p-ISSN: 2184-2205 e-ISSN: 2184-3414 ISBN: 978-989-35728-4-9 © 2025
https://doi.org/10.36315/2025inpact138

THE COGNITIVE COMPLEXITY IN READING THE ARABIC SCRIPT OF MALAY

Khazriyati Salehuddin, Nurjanah Mohd Jaafar, & Sharifah Raihan Syed Jaafar
Centre for Research in Language Studies and Linguistics, Universiti Kebangsaan Malaysia (Malaysia)

Abstract

The Malay language is a synchronous digraphic language as it is written in both the Roman Script (or Latin Script, henceforth, Rumi) and the Arabic Script (henceforth, Jawi). Unfortunately, the majority of the Malay speakers are not Rumi-Jawi biscriptals; in fact, Jawi is marginalized despite being the other official script for the Malay language. Jawi has been described as cognitively more complex than Rumi due to the inconsistent cognitive mapping between the Jawi symbols and the sound each symbol is said to represent. There are 4 spelling patterns in the current Jawi spelling system and this further complicates the process of reading Jawi. However, more scientific evidence is needed to support this claim. Hence, this study was conducted to illustrate how complex the process of reading Jawi is. Three experiments were conducted using the E-Prime Software that is linked to the TOBII TX300 eye tracker. Experiment 1, Experiment 2, and Experiment 3 were conducted on 30 Jawi readers to investigate the complexity in reading all four spelling patterns. Experiment 1 measured participants' correct responses and their Reaction Times in choosing the correct spelling for the words that the participants heard whereas Experiment 2 and Experiment 3 measured the participants' correct responses, reaction times (Onset), and the reading duration when two-syllable Malay words written in current Jawi spelling system were presented to them. Results show that different spelling patterns result in different mean number of correct responses and in different mean reaction times (in Experiment 1, Experiment 2, and Experiment 3). This suggests that some spelling patterns in the current Jawi spelling system are easy whereas some are not. This study hence proposes that a more consistent spelling system should be introduced to Jawi so that the number of Rumi-Jawi biscriptals can increase in the future.

Keywords: *Arabic script, cognitive complexity, Malay, reading, spelling.*

1. Introduction

Writing system refers to a set of graphic symbols that enables visual information to be understood. The Roman Script uses the alphabetic writing system as each of its symbols – which are typically called letters – represents the individual speech sound that exists in the language (Yule, 2016). The Arabic Script, on the other hand, uses the abjad writing system as each of its symbols – which are called consonant letters – represents a syllable as the vowel sounds are not clearly manifested in the syllables of the abjad writing system (Yule, 2016). Two other writing systems, i.e., the Logographic, and the Syllabic writing systems; however, they are not elaborated in this manuscript for relevance reasons.

The Malay language is a language spoken mainly in Southeast Asia and is the official language of Malaysia, Brunei Darussalam, dan Singapura (Salehuddin & Mahmud, 2023). The Malay language (henceforth, Malay) is a synchronous digraphic language as it is written in both the Roman Script (or Latin Script, henceforth, Rumi) and the Arabic Script (henceforth, Jawi). Despite the fact that both Rumi and Jawi are official writings of the Malay language (The National Language Act 1963/67, 2006), and despite the fact that the Malay language is the national language of Malaysia (The National Language Act 1963/67, 2006), the majority of the speakers of Malay are not Rumi-Jawi biscriptals. As a matter of fact, Jawi today is very much marginalized by the speakers of Malay that the Malaysian Government's initiative to reintroduce Jawi in the school curriculum is opposed by many of its own citizens (Mohd Salleh 2019). Those who oppose to this idea believe that Jawi is more relevant to the more conservative individuals and they strongly believe that the script should only be recognized as a national heritage. These justifications are unacceptable because as a synchronous digraphic language, the wide usage of Rumi does not mean Jawi is to be replaced by Rumi; it means that both Rumi and Jawi must co-exist in the linguistic landscape of the nation and be used widely and comfortably by all speakers of Malay as the National Language of Malaysia.

One of the possible reasons why Jawi is marginalized by a majority of its citizens is perhaps due to its cognitive complexity in comparison to Rumi (Salehuddin, 2012). The cognitive mapping between the Jawi symbols and the sound each symbol is said to represent is very inconsistent – unlike that of Rumi. For example, while the five vowel letters of Rumi are almost adequate to represent all the six vowel sounds of the Malay language, the three vowel letters of Jawi is far from enough to represent all the six vowel sounds of the Malay language. The use of consonant letters (e.g., <ق>, which, in Arabic is used to represent the consonant /q/) that do not reflect the sounds in Malay (e.g., <مق> ('mak') to represent the consonant /ʔ/) baffles many as the sound /q/ is nonexistent in Malay. The inconsistent manifestations of vowel letters to represent vowel sounds in its spelling (e.g., <جك> (C·C), <توا> (C·CV), <اين> (CV·C), and <حاتي> (CV·CV) for two-syllable words – when vowel sounds are obligatory sounds in all syllables) complicates the process of reading Jawi. Such an inconsistent spelling system is almost nonexistent in Rumi as all Malay vowel sounds are manifested in the form of vowel letters in Rumi. As a matter of fact, there are four spelling patterns in the current Jawi spelling system and they each come with a long list of rules and exceptions. All these are evidences that can explain why Jawi is less preferred by speakers of Malay than Rumi. However, more scientific evidence is needed to support this claim.

2. Objectives

This study was conducted to illustrate how complex the process of reading Jawi is. Specifically, it hopes to

a. outline the difficulty in identifying the correct Jawi spelling in a listening test;
b. identify which of the four current Jawi spelling pattern, is easy for readers to read; and
c. examine if there is another spelling pattern that is easier for readers to read than the one in (b) above.

Based on past studies, it is assumed that words are regarded as easy to read if they have 1) the highest mean of correct responses, 2) the fastest mean of onset to the correct responses, and 3) the shortest mean of word reading duration (e.g., McLelland, 2009; Ehri, 1995; Share, 2008).

3. Method

This study was conducted to illustrate how complex the process of reading Jawi is. Three experiments were conducted using the E-Prime Software that is linked to the TOBII TX300 eye tracker. Experiment 1 was conducted to achieve Research Objective 1, Experiment 2 was conducted to achieve Research Objective 2, and Experiment 3 was conducted to achieve Research Objective 3. Data from the eye tracker were not obtained for this study as it is not within the scope of the study. All three experiments were conducted one after another for each participant with a short break (10-15 minutes) in between. The experiments were conducted at the Eye Tracking Lab at the Centre for Research in Language and Linguistics, Universiti Kebangsaan Malaysia. Research Ethics Approval (JEP-2022-769) was obtained prior to the data collection stage.

3.1. Participants

Thirty (20 male, 10 female) native speakers of Malay who were between 17 and 58 years old participated in all three experiments. They were all Malay-English bilinguals who can read the Qur'an, which is written in the Arabic script. This inclusion criterion is important as all the words used as stimuli in all three experiments were written in the Arabic script. Their familiarity with the Arabic letters, the different forms of the letters, and what sound each Arabic letter represents are important; however, ability to communicate in Arabic, both in speaking and writing is not a requirement as being able to read the Qur'an does not mean one is able to communicate in Arabic. The participants were shortlisted from 35 individuals who responded to the call for participation through the social media. The experiments were conducted on one person at a time. All participants read the information sheet once they reached the eye tracking lab and signed the Informed Consent Form prior to participating in the experiment.

3.2. Experiment 1

Experiment 1 was conducted to outline the difficulty in identifying the correct Jawi spelling in a listening test.

3.2.1. Stimuli. 36 high-frequency, two-syllable Malay words in audio forms were used as stimuli in Experiment 1. All the words were read by a male native speaker of the Malay language, prerecorded at the developmental stage of this experiment. The words were presented to them individually via the E-Prime

software and participants were to choose, from four spelling options presented to them visually on the Eye Tracking Monitor via E-Prime, the spelling that matches the word they hear.

The four options were variations of a Malay word that was spelled in the four possible spelling patterns for Malay two-syllable words. The patterns include 1) Consonant letters only in the first and the second syllables (i.e., no vowels in both syllables); 2) Consonant letter(s) only in the first syllable; consonant letter(s) and a vowel letter in the second syllable (i.e., no vowel letter in the first syllable); 3) Consonant letters(s) and a vowel letter in the first syllable only; consonant letters only in the second syllable (i.e., no vowel letter in the second syllable); and 4) consonant letter(s) and a vowel letter in both the first and second syllables (i.e., vowel letters in both syllables). However, only one of the four options was correct. All the four options were positioned at the left, top, right, and at the bottom of the monitor.

3.2.2. Procedure. Once the informed consent form was signed, participants went through a practice trial session with 10 one-syllable words to be familiar with the experiment. Participants were asked to choose either the left, up, right, or the down arrow key to select the answers. The position of each arrow key matches with the position of the options displayed on the monitor, hence, if they think the spelling of the word, they hear matches with the word they see at the bottom of the monitor, they need to press the down arrow key. They were briefed orally, but the written form of the instruction was also displayed at the beginning of the experiment.

A fixation screen appeared for a second to keep the participants focused. A two-second audio prompt in the form of a one-syllable word was presented. This was followed by a five-second display of the four spelling options. The participants selected the spelling pattern that matches the audio display and the E-Prime automatically recorded their responses as “correct” or “incorrect”, and their reaction time to the prompt. Their focus was then brought back to the fixation screen, after which, another one-syllable word appeared. If, after 5 seconds they provided no response, they will be brought to the fixation screen, and then another one syllable word. This procedure repeated until all ten one-syllable Malay words were presented. The actual experiment took place after the completion of the practice trial, following the same procedure, only now with 36 two-syllable Malay words.

3.2.3. Data Analysis. Accuracy and reaction times data from all 30 participants were collected. However only the reaction time for the correct responses were analysed. Descriptive data were analyzed using means and SD. Repeated-measure ANOVA was conducted on the data, with four factors (Spelling Pattern: 1, 2, 3 and 4).

3.3. Experiment 2

Experiment 2 was conducted after the completion of Experiment 1 to identify which, out of the four spelling patterns of the current Jawi spelling system, is easy for readers to read.

3.3.1. Stimuli. Thirty-six high-frequency, two-syllable Malay words written in Jawi were displayed for five seconds on the eye tracking monitor via E-Prime, one word at a time, for them to read aloud. The words are not the same as those in Experiment 1.

3.3.2. Procedure. After the completion of Experiment 1, participants went through a practice trial session with 10 one-syllable words, displayed one at a time, to be familiar with the experiment. Participants were asked to read aloud each word using a Desktop USB Microphone and their audio responses were automatically recorded via the E-Prime software. A one-minute fixation screen separates the display of each word. This procedure repeated until all ten one-syllable words were presented. The actual experiment took place after the completion of the practice trial, following the same procedure, only now with 36 two-syllable Malay words.

3.3.3. Data Analysis. The audio recordings of each word for each participant were examined for accuracy and only those that were read aloud accurately were further analysed using Praat (Boersma & Weenink, 1992-2025) to measure the reaction time and the duration (in second) taken to complete reading each word. All statistical analyses were conducted using SPSS version 25.0 (IBM Corp., Armonk, NY, USA). Descriptive data were analysed using means and SD. A repeated-measure ANOVA on the data was conducted, with four factors (Spelling Pattern: 1, 2, 3 and 4).

3.4. Experiment 3

Experiment 3 was conducted after the completion of Experiment 2 to examine if there is another spelling pattern that is easier for readers to read than the one in (b) above.

3.4.1. Stimuli. Five two-syllable pseudowords (which are phonologically and orthographically plausible as Malay words) written in Jawi, were displayed for five seconds on the eye tracking monitor via E-Prime, one word at a time, for participants to read aloud. The 5 pseudowords were spelled following the four spelling patterns of the current Jawi spelling system. A fifth spelling pattern, which reflects the pattern in the Qur'an (i.e., only consonant letters with diacritics to represent vowels sounds) was also included to spell the pseudowords. The fifth spelling pattern was included because all of them can read the Qur'an fluently. Altogether there were 25 pseudowords for the participants to read.

3.4.2. Procedure. The experimental procedure for Experiment 3 was the same as the one in Experiment 2, using the same instruments.

3.4.3. Data Analysis. The audio recordings of each word for each participant were examined for accuracy. Although they were pseudowords, their accuracy was based on how they matched the expected production by the researcher. Only those that were read aloud accurately were further analysed using Praat to measure the reaction time and the duration taken to complete reading each word. All statistical analyses were conducted using SPSS version 25.0 (IBM Corp., Armonk, NY, USA). Descriptive data were analysed using means and SD. A repeated-measure ANOVA on the data was conducted, with four factors (Spelling Pattern: 1, 2, 3, 4, and 5)

4. Results

4.1. Experiment 1

ANOVA on the number of correct responses revealed that the effect of Spelling Pattern is not significant, indicating the mean accuracy was more or less the same for all spelling patterns. ANOVA on the mean reaction times revealed that the effect of spelling patterns is significant, suggesting that more time was needed to process Pattern 2 and Pattern 4. Shorter time was needed for Pattern 1 and Pattern 3.

4.2. Experiment 2

ANOVA on the number of correct responses revealed that the effect of Spelling Pattern is significant. Post-hoc comparisons suggests that words with vowel letters in both syllables were read most accurately than other patterns.

ANOVA on the mean reaction time suggests that participants took longer time to read words that have no vowels in the first syllable.

ANOVA on the mean duration suggests that the amount of time taken to completely read the all the two-syllable Malay words correctly was the same for all spelling types.

4.3. Experiment 3

ANOVA on the mean number of correct responses revealed that the effect of Spelling Pattern is significant. Post-hoc comparisons suggest that the presence of diacritics in the spelling of the pseudowords helps participants to read the pseudowords more accurately.

ANOVA on the mean reaction time revealed that the effect of Spelling Pattern is significant. Post-hoc comparisons suggest that Pattern 1 (c·c) and Pattern 5 (ć·ć) had the fastest reaction time, which indicates least processing is needed to read these two spelling patterns whereas Pattern 4 requires more processing than Pattern 5.

ANOVA on the mean duration revealed that the effect of Spelling Pattern is significant. Post-hoc comparisons showed that, Pattern B had shorter duration than Pattern 4 and Pattern 5 but were not significantly different from Pattern 4 and Pattern 3. Pattern 1, Pattern 3, Pattern 4, and Pattern 5 were not significantly different from one another

5. Discussion and conclusion

This study was conducted based on the assumption that spelling pattern that gets the highest number of correct responses, that gets the fastest reaction times, and that is read in the shortest duration is the easiest spelling pattern. Experiment 1 shows that participants were not able to identify the correct Jawi spelling for all the Malay words they heard efficiently; less than 50% of the responses were correct. This experiment also shows that although the words were common words in the participants' environment, the participants needed more time to read words with more items. This suggests that in a listening test, all four spelling patterns of the current Jawi are equally difficult.

Experiment 2 shows that participants were able to read words most accurately when there are vowels letters in both syllables, and knowledge of the Malay vocabulary helps to speed up the mean reaction times. However, the duration when reading those words were not significantly different from each other as their vocabulary knowledge helps them to complete reading the words quickly.

Experiment 3 shows that participants were able to read pseudowords most accurately when there are vowel diacritics in both syllables. This is followed by pseudowords with vowel letters. This suggests that knowledge in vocabulary is negligible when the objective is to enable pseudowords to be read accurately. However, although the presence of vowel marks increases the "decipherability" of the words (Share, 2008), they took more time to start to read those words and spent the longest to complete them as there were more items to look at.

This study has shown that only one of the spelling patterns in the current Jawi spelling system is easy whereas the rest are not. Although as mentioned earlier, 'easiest' refer to the highest mean number of correct responses, the fastest reaction times, or the shortest duration, for this study, 'easiest' was only indicated by the mean number of correct responses. Although in most cases, longer time indicates more difficulty in reading, the presence of either diacritics or vowel letters in both syllables actually help readers to be more careful in their readings, which results in the two highest correct responses for these two types. This suggests that accuracy is generally more important than speed.

This study has shown that reading Jawi is not entirely difficult. If vowel letters or diacritics are consistently present in all of its syllables, reading Jawi will be much easier. Readers do not have to possess knowledge of the Malay vocabulary and grammar to be able to read Jawi if each of the symbols used in Jawi can be matched to a particular speech sound in the Malay language. If such spelling pattern consistently exist in Jawi, the aim to increase the number of Rumi-Jawi biscriptals among speakers of the Malay language can be materialized.

Acknowledgments

This study is funded by Universiti Kebangsaan Malaysia (UKM) through the Research University Grant (Geran Universiti Penyelidikan (GUP), GUP-2022-055). Acknowledgements are given to all the 30 participants who participated in this study. Special acknowledgements are also given to the following individuals: Nur Ad-Diina Shamsudin, Ahmad Naiimuddin Shamsudin, Ahmad Najjmuddin Shamsudin, Nur Aiman Noor Mu'azam, Fatin Nadiah Mahmud, Yin Hanmo, and Zhang Minwen for assisting us in various ways throughout this study.

References

Boersma, P. & Weenink, D. (1992-2025). *Praat: doing phonetics by computer [Computer Program]*. Version 6427. Retrieved January 28, 2025, from https://www.praat.org

Ehri, L. C. (2005). Learning to Read Words: Theory, Findings, and Issues. *Scientific Studies of Reading*, *9*(2), 167-188.

Mahmud, F. N. & Salehuddin, K. (2023). How Bilingual Are Malaysian Undergraduates? A Snapshot of the Different Bilingual Categories in Malaysia. *GEMA Online® Journal of Language Studies, 23*(2), 144-164. http://doi.org/10.17576/gema-2023-2302-08 (Scopus, WoS-ESCI)

McLelland, N. (2009). Linguistic Purism, Protectionism, and Nationalism in the Germanic Languages Today. *Journal of Germanic Linguistics, 21*(2), 93-112. https://doi.org/10.1017/S1470542709000208

Mohd. Salleh, N. H. (2019, December 19). Takutkan Islamisasi punca Jawi ditolak, kata Dong Zong. *Free Malaysia Today*. Retrieved May 31, 2020, from https://www.freemalaysiatoday.com/category/bahasa/2019/12/26/takutkan-islamisasi-punca-tulisan-jawi-ditolak-kata-dong-zong/

Salehuddin, K. (2012). Penilaian ke atas kerumitan kognitif dalam proses membaca Jawi. *GEMA Online® Journal of Language Studies, 12*(4), 1181-1194.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of over-reliance on an "outlier" orthography. *Psychological Bulletin, 134*, 584-615. https://doi:10.1037/0033-2909.134.4.584

The 1963/67 National Language Act. (2006). *Undang-undang Malaysia 32*. Retrieved May 22, 2023, from https://www.lawyerment.com/library/legislation/acts/1963/32/cite/?s=2

Yule, G. (2019). *The Study of Language* (6th Ed.). Cambridge: Cambridge University Press.