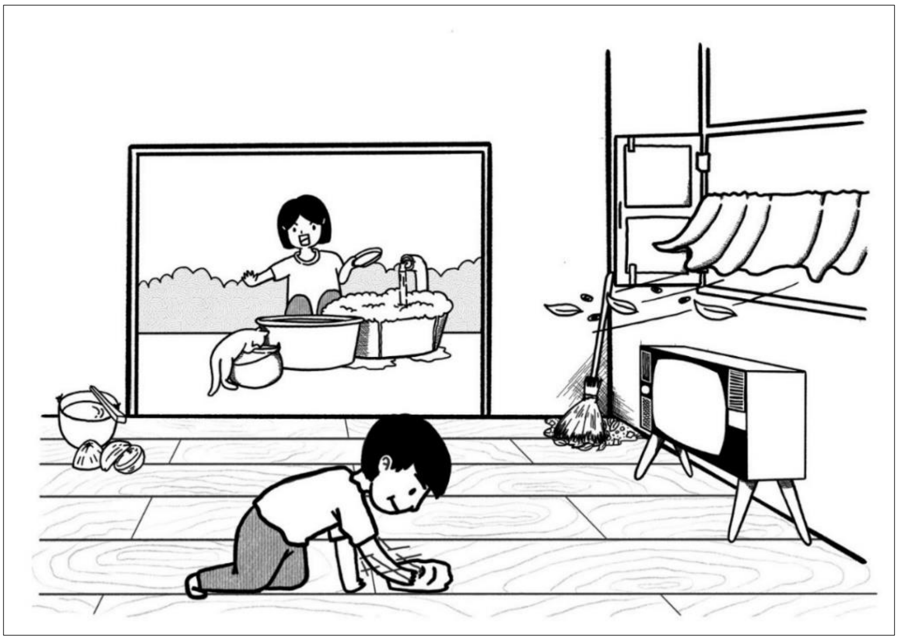


Figure 18 Picture stimulus of TPD

The administration was applied by Giles et al. (1996) and Marshall and Wright (2007) while scoring criteria based on Nicholas and Brookshire (1993). The administration of TPD allowed the participant to spontaneously generate the story without interruption but with very less prompt at the beginning. Participants would be told to narrate the story out of the illustrated stimulus rather than describe it (Bayles et al., 1999). Physical encouragement was suggested for an examiner. When participants gave a sign of ending the task, they should be asked to generate more before confirming the end of the story (Tomoeda et al., 1996). The verbal performances would be recorded and transcribed for scoring. Transcribe notes of all participants would be segmented into content units by Thai segmentation software developed by Aroonmanakun (2002). Then each word was examined and scored based on the adapted scoring criteria. Scores were derived from the correct information units (CIU). The details of the final version used in this study were presented in the instruments section (see Section i. 3.2.3.3).

4.1.3.2 Thai Story Recall Task (TSR)

Elements composing the story were derived from synthesized components, standardized neuropsychological assessments, and the working memory model. The storyline of this stimulus was created based on the prepared words and objects. The first source of words was the Thai version of the Boston Naming Test (Aniwattanapong et al., 2018). Two words from the medium frequency category were selected, including /wong-wia:n/ 'compass' or 'round about' and /lú:k dɔ̀:k/ 'dart'. The approach for creating words was the model of working memory presented by Baddeley (2000). The researcher selected Thai words that sound similar in phonetic consonants and vowels to demonstrate the phonological similarity effect of the phonological loop. Since working memory is affected by Alzheimer's disease, those targeted words were aimed to magnify the ability or deficiency of working memory. Thai names of people, provinces and city were selected based on similar sounding in each component which are described in Table 5.

Eventually, the brief story and administration were invented. The story stimulus was composed of 21 scorable units and 70 words (see Section ii. 3.2.3.3). The Thai Story Recall Task was designed to be a single-read protocol. The story was read at the beginning of the administration. The first recall was immediately after the examiner had finished reading. When participants had finished an immediate recall, they were then asked to make five factual inquiries. It should be noted that before the content validity process, there were ten questions. The questions were reduced to five questions for the final version of the test (see Appendix B).

Table 5 Element composing in the short story based on the phonological loop

Word	First syllable	First consonant	Vowel	Final consonant	Tone
• **phong** • **phon**	none	/ph/ พ	/o/ โอะ	- /ng/ ง - /n/ น (both are nasal sounds)	mid
• **u-bon** • **u-dɔ:n**	/u/ อุ	- /b/ บ - /d/ ด	- /o/ โอะ - /ɔ:/ ออ	/n/ น	mid
• **u-bon** • **muea:ng phon**	- /u/ อุ - /muea:ng/ เมือง	- /b/ บ - /ph/ พ	/o/ โอะ	/n/ น	mid
• **u-dɔ:n** • **bɔ:n**	- /u/ อุ - none	- /d/ ด - /b/ บ	/ɔ:/ ออ	/n/ น	mid

Moreover, the expected interval of the delayed period was 20 minutes; the professionals suggested decreasing the proposed interval (see further details in Section 4.1.4). The delayed interval is consequently expected to be 7 – 10 minutes. This adjustment resulted in the selection of the filler tests in a delayed period. Digit span tests were moved from the set of filler tests to the end of the delayed recall test. Scoring criteria followed the standardized story recall which one score was given to a correctly recalled unit or item.

4.1.3.3 Semi-structured Interview for Thai (SIT)

In this initial development of Thai spontaneous speech tasks, the semi-structured interview was selected to balance the pros and cons of structured and unstructured interviews. An unstructured interview can elicit everyday conversation and is considered the most connected spoken form, yet a novice or quiet examiner may struggle with how to generate questions. While a structured interview can ensure an equivalent protocol among each individual session, it limits the natural responses of a participant. The theme of the interview was based on personal experiences in adolescence and early adulthood. Three topics were prepared for the content validity assessment; eventually, the topic of tourist attraction was only selected. Questions related to the topic were prepared for an examiner to follow (see Section iii. 3.2.3). The examiner can select the question from the list to ask regarding the presenting contents that participants were talking about.

A score was not given for SIT. Unlike the AMI, the responses are scored for accuracy in the personal semantic schedule and descriptive richness in the autobiographical incident schedule. The spoken responses of AMI are scored in terms of the richness and specificity of the descriptive information (Hodges, 2007). The SIT is identical to the autobiographical incident part of the AMI, as both ask for the specific incident from a certain life period. On the other hand, the SIT analyzes the linguistic abilities of the participants based on an acoustic approach and no scoring.

4.1.3.4 Verbal fluency

In Thailand, Both letter and category fluency tasks were recommended as a screening test in the history-taking process and a specific cognitive domain test for differential diagnosis (Neurological Institute of Thailand, 2014). In the Thai version, the letter 'Kor' (ก) was suggested for phonemic fluency, and the animal was used in the category fluency (Charernboon, 2018). This study assessed both phonemic and semantic abilities. The conventional administration was applied to this study by asking the participants to generate the words regarding the instruction (letter or category) within 60 seconds. An instruction was modified from the latest version of MoCA (Nasreddine, 2017). Repeated words are not scored, so repetition is needed to clarify and present Thai sample words. The used protocol was described in the instruments Section i. 3.2.3.2.

4.1.3.5 Digit span

This test is considered a specific cognitive domain assessment that reflects the ability to allocate attention (Muangpaisan et al., 2010). There are two versions of the digit span task, which are the forward and backward versions. Straightforwardly, the forward version asks the participants to pronounce a list of listened digits in the same order. On the other hand, the participants need to reproduce the reverse order from a given list. Stimulus spans are derived from Monaco et al. (2013). The longest span is nine for the forward version and eight for the backward version. Administration and scoring implementing in this study were adopted from the same study by Monaco et al. (2013). While instruction for forward and backward tasks was applied from the latest version of MoCA (Nasreddine, 2017). The developed protocol was described in the instruments Section ii. 3.2.3.2.

4.1.4 Evaluated psychometric index

Content validity was rated by professionals in multidisciplinary; the proposed instruments received acceptable scores. Three professionals were invited to investigate the three developed stimuli and administration protocols. The first professional was a psychologist who had worked with neurodegenerative patients in the neurological ward for more than ten years of experience. The second rater was the Assistant Professor of Nurse with a specialization in community and aging and graduated with Ph.D. in Research and Statistics in Cognitive Science. The third expert was an Assistant Professor of Educational Science with a specialization in measurement and a lecturer of psychological testing. The topic of content validity index (CVI) was composed of five sections which are (1) consistency of the protocol with the theory and approaches, (2) characteristics of the stimuli, (3) appropriateness of the administration and instruction, (4) reasonability of the scoring criteria, and (5) the overall and combination of three tests. Overall CVI scored one point as well as the point on each tool which was an acceptable score for three experts' evaluations (Yusoff, 2019) (see Appendix F).

The professionals mostly suggested revising the story recall task and picture description task. The administration of the story recall was criticized that the delayed period was too long, and ten content-inquiry questions were overloaded. The final administration of the story recall test was 7 – 10 minutes delayed, and only five factual inquiries remained. In addition, the suggestions for the story recall emphasized clearly pronouncing the words and avoiding ambiguous terms such as ‘Chon’ (ชลฯ),

which was replaced with 'Chonburi' (ชลบุรี), a full term of 'Chon' province. Because /chon/ in Thai can refer to the province and crashing. The original short story was then revised to be clearer and removed some excessive words to make a compact paragraph. The final version of the short story is about a man who is going to give a toy to his nephew; the earthquake causes a delay in his train during his trip. The final version of the short story is shown in Figure 9. For the picture description, the concerns were on an unlimited telling time, and the participants might describe the picture instead of telling the story. In this initial study, the research wanted to collect the telling time of Thai older adults pertaining to the picture stimulus and explore the mean, so the telling time was not indicated. In order to prevent describing the picture, a prompt was suggested to give when the participants began to describe the picture. Besides, scoring criteria only considered the CIUs, which were from the telling, not describing part. Lastly, the professionals suggested selecting only one topic to be administered and shown to an examiner. The final version of three tests that were implemented in the data-collecting process is presented in Appendix A and B. The sequence of the proposed speech assessments can be found in Figure 12.

4.2 Data analysis

The findings are presented in four sections, including characteristics of the participants, analysis of the content variables, analysis of the acoustic features, and all variables as predictors. The *p*-value chosen for significant effects was $p < .05$ for all analysis.

4.2.1 Characteristics of the participants

The initial sample consisted of 102 participants who were native Thai-speaking, aged between 55 – 80 years, with normal hearing and sight. After the screening, three of them were shown not to meet the inclusion criteria, as they had epilepsy, Parkinson's disease, and a history of drug abuse. Moreover, one participant was unable to finish all the programmed tasks. The total sample consequently was 98. The demographic characteristics of the final sample of participants are presented in Table 6.

The age range of the participants was 56 – 79 years. The majority of participants were female, as male participants were approximately one-fifth of the sample. The mean years of education (YoE) were ten years, and most of them acquired a primary level of education. Their occupations varied; the top two occupations were merchandisers and professionals such as a teacher and nurse, respectively. All participants were Thai-native speakers; most reported that they could not communicate in a second language. Of those who could, 16% had English as a second language and 9.18% had Chinese. Given the period the data were collected, we investigated about the Coronavirus disease (COVID-19). About one-tenth reported that they had been infected and had recovered before the experiment day. Other reported diseases were Hypertension (approximately one-third of the sample), a quarter sample with Hyperlipidemias (approximately one-fourth of the sample), and one-fifth of the sample with Diabetes Mellitus (approximately one-fifth of the sample). The results of the PHQ-9 test showed that approximately 23% of the elderly participants had mild to severe depression symptoms. However, they had never been diagnosed or had a history of major psychiatric illness. Since there were 13 missing data in the ADL questionnaire, only the complete data were analyzed and are reported here. No participant demonstrated a dependence on either ADL or iADL. The hearing ability was examined by the self-report questionnaires and confirmed by an inquiries interview. Although 14% of the participants reported that they had a mild problem with hearing, they were able to hear the examiners at normal speaking volume.

Table 6 Characteristics of the participants

Characteristic	HC (n = 32)		MCI (n = 32)		AD (n = 34)		All participants (n = 98)	
	n	%	n	%	n	%	n	%
Gender								
Male	7	21.88	6	18.75	8	23.53	21	21.43
Female	25	78.13	26	81.25	26	76.47	77	78.57
Education								
No schooling	0	0	0	0	1	2.94	1	1.0
Primary	6	18.75	10	31.25	19	55.88	35	35.7

Characteristic	HC (n = 32)		MCI (n = 32)		AD (n = 34)		All participants (n = 98)	
	n	%	n	%	n	%	n	%
Lower secondary	4	12.50	2	6.25	6	17.65	12	12.2
Upper secondary	5	15.63	5	15.63	0	0	10	10.2
Higher level /diploma	1	3.13	1	3.13	1	2.94	3	3.1
Bachelor's degree	15	46.88	9	28.13	4	11.76	28	28.6
Higher than a bachelor's degree	1	3.13	5	15.63	3	8.82	9	9.2
<u>Occupation</u>								
Armed forces	0	0	2	6.25	0	0	2	2.04
Professionals e.g., teacher	10	31.25	7	21.88	3	8.82	20	20.41
Clerical support workers	1	3.13	2	6.25	1	2.94	4	4.08
Skilled agricultural, forestry, and fishery workers	2	6.25	1	3.13	3	8.82	6	6.12
Plant and machine operators and assemblers	0	0	1	3.13	0	0	1	1.02
Elementary occupations	3	9.38	6	18.75	6	17.65	15	15.31
Managers	3	9.38	2	6.25	2	5.88	7	7.14
Technicians and associate professionals	2	6.25	0	0	1	2.94	3	3.06

Characteristic	HC (n = 32)		MCI (n = 32)		AD (n = 34)		All participants (n = 98)	
	n	%	n	%	n	%	n	%
Services and sales workers	6	18.75	5	15.63	16	47.06	27	27.55
Craft and related trades workers	4	12.50	5	15.63	0	0	9	9.18
Unemplyed	1	3.13	1	3.13	2	5.88	4	4.08
Second language								
Chinese	3	9.38	4	12.50	2	5.88	9	9.18
English	7	21.88	7	21.88	2	5.88	16	16.33
Others	0	0	1	3.13	2	5.88	3	3.06
No	22	68.75	20	62.50	28	82.35	70	71.43
Health status								
Hyperlipidemias	5	15.63	12	37.50	8	23.53	25	25.51
Hypertension	8	25	7	21.88	17	50	32	32.65
Diabetes Mellitus	4	12.50	2	6.25	15	44.12	21	21.43
COVID-19	3	9.38	3	9.38	4	11.76	10	10.20
Depression								
No symptom	27	84.38	25	78.13	23	67.65	75	76.53
Mild	5	15.63	7	21.88	9	26.47	21	21.43
Moderate	0	0	0	0	1	2.94	1	1.02
Severe	0	0	0	0	1	2.94	1	1.02
ADL								
Missing data	7	21.90	4	12.50	2	5.88	13	13.3
No dependence	25	100.00	28	100.00	32	100.00	85	100.00
iADL								
No dependence	32	100.00	32	100.00	34	100.0	98	100.00
Hearing								
No problem	15	46.88	18	56.25	18	52.94	51	52.04
Mild problems	13	40.63	9	28.13	11	32.35	33	33.67

Characteristic	HC (n = 32)		MCI (n = 32)		AD (n = 34)		All participants (n = 98)	
	n	%	n	%	n	%	n	%
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age	63.84	4.15	65.97	4.88	67.00	6.13	65.63	5.26
YoE	12.28	4.58	11.66	5.27	8.09	5.01	10.62	5.26
PHQ-9	3.13	2.32	3.78	3.31	4.94	4.83	3.97	3.71
ADL	19.80	0.58	19.71	0.60	19.41	0.95	19.62	0.76
iADL	11.97	0.18	11.72	0.52	11.65	0.73	11.78	0.55
Hearing	6.03	3.63	5.9 7	4.93	5.41	4.38	5.80	4.31

Note: ADL = Barthel's Activities of Daily Living Questionnaire; i-ADL = Instrumental Activities of Daily Living Scale; Hearing = Five-minute Hearing Test; YoE = Years of education; PHQ-9 = Patient Health Questionnaire.

Neuropsychological characteristics were represented by five tests, namely the MoCA, LF, CF, FDS, and BDS. The means and standard deviations of the three groups and all participants are presented in Table 7. The standard deviation of the MoCA in AD group was higher than the other two groups, accordingly the range of scores was the widest. Since there is no minimal limitation of the MoCA for participant recruitment. In general, the patterning of the mean scores was quite systematic, with better performance by HC, followed by MCI, and followed by AD. The only exception was the BDS scores of MCI (*M* = 3.41, *SD* = 1.62), that were slightly lower than the mean scores of AD (*M* = 3.44, *SD* = 1.83).

The opposite trend of mean scores was found in the repeated words of CF and LF, which AD produced the highest repeated words than MCI and HC in both tests of verbal fluency. Considering the number of responses of verbal fluency, the mean score of CF (*M* = 17.72) is higher than LF (*M* = 10.87). For digit span tests, most participants could remember the digits of FDS (Mode = 6) more than BDS (Mode = 4). The different versions of verbal fluency and digit span tests rely on particular cognitive functions. Notably, there was missing data in both versions of the digit span. The total missing cases were 25 people (25.51% of all participants), with eight people in HC (25.00% of all HC), ten people in MCI (31.25% of all MCI), and seven people in AD (20.59% of all AD). Due to the technical problem in administering ang audio recording, the data of the first 25 participants could not be included in the analysis.

Table 7 Descriptive statistics of the neuropsychological tests

	HC					MCI					AD					All participants						
	Range	Mode	Median	*M*	*SD*	Range	Mode	Median	*M*	*SD*	Range	Mode	Median	*M*	*SD*	Range	Mode	Median	*M*	*SD*	*P*. 25	*P*. 75
MoCA	25-29	25	26	26.28	1.30	22-24	22	23	22.94	0.95	11-21	20	20	18.50	2.55	11-29	22	22.5	22.49	3.66	20	25
LF	3-31	12	12.00	13.50	6.91	2-23	8	10.00	11.25	4.89	1-25	10	8.00	8.03	5.04	1-31	9	10.00	10.87	6.06	7	14
LF Rep	0-4	0	0.00	0.59	0.87	0-7	0	1.00	1.03	1.43	0-8	0	0.00	1.35	2.13	0-8	0	0.50	1.00	1.59	0	1
CF	7-37	21	20.00	19.63	5.56	10-29	20	20.00	18.66	4.43	3-26	7	15.00	15.06	5.74	3-37	20	18.00	17.72	5.60	14	21
CF Rep	0-6	0	1.00	1.84	1.90	0-9	0	1.00	1.56	2.03	0-10	1	1.50	2.21	2.48	0-10	0	1.00	1.88	2.16	0	2
FDS	5-9	7	7.00	6.88	1.15	3-8	6	6.00	6.05	1.13	4-9	5	5.00	5.74	1.13	3-9	6	6.00	6.21	1.22	5	7
BDS	1-11	5	5.00	5.29	1.27	0-6	4	4.00	3.41	1.62	0-7	3	3.00	3.44	1.83	0-8	4	4.00	4.04	1.81	3	5

Note: LF = letter fluency; LF Rep = repeated words in letter fluency; CF = category fluency; CF Rep = repeated words in category fluency; FDS = forward digit span; BDS = backward digit span.

4.2.1.1 Normal distribution

The Shapiro-Wilk test was performed to examine a normal distribution of continuous variables (Blanca et al., 2017; Field, 2018, p. 249). According to demographic variables (See Table 8), the results showed that all six variables were non-normally distributed.

Table 8 Normality test by Shapiro-Wilk test of demographic data

Variables	Skewness	Kurtosis	*H*	*p*
Age	0.58	-0.12	*0.96*	*<.01*
Years of education	0.02	-1.65	*0.84*	*<.001*
PHQ-9	1.62	3.95	*0.86*	*<.001*
ADL	-2.30	5.96	*0.56*	*<.001*
iADL	-2.78	8.40	*0.47*	*<.001*
Five-minutes hearing test	1.13	2.01	*0.92*	*<.001*

As for the neuropsychological tests (see Table 9), four variables, except CF, were significantly different from the normal distribution (Mishra et al., 2019).

Table 9 Normality test by Shapiro-Wilk test of the neuropsychological tests

Variables	Skewness	Kurtosis	*H*	*p*
MoCA	-0.62	0.30	*0.96*	*<.01*
LF	1.14	2.03	*0.92*	*<.001*
CF	0.02	0.70	0.98	.15
FDS	0.20	-0.05	*0.93*	*<.01*
BDS	-0.53	0.62	*0.91*	*<.001*

Note: LF = letter fluency; CF = category fluency; FDS = forward digit span; BDS = backward digit span.

4.2.1.2 Equivalent groups

A chi-square test was performed to compare ten categorical variables among three groups of participants. The statistical and significant tests are presented in Table 10. Three characteristics of the sample revealed a significant difference. The first variable was education level (χ^2 = 25.42, *df* = 12, *p* <.05). Most of the AD group graduated in primary education, while a majority of the HC group acquired a bachelor's degree. The other two variables were hypertension (χ^2 = 7.20, *df* = 2, *p* <.05) and diabetes mellitus (χ^2 = 16.29, *df* = 2, *p* <.001), which were more frequent in the AD group than in HC and MCI groups.

Table 10 Chi-square test for categorical variables of demographic data

Variables	*df*	χ^2	*p*
Sex	2	0.23	.89
Education level	12	*25.42*	*<.05*
Occupation	20	27.82	.11
Hyperlipidemia	2	4.14	.13
Hypertension	2	7.20	<.05
Diabetes Mellitus	2	*16.29*	*<.001*
Second language	6	7.09	.31
Covid	2	0.14	.93
Depression	6	5.29	.51
Hearing	4	1.16	.88

Due to non-normality distribution in six variables of demographic data, a ranking test is appropriate for these variables. The Kruskal – Wallis *H* test was performed to compare the independent variables between the three groups as a nonparametric technique or ranking test (Privitera, 2015, p. 618). The *p*-value chosen for significant effects was $p < .05$. The results are presented in Table 11. Mean years of age between the three groups were not different, with a *p*-value at the critical point. The categorical variable "YoE" differed significantly among the three groups (*H* = 9.58, *df* = 2, *p* <.01). Specifically, the mean number of years of education for the HC group is

12.28 years ($SD = 4.58$), with 46.88 % of them having a bachelor's degree; for the MCI is 11.66 ($SD = 5.27$), with a few having a bachelor's degree; for the AD group is 8.09 years ($SD = 5.01$). According to the ranking test, iADL showed a significant difference among the three groups ($H = 6.71$, $df = 2$, $p < .05$), with a mean rank of 56.53 for HC, 46.00 for MCI, and 46.18 for AD. The difference was due to the higher iADL scores of the HC group with respect to the two other groups. The mean rank of the three screening tests was not different between the three groups of participants, namely, PHQ-9, ADL, and the five-minute hearing test.

Table 11 Ranking test by Kruskal – Wallis *H* test of demographic data

Variables	*df*	*H*	*p*
Age	2	6.01	.05
Years of education	2	*9.58*	*<.01*
PHQ-9	2	1.16	.56
ADL	2	4.86	.09
iADL	2	*6.71*	*<.05*
Five-minutes hearing test	2	0.98	.61

With regard to the normality assumption of ANOVA, F-test was proved to be robust with non-normal distributed and unequal variance data (Blanca et al., 2017). Hence, one-way ANOVA was used to determine the mean differences in neuropsychological scores between the three groups (see Table 12). The F tests showed that the mean scores of the three groups in each neuropsychological test were significant differences, except for the repeated words of both verbal fluency tests. Turkey's HSD test found that there was no statistically significant difference in mean scores between HC and MCI in LF ($p = .26$), and CF ($p = .74$). In addition, the post-hoc test by Turkey's HSD test revealed no statistically significant difference in mean scores between MCI and AD in LF ($p = .06$), FDS ($p = .62$), and BDS ($p = .99$). There were significant differences between HC and AD in LF ($p < .001$), CF ($p < .01$), FDS ($p < .05$), and BDS ($p < .001$).

Table 12 ANOVA Neuropsychological scores of the participants

Variables	*df*	*F*	*p*	η_p^2
MoCA	2	*142.18*	< *.001*	.77
LF	2	*7.77*	< *.01*	.14
LF Rep	2	1.92	.15	.04
CF	2	*6.89*	< *.01*	.13
CF Rep	2	0.74	.48	.02
FDS	2	*6.61*	< *.01*	.16
BDS	2	*10.94*	< *.001*	.24

Note: LF = letter fluency; LF Rep = repeated words in letter fluency; CF = category fluency; CF Rep = repeated words in category fluency; FDS = forward digit span; BDS = backward digit span.

4.2.2 Analysis of the content variables

For two of the spontaneous speech tasks, namely the TPD and TSR some dependent variables were identified. This study did not quantify the semi-structured interview task into content variable, but for acoustic features which are presented in Section 4.2.3. The following section presents four analyses pertaining to the content variables, including descriptive data, means comparison, predictive model, and discriminant ability.

4.2.2.1 Characteristics of the content variables among three groups

Table 13 presents descriptive information on the content variables. The variable of TPD consisted of the amount of correct information units (CIU) of the story reported by participants. The three variables of the TSR were the amount of correct information reported in the immediate recall (Imm), the delayed recall (Del), and the number of correct answers to the five-factual inquires (Ans). For each, the range, the mean, the median, and mode was computed for each of the three groups. All three statistics of AD tended to be the lowest scores, followed by MCI and HC, respectively. The mean CIUs of HC (*M* = 60.44) was approximately double of AD (*M* = 32.03), and MCI (*M* = 56.28) was close to HC. The mean scores of both recall tasks of TSR showed that HC (Imm *M* = 5.28, and Del *M* = 5.97) scored the highest, and MCI's scores (Imm *M* = 4.47, and Del *M* = 4.28) were close to AD (Imm *M* = 4.24, and Del *M* = 3.64). However, the groups comparison was examined carefully by the

Turkey's HSD afterward. The different patterning of mean scores was found in the scores of correct answers of TSR which MCI's scores (*M* = 1.34) were slightly lower than AD (*M* = 1.56). Interestingly, the CIU of TPD has a very wide range of scores since it was not limited by the number of items (as in recall tasks) and response time. The minimum score of CIU in AD is zero. This was a result of incorrect narration due to one of the participants being unable to tell the proper story or their responses did not meet the CIU criteria. The range of both recall tasks was 0 - 11 (Imm) and 0 - 10 (Del), the upper limit being approximately half of the total items (21 items) and was quite similar in all three groups.

Pearson's correlation in Table 14 revealed that the content variables had a significant positive relationship with years of education but not with age. As for CIU, the statistically significant correlations were found at moderate positive level with years of education ($r = .46, p < .01$), and CF ($r = .40, p < .01$), and weak positive correlation with MoCA ($r= .37, p < .01$), with FDS ($r= .28, p < .05$), with LF ($r = .27, p < .01$), and with Del ($r = .20, p < .05$). According to TSR tasks, Imm were found to have statistically significant correlation at strong positive level with Del ($r= .75, p < .01$), and moderate positive level with Ans ($r= .63, p < .01$). Furthermore, Imm had weak positive correlation with the MoCA ($r= .28, p < .01$), and CF ($r= .22, p < .05$). The significant correlation among TSR tasks was also found between Del and Ans with strong positive level ($r= .55, p < .01$). Del additionally had significant correlation with MoCA at moderate positive level ($r = .43, p < .01$), and weak positive correlation with BDS ($r = .33, p < .01$), with CF ($r= .27, p < .01$), with FDS ($r = .24, p < .05$) and YoE ($r= .20, p < .05$). For the last TSR parameter, statistically significant correlations of Ans were only found between the two recalled scores as aforementioned.

Table 13 Descriptive statistics of the content variables

	HC					MCI					AD					All participants						
	Range	Mode	Mediar	*M*	*SD*	Range	Mode	Median	*M*	*SD*	Range	Mode	Median	*M*	*SD*	Range	Mode	Median	*M*	*SD*	*P. 25*	*P. 75*
<u>TPD</u>																						
CIU	13-141	79	52.00	60.44	33.19	6-151	61	50.50	56.28	36.10	0-122	27	28.00	32.03	25.57	0-151	31	41.50	49.22	33.95	26.75	66.50
<u>TSR</u>																						
Imm	0-10	5	5.00	5.28	2.44	2-10	3	4.50	4.47	1.85	1-9	2	4.00	4.24	2.23	1-11	5	5.00	4.65	2.21	3	6
Del	0-10	5	5.50	5.97	2.65	0-9	4	4.00	4.28	2.17	0-9	3	3.00	3.64	2.70	0-10	3	5.00	4.62	2.68	3	7
Ans	0-4	1	2.00	1.88	1.24	0-4	1	1.00	1.34	1.00	0-3	1	1.00	1.56	0.96	0-4	1	1.00	1.59	1.08	1	2

Note: TPD = Thai Picture Description task; CIU = correct information unit; TSR = Thai Story Recall task; Imm = immediate recall scores; Del = delayed recall scores; Ans = correct answers to the five-factual inquires.

Table 14 Pearson’s correlation of age, years of education, the neuropsychological tests, and the content variables

	Age	YoE	MoCA	CIU	Imm	Del	Ans	LF	LF Rep	CF	CF Rep	FDS	BDS
Age	1.00												
YoE	-.17	1.00											
MoCA	-.26**	.33**	1.00										
CIU	-.18	.46**	.37**	1.00									
Imm	.05	.20	.28**	.11	1.00								
Del	-.05	.20*	.43**	.20*	.75**	1.00							
Ans	-.05	.13	.15	.11	.63**	.55**	1.00						
LF	-.18	.26**	.40**	.27**	.03	.14	.01	1.00					
LF Rep	-.04	.09	-.17	-.06	-.08	-.14	-.06	.14	1.00				
CF	-.25*	.34**	.43**	.40**	.22*	.27**	.14	.23*	.05	1.00			
CF Rep	-.12	.17	-.07	.00	.12	.11	.09	.05	.32**	.23*	1.00		
FDS	-.26*	.33**	.43**	.28*	.14	.24*	.12	.40**	-.05	.31**	.17	1.00	
BDS	.01	.26*	.41**	.11	.15	.33**	.13	.18	.02	.04	.02	.31**	1.00

* $p < .05$, ** $p < .01$

Note: YoE = years of education; CIU = correct information unit; Imm = immediate recall scores; Del = delayed recall scores; Ans = correct answers to the five-factual inquires; LF = letter fluency; LF Rep = repeated words in letter fluency; CF = category fluency; CF Rep = repeated words in category fluency; FDS = forward digit span; BDS = backward digit span.

The Shapiro-Wilk test was performed to examine a normal distribution of four content variables (Blanca et al., 2017; Field, 2018, p. 249). Table 15 presents the values of skewness, kurtosis, and normality test of TPD and TSR scores. The Shapiro-Wilk test reveals that the data of four variables are not normally distribution. However, One-way ANOVA is able to examine mean differences of these non-normal distributed data (Blanca et al., 2017).

Table 15 Normality Shapiro-Wilk test of the content variables

Variable	Skewness	Kurtosis	*Shapiro-Wilk*	*p*
TPD				
CIU	1.07	1.00	*0.92*	< *.001*
TSR				
Imm	0.50	-0.08	*0.96*	< *.01*
Del	0.12	1.59	*0.97*	< *.05*
Ans	0.33	-0.64	*0.90*	< *.01*

Note: TPD = Thai Picture Description task; CIU = correct information unit; TSR = Thai Story Recall task; Imm = immediate recall scores; Del = delayed recall scores; Ans = correct answers to the five factual inquires.

One-way ANOVA was used to determine the mean differences in the content variables between the three groups. Table 16 presents the statistics and significant tests of the one-way ANOVA for the dependent variables, which are four content variables from the spontaneous speech tests. Mean scores of CIU and of the delayed recall task were significantly different among the three groups ($p < .01$). Both variables obtained a similar tendency which was HC had the highest scores, followed by MCI and AD consecutively (See mean scores in Table 13). As for CIU, Turkey's HSD test found that the mean scores were significantly different between MCI and AD ($p < .01$) and HC and AD ($p < .01$). There was no statistically significant difference in mean CIU between HC and MCI ($p = .86$). Considering Del, Turkey's HSD test found that the mean scores were significantly different between HC and

MCI ($p < .05$) and HC and AD ($p < .01$). There was no statistically significant difference in mean CIU between MCI and AD ($p = .56$).

Table 16 ANOVA of the content variables

Variable	*df*	*F*	*p*	η_p^2
<u>TPD</u>				
CIU	2	*7.74*	*< .01*	.14
<u>TSR</u>				
Imm	2	2.05	.13	.04
Del	2	*7.39*	*< .01*	.14
Ans	2	1.99	.14	.04

Note: TPD = Thai Picture Description task; CIU = correct information unit; TSR = Thai Story Recall task; Imm = immediate recall scores; Del = delayed recall scores; Ans = correct answers to the five factual inquires.

According to the Chi-square test (Table 10) and Kruskal-Wallis *H* test (Table 11), both education level and years of education are significantly different among the three groups of participants. Therefore, further analyses were conducted to explore the effect of education on the participants' performance. The seven educational levels were grouped into a new, three-level variable called "education groups" distinguishing primary school (no schooling and primary level), high school (lower secondary and upper secondary), and university (higher level/diploma, bachelor's degree, and higher than bachelor's degree). The Spearman' rho reveals a significant correlation between education categorical variable (three-level education groups) and pathological characteristics of the participants (three experimental groups) at –.33 (p <.01). One-way ANOVA was performed to examine the variables which are different among the three groups of participants for each education level.

Table 17 presents the ANOVA statistics of the neuropsychological and content variables for each of the three education groups. No variable was found to differ in the group of primary school. For the high school, three variables are significantly different among the HC, MCI, and AD groups, i.e., LF (F(2) = 4.86, p <.05, η_p^2 = .34), BDS (F(2) = 6.23, p <.05, η_p^2 = .42), and Del (F(2) = 4.94, p <.05, η_p^2 = .34). For the university group, three variables show significant differences among the HC, MCI, and AD groups, i.e., LF Rep (F(2) = 4.65, p <.05, η_p^2 = .20), CF (F(2) = 5.38, p <.05, η_p^2 = .23), and BDS (F(2) = 4.17, p <.05, η_p^2 = .32).

Table 17 ANOVA of the neuropsychological tests and the content variables separated by education three-level

Variable	Education Group	*df*	*F*	*p*	η_p^2
LF	Primary school	2	1.86	.17	.10
	High school	2	*4.86*	*<.05*	.34
	University	2	0.71	.50	.04
LF Rep	Primary school	2	0.78	.47	.05
	High school	2	0.73	.49	.07
	University	2	*4.65*	*<.05*	.20
CF	Primary school	2	0.69	.51	.04
	High school	2	1.37	.28	.13
	University	2	*5.38*	*<.05*	.23
CF rep	Primary school	2	0.12	.89	.01
	High school	2	0.62	.55	.06
	University	2	1.05	.36	.05
FDS	Primary school	2	0.28	.76	.02
	High school	2	2.16	.15	.20
	University	2	2.29	.13	.20
BDS	Primary school	2	1.41	.26	.09
	High school	2	*6.23*	*<.05*	.42
	University	2	*4.17*	*<.05*	.32
Imm	Primary school	2	1.96	.16	.11
	High school	2	2.07	.15	.18
	University	2	1.68	.20	.08
Del	Primary school	2	1.29	.29	.07
	High school	2	*4.94*	*<.05*	.34

Variable	Education Group	*df*	*F*	*p*	η_p^2
	University	2	2.71	.08	.13
Ans	Primary school	2	0.13	.88	.01
	High school	2	2.78	.09	.23
	University	2	1.04	.36	.05
CIU	Primary school	2	2.13	.13	.11
	High school	2	2.02	.16	.18
	University	2	0.82	.45	.04

Note: LF = letter fluency; LF Rep = repeated words in letter fluency; CF = category fluency; CF Rep = repeated words in category fluency; FDS = forward digit span; BDS = backward digit span; Imm = immediate recall scores; Del = delayed recall scores; Ans = correct answers to the five-factual inquires; CIU = correct information unit.

4.2.2.2 Multivariate analysis of variance of the content variables

A multivariate analysis of variance (MANOVA) was performed across a group of content variables to examine the effect of the independent variable (three experimental grouping based on dementia pathological stages) on the dependent variable which are the content and neuropsychological variables. Variables selection and grouping were based on an evidence-based approach which was theory and statistical results in this study.

Three sets were (1) four content variables (i.e., CIU, Imm, Del, and Ans) and four neuropsychological variables (i.e., LF, CF, FDS, and BDS, while repeated scores of two verbal fluency tests were not included. Due to no correlation of two repeated scores with any of four neuropsychological variables and no significant difference among three groups of participants); (2) three variables of TSR and two scores of digit span tests; and (3) CIU and two scores of verbal fluency tests. The first set was selected based on the relationship between two spontaneous speech tests that CIU had positive weak correlation with Del. While the two proposed spontaneous speech tests showed relationship with the speech-based neuropsychological tests. The second set comprised of the story recall tasks and digit span tests. Story recall tests and digit span tests were found to share common cognitive function on auditory reception and memory, with the collaboration of the phonological loop and the central executive function (Monaco et al., 2013). According to the correlation (see Table 14), Del showed significant correlation with both digit span scores in weak level. The third set was language assessments. Verbal fluency tests were considered as an isolated test and picture description tests reflected communication abilities (Slegers et al., 2018). Also, CIU and both verbal fluency scores had significant correlations (see Table 14).

The assumptions concerning multivariate analysis were investigated before conducting the analysis. According to Hair et al. (2014), three assumptions must be verified before conducting MANOVA (pp. 684 – 686). These assumptions are a further prerequisite for discriminant analysis which are suggested to perform and report after MANOVA (Field, 2018, pp. 753 - 754; Hair et al., 2014, pp. 249 - 251). The testing procedure and results are described as follows:

a) Independence

Each group needs to respond independently and not be influenced by any other group. However, possible extraneous effects can occur uncontrollably. Hair et al. (2014) suggested the two most common violations of independence, which were time-ordered effects and setting of data collecting. The two phenomena were taken into consideration and assessed. For the experimental setting, all participants were interviewed in the same room, which was quiet with appropriate light and temperature at the Center of Excellence in Cognitive Science (CECoS), Burapha University. For the time-order effect, two factors were considered, i.e., time of the day (morning and afternoon session) and test sequence. Since two tests in the speech tasks (verbal fluency tests and digit span tests) and in one of the subtest in MoCA were similar, a possible influence might happen because a participant potentially became fluent after performing similar tests. Test sequences were thus conducted in two series, including having the MoCA test prior to the speech tasks or vice versa (see section 3.2.4 Procedure and data collecting). The direct effects of the time of the day and its interaction with test sequence were evaluated by a two-way ANOVA.

Table 18 presents the results of the univariate test for eight dependent variables and MoCA. The test sequence had no direct effects. However, a significant effect was found for both the time of the day and the interaction between time of the day and test sequence. Simple direct effects analysis revealed that the time of the day affects CF (*p* = .04, effect size = .05). Participants performing CF in the morning session (*M* = 19.02, *SD* = 5.38) generated more words than participants performing the task in the afternoon session (*M* = 16.64, *SD* = 5.71). The interaction effect between time of the day and test sequence (*F(1, 64)* = 5.72, *p* = .02, effect size = .08) was due to the fact that the BDS scores were higher in the morning session when the speech test were conducted before the MoCA (*M* = 5.15, *SD* = 0.48), followed by the afternoon session with testing of the MoCA prior the speech test (*M* = 4.12, *SD* = 0.42), and followed by the morning session with the MoCA preceding the speech test (*M* = 3.79, *SD* = 0.40).

Table 18 Univariate two-way ANOVA of the neuropsychological tests and the content variables

Variable	Time of the day		Test sequence		Interaction effect	
	F	*p*	*F*	*p*	*F*	*p*
Neuropsychological tests						
MoCA	2.12	.15	0.15	.70	0.23	.64
LF	1.88	.17	0.26	.61	2.42	.12
CF	*4.45*	*.04*	1.25	.27	1.31	.26
FDS	0.62	.44	0.07	.79	1.51	.22
BDS	2.59	.11	0.74	.39	*5.72*	*.02*
Spontaneous speech tests						
CIU	2.60	.11	0.14	.71	0.04	.85
Imm	0.07	.80	0.70	.41	2.49	.12
Del	2.37	.13	0.64	.43	1.76	.19
Ans	0.22	.64	0.53	.47	2.07	.15

Note: LF = letter fluency; CF = category fluency; FDS = forward digit span; BDS = backward digit span; CIU = correct information unit; Imm = immediate recall scores; Del = delayed recall scores; Ans = correct answers to the five factual inquires.

The afternoon session with testing of the speech tasks prior to the MoCA showed the lowest score (*M* = 3.47, *SD* = 0.40). Although, the main effect was found in CF and the interaction effect had an effect on BDS, the participants were randomly assigned to the experiment sessions regardless of time of the day or test sequence. In addition, there was no correlation between the participants' condition of the disease and the time of the day or test sequence. Also, the effect size of the two statistically significant effects was very small (Lakens, 2013). Therefore, we may assume that the three groups of participants were in compliance with the assumption of independence.

b) Equality of Variance–Covariance Matrices

In the MANOVA, all matrices of variance and covariance produced by dependent variables are expected to be equal across the groups. Homoscedasticity is assumed when there is a homogeneity of variance across groups of independent variables and no difference between the groups on dependent variables collectively. When this is not the case, the violation is called heteroscedasticity. SPSS provides the statistics for the test of homoscedasticity assumption in the form of the Box's M test. This test should be nonsignificant differences; the equality of covariance thus can be assumed. The suggested threshold of the *p*-value for the Box's M test was .01 (instead of .05) for a very conservative level of significant differences (Hair et al., 2014, p. 251 and 685). The results of the Box's M test in Table 19 can imply that three proposed sets of dependent variables were in accordance with the homoscedasticity assumption.

Table 19 Box's test of equality of covariance matrices for three sets of neuropsychological tests and content variables

Set of variables	Box's M	*F*	*df 1*	*df 2*	*p*
Speech tasks	80.65	0.91	72	10084.40	.69
Del & DS	9.93	0.78	12	22083.50	.68
CIU & VF	16.80	1.33	12	32200.05	.19

Note: Del = delayed recall scores; DS = two versions of digit span tests; CIU = correct information unit; VF = two versions of verbal fluency tests.

c) Normality

This assumption requires multivariate normal distribution in all variables. However, there is no specific test for multivariate normality. Hair et al. (2014, p. 251 and 686) mention that a test for univariate normality can be used for each variable, but it does not guarantee multivariate normality. The univariate normality tests of the dependent variables can be seen in Tables 9 and 15, that only CF is normally distributed. Although this assumption is violated for most variables, little impact would occur, especially with a large sample size.

For moderate sample size such this study, outliers should be concerned and corrected. Since they can cause violations of the normality assumption and lead to the Type I error (Blanca et al., 2017; Hair et al., 2014). This study identified the outlier by an interquartile method. The range of accepted values of each variable was calculated as a result of an interquartile multiplied by 1.5 (Taylor, 2020). The lower bound was created by taking 1.5 times interquartile to subtract from values at the 25th percentile. At the same time, the upper bound was adding 1.5 times of interquartile to value at the 75th percentile. Any values lower or higher than the boundaries were considered outliers.

Table 20 Outliers in the neuropsychological tests and the content variables

Variable	Lower bound	Upper bound	Outlier value	Number of cases with outliers
CIU	–32	125	128, 133, 141, 149, 151	5
Imm	–2	11		
Del	–3	13.0		
Ans	–1	4		
MoCA	13	33	11	1
LF	–4	25	28, 30, 31	4
CF	4	32	37	1
FDS	2	10		
BDS	0	8		

Note: LF = letter fluency; CF = category fluency; FDS = forward digit span; BDS = backward digit span; CIU = correct information unit; Imm = immediate recall scores; Del = delayed recall scores; Ans = correct answers to the five factual inquires.

Table 20 depicts the valid range of scores and the number of outliers for each variable. Outliers were found in four variables, i.e., CIU (5 cases), MoCA (1 case), LF (4 cases), and CF (2 cases). These cases were removed before running the MANOVA.

The following section presents three MANOVAs on the set of variables.

i. MANOVA on the set of speech tasks

Set 1 of dependent variables consisted of CIU, Imm, Del, Ans, LF, CF, FDS, and DBS. The Box's M test of this set presents a nonsignificant value (p = .69), which implies equal covariance matrices between three groups on the eight dependent variables collectively (see Table 19). There was a significant main effect of dementia pathological stages on the set of speech tasks with moderate effect size (F(16, 116) = 3.33, p <.001, effect size = .32) (see Table 21). The valid samples for this set of variables were 67 due to the missing data in digit span tests and removing outlier cases. However, the univariate tests indicated a nonsignificant difference on Imm (F(2) = 0.64, p = .53, η_p^2 = .02) and Ans (F(2) = 2.16, p = .12, η_p^2 = .06) between three groups (see Table 22). Since there is no direct effect of dementia pathological stage on Imm and Ans, only Del would be included in the following analysis.

Table 21 Multivariate tests for three sets of neuropsychological tests and content variables

Set of variables	*Pillai's Trace*	*F*	*Hypothesis df*	*Error df*	*p*	η_p^2
Speech tasks	0.63	*3.33*	16.00	116.00	*<.001*	.32
Del & DS	0.37	*5.08*	6.00	136.00	*<.001*	.18
CIU & VF	0.24	*3.77*	6.00	170.00	*<.01*	.12

Note: Del = delayed recall scores; DS = two versions of digit span tests; CIU = correct information unit; VF = two versions of verbal fluency tests.

Table 22 Tests of between-subjects effects for set 1 - the speech tasks (n = 67)

Variable	*df*	*F*	*p*	η_p^2
CIU	2	*9.27*	*<.001*	.23
Imm	2	0.64	.53	.02
Del	2	*6.99*	*<.01*	.18
Ans	2	2.16	.12	.06
LF	2	*4.74*	*<.05*	.13
CF	2	*4.62*	*<.05*	.13
FDS	2	*4.35*	*<.05*	.12
BDS	2	*10.38*	*<.001*	.25

Note: CIU = correct information unit; Imm = immediate recall scores; Del = delayed recall scores; Ans = correct answers to the five-factual inquires; LF = letter fluency; CF = category fluency; FDS = forward digit span; BDS = backward digit span.

ii. MANOVA on the set of Del and digit span tests

According to Pearson correlation (Table 14), Del has a significant relationship with two variables of digit span tests, consequently the set Del, FDS, and BDS was examined in MANOVA. The Box's M test of this set reveals a nonsignificant value ($p = .68$); consequently, homoscedasticity is assumed (see Table 19). There was a significant main effect of groups on the set of Del and digit span tests ($F(6, 136) = 5.07$, $p < .001$, $\eta_p^2 = .18$) (see Table 21). The valid samples for this set were 72 due to the missing data in digit span tests and removing outlier cases. The univariate tests confirm a significant difference in each dependent variable between the three groups of participants (see Table 23).

Table 23 Tests of between-subjects effects for set 2 Del and digit span tests (n = 72)

Variable	*df*	*F*	*p*	η_p^2
Del	2	*7.23*	*<.01*	0.17
FDS	2	*6.22*	*<.01*	0.15
BDS	2	*11.25*	*<.001*	0.25

Note: Del = delayed recall scores; FDS = forward digit span; BDS = backward digit span.

iii. MANOVA on the set of CIU and verbal fluency tests

According to Pearson correlation (Table 14), CIU has a significant relationship with two variables of verbal fluency tests, the set of CIU, LF and CF thus was grouped for MANOVA. The Box's M test of this set reports a nonsignificant value ($p = .19$); consequently, homoscedasticity is assumed (see Table 19). There was a significant main effect of groups on the set of CIU and verbal fluency tests ($F(6, 170) = 3.77$, $p < .01$, effect size = .12) (see Table 21). The valid samples for this set of variables were 89 due to removing outlier cases. The univariate tests confirm a significant difference in each dependent variable between the three groups of participants (see Table 24).

Table 24 Tests of between-subjects effects for set 3 CIU and VF (n = 89)

Variable	*df*	*F*	*p*	η_p^2
CIU	2	*6.29*	*<.01*	0.13
LF	2	*4.98*	*<.01*	0.10
CF	2	*7.66*	*<.01*	0.15

Note: CIU = correct information unit; LF = letter fluency; CF = category fluency.

4.2.2.3 Discriminant analysis of the content variables

Multiple discriminant analysis (MDA) is suggested as a follow-up analysis after the MANOVA (Field, 2018, p. 765). This analysis has the objective of obtaining the model of the content variables and neuropsychological variables for predicting dementia pathological stages. The assumptions of MDA are similar to MANOVA, which consists of homoscedasticity between individual variables (Box's M test for MDA) and multivariate normality. The MDA assumption regarding sample size is at least 20 samples for one category of the dependent variable and at least five observations per independent variable. Moreover, MDA requires a lack of multicollinearity; if two or more predictive variables are highly correlated, the MDA cannot proceed with an estimation. (Hair et al., 2014, pp. 250 - 251). For stepwise estimation, multicollinearity is measured by tolerance (see Section 4.2.4.1). MDA in this section is performed with a simultaneous procedure.

Only six predictor variables were entered into MDA (independent variables in this analysis). Since the scores of Imm and Ans were not different between the three groups based on the univariate test in Table 22 (MANOVA 1) they were not selected. The three groups of participants are categorical dependent variables in this MDA. MDA was conducted with simultaneous estimation. Table 25 reports the mean scores of the six predictors in each group and the tests of equality of group means. The mean scores of each variable were significantly different between the three groups. The Box's M test indicates that the covariance matrices of the six predictors do not differ between groups since this test presents a non-significant statistic value ($p = .86$). Table 26 reveals that two discriminant functions are estimated. Function 1 achieves the eigenvalue of 0.77, and it accounts for 84.5% of the explained variance. The discriminant function 1 returns the higher canonical correlation, which is .66. The square of this coefficient is .43. It can be inferred that 43.56% of the variance in the dementia pathological variable is explained by model 1. Wilk's lambda reports a significant difference resulting from two discriminant functions together (χ^2 (12) = 43.07, $p < .001$). Then the first function is removed, and the second function alone is unable to discriminate among the three groups of participants (χ^2 (5) = 8.07, $p = .15$). In other words, the first function of six variables significantly differentiates between the three groups of participants. According to the unstandardized coefficients, the discriminant function (equation) of the six predictors can be written as follows:

$$\textit{Discriminant scores} = 0.36(BDS) + 0.15(FDS) + 0.14(Del) + 0.03(CF) + 0.03(CIU) + 0.02(LF) - 4.67$$

Table 25 Tests of equality of group means for the content variable (n = 67)

Predictor	Group	*M*	*SD*	*Wilks' Lambda*	*F*	*df1*	*df2*	*p*
CIU	HC	51.74	23.48	0.78	*9.27*	2	64	*<.001*
	MCI	41.86	18.19					
	AD	25.69	20.21					
Del	HC	6.16	2.69	0.82	*6.99*	2	64	*<.01*
	MCI	4.09	1.87					
	AD	3.46	2.67					
LF	HC	10.79	4.14	0.87	*4.74*	2	64	*<.05*
	MCI	11.14	5.17					
	AD	7.42	4.45					
CF	HC	19.05	4.50	0.87	*4.62*	2	64	*<.05*
	MCI	18.27	4.00					
	AD	15.12	5.32					
FDS	HC	6.79	1.23	0.88	*4.35*	2	64	*<.05*
	MCI	6.05	1.13					
	AD	5.77	1.14					
BDS	HC	5.42	1.35	0.76	*10.38*	2	64	*<.001*
	MCI	3.41	1.62					
	AD	3.38	1.83					

Note: CIU = correct information unit; Del = delayed recall scores; LF = letter fluency; CF = category fluency; FDS = forward digit span; BDS = backward digit span.

Table 26 Eigenvalue and significant tests of the discriminant functions for the content variable (n = 67)

Function	Eigenvalue	Percentages of Variance	Canonical Correlation	After function	Wilks' Lambda	χ^2	*df*	*p*
				0	0.50	*43.07*	12	*<.001*
1	0.77	84.54	.66	1	0.88	8.07	5	.15
2	0.14	15.46	.35					

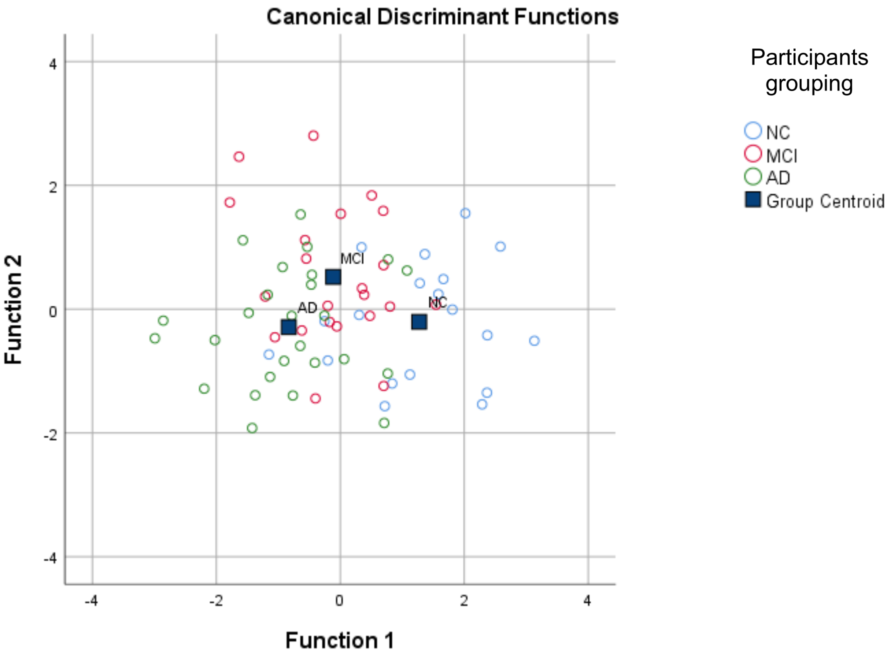


Figure 19 Group centroids of canonical discriminant functions of the content variables

Table 27 Classification results of the discriminant functions for the content variable (n = 67)

Actual groups membership	Predicted Group Membership			Total
	HC	MCI	AD	
HC (*n*)	14	2	3	19
MCI (*n*)	3	12	7	22
AD (*n*)	3	5	18	26
HC (%)	73.68	10.53	15.79	100
MCI (%)	13.64	54.55	31.82	100
AD (%)	11.54	19.23	69.23	100

65.7% of original grouped cases correctly classified.

Table 28 Structure matrix of the discriminant functions for the content variable (n = 67)

Variable	Function 1	Function 2
BDS	0.61*	-0.52
CIU	0.58*	0.45
Del	0.53*	-0.15
FDS	0.42*	-0.08
LF	0.32	0.71*
CF	0.38	0.49*

* Largest absolute correlation between each variable and any discriminant function

Note: BDS = backward digit span; CIU = correct information unit; Del = delayed recall scores; FDS = forward digit span; LF = letter fluency; CF = category fluency.

The model of six variables provided 65.7% of correct classification (see Table 27). The level of accuracy varied in each group of dementia pathological stages, i.e., 73.7% of HC, 54.5% of MCI, and 69.2% of AD. Only half of MCI was correctly classified.

4.2.2.4 Receiver Operating Characteristic curve analysis for content variables

The discriminant ability was validated by Receiver Operating Characteristic (ROC) and area under the curve (AUC) method. In this step, the discriminant function of the six variables will be evaluated. Discriminant scores derived from the previous process were assessed for the AUC, sensitivity, and specificity in differentiating the four pairs of experimental groups (HC, MCI, and AD). The illustrations of ROC curve are presented to show varying cut-offs generated by the plot of sensitivity and 1-specificity. The reported sensitivity and specificity were selected regarding the optimal cut-off scores based on the Youden Index.

i. HC vs. MCI

The results of various cut-off values of the discriminant scores derived from the six variables for MCI screening among non-AD groups (HC and MCI) are given in Figure 20. The corresponding ROC curve generates the AUC of .84 (*p* <.001, 95 % confidence interval (CI), .71–.97). In this analysis, the optimal cut-off scores in detecting MCI from HC return the sensitivity of 73.68% and specificity of 90.91%.

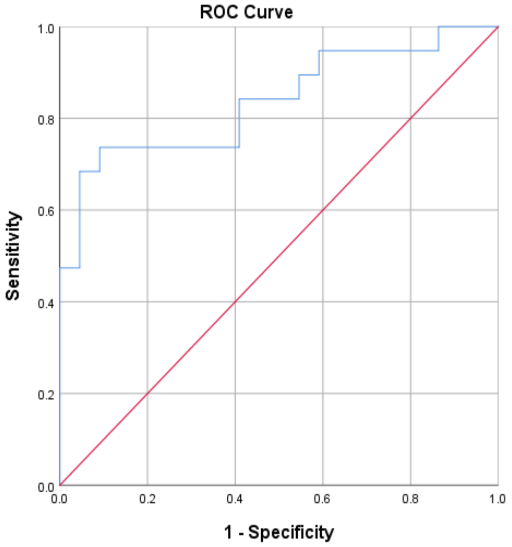


Figure 20 AUC of the discriminant model of the content variables in detecting MCI among non-AD groups

ii. MCI vs. AD

Figure 21 illustrates the ROC curve analysis of the six variables in detecting MCI among cognitive impairment groups (MCI and AD). The AUC of the discriminant scores was .70 (*p* <.05, 95 % CI, .55–.86). The optimal cut-off point presents a sensitivity of 81.82% and specificity of 61.54%.

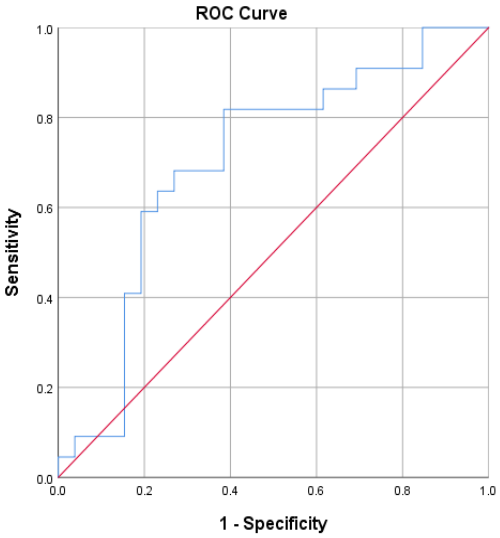


Figure 21 AUC of the discriminant model of the content variables in detecting MCI among the two cognitive impairment groups

iii. HC vs. AD

The AUC of the pair of HC and AD is depicted in Figure 22 with high accuracy at .92 (*p* <.001, 95 % CI, .84–1.00). Due to an outstanding AUC, the sensitivity is 94.74%, and the specificity is 80.77%.

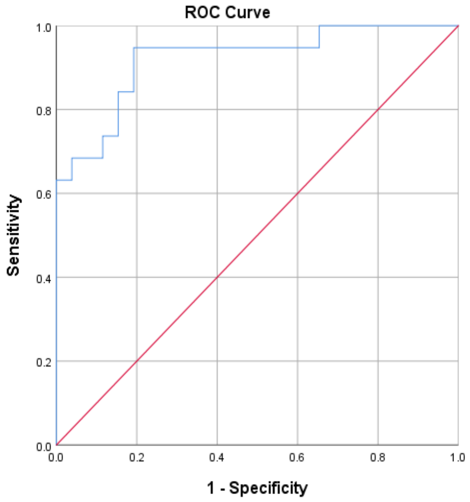


Figure 22 AUC of the discriminant model of the content variables in detecting AD from HC

iv. HC vs. Cognitive impairment

The fourth pair is the result of differentiating HC and persons with cognitive impairment; MCI and AD were considered in the same group. An excellent level of AUC is derived, an AUC of .88 (*p* <.001, 95 % CI, .79–.98) (see Figure 23). The optimal cut-off point presents a sensitivity of 94.74% and specificity of 62.50% in detecting a person with cognitive impairment.

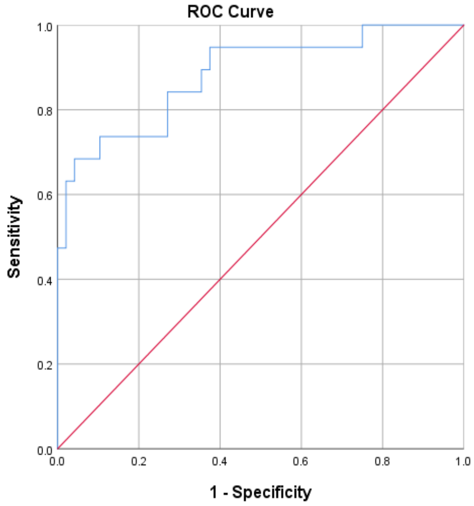


Figure 23 AUC of the discriminant model of the content variables in detecting the persons with cognitive impairment among the total sample

4.2.3 Analysis of the acoustic features

This study considered the acoustic features of frequency and temporal dynamics. Each domain was extracted by different techniques (see section 3.2.5 Speech extraction).

4.2.3.1 Characteristics of the acoustic variables among three groups

i. Frequency features

Table 29 shows the descriptive statistics of frequency-related variables for the spontaneous speech tasks presenting for the total samples and categorized by the three experimental groups. All acoustic features in the frequency domain were found to have no consistent tendency among the three groups. Table 30 displays Pearson's correlation of age, years of education, MoCA scores, and 20 frequency variables. Years of education significantly correlated with NVB of PD (r = 0.43, $p < .01$) and SI (r = 0.40, $p < .01$) at a moderate level. Apart from the NVB of the two tasks, all of the frequency-related variables obtained very weak correlation magnitude, approximately at 0 – 0.1 with age, YoE, and MoCA. Each frequency-related variable further showed inconsistent correlation direction with the three independent variables.

Table 29 Descriptive statistics of the frequency variables (n=98)

	HC					MCI					AD					All participants						
	Range	Mode	Median	*M*	*SD*	Range	Mode	Median	*M*	*SD*	Range	Mode	Median	*M*	*SD*	Range	Mode	Median	*M*	*SD*	P. 25	P. 75
TPD																						
F_0	105.45-226.98	105.45	179.27	169.48	33.33	126.52-217.14	126.52	163.08	164.40	25.00	104.69-217.56	104.69	172.30	170.06	28.16	104.69-226.98	104.69	171.86	168.02	28.83	142.74	191.24
Jitter	2.06-5.07	2.06	2.83	3.03	0.68	2.00-5.49	2.00	3.13	3.25	0.80	2.16-5.56	2.16	2.96	3.16	0.80	2.00-5.56	2.00	2.96	3.15	0.76	2.60	3.48
Shimmer	0.12-19.94	0.12	12.85	12.90	4.82	8.74-20.44	8.74	14.13	14.21	2.60	0.15-19.99	0.15	14.11	13.28	4.91	0.12-20.44	0.12	13.80	13.46	4.25	12.13	16.01
NVB	9.00-217.00	9.00	53.50	71.24	48.70	12.60-322.00	40.00	64.23	84.90	68.71	8.67-160.67	40.67	41.42	51.85	37.19	8.67-322.00	40.67	49.88	68.97	54.09	36.42	92.13
HNR	3.79-11.03	3.79	8.00	7.81	1.73	3.89-12.15	3.90	7.41	7.36	2.12	3.79-10.89	3.79	7.43	7.39	1.69	3.79-12.15	6.47	7.54	7.51	1.84	6.18	8.73
Imm																						
F_0	101.99-236.13	101.99	174.39	167.98	34.79	119.33-259.69	119.33	158.51	162.53	30.96	97.44-215.62	97.44	168.11	166.24	28.74	97.44-259.69	97.44	164.98	165.59	31.29	144.27	188.58
Jitter	2.12-5.21	2.12	2.99	3.28	0.90	2.01-5.43	2.01	3.31	3.53	0.85	1.74-5.66	1.74	3.28	3.50	1.00	1.74-5.66	5.43	3.18	3.44	0.91	2.75	4.18
Shimmer	10.28-22.41	14.29	14.52	15.08	2.44	10.63-19.56	10.63	16.31	16.07	2.50	8.05-21.05	8.05	15.30	15.02	2.71	8.05-22.41	14.29	15.17	15.38	2.58	13.86	17.17
NVB	6.50-92.00	38.00	38.67	42.39	22.66	10.50-128.00	31.25	39.92	47.79	28.58	6.00-138.50	6.00	32.04	37.95	26.18	6.00-138.50	59.00	36.75	42.61	25.97	25.44	55.42
HNR	2.82-10.64	2.82	7.55	7.50	2.10	4.32-11.19	4.33	5.91	6.63	1.88	4.11-12.54	4.11	6.62	7.04	1.99	2.82-12.54	2.82	6.64	7.06	2.00	5.57	8.79

	HC					MCI					AD					All participants						
	Range	Mode	Median	*M*	*SD*	Range	Mode	Median	*M*	*SD*	Range	Mode	Median	*M*	*SD*	Range	Mode	Median	*M*	*SD*	P. 25	P. 75
<u>Del</u>																						
F_0	102.94-267.88	102.94	179.39	176.79	38.99	114.68-207.10	114.68	171.13	167.68	25.43	111.97-250.03	111.97	168.64	171.78	34.97	102.94-267.88	102.94	174.40	172.08	33.50	149.56	191.19
Jitter	2.21-6.94	2.21	3.18	3.46	1.15	1.88-6.52	1.88	3.02	3.35	0.99	2.00-5.89	2.00	3.23	3.46	0.96	1.88-6.94	1.88	3.13	3.42	1.02	2.74	3.83
Shimmer	10.08-19.42	10.08	16.24	15.90	2.52	9.43-21.25	9.43	15.37	15.39	2.97	9.13-19.81	9.13	15.27	15.05	2.70	9.13-21.25	9.13	15.55	15.44	2.73	13.99	17.23
NVB	5.67-92.00	29.50	29.50	35.72	20.75	5.75-205.00	35.00	34.13	43.10	37.44	4.50-128.00	17.00	17.25	27.60	26.22	4.50-205.00	29.50	29.25	35.32	29.33	16.19	43.00
HNR	2.88-9.84	2.88	6.93	6.97	1.95	3.89-11.56	3.89	7.02	7.01	2.04	4.58-12.25	4.58	6.80	7.09	1.78	2.88-12.25	2.88	6.87	7.02	1.90	5.55	8.34
<u>SIT</u>																						
F_0	103.24-217.14	103.24	175.14	170.35	32.21	127.45-215.41	127.45	169.60	170.81	25.34	110.65-228.10	110.65	179.89	176.08	32.24	103.24-228.10	168.10	175.44	172.49	29.95	151.41	195.91
Jitter	2.44-4.06	2.44	2.95	2.98	0.37	1.93-3.98	1.93	3.11	3.07	0.42	2.08-3.94	2.08	2.93	2.95	0.40	1.93-4.06	2.99	2.98	3.00	0.40	2.73	3.27
Shimmer	1.17-2.20	1.17	1.46	1.50	0.21	0.86-2.11	0.87	1.54	1.55	0.25	0.95-2.02	0.95	1.43	1.46	0.23	0.86-2.20	1.48	1.48	1.50	0.23	1.34	1.67
NVB	7.59-34.09	13.55	17.75	18.85	7.49	7.35-50.50	7.35	18.64	20.77	10.57	6.76-33.33	6.76	15.95	15.76	5.98	6.76-50.50	13.55	17.03	18.40	8.38	12.95	22.93
HNR	3.61-9.20	3.61	7.32	7.20	1.28	4.13-10.96	4.13	7.26	7.31	1.61	5.00-11.35	5.00	7.37	7.70	1.47	3.61-11.35	7.26	7.32	7.41	1.46	6.59	8.34

Note: TPD = Thai Picture Description task; Imm = immediate recall task; Del = delayed recall task; SIT = Semi-structured Interview for Thai; F_0 = fundamental frequency in Hz; Jitter = relative jitter in percentages; Shimmer = relative shimmer in percentages; NVB = number of voice breaks; HNR = mean harmonics-to-noise ratio in decibel

Table 30 Pearson's correlation of age, years of education, and the frequency variables

	Age	YoE	MoCA	PD F_0	PD Jitter	PD Shimmer	PD NVB	PD HNR	IM F_0	IM Jitter	IM Shimmer	IM NVB	IM HNR	DL F_0	DL Jitter	DL Shimmer	DL NVB	DL HNR	SI F_0	SI Jitter	SI Shimmer	SI NVB	SI HNR
Age	1																						
YoE	-.17	1																					
MoCA	-.26**	.33**	1																				
PD F_0	.01	-.06	-.03	1																			
PD Jitter	.13	-.04	-.07	-.23*	1																		
PD Shimmer	-.09	0	-.06	-.18	.20*	1																	
PD NVB	-.01	.43**	.12	-.1	.15	.15	1																
PD HNR	-.18	.02	.06	.48**	-.75**	-.12	-.19	1															
IM F_0	-.11	.04	.04	.77**	-.11	-.12	0	.28**	1														
IM Jitter	.03	-.08	-.14	-.25*	.63**	.12	-.02	-.48**	-.15	1													
IM Shimmer	.08	.07	.04	-.15	.51**	.03	-.01	-.60**	-.09	.65**	1												
IM NVB	.02	.08	.07	-.06	.16	.08	.26*	-.2	-.07	.16	.2	1											
IM HNR	-.12	-.02	.09	.46**	-.60**	-.16	-.07	.78**	.41**	-.74**	-.82**	-.27**	1										
DL F_0	-.01	.01	.14	.75**	-.21*	-.25*	-.03	.34**	.73**	-.23*	-.15	-.11	.42**	1									
DL Jitter	.11	-.13	-.05	-.19	.55**	.03	-.04	-.41**	-.23*	.56**	.41**	.23*	-.43**	-.16	1								
DL Shimmer	.1	.03	.07	.01	.37**	-.1	.04	-.46**	.06	.33**	.60**	.12	-.42**	-.05	.59**	1							
DL NVB	.16	.05	.06	.04	.18	.07	.07	-.1	0	.22*	.26*	.47**	-.20*	-.05	.23*	.27**	1						
DL HNR	-.11	.01	.02	.38**	-.58**	-.03	-.09	.74**	.30**	-.51**	-.61**	-.23*	.70**	.39**	-.72**	-.78**	-.24*	1					
SI F_0	.03	-.02	-.05	.83**	-.36**	-.1	-.09	.50**	.77**	-.26**	-.17	-.07	.47**	.75**	-.32**	-.08	.06	.48**	1				
SI Jitter	.17	.14	.04	-.22*	.69**	.08	.21*	-.55**	-.09	.59**	.58**	.25*	-.54**	-.15	.51**	.40**	.25*	-.52**	-.21*	1			

	Age	YoE	MoCA	PD F_0	PD Jitter	PD Shimmer	PD NVB	PD HNR	IM F_0	IM Jitter	IM Shimmer	IM NVB	IM HNR	DL F_0	DL Jitter	DL Shimmer	DL NVB	DL HNR	SI F_0	SI Jitter	SI Shimmer	SI NVB	SI HNR
SI Shimmer	.16	.15	.08	-.11	.59**	-.02	.23*	-.49**	.01	.53**	.61**	.25*	-.48**	-.03	.47**	.49**	.27**	-.50**	-.06	.95**	1		
SI NVB	.11	.40**	.14	-.16	.2	.05	.36**	-.21*	.01	.05	.11	.29**	-.14	-.12	0	.11	.28**	-.13	-.13	.24*	.22*	1	
SI HNR	-.13	-.14	-.17	.43**	-.51**	.05	-.19	.81**	.27**	-.32**	-.60**	-.27**	.70**	.32**	-.36**	-.54**	-.13	.76**	.50**	-.56**	-.54**	-.21*	1

* $p < .05$, ** $p < .01$

Note: PD = Thai Picture Description task; IM = immediate recall task; DL = delayed recall task; SI = Semi-structured Interview for Thai; F_0 = fundamental frequency; Jitter = relative jitter; Shimmer = relative shimmer; NVB = number of voice breaks; HNR = mean harmonics-to-noise ratio.

One-way ANOVA was performed to explore the difference between groups across 20 variables. Although NVB in TPD and SI are found to be significant differences between the three groups, both variables reveal NVB tendency by MCI > HC > AD. Moreover, the p-value of NVB in TPD ($F = 3.27$, $p < .05$) and SIT ($F = 3.15$, $p < .05$) just pass critical values (see Table 31). Eventually, the post-hoc test reported that there were statistically significant differences only in the pair of MCI and AD for both NVB in TPD ($p < .05$) and NVB in SIT ($p < .05$).

Table 31 P-values in ANOVA of the frequency variables

Variable	TPD	Imm	Del	SIT
F_0 (Hz)	.69	.78	.56	.69
Jitter (%)	.49	.52	.88	.46
Shimmer (%)	.45	.18	.45	.35
NVB (no.)	*<.05*	.31	.10	*<.05*
HNR (dB)	.56	.22	.97	.35

Note: TPD = Thai Picture Description task; Imm = immediate recall task; Del = delayed recall task; SIT = Semi-structured Interview for Thai; F_0 = fundamental frequency; Jitter = relative jitter; Shimmer = relative shimmer; NVB = number of voice breaks; HNR = mean harmonics-to-noise ratio

Due to the significant differences in education level among three groups of the participants, the following analyses were conducted to examine the effect of education on the participant's acoustic features. Three education groups are separately examined with One-way ANOVA, and the results are shown in Table 32. Two frequency features are selected because these two variables found a significant difference between three groups of participants in overall samples. However, NVB of TPD and SIT present non-significant differences in every three groups of education.

Table 32 ANOVA of the frequency variables separated by education three-level

Variable	Education Group	*df*	*F*	*p*	η_p^2
NVB TPD	Primary school	2	0.29	.75	.02
	High school	2	0.20	.82	.02
	University	2	1.85	.17	.09
NVB SIT	Primary school	2	0.87	.43	.05
	High school	2	2.37	.12	.20
	University	2	0.70	.50	.04

Note: NVB TPD = number of voice breaks in the Thai Picture Description task; NVB SIT = number of voice breaks in the Semi-structured Interview for Thai.

ii. Temporal features

Nine temporal variables of the spontaneous speech tasks are presented in Table 33 for each group. There is no consistent pattern of temporal features among the three groups of participants. According to locution, the participants spent approximately one minute and a half telling the story in the picture description (*M* = 90.32, *SD* = 38.35). The response time in the recall tasks was similar in Imm (*M* = 51.09, *SD* = 21.77) and Del (*M* = 42.98, *SD* = 22.38). The mean locution of SIT was approximately four minutes (*M* = 246.90, *SD* = 66.40).

Table 34 presents Pearson's correlation between age years of education, MoCA scores, and the primary set of temporal variables. Interestingly, age does not have a relationship with any temporal features. YoE significantly correlates with a total time of utterance in TPD (*r* = .25) and SI (*r* = .22) at a weak level. This demographic variable further has a weak negative correlation with silence variables in Del, i.e., number of silence segments (*r* = –.23) and total duration of silence (*r* = –.27). While MoCA positively correlates with three utterance variables in a weak level, i.e., number of utterance segments of Imm (*r* = .20), total duration of utterance in Imm (*r* = .20), and total utterance of duration in Del (*r* = .30).

Pearson's correlations of the transformed variables of temporal features and three independent variables are displayed in Table 35. The relatively high magnitudes of correlations are found in two spontaneous speech tasks, namely TPD and Del. Years of education weakly correlate with four temporal features of TPD, i.e., utterance proportion (*r* = .28), silence proportion (*r* = –.28), pause rate (*r* = –.23), and hesitation rate (*r* = –.22). This education variable also significantly correlates, albeit weakly with the same variables in Del, i.e., utterance proportion (*r* = .30), silence proportion (*r* = –.30), pause rate (*r* = –.22), and hesitation rate (*r* = –.20). Moreover, two variables in Del reveal a significant relationship with MoCA, namely utterance proportion (*r* = .23), and silence proportion (*r* = –.23).

Table 33 Descriptive statistics of the temporal variables

	HC					MCI					AD					All participants						
	Range	Mode	Median	M	SD	Range	Mode	Median	M	SD	Range	Mode	Median	M	SD	Range	Mode	Median	M	SD	P. 25	P. 75
TPD																						
Utt seg	9-42	14.00	23.00	23.56	8.30	11-62	23.00	22.50	26.84	14.25	6-63	10.00	18.50	21.76	12.85	6-63	22.00	22.00	24.01	12.15	15.00	31.25
Utt time	21.28-115.54	21.28	56.34	57.41	26.01	18.69-137.99	18.69	48.99	58.36	29.11	8.76-163.27	8.76	35.84	49.83	39.19	8.759-163.27	8.76	46.48	55.09	32.01	31.98	73.81
Sil seg	2-25	12.00	12.50	13.16	6.10	4-34	8.00	14.00	15.06	7.83	0-29	8.00	11.50	12.32	7.10	0-34	8.000	13.00	13.49	7.07	8.00	19.00
Sil time	3.57-62.53	3.57	34.86	32.29	16.25	4.84-90.20	4.82	37.14	39.28	24.68	0-119.88	0	29.63	34.18	24.83	0-119.88	0	32.11	35.23	22.30	19.48	49.32
Loc	46.72-146.49	46.72	85.55	89.70	26.92	41.30-191.65	41.30	86.56	97.64	38.49	15.04-192.96	15.04	72.56	84.01	46.43	15.04-192.96	15.04	83.34	90.32	38.35	58.79	115.49
Utt prop	33.10-96.33	33.10	61.77	63.00	17.17	18.85-96.62	18.85	60.45	60.67	18.23	12.03-100	12.03	55.41	56.83	21.08	12.03-100	12.03	61.25	60.10	18.93	47.74	72.74
Sil prop	3.67-66.90	3.67	38.23	37.00	17.17	3.38-81.15	3.38	39.55	39.33	18.23	0-87.97	0	44.59	43.17	21.08	0-87.97	0	38.75	39.90	18.93	27.26	52.26
Ps rate	0.02-0.25	0.02	0.15	0.15	0.06	0.03-0.22	0.03	0.16	0.15	0.05	0-0.23	0	0.15	0.15	0.05	0-0.25	0	0.16	0.15	0.05	0.13	0.19
Hes rate	0.08-1	0.08	0.59	0.57	0.21	0.18-1	0.70	0.63	0.58	0.18	0-0.9	0.67	0.59	0.58	0.20	0-1	0.67	0.60	0.58	0.20	0.47	0.72
Imm																						
Utt seg	5-35	14.00	14.00	16.00	7.33	3-33	12.00	12.50	15.59	7.29	4-41	13.00	12.50	12.71	6.93	3-41	12.00	13.00	14.72	7.26	10.00	18.00
Utt time	7.64-78.13	7.63	28.32	30.05	15.41	10.84-64.1	10.84	23.58	27.80	13.61	8.15-65.85	8.15	22.42	23.29	11.99	7.63-78.13	7.63	24.48	26.97	13.86	17.06	34.90
Sil seg	2-20	8.00	8.00	9.66	4.92	2-24	7.00	9.00	10.16	5.28	0-24	10.00	7.00	7.71	4.88	0-24	7.00	8.00	9.14	5.09	5.75	12.00
Sil time	2.90-74.01	2.90	24.05	25.79	15.83	4.23-58.41	4.23	19.36	24.22	14.18	0-63.12	0	17.68	22.46	16.01	0-74.01	0	19.92	24.12	15.28	11.81	33.68
Loc	17.94-103.13	17.94	51.07	55.83	21.73	19-97.90	19.00	47.04	52.02	22.13	17.13-119.31	17.13	43.48	45.76	20.93	17.13-119.31	17.13	45.70	51.09	21.77	37.22	62.59
Utt prop	15.23-92.69	15.22	55.95	55.02	19.92	27.05-89.21	27.05	53.18	54.50	15.37	18.90-100	18.90	53.93	53.32	19.68	15.22-100	15.22	53.97	54.26	18.29	41.29	66.34

	HC					MCI					AD					All participants						
	Range	Mode	Median	M	SD	Range	Mode	Median	M	SD	Range	Mode	Median	M	SD	Range	Mode	Median	M	SD	P. 25	P. 75
Imm																						
Sil prop	7.31-84.78	7.31	44.05	44.98	19.92	10.79-72.95	10.79	46.82	45.50	15.37	0-81.10	0	46.07	46.68	19.68	0-84.78	0	46.03	45.74	18.29	33.66	58.71
Ps rate	0.04-0.27	0.04	0.17	0.17	0.05	0.07-0.26	0.07	0.20	0.19	0.04	0-0.23	0	0.18	0.16	0.06	0-0.27	0	0.18	0.18	0.05	0.15	0.22
Hes rate	0.14-1	0.67	0.67	0.63	0.23	0.2-1	0.75	0.69	0.66	0.18	0-1	0.54	0.59	0.59	0.20	0-1	0.67	0.64	0.62	0.21	0.50	0.76
Del																						
Utt seg	1-44	9.00	12.00	13.94	8.42	5-27	6.00	8.00	10.75	5.83	4-28	12.00	12.50	13.50	6.17	1-44	7.00	12.00	12.74	6.96	8.00	16.00
Utt time	2.85-77.60	2.85	26.18	32.68	18.59	9.15-57.94	9.15	21.96	23.58	12.69	2.17-44.33	2.17	24.97	24.29	9.99	2.17-77.60	2.17	23.67	26.80	14.59	17.42	32.46
Sil seg	0-29	4.00	6.00	7.25	5.91	0-16	5.00	4.50	5.06	4.10	0-21	7.00	7.00	7.38	5.30	0-29	5.00	5.00	6.58	5.22	3.00	9.00
Sil time	0-58.05	0	12.18	16.64	14.38	0-57.36	0	7.74	12.51	13.02	0-63.02	0	15.89	19.21	16.58	0-63.02	0	11.45	16.18	14.89	5.79	22.31
Loc	10.23-135.65	10.23	41.47	49.32	26.16	17.52-83.96	17.52	30.64	36.08	18.10	15.92-95.38	15.92	40.81	43.50	20.89	10.23-135.65	10.23	37.23	42.98	22.38	26.73	54.66
Utt prop	6.34-100	6.34	70.91	67.90	19.82	26.39-100	100.00	74.35	68.76	21.80	7.17-100	100.00	63.03	60.93	23.92	6.34-100	100.00	69.04	65.76	22.02	50.55	80.57
Sil prop	0-93.66	0	29.09	32.10	19.82	0-73.61	0	25.65	31.24	21.80	0-92.83	0	36.97	39.07	23.92	0-93.66	0	30.96	34.24	22.02	19.43	49.45
Ps rate	0-0.26	0	0.13	0.14	0.06	0-0.32	0	0.13	0.13	0.07	0-0.41	0	0.17	0.16	0.08	0-0.41	0	0.15	0.14	0.07	0.10	0.19
Hes rate	0-1.67	0.38	0.47	0.52	0.30	0-1	0.43	0.43	0.46	0.27	0-1.25	0	0.54	0.53	0.31	0-1.67	0	0.47	0.50	0.29	0.29	0.67
SIT																						
Utt seg	43-130	57.00	65.50	67.75	18.75	44-151	62.00	70.50	77.13	22.10	35-122	62.00	73.00	74.79	18.09	35-151	62.00	69.00	73.26	19.89	60.75	79.25
Utt time	90.71-496.61	90.71	194.77	211.39	74.53	119.57-384.86	119.57	200.28	215.12	57.99	103.68-364.08	103.68	199.56	201.36	54.62	90.71-496.61	90.71	200.18	209.13	62.45	168.42	246.18
Sil seg	4-59	14.00	18.00	19.84	12.16	6-58	15.00	19.50	22.28	13.17	7-46	13.00	21.00	22.56	10.78	4-59	13.00	19.00	21.58	11.99	13.00	26.25
Sil time	5.18-95.74	5.18	34.07	35.34	23.30	6.45-110.21	6.45	35.08	39.30	26.39	9.97-85.99	9.968	33.63	38.62	21.11	5.18-110.21	5.18	34.24	37.77	23.46	20.49	48.57

	HC					MCI					AD					All participants						
	Range	Mode	Median	M	SD	Range	Mode	Median	M	SD	Range	Mode	Median	M	SD	Range	Mode	Median	M	SD	P. 25	P. 75
<u>SIT</u>																						
Loc	106.58-503.16	106.58	233.26	246.73	72.81	135.86-485.43	135.86	240.09	254.42	70.41	130.54-415.89	130.54	233.94	239.98	56.71	106.58-503.16	106.58	236.57	246.90	66.40	205.13	277.77
Utt prop	62.08-98.70	62.08	85.30	84.96	8.97	63.51-97.26	63.51	85.61	84.89	8.05	57.64-95.97	57.64	85.58	83.58	9.02	57.64-98.70	57.64	85.50	84.46	8.63	79.42	90.77
Sil prop	1.30-37.92	1.30	14.70	15.04	8.97	2.74-36.49	2.74	14.39	15.11	8.05	4.031-42.36	4.03	14.42	16.42	9.02	1.30-42.36	1.30	14.50	15.54	8.63	9.23	20.58
Ps rate	0.01-0.19	0.01	0.09	0.08	0.04	0.03-0.19	0.03	0.09	0.09	0.04	0.03-0.18	0.03	0.09	0.09	0.04	0.01-0.19	0.01	0.09	0.09	0.04	0.06	0.12
Hes rate	0.07-0.52	0.07	0.28	0.28	0.12	0.10-0.52	0.10	0.27	0.28	0.11	0.10-0.51	0.27	0.31	0.30	0.12	0.07-0.52	0.23	0.28	0.29	0.11	0.19	0.37

Note: TPD = Thai Picture Description task; Imm = immediate recall task; Del = delayed recall task; SIT = Semi-structured Interview for Thai; Utt seg = number of utterance segments; Utt time = total duration of utterance; Sil seg = number of silent segments; Sil time = total duration of silences; Loc = locution; Utt prop = utterance proportion; Sil prop = silence proportion; Ps rate = pause rate; Hes rate = hesitation rate.

Table 34 Pearson’s correlation of age, years of education, and the primary temporal variables

	Age	YoE	MoCA	PD Utt seg	PD Utt time	PD Sil seg	PD Sil time	IM Utt seg	IM Utt time	IM Sil seg	IM Sil time	DL Utt seg	DL Utt time	DL Sil seg	DL Sil time	SI Utt seg	SI Utt time	SI Sil seg	SI Sil time
Age	1.00																		
YoE	-.17	1.00																	
MoCA	-.26**	.33**	1.00																
PD Utt seg	.06	.13	.12	1.00															
PD Utt time	.03	.25*	.13	.72**	1.00														
PD Sil seg	.05	.01	.12	.78**	.31**	1.00													
PD Sil time	.03	-.07	.05	.43**	-.04	.79**	1.00												
IM Utt seg	-.05	.01	.20*	.34**	.18	.30**	.14	1.00											
IM Utt time	.00	.09	.22*	.35**	.52**	.08	-.16	.77**	1.00										
IM Sil seg	-.04	-.03	.16	.24*	-.03	.37**	.30**	.85**	.46**	1.00									
IM Sil time	-.01	-.08	.06	.06	-.20	.30**	.44**	.56**	.11	.82**	1.00								
DL Utt seg	.08	-.17	.08	.26*	.12	.31**	.18	.53**	.40**	.500	.41**	1.00							
DL Utt time	-.01	.07	.30**	.38**	.40**	.23*	-.02	.51**	.61**	.32**	.12	.78**	1.00						
DL Sil seg	.13	-.23*	.00	.10	-.08	.30**	.30**	.35**	.15	.46**	.50**	.81**	.38**	1.00					
DL Sil time	.19	-.27**	-.14	.03	-.15	.27**	.36**	.24*	.01	.40**	.55**	.62**	.15	.90**	1.00				
SI Utt seg	.05	.06	-.09	.13	-.02	.19	.17	.12	-.01	.10	.10	.11	.00	.10	.06	1.00			
SI Utt time	-.01	.22*	.09	.14	.31**	.01	-.06	.07	.25*	-.07	-.13	.00	.17	-.10	-.18	.56**	1.00		
SI Sil seg	.05	-.05	-.05	-.01	-.25*	.20*	.29**	.12	-.09	.20	.26**	.17	-.08	.28**	.30**	.72**	.08	1.00	
SI Sil time	.05	-.07	-.02	-.06	-.28**	.17	.28**	.08	-.14	.19	.28**	.14	-.12	.29**	.33**	.61**	-.01	.95**	1.00

* $p < .05$, ** $p < .01$

Note: PD = Thai Picture Description task; IM = immediate recall task; DL = delayed recall task; SI = Semi-structured Interview for Thai; Utt seg = number of utterance segments; Utt time = total duration of utterance; Sil seg = number of silent segments; Sil time = total duration of silences.

Table 35 Pearson’s correlation of age, years of education, and the transformed temporal variables

	Age	YoE	MoCA	PD Loc	PD Utt prop	PD Sil prop	PD Ps rate	PD Hes rate	IM Loc	IM Utt prop	IM Sil prop	IM Ps rate	IM Hes rate	DL Loc	DL Utt prop	DL Sil prop	DL Ps rate	DL Hes rate	SI Loc	SI Utt prop	SI Sil prop	SI Ps rate	SI Hes rate
Age	1.00																						
YoE	-.17	1.00																					
MoCA	-26**	.33**	1.00																				
PD Loc	.04	.17	.14	1.00																			
PD Utt prop	-.06	.28**	.12	.11	1.00																		
PD Sil prop	.06	-.28**	-.12	-.11	-1.00**	1.00																	
PD Ps rate	.08	-.23*	-.06	-.11	-.65**	.65**	1.00																
PD Hes rate	.08	-.22*	-.01	-.09	-.88**	.88**	.75**	1.00															
IM Loc	.00	.00	.18	.29**	.00	.00	.03	.01	1.00														
IM Utt prop	.03	.14	.07	.18	.67**	-.67**	-.44**	-.60**	-.19	1.00													
IM Sil prop	-.03	-.14	-.07	-.18	-.67**	.67**	.44**	.60**	.19	-1.00**	1.00												
IM Ps rate	-.09	.01	.06	-.16	-.32**	.32**	.57**	.32**	.14	-.52**	.52**	1.00											
IM Hes rate	.01	-.01	.08	-.17	-.49**	.49**	.46**	.51**	.16	-.83**	.83**	.68**	1.00										
DL Loc	.12	-.13	.11	.27**	-.09	.09	.18	.12	.57**	-.10	.10	.03	.05	1.00									
DL Utt prop	-.11	.30**	.23*	.07	.55**	-.55**	-.32**	-.52**	-.07	.53**	-.53**	-.09	-.30**	-.30**	1.00								
DL Sil prop	.11	-.30**	-.23*	-.07	-.55**	.55**	.32**	.52**	.07	-.53**	.53**	.09	.30**	.30**	-1.00**	1.00							
DL Ps rate	.05	-.22*	-.10	-.08	-.42**	.42**	.34**	.44**	.01	-.37**	.37**	.17	.24*	.24*	-.75**	.75**	1.00						
DL Hes rate	.12	-.20*	-.10	-.04	-.45**	.45**	.26*	.48**	.03	-.44**	.44**	.03	.28**	.25*	-.92**	.92**	.77**	1.00					
SI Loc	.00	.18	.08	.19	.14	-.14	-.18	-.13	.10	.12	-.12	-.14	-.12	.04	.12	-.12	-.01	-.05	1.00				
SI Utt prop	-.03	.17	.06	.20*	.50**	-.50**	-.34**	-.42**	-.05	.46**	-.46**	-.24*	-.30**	-.10	.47**	-.47**	-.35**	-.36**	.11	1.00			
SI Sil prop	.03	-.17	-.06	-.20*	-.50**	.50**	.34**	.42**	.05	-.43**	.46**	.24*	.30**	.10	-.47**	.47**	.35**	.36**	-.11	-1.00**	1.00		
SI Ps rate	.05	-.15	-.10	-.19	-.48**	.48**	.38**	.41**	.06	-.41**	.41**	.25*	.28**	.11	-.46**	.46**	.36**	.33**	-.10	-.92**	.92**	1.00	
SI Hes rate	.06	-.14	-.05	-.15	-.45**	.45**	.32**	.40**	.11	-.37**	.37**	.19	.26*	.16	-.44**	.44**	.34**	.33**	-.02	-.92**	.92**	.95**	1.00

* $p < .05$, ** $p < .01$

Note: PD = Thai Picture Description task; IM = immediate recall task; DL = delayed recall task; SI = Semi-structured Interview for Thai; Loc = locution; Utt prop = utterance proportion; Sil prop = silence proportion; Ps rate = pause rate; Hes rate = hesitation rate.

One-way ANOVA was performed to explore the difference between groups across 36 variables. Table 36 reports the ANOVA of all temporal features in each block; only one variable is found to have a significant difference. The total duration of utterance in the delayed recall is significantly different between the three groups (*F* (2) = 4.13, *p* = .02), with the highest mean time for HC (*M* = 32.68), followed by AD (*M* = 24.29) and MCI (*M* = 23.58). The post hoc test by Turkey shows that HC differs significantly from both MCI (*p* < .05) and HC (*p* < .05) while MCI and AD do not differ (*p* = .98).

Table 36 *P*-values in ANOVA of the temporal variables

Variable	TPD	Imm	Del	SIT
Utterance segments	.23	.13	.14	.14
Utterance time	.50	.13	<*.05*	.65
Silence segments	.28	.12	.13	.61
Silence time	.43	.68	.18	.77
Locution	.35	.16	.06	.68
Utterance proportion	.41	.93	.29	.77
Silence proportion	.41	.93	.29	.77
Pause rate	.85	.07	.37	.55
Hesitation rate	.96	.34	.53	.71

Note: TPD = Thai Picture Description task; Imm = immediate recall task; Del = delayed recall task; SIT = Semi-structured Interview for Thai.

Due to the significant differences in education level among three groups of the participants, the following analyses were conducted to examine the effect of education on the participant’s temporal variables. One-way ANOVA was performed to examine the trend of the utterance time in Del in each education group. Table 37 shows that the total utterance time in Del is significantly different among the three groups of participants only in the group of high school (*F*(2) = 5.63, *p* <.05, η_p^2 = .37).

Table 37 ANOVA of the temporal variables separated by education three-level

Variable	Education Group	*df*	*F*	*p*	η_p^2
Utterance time in the delayed recall task	Primary school	2	0.51	.60	.03
	High school	2	*5.63*	*<.05*	.37
	University	2	0.07	.93	.00

4.2.3.2 Multivariate analysis of variance of the acoustic features

MANOVA was conducted to examine possible differences among the three groups with respect to the acoustic features. The acoustic variables were selected with theoretical statistical support. According to Spearman's correlation, group performance significantly correlates with six acoustic variables, including the number of voice breaks in Del ($r = -.24, p < .05$), utterance duration in TPD ($r = -.20, p < .05$), number of utterance segments in Imm ($r = -.20, p < .05$), utterance duration in Imm ($r = -.21, p < .05$), locution in Imm ($r = -.21, p < .05$), and number of utterance segments in SIT ($r = -.21, p < .05$). Since the ANOVA (see Table 36) showed that utterance duration in Del is significantly worst in the MCI and AD groups than in the HC group, this variable should also be included in this analysis. The seven acoustic variables were included in the MANOVA without covariate.

Three assumptions of multivariate analysis were primarily examined (see section 4.2.2.2 for details about the underlying logic). The results are reported below.

a) Independence

Table 38 reports the univariate test and interaction effect of time of the day, test sequence, and seven acoustic variables. There are no direct or interaction effects of the time of the day and test sequence on the seven acoustics features. In conclusion, the data were independent and not influenced by extraneous factors.

Table 38 Univariate two-way ANOVA of the acoustic features

Variable	Time of the day		Test sequence		Interaction effect	
	F	*p*	*F*	*p*	*F*	*p*
<u>Frequency domain</u>						
Number of voice breaks in Del	0.48	.49	0.00	1.00	0.21	.65
<u>Temporal domain</u>						
Total utterance time in TPD	0.05	.82	0.00	1.00	1.61	.21
Total utterance time in Imm	0.85	.36	1.47	.23	1.78	.19
Total utterance time in Del	0.06	.80	0.05	.82	0.57	.45
Number of utterance segments in Imm	1.03	.31	2.26	.14	0.00	.96
Number of utterance segments in SIT	0.06	.80	1.11	.30	1.84	.18
Locution in Imm	0.59	.45	1.57	.21	0.03	.86

Note: TPD = Thai Picture Description task; Imm = immediate recall task; Del = delayed recall task; SIT = Semi-structured Interview for Thai.

b) Normality

Table 39 presents the Shapiro-Wilk test for normal distribution. All variables are significantly different from a normal distribution. However, the absolute value of skewness and kurtosis of locution in IM is not exceeded 1; the data in this variable distributes approximately normally (Mishra et al., 2019). The number of voice breaks in DL shows extreme values of skewness and kurtosis; this variable would not be included in further analysis. The non-normal distributed data might cause a slight impact, so the Type I error should be a concern for an interpretation (Hair et al., 2014, p. 686). Furthermore, the Box's M test can reflect a severe violation of normality distribution which can be considered in the following section. Eventually, six variables in the temporal domain were included in the MANOVA and the MDA.

Table 39 Normality by Shapiro-Wilk test for the acoustic features as predictors

Variables	Skewness	Kurtosis	*Shapiro-Wilk*	*p*
Frequency domain				
Number of voice breaks in Del	2.73	11.47	*0.76*	*<.001*
Temporal domain				
Total utterance time in TPD	1.09	1.03	*0.92*	*<.001*
Total utterance time in Imm	1.18	1.56	*0.92*	*<.001*
Total utterance time in Del	1.27	1.80	*0.91*	*<.001*
Number of utterance segments in Imm	1.23	1.70	*0.91*	*<.001*
Number of utterance segments in SIT	1.27	2.34	*0.91*	*<.001*
Locution in Imm	0.99	0.67	*0.93*	*<.001*

Note: TPD = Thai Picture Description task; Imm = immediate recall task; Del = delayed recall task; SIT = Semi-structured Interview for Thai.

c) Equality of Variance–Covariance Matrices

The six acoustic features (only temporal domain) used for the MONOVA were examined by the Box’s M test. Table 40 shows a nonsignificant value (p = .09) of the Box’s M test. This result suggests equal covariance matrices of the six dependent variables among the three groups.

Table 40 Box's test of equality of covariance matrices for the acoustic features.

Set of variables	*Box's M*	*F*	*df1*	*df2*	*p*
Six temporal variables	60.55	1.31	42	26626.96	.09

Since all the qualified acoustic variables are in the temporal domain, the set of variables used in the MONOVA is called six temporal variables. The multivariate test in Table 41 shows that the set of temporal variables size do not differ among the three groups (F (12, 182) = 1.31, p = .09). The univariate test reported in Table 42 shows that only total time of utterance in Del is significantly different among the three groups (F (2) = 4.13, p <.05, η_p^2 = .08). Since the three groups differ only with respect to the main effect of only one acoustic feature, MDA would not be conducted.

Table 41 Multivariate tests for the acoustic features

Set of variables	*Pillai's Trace*	*F*	*Hypothesis df*	*Error df*	*p*	η_p^2
Six temporal variables	0.19	1.60	12.00	182.00	.09	.10

Table 42 Tests of between-subjects effects for the acoustic features

Variables	*df*	*F*	*p*	η_p^2
Total utterance time in TPD	2	0.71	.47	.02
Total utterance time in Imm	2	2.09	.13	.04
Total utterance time in Del	2	*4.13*	*<.05*	.08
Number of utterance segments in Imm	2	2.08	.13	.04
Number of utterance segments in SIT	2	1.97	.15	.04
Locution in Imm	2	1.84	.16	.04

Note: TPD = Thai Picture Description task; Imm = immediate recall task; Del = delayed recall task; SIT = Semi-structured Interview for Thai.

4.2.3.3 Receiver Operating Characteristic curve analysis for acoustic features

The ROC and AUC examine the discriminant ability of only one acoustic variable, which is utterance time in Del. The cut-off scores for utterance time are selected based on the Youden Index. The optimal cut-off time in seconds is reported below.

i. HC vs. MCI

The ROC of the utterance time in Del in detecting MCI among non-AD groups is illustrated in Figure 24. The AUC of this pair is .66 (p <.05, 95 % CI, .53–.79), which is lower than an acceptable level (Mandrekar, 2010). The optimal utterance time in differentiating MCI from HC is 23.25 sec. This duration obtains the sensitivity and specificity of 62.50 %.

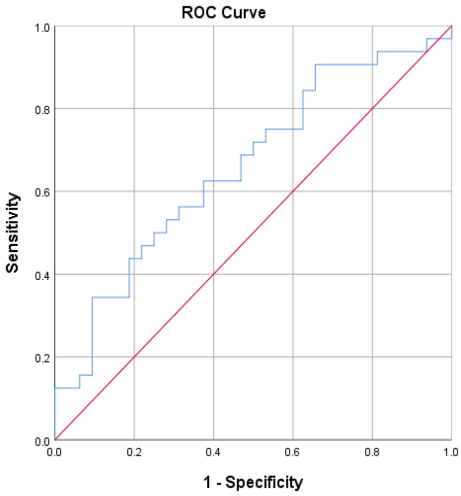


Figure 24 AUC of the total utterance time in Del in detecting MCI among non-AD groups

ii. MCI vs. AD

Figure 25 displays the ROC of utterance time in Del for a pair of MCI and AD. Since the mean time of utterance in Del of AD (M = 24.29, SD = 9.99) is higher than MCI (M = 23.58, SD = 12.69), the curve lies under the reference line. The AUC is only .42, which indicates no difference from 50% by chance in distinguishing AD from MCI (*p* = .27, 95 % CI, .28–.56). Hence, the total utterance time in DL is not appropriate for differentiating AD from MCI.

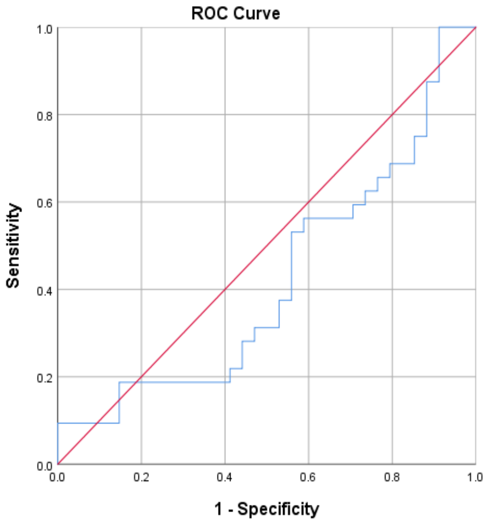


Figure 25 AUC of the total utterance time in Del in detecting AD among cognitive impairment groups

iii. HC vs. AD

In differentiating AD from HC, the utterance duration in Del obtains the AUC of .61 (*p* = .14, 95 % CI, .47–.74), as shown in Figure 26. The discriminant ability of this variable does not significantly differ from 50% by chance. Therefore, the utterance time in DL alone does not differentiate HC and AD.

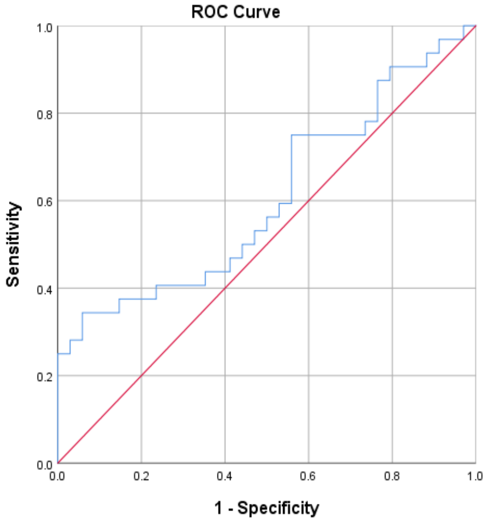


Figure 26 AUC of the total utterance time in Del in differentiating AD from HC

iv. HC vs. Cognitive impairment

Figure 27 portrays the ROC corresponding to the utterance time in Del for distinguishing HC and persons with cognitive impairment (MCI and AD). Although the AUC of this pair is significantly different from 50%, the value of .63 is below an acceptable level (*p* <.05, 95 % confidence interval (CI), .51–.75). Moreover, the optimal cut-off score regarding Youden’s Index achieved very low sensitivity at 34.38% with a specificity of 92.42%.

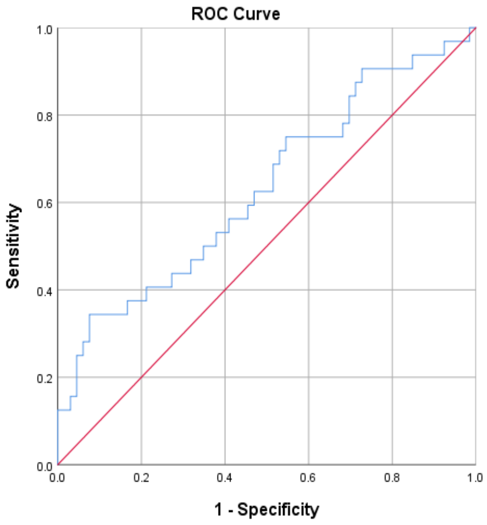


Figure 27 AUC of the total utterance time in Del in detecting the persons with cognitive impairment among the total sample

4.2.4 All variables as predictors

In response to the third objective of this study, the potential variables were validated by their ability to classify HC, MCI, and AD. All possible predictive variables were entered into the stepwise MDA, then evaluated discriminant ability by ROC and AUC analysis. Finally, all variables and the discriminant model, which significantly differentiates three groups of participants, were compared for their accuracy and appropriateness in detecting three pathological stages.

4.2.4.1 Discriminant analysis with the stepwise estimation of the potential predictors

MDA with stepwise estimation has the objective of obtaining the model of the best set of predictors for predicting dementia pathological stages. The predictive variables were selected based on theoretical background, previous literature, and the statistical results of this study. The eight predictors are years of education, iADL, hypertension, hyperlipidemia, BDS, CIU, Del, and total utterance time in DL. The former four variables are demographic and health characteristics. These four variables are significantly different among the three groups of participants. Three content variables, i.e., BDS, CIU, and Del, obtain the three highest discriminant loadings, and their coefficients are higher than .5 (see Table 28). Then only one acoustic feature included is the total utterance time in Del.

Stepwise discriminant analysis was performed to find the best set of predictors among the eight aforementioned variables. The entering and removing criteria were determined by F values between 2.71 – 3.84. The Box's M in Table 43 reveals a non-significant value ($p = .24$); thus, equality of the variance-covariance matrix can be assumed. Table 44 presents the Stepwise statistics of two steps which means two functions are estimated. A tolerance of .99 confirms a lack of multicollinearity among two variables (Hair et al., 2014, p. 234). The qualified variables are BDS and CIU. The discriminant function 1 obtains the eigenvalue of 0.65, and this model accounts for 93.2% of the explained variance (see Table 45). The canonical correlation of this model is .63. The square of this coefficient is .39, which indicates that 39.69% of the variance in the dementia pathological variable is explained by the discriminant function 1. The Wilks' lambda in the same table indicates a significant value of two functions together (χ^2 (4) = 36.95, $p < .001$). However, function two alone is unable to discriminate among the three groups of participants (χ^2 (1) = 3.12, $p < .001$). Only the discriminant function 1 of two variables significantly differentiates between the three groups of dementia pathological stages. The discriminant function can be written as the following equation. The group centroids regarding this equation are 1.10 for HC, –0.23 for MCI, and –0.77 for AD.

$$\textit{Discriminant scores} = 0.48(BDS) + 0.03(CIU) - 3.23$$

Table 43 Box's test of equality of covariance matrices for the stepwise estimation of the eight predictors.

Box's M	*F*	*df1*	*df2*	*p*
8.319	1.33	6	101563.83	0.241

Table 44 Variables entered in the analysis for the stepwise estimation

Step	Variables entered	Tolerance	F to Remove	Wilks' Lambda
1	BDS	1.00	11.53	
2	BDS	0.99	11.12	0.77
	CIU	0.99	9.75	0.75

Note: BDS = backward digit span; CIU = correct information unit

Table 45 Summary of canonical discriminant functions for the stepwise estimation

Function	Eigenvalue	Percentages of Variance	Canonical Correlation	After step	*Wilks' Lambda*	χ^2	*df*	*p*
				0	0.58	*36.96*	4	*<.001*
1	0.65	93.21	0.63	1	0.96	3.12	1	0.08
2	0.05	6.79	0.21					

Figure 28 depicts the scatter plot of discriminant values and group centroids of the three groups. The level of accuracy in classification regarding the discriminant function of BDS and CIU is 61.1 % (see Table 46). The accuracy varies in each group, i.e., it is 73.9% in HC, 36.4% in MCI, and 70.4% in AD. The classified accuracy of MCI is very low.

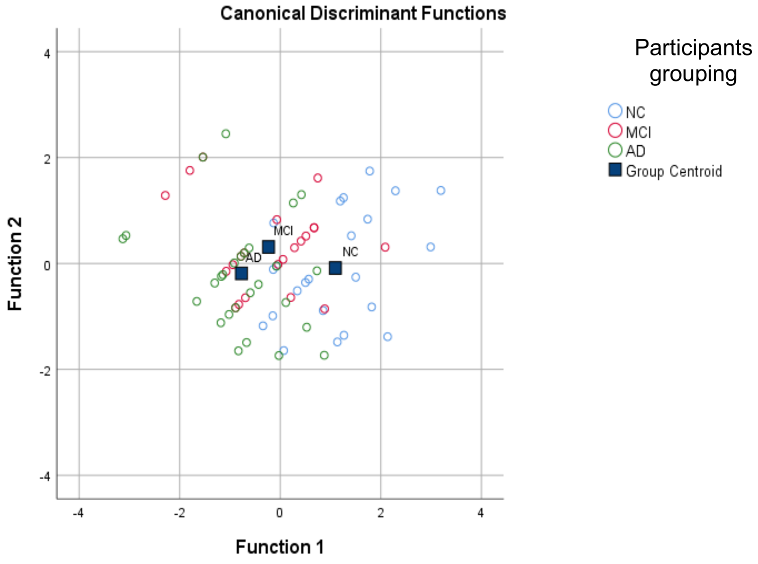


Figure 28 Group centroids of canonical discriminant functions of the stepwise estimation

Table 46 Classification results of the discriminant functions for the stepwise estimation

Actual groups membership	Predicted Group Membership			Total
	HC	MCI	AD	
HC (*n*)	17	2	4	23
MCI (*n*)	4	8	10	22
AD (*n*)	3	5	19	27
HC (%)	73.9	8.7	17.4	100.0
MCI (%)	18.2	36.4	45.5	100.0
AD (%)	11.1	18.5	70.4	100.0

61.1% of original group cases correctly classified.

4.2.4.2 Receiver Operating Characteristic curve analysis for the potential predictors

ROC and AUC were performed to evaluate the discriminant ability of the discriminant scores generated from the model of BDS and. The optimal cut-off scores of each pair are suggested based on Youden Index.

i. HC vs. MCI

An excellent level of AUC is found in differentiating HC and MCI, with an AUC of .81 (*p* <.001, 95 % CI, .68 – .93) (see Figure 29). The optimal cut-off discriminant scores reveal a sensitivity of 56.52% and a specificity of 95.46%. Although the specificity is at an outstanding level, the sensitivity is lower than an acceptable level.

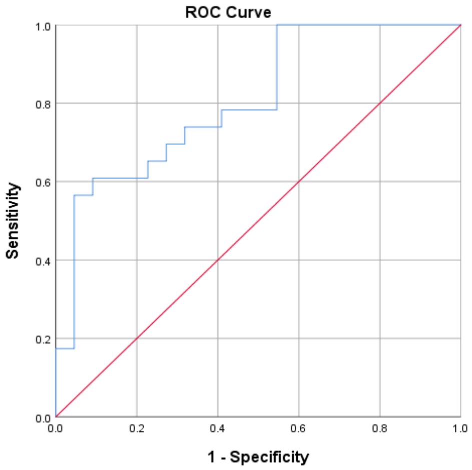


Figure 29 AUC of the discriminant model of BDS and CIU in detecting MCI among non-AD groups

ii. MCI vs. AD

The AUC in discriminating AD from MCI is lower than the acceptable level and shows no difference with the randomly differentiating these two groups of participants (see Figure 30). The ROC regarding the discriminant function of BDS and CIU provides the AUC of .64 (p = .11, 95 % CI, .48 – .79).

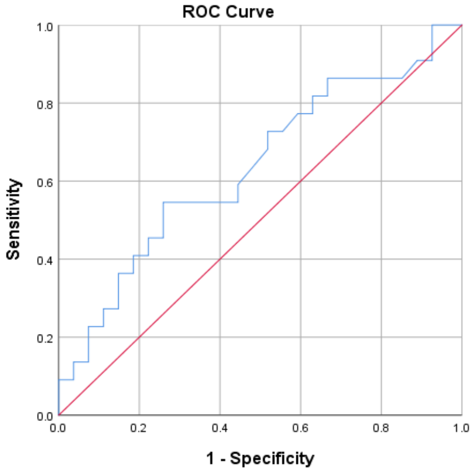


Figure 30 AUC of the discriminant model of BDS and CIU in detecting AD among the cognitive impairment groups

iii. HC vs. AD

Discriminant ability in differentiating HC and AD is at an outstanding level, with an AUC of .91 (p <.001, 95 % CI, .84 – .99) (see Figure 31). The optimal cut-off point is determined by the highest Youden Index at 1.70. This point provides a sensitivity of 100% and a specificity of 70.37%.

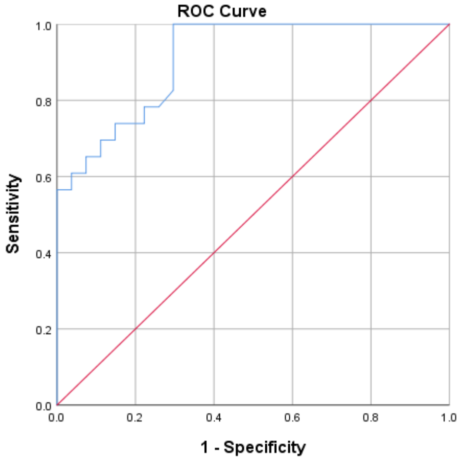


Figure 31 AUC of the discriminant model of BDS and CIU in differentiating AD from HC

iv. HC vs. Cognitive impairment

Considering MCI and AD as a cognitive impairment group yelds a slightly higher AUC values than differentiating HC and MCI (see Figure 32). The discriminant scores of the BDS and CIU model can significantly discriminate HC and people with cognitive impairment by showing an AUC of .86 (p <.001, 95 % CI, .78 – .95). The coordinate points of the ROC curve obtain the highest Youden Index at 1.59. At this optimal cut-off score, the sensitivity is 100%, and the specificity is 59.18%.

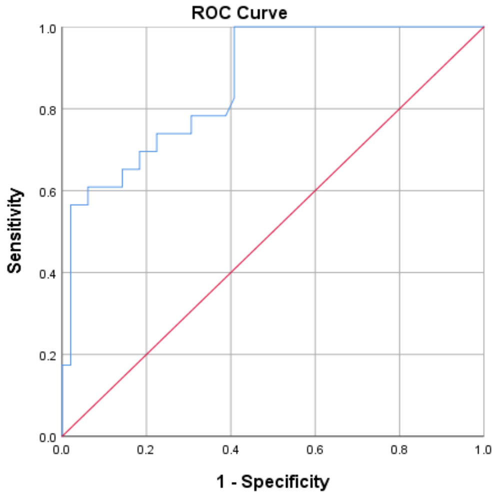


Figure 32 AUC of the discriminant model of BDS and CIU in detecting the persons with cognitive impairment among the total sample

4.2.4.3 Potential predictors

Table 47 demonstrates the psychometric index of three predictors, including the discriminant function of six neuropsychological and content variables, i.e., Del, CIU, LF, CF, FDS, and BDS (from Section 4.2.2.3), total utterance time in Del (from Section 4.2.3.3), and the discriminant function of BDS and CIU (from Section 4.2.4.1). This section explores the best differentiating model for four pairs of comparison, i.e., HC vs MCI, MCI vs AD, HC vs AD, and HC vs (MCI + AD: the cognitive impairment group). The AUC, sensitivity, specificity, and YI are expected to show the highest values. Furthermore, the number of variables or assessments is also considered. The lesser number of variables is considered as a suggested method. The suggested model in each of the following pairs is accordance with the criteria of predictive test, which requires the sum of sensitivity and specificity to be at least 1.5 (Power et al., 2013).

i. HC vs. MCI

The six variables model (AUC = .84) obtained slightly higher AUC than the BDS and CIU model (AUC = .81). Although the specificity of the six-variables model (90.91%) is slightly lower than the BDS and CIU model (95.46%), the sensitivity of the six models (73.68%) is distinctively higher than the two (56.52%) and considered as an acceptable level. The value of sensitivity and specificity of the discriminant model of six variables is 1.65, which is higher than 1.5 (Power et al., 2013). Therefore, the six-variable model is suitable for differentiating HC and MCI.

ii. MCI vs. AD

Differentiating MCI and AD obtains a low overall psychometric index in comparison with the other three comparisons. The only significant AUC of this pair is found in the six-variable model with an AUC of .70. This value is just at an acceptable level. Even though the sensitivity of this model is acceptable, specificity at 61.54% is at a low level. Considering the scores of test performance, the sum of sensitivity and specificity is only 1.43. This model is not appropriate for implementation.

iii. HC vs. AD

An outstanding AUC is found in differentiating HC and AD by the discriminant function of neuropsychological and content variables (AUC = .92) and the discriminant function of BDS and CIU (AUC = .91). Sensitivity of the BDS and CIU model (100%) is higher than the neuropsychological and content variables model (94.74%). The specificity of the BDS and CIU model (70.37%) is lower than the neuropsychological and content variables model (80.77%). This phenomenon is caused by the trade-off property of sensitivity and specificity (Larner, 2017). However, the YI of the neuropsychological and content variables (0.75) is slightly higher than the BDS and CIU model (0.70). These two models are marginally different in the aforementioned psychometric indexes, but the number of variables in the two models are quite comparable. According to the number of variables, the model of BDS and CIU requires lesser tests than the model of neuropsychological and content variables, which need six assessments. In addition, the test performance value is 1.7, higher than the suggested value of 1.5 (Power et al., 2013). The discriminant function of BDS and CIU is thus the best model for differentiating HC and AD.

iv. HC vs. Cognitive impairment

Two predictors present marginal differences in AUC and Youden Index, i.e., the six-variable model (AUC = .88, YI = .57) and the model of BDS and CIU (AUC = .86, YI = .59). However, the discriminant function of BDS and CIU demands fewer number of variables. Hence, this model is appropriate for detecting a person with cognitive impairments out of HC. Nevertheless, the specificity is very low (59.18%), and the test performance value is 1.59 due to the 100% sensitivity.

Table 47 AUC, sensitivity and specificity of the potential predictors

Comparison groups	AUC	Sensitivity (%)	Specificity (%)
The discriminant function of neuropsychological and content variables			
HC vs. MCI	.84**	73.68	90.91
MCI vs. AD	.70*	81.82	61.54
HC vs. AD	.92**	94.74	80.77
HC vs. Cognitive impairment	.88**	94.74	62.50
Total utterance time in DL			
HC vs. MCI	.66*	62.50	62.50
MCI vs. AD	.42	*NA*	*NA*
HC vs. AD	.61	*NA*	*NA*
HC vs. Cognitive impairment	.63*	34.38	92.42
The discriminant function of BDS and CIU			
HC vs. MCI	.81**	56.52	95.46
MCI vs. AD	.64	*NA*	*NA*
HC vs. AD	.91**	100.00	70.37
HC vs. Cognitive impairment	.86**	100.00	59.18

* $p < .05$, ** $p < .001$, *NA* = not available

In conclusion, the six-variables model is suitable for differentiating HC and MCI. However, the sensitivity of this model is lower than 80%, which is the level considered acceptable. When distinguishing AD and HC, and persons with cognitive impairments and HC, the discriminant function of BDS and CIU is the best model. However, specificity in detecting persons with cognitive impairment from HC is sub-optimal. Finally, there is no appropriate predictor in differentiating MCI and AD.

4.2.5 General conclusion

This study developed three spontaneous speech tasks, including the Thai Picture description (TPD), Thai Story Recall (TSR), and Semi-structured Interview for Thai (SIT) from which several dependent variables were extracted. The three tasks receive one score for the overall content validity index (CVI). Although trying to get equivalent groups on demographic variables, the actual sample shows significant differences among the three experimental groups (HC, MCI, and AD) for years of education, with AD having fewer years, followed by MCI and HC. As for the administered neuropsychological tests, i.e., MoCA, letter fluency, category fluency, forward digit span, and backward digit span. ANOVAs show significant differences among three groups of participants in the respected direction.

Content variables were derived from two spontaneous speech tasks, TPD and TSR. Two of the variables show significant differences among three groups of participants, i.e., correct information unit (CIU) and delayed recall scores (Del). These two variables significantly correlate with the MoCA at a moderate positive level. Three sets of MONOVA were further performed, i.e., the combination of content variables and neuropsychological tests, CIU and two tests of verbal fluency, and Del and two tests of digit span. All of them show significant main effects of the dementia pathological stage. Furthermore, MDA was conducted. The discriminant function of the content variables (i.e., CIU and Del) and four neuropsychological tests (i.e., letter fluency, category fluency, forward digit span, and backward digit span) provides 65.7% of correct classification. In the ROC and AUC analyses, the discriminant function from MDA yields an acceptable level of AUC in differentiating three theoretically relevant comparisons i.e., HC vs. MCI, HC vs. AD, and HC vs. Cognitive impairment (AD+MCI).

Acoustic features did not show the expected discriminant results. The descriptive statistics in both the frequency-related and the temporal domains are found to have no consistent tendency among the three groups. None of the frequency-related variables has a correlation with the MoCA while three variables in the temporal domain show a significant positive correlation with the MoCA, i.e., number of utterance segments in immediate recall, total utterance time in immediate recall, and total utterance time in delayed recall. ANOVAs reveal that three acoustic features are significantly different among the three groups, i.e., number of voice breaks in the TPD, number of voice breaks in the SIT, and total utterance time in delayed recall. A MONOVA conducted with six qualified variables did not yield significant results. Hence, ROC and AUC analysis was only examined for total utterance time in delayed recall. This temporal variable fails to reach an acceptable level of AUC in differentiating any pairs of the participant groups.

To find the optimal combination of predictive variables, a stepwise estimation in MDA was conducted. The result shows that the model combining CIU, and backward digit span provides 61.1% of correct classification. The potential predictors were examined by ROC and AUC analysis. The six-variables discriminant function is suitable for differentiating HC and MCI. The discriminant function of CIU and backward digit span is suitable for detecting AD and persons with cognitive impairments, differentiating them from HC. However, there is no appropriate predictor in differentiating MCI and AD.

CHAPTER 5
DISCUSSION

This chapter consists of five sections. The first section summarizes the findings in this study. In the second section, the research findings are discussed in relation to the three objectives of this research. Then the limitations are presented. The implication and recommendation for future study are separated into two sections for clearly different purposes of utilization.

Summary of the study

The cognitive tests in this study were developed based on theoretical approaches and closely matching the established assessment tools. Three newly developed tests in Thai aimed at eliciting spontaneous speech from Thai participants, i.e., the Thai Picture description (TPD), the Thai Story Recall (TSR), and the Semi-structured Interview for Thai (SIT). This study was interested in exploring the speech profile of older Thai samples taking into account aspects of acoustic features. Participants belonged to one of three groups, i.e., Healthy Control (HC), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD). The three groups were compared on a range of language and memory variables. As a final goal, three spontaneous speech tasks were validated, and their eligibility to classify the three groups was determined.

This study is a cross-sectional, quasi-experimental study with a comparative design. A total of 98 Thai-speaking older adults aged between 56 – 79 years in Chonburi met the inclusion criteria and completed the experimental sessions. They were invited to participate by the staff and the village health volunteers of community healthcare units and the hospitals in Chonburi province. They underwent the screening assessments and were classified into one of three groups, i.e., people with no dementia (HC group), people with mild cognitive impairment (MCI group), and people with Alzheimer dementia (AD group) on the bases of standardized criteria and neuropsychological tests. In the experimental session, the examiners were blinded to which groups participants belonged to. Recordings of the three spontaneous speech tasks and of the oral responses to the cognitive tests were conducted with the participants' consent.

The independent variables obtained from the data in the screening tests included the dementia pathological stage (No dementia, MCI, Alzheimer's disease), the MoCA scores, and the educational level. The dependent variables were the performances of the cognitive tests and the content and acoustic variables from the three spontaneous speech tasks. Speech extraction was performed manually by segmenting the speaker's boundaries, after which the participant's verbal performance was automatically extracted considering two types of acoustic features, i.e., frequency-based and temporal variables. To explore the acoustic features' profile, ANOVA and correlation analyses were conducted. Different statistical analyses were conducted to examine the tests validity and discriminant ability, including MANOVA, MDA, and AUC of the ROC curve.

Discussion of the research findings

The results of this research are concluded and elaborated separately into three main sections corresponding to the research objectives.

1. Spontaneous speech analysis of participants' performance in the spontaneous speech tasks.

The participants struggled when asked to create a story out of the picture presented to them in the TPD. They tended to very descriptive, mentioning what they saw by saying, 'There is/are…' or 'A person is doing/(acting)…'. Many felt the need to excuse themselves by saying that they could not make a story. The examiner often prompted them to imagine possible situations or possible relationships among the objects, persons, and actions in the picture. The participants who could easily tell the story tended to have higher years of education, as shown by the correlation of CIU with years of education ($r = .46, p < .01$). Furthermore, the participants often complained about the picture presenting ambiguous items and/or perspectives, such as being difficult to identify the depicted animal as a cat or a dog or a rat, distinguishing between a girl or a woman, and whether the woman was sitting inside or outside of the house. In the experiment session, two older participants mentioned that the dimension of the picture was confusing. One participant said that the woman was washing the dishes inside the house because it was impossible to do this activity outdoor. The other participant recognized the figures in the door frame as a photograph hanging on the house's wall.

The fact that the drawings were considered unclear and "puzzling" and failed to elicit the narrative behavior they were intended to, leads to issue. It may be the case that for participants with a lower education level, the sketched pictures are difficult to decode and understand, and thus fail to elicit the verbal behavior they are meant to. A picture is not a straightforward representation of reality but it is a symbolic system (Gombrich, 1972) that requires prior knowledge and experience with such medium. This being the case, a more explicit stimulus, such as a photograph, may be a better experimental material. On the other hand, it may be that some aspects of the actual drawing are confusing. The fact that even the elements of the pictures that were intentionally made to elicit imagination, such as the relationship between two people, were not able to elicit a narrative renders the stimulus dubious. This being the case, a preliminary investigation with a large sample of a different version of the drawing should be performed.

The short story in the TSR characterizes more syllables and are taken longer time to verbally present than the short story in the established tests; it thus affected on low scores of both recalls. In administrating the TSR, the attention of participants was very important. Several of the participants were lost while listening to the short story. They complained that the story was long, so they could keep their concentration only half of listening. Although, the examiners tried to tell the older adults to focus on the story at the beginning of the administration. The proportion of recall units to the total number of scorable units in both rounds is 22%. This recall proportion is considered as low proportion in comparison with similar studies. In the Korean story recall study of Park et al. (2017), there were 24 scoring items; the proportion of immediate recall units is 62%, and of delayed recall units is 18%. In the TSR, the story contains 70 words or 96 syllables with approximately 90 seconds of reading. In comparison, the story in WLM is composed of 62 words and is estimated to read approximately 30 seconds (Hodges, 2007; National Institute on Aging, 2006). The long passage in the TSR is too hard for the participants to remember, thus resulting in a floor effect (Baek et al., 2011). However, it is interesting that the percentages of recall proportion of immediate round and delayed round are not different. The older adults remembered two-fifths of the story in the first recall and retained similar information until the second round. The similar number of remaining information units may possibly be due to a short delay, approximately nine minutes.

Non-continued talk in the SIT caused problems in administration and speech segmenting. During the semi-structured interview, the examiners needed to reinforce the participants to talk for a long duration. Sometimes, they could not generate context-related questions to continue the conversation, or some participants preferred to answer shortly. An interrupting piece of speech was used in the previous studies. Ambrosini et al. (2019) recorded 2 minutes of uninterrupted talk; this was a face-to-face session. János et al. (2022) collected the speech data by calling the mobile phone and asking the participant to talk about themselves. Their administration asked one open-ended question and did not allow the examiners to repeat or give verbal prompts after delivering the question. In their study, the mean utterance length of healthy control group was 275.33 seconds (approximately 4.6 minutes) and the mean of MCI group was 201.94 seconds (approximately 3.4 minutes). While the mean utterance duration in this study was 209.13 seconds (SD = 62.45). Nonetheless, this parameter was the sum of the participants' talking turns from the total duration of 5-minute interview. The SIT administration should be adjusted to reduce time consuming in the session and increase standardization of protocol.

A similar issue across the three new tasks was the non-continuity of utterance that had an influence on the acoustic features extraction. It was a difficult and time-consuming process for marking the boundaries of the participant and the examiner. Although this study applied the segmenting mechanism of Khodabakhsh et al. (2015), they used a different technique for a voice activity detector (VAD), namely the expectation-maximization (EM) algorithm. In contrast, this study used the function in Praat, namely To TextGrid (silences). The sophistication of the VAD played an important role in the feature extraction. The VAD and non-continued speech apparently confounded the speech analysis. Especially analysis of pauses, the silent pauses serve as turn-taking marks as well as the time of cognitive processing (Stenström, 2012, p. 539). Intervening with the examiners' questions allowed the speakers to spend this duration for cognitive processing, such as planning for the answer or retrieval of the information. Although the examiners tried not to interrupt in between the verbal responses of the participants, they needed to query, encourage, and ask questions to continue the conversation spontaneously. The cognitive process of the participants thus did not totally reflect in their pieces of verbal responses.

This study developed three spontaneous speech tasks that are suitable for older Thai adults. Three tasks obtain difference structural and liberal levels of verbal output. The most liberal task is the semi-structured interview (SIT), picture description (TPD), and story recall (TSR), consecutively. The TPD required the participants to remember the given story and retell the story as similarly as they could. The scores thus rely on the correctness of the participants' memorable information. While the TPD accepts a more liberal response, the scores depend on how much the participants can include the stimuli in the picture to their story. The participants can be scored as much as they can tell the relevant story, not memorable stimuli. For the SIT, there was no scoring for this task. The stimuli questions thus can mostly elicit spontaneous speech. The participants' responses were analyzed as acoustic features. The constrained level of protocol influences a limited speech output, and the restricted extent of speech is directly related to reliability in scoring. The three newly developed tasks consequently acquired a trade-off level between spontaneous speech and difficulty in scoring standardization.

2. Comparison of the patterns of acoustic features profile in Thai older adults with MCI, AD, and cognitively intact persons.

Descriptive statistics revealed that the profile of the frequency domain in this study is different from the previous studies, but previous studies also reported inconsistency results among them. The highest values of the acoustic features were mostly found in the MCI group; nine out of 20 variables (45%) in the frequency domain and 13 out of 36 variables (36%) in the temporal domain. Considering the similarity and contrast in the previous study, there are inconsistency results in the acoustic analysis, e.g., pitch or mean of fundamental frequency (F0). In the study of Ambrosini et al. (2019), the pathological group of MCI demonstrated values of F0 higher than the control group. This study found that the mean value of F0 in AD>HC>MCI respectively in the picture description task and semi-structured interview. On the other hand, Meilán et al. (2014) reported that the F_0 of healthy control was higher than AD. Also, the lower pitch in the pathological groups (AD and MCI) was found in both recall rounds of the story recall task. This mixed trend of acoustic features was further found in the values of relative jitter, e.g., HC>AD in Meilán et al. (2014) but AD>HC in Asiaee et al. (2020). In this study, the trend of jitter values was MCI>AD>HC, obviously in the picture description and immediate recall, but not in the interview and delayed recall. However, all the sample studies did not find a significant difference in the mentioned variables. The frequency domain in acoustic features may not be an acceptable distinguishing variable or predictor of dementia pathology. After all, feature extraction should be considered as a factor affecting the diverse pattern of frequency variables. The technique and mechanism issues are mentioned in the first objective's result.

The contrasting results of frequency-based markers are concerned with the differences in methodology and non-cognitive factors. Regarding the significant variable, the number of voice breaks (NVB) was identified to have significant differences between pathological groups and healthy control in this study and the previous studies. However, the tendency of mean NVB in this study was incongruent with MCI>HC>AD. One-way ANOVA revealed significant differences in NVB among the three groups in two tasks, namely, the picture description and interview. A similar study with a free-speech task was the work of Ambrosini et al. (2019). Although the pitch floor in that study was set at 70 Hz and the current study identified the minimum F0 at 60 Hz. They found a significant difference in the NVB between MCI and healthy control by MCI>HC, wherein this study found a similar trend. Despite the fact that the tendency of NVB between HC and MCI in the interview of this study replicated the result in Ambrosini et al. (2019), the tendency between HC and AD did not congruent with the previous study and explanation. Meilán et al. (2014) compared the NVB from the reading task of HC and AD; they reported a significant difference in which AD>HC. The incongruent results between the pathological groups and control group in this study can be attributed to different feature extracting, speech tasks, and administration. Besides, in the review of Martínez-Nicolás et al. (2021) mentioned that voice breaks were one of the acoustic features which was affected by AD. An explanation of the relationship between NVB and AD is stated that the voice of AD is contaminated with a tremulous voice and less uncontrollable airflow than the HC. Therefore the, sudden changes in voiced frequency occur several times (Boersma & Weenink, 2023c; Meilán et al., 2014). The spectral properties of a voice, such as pitch and voice breaks, are associated with the physical health of vocal organs, especially the vibration of the vocal folds (Ladefoged, 2006, p. 4). Consequently, the frequency-based variables of the acoustic features are not only influenced by cognitive processes but physical properties of articulatory organs.

The temporal variable and recalled units in the delayed recall task reflect preserved cognitive ability in the HC group. Despite the fact that six temporal variables demonstrated a significant correlation with the dementia pathological stage, only the total duration of utterance in Del indicated significant differences between the three groups. The HC spent the longest time in Del (*M* = 32.68 sec.), followed by AD (*M* = 24.29) and MCI (*M* = 23.58), but Turkey's post-hoc analysis report non-significant difference in a pair of AD and MCI. In the study of Roark et al. (2011), they reported significant differences in total phonation (speech events) time between HC and MCI in both the Wechsler Logical Memory I and II. Similarly, the utterance time of HC was higher than MCI group. Since recalling tasks rely on episodic and semantic memory, wherein memory deficit is the hallmark of AD (Caine & Crutch, 2016; Szatloczki et al., 2015). The results of language measures in this study further support the relationship between dementia pathology and language ability. The letter fluency and category fluency yield a significant difference between the three groups of participants. Noticeably, the scores of both recall rounds significantly correlate with the category fluency answers at a weak positive level. Pearson's correlation indicated a moderate positive correlation between recalled units in Del and utterance time in Del ($r = .49, p < .01$). The preserved abilities of memory and language of HC allowed them to generate more retrained information and spend more time to recall.

After all, the difference results in temporal variables of the current study and previous studies occur in the form of a non-significant difference and dissimilar direction of values in the acoustic features. The opposite values were spotted at only one temporal variable that showed a significant difference between three experimental groups. As mentioned above, HC obtained more utterance time in delayed recall in this study and Roark et al. (2011). Regarding the recall task, Toth et al. (2018) and

Gosztolya et al. (2019) oppositely reported that MCI spent longer speech time than MC. Martínez-Nicolás et al. (2021) provided evidence supporting the higher values of speech time in the AD pathological group. They stated that the temporal parameters of the early stages of AD patients have changed in more advanced stages of dementia. Regardless of speech tasks, the pattern of AD patients speaking longer than healthy control was exhibited in the picture description task of König et al. (2015) and the interview of Hoffmann et al. (2010). One possibility is due to the difference in the 'silence' operational definition. The short silence parts at the beginning, between, and ending of utterance will be included in speech events or silence events depending on how long the silent duration is identified in milliseconds. The definition of pauses in this study was silent events longer than 1,000 msec (1 sec) which was endorsed from Singh et al. (2001) and Roark et al. (2011). Consequently, the shorter silent segments were combined with the utterance segments, resulting in longer utterance segments in comparison with the case of silent segments shorter than 1,000 msec. The different operational definition across studies probably causes dissimilar speech-extracting results and acoustic features analysis.

3. Validating the speech analysis measure in classifying healthy controls, MCI, and AD.

Delayed recall is the most promising task among the three variables of the TSR in distinguishing three experimental groups. Mueller et al. (2022) proved that the items in delayed recall acquired a higher discriminant index than immediate recall in distinguishing amyloid positive and negative groups. Plenty of evidence reported that the scores of both recall rounds in various languages were found to yield significant differences between the three groups of the dementia pathological stage. Both scores of immediate and delayed recalls in the AD group are significantly lower than MCI and HC, respectively, in at least three versions of the story recall which include but are not limited to Chinese version of the Adult Memory and Information Processing Battery in Jing et al. (2014), (Coutinho et al., 2015), English version of Wechsler Logical Memory I and II (WMS-III) in Coutinho et al. (2015), the Korean story recall test in Park et al. (2017). Although, the TSR has similar results with the previous studies in only the delayed recall scores. Nonetheless, the delayed recall alternatively showed a convergent validation by having a significant positive correlation with the

MoCA at a moderate level ($r = .43, p < .01$). The performance of the delayed recall is sensitive to AD; it thus can separate AD cases from older adults with cognitively intact and from persons with MCI (Coutinho et al., 2015; Mitchell & Malladi, 2010). It is possible to use delayed recall as a screening tool rather than a diagnostic test. (Jing et al., 2014). In order to maximize the discriminant ability of the delayed recall, it must be combined with various related information and neuropsychological test data. The suggested combination in this study was a set of delayed recall scores and two versions of digit span, which was indicated a significant main effect of dementia pathological stages by MANOVA.

The scoring method of TPD presented similar results to the previous studies and revealed the linguistic deficits in the advance stage of AD. In Mueller, Hermann, et al. (2018), the CIUs were reported to have a significant difference between AD and HC. Several studies that applied similar criteria, namely information unit, also provided similar findings with the CIUs. In the review of Boschi et al. (2017), 13 studies that used picture description tests found significant differences in information units between AD and HC. In the same review, only one study reported a significant difference in a pair of AD and aMCI, and none of the studies found a difference in information units between HC and MCI. At the same time, Pearson's correlations in this study revealed that the CIU had a significant positive correlation with both tasks of verbal fluency. The discriminant results and correlations pertaining to the TPD support the previous findings that lexical-semantic functions are one of the most linguistic impairments in AD (Cummings, 2020, p. 7). People with AD tend to struggle with the tasks that rely on semantic knowledge, such as naming tests, verbal fluency tests, and picture description tests (Bradley et al., 2010; Szatloczki et al., 2015). Although, this study did not find a difference in the CIUs in a pair of HC and MCI when further conducted Turkey's post-hoc analysis. Notably, the deficits and impairments in MCI are not stable (Cummings, 2020, p. 75). One possible explanation is cognitive impairments in MCI related to the subtype, which this study did not classify people with MCI into subtypes (Petersen et al., 2014).

Noteworthy, the combination of CIU from the TPD and backward digit span in differentiating the dementia pathological stages adds evidence to the knowledge of the relationship between working memory and language processing. According to the

AUC (Table 4 - 43), the best combination in distinguishing the dementia stages is the discriminant function of six variables (Del, CIU, LF, CF, FDS, and BDS), which provides an acceptable AUC in four pairs, including, HC vs. MCI, MCI vs. AD, HC vs. AD, and HC vs. Cognitive impairment. Nevertheless, the set of six variables practically requires more administration than the combination of CIU and BDS. Although the discriminant of CIU and BDS cannot reach an acceptable AUC in differentiating MCI vs. AD, these two variables have an interesting relationship that can be explored furthermore. Regarding Pearson's correlation, CIU has a significant positive correlation with letter fluency ($r = .257, p < .01$) and category fluency ($r = .40, p < .01$). Alyahya et al. (2021) indicated that picture description as a discourse task accurately reflect semantic fluency in term of content measure or information unit during spontaneous speech. Telling the story from the picture requires several linguistic abilities, such as semantic knowledge, pragmatic skills, and elaboration of information (Lee & Kim, 2021). The TPD demands discourse production, which is sensitive to working memory capacity in order to integrate language ability (Lee & Kim, 2021). Further supportive explanation, Slegers et al. (2018) explained that changes in working memory are suspected of causing deterioration in comprehension and expression of syntax in AD patients. Notably, BDS is accounted as a working memory measure, especially in the central executive system (Monaco et al., 2013). Hence, the combination of CIU and BDS is not only contributed by language ability but also working memory capacity.

The findings in this study conformably found literacy bias with the previous studies in the dementia group that reduces discriminant validity of the langued-based task. Years of education had significant positive correlations with main variables, namely, MoCA, CIU, Del, verbal fluency tasks, and digit span tasks. In the study of Jing et al. (2014), the delayed recall scores showed a positive correlation with education level. For the picture description task, the number of accurate and complete information increased with the amount of education (Mackenzie et al., 2007). The results of ANOVA separated by education group confirmed this limitation (Table 4 – 13, 4 – 28, and 4 -33). In the high school group, four variables presented significant differences between three groups of the participants who acquired high school as the highest education, namely, letter fluency, backward digit span, delayed recall, and utterance time in the delayed recall. While category fluency, repeated scores in letter fluency and backward digit span indicated significant differences between the three groups of participants who graduated from university. None of the three spontaneous speech tasks showed significant differences between the three groups of participants with primary school and university level. One suggestion to reduce literacy bias is an adjusted score for education.

Limitations

The limitations in this study are separated into three sections, including theoretical-based reason, methodology, and practical application.

Theoretical-based reason

1. Word tokenization in the Thai language is different from English or those languages that form a sentence with discrete words. This study manually transcribed verbal responses in the picture description (TPD) into texts before separating them into discrete words. Since Thai written format is *scriptio continua* language or writing system in which a sentence or clause consists of words without spaces in between (Kasisopa et al., 2013). Delimiting Thai words in written format requires professional skills or a well-developed segmentation program. The program used for delimiting words in this study was Thai tokenization which was developed by Aroonmanakun (2002). At the same time, the other studies of picture description with Thai older adults used different methods, such as labeling words by linguists in the studies of Nagarachinda et al. (2020) and Amonlaksananon et al. (2021) or fully automated transcribing and tagging texts in the project of Sangchocanonta et al. (2021). The fact that the Thai language is scriptio continua language possibly causes a different number of words or information units due to different techniques and algorithms of tokenization. This limitation reduces standardization in the scoring process of the TPD. For example, tokenization with the automatic program may return a higher number of information units or words than manual delimiting by the linguist. The data with more information units has relatively higher opportunity to be scored for the correct information units.

Methodology

2. The subjective memory complaint and other behaviors regarding the CDR caused some incompliant classification. Some older adults in the control group scored 0.5 in CDR because they reported forgetful behavior, which matched with the CDR criteria. They may not be a proper representative of the cognitively intact population and limit the generalizability of the findings. Since the CDR apparently depended on the subjective observation of the participants, the scores of this measure were sometimes not compliant with the other objective scales, such as the MoCA.

3. The scoring criteria of the TPD rely on the opinions and experiences of an examiner. The scoring system used in this task allows the participants to earn scores as much as they can generate in the relevant story. At the same time, an examiner needs to decide whether to give a score to an ambiguous unit or not. This process may cause difficulty for inexperienced examiners. Also, scoring may vary across different examiners. Considerably, one scoring criterion of the picture description is scoring on target words or content units (Alyahya et al., 2021; Jensen et al., 2006). A score is given for each correctly identified information unit which is the most frequently generated by healthy control (Catricalà et al., 2017; de Vries et al., 2020). Scoring criteria on target words need a big sample of healthy controls or non-brain damage adults of different sex, age range, and education level for creating a normative list of target words (Catricalà et al., 2017; Marshall & Wright, 2007). The target word scoring was applied in several studies with neurological disorders, including but not limited to Marshall and Wright (2007), Catricalà et al. (2017), and Alyahya et al. (2021).

However, different scoring criteria can be applied to the same stimulus picture. The divided attention picture of Marshall and Wright (2007) was initially developed for determining a further examination; a number of target words thus were an indicator of the severity of aphasia. Whereas the same picture was used to elicit connected speech in older Thai adults. Nagarachinda et al. (2020) examined the differences in a cohesive device (discourse) between MCI and AD; they did not use the scoring of target words. Incidentally, the scoring rules for picture description depend on the purpose of the research.

4. Notably, the five factual inquiries in the TSR did not show differences between the three groups and also presented the floor effect. This study experimented with recognition questions. The first adaptation was asking questions without a cue. Unlike the established tests, two kinds of recognition questions were conducted, namely, yes/no questions and multiple questions (Foldi, 2011; Song et al., 2019; Wechsler, 2009). The mean score of the five questions is 1.59, and the mode is 1. The mode score can be inferred that the question without cue in this study was too difficult; the floor effect thus happened. The second adaptation with the recognition question was placing the questions after the immediate recall, not at the end of the delayed recall. In WLM and MMSE-2: EV, the yes/no questions are queried after the delayed recall (Song et al., 2019; Wechsler, 2009). In the Cowboy story, the multiple choice questions are presented after the first recall, with no delayed recall (Foldi, 2011). When the participants answered the five questions, they also rehearsed the story, which allowed them to strengthen the memory trace. Presenting questions after the immediate recall was one factor causing the similar number of recall units in both recall rounds.

Another weakness is the TSR was only validated by the CVI, the other methods should be considered. For instance, a pilot study can alternatively provide a small statistical analysis and simulation. It could allow us to demonstrate the difficulties of stimuli, such as a few correct responses to factual inquiries. The researcher would then be able to adjust the test by opting for the recognition method as well as revising the segments of the story to acquire more refined scores.

Practical application

5. A lack of technological measurement to identify speakers' boundaries resulted in a time-consuming process and inaccurate and unstandardized speech extraction. The literature on speech analysis with dementia groups rarely mentioned how to identify the boundaries between participants and examiners in the manual segmenting method. For those who performed the automatic speaker identification, specifically developed or applied the automated tools for their languages of interest, such as Brazilian Portuguese language in the study of Treviso et al. (2018) and Hungarian language in the work of Toth et al. (2018).

6. Scoring for the TPD was time-consuming and relied on examiners' justification. Manually converting the speech into texts provided accurate texts with fewer errors, but this method consumed the amount of time. After having the written form of responses, the examiners decided whether each word was correct or incorrect information unit. Even though the examiners tried to strictly follow the rules for CIUs, some units were ambiguous and difficult to decide. The subjective decision thus decreases the standardization of CIUs scoring criteria.

Implications

The finds in this study can be beneficial for practical implications and research with similar purposes. Also, this study emphasized the importance of dementia screening, which should be involved in Thailand's healthcare policy. Four implications are suggested below.

Clinical implication

1. According to the developed spontaneous speech tasks, the TPD and TSR can be used as screening tools incorporated with other neurocognitive assessments. Although the scoring criteria of the TPD require time and several processes, the CIU yields good validity by showing a moderate positive correlation with the MoCA. A backward digit span was recommended to pair with the CIU in distinguishing people with cognitively intact and those with cognitive impairment. In contrast, the TSR appears to have respectively higher structural administration and scoring criteria than the TPD. The delayed recall task is able to elicit the different cognitive deficiencies between three experimental groups.

Research implication

2. The essential components of the three spontaneous tasks presented in this study are beneficial for further study to develop administration and scoring protocol. Also, the two stimuli can be adapted for specific or similar purposes. Especially the short story of the TSR that was created based on the phenomenon in the phonological loop. Furthermore, the findings regarding the TPD and TSR provided guidelines on how the stimulus picture and story should be improved.

3. To the researcher's knowledge, this study is the pioneer in detecting AD and MCI in Thai older adults with respectively inclusive acoustic features from

spontaneous speech. This study aimed to explore the acoustic features in both frequency-related and temporal domains. Although, the findings regarding the acoustic features are not in line with the international studies in a similar methodology. The developed spontaneous speech tasks shed light on the speech analysis for detecting dementia in older Thai adults. The acoustic features extracted from the spontaneous speech in this study provide uniqueness and specification. Meanwhile, the feature extraction suitable for Thai speech need to be improved. As mentioned in the limitation, that *scriptio continua* feature of the Thai language decreases generalizability and practical application. This limitation thus constitutes a potential investigation and invention of the ready-to-use and inclusive program or mobile application for detecting AD and MCI with spontaneous speech responses.

Policy implication

4. Accessible and cheap dementia screening should be progressively developed for older Thai adults. Healthcare policymakers should take action corresponding to the changing of the older Thai population. Research in dementia screening should be invested in and receive more attention. This study proposed spontaneous speech tasks which can be applied in clinical settings and further improved for telemedicine schemes. At the same time, reducing the gap in automatic speech recognition should be parallelly studied and targeted at dementia speech samples.

Recommendations for future research

1. Different methods of scoring for the picture description should be studied. To increase the generalizability of the TPD, more structured and standardized scoring criteria are required. The recommended scoring methods are target word scoring and counting words or syllables. The target words can be derived from the salient objects or actions in a picture or the normative data from responses of people without cognitive impairment (healthy control). When participants tell the story containing the target words, they receive one point per target word. Slegers et al. (2018) demonstrated that 85 % of the studies with picture descriptions found that people with AD tended to generate fewer information units. The 'information units' in the review of Sleger and team included content units, total semantic units, subjects, objects,

actions, component measures, quantity of essential material, locations, correct information units, essential units, information conveyed, information content, number of content units, repetition of expected ideas, pictorial themes, number of relevant descriptions, keywords, places, main concept score, and localizations. Future studies may also try to score with those methods. Another suggested method is counting the generated words or syllables. This method was found to be applied in the study with older Thai adults of Amonlaksananon et al. (2021). Both recommended scoring methods reduce an examiner's justification.

2. Various psychometric evaluations should be conducted. The developed tasks in this study were only validated by the CVI. Alternatively, a pilot study is recommended since it can provide a small statistical analysis and simulation of practical implications. The small set of data from a pilot study can be advantageous for adjusting the created tasks, such as controlling administration time or solving problems regarding a floor effect. Different reliability assessments can be performed, such as test-retest or inter-rater. However, the test-retest method should be concerned with a learning effect. For validity evaluation, this study proposed a concurrent validity regarding AUC, examined the accuracy in differentiating known groups, and a convergent validity by investigating Pearson's correlation of the interested parameters of the spontaneous speech tasks with the MoCA. The future study may employ predictive validity one year after the current study or design a longitudinal study for this purpose.

3. Heterogeneous samples should be further investigated to increase the representativeness of older Thai adults. First, the sample in this study is community-based participants, and the clinical patients should be invited more. The participants in the community can resemble older adults in the preclinical stage, but this sample lacks external control. Otherwise, biomarkers can be added to increase the objectivity of a differential diagnosis. The second factor which should be concerned is the different dialects in Thai. Generally, the participants in this study speak in the Central Thai dialect. In Thailand, there are four main dialects, namely, Central, Northern, North-Eastern, and Southern dialects. Last but not least, MCI subtypes should be differentiated or specifically selected to optimally control the influences of prominent cognitive impairment in each subtype.

REFERENCES

Abhang, P. A., Gawali, B. W., & Mehrotra, S. C. (2016). Chapter 3 - Technical Aspects of Brain Rhythms and Speech Parameters. In P. A. Abhang, B. W. Gawali, & S. C. Mehrotra (Eds.), *Introduction to EEG- and Speech-Based Emotion Recognition* (pp. 51-79). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-804490-2.00003-8

Acheson, D. J., & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychol Bull*, *135*(1), 50-68. https://doi.org/10.1037/a0014411

Ahmed, R. M., Paterson, R. W., Warren, J. D., Zetterberg, H., O'Brien, J. T., Fox, N. C., Halliday, G. M., & Schott, J. M. (2014). Biomarkers in dementia: clinical utility and new directions. *Journal of Neurology, Neurosurgery & Psychiatry*, *85*(12), 1426-1434. https://doi.org/10.1136/jnnp-2014-307662

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, *18*(3), 91-93. https://doi.org/https://doi.org/10.1016/j.tjem.2018.08.001

Al-Hameed, S., Benaissa, M., & Christensen, H. (2017). *Detecting and Predicting Alzheimer's Disease Severity in Longitudinal Acoustic Data* Proceedings of the 4th International Conference on Bioinformatics Research and Applications, Barcelona, Spain. https://doi.org/10.1145/3175587.3175589

Al-Hameed, S., Benaissa, M., Christensen, H., Mirheidari, B., Blackburn, D., & Reuber, M. (2019). A new diagnostic approach for the identification of patients with neurodegenerative cognitive complaints. *PLoS ONE*, *14*(5), Article e0217388. https://doi.org/10.1371/journal.pone.0217388

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*, *7*(3), 270-279. https://doi.org/10.1016/j.jalz.2011.03.008

Alyahya, R. S. W., Conroy, P., Halai, A. D., & Lambon Ralph, M. A. (2021). An efficient, accurate and clinically-applicable index of content word fluency in Aphasia. *Aphasiology*, 1-19. https://doi.org/10.1080/02687038.2021.1923946

Alzheimer's Association. (2016). 2016 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, *12*(4), 459-509. https://doi.org/10.1016/j.jalz.2016.03.001

Alzheimer's Association. (2020a). 2020 Alzheimer's disease facts and figures. *Alzheimer's & Dementia, 16*(3), 391-460. https://doi.org/10.1002/alz.12068

Alzheimer's Association. (2020b). *Earlier Diagnosis*. Retrieved 20 Aug 2020 from https://www.alz.org/alzheimers-dementia/research_progress/earlier-diagnosis

Alzheimer's Association. (2022). *2022 Alzheimer's Disease Facts and Figures*. https://www.alz.org/media/Documents/alzheimers-facts-and-figures.pdf

Alzheimer's Society. (2016). *What is dementia with Lewy bodies (DLB)?* https://www.alzheimers.org.uk/sites/default/files/migrate/downloads/what_is_dementia_with_lewy_bodies.pdf

Ambrosini, E., Cid, M., de Isla, C. G., Salamanca, P., Borghese, N. A., Ferrante, S., Caielli, M., Milis, M., Loizou, C., Azzolino, D., Damanti, S., Bertagnoli, L., Cesari, M., & Moccia, S. (2019). Automatic speech analysis to early detect functional cognitive decline in elderly population. *Conf Proc IEEE Eng Med Biol Soc, 2019*, 212-216. https://doi.org/10.1109/embc.2019.8856768

Amezquita-Sanchez, J. P., Mammone, N., Morabito, F. C., Marino, S., & Adeli, H. (2019). A novel methodology for automated differential diagnosis of mild cognitive impairment and the Alzheimer's disease using EEG signals. *Journal of Neuroscience Methods, 322*, 88-95. https://doi.org/https://doi.org/10.1016/j.jneumeth.2019.04.013

Amonlaksananon, A., Onsuwan, C., & Tantibundhit, C. (2021). Speech disfluencies in elderly Thais: A comparison among Mild Alzheimer's patients, Mild Cognitive Impairment participants, and cognitively normal participants. *Journal of Liberal Arts, Thammasat University, 21*(2), 299-334.

Andric, M., & Small, S. (2015). fMRI methods for studying the neurobiology of language under naturalistic conditions. In R. M. Willems (Ed.), *Cognitive neuroscience of natural language use* (pp. 8-28). Cambridge University Press. https://doi.org/10.1017/CBO9781107323667.002

Aniwattanapong, D., Tangwongchai, S., Supasitthumrong, T., Hemrunroj, S., Tunvirachaisakul, C., Tawankanjanachot, I., Chuchuen, P., Snabboon, T., Carvalho, A. F., & Maes, M. (2018). Validation of the Thai version of the short Boston Naming Test (T-BNT) in patients with Alzheimer's dementia and mild cognitive impairment: clinical and biomarker correlates. *Aging Ment Health, 23*(7), 840-850. https://doi.org/10.1080/13607863.2018.1501668

Aramaki, E., Shikata, S., Miyabe, M., & Kinoshita, A. (2016). Vocabulary Size in Speech May Be an Early Indicator of Cognitive Impairment. *PLoS ONE, 11*(5), e0155195. https://doi.org/10.1371/journal.pone.0155195

Ardila, A. (2005). Cultural Values Underlying Psychometric Cognitive Testing. *Neuropsychology Review, 15*(4), 185. https://doi.org/10.1007/s11065-005-9180-y

Armstrong, L., Borthwick, S. E., Bayles, K. A., & Tomoeda, C. K. (1996). Use of the Arizona Battery for Communication Disorders of Dementia in the UK. *International Journal of Language & Communication Disorders, 31*(2), 171-180. https://doi.org/10.3109/13682829609042219

Aroonmanakun, W. (2002). Collocation and Thai Word Segmentation.

Arvanitakis, Z., Shah, R. C., & Bennett, D. A. (2019). Diagnosis and Management of Dementia: Review. *Jama, 322*(16), 1589-1599. https://doi.org/10.1001/jama.2019.4782

Asgari, M., Kaye, J., & Dodge, H. (2017). Predicting mild cognitive impairment from spontaneous spoken utterances. *Alzheimer's & Dementia: Translational Research & Clinical Interventions, 3*(2), 219-228. https://doi.org/https://doi.org/10.1016/j.trci.2017.01.006

Asiaee, M., Vahedian Azimi, A., Atashi, S., Keramatfar, A., & Nourbakhsh, M. (2020). Voice Quality Evaluation in Patients With COVID-19: An Acoustic Analysis. *Journal of Voice, 36*. https://doi.org/10.1016/j.jvoice.2020.09.024

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189-208. https://doi.org/10.1016/S0021-9924(03)00019-4

Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the Articulatory Loop. *The Quarterly Journal of Experimental Psychology Section A, 36*(2), 233-252. https://doi.org/10.1080/14640748408402157

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 575-589. https://doi.org/https://doi.org/10.1016/S0022-5371(75)80045-4

Baek, M. J., Kim, H. J., Ryu, H. J., Lee, S. H., Han, S. H., Na, H. R., Chang, Y., Chey, J. Y., & Kim, S. (2011). The usefulness of the story recall test in patients with mild cognitive impairment and Alzheimer's disease. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn, 18*(2), 214-229. https://doi.org/10.1080/13825585.2010.530221

Bayles, K. A., Azuma, T., Cruz, R. F., Tomoeda, C. K., Wood, J. A., & Montgomery, E. B., Jr. (1999). Gender differences in language of Alzheimer disease patients revisited. *Alzheimer Dis Assoc Disord*, *13*(3), 138-146. https://doi.org/10.1097/00002093-199907000-00005

Bayles, K. A., McCullough, K., & Tomoeda, C. K. (2020). *Cognitive-Communication Disorders of MCI and Dementia: Definition, Assessment, and Clinical Management* (3rd ed.). Plural Publishing, Inc.

Beltrami, D., Calzà, L., Gagliardi, G., Ghidoni, E., Marcello, N., Favretti, R., & Tamburini, F. (2016). Automatic Identification of Mild Cognitive Impairment through the Analysis of Italian Spontaneous Speech Productions. 10th International Conference on Language Resources and Evaluation, Portorož, Slovenia.

Beltrami, D., Gagliardi, G., Rossini Favretti, R., Ghidoni, E., Tamburini, F., & Calzà, L. (2018). Speech Analysis by Natural Language Processing Techniques: A Possible Tool for Very Early Detection of Cognitive Decline? *Frontiers in Aging Neuroscience*, *10*(369), 1-13. https://doi.org/10.3389/fnagi.2018.00369

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, *29*(4), 552-557. https://doi.org/10.7334/psicothema2016.383

Boersma, P., & Weenink, D. (2022). *Praat: doing phonetics by computer*. https://www.fon.hum.uva.nl/praat/

Boersma, P., & Weenink, D. (2023a). *Harmonicity*.

Boersma, P., & Weenink, D. (2023b). *Sound: To TextGrid (silences)...* https://www.fon.hum.uva.nl/praat/manual/Sound__To_TextGrid__silences____.html

Boersma, P., & Weenink, D. (2023c). *Voice 1. Voice breaks*.

Boersma, P., & Weenink, D. (2023d). *Voice 2. Jitter*. https://www.fon.hum.uva.nl/praat/manual/Voice_2__Jitter.html

Boersma, P., & Weenink, D. (2023e). *Voice 3. Shimmer*.

Boonpeng, A. (2014). *Care needs in older persons with early, moderate, and late stage of dementia* [Master's thesis, Chulalongkorn University]. Bangkok. https://doi.nrct.go.th//ListDoi/listDetail?Resolve_DOI=10.14457/CU.the.2014.521

Boschi, V., Catricalà, E., Consonni, M., Chesi, C., Moro, A., & Cappa, S. F. (2017). Connected Speech in Neurodegenerative Language Disorders: A Review. *Frontiers in Psychology, 8*(269). https://doi.org/10.3389/fpsyg.2017.00269

Bradley, V., Kapur, N., & Marshall, J. (2010). Neuropsychological assessment of memory disorders. In J. Gurd & U. Kischka (Eds.), *The Handbook of Clinical Neuropsychology* (pp. 159-183). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199234110.003.009

Brown, R. K. J., Bohnen, N. I., Wong, K. K., Minoshima, S., & Frey, K. A. (2014). Brain PET in Suspected Dementia: Patterns of Altered FDG Metabolism. *RadioGraphics, 34*(3), 684-701. https://doi.org/10.1148/rg.343135065

Buchsbaum, B. R. (2016a). Chapter 69 - Working Memory and Language. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 863-875). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-407794-2.00069-9

Buchsbaum, B. R. (2016b). Working Memory and Language. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 863-875). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-407794-2.00069-9

[Record #317 is using a reference type undefined in this output style.]

Budson, A. E., Solomon, P. R., & Budson, A. E. (2016). *Memory loss, Alzheimer's disease, and dementia : a practical guide for clinicians*. https://www.clinicalkey.com/dura/browse/bookChapter/3-s2.0-C20130000011

Bung, M. A. (2016). *A quantitative approach to analysing conversational turn-taking in alzheimer's disease and frontotemporal dementia* [Ph.D., University of Southampton (United Kingdom)]. ProQuest Dissertations & Theses Global. Ann Arbor. https://search.proquest.com/docview/1985579578?accountid=44783

Bürki, A. (2018). Variation in the speech signal as a window into the cognitive architecture of language production. *Psychon Bull Rev, 25*(6), 1973-2004. https://doi.org/10.3758/s13423-017-1423-4

Caine, D., & Crutch, S. J. (2016). Neuropsychological assessment. In M. Husain & J. M. Schott (Eds.), *Oxford Textbook of Cognitive Neurology and Dementia*. Oxford University Press.

Catricalà, E., Gobbi, E., Battista, P., Miozzo, A., Polito, C., Boschi, V., Esposito, V., Cuoco, S., Barone, P., Sorbi, S., Cappa, S. F., & Garrard, P. (2017). SAND: a Screening for Aphasia in NeuroDegeneration. Development and normative data. *Neurol Sci, 38*(8), 1469-1483. https://doi.org/10.1007/s10072-017-3001-y

Celine De, L., Finnian, K., Lisa, C., Aisling, V., Robert, F. C., Cathal, W., Brian, A. L., & Richard, B. R. (2018). Changes in Speech Chunking in Reading Aloud is a

Marker of Mild Cognitive Impairment and Mild-to-Moderate Alzheimer's Disease. *Current Alzheimer Research, 15*(9), 828-847. https://doi.org/http://dx.doi.org/10.2174/1567205015666180404165017

Chapman, S. B., Highley, A. P., & Thompson, J. L. (1998). Discourse in fluent aphasia and Alzheimer's disease: Linguistic and pragmatic considerations. *Journal of Neurolinguistics, 11*(1), 55-78. https://doi.org/https://doi.org/10.1016/S0911-6044(98)00005-0

Charernboon, T. (2018). Verbal fluency in the Thai elderly, elderly with mild cognitive impairment and elderly with dementia. *Journal of Mental Health of Thailand, 26*(2), 91-102.

Charernboon, T. (2019). Diagnostic accuracy of the Thai version of the Mini-Addenbrooke's Cognitive Examination as a mild cognitive impairment and dementia screening test. *Psychogeriatrics*, *19*(4), 340-344. https://doi.org/10.1111/psyg.12417

Chari, D., Ali, R., & Gupta, R. (2015). Reversible dementia in elderly: Really uncommon? *J Geriatr Ment Health, 2*(1), 30-37. http://www.jgmh.org/text.asp?2015/2/1/30/161378

Chen, S.-H., Bu, X.-L., Jin, W.-S., Shen, L.-L., Wang, J., Zhuang, Z.-Q., Zhang, T., Zeng, F., Yao, X.-Q., Zhou, H.-D., & Wang, Y.-J. (2017). Altered peripheral profile of blood cells in Alzheimer disease: A hospital-based case-control study. *Medicine, 96*(21), e6843-e6843. https://doi.org/10.1097/MD.0000000000006843

Chertkow, H., Nasreddine, Z., Joanette, Y., Drolet, V., Kirk, J., Massoud, F., Belleville, S., & Bergman, H. (2007). Mild cognitive impairment and cognitive impairment, no dementia: Part A, concept and diagnosis. *Alzheimer's & Dementia*, *3*(4), 266-282. https://doi.org/https://doi.org/10.1016/j.jalz.2007.07.013

Chetelat, G., Eustache, F., Viader, F., De La Sayette, V., Pelerin, A., Mezenge, F., Hannequin, D., Dupuy, B., Baron, J. C., & Desgranges, B. (2005). FDG-PET measurement is more accurate than neuropsychological assessments to predict global cognitive deterioration in patients with mild cognitive impairment. *Neurocase, 11*(1), 14-25. https://doi.org/10.1080/13554790490896938

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education*. Routledge.

Conrad, R., & Hull, A. J. (1964). INFORMATION, ACOUSTIC CONFUSION AND MEMORY SPAN. *British Journal of Psychology*, *55*(4), 429-432. https://doi.org/https://doi.org/10.1111/j.2044-8295.1964.tb00928.x

Coutinho, G., Drummond, C., De Oliveira-Souza, R., Moll, J., Tovar-Moll, F., & Mattos, P. (2015). Immediate story recall in elderly individuals with memory complaints: How much does it contribute to memory assessment? *International Psychogeriatrics, 27*(10), 1679-1686. https://doi.org/10.1017/S1041610215000307

Creswell, J. W., & Creswell, J. D. (2018). *Research design : qualitative, quantitative, and mixed methods approaches*.

Crosson, B., Ford, A., McGregor, K. M., Meinzer, M., Cheshkov, S., Li, X., Walker-Batson, D., & Briggs, R. W. (2010). Functional imaging and related techniques: an introduction for rehabilitation researchers. *Journal of rehabilitation research and development, 47*(2), vii-xxxiv. https://doi.org/10.1682/jrrd.2010.02.0017

Cummings, J. (2012). Alzheimer's disease diagnostic criteria: practical applications. *Alzheimers Res Ther, 4*(5), 35. https://doi.org/10.1186/alzrt138

Cummings, L. (2019). Describing the Cookie Theft picture: Sources of breakdown in Alzheimer's dementia. *Pragmatics and Society, 10*, 151-174. https://doi.org/10.1075/ps.17011.cum

Cummings, L. (2020). *Language in Dementia*. Cambridge University Press. https://doi.org/DOI: 10.1017/9781108587921

Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication, 71*, 10-49. https://doi.org/https://doi.org/10.1016/j.specom.2015.03.004

de Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods, 41*(2), 385-390. https://doi.org/10.3758/BRM.41.2.385

de Vries, S. M., Tucha, O., Melis-Dankers, B. J. M., Vrijling, A. C. L., Ribbers, S., Cornelissen, F. W., & Heutink, J. (2020). The Birthday Party Test (BPT); A new picture description test to support the assessment of simultanagnosia in patients with acquired brain injury. *Applied Neuropsychology:Adult*. https://doi.org/10.1080/23279095.2020.1763998

Dementia Australia. (2017). *Mild Cognitive Impairment (MCI)*. https://www.dementia.org.au/sites/default/files/helpsheets/Helpsheet-DementiaQandA13-MildCognitiveImpairment_english.pdf

Dementia Australia. (2018). *Diagnostic criteria for dementia*. https://www.dementia.org.au/information/for-health-professionals/clinical-resources/diagnostic-criteria-for-dementia

Dharmasaroja, P. A., Assanasen, J., Pongpakdee, S., Jaisin, K., Lolekha, P., Phanasathit, M., Cheewakriengkrai, L., Chotipanich, C., Witoonpanich, P., Pitiyarn, S., Lertwilaiwittaya, P., Dejthevaporn, C., Limwongse, C., & Phanthumchinda, K. (2021). Etiology of Dementia in Thai Patients. *Dementia and Geriatric Cognitive Disorders Extra*, *11*(1), 64-70. https://doi.org/10.1159/000515676

Diessel, H. (2019). Cognitive Processes and Language Use. In *The Grammar Network: How Linguistic Structure Is Shaped by Language Use* (pp. 23-40). Cambridge University Press. https://doi.org/DOI: 10.1017/9781108671040.003

Dronkers, N. F., Ivanova, M. V., & Baldo, J. V. (2017). What Do Language Disorders Reveal about Brain-Language Relationships? From Classic Models to Network Approaches. *J Int Neuropsychol Soc*, *23*(9-10), 741-754. https://doi.org/10.1017/s1355617717001126

Drozdick, L. W., Holdnack, J. A., Weiss, L. G., & Zhou, X. (2013). Overview of the WAIS-IV/WMS-IV/ACS. In J. A. Holdnack, L. W. Drozdick, L. G. Weiss, & I. G. L. (Eds.), *WAIS-IV, WMS-IV, and ACS : Advanced Clinical Interpretation*. Academic Press. http://www.123library.org/book_details/?id=102340

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. https://doi.org/10.3758/BF03193146

Fernandéz, E. M., & Cairns, H. S. (2010). *Fundamentals of psycholinguistics*. Wiley-Blackwell.

Ferris, S. H., & Farlow, M. (2013). Language impairment in alzheimer's disease and benefits of acetylcholinesterase inhibitors. *Clinical Interventions in Aging*, *8*, 1007-1014. https://doi.org/10.2147/CIA.S39959

Field, A. P. (2018). *Discovering statistics using ibm spss statistics* (5th ed.). SAGE Publications.

Filiou, R.-P., Bier, N., Slegers, A., Houzé, B., Belchior, P., & Brambati, S. M. (2019). Connected speech assessment in the early detection of Alzheimer's disease and mild cognitive impairment: a scoping review. *Aphasiology*, *34*(6), 723-755. https://doi.org/10.1080/02687038.2019.1608502

Flinker, A., Korzeniewska, A., Shestyuk, A. Y., Franaszczuk, P. J., Dronkers, N. F., Knight, R. T., & Crone, N. E. (2015). Redefining the role of Broca's area in speech. *Proceedings of the National Academy of Sciences*, *112*(9), 2871-2875. https://doi.org/10.1073/pnas.1414491112

Foldi, N. S. (2011). Getting the hang of it: preferential gist over verbatim story recall and the roles of attentional capacity and the episodic buffer in Alzheimer disease. *J Int Neuropsychol Soc*, *17*(1), 69-79. https://doi.org/10.1017/s1355617710001165

Foundation of Thai Gerontology Research and Development Institute. (2016). *Situation of the Thai Elderly*. http://thaitgri.org/?p=38427

Foundation of Thai Gerontology Research and Development Institute. (2019). *Situation of the Thai Elderly 2019*. https://thaitgri.org/?wpdmpro=situation-of-the-thai-eldery-2019

Fraser, K. C., Lundholm Fors, K., Eckerström, M., Öhman, F., & Kokkinakis, D. (2019). Predicting MCI Status From Multimodal Language Data Using Cascaded Classifiers. *Frontiers in Aging Neuroscience, 11*, 205-205. https://doi.org/10.3389/fnagi.2019.00205

Fraser, K. C., Meltzer, J. A., Graham, N. L., Leonard, C., Hirst, G., Black, S. E., & Rochon, E. (2014). Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex*, *55*, 43-60. https://doi.org/https://doi.org/10.1016/j.cortex.2012.12.006

Frisoni, G. B., Fox, N. C., Jack, C. R., Jr., Scheltens, P., & Thompson, P. M. (2010). The clinical use of structural MRI in Alzheimer disease. *Nature reviews. Neurology*, *6*(2), 67-77. https://doi.org/10.1038/nrneurol.2009.215

Gagliardi, G., & Tamburini, F. (2021). Linguistic biomarkers for the detection of Mild Cognitive Impairment. *Lingue e Linguaggio, 20*(1), 3-31. https://doi.org/10.1418/101111

Gandour, J., Dardarananda, R., Buckingham, H., & Viriyavejakul, A. (1986). A THAI ADAPTATION OF THE BOSTON DIAGNOSTIC APHASIA EXAMINATION. *Crossroads: An Interdisciplinary Journal of Southeast Asian Studies*, *2*(3), 1-39. www.jstor.org/stable/40860213

Gandour, J., Ponglorpisit, S., Khunadorn, F., Dechongkit, S., Boongird, P., & Satthamnuwong, N. (2000). Speech timing in Thai left- and right-hemisphere-damaged individuals. *Cortex, 36*(2), 281-288. https://doi.org/10.1016/s0010-9452(08)70529-x

Gandour, J., Potisuk, S., Ponglorpisit, S., Dechongkit, S., Khunadorn, F., & Boongird, P. (1996). Tonal Coarticulation in Thai after Unilateral Brain Damage. *Brain and Language, 52*(3), 505-535. https://doi.org/https://doi.org/10.1006/brln.1996.0027

Gaubert, S., Raimondo, F., Houot, M., Corsi, M. C., Naccache, L., Diego Sitt, J., Hermann, B., Oudiette, D., Gagliardi, G., Habert, M. O., Dubois, B., De Vico Fallani, F., Bakardjian, H., & Epelbaum, S. (2019). EEG evidence of compensatory mechanisms in preclinical Alzheimer's disease. *Brain*, *142*(7), 2096-2112. https://doi.org/10.1093/brain/awz150

Giles, E., Patterson, K., & Hodges, J. R. (1996). Performance on the Boston Cookie Theft picture description task in patients with early dementia of the Alzheimer's type: Missing information. *Aphasiology*, *10*(4), 395-408. https://doi.org/10.1080/02687039608248419

Gindri, G., & Fonseca, R. P. (2012). Rehabilitation of post-stroke communication impairments. *Revista da Sociedade Brasileira de Fonoaudiologia, 17*(2), 363-369. https://doi.org/10.1590/S1516-80342012000300021.

Global Burden of Disease 2019 Ageing Collaborators. (2022). Global, regional, and national burden of diseases and injuries for adults 70 years and older: systematic analysis for the Global Burden of Disease 2019 Study. *Bmj*, *376*, e068208. https://doi.org/10.1136/bmj-2021-068208

Goldstein, E. B. (2019). *Cognitive psychology : connecting mind, research and everyday experience* (5 ed.). Cengage learning.

Goldstein, L. H., & McNeil, E. (2012). *Clinical Neuropsychology : A Practical Guide to Assessment and Management for Clinicians* (2 nd. ed.). John Wiley & Sons, Incorporated. http://ebookcentral.proquest.com/lib/trento-ebooks/detail.action?docID=7104535

Gombrich, E. H. (1972). The visual image. *Sci Am*, *227*(3), 82-96. https://doi.org/10.1038/scientificamerican0972-82

Gomez, R. G., & White, D. A. (2006). Using verbal fluency to detect very mild dementia of the Alzheimer type. *Archives of Clinical Neuropsychology*, *21*(8), 771-775. https://doi.org/https://doi.org/10.1016/j.acn.2006.06.012

Gomperts, S. N. (2016). Lewy Body Dementias: Dementia With Lewy Bodies and Parkinson Disease Dementia. *Continuum (Minneapolis, Minn.)*, *22*(2 Dementia), 435-463. https://doi.org/10.1212/CON.0000000000000309

Goswami, U. (2012). Phonological Representation. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 2625-2627). Springer US. https://doi.org/10.1007/978-1-4419-1428-6_148

Gosztolya, G., Vincze, V., Tóth, L., Pákáski, M., Kálmán, J., & Hoffmann, I. (2019). Identifying Mild Cognitive Impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. *Computer Speech and Language*, *53*, 181-197. https://doi.org/10.1016/j.csl.2018.07.007

Grabowski, T., & Damasio, A. (2004). The Neuropathology of Dementia: Definition, clinical features and neuroanatomical basis of dementia. In M. M. Esiri, V. M. Y. Lee, & J. Q. Trojanowski (Eds.), *The neuropathology of dementia* (pp. 1-33). Cambridge University Press.

Greicius, M. D., Srivastava, G., Reiss, A. L., & Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. *Proceedings of the National Academy of Sciences of the United States of America, 101*(13), 4637-4642. https://doi.org/10.1073/pnas.0308627101

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Pearson Education Limited.

Hajian-Tilaki, K. (2013). Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian J Intern Med, 4*(2), 627-635.

Hampel, H., & Lista, S. (2016). The rising global tide of cognitive impairment. *Nature Reviews Neurology, 12*(3), 131-132. https://doi.org/10.1038/nrneurol.2015.250

Hanağası, H. A., Bilgiç, B., & Emre, M. (2016). Dementia with Lewy bodies and Parkinson’s disease dementia. In M. Husain & J. M. Schott (Eds.), *Oxford Textbook of Cognitive Neurology and Dementia*. Oxford University Press.

Hemrungrojn, S., Charernboon, T., Phannasathit, M., Akkayagorn, L., Tangwongchai, S., Phanthumchinda, K., & Nasreddine, Z. S. (2009). *Cognitive domain from Thai-Montreal Cognitive Assessment to demonstrate between amnestic MCI and mild AD from normal aging* International Psychogeriatric Association 14th International Congress, Montreal, Canada. https://www.researchgate.net/publication/315741092_Cognitive_domain_from_Thai-Montreal_Cognitive_Assessment_to_demonstrate_between_amnestic_MCI_and_mild_AD_from_normal_aging

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia, 42*(9), 1212-1222. https://doi.org/10.1016/j.neuropsychologia.2004.02.001

Henry, L. (2011). *The Development of Working Memory in Children* (1st ed.). SAGE Publications.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience, 13*(2), 135-145. https://doi.org/10.1038/nrn3158

Hodges, J. R. (2007). *Cognitive Assessment for Clinicians* (2 ed.). OUP Oxford. http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=1591080

Hoffmann, I., Nemeth, D., Dye, C. D., Pákáski, M., Irinyi, T., & Kálmán, J. (2010). Temporal parameters of spontaneous speech in Alzheimer's disease. *International Journal of Speech-Language Pathology, 12*(1), 29-34. https://doi.org/10.3109/17549500903137256

Holdnack, J. A., & Drozdick, L. W. (2010). CHAPTER 9 - Using WAIS-IV with WMS-IV. In L. G. Weiss, D. H. Saklofske, D. L. Coalson, & S. E. Raiford (Eds.), *WAIS-IV Clinical Use and Interpretation* (pp. 237-283). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-375035-8.10009-6

Hugo, J., & Ganguli, M. (2014). Dementia and cognitive impairment: epidemiology, diagnosis, and treatment. *Clinics in geriatric medicine*, *30*(3), 421-442. https://doi.org/10.1016/j.cger.2014.04.001

Hussain, R., Zubair, H., Pursell, S., & Shahab, M. (2018). Neurodegenerative Diseases: Regenerative Mechanisms and Novel Therapeutic Approaches. *Brain sciences*, *8*(9), 177. https://doi.org/10.3390/brainsci8090177

Iadecola, C. (2013). The Pathobiology of Vascular Dementia. *Neuron*, *80*(4), 844-866. https://doi.org/https://doi.org/10.1016/j.neuron.2013.10.008

Itprasert, S. (2017). *Quality of life of the elderly in Klaeng Kachet municipality, Mueang district, Rayong province* Burapha University].

János, K., Devanand, D., Gosztolya, G., Balogh, R., Imre, N., Tóth, L., Hoffmann, I., Kovács, I., Vincze, V., & Pákáski, M. (2022). Temporal Speech Parameters Detect Mild Cognitive Impairment in Different Languages: Validation and Comparison of the Speech-GAP Test® in English and Hungarian. *Current Alzheimer Research*, *19*. https://doi.org/10.2174/1567205019666220418155130

Jantapo, A., & Kusoom, W. (2021). Lifestyle and Cultural Factors Related to Longevity Among Older Adults in the Northeast of Thailand. *Journal of Transcultural Nursing*, *32*(3), 248-255. https://doi.org/10.1177/1043659620921225

Jensen, A. M., Chenery, H. J., & Copland, D. A. (2006). A comparison of picture description abilities in individuals with vascular subcortical lesions and Huntington's Disease. *Journal of Communication Disorders*, *39*(1), 62-77. https://doi.org/https://doi.org/10.1016/j.jcomdis.2005.07.001

Jing, S., Mingqing, W., Jinzhou, T., Snowden, J., Xuekai, Z., Jingnian, N., Ting, L., Wenjia, J., Congcong, M., Yanping, T., Jianping, L., Tonghua, L., Pengwen, W., & Yongyan, W. (2014). The Chinese version of story recall: a useful screening tool for mild cognitive impairment and Alzheimer's disease in the elderly. *BMC Psychiatry*, *14*(1), 1-10. https://doi.org/10.1186/1471-244X-14-71

Jokel, R., Seixas Lima, B., Fernandez, A., & Murphy, K. J. (2019). Language in Amnestic Mild Cognitive Impairment and Dementia of Alzheimer’s Type: Quantitatively or Qualitatively Different? *Dementia and Geriatric Cognitive Disorders Extra*, *9*(1), 136-151. https://doi.org/10.1159/000496824

Jongsiriyanyong, S., & Limpawattana, P. (2018). Mild Cognitive Impairment in Clinical Practice: A Review Article. *American Journal of Alzheimer's Disease & Other Dementiasr*, *33*(8), 500-507. https://doi.org/10.1177/1533317518791401

Julayanont, P., Tangwongchai, S., Hemrungrojn, S., Tunvirachaisakul, C., Phanthumchinda, K., Hongsawat, J., Suwichanarakul, P., Thanasirorat, S., & Nasreddine, Z. S. (2015). The Montreal Cognitive Assessment-Basic: A Screening Tool for Mild Cognitive Impairment in Illiterate and Low-Educated Elderly Adults. *J Am Geriatr Soc*, *63*(12), 2550-2554. https://doi.org/10.1111/jgs.13820

Jung, Y., Duffy, J. R., & Josephs, K. A. (2013). Primary progressive aphasia and apraxia of speech. *Seminars in Neurology*, *33*(4), 342-347. https://doi.org/10.1055/s-0033-1359317

Kalaria, R. N. (2016). Neuropathological diagnosis of vascular cognitive impairment and vascular dementia with implications for Alzheimer's disease. *Acta Neuropathol*, *131*(5), 659-685. https://doi.org/10.1007/s00401-016-1571-z

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., & Hudspeth, A. J. (2013). *Principles of Neural Science*. McGraw-Hill. https://www.journals.uchicago.edu/doi/abs/10.1086/670559

Kasisopa, B., G. Reilly, R., Luksaneeyanawin, S., & Burnham, D. (2013). Eye movements while reading an unspaced writing system: The case of Thai. *Vision Research*, *86*, 71-80. https://doi.org/https://doi.org/10.1016/j.visres.2013.04.007

Kellar, S. P., Kelvin, E. A., & Munro, B. H. (2013). *Munro's statistical methods for health care research* (6 th. ed.). Wolters Kluwer Health/Lippincott Williams & Wilkins.

Kemmerer, D. L. (2015). *Cognitive Neuroscience of Language*. Psychology Press. https://books.google.co.th/books?id=nYi9oAEACAAJ

Kent, P. (2013). The Evolution of the Wechsler Memory Scale: A Selective Review. *Applied neuropsychology. Adult*, *20*(4), 277-291. https://doi.org/10.1080/09084282.2012.689267

Khan, A. U. (1986). Memory Assessment. In *Clinical Disorders of Memory* (pp. 41-65). Springer US. https://doi.org/10.1007/978-1-4684-5128-3_2

Khodabakhsh, A., Yesil, F., Guner, E., & Demiroglu, C. (2015). Evaluation of linguistic and prosodic features for detection of Alzheimer’s disease in Turkish conversational speech. *EURASIP Journal on Audio, Speech, and Music Processing*, *2015*(1), 9. https://doi.org/10.1186/s13636-015-0052-y

Kim, N., Kim, J.-H., Wolters, M. K., MacPherson, S. E., & Park, J. C. (2019). Automatic Scoring of Semantic Fluency. *Frontiers in Psychology, 10*(1020). https://doi.org/10.3389/fpsyg.2019.01020

Kirshner, H. S. (2012). Primary progressive aphasia and Alzheimer's disease: Brief history, recent evidence. *Current Neurology and Neuroscience Reports, 12*(6), 709-714. https://doi.org/10.1007/s11910-012-0307-2

Klimova, B., & Kuca, K. (2016). Speech and language impairments in dementia. *Journal of Applied Biomedicine*, *14*(2), 97-103. https://jab.zsf.jcu.cz/artkey/jab-201602-0003.php
http://dx.doi.org/10.1016/j.jab.2016.02.002

König, A., Linz, N., Zeghari, R., Klinge, X., Tröger, J., Alexandersson, J., & Robert, P. H. (2019). Detecting apathy in older adults with cognitive disorders using automatic speech analysis. *Journal of Alzheimer's Disease, 69*(4), 1183-1193. https://doi.org/10.3233/JAD-181033

König, A., Satt, A., Sorin, A., Hoory, R., Derreumaux, A., David, R., & Robert, P. H. (2018). Use of Speech Analyses within a Mobile Application for the Assessment of Cognitive Impairment in Elderly People. *Curr Alzheimer Res*, *15*(2), 120-129. https://doi.org/10.2174/1567205014666170829111942

König, A., Satt, A., Sorin, A., Hoory, R., Toledo-Ronen, O., Derreumaux, A., Manera, V., Verhey, F., Aalten, P., Robert, P. H., & David, R. (2015). Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 1*(1), 112-124. https://doi.org/https://doi.org/10.1016/j.dadm.2014.11.012

Koo, B. M., & Vizer, L. M. (2019). Mobile Technology for Cognitive Assessment of Older Adults: A Scoping Review. *Innovation in aging, 3*(1), igy038-igy038. https://doi.org/10.1093/geroni/igy038

Kulkantrakorn, K. (2018). Dementia and Alzheimer's Disease. *Thammasat Medical Journal*, *18*(3), 442-468.

Ladefoged, P. (2006). *satthasat̄ [A Course in Phonetics]* (A. Tumtavitikul, Trans.; 1 ed.). Thomson Learning.

Ladefoged, P., & Johnson, K. (2015). *A course in phonetics* (7th edition ed.). Wadsworth Australia.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00863

Larner, A. J. (2017). *Cognitive screening instruments: A practical approach* (2nd ed.). Springer International Publishing https://doi.org/10.1007/978-3-319-44775-9

Leal, S. L., Noche, J. A., Murray, E. A., & Yassa, M. A. (2021). Positivity effect specific to older adults with subclinical memory impairment. *Learn Mem, 23*(8), 415-421. https://doi.org/10.1101/lm.042010.116

Lee, M. S., & Kim, B. S. (2021). Effects of working memory intervention on language production by individuals with dementia. *Neuropsychological Rehabilitation, 31*(10), 1557-1581. https://doi.org/10.1080/09602011.2020.1789479

Lee, S., & Lee, D. K. (2018). What is the proper way to apply the multiple comparison test? *Korean J Anesthesiol, 71*(5), 353-360. https://doi.org/10.4097/kja.d.18.00242

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav Brain Sci, 22*(1), 1-38; discussion 38-75. https://doi.org/10.1017/s0140525x99001776

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). Oxford University Press.

Lim, L., Ng, T. P., Ong, A. P., Tan, M. P., Cenina, A. R., Gao, Q., Ng, A., & Kandiah, N. (2018). A novel language-neutral Visual Cognitive Assessment Test (VCAT): validation in four Southeast Asian countries. *Alzheimers Res Ther, 10*(1), 6. https://doi.org/10.1186/s13195-017-0333-z

Limpawattana, P., Tiamkao, S., Sawanyawisuth, K., & Thinkhamrop, B. (2012). Can Rowland Universal Dementia Assessment Scale (RUDAS) replace Mini-mental State Examination (MMSE) for dementia screening in a Thai geriatric outpatient setting? *Am J Alzheimers Dis Other Demen, 27*(4), 254-259. https://doi.org/10.1177/1533317512447886

Lin, J. S., O'Connor, E., Rossom, C., Perdue, L. A., & Eckstrom, E. (2013). Screening for cognitive impairment in older adults: A systematic review for the U.S. preventive services task force. *Annals of Internal Medicine, 159*(9), 601-612. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84887862604&partnerID=40&md5=084d50c720c4237a377430c59bcf5dfb

Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., Ballard, C., Banerjee, S., Burns, A., Cohen-Mansfield, J., Cooper, C., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Larson, E. B., Ritchie, K., Rockwood, K., Sampson, E. L., Samus, Q., Schneider, L. S., Selbaek, G., Teri, L., & Mukadam, N. (2017). Dementia prevention, intervention, and care. *Lancet, 390*(10113), 2673-2734. https://doi.org/10.1016/s0140-6736(17)31363-6

Lopez-De-Ipina, K., Martinez-De-Lizarduy, U., Barroso, N., Ecay-Torres, M., Martinez-Lage, P., Torres, F., & Faundez-Zanuy, M. (2015). Automatic analysis of Categorical Verbal Fluency for Mild Cognitive impartment detection: A non-linear language independent approach. IWOBI 2015 - 2015 International Work Conference on Bio-Inspired Intelligence: Intelligent Systems for Biodiversity Conservation, Proceedings,

Lotrakul, M., Sumrithe, S., & Saipanish, R. (2008). Reliability and validity of the Thai version of the PHQ-9. *BMC Psychiatry*, *8*, 46. https://doi.org/10.1186/1471-244x-8-46

Mack, W. J., Freed, D. M., Williams, B. W., & Henderson, V. W. (1992). Boston Naming Test: Shortened versions for use in Alzheimer's disease. *Journals of Gerontology*, *47*(3), P154-P158. https://doi.org/10.1093/geronj/47.3.P154

Mackenzie, C., Brady, M., Norrie, J., & Poedjianto, N. (2007). Picture description in neurologically normal adults: Concepts and topic coherence. *Aphasiology*, *21*(3-4), 340-354. https://doi.org/10.1080/02687030600911419

Mahon, E., & Lachman, M. E. (2022). Voice biomarkers as indicators of cognitive changes in middle and later adulthood. *Neurobiology of Aging*, *119*, 22-35. https://doi.org/https://doi.org/10.1016/j.neurobiolaging.2022.06.010

Mahrt, T. (2016). *PraatIO*. https://github.com/timmahrt/praatIO

Malek-Ahmadi, M. (2016). Reversion From Mild Cognitive Impairment to Normal Cognition: A Meta-Analysis. *Alzheimer Dis Assoc Disord*, *30*(4), 324-330. https://doi.org/10.1097/wad.0000000000000145

Mandrekar, J. N. (2010). Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *Journal of Thoracic Oncology*, *5*(9), 1315-1316. https://doi.org/10.1097/JTO.0b013e3181ec173d

Manochiopinig, S., Thubthong, N., & Kayasith, P. (2008). Dysarthric speech characteristics of Thai stroke patients. *Disability and Rehabilitation: Assistive Technology*, *3*(6), 332-338. https://doi.org/10.1080/17483100802281368

Marczinski, C. A., & Kertesz, A. (2006). Category and letter fluency in semantic dementia, primary progressive aphasia, and Alzheimer's disease. *Brain and Language*, *97*(3), 258-265. https://doi.org/10.1016/j.bandl.2005.11.001

Marshall, R. C., & Wright, H. H. (2007). Developing a clinician-friendly aphasia test. *Am J Speech Lang Pathol*, *16*(4), 295-315. https://doi.org/10.1044/1058-0360(2007/035)

Martínez-Nicolás, I., Llorente, T. E., Martínez-Sánchez, F., & Meilán, J. J. G. (2021). Ten Years of Research on Automatic Voice and Speech Analysis of People With

Alzheimer's Disease and Mild Cognitive Impairment: A Systematic Review Article. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.620251

Mazrooei Rad, E., Azarnoosh, M., Ghoshuni, M., & Khalilzadeh, M. M. (2021). Diagnosis of mild Alzheimer's disease by EEG and ERP signals using linear and nonlinear classifiers. *Biomedical Signal Processing and Control, 70*, 103049. https://doi.org/https://doi.org/10.1016/j.bspc.2021.103049

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack Jr, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's and Dementia, 7*(3), 263-269. https://doi.org/10.1016/j.jalz.2011.03.005

McMillan, J. H., & Schumacher, S. (2014). *Research in education : evidence-based inquiry*. Pearson.

Meeus, B., Theuns, J., & Van Broeckhoven, C. (2012). The genetics of dementia with Lewy bodies: what are we missing? *Arch Neurol, 69*(9), 1113-1118. https://doi.org/10.1001/archneurol.2011.3678

Meilán, J. J. G., Martínez-Sánchez, F., Carro, J., López, D. E., Millian-Morell, L., & Arana, J. M. (2014). Speech in Alzheimer's Disease: Can Temporal and Acoustic Parameters Discriminate Dementia? *Dementia and Geriatric Cognitive Disorders, 37*(5-6), 327-334. https://doi.org/10.1159/000356726

Miller, B., & Yoon, S. J. (2016). Frontotemporal dementia. In M. Husain & J. M. Schott (Eds.), *Oxford Textbook of Cognitive Neurology and Dementia*. Oxford University Press.

Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Ann Card Anaesth, 22*(1), 67-72. https://doi.org/10.4103/aca.ACA_157_18

Mitchell, A. J., & Malladi, S. (2010). Screening and case-finding tools for the detection of dementia. Part II: Evidence-based meta-analysis of single-domain tests. *American Journal of Geriatric Psychiatry, 18*(9), 783-800. https://doi.org/10.1097/JGP.0b013e3181cdecd6

Monaco, M., Costa, A., Caltagirone, C., & Carlesimo, G. A. (2013). Forward and backward span for verbal and visuo-spatial data: standardization and normative data from an Italian adult population. *Neurol Sci, 34*(5), 749-754. https://doi.org/10.1007/s10072-012-1130-x

Morris, J. C. (1993). The clinical dementia rating (CDR): Current version and scoring rules. *Neurology*, *43*(11), 2412-2414. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0027425211&partnerID=40&md5=c10173b4a72a756955fc7af87eddd6f8

Morris, J. H., & Nagy, Z. (2004). Alzheimer’s disease. In M. M. Esiri, V. M. Y. Lee, & J. Q. Trojanowski (Eds.), *The neuropathology of dementia* (pp. 161-206). Cambridge University Press.

Muangpaisan, W., Intalapaporn, S., & Assantachai, P. (2010). Digit span and verbal fluency tests in patients with mild cognitive impairment and normal subjects in Thai-community. *Journal of the Medical Association of Thailand*, *93*(2), 224-230. https://www.scopus.com/inward/record.uri?eid=2-s2.0-77949752838&partnerID=40&md5=03665076d4d834f0ad03aedaf65fa49a

Muangpaisan, W., Petcharat, C., & Srinonprasert, V. (2012). Prevalence of potentially reversible conditions in dementia and mild cognitive impairment in a geriatric clinic. *Geriatr Gerontol Int*, *12*(1), 59-64. https://doi.org/10.1111/j.1447-0594.2011.00728.x

Mueller, K. D., Du, L., Bruno, D., Betthauser, T., Christian, B., Johnson, S., Hermann, B., & Koscik, R. L. (2022). Item-Level Story Recall Predictors of Amyloid-Beta in Late Middle-Aged Adults at Increased Risk for Alzheimer’s Disease. *Frontiers in Psychology*, *13*, Article 908651. https://doi.org/10.3389/fpsyg.2022.908651

Mueller, K. D., Hermann, B., Mecollari, J., & Turkstra, L. S. (2018). Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks. *J Clin Exp Neuropsychol*, *40*(9), 917-939. https://doi.org/10.1080/13803395.2018.1446513

Mueller, K. D., Koscik, R. L., Du, L., Bruno, D., Jonaitis, E. M., Koscik, A. Z., Christian, B. T., Betthauser, T. J., Chin, N. A., Hermann, B. P., & Johnson, S. C. (2020). Proper names from story recall are associated with beta-amyloid in cognitively unimpaired adults at risk for Alzheimer's disease. *Cortex*, *131*, 137-150. https://doi.org/10.1016/j.cortex.2020.07.008

Mueller, K. D., Koscik, R. L., Hermann, B. P., Johnson, S. C., & Turkstra, L. S. (2018). Declines in Connected Language Are Associated with Very Early Mild Cognitive Impairment: Results from the Wisconsin Registry for Alzheimer's Prevention. *Front Aging Neurosci*, *9*, 437. https://doi.org/10.3389/fnagi.2017.00437

Mueller, K. D., Koscik, R. L., Turkstra, L. S., Riedeman, S. K., LaRue, A., Clark, L. R., Hermann, B., Sager, M. A., & Johnson, S. C. (2016). Connected Language in Late Middle-Aged Adults at Risk for Alzheimer's Disease. *J Alzheimers Dis*, *54*(4), 1539-1550. https://doi.org/10.3233/jad-160252

Munthuli, A., Vongsurakrai, S., Anansiripinyo, T., Ellermann, V., Sroykhumpa, K., Onsuwan, C., Chutichetpong, P., Hemrungrojn, S., Kosawat, K., & Tantibundhit, C. (2021). Thammasat-NECTEC-Chula's Thai Language and Cognition Assessment (TLCA): The Thai Alzheimer's and Mild Cognitive Impairment Screening Test. *Annu Int Conf IEEE Eng Med Biol Soc, 2021*, 690-694. https://doi.org/10.1109/embc46164.2021.9630779

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78*, 679-684.

Na Chiangmai, N., & Wongupparaj, P. (2020). Dementia screening tests in Thai older adults: a systematic review. *Journal of Mental Health of Thailand, 3*(28), 252-265.

Nagarachinda, T., Anansiripinyo, T., Munthuli, A., Tantibandhit, C., Hemrungrojn, S., Kosawat, K., & Onsuwan, C. (2020). Patterns of Thai cohesive devices in elderly with Alzheimer. *Journal of Language and Culture, 39*(1), 59-88. https://so03.tci-thaijo.org/index.php/JLC/article/view/241585/163965

Nasreddine, Z. S. (2017). *MoCA-Paper English 8.3*. https://www.mocatest.org/paper#paper_form_full

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society, 53*(4), 695-699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

National Institute on Aging. (2006). *Rush Alzheimer's Disease Center Load Cognitive Assessment Manual*. Retrieved 14 Dec from https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/document.cgi?study_id=phs000168.v2.p2&phd=1474

Neurological Institute of Thailand. (2014). *Clinical Practice Guidelines: Dementia*. Prasat Neurological Institute.

Neurological Institute of Thailand. (2021). *Clinical Practice Guidelines: Dementia*. Tanapress.

Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *J Speech Hear Res, 36*(2), 338-350. https://doi.org/10.1044/jshr.3602.338

Nichols, E., Steinmetz, J. D., Vollset, S. E., Fukutaki, K., & al, e. (2022). Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. *The Lancet Public Health, 7*(2), e105-e125. https://doi.org/10.1016/S2468-2667(21)00249-8

NIH National Institute on Aging. (2020, 21 October 2020). *Memory, Forgetfulness, and Aging: What's Normal and What's Not?* https://www.nia.nih.gov/health/memory-forgetfulness-and-aging-whats-normal-and-whats-not

Ouchi, Y., Kasai, M., Nakamura, K., Nakatsuka, M., & Meguro, K. (2016). Qualitative Assessment of Instrumental Activities of Daily Living in Older Persons with Very Mild Dementia: The Kurihara Project. *Dementia and Geriatric Cognitive Disorders Extra, 6*(2), 374-381. https://doi.org/10.1159/000446769

Paek, E. J., Murray, L. L., Newman, S. D., & Kim, D.-J. (2019). Test-retest reliability in an fMRI study of naming in dementia. *Brain and Language, 191*, 31-45. https://doi.org/https://doi.org/10.1016/j.bandl.2019.02.002

Panmeta, N. (2015). *Wiyakorn Thai [Thai grammar]*. Research and publication division, faculty of Arts, Chulalongkorn University.

Park, J.-H., Park, H., Sohn, S. W., Kim, S., & Park, K. W. (2017). Memory performance on the story recall test and prediction of cognitive dysfunction progression in mild cognitive impairment and Alzheimer's dementia. *Geriatrics & Gerontology International, 17*(10), 1603-1609. https://doi.org/10.1111/ggi.12940

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine, 256*(3), 183-194. https://doi.org/10.1111/j.1365-2796.2004.01388.x

Petersen, R. C. (2016). Mild Cognitive Impairment. *Continuum (Minneap Minn), 22*(2 Dementia), 404-418. https://doi.org/10.1212/con.0000000000000313

Petersen, R. C., Caracciolo, B., Brayne, C., Gauthier, S., Jelic, V., & Fratiglioni, L. (2014). Mild cognitive impairment: a concept in evolution. *J Intern Med, 275*(3), 214-228. https://doi.org/10.1111/joim.12190

Petersen, R. C., Jack Jr, C. R., Xu, Y. C., Waring, S. C., O'Brien, P. C., Smith, G. E., Ivnik, R. J., Tangalos, E. G., Boeve, B. F., & Kokmen, E. (2000). Memory and MRI-based hippocampal volumes in aging and AD. *Neurology, 54*(3), 581-587. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0034620566&partnerID=40&md5=5ac15be6eb676d4cd80f6f9519fe1760

Petersen, R. C., Lopez, O., Armstrong, M. J., Getchius, T. S. D., Ganguli, M., Gloss, D., Gronseth, G. S., Marson, D., Pringsheim, T., Day, G. S., Sager, M., Stevens, J., & Rae-Grant, A. (2018). Practice guideline update summary: Mild cognitive impairment: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology, 90*(3), 126-135. https://doi.org/10.1212/WNL.0000000000004826

Petersen, R. C., & Morris, J. C. (2005). Mild cognitive impairment as a clinical entity and treatment target. *Archives of neurology*, *62*, 1160–1163. https://doi.org/http://dx.doi.org/10.1001/archneur.62.7.1160

Petersen, R. C., Thomas, R. G., Grundman, M., Bennett, D., Doody, R., Ferris, S., Galasko, D., Jin, S., Kaye, J., Levey, A., Pfeiffer, E., Sano, M., van Dyck, C. H., & Thal, L. J. (2005). Vitamin E and donepezil for the treatment of mild cognitive impairment. *N Engl J Med*, *352*(23), 2379-2388. https://doi.org/10.1056/NEJMoa050151

Pinyopodjanard, S., Suppakitjanusant, P., Lomprew, P., Kasemkosin, N., Chailurkit, L., & Ongphiphadhanakul, B. (2019). Instrumental Acoustic Voice Characteristics in Adults with Type 2 Diabetes. *Journal of Voice*. https://doi.org/https://doi.org/10.1016/j.jvoice.2019.07.003

Pistono, A., Senoussi, M., Guerrier, L., Rafiq, M., Giméno, M., Péran, P., Jucla, M., & Pariente, J. (2021). Language Network Connectivity Increases in Early Alzheimer's Disease. *J Alzheimers Dis*, *82*(1), 447-460. https://doi.org/10.3233/jad-201584

Poil, S.-S., De Haan, W., van der Flier, W., Mansvelder, H., Scheltens, P., & Linkenkaer-Hansen, K. (2013). Integrative EEG biomarkers predict progression to Alzheimer's disease at the MCI stage. *Frontiers in Aging Neuroscience*, *5*. https://doi.org/10.3389/fnagi.2013.00058

Postma, A. (2000). Detection of errors during speech production: a review of speech monitoring models. *Cognition*, *77*(2), 97-132. https://doi.org/https://doi.org/10.1016/S0010-0277(00)00090-1

Power, M., Fell, G., & Wright, M. (2013). Principles for high-quality, high-value testing. *Evidence Based Medicine*, *18*(1), 5. https://doi.org/10.1136/eb-2012-100645

Privitera, G. J. (2015). *Statistics for the behavioral sciences, 2nd ed*. Sage Publications, Inc.

Prud'hommeaux, E. T., Mitchell, M., & Roark, B. (2011). Using Patterns of Narrative Recall for Improved Detection of Mild Cognitive Impairment. Rehabilitation Engineering and Assistive Technology Society of North America (RESNA) and 3rd International Conference on Technology and Aging (ICTA),

Pulido, M. L. B., Hernández, J. B. A., Ballester, M. Á. F., González, C. M. T., Mekyska, J., & Smékal, Z. (2020). Alzheimer's disease and automatic speech analysis: A review. *Expert Systems with Applications*, *150*, Article 113213. https://doi.org/10.1016/j.eswa.2020.113213

Qiao, Y., Xie, X. Y., Lin, G. Z., Zou, Y., Chen, S. D., Ren, R. J., & Wang, G. (2020). Computer-Assisted Speech Analysis in Mild Cognitive Impairment and Alzheimer's Disease: A Pilot Study from Shanghai, China. *J Alzheimers Dis, 75*(1), 211-221. https://doi.org/10.3233/jad-191056

Ramoo, D. (2021). *Psychology of language* (7th ed.). BCcampus. https://opentextbc.ca/psyclanguage/

Rapčan, V., D'Arcy, S., Penard, N., Robertson, I., & Reilly, R. (2009). *The use of telephone speech recordings for assessment and monitoring of cognitive function in elderly people*. https://doi.org/10.21437/Interspeech.2009-282

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk , Kolmogorov-Smirnov , Lilliefors and Anderson-Darling tests.

Roark, B., Hosom, J.-P., Mitchell, M., & Kaye, J. A. (2007, June). Automatically derived spoken language markers for detecting mild cognitive impairment. 2nd International Conference on Technology and Aging (ICTA), Toronto, ON.

Roark, B., Mitchell, M., Hosom, J., Hollingshead, K., & Kaye, J. (2011). Spoken Language Derived Measures for Detecting Mild Cognitive Impairment. *IEEE Transactions on Audio, Speech, and Language Processing, 19*(7), 2081-2090. https://doi.org/10.1109/TASL.2011.2112351

Robin, J., Xu, M., Balagopalan, A., Novikova, J., Kahn, L., Oday, A., Hejrati, M., Hashemifar, S., Negahdar, M., Simpson, W., & Teng, E. (2023). Automated detection of progressive speech changes in early Alzheimer's disease. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring, 15*(2), Article e12445. https://doi.org/10.1002/dad2.12445

Rohrer, J. D., Rossor, M. N., & Warren, J. D. (2012). Alzheimer's pathology in primary progressive aphasia. *Neurobiology of Aging, 33*(4), 744-752. https://doi.org/10.1016/j.neurobiolaging.2010.05.020

Rohrer, J. D., & Warren, J. D. (2016). Primary progressive aphasia. In M. Husain & J. M. Schott (Eds.), *Oxford Textbook of Cognitive Neurology and Dementia*. Oxford University Press. https://doi.org/10.1093/med/9780199655946.003.0041

Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., & Petersen, R. C. (2014). Classifying neurocognitive disorders: the DSM-5 approach. *Nature Reviews Neurology, 10*(11), 634-642. https://doi.org/10.1038/nrneurol.2014.181

Sánchez Cabaco, A., De La Torre, L., Alvarez Núñez, D. N., Mejía Ramírez, M. A., & Wöbbeking Sánchez, M. (2023). Tele neuropsychological exploratory assessment of indicators of mild cognitive impairment and autonomy level in

Mexican population over 60 years old. *PEC Innovation*, *2*, 100107. https://doi.org/https://doi.org/10.1016/j.pecinn.2022.100107

Sangchocanonta, S., Vongsurakrai, S., Sroykhumpa, K., Ellermann, V., Munthuli, A., Anansiripinyo, T., Onsuwan, C., Hemrungrojn, S., Kosawat, K., & Tantibundhit, C. (2021). Development of Thai Picture Description Task for Alzheimer's Screening using Part-of-Speech Tagging. *Annu Int Conf IEEE Eng Med Biol Soc*, *2021*, 2104-2109. https://doi.org/10.1109/embc46164.2021.9629861

Satt, A., Sorin, A., Toledo-Ronen, O., Barkan, O., Kompatsiaris, I., Kokonozi, A., & Tsolaki, M. (2013). *Evaluation of Speech-Based Protocol for Detection of Early-Stage Dementia*.

Schmand, B., Eikelenboom, P., van Gool, W. A., & Initiative, t. A. s. D. N. (2011). Value of Neuropsychological Tests, Neuroimaging, and Biomarkers for Diagnosing Alzheimer's Disease in Younger and Older Age Cohorts. *Journal of the American Geriatrics Society*, *59*(9), 1705-1710. https://doi.org/https://doi.org/10.1111/j.1532-5415.2011.03539.x

Schmid, H.-J. o. (2007). Entrenchment, Salience, and Basic Levels. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of Cognitive Linguistics* (pp. 117-138). Oxford University Press.

Schweitzer, A. (2016). *Praat Scripting*. https://www.ims.uni-stuttgart.de/institut/arbeitsgruppen/ehemalig/ep-dogil/helps/praat-scripting/praat.tutorial.pdf

Schwering, S. C., & MacDonald, M. C. (2020). Verbal Working Memory as Emergent from Language Comprehension and Production. *Frontiers in Human Neuroscience*, *14*. https://doi.org/10.3389/fnhum.2020.00068

Seghier, M. L. (2013). The angular gyrus: multiple functions and multiple subdivisions. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*, *19*(1), 43-61. https://doi.org/10.1177/1073858412440596

Shim, Y. S., Yang, D. W., Kim, H.-J., Park, Y. H., & Kim, S. (2017). Characteristic differences in the mini-mental state examination used in Asian countries. *BMC neurology*, *17*(1), 141-141. https://doi.org/10.1186/s12883-017-0925-z

Silpakit, O. (2013). Review article: Dementia screening. *Journal of Mental Health of Thailand*, *21*(1), 34-47.

Silpakit, O., Chomchuen, S., & Kittipongpisal, S. (2018). The diagnostic accuracy of the Thai dementia screening scale for primary care in the community. *Neurology Asia*, *23*(4), 333-341. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85059591917&partnerID=40&md5=225ef04f394ae6b5e15bb0c03cba0b8c

Silpakit, O., Silpakit, C., & Pukdeenaul, P. (2007). A Comparison Study of Cognitive Impairment Screening Tools: CDT, IQCODE VS MMSE. *Siriraj Med J, 59*, 361-363. https://www.academia.edu/28571211/A_Comparison_Study_of_Cognitive_Impairment_Screening_Tools_CDT_IQCODE_VS_MMSE

Silpakit, O., Silpakit, P., Kittipongpisarn, S., & Chomcheun, R. (2017). The diagnostic accuracy of the Thai version of Brief Community Screening Instrument for Dementia (CSI-D). *J Ment Health Thai*, *25*(1), 32-46. https://he01.tci-thaijo.org/index.php/jmht/article/view/129516/97333

Silpakit, O., Sukying, C., Chomchuen, R., & Kittipongpisarn, S. (2017). The assessment tools for Thai dementia in and above primary care. *Journal of Mental Health of Thailand*, *25*(2), 69-83.

Singh, S., Bucks, R. S., & Cuerden, J. M. (2001). Evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech. *Aphasiology, 15*(6), 571-583. https://doi.org/10.1080/02687040143000041

Skodda, S., Groenheit, W., Schlegel, U., Suedmeyer, M., Schnitzler, A., & Wojtecki, L. (2014). Effect of Subthalamic Stimulation on Voice and Speech in Parkinson’s Disease: For the Better or Worse? *Frontiers in Neurology, 4*(218). https://doi.org/10.3389/fneur.2013.00218

Slegers, A., Filiou, R. P., Montembeault, M., & Brambati, S. M. (2018). Connected Speech Features from Picture Description in Alzheimer's Disease: A Systematic Review. *Journal of Alzheimer's disease : JAD, 65*(2), 519-542. https://doi.org/10.3233/JAD-170881

Smith, G. E., & Bondi, M. W. (2013). *Mild cognitive impairment and dementia: Definitions, diagnosis, and treatment*. Oxford University Press.

Sofia, S., Paliokas, I., Tzovaras, D., Karagiannidis, C., & Tsolaki, M. (2016). *A New Method for the Assessment of Discourse Comprehension Abilities of Older Adults with Mild Cognitive Impairment*. https://doi.org/10.1145/2910674.2935848

Song, M., Lee, S. H., Yu, K.-H., & Kang, Y. (2019). Development and Validation of the Full Version of Story Memory in the Korean-Mini Mental State Examination, 2nd Edition: Expanded Version (K-MMSE-2: EV). *Dement Neurocogn Disord, 18*(3), 96-104. https://doi.org/10.12779/dnd.2019.18.3.96

Sonnen, J. A., Montine, K. S., Quinn, J. F., Kaye, J. A., Breitner, J. C. S., & Montine, T. J. (2008). Biomarkers for cognitive impairment and dementia in elderly people. *The Lancet. Neurology, 7*(8), 704-714. https://doi.org/10.1016/S1474-4422(08)70162-5

Steeman, E., de Casterlé, B. D., Godderis, J., & Grypdonck, M. (2006). Living with early-stage dementia: a review of qualitative studies. *J Adv Nurs, 54*(6), 722-738. https://doi.org/10.1111/j.1365-2648.2006.03874.x

Stenström, A.-B. (2012). Pauses and hesitations. In A. Gisle & A. Karin (Eds.), *Pragmatics of Society* (pp. 537-568). De Gruyter Mouton. https://doi.org/doi:10.1515/9783110214420.537

Suanpirintr, S., Thubthong, N., & Suputtitada, A. (2008). *A preliminary study: Acoustic analysis for Parkinson's disease patients focusing on disease state*.

Szatloczki, G., Hoffmann, I., Vincze, V., Kalman, J., & Pakaski, M. (2015). Speaking in Alzheimer's Disease, is That an Early Sign? Importance of Changes in Language Abilities in Alzheimer's Disease. *Frontiers in Aging Neuroscience, 7*, 195-195. https://doi.org/10.3389/fnagi.2015.00195

Tachibana, M., Holm, M. L., Liu, C. C., Shinohara, M., Aikawa, T., Oue, H., Yamazaki, Y., Martens, Y. A., Murray, M. E., Sullivan, P. M., Weyer, K., Glerup, S., Dickson, D. W., Bu, G., & Kanekiyo, T. (2019). APOE4-mediated amyloid-β pathology depends on its neuronal receptor LRP1. *J Clin Invest, 129*(3), 1272-1277. https://doi.org/10.1172/jci124853

Tangwongchai, S., Charernboon, T., Phannasathit, M., Akkayagorn, L., Hemrungrojn, S., Phanthumchinda, K., & Nasreddine, Z. S. (2009). *The validity of thai version of the Montreal Cognitive Assessment (MoCA-T)* International Psychogeriatric Association 14th International Congress, Montreal, Canada. http://demneuropsy.com.br/detalhe_artigo.asp?id=164

Tangwongchai, S., Hemrungrojn, S., Likitjaroen, Y., Worakul, P., Tepmongkol, S., & Tunvirachaisakul, C. (2015). *Development of the CERAD for assessment and registry of Thai elderly with Alzheimer's desease*. https://dric.nrct.go.th/index.php?/Search/SearchDetail/297698

Taylor, C. (2020). *What Is the Interquartile Range Rule?* ThoughtCo. thoughtco.com/what-is-the-interquartile-range-rule-3126244

Themistocleous, C., Eckerstrom, M., & Kokkinakis, D. (2018). Identification of Mild Cognitive Impairment From Speech in Swedish Using Deep Sequential Neural Networks. *Front Neurol, 9*, 975. https://doi.org/10.3389/fneur.2018.00975

Themistocleous, C., Eckerström, M., & Kokkinakis, D. (2020). Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from Healthy Controls. *PLoS ONE, 15*(7), e0236009. https://doi.org/10.1371/journal.pone.0236009

Tomoeda, C. K., Bayles, K. A., Trosset, M. W., Azuma, T., & McGeagh, A. (1996). Cross-sectional analysis of Alzheimer disease effects on oral discourse in a picture description task. *Alzheimer Dis Assoc Disord, 10*(4), 204-215. https://doi.org/10.1097/00002093-199601040-00006

Toth, L., Hoffmann, I., Gosztolya, G., Vincze, V., Szatloczki, G., Banreti, Z., Pakaski, M., & Kalman, J. (2018). A Speech Recognition-based Solution for the Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech. *Curr Alzheimer Res, 15*(2), 130-138. https://doi.org/10.2174/1567205014666171121114930

Train The Brain Forum Commitee. (1993). Thai mental state examination (TMSE). *Siriraj Hospital Gazette, 45*(6), 359-374.

Treviso, M., Borges dos Santos, L., Shulby, C., Hubner, L., Mansur, L., & Aluisio, S. (2018). Detecting mild cognitive impairment in narratives in Brazilian Portuguese: first steps towards a fully automated system. *Letras de Hoje, 53*(1), 48-58. https://doi.org/10.15448/1984-7726.2018.1.30955

Tröger, J., Linz, N., König, A., Robert, P., & Alexandersson, J. (2018, May 2018,). Telephone-based Dementia Screening I: Automated Semantic Verbal Fluency Assessment. 12th EAI International Conference on Pervasive Computing Technologies for Healthcare (Pervasive Health), New York, NY, USA.

van den Heuvel, M. P., & Hulshoff Pol, H. E. (2010). Exploring the brain network: A review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology, 20*(8), 519-534. https://doi.org/10.1016/j.euroneuro.2010.03.008

Vecchio, F., Babiloni, C., Lizio, R., De Vico Fallani, F., Blinowska, K., Verrienti, G., Frisoni, G., & Rossini, P. M. (2013). Chapter 15 - Resting state cortical EEG rhythms in Alzheimer's disease: toward EEG markers for clinical applications: a review. In E. Başar, C. Başar-Eroĝlu, A. Özerdem, P. M. Rossini, & G. G. Yener (Eds.), *Supplements to Clinical Neurophysiology* (Vol. 62, pp. 223-236). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-7020-5307-8.00015-6

Vichitvejpaisal, P., Preechakoon, B., Supaprom, W., Sriputtaruk, S., Rodpaewpaln, S., Saen-Ubol, R., & Vessuwan, K. (2015). The Montreal Cognitive Assessment as a Screening Tool for Preoperative Cognitive Impairment in Geriatric Patients. *J Med Assoc Thai, 98*(8), 782-789.

Wechsler, D. (2009). *WMS-IV: Wechsler memory scale-fourth edition*. Pearson.

Weekes, B. S. H. (2020). Aphasia in Alzheimer’s Disease and Other Dementias (ADOD): Evidence From Chinese. *American Journal of Alzheimer's Disease & Other Dementias®, 35*, 1533317520949708. https://doi.org/10.1177/1533317520949708

Weiner, J., Engelbart, M., & Schultz, T. (2017). *Manual and Automatic Transcriptions in Dementia Detection from Speech*. https://doi.org/10.21437/Interspeech.2017-112

Weiner, J., Herff, C., & Schultz, T. (2016). *Speech-Based Detection of Alzheimer's Disease in Conversational German*. https://doi.org/10.21437/Interspeech.2016-100

Willems, R. M. (2015). Cognitive neuroscience of natural language use. http://ezproxy.uniandes.edu.co:8080/login?url=https://doi.org/10.1017/CBO9781107323667

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L. O., Nordberg, A., Backman, L., Albert, M., Almkvist, O., Arai, H., Basun, H., Blennow, K., de Leon, M., DeCarli, C., Erkinjuntti, T., Giacobini, E., Graff, C., Hardy, J., Jack, C., Jorm, A., Ritchie, K., van Duijn, C., Visser, P., & Petersen, R. C. (2004). Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J Intern Med*, *256*(3), 240-246. https://doi.org/10.1111/j.1365-2796.2004.01380.x

Wongupparaj, P., & Wongupparaj, R. (2012). 39 Years of Working Memory Model: Research and Its Applications. *Research Methodology & Cognitive Science*, *10*(2), 1-16.

World Health Organization. (2017). *Global action plan on the public health response to dementia 2017–2025*.

World Health Organization. (2018). *The Global Dementia Observatory Reference Guide* https://www.who.int/publications/i/item/the-global-dementia-observatory

World Health Organization. (2023, 15 March 2023). *Dementia*. https://www.who.int/news-room/fact-sheets/detail/dementia

Xiu, N., Vaxelaire, B., Li, L., Ling, Z., Xu, X., Huang, L., Sun, B., Huang, L., & Sock, R. (2022). A Study on Voice Measures in Patients With Alzheimer's Disease. *Journal of Voice*. https://doi.org/https://doi.org/10.1016/j.jvoice.2022.08.010

Yamada, M., Komatsu, J., Nakamura, K., Sakai, K., Samuraki-Yokohama, M., Nakajima, K., & Yoshita, M. (2020). Diagnostic Criteria for Dementia with Lewy Bodies: Updates and Future Directions. *Journal of movement disorders*, *13*(1), 1-10. https://doi.org/10.14802/jmd.19052

Yanick Jadoul. (2022). *Parselmouth*. https://parselmouth.readthedocs.io/en/stable/

Yimtae, K., Kasemsiri, P., Thanawirattananit, P., & Siriphaopradit, P. (2011). *The Evaluation of Thai-version Five Minute Hearing Test to Screen Hearing in*

Community. http://digital.nlt.go.th/items/show/9120?advanced%5B%5D%5Belement_id%5D=49&advanced%5B%5D%5Btype%5D=is+exactly&advanced%5B%5D%5Bterms%5D=%E0%B9%80%E0%B8%A7%E0%B8%8A%E0%B8%9B%E0%B8%8F%E0%B8%B4%E0%B8%9A%E0%B8%B1%E0%B8%95%E0%B8%B4

Yingthawornsuk, T. (2016, 28 Nov.-1 Dec. 2016). Spectral Entropy in Speech for Classification of Depressed Speakers. 2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS),

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal*, *11*(2), 49-54. https://doi.org/10.21315/eimj2019.11.2.6

Zeki, D. A., & Hillis, A. E. (2016). Acquired disorders of language and speech. In M. Husain & J. M. Schott (Eds.), *Oxford Textbook of Cognitive Neurology and Dementia*. Oxford University Press. https://doi.org/10.1093/med/9780199655946.003.0041

Zhou, M., Huang, T., Collins, N., Zhang, J., Shen, H., Dai, X., Xiao, N., Wu, X., Wei, Z., York, J., Lin, L., Zhu, Y., LaDu, M. J., & Chen, X. (2016). APOE4 Induces Site-Specific Tau Phosphorylation Through Calpain-CDK5 Signaling Pathway in EFAD-Tg Mice. *Curr Alzheimer Res*, *13*(9), 1048-1055. https://doi.org/10.2174/1567205013666160415154550

APPENDICES

APPENDIX A

Instruction and scoring rules for the Thai Picture Description task (TPD)

The Thai Picture Description Task (TPD)

Instructions for administration

1. An examiner hands the picture (printed on A4 size paper) to a participant.
2. An examiner asks a participant to tell the story pertaining to the portrayed objects, persons, and actions by saying, ‘Please tell me what is going on in this picture as much as you can?’ ‘ช่วยเล่าเรื่องจากภาพที่เห็นให้ได้มากที่สุด’.
3. Allow a participant to verbally create the story without disruption, inquiry, hint, or pointing at the picture from an examiner.
    - If a participant is unable to tell the story, an examiner can encourage the participant by saying, ‘What do you see in the picture? You may create the story from what you see.’ ‘เห็นอะไรในรูปบ้าง เล่าเรื่องจากสิ่งที่เห็นได้เลย’
    - If a participant starts to mention each object, e.g., ‘This is a boy. This is a girl’ or ‘There are TV, curtain, broom, and coconut,’ an examiner should ask them not to describe the picture but tell the story in general.
    - The prompt can only deliver once for a helping person who might misunderstand the prior instruction. Despite the participants still pointing at each object after receiving this prompt, the session can continue.
4. While a participant is telling the story, an examiner may provide physical encouragement by nodding simultaneously to the context and avoiding making noise.
    - Nodding should not be used as an answer to the inquiries asked by a participant.
5. When a participant mentions the ending of his/her story or is quiet for more than 10 seconds, an examiner may ask a question to confirm the ending of the story by saying, ‘Is there anything you want to say’ ‘มีอะไรอยากเล่าเพิ่มอีกหรือไม่’ or ‘What else is happening’ ‘มีอะไรเกิดขึ้นอีกในภาพ.’
    - A participant may finish or continue the story after an examiner queries.
    - This activity ends when a participant wants to finish the story

- If a participant tells the story for more than five minutes, an examiner may ask the participant to conclude his/her story.

Rules for counting correct information units (CIUs)

Scoring instruction: Mark a color highlight or circle on the words that are included in the CIUs.

E.g., ครอบครัว | หนึ่ง | น่าจะ | มี | ลูก | 2 | คน

‘One family seems to have two children.’

Correct information units (CIUs) are the smallest unit that contains meaning (can rely on the Thai tokenization program of Aroonmanakun (2002)), and are relevant to the picture. CIUs are characterized by comprehensible context, informative content, and accuracy in relation to the picture. Every single unit featured in the aforementioned characteristics is included in the CIUs. The total number of CIUs is considered as scores for the TPD.

Incorrect information units are words that would not be counted as CIUs. The characteristics of non-scoring words are established from the study of Nicholas and Brookshire (1993) and the Thai grammar of Panmeta (2015). In this section, words that are not given a score are put a diagonal slash through.

1. Words that inaccurately portray what happens or depicts in the given illustration.

E.g., ~~พ่อแม่ | ของ | เด็ก | ใจร้าย~~

‘~~Their parents~~ (of the girl and the boy) ~~are mean~~.’

2. Repetition of words or ideas that do not add new information to the story is not necessary and are not intensify meaning.

E.g., เด็กผู้ชาย | กำลัง | ถู |บ้าน | ~~ทำความสะอาด | บ้าน~~

‘The boy is wiping the floor, ~~cleaning the house~~.’

3. Attempt to correct articulating errors except for the final correct attempt.

E.g., เด็กผู้ชาย | กำลัง | ถู | ~~ฟื้น~~ | พื้น |

‘The boy is wiping the ~~four~~, floor.’

4. Dead ends, false starts, or revisions that are incomplete or informative.

E.g., ~~พี่สาว | ไม่ | ไม่~~ | แม่ | กำลัง | ล้าง | จาน |

‘~~Sister, no no.~~ The mother is washing the dishes.

5. Vague or nonspecific words that are not necessary for the completeness of a sentence and for which the subject has not provided a clear referent.

E.g., ~~อะไรประมาณนั้น~~, ~~อันเนี่ย~~

something like that, this stuff

6. Conjunctive terms that are not used as cohesive ties but as a filler or continuant.

E.g., particularly ‘so’, ‘and’ and ‘then’ in Thai ~~แล้ว~~, ~~แล้วก็~~, ~~ในส่วนของ~~

7. Filler words or interjections that do not convey information about the story regarding the picture and tag questions.

E.g., ~~ก็~~, ~~แล้วก็~~, ~~แบบว่า~~, ~~ใช่ไหม~~, ~~ค่ะ~~, ~~ครับ~~

then, and then, sort of, isn’t it, a polite ending particle for females and males

8. Quantifiers and modifiers that are used as a filler or added unnecessarily to clear the content of the picture.

E.g., ~~จริง ๆ แล้ว~~, ~~อันที่จริง~~, ~~คิดว่านะ~~, ~~เหมือนกับ/จะ~~, ~~ประมาณว่า~~

apparently, evidently, I think that, seems like, sort of

9. Additive words that are placed frequently at the end of sentences, especially in Thai spoken statements.

E.g., ~~ละ~~, ~~นะ~~, ~~ซิ~~, ~~เถอะ~~, ~~มั้ง~~, ~~หรอก~~, ~~แน่ะ~~, ~~เนี่ย~~, ~~แหละ~~

10. Commentary on the task and lead-in phrases that do not give information about the picture or story.

E.g., ~~ในรูปนี้...~~, ~~รูปนี้เป็นลายเส้น~~, ~~อย่างที่เห็นก็คือ...~~

In this picture…, This is a line drawing picture., As you see in this picture…

11. Commentary on the participant’s performance or personal experiences.

E.g., ~~สมัยก่อนนะป้า/ลุงต้องทำงานบ้าน~~, ~~เล่าไม่ได้~~, ~~ให้เล่ายังไง~~, ~~ไม่รู้ว่าทำไมบ้านสกปรกขนาดนี้~~

In the past, I had to do chores; I could not tell a story; how to tell a story; I don’t know why this house is so dirty.

APPENDIX B

Instruction and scoring criteria for the Thai Story Recall task (TSR)

The Thai Story Recall Task (TSR)

Instructions for administration

1. An examiner tells a participant that he/she needs to listen carefully to the short story because he/she will be asked to tell the whole story at the end of the listening.
2. When a participant is ready and pays attention to an examiner, an examiner reads the story to the participant at a normal to slow pace. It takes approximately 90 seconds to finish reading.
3. Then, an examiner asks a participant to recall the story as much as he/she can by saying, ‘Please tell the story you have listened to me as much as you can.’ ‘ช่วยเล่าเรื่องที่ฟังไปเมื่อสักครู่ให้ได้มากที่สุด’.

    - While a participant retells the story, an examiner should not interrupt or provide a cue.
    - There is no time limit for recalling.
    - An examiner can encourage by saying, ‘You may say a single word that you can remember.’ ‘พูดเป็นคำ ๆ ที่จำได้ก็ได้’.
    - This is an ‘immediate recall round.’

4. When a participant finishes his/her recall, an examiner asks five factual inquiries one at a time and in the same order. An examiner is not allowed to provide a cue of the answers.

    1) ใครกำลังเดินทาง Who is travelling?
    2) เขากำลังไปที่ไหน Where is he going?
    3) หลานอายุเท่าไร How old is the kid?
    4) ภูเขาเป็นแบบไหน What kind of mountain?
    5) สาเหตุที่รถช้าเพราะอะไร What causes a delay of the train?

5. After finishing the five questions, an examiner prompts a participant that he/she will be asked to tell the same story again by saying, ‘Please remember the story because I will ask you to tell the story again’ ‘กรุณาจำเรื่องที่ฟังไปเมื่อสักครู่ไว้เพราะเดี๋ยวจะขอให้เล่าอีกครั้งหนึ่ง.’
6. During a delayed period, an examiner conducts filler tasks, including animal fluency, letter fluency (‘ก’ Kor), and semi-structured interview. A delayed period should last 10 minutes.
7. After 10 minutes of delayed intervals, an examiner asks a participant to tell the story again by following the instruction in item no. 3.
    - This is a ‘delayed recall round.’

Scoring rules for recall units

Scoring instruction: Mark the underlying words (important words) when a participant says those words. The total number of correctly recalled units is the score for each round.

✓ The important word will be scored when it is told with the correct storyline or said a single word without context.

✗ The important word will not be scored when it is told with an incorrect storyline, e.g., he departs from Chonburi. No score is given to Chonburi.

- In case a participant tells the same content twice, and the information is different, only the last telling is considered.

Scoring guidelines

Unit	Story detail	Scorable words
1	พง phong	พง Phong
2	เป็นคนชลบุรี pen khon Chonburi	ชลบุรี Chonburi, ชลฯ Chon
3	ย้ายมาอยู่ที่อีสาน yá:y ma: yù: thî: i:-sǎ:n	อีสาน i:-sǎ:n

Unit	Story detail	Scorable words
4	ได้ 4 ปี daî sì: pi:	4 ปี sì: pi:, 4 ขวบ sì: khùa:p
5	เขาออกจากอุบลฯ khăo ɔ̀:k jà:k u-bon	อุบลฯ u-bon, อุบลราชธานี u-bon rat cha tha ni
6	จะเอาลูกดอก jà ao lú:k dɔ̀:k	ลูกดอก lú:k dɔ̀:k
7	ของเล่น khɔ̌:ng lên	ของเล่น khɔ̌:ng lên
8	ไปให้น้องบอล pai haî nɔ́:ng bɔ:n	น้องบอล nɔ́:ng bɔ:n, คนชื่อบอล bɔ:n
9	หลาน lă:n	หลาน lă:n
10	อายุ 3 ขวบ a:-yú să:m khùa:p	3 ขวบ să:m khùa:p, 3 ปี să:m pi:
11	ขณะที่นั่งรถไฟไปนั้น khà-nà thî: nâng rót-fai pai nán	รถไฟ rót-fai
12	เกิดแผ่นดินไหว kə̀:t phàe:n-din waĭ	แผ่นดินไหว phàe:n-din waĭ
13	ที่เมืองพล thî: muea:ng phon	เมืองพล muea:ng phon, อำเภอพล amphoe phon
14	รถไฟขยับได้ช้า ๆ rót-fai khà-yàp daî chá: chá:	ช้า ๆ chá: chá:, ไปช้า ๆ pai chá: chá:
15	ผ่านภูเขา phà:n phu:-khăo	ภูเขา phu:-khăo
16	หินดินดานไป hĭn din-da:n pai	หินดินดาน hĭn din-da:n, ดินดาน din-da:n
17	พอเริ่มเข้าตัวเมือง phɔ́: rə̂:m khâo tua muea:ng	เข้าเมือง khâo muea:ng , เข้าตัวเมือง khâo tua muea:ng, เห็นเมือง hen muea:ng, เห็นตัวเมือง hen tua muea:ng
18	มองเห็นวงเวียนแล้ว mɔ:ng hĕn wong-wia:n láe:w	วงเวียน wong-wia:n

Unit	Story detail	Scorable words
19	จึงรู้ว่าใกล้ถึงสถานี jueng rú: wâ: klaî thŭeng sa-thă:-ni:	สถานี sa-thă:-ni:, ถานี thă:-ni:
20	กว่าจะถึงอุดรฯ kwà: jà thŭeng u-dɔ:n	อุดรฯ u-dɔ:n, อุดรธานี u-dɔ:n tha ni
21	ก็เสียเวลาไปแล้ว 2 ชั่วโมง kɔ̂: sĭa: we:-la: pai láe:w sɔ̌:ng chûa-mo:ng	2 ชั่วโมง sɔ̌:ng chûa-mo:ng

Scoring rules for five factual inquiries

Scoring instruction: Write down the answers generated by a participant. Give one point to a correct answer. Correct or acceptable answers to each question are listed here.

1) ใครกำลังเดินทาง Who is travelling? - พง phong
2) เขากำลังไปที่ไหน Where is he going? – อุดรฯ u-dɔ:n, อุดรธานี u-dɔ:n tha ni
3) หลานอายุเท่าไร How old is the kid? – 3 ขวบ, 3 ปี, 3 (three years or three)
4) ภูเขาเป็นแบบไหน What kind of the mountain? – หินดินดาน hĭn din-da:n, ดินดาน din-da:n
5) สาเหตุที่รถช้าเพราะอะไร What causes a delay of the train? – แผ่นดินไหว phàe:n-din waĭ

APPENDIX C

Materials for extracting the frequency-based variables

Python Script for Extracting the Frequency-based Variables

Install Parselmouth

https://parselmouth.readthedocs.io/en/stable/

Script of ‘speech_analysis_praat.py’

```
import parselmouth
from parselmouth.praat import call , run_file
import numpy as np
import librosa
import argparse

parser = argparse.ArgumentParser(description="Speech analysis for detecting dementia")
parser.add_argument("--src", type=str,  required=True,  help="Source signal",    )
parser.add_argument("--verbose",  action='store_true', help="verbose")
parser.add_argument("--seg",type=str,   default=None, help="segmentation" )
parser.add_argument("--rep",type=str,   default="rep.txt", help="segmentation"  )
parser.add_argument("--process_label",type=str,   default="pd", help="segmentation"  )
parser.add_argument("--sr",type=int,   default=16000, help="sampling rate" )
args = parser.parse_args()
src_wav_path  = args.src
sr=args.sr
seg_path = args.seg
process_label=args.process_label
report_file=args.rep
F0min=60
F0max=600

print(f"Reading {src_wav_path}")
waveform= parselmouth.Sound(src_wav_path).convert_to_mono().resample(sr)
```

```
with open(report_file, 'w') as f:
    if seg_path is not None:
        print(f"Reading {seg_path}")
        seg_file = open(seg_path, 'r')
        for seg in seg_file.read().splitlines():
            start_sec,end_sec,label = seg.split()
            if label == process_label:
                print(f"Processing segment [{start_sec}-{end_sec}] with
label {label}")
                print(f"Processing segment [{start_sec}-{end_sec}] with
label {label}",file=f)

snd=waveform.extract_part(from_time=float(start_sec),to_time=float(end_sec)
)
                pitch = call(snd, "To Pitch", 0.0, F0min, F0max)
                pulse = call([snd, pitch], "To PointProcess (cc)")
                voice_report = call([snd, pitch, pulse], "Voice report",
0.0, 0.0, F0min, F0max, 1.3, 1.6, 0.03, 0.45).split(chr(10))
                print("Standart Praat voice report",file=f)
                for single in voice_report:
                    print(single,file=f)
    else:
         snd=waveform
         pitch = call(snd, "To Pitch", 0.0, F0min, F0max)
         pulse = call([snd, pitch], "To PointProcess (cc)")
         voice_report = call([snd, pitch, pulse], "Voice report", 0.0, 0.0,
F0min, F0max, 1.3, 1.6, 0.03, 0.45).split(chr(10))
         print("Standart Praat voice report",file=f)
         for single in voice_report:
             print(single,file=f)

print(f"Written {report_file}")
```

Usage of ‘speech_analysis_praat.py’

```
python speech_analysis_praat.py --src seg01.wav --seg seg01.txt --
process_label "pd" --rep seg01.rep
```

This script extracts segments with the label ‘pd’ (picture description) and writes the related statistics on file ‘seg01.rep’. File ‘seg01.wav’ is an audio file with .wav format of participant no.1.

APPENDIX D

Script for the Voice Activity Detection technique (VAD) in Praat

Script for Making a Textgrid Object in Praat by Using ‘To TextGrid (silences)’

```
Read from file: "D:\WAV\dn01.wav"
Convert to mono
selectObject: "Sound dn01_mono"
To TextGrid (silences): 100, 0, -30, 0.5, 0.2, "P", "U5 and 6"
selectObject: "Sound dn01_mono"
plusObject: "TextGrid dn01_mono"
View & Edit
```

This script converts the ‘dn01.wav’ from stereo to a mono audio file, then makes the textgrid object from the provided parameter in line 4. The file name should be addressed at lines 1, 5, and 6 correspondingly.

APPENDIX E

Materials for making intersection in textgrid object

Python script for making intersection in textgrid object

Install Praatio

https://pypi.org/project/praatio/3.6.8/

Script of ‘intersect.py’

```
from praatio import textgrid
from os.path import join

inputFN="C:\Users\pyintersect\merged03.TextGrid"
tg = textgrid.openTextgrid(inputFN,False)

firstTier = tg.tierDict['Label-Track-1']

silenceTier = tg.tierDict['silences']

input("Intertier, Press Enter to continue...")
interTier = firstTier.intersection(silenceTier)
interTier = interTier.new(name="intersection")

tg.addTier(interTier)
tg.save("C:\Users\pyintersect\intersect03.TextGrid",'long_textgrid',True)
f = open("tier3.txt", "w")
for start, stop, label in interTier.entryList:
    print("%f, %f, %s" % (start, stop, label))
    f.write(str(start)+','+str(stop)+','+label+'\n')
f.close()
```

The file name should be addressed at line 4. The save destination should be addressed at line 16.

Usage of ‘intersect.py’

`python intersect.py` (Enter)

`Intertier, Press Enter to continue …` (press Enter again)

APPENDIX F

The results of the content validity index (CVI) on the three spontaneous speech tasks from three experts

The evaluated results of content validity index on the three spontaneous speech tasks for distinguishing persons with Alzheimer's disease, persons with Mild Cognitive Impairment and cognitively intact older adults

Evaluation criteria	**Expert no.**			**I-CVI**
	1	**2**	**3**	
1. Consistency with the basic theory and approaches				
1.1 Thai Picture Description Task (TPD)	4	4	3	1
1.2 Thai Story Recall Task (TSR)	4	4	3	1
1.3 Semi-structured Interview for Thai (SIT)	4	4	3	1
2. Appropriation of stimuli				
2.1 Thai Picture Description Task	3	3	3	1
2.2 Thai Story Recall Task	3	3	3	1
2.3 Semi-structured Interview for Thai	4	3	3	1
3. Administration and construction				
3.1 Thai Picture Description Task	3	4	3	1
3.2 Thai Story Recall Task	4	3	3	1
3.3 Semi-structured Interview for Thai	4	3	3	1
4. Reasonable scoring criteria				
4.1 Thai Picture Description Task	4	3	3	1
4.2 Thai Story Recall Task	4	3	3	1
5. Overall				
5.1 Thai Picture Description Task	4	3	3	1
5.2 Thai Story Recall Task	4	3	3	1
5.3 Semi-structured Interview for Thai	4	3	3	1
S-CVI				**1**

BIOGRAPHY

NAME	Natinee Na Chiangmai
DATE OF BIRTH	28 Apr 1988
PLACE OF BIRTH	Chiang Mai
PRESENT ADDRESS	RMCS
EDUCATION	MSc