2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Thammasat-NECTEC-Chula's Thai Language and Cognition Assessment (TLCA): The Thai Alzheimer's and Mild Cognitive Impairment Screening Test

A. Munthuli, S. Vongsurakrai, T. Anansiripinyo, V. Ellermann, K. Sroykhumpa, C. Onsuwan, P. Chutichetpong, S. Hemrungrojn, K. Kosawat, and C. Tantibundhit*

***Abstract*— Thammasat-NECTEC-Chula's Thai Language and Cognition Assessment (TLCA) is a cognitive paper-based test consisting of 21 tasks that cover 3 domains: memory, language, and other cognitive abilities. The TLCA follows some aspects of the existing tests (Thai Addenbrooke's Cognitive Examination-Revised (Thai-ACE-R) and the Thai Montreal Cognitive Assessment Test (Thai-MoCA)) and many parts were reconstructed to be more adapted to the Thai culture. Data obtained from the test will be able to precisely distinguish between patients with Mild Cognitive Impairment (MCI), Alzheimer's Disease (AD), and Normal healthy Controls (NC). The TLCA was tested on 90 participants (32 on the paper-based version and 58 on the computerized version) using a scoring procedure and speech features from verbal responses with machine learning classification. The scoring results showed significant difference between non-AD (NC + MCI) vs AD participants in 3 domains and could differentiate between NC and MCI, while machine classification could classify in three settings: NC vs non-NC (MCI + AD), AD vs non-AD and NC vs MCI vs AD. These promising results suggest that TLCA could be further verified and used as an efficient assessment in MCI and AD screening for Thais.**

***Clinical relevance*— The speech feature analysis of TLCA showed promising result for screening MCI and AD for Thais.**

I. INTRODUCTION

The world's population is transitioning into an aging society. Between 2000 and 2050, the world population that is above 60 years is predicted to double [1]. With this increase in the elderly population, a concerning hurdle must be solved: the increase in prevalence of dementia. Global prevalence of this dementia for adult ages above 60 is approximately $5-7\%$ while the prevalence is more than 50% for people aged 90 or above [2]. Alzheimer's disease (AD) is the most common type of dementia, with symptoms including memory loss, cognitive deterioration, and behavioral impairments. In addition, AD may include patients suffering Mild Cognitive Impairment (MCI) [3]. Although there is no known cure for AD, the progression of the disease can be delayed if given the appropriate treatment. As a result, an early and accurate diagnosis of AD is crucial for patients, as it enables them to start an early treatment before the symptoms progress towards critical stages [4].

Current clinical approach to the diagnosis of AD revolves around the use of brain-imaging analysis tests such as magnetic resonance imaging, computed tomography, and positron emission tomography scans [5]. However, these methods are complicated, expensive, and require specialized personnel, which make them inconvenient in low and middle-income countries. While neuropsychological tests such as Mini Mental State Examination (MMSE) [6], Addenbrooke's Cognitive Examination-III (ACE-III) [7], and Montreal Cognitive Assessment Test (MoCA) [8] can also be used, these examinations still require specific technicians and are extremely time consuming. Therefore, a diagnosis method that is inexpensive, accurate, and fast is still needed.

In recent years, there has been a shift towards the use of spoken language features as a marker of AD as it has been shown that AD affects the use of language in a variety of ways [9], such as the reduction in lexical content, fluency, semantic content, and syntactic complexity [10], [11]. This promising use of language features as a viable marker for AD combined with the improvement of powerful computational linguistic tools has resulted in new diagnostic methods such as the Automated Spontaneous Speech Analysis [12]. These techniques involve an extraction of features from speech samples for machine learning classification algorithm to distinguish between Normal healthy Controls (NC) and AD patients. A study done by Thomas et al. [13] employed a "common n-grams" approach to classify English spontaneous speech samples from ADs. The proposed method was able to distinguish between NCs and ADs with an accuracy of 94.5%, and between NC and MCI patients with an accuracy of 75.3%. Another study by Habash and Guinn [14] used English conversational speech samples from 31 ADs and 57 NCs and achieved the best accuracy of 79.5% by measuring pauses, repetition of words, and incomplete enumeration of words.

Linguistic and text-based features have also been explored by various studies based on the use of acoustic and prosodic features [8], [15], [16]. Meilan et al. [15] studied 30 ADs and 36 NCs reading a relatively long paragraph. Their results revealed that acoustic features, primarily the percentage of voiceless segments in patients' speech, was strongly related

*C. Tantibundhit (e-mail: `tchartur@engr.tu.ac.th`), A. Munthuli, V. Ellermann, and K. Sroykhumpa are with Center of Excellence in Intelligence Informatics, Speech, and Language Technology, and Service Innovation (CILS), Faculty of Engineering, S. Vongsurakrai and P. Chutichetpong are with CILS, and T. Anansiripinyo and C. Onsuwan, are with CILS and Faculty of Liberal Arts, Thammasat University. S. Hemrungrojn is with Cognitive Fitness Research Group, Faculty of Medicine, Chulalongkorn University. K. Kosawat is with National Electronics and Computer Technology Center, NSTDA, Thailand.

TABLE I: Percentage of different subtasks distributed across memory, language and other cognition domains of 3 existing screening tools: Thai ACE-R (translated from ACE-R [20]), Thai-MoCA [18], and TLCA.

	Task	Thai ACE-R	Thai MoCA	TLCA
Memory	Registration (REG)	1.56%	-	1.25%
	Recall (REC)	1.56%	-	1.25%
	Anterograde Memory (AM)	1.56%	4.54%	1.25%
	Retrograde Memory (RM)	6.25%	-	7.50%
	Recall2 (REC2)	1.56%	4.54%	1.25%
	Recognition (REC3)	1.56%	-	1.25%
	Sum	**14.05%**	**9.08%**	**13.75%**
Language	Abstraction (ABS)	-	4.54%	-
	Verbal Fluency (VF)	3.12%	4.54%	3.75%
	Language Comprehension (LC)	3.12%	-	2.50%
	Language Writing (LW)	1.56%	-	1.25%
	Language Repetition (LRP)	9.38%	9.10%	12.50%
	Language Naming (LN)	18.75%	13.64%	15.00%
	Language Comprehension2 (LC2)	6.25%	-	5.00%
	Language Reading (LRD)	7.82%	-	7.50%
	Verbal Description Abilities (VDA)	-	-	1.25%
	Verbal Description Abilities2 (VDA2)	-	-	1.25%
	Sum	**50.00%**	**31.82%**	**50.00%**
Cognition	Orientation (ORI)	15.63%	27.28%	12.50%
	Attention & Concentration (AC)	3.13%	18.19%	6.25%
	Visuospatial Abilities (VA)	4.69%	9.09%	3.75%
	Perceptual Abilities (PA)	6.25%	-	5.00%
	Perceptual Abilities2 (PA2)	6.25%	-	7.50%
	Executive Function (EF)	-	4.54%	1.25%
	Sum	**35.95%**	**59.10%**	**36.25%**

to AD. Weiner et al. [16] conducted a speech-based detection for German conversation and based the machine learning classification algorithm on acoustic features revolving around the statistical relationship between silence and transcription segments such as mean silence duration, silence rate, and word rate. The linear discriminant analysis classification system implemented was found to give an accuracy of 85.7%.

Despite these advancements in methods of diagnosing AD, there is still a huge research gap as these computational methods are not applicable to the Thai population due to the fact that their primary language is Thai − a language that is linguistically and semantically different, compared to English [17]. As a result, diagnosis of AD in Thailand still relies on conventional methods and standardized examinations such as the Thai MoCA (Thai-MoCA) [18] and the Thai ACE-III (Thai-ACE-III) [19]. This is extremely problematic in Thailand, especially in the rural regions, where AD patients far outnumber medical specialists and diagnostic equipment are lacking. Because of this, many Thais are left unaware about their condition, resulting in the inevitable progression of the disease.

This study aims to develop a new paper-based test namely TLCA that conforms to the Thai language and an algorithm that can automatically screen for AD and MCI using patient's verbal responses to TLCA. By automating the process, this study hopes to create a simple, fast, and accurate screening tool of AD and MCI that can be used anywhere especially in rural areas, where well-trained medical staff is not available or is very limited. This may solve the expensive nature of conventional examinations, which have been a major hurdle in Thailand.

II. Development of TLCA

In medical service, there are effective paper-based screening tools that can identify between AD, MCI, and healthy patients such as Addenbrooke's Cognitive Examination-Revised (ACE-R) [20] (the previous version of ACE-III [7]) and MoCA [21]. However, most of them are designed for the English-speaking population. Although there are many attempts to translate these tests to various languages, including Thai, the difference in culture and language aspects makes these tests (Thai-ACE-R, translated from ACE-R [20], and Thai-MoCA [18]) unsuitable for Thai patients. Therefore, we intend to develop a Thai paper-based test that adheres to the Thai language and culture.

TABLE II: Comparison of tasks in TLCA, Thai-ACE-R, and Thai-MoCA (Italic). Abbreviations refer to Table 1.

Task	TLCA	Thai-ACE-R / *Thai-MoCA*
ORI	Answer 10 questions (followed Thai-ACE-R).	Answer 10 questions about date, time, and place. *Answer 6 questions about date, time, and place.*
REG	Repeat and memorize 3 words (followed Thai-ACE -R).	Repeat and memorize 3 words. *N/A*
AC	- Take 7 away 5 times starting from 100 (followed Thai-ACE-R). - Spell the word "มะนาว" 'lemon' backward (followed Thai-ACE-R). - Repeat digits in the forward order (followed Thai-MoCA). - Repeat digits in the backward order (followed Thai-MoCA). - Tap when heard "1" (followed Thai-MoCA).	- Take 7 away 5 times starting from 100. - Spell the word "มะนาว" 'lemon' backward. *- Take 7 away 5 times starting from 100.* *- Repeat digits in the forward order.* *- Repeat digits in the backward order.* *- Tap when heard "1".*
REC	Recall 3 words from REG task (followed Thai-ACE -R).	Recall 3 words from REG task. *N/A*
AM	Repeat and memorize a name and address (followed Thai-ACE-R).	Repeat and memorize a name and address. *Repeat and memorize 5 words.*
RM	Answer 6 questions, 3 items (followed Thai-ACE-R) and 3 items are about famous places and flower of Thailand.	Answer 4 questions about famous person, capital, and the important days of Thailand. *N/A*
VF	Name maximum number of words in one minute: - the letter "ส" - the letter "ก" (followed Thai-MoCA) - Animals (followed Thai-ACE-R)	Name maximum number of words in one minute: - the letter "อ" - Animals *Name maximum number of words begin with the letter "ก" in one minute.*
LC	Follow 2 instructions (followed Thai-ACE-R).	Follow 2 instructions. *N/A*
LW	Write a sentence (followed Thai-ACE-R).	Write a sentence. *N/A*
LRP	Repeat 8 words and 2 phrases.	Repeat 4 words and 2 phrases. *Repeat 2 sentences.*
LN	Name 12 pictures, 9 of them are followed Thai-ACE-R and 3 pictures (adapted from Thai-ACE-R).	Name 12 pictures. *Name 3 pictures.*
LC2	Answer 4 questions about 12 pictures from LN task, 3 items are followed Thai-ACE-R.	Answer 4 questions about 12 pictures from LN task. *N/A*
LRD	Read 6 words.	Read 5 words. *N/A*
VA	- Copy 2 drawings (followed Thai-ACE-R). - Draw a clock (followed Thai-ACE-R).	- Copy 2 drawings: overlapping pentagons and cube. - Draw a clock (ten past five). *- Copy 1 drawing: a cube.* *- Draw a clock (ten past eleven).*
PA	Identify the number of dots in each square: 4 items (followed Thai-ACE-R).	Identify the number of dots in each square: 4 items. *N/A*
PA2	Identify 6 Thai letters: 4 consonants (followed Thai -ACE-R) 1 vowel, and 1 tone.	Identify 4 Thai letters. *N/A*
REC2	Recall a name and address from AM task (followed Thai-ACE-R).	Recall a name and address from AM task. *Recall 5 words.*
REC3	Recognize name and address from AM task (followed Thai-ACE-R).	Recognize name and address from AM task. *N/A*
EF	- Draw a line, going from a number to dot(s) in ascending order (followed Thai-MoCA-B).	N/A *Draw a line, going from a number to Thai letter in ascending order.*
VDA/ VDA2	Picture description task: 2 items	N/A *N/A*

A. Thai Standard Screening Tools

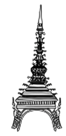
Thai-ACE-R consists of 18 tasks, while Thai-MoCA contains 10 tasks. The tasks in each test can be grouped into 3 domains: memory, language, and other cognitive abilities as shown in Table 1. By comparing the proportion of domains between each screening tool, it shows that Thai-ACE-R contains 14.05% memory, 50.00% language , and 35.95% cognition while Thai-MoCA consists of 9.08% memory, 31.82% language, and 59.10% cognition. Moreover, the proportion of the newly created test, TLCA, is also shown in Table 1, including 13.75% memory, 50.00% language, and 36.25% cognition. Apparently, Thai-ACE-R and TLCA gives higher priority to the language domain. However, the language tasks in TLCA are of more variety.

B. TLCA

TLCA has been adapted and developed from Thai-ACE-R and Thai-MoCA with newly created items. The proportion of

(a) Owl (LN)

(b) Glazed water jar with dragon patterns (LN)



(c) Headdress (LN)

(d) Thai consonant letter (PA2)

(e) Thai vowel letter (PA2)

(f) Thai tone marker (PA2)

Fig. 1: Samples from subtests of TLCA.

the newly created tasks in TLCA is 15.40%, while 15.87% is adapted from Thai-ACE-R, and some parts of Thai-MoCA, while 68.73% followed Thai-ACE-R. Some parts of Thai-MoCA and Thai-MoCA-B also included [22]. The differences in detail of each task in TLCA compared to Thai-ACE-R and Thai-MoCA are shown in Table 2. Note that the last column in Table 2 combines the information about Thai-ACE-R and Thai-MoCA, where the details of Thai-MoCA are presented in italics.

Most tasks created and reorganized are related to the areas of language and culture. As we prioritize language features for screening MCI and AD, we separately introduced another linguistic test by creating a picture description task [23] in order to get data in the form of spontaneous speech. This task consists of two scenario pictures that are appropriate for the Thai context. Participants were asked to describe the information as much as they could provide.

Another task which was redesigned is Language Naming (LN) task. In this task, participants were asked to name objects presented in the picture. The task consists of 12 pictures: 6 pictures are from ACE-R; 3 pictures were redesigned based on ACE-R; and 3 pictures were newly created to conform to the Thai culture. A picture of an owl (Fig. 1a)) "นกฮูก" was created to replace the picture of a penguin in ACE-R, since penguin is not commonly seen in a tropical country like Thailand. Moreover, an orchid and a crown in ACE-R were substituted by a glazed water jar with dragon patterns "โอ่ง" (Fig. 1b)) and a Thai theatrical crown "ชฎา" (Fig. 1c)); these objects are more familiar for the Thais.

In addition, in the Perceptual Abilities (PA) task, participants were asked to name Thai alphabets, presented in an incomplete form. This task contains 6 items: 4 were from Thai-ACE-R and 2 were newly developed to make the test cover 3 categories of the Thai alphabets: consonant, vowel, and tone. Figure 1 shows some of these items: d) is a consonant symbol "จ" named "จอ จาน" [tɕɔː tɕaːn] that was redesigned from Thai-ACE-R; e) is a vowel symbol "ใ" named "สระใอไม้ม้วน" [sàʔ ràʔ ʔaj máːj múan]; and f) is a tone marker "้" named "ไม้โท" [máːj tʰoː].

Moreover, in the Language Reading (LRD) task, participants were asked to read aloud words. This task is made up of 6 words that were newly selected' i.e., "ฤกษ์" [rɤ̂ːk] 'auspicious time', "แสตมป์" [sàʔ tæm] 'stamp', "สลึง" [sàʔ lɯ̌ŋ] 'quarter of a baht', "ไทร" [saj] 'banyan tree', "ครรภ์" [kʰan] 'womb', and "หลา" [lǎː] 'yard'.

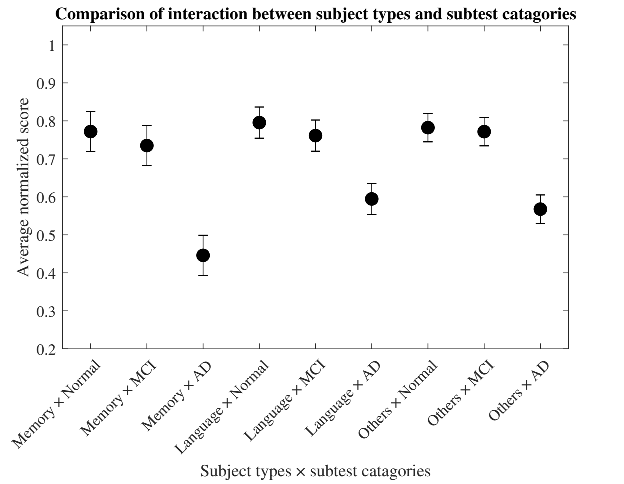


Fig. 2: Comparison from two-factor balanced ANOVA with factors TYPE and DOM $[F(4,89) = 2.42, p = 0.054]$.

III. Data Collection

A. Participants

Ninety people (32 from the paper-based and 58 from the computerized version) who are natively Thai participated in this study. They were classified into 3 groups: NCs (30) divided into 23 females and 7 males, with an average age of 65.80 ± 4.12 (mean ± SD), MCIs, (30) divided into 17 females and 13 males, with an average age of 71.87 ± 6.34 (mean ± SD), and ADs, (30) divided into 17 females and 13 males, with an average age of 71.17 ± 7.25 (mean ± SD). All of the participants were diagnosed by a team of psychiatrists from the Department of Psychiatry, Faculty of Medicine, Chulalongkorn University. More importantly, this study was conducted under the institutional review board (IRB) review and approval from Chulalongkorn Memorial Hospital no. 206/59.

B. Experimental Procedures

Participants were asked to perform paper-based or computer-based TLCA (chose one of the formats) their speaking, writing, and practicing abilities. Participants might assign to perform either the paper-based or computerized version but not both. The test consists of 21 tasks divided into 16 speech-related tasks and 5 non-speech related tasks. The answers were recorded in 2 forms, sound recording and manual scoring by researchers. The experiment was held individually in a quiet room at Cognitive Fitness Center, Chulalongkorn Memorial Hospital to control the participants' sound quality. The test took 30 to 60 minutes depending on individual performance.

C. *Data Statistic of Task Scores*

A two-factor balanced Analysis of Variance (ANOVA) was conducted with scoring from 90 participants to examine the effects of the factor "subject types" between NCs, MCIs, and ADs (called TYPE) and the factors "domains" between memory, language, and other cognition (called DOM). The results from the individual factor of TYPE revealed that there is a significant difference $[F(2, 89) = 6.14, p < 0.05]$, where non-ADs are different from ADs. The highest performance came from the group of NCs followed by MCIs, then ADs by the factor DOM (as illustrated in Fig. 2). The results from the interaction between the effects of TYPE $\times$ DOM also show significant difference $[F(4, 89) = 2.42, p = 0.054]$ conveying that the trends of classification between NCs vs MCIs vs ADs for each domain are not different. Figure 2 illustrates that for the memory domain, there is the same trend in each group, where non-ADs vs ADs are varied while there is no difference within the group of non-ADs.

IV. Speech Data for Machine Learning

A. *Speech Materials*

Spontaneous speech data from 90 participants were collected. Although the sound environment was controlled, the data still had some background noise and long periods of silence that reflects the real situation. The recorded materials were separated into 21 segments according to the number of tasks in TLCA. The participants' verbal responses elicited from the relevant tasks result in a total of 1,440 sound files (90 participants $\times$ 16 tasks). These 16 speech-related tasks include REG, REC, AM, RM, REC2, REC3, VF, LRP, LN, LRD, VDA, VDA2, ORI, AC, PA, and PA2.

B. *Feature Extraction and Classification*

Verbal responses were selected to perform an automatic feature extraction based on speech signal using the Emobase feature set [24] from openSMILE [25]. These sets consisted of 988 features derived from 19 functionals of 26 low-level-descriptors (e.g., intensity, loudness, Mel-Frequency Cepstral Coefficients (MFCCs), Zero Crossing Rate (ZCR), etc.) with their delta regression coefficients. The data from feature extraction result in 988 features of 1,440 files in Attribute-Relation File Format (ARFF). These features from the 16 tasks were then combined into a single ARFF file for each participant (90 in total). The attributes from feature extraction were then analyzed with machine learning using the 5-fold cross-validation technique with hold-out data. The data obtained were randomly split into 75 participants for cross-validation (15 participants for each fold with an equal ratio of NC, MCI, and AD) and 15 hold-out participants with equal ratio of NC, MCI, and AD for testing.

To analyze speech features for machine learning classification, the feature selection process is crucial for obtaining accurate results. Feature selection was performed using correlation-based feature selection along with the best-first search method by searching subspace using greedy hill climbing augmented with a backtracking facility [26] to get the more reliable features for each data set. The experiments were performed to find the most effective set of features for AD screening by dividing experiment into 3 settings: 1) NC vs MCI vs AD, 2) NC vs non-NC (MCI and AD), and 3) non-AD vs AD. Then, the classification process was performed using Scikit-learn, machine learning in Python, with 6 classifiers: Linear Discriminant Analysis (LDA), Support Vector Classification (SVC), Logistic Regression (LR), Adaptive Boosting (AdaBoost), Random Forest (RF), and Multi-Layer Perceptron (MLP) [27].

TABLE III: Results from the best validation of each settings.

	NC vs non-NC	AD vs non-AD	NC vs MCI vs AD NC	MCI	AD
Classifier	MLP (fold 4)	RF (fold 2)		LR (fold 1)	
Best AUC	0.94	0.94	1.00	0.94	1.00
Sensitivity (%)	100.00	100.00	100.00	100.00	100.00
Specificity (%)	83.33	83.33	100.00	100.00	100.00
Accuracy (%)	93.33	93.33	100.00	100.00	100.00

V. Experimental Results

In order to determine the performance difference of NC, MCI or AD participants, transcription data from 135 participants were evaluated from both paper-based and computerized versions. Statistical results suggest that the test can be used to classify patients accurately. Every domain has the same trend for classifying between NC vs MCI vs AD $[F(4, 89) = 2.42, p = 0.054]$, but NCs and MCIs were not classified. However, there is a significant difference between non-ADs vs ADs in the scoring results $[F(2, 89) = 6.14, p < 0.05]$. The results from the scoring were conformed with the assumption that NCs generally have the highest performance, followed by MCIs and ADs, respectively. However, the results are unable to differentiate between NC and MCI patients.

Participants' verbal responses were analyzed with machine learning classification models using the 5-fold cross-validation method. Table 3 shows the selected model and number of validation folds that yields the highest Area Under the Receiver operating characteristic Curve (AUC) among 30 models (number of folds $\times$ models $= 5 \times 6 = 30$). The results were 1) NC vs non-NC setting of MLP classifier (fold 4) with AUC of 0.94, 2) AD vs non-AD of RF classifier (fold 2) with AUC of 0.94, and 3) NC vs MCI vs AD of LR classifier (fold 1) with AUC of class NC of 1.00, MCI of 0.94, and AD of 1.00.

The sensitivity, specificity, and accuracy of the best-performing classifier from 6 classifiers stated earlier are compared between 3 settings of feature selection. As shown in Table 3, the classification of NC vs non-NC model and AD vs non-AD model have similar results and the misclassification from both NC vs non-NC and AD vs non-AD participants came from MCI. The NC vs MCI vs AD model was perfectly accurate for all participants in the three classes. Unless the results from speech features are different from ANOVA results in scoring method, the misclassification from both NC vs non-NC model and AD vs non-AD show agreement that MCI tends to be misclassified from NC rather than AD participants.

.

VI. Discussions and Future Work

The new memory, language, and cognition assessment or TLCA that conforms to the Thai language was designed to screen Thai patients with AD and MCI. TLCA consists of 21 tasks, focusing on 3 domains, which are memory, language, and other cognition abilities. This test can be used to elicit the differences of language usage between Thai AD patients and non-AD patients.

Verbal responses from the 16 tasks that elicit various speech signal features revealed that MCIs can be classified through speech feature analysis. The three test settings of NC vs non-NC, AD vs non-AD, and NC vs MCI vs AD were performed and showed promising results. Although most of the misclassifications came from MCI (misclassification between NC vs MCI and MCI vs AD), the results showed that speech features from the verbal responses can be used to distinguish between NC vs MCI vs AD. A combination of scoring method and verbal response (with speech signal features) can be efficiently used for screening NCs, MCIs, and ADs. It is noteworthy that in every group of NC, MCI, and AD, the number of hold-out participants was five, as the number of participants is limited.

In the future, we intend to increase the number of participants and also take additional characteristics into account such as age, gender, educational level, and occupation. Furthermore, the reliability and validity of the TLCA and other standard screening tests will be evaluated, and more acoustic features will be explored, which might be relevant for classification of MCI patients. Finally, we will apply the automatic speech recognition system to convert speech into text to develop an automatic scoring system in the next phase.

Acknowledgment

The authors would like to thank Mr. Kantapon Pornprasertsakul and Miss Sutanya Phuechpanpaisal for the data collection of speech samples.

References

[1] World Health Organization (WHO), "Facts about ageing," 2018. [online]. Available: http://www.who.int/ageing/about/facts/en/. Accessed: 2020-03-25.

[2] M. Prince, R. Bryce, E. Albanese, A. Wimo, W. Ribeiro, and C. P. Ferri, "The global prevalence of dementia: A systematic review and metaanalysis," *Alzheimers. Dement.*, vol. 9, no. 1, pp. 63–75, 2013.

[3] M. L. B. Pulido, J. B. A. Hernández, M. Á. F. Ballester, C. M. T. González, J. Mekyska, and Z. Smékal, "Alzheimer's disease and automatic speech analysis: A review," *Expert Syst. Appl.*, p. 113213, 2020.

[4] Z. S. Khachaturian, "Diagnosis of Alzheimer's disease," *Arch. Neurol.*, vol. 42, no. 11, pp. 1097–1105, 1985.

[5] Alzheimer's Association, "Early diagnosis," Alzheimer's association 2018. [online]. Available: https://www.alz.org/research/science/earlier_alzheimers_diagnosis.asp. Accessed: 2020-04-02.

[6] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "Mini-mental state: A practical method for grading the cognitive state of patients for the clinician," *J. Psychiatr. Res.*, vol. 12, no. 3, pp. 189–198, 1975.

[7] J. A. Matias-Guiu, A. Cortés-Martínez, M. Valles-Salgado, T. Rognoni, M. Fernández-Matarrubia, T. Moreno-Ramos, and J. Matías-Guiu, "Addenbrooke's cognitive examination III: Diagnostic utility for mild cognitive impairment and dementia and correlation with standardized neuropsychological tests," *Int. Psychogeriatr.*, vol. 29, no. 1, pp. 105–113, 2017.

[8] K. López-de Ipiña, J. B. Alonso, C. M. Travieso, J. Solé-Casals, H. Egiraun, M. Faundez-Zanuy, A. Ezeiza, N. Barroso, M. Ecay-Torres, P. Martinez-Lage *et al.*, "On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis," *Sensors*, vol. 13, no. 5, pp. 6730–6745, 2013.

[9] J. Appell, A. Kertesz, and M. Fisman, "A study of language functioning in Alzheimer patients," *Brain Lang.*, vol. 17, no. 1, pp. 73–91, 1982.

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

[11] S. Ahmed, C. A. de Jager, A. Haigh, and P. Garrard, "Semantic processing in connected speech at a uniformly early stage of autopsy-confirmed Alzheimer's disease." *Neuropsychology*, vol. 27, no. 1, p. 79, 2013.

[12] K. C. Fraser, J. A. Meltzer, and F. Rudzicz, "Linguistic features identify Alzheimer's disease in narrative speech," *J. Alzheimers Dis.*, vol. 49, no. 2, pp. 407–422, 2016.

[13] C. Thomas, V. Keselj, N. Cercone, K. Rockwood, and E. Asp, "Automatic detection and rating of dementia of Alzheimer type through lexical analysis of spontaneous speech," in *IEEE Int. Conf. Mech. Autom.*, vol. 3, 2005, pp. 1569–1574.

[14] C. I. Guinn and A. Habash, "Language analysis of speakers with dementia of the Alzheimer's type," in *AAAI Fall Symposium: AI Gerontech.*, 2012, pp. 8–13.

[15] J. J. Meilán, F. Martínez-Sánchez, J. Carro, J. A. Sánchez, and E. Pérez, "Acoustic markers associated with impairment in language processing in Alzheimer's disease," *Span J. Psychol*, vol. 15, no. 2, pp. 487–494, 2012.

[16] J. Weiner, C. Herff, and T. Schultz, "Speech-based detection of Alzheimer's disease in conversational German." in *Proc. 17th Annu. Conf. Int. Speech Commun. Assoc. (Interspeech'16)*, 2016, pp. 1938–1942.

[17] B. Comrie, *The world's major languages*. Abingdon, UK: Routledge, 2009.

[18] S. Tangwongchai, M. Phanasathit, T. Charernboon, L. Akkayagorn, S. Hemrungrojn, and K. Phanthumchinda, "The validity of Thai version of the Montreal Cognitive Assessment (MoCA-T)," *Dement. Neuropsychol.*, vol. 3, no. 2, pp. 136–178, 2009.

[19] T. Charernboon, K. Jaisin, and T. Lerthattasilp, "The Thai version of the Addenbrooke's cognitive examination III," *Psychiat. Invest.*, vol. 13, no. 5, p. 571, 2016.

[20] E. Mioshi, K. Dawson, J. Mitchell, R. Arnold, and J. R. Hodges, "The Addenbrooke's cognitive examination revised (ACE-R): A brief cognitive test battery for dementia screening," *Int. J. Geriatr. Psych.*, vol. 21, no. 11, pp. 1078–1085, 2006.

[21] Z. S. Nasreddine, N. A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings, and H. Chertkow, "The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment," *J. Am. Geriatr. Soc.*, vol. 53, no. 4, pp. 695–699, 2005.

[22] P. Julayanont, S. Tangwongchai, S. Hemrungrojn, C. Tunvirachaisakul, K. Phanthumchinda, J. Hongsawat, P. Suwichanarakul, S. Thanasirorat, and Z. S. Nasreddine, "The Montreal Cognitive Assessment—Basic: A screening tool for mild cognitive impairment in illiterate and low-educated elderly adults," *J. Am. Geriatr. Soc.*, vol. 63, no. 12, pp. 2550–2554, 2015.

[23] K. Pornprasertsakul and S. Phuechpanpaisal, "Alzheimer's and mild cognitive impairment screening test using speech signal features," *B.E. Thesis, Faculty of Engineering, Thammasat University, Pathum Thani, TH*, 2016.

[24] C. Parlak, B. Diri, and F. Gürgen, "A cross-corpus experiment in speech emotion recognition." in *Proc. 14th Int. Workshop Speech Lang. Audio Multimedia (SLAM'14)*, 2014, pp. 58–61.

[25] F. Eyben, M. Wöllmer, and B. Schuller, "Opensmile: the munich versatile and fast open-source audio feature extractor," in *Proc. 18th ACM Int. Conf. Multimedia*, 2010, pp. 1459–1462.

[26] K. Selvakuberan, M. Indradevi, and R. Rajaram, "Combined feature selection and classification–A novel approach for the categorization of web pages," *J. Inf. Comput. Sci.*, vol. 3, no. 2, pp. 83–89, 2008.

[27] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.