

The Effect of Final Consonants in the Cognitive Process of Thai Language

Phenpimon Wilairatana ^{*}, Koichi Mizutani [‡],
Pichit Ngamjarussrichai [‡], Chanongkorn Kuntonbutr [†],
Konosu Tsutomu ^{*}

Abstract

Written nature of Thai language does not have spaces in a sentence like English and does not have Kanji characters as in Japanese. Words will be written continuously until the end of the sentence the judgement for separate words depends on human decisions. Therefore, cognitive processing is very important to separate words in the Thai language. This research studied the cognitive processing and influence of final consonants in the Thai language by using word identification tasks to verify the difference in time spent answering and correcting data, as there is no direct measurement. Therefore, the time spent on selected word and correctness ratio was the most important variable in the cognitive processing test. The differences of average duration for selected words were analyzed by ANOVA and t-test that is statistical analysis. The results showed that final consonants work as the space in English is used for separating words and sentences and can make the Thai language easier for detect words in Thai language and also reduce the confusion when separating words in the long run on sentences. It works as the accelerator of competency in cognition of Thai language and helps learners to receive more precise and faster information.

Keywords: Thai language, Final Consonants, Information Processing, Cognitive Processing

1 Introduction

In the era of globalization, technology plays as an important role in the communication and transfer of information with no limits on time or distance.

The user interface is an important factor in the development of up-to-date technology, representing a medium whereby connections between devices and users occur focusing on the user's experience and interaction by reducing high personal skills and allowing proficiency after a short period of training that makes the user's interaction as simple and efficient as possible, in terms of accomplishing user goals.

^{*} Chiba Institute of Technology, Chiba, Japan

[‡] Thai-Nichi Institute of Technology, Bangkok, Thailand

[†] Rajamangala University of Technology, Pathumtani, Thailand

Moreover, the development in user interface creates High value add in products. Not only avoiding the cost competition, but also creating new opportunities in expanding the markets. For the current situation in Thailand, user interface design in Thai electronic industry relies on foreign investor's know-how, whether it be Japan or USA. There is no own interface design on electronic products, whether it be command buttons, software and so on. Thai users need to rely on translated data which is not designed for Thai users and cannot make the most of them. This reflects the lack of development interface design in electronic industry of Thailand. Therefore, researchers realize the problem of Interface designs which are not made for Thai users and intend to create Interface designs which specially work for Thai users. The previous study pointed out that to create a proper interface design, it is necessary to start from studying the cognitive process of human. Moreover, user interfaces are created to strengthen the relationship between human language and computers through Natural Language Processing (NLP). NLP is used to analyze text, allowing machines to understand how humans speak [1]. The result is that it is advantageous to study NLP by developing the cognitive process. Cognitive processing is a method to explore how human beings perceive the visual sensation and cognition which occurs in sequence[2].

Each language in the world has its own unique characteristics and ways of study which cause differences in cognition. For example, English has spaces, Japanese has Kanji which have their own meaning, while Thai has no spaces between words. Due to the special characteristics of these 3 languages, the speed of visual cognition is likewise different. The differences in the capacity to receive information in the English, Japanese, and German languages. The result indicated that the capacity to receive data amounted to 50 bits per fixation [3]. On the contrary, Thai language reading in the same way and the result turned out to be that the Thai language can receive data amounting to 88 bits per fixation, which is higher than the English, Japanese, and German languages [4]. The Thai language is one of the tonal language group with respect to speech recognition. According to Thai pronunciation standard has a syllable which is represented in the form of $C_iV(C_f)(T)$ when C_i denotes an initial consonant, V denotes a vowel, C_f denotes a final consonant, and the last T denotes a tonal marker [5]. There are 44 consonants and all of them can be an initial consonant; however, not all Thai consonants can be a final consonant [6]. Moreover, only 21 phonemes (k, kh, ng, c, ch, s, j, d, t, th, n, b, p, ph, f, m, r, l, w, h, and z) are represented as initial consonants and 8 phonemes (k^{\wedge} , ng^{\wedge} , t^{\wedge} , j^{\wedge} , n^{\wedge} , p^{\wedge} , m^{\wedge} , and n^{\wedge}) are represented as final consonants as shown in Table 1. Furthermore, the consonant letter and pronunciation relationship will change according to very complex set of Thai language phonemic rules, or the position in syllable form as shown in Table 1.

Consonant	Phoneme		Consonant	Phoneme	
	Initial (Ci)	Final (Cf)		Initial (Ci)	Final (Cf)
ก	k	k^{\wedge}	บ	b	p^{\wedge}
ข, กข, ข	kh	k^{\wedge}	ป	p	p^{\wedge}
ง	ng	ng^{\wedge}	ฝ, ฟ, ฟ	ph	p^{\wedge}
จ	c	t^{\wedge}	ฟ, ฟ	f	p^{\wedge}
ฉ, ฉ, ฉ	ch	t^{\wedge}	ม	m	m^{\wedge}
ซ, ซ, ซ, ส	s	t^{\wedge}	ร	r	n^{\wedge}
ฉ, ฉ	j	j^{\wedge}	ล, ล	l	n^{\wedge}
ด, ด	d	t^{\wedge}	ว	w	w^{\wedge}
ต, ต	t	t^{\wedge}	ห, ห	h	-
ท, ท, ท, ท, ท, ท	th	t^{\wedge}	อ	z	-
น, น	n	n^{\wedge}	Foreign lang.	br, bl, fr, fl, dr	$f^{\wedge}, s^{\wedge}, ch^{\wedge}, l^{\wedge}$

For example, the character “ส” in สม /soʔm/ is pronounced as /s/ and in รส /roʔt/ is pronounced as /t/. The problem for studying the Thai language is divided into 2 main points. First, the complication of separating the sentences from any paragraph and segmenting the words from a sentence. Since the Thai language is written continuously without spacing, the judgement for separate words depends on human decisions. Therefore, cognitive processing is very important to separate words in the Thai language. This process of making decisions helps readers to separate the words in the correct way, but if words are separated in the wrong position then it leads to a change in meaning. For example, “อาหารอร่อยหมดทุกอย่าง” which is pronounced as /aa-/haʔn-/a-/ /roʔi-/moʔt-/tuʔk-/yaʔng/. When a native speaker reads this sentence, it can be interpreted into 2 meanings: อาหาร / อร่อยหมดทุกอย่าง” means *all dishes are so delicious*. while

“อาหารอร่อย / หมดทุกอย่าง” means *all the delicious dishes are finished*. As a consequence, studying the differences in cognition between words with and without final consonant can reduce the mistakes in word separation. The second problem is the phonetic transcription. It is necessary to remember the rules of pronunciation in order to pronounce the Thai alphabet as shown in Table 1. To be able to pronounce correctly, a dictation test is applied in the learning process for the Thai language. It helps learners to remember consonants accurately, which are the components of each word. When compared with the Japanese language, transition from the alphabet to pronunciation is not as difficult because one character stands for one phonetic pronunciation. Moreover, Kanji in Chinese or character images which have their own meaning can be recognized immediately after you read it. Therefore, in order to create the correct memory data, the dictation test is deemed to be a significant process in learning the Thai language. According to the Two-Process Theory of Memory, the human memory process will start when the brain has acknowledged the new information; the brain will not keep it immediately, but it will transform the information to be in the form of codes [7]. For example, when students hear a new lexical item, the brain does not memorize the sound, but keeps it as code before transferring to the STM (short-term memory), but the items that are stored in the short-term memory must be reviewed and refreshed continuously otherwise the information will fade away and be lost. On the other hand, if students use this word frequently in their everyday lives, the new vocabulary becomes stored in LTM (long-term memory) and tends to be permanent [8]. Thus, to study the Thai language, memory also becomes an important factor in the cognitive information process.

2 Objective

To study the effect of final consonants in the Thai language by using word identification tasks to verify the difference in cognition between words with and without final consonants. The time spent to answer each question and the correctness ratio are important variables in the cognitive processing test in the Thai language.

The most significant hypothesis of this research was divided into 3 topics:

- I. To study the influence of final consonants in affecting cognition.
The assumption is that words with final consonants will be recognized faster and better corrected than words without final consonants because the final consonants work as a stimulator to improve the speed of cognition.
- II. To study the influence of word difficulty level in affecting cognition.
The assumption is that word difficulty at each level will have different results in both timing and accuracy in answering questions.
- III. To study the influence of kind of word which affect cognition.
The assumption is that four kinds of words have different results.

3 Methodology

This research studied cognitive processing and the influence of final consonants in the Thai language through word identification tasks to verify the difference in time spent answering and correcting data because as there is no direct measurement of this. Therefore, the time spent on selected words and the correctness ratio were the most important variables in the cognitive processing test. The test will be undertaken via websites on a computer system to enable accurate time recording for the duration spent answering questions, and for greater data collection convenience from the sample group. The system will record the amount of time spent on each question, starting with when a question appears and ending when the participants select an answer. After that, the system will save the data and summarize all 48 questions to verify the cognition differences in each condition.

3.1 Test Conditions

The conditions used for testing was divided into 3 patterns, including:

- (1) Types of word options, or word choices divided by words with and without consonants. The standard for words without consonants in the syllable was represented in the form of $C_iV(T)$. For example, the word “ต” means “eye”, and is pronounced /ta:/, with the syllable consisting of C_i as /t/ and V is /-a:/. In contrast, words with final consonants were represented in the form of $C_iVC_f(T)$. For example, the word “ตก” means “to dry in the sun” and is pronounced /taːk/ with the syllable consisting of C_i as /t/, while V is /-a:/ and C_f is /k/ [9].
Moreover, this study examined the competency of the Thai language cognition test by using an eye tracking camera to observe the reader’s eyes while they read, and found that their eyes focused on the final consonants [10]. Therefore, the final consonant is the important factor in Thai language cognition. In addition, the ability of final consonants by using image identification tasks on mobile applications by recording the speed and accuracy of responses to the words appearing on the screen. The test results indicate that words with final consonants had a higher ratio of response than words without final consonants, even though the average correctness ratio was quite similar. Hence, final consonants work as stimulators in Thai language cognition.
- (2) The difficulty level of the word is divided by the level of achieved education in Thai language. Nowadays, media in everyday life does not limit the difficulty because audiences themselves have the right to choose what they want to see. Thereby, this research categorizes the word difficulty into 3 levels of Thai textbooks and media, which were then divided by the age of the learner.
 - A) **Easy level** was tested by using the words that appeared in primary school textbooks and novels recommended for children between 6-12 years old.
 - B) **Middle level** was tested by words that appeared in normal magazines or novels that are recommended for people aged between 13-20 years old.
 - C) **Hard level** was tested by words that appeared in academic writing and journals recommended for people who study at university or in further education.
- (3) The different kinds of words are divided into 4 kinds: Uncontrolled, Homophone, Homograph, and Semantic. The difficulty of the Thai language comes from the characteristics of words which can cause ambiguity, whether it be Homophones (same pronunciation but different meaning), Homographs (same spelling but different pronunciation and meaning), or Semantic (similar meaning that can lead to misunderstanding). All of these words are specified in this experiment to imitate reading Thai in real situations as closely as possible.

3.2 Determining Sample Size

The simplified formula [11] (1) was applied to determine the sample size in this study, where n is the sample size, N is the population size, and e is the level of precision.

$$n = \frac{N}{1+(e)^2} \quad (1)$$

$$n = \frac{134}{1 + (0.05)^2}$$

$$n = 100$$

Thus, 100 responses were used as Thai native speakers. All of those 100 responses were analyzed after they were verified to be valid and complete without any unusual data or multiple responses. The number of valid responses were conformed to a finite population sampling formula, which gave a 95% confidence level and a 5% precision level.

Before beginning the experiment, the purpose of the test and instructions were explained to all the participants so that they all had the same level of understanding. Participants were then asked to do the test in a room isolated from any disturbance. The test was undertaken continuously to evaluate variables such as time and the accuracy in answering the questions. Each participant had to answer 48 questions and then all the data was statistically analyzed.

3.3 Experiment Procedure

- (1) Participants were required to follow the test procedure at <http://www.touchthailanguage.com>.
- (2) The first screen shows a registration screen which explains the purpose and procedure of the experiment to make sure the participants and researchers share the same understanding. The participants must read and endeavor to understand the information on the registration screen.
- (3) Participants are required to fill in their name to register for the experiment and then press the start button.
- (4) The experiment screen then appears meaning that the test had started.
- (5) When the experiment screen appears, a word target pops up in the center of the screen. Participants were required to choose only one answer out of 4 option words or word choices, which is related to the word target in the center, as shown in Figure 1. This should be completed as quickly as possible by tapping the selected words. The measurement of cognition starts after the experiment screen appears and ends after the participant chooses the answers, with the experiment ending when the participant completes all 48 questions.

เล่น play		กิน eat
	อาหาร food	
นอน sleep		ดื่ม drink

Figure 1: Example of word choices in the experiment screen.

- (6) The experiment continues until the summary score screen appears, which means the experiment has finished.

4 Results

The researchers applied statistical tooling in order to analyze the relationship between the independent and dependent variables. The statistical tooling included Graphical, two-way ANOVA, one-way ANOVA and t-tests for clarification and summarization of the experiment results. The results are shown as follows:

4.1 Analysis of the relationship between the independent variables and the correctness ratio in answering questions

The value of the mean shows that words with final consonants have a higher correctness ratio in answering questions than words without final consonants, as shown in Table 2. The relationship of the correctness ratio in answering questions which change following the difficulty of words depends on the difficulty levels as shown in Figure 2. Regarding to 48 questions, the final consonants were set up in 24 question in each part (without consonant question part and with consonant question part). The results indicate that consonant and word level affect the accuracy of answering questions. With a 95% confidence level, there was a significant effect for consonants, $F(1, 36) = 7.21$, $p = .01$, and a significant word level, $F(2, 36) = 4.00$, $p = .02$. It can be summarized that both variables directly affect the accuracy of question answers.

Table 2: The average correctness ratio in answering questions, split between words with and without final consonants.

Consonant		N	Mean
Correct item	without	24	74.42 %
	with	24	87.00 %

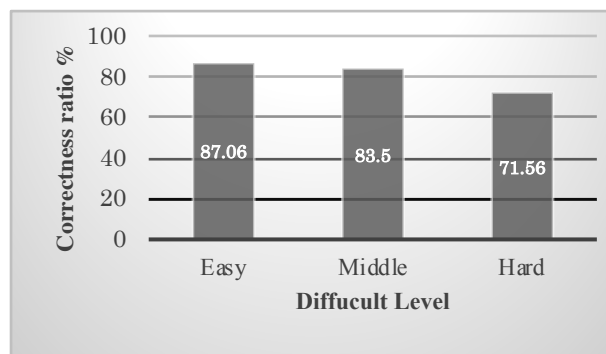


Figure 2: The relationship between the correctness ratio of answering questions which change with increasing word difficulty.

Moreover, differences in the group of participants between words with and without final consonants was analyzed by t-test. As the confidence level is 95%, there was a significant difference in the score for words without a final consonant ($M = 74.42$, $SD = 17.03$) and words with a final consonant ($M = 87.00$, $SD = 16.42$) condition; $t(46) = -2.61$, $p = .01$. Thus, it can be summarized that the factor for words with final a consonant has an influence on correctly answering questions in comparison to words without a final consonant.

The analysis by the statistical tooling indicated that words with final consonants work as an

accelerator of cognition competency in the Thai language. This means a final consonant presents itself as a ‘trigger’ in the Thai language and supports cognition. Furthermore, it functions like a ‘space’ in English, which is used to separate the words in a sentence in Thai, it helps readers detect words more easily and also reduces confusion when separating words in long run-on sentences.

4.2 Analysis of the timing ratio statistics in answering questions

The results indicate that consonant and word difficulty level affect the timing ratio statistics for answering questions. At a 95% confidence level, there was not a significant difference for consonants, $F(1, 36) = 0.89, p = .35$, on the contrary, word level was significantly different, $F(2, 36) = 1.49, p = .02$.

Table 3: Average of the timing ratio in answering questions split between words with and without final consonants.

	Consonant	N	Mean
Time	without	24	4.59 Sec.
	with	24	4.24 Sec.



Figure 3: The relationship of the timing ratio in answering questions which change with increasing word difficulty.

It can be summarized that words with final a consonant has an influence on the timing ratio in answering questions than words without final a consonant. Regardless of the analysis by statistical tooling, the time spent answering questions with independent variables showed no significant difference. In the case of considering the value of the mean in each group, it was found that words with final consonants also required less time for answering questions than words without final consonants, as shown in Table 3. Moreover, the timing ratio of answering questions varied according to the difficulty of the words, as shown in Figure 3. To summarize, word difficulty level is one of the factors which supports Thai language cognition.

4.3 Analysis of different kinds of words

Every language in the world is unique. However, one thing that is considered to be universal due to the limits of language are words. Words can be divided into three groups, including semantic words, homophones and homographs. Some assumptions indicated that differences in words affect the time taken to answer questions and the level of accuracy. Therefore, this research analyzed the relationship between the groups with different kinds of words by using statistical tooling, called one-way ANOVA, to explain and summarize the experiment. With a confidence level of 95%, the results showed as the following:

(1) The correctness ratio statistic for answering questions

There was a significant effect of the correctness ratio for answering among the 4-word groups at the $p < .05$ level for the 4 conditions [$F(3, 44) = 3.11, p = 0.03$]. It can be summarized that the average correctness ratio statistic to answer questions among the 4-word groups had different mean values: homographs and semantic had a lower than average correctness ratio in answering questions than homophone and uncontrolled, as shown in Figure 4.

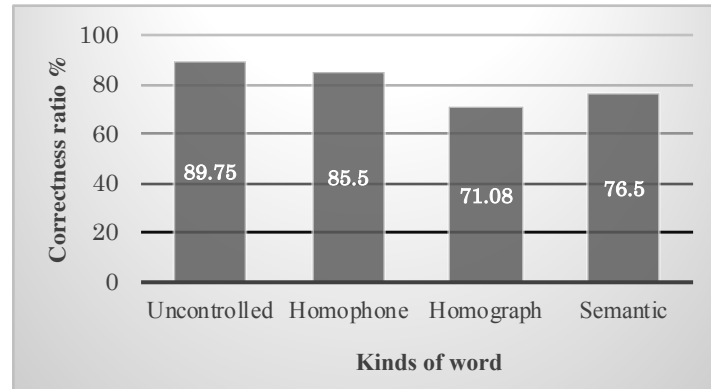


Figure 4: Average of the correctness ratio in answering questions depending on the kind of words.

(2) Timing ratio in answering questions

There was a significant effect of timing ratio for answering questions among the 4-word groups at the $p < .05$ level for the 4 conditions [$F(3, 44) = 3.28, p = 0.03$]. It can be summarized that homophones and uncontrolled had the lowest average timing ratio in answering questions than homograph and semantic, as shown in Figure 5.

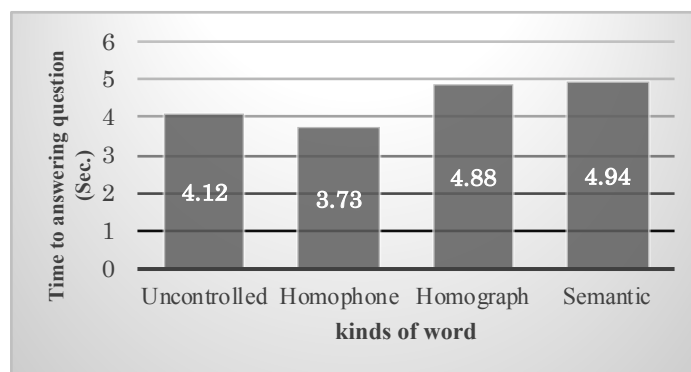


Figure 5: Average of the timing ratio in answering questions depending on the kind of words.

5 Conclusion

According to the analysis by statistic tooling, it can be summarized the very useful result using for improving and developing Thai interface design. The result can be divided into three main topics as below.

- (1) Words with final consonants showed considerable importance in cognition in comparison to words without final consonants. As final consonants in Thai work in a similar way to the 'space' in English to separate between words and make reading easier, the final consonant works as an accelerator of competency in cognition of the Thai language, whether in competency to receive precise information or learning those words faster than words without final consonants. Therefore, the final consonant is deemed to be one of the most important factors to aid in the design of a Thai interface for devices and tools with command buttons or signs in Thai.
- (2) The results of the time and accuracy analysis varied according to the word level difficulty. It shows that a good interface needs to consider the difficulty of the words when explaining information on the device. Poorly selected word choice beyond the user's competency will negatively affect the cognitive competency, in terms of both the time taken and accuracy in answering the questions.
- (3) Word choices for designing a Thai user interface should avoid semantic and homographs words which can cause more cognitive errors and increased the amount of time participants took to answer the questions. The design of a good interface should limit the potential for confusing the user because this can cause misunderstanding when they use the device.

To sum up, the application and improvement of these methods should be used to create an interface designed specifically for Thai users. This will help to increase value and reduce cost competition because the device will be easier and more convenient to use than an international interface design. Writing the instructions in Thai will also decrease the possibility of user error. Consequently, it can also affect consumer decision making and gives an advantage over imported goods with English language instructions. This will help to increase internal market sales and upgrade the Thai electronic industry's ability to compete with those of other countries.

6 Future plan

In the currently, Thai electrical devices and equipment program use designed from foreign manufacturers such as Japanese and America. Moreover, Thai people are still unaware of the inconvenient using of user interface which is not designed for Thai people. If no one solve this problem, Thai interface design will not be improved and never been used in its best conditions. Thereby, this research is to analyze the problems and obstacle in User interface in Thailand from government, manufacturing, education and consumers' point of views to improve and develop products. Moreover, this research is to support Thailand to have own user interface which shows the improvement in technology competency that can respond more consumers demand and to support Thailand to be one of the leading in user interface designing universally.

Acknowledgment

I would like to appreciate all professor and student at Chiba Institute of Technology, have made valuable comment suggestions on this paper, which gave an inspiration to improve data and thank all the people for their help directly and indirectly to complete the paper. Especially, I would also like to thank the all professor and student of Thai-Nichi Institute of Technology who were involved in the validation experiment for this research. Without their passionate participation and input, the validation output could not have been successfully conducted.

References

- [1] Subasic, P., & Huettner, A. (2001). Affect analysis of text using fuzzy semantic typing. *IEEE Transactions on Fuzzy systems*, 9(4), 483-496.
- [2] Thomas, E. A., & Weaver, W. B. (1975). Cognitive processing and time perception. *Perception & Psychophysics*, 17(4), 363-367.
- [3] Fukuda, T. (1992). Visual capability to receive character information: I. How many characters can we recognize at a glance? *Ergonomics*, 35(5), 617-627.
- [4] Ikeda, M., Konosu, T. (2012). Ergonomic study on the information capacity of isolating language without Kanji characters. 4th International Conference on Applied Human factors and Ergonomics, 666.
- [5] Kasuriya, S., Somlertlamvanich, V., Cotsomrong, P., Kanokphara, S., & Thatphithakkul, N. (2003). Thai Speech Recognition Corpora. *Journal of Chinese Language and Computing*, 14(4), 279-293.
- [6] Kasuriya, S., Somlertlamvanich, V., Cotsomrong, P., Kanokphara, S., & Thatphithakkul, N. (2003). Thai speech corpus for Thai speech recognition. In *Proceedings of Oriental CO-COSDA*, 54-61.
- [7] Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes I. In *Psychology of learning and motivation* 2, 89-195.
- [8] Hinrichs, J. V. (1970). A two-process memory-strength theory for judgment of recency. *Psychological Review*, 77(3), 223.
- [9] Wilpon, J., & Rabiner, L. (1985). A modified K-means clustering algorithm for use in isolated word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(3), 587-594.
- [10] Aso, H., & Konosu, T. (2014). Experimental Study on the Effect of Word Dominance of Thai Language. *Proceedings of International Conference on Business and Industrial Research*, 189- 193.
- [11] Yamane, T. (1973). *Statistics: An introductory analysis*, 2nd Ed., New York: Harper and Row.