The effectiveness of phonological reading aids and the role of self-regulation in L2 Japanese kanji word learning

Brendan Joshua Allen

Please note that some graphs/tables/images may be removed in order to comply with copyright restrictions.

MSc in Applied Linguistics and Second Language Acquisition (ALSLA), 2023

DECLARATION BY THE CANDIDATE AS AUTHOR OF THE DISSERTATION

1. I understand that I am the owner of this dissertation and that the copyright rests with me unless I specifically transfer it to another person.

2. I allow the Department to deposit on my behalf a copy of this dissertation in the Oxford University Research Archive ('ORA') where it shall be freely available online for use in accordance with ORA's Terms and Conditions of Use [https://ora.ox.ac.uk/terms_of_use].

3. I understand that this dissertation should not contain material that can be used to personally identify individuals or specific groups of individuals (unless permission has been obtained from the individuals) and that such material should be removed before this dissertation is deposited in ORA.

4. I agree to be bound by the terms of the ORA Grant of Non-exclusive Licence [https://ora.ox.ac.uk/deposit_agreements] and I warrant that to the best of my knowledge, making my thesis available on the internet will not infringe copyright or any other rights of any other person or party, nor contain defamatory material.

5. I agree that my dissertation shall be available for download in ORA in accordance with paragraphs 2, 3 and 4 above.

Signed [an electronic signature is sufficient]:	Brendan J Allen
Date:	28/09/2023

The effectiveness of phonological reading aids and the role of self-regulation in L2 Japanese kanji word learning





Brendan Joshua Allen

St Anne's College

Dissertation submitted to the University of Oxford in partial fulfilment of the requirements for the degree;

Master of Science in Applied Linguistics and Second Language Acquisition

Trinity Term, 2023

Word count: 19,998

ABSTRACT

Accessing Japanese kanji word pronunciation is particularly challenging for Japanese second language (L2) learners due to the phonologically opaque nature of the Japanese writing system. As a result, this 'phonological inaccessibility' can negatively impact second language (L2) Japanese kanji word decoding, a vital skill in the learning and retaining L2 vocabulary. 98 adult beginner-to-intermediate level participants were recruited across divergent first language (L1) writing systems to investigate the effectiveness of using romaji and furigana as phonological reading aids in L2 Japanese kanji word learning.

Results from an ANCOVA analysis found a *significant* overall main effect between the two phonological learning conditions, with a higher recall of novel kanji word pronunciation using the Romanised version of Japanese. In terms of phonological decoding of L2 Japanese kanji words, the results of this study provide further empirical support of the influence of cross-linguistic transfer in L2 learning. The effectiveness of phonological reading aids in kanji decoding by proficiency level found *no* statistically significant differences between learning groups. The implications for pedagogy are that early L2 learners of Japanese may benefit from using a Romanised version of Japanese to reinforce L2 kanji learning before moving on to more authentic writing systems.

This study also explored L2 kanji word learning under the framework of self-regulation. Differences between L2 Japanese learners across divergent writing systems and proficiency levels were explored using a modified version of Tseng et. al's. (2006) SRCVoc, called the 'Self-Regulatory Capacity for learning Kanji' developed by Rose (2017). Results of the cross-orthographic self-regulation analysis found *no* significant differences between different L1 learners, but *did* find significant differences in *commitment control* and *metacognitive strategies* by L2 proficiency level. Beginner-level learners scored higher than more advanced Japanese learners, suggesting that more advanced learners may suffer from a breakdown in self-regulation as they progress through higher stages of the language.

ACKNOWLEDGEMENTS

First of all, I'd like to thank my supervisor, Professor Heath Rose. Heath first inspired me to apply to Oxford after meeting him at Rikkyo University in Tokyo more than 10 years ago. Heath has been a wonderful and supportive supervisor. Without his expert guidance, critical feedback, and general encouragement, this dissertation would not have been possible.

Next, I wish to express my eternal gratitude to my almost 90-year-old grandmother, Jennifer Allen-Williams. As a former academic, teacher, and sole legal guardian to me since the age of four, my grandmother has been a tremendous driving force in my life. She has always encouraged me to pursue my academic interests. Without her guidance and generous financial support, completing this master's would not have been possible. Therefore, this dissertation is dedicated to her.

I'd like to thank all of my friends and colleagues who generously donated to my GoFundMe campaign. In total, £2,425 was raised from 52 donors. Without all of your support, I would not have been able to take-up my place at Oxford. I'd like to give special wish to my former boss Mark Annear. Mark not only helped me a great deal personally, he was also extremely positive in supporting my academic endeavours. So, I owe a large part of this journey to him.

I wish to acknowledge the generous financial support received from the Huxter Fund. The Huxter scholarship allowed me to dedicate myself fully on my studies without the burden of financial worry. I am grateful to Dr and Mr Huxter for their continued support of less-fortunate students at St Anne's College, Oxford.

I would also like to give a special mention to Dr Matthew Williams, Access Fellow, Jesus College, Oxford. Matt is doing tremendous work to raise access awareness through his YouTube channel, and he was extremely generous with his time to personally show me around Jesus College during the application process. His help was invaluable in answering my questions about the postgraduate application process.

Finally, to all the teachers in my life who pushed me and inspired me to aim higher. Educators possess a special gift. They hold the ability to profoundly shape the lives of young minds, inspiring them to go forward, 'purposely and boldly' [c*onsulto et audacter*] – St Anne's College motto.

"*We all have opportunities in life, it's whether or not you take them.*"

TABLE OF CONTENTS

ABSTRACT……1
ACKNOWLEDGEMENTS…… 2
TABLE OF CONTENTS…….. 3
LIST OF TABLES AND FIGURES……5
LIST OF ABBREVIATIONS…….. 7
GLOSSARY……8
1. INTRODUCTION
1.1 Statement of the problem…… 10
1.2 The Japanese writing system……10
1.3 Challenges in L2 Japanese kanji learning…… 11
1.4 Phonological inaccessibility in L2 kanji learning…… 12
1.5 Phonological reading aids……13
1.6 Rationale for the study……14
1.7 Aims of the study…… 15
1.8 Research questions……16
2. LITERATURE REVIEW
2.1 Phonological reading aids in L2 Japanese word learning……17
2.11 Romaji……17
2.12 Furigana……22
2.2 Japanese L2 kanji word processing……25
2.3 Cross-linguistic transfer……27
2.4 L2 Japanese kanji flashcard learning …… 28
2.5 Memory and L2 word learning…… 29
2.6 Self-regulation and kanji word learning……30
2.7 Sumary and relevance to the study……33
3. SAMPLE
3.1 The sample……34
3.2 L2 Japanese proficiency assessment……38
4. METHODOLOGY
4.1 Study design……42
4.11 Quasi-experimental design…… 43
4.12 Data analysis methods…… 44
4.2 Data collection tools ……44
4.21 Selecting target vocabulary …… 44
4.22 Measuring L2 kanji word knowledge……46
4.23 Measuring L2 self-regulation…… 47
4.24 Materials…… 49
4.3 Experiment procedure……50
4.31 Pre-test…… 50
4.32 Flashcard learning cycle…… 51
4.33 Immediate post-test…… 53
4.34 Delayed post-test …… 54
4.4 Research ethics……54
5. RESULTS
5.1 Analysis and interpretation of data……55

5.11 Outlier detection and treatment……56
5.12 Normality testing and analysis…… 56
5.13 Descriptive statistics……..58
5.14 Inferential statistics…… 59
6. DISCUSSION
6.1 Interpretation of results……..62
6.11 *RQ1*
To what extent do different phonological reading aids (romaji and furigana) help in the learning and retention of kanji word pronunciation?……..**63**
6.12 *RQ2*
What is the relationship between self-regulation, L1 background, and L2 proficiency in Japanese kanji word learning?……**64**
6.2 Implications for pedagogy…… 65
6.3 Limitations of the study…….. 66
6.31 Methodological limitations……66
6.32 Pedagogical limitations…… 68
7. CONCLUSION
7.1 Summary of findings……..70
7.2 Directions for future research…… 70
REFERENCES……..71
APPENDICIES……89
Appendix A Background screening survey…… 90
Appendix B Participants by location, proficiency and L1 background….91
Appendix C SPOT90 scores of all subjects tested……..92
Appendix D SPOT90 score conversion table (TTBJ) …… 92
Appendix E Online marketing poster……93
Appendix F Marketing post on Reddit……..94
Appendix G Gorilla experiment flowchart……95
Appendix H Pilot vocabulary items (10) ……..98
Appendix I Target vocabulary items (16)……99
Appendix J SRCKanji survey (adapted from Tseng et. al. (2006))……..101
Appendix K Examples of romaji transliteration systems ……..103
Appendix L Boxplot of SPOT90 scores by proficiency level…… 104
Appendix M Boxplot of Test scores by proficiency level…… 105
Appendix N Boxplot of Test scores by treatment group……..106
Appendix O Z-score calculations of SPOT90 …… 107
Appendix P Z-score calculations of test scores, by proficiency…… 108
Appendix Q Z-score calculations of test scores, by treatment group……109
Appendix R ANOVA and ANCOVA output from SPSS……..111
Appendix S Kruskal-Wallis output from SPSS……..114
Appendix T CUREC approval…… 115
Appendix U Participant online consent form…… 116
Appendix V Online consent form…… 119
Appendix W Online marketing poster……120
Appendix X Dual language business cards…… 121

LIST OF TABLES AND FIGURES

LIST OF TABLES

Table 1.1 Types of reading aids in L2 Japanese pedagogy with example texts.

Table 1.2 JLPT examinees by country, 2019.

Table 3.1 Breakdown of final sample (N=75) by proficiency and treatment group.

Table 3.2 Descriptive statistics of the sample (N=75).

Table 3.3 List of universities and institutions where participants were recruited who were enrolled on formal Japanese language programmes.

Table 3.4 Examples of TTBJ SPOT90 questions.

Table 4.1 Summary of target 16 words and their corresponding word frequency bands.

Table 4.2 Sample of target Japanese kanji words used in the experiment by word type and word frequency.

Table 4.3 Cronbach's alpha internal reliability of the SRCKanji (N=98).

Table 4.4 Example of how novel L2 Japanese vocabulary were presented to participants under each treatment condition.

Table 5.1 SPOT90 descriptive statistics across L2 proficiency level (N=75).

Table 5.2 Descriptive statistics showing Japanese kanji word pre, post and delayed post-test raw test scores by proficiency (N=75).

Table 5.3 Descriptive statistics for the SRCKanji sub-scale coefficients by L2 proficiency and L1 background (N=98).

LIST OF FIGURES

Figure 1.1 Differences between major writing systems (with examples written in their given native script).

Figure 1.2 Sample text illustrating the prevalence of kanji in modern Japanese written text.

Figure 2.1 Media advertisement which shows the use of the Roman alphabet in modern Japanese.

Figure 2.2 Excerpt from early L2 Japanese textbook 'GENKI I' 「げんき」 which utilises the romaji script alongside Japanese kana

Figure 3.1 Chart showing participant country of residence of the final sample (N=75).

Figure 3.2 SPOT90 test user interface, showing close target item, responses, and timer.

Figure 4.1 Flow diagram of the quantitative design of the study.

Figure 4.2 Illustrative representation of the learning task on Gorilla (group B – romaji condition).

Figure 4.3 User interface on Gorilla of the immediate *post-test* (group B – romaji condition).

Figure 5.1 Graph showing gain scores in post-test and delayed post-test by treatment group (N=75).

Figure 5.2 Stacked bar-chart showing raw mean test scores by group and proficiency level.

Figure 5.3 Interaction plot showing the estimated marginal means of the post and delayed post-test scores by treatment group controlling for pre-test scores.

LIST OF ABBREVIATIONS

ESL	English as a Second Language, where the learning environment is in the context of an English-speaking country.
EFL	English as a Foreign Language, refers to the teaching of English in where English is not the dominant first language (i.e. Japan)
JFL	Japanese as a Foreign Language (in a non-Japanese environment. JFL education tends to focus more on essential skills to communicate
JLPT	The Japanese Language Proficiency Test (日本語能力試験)[1] is a test aimed at second language speakers of Japanese, consisting of five proficiency levels, ranging from N5 (lowest) to N1(highest).
L1	First language.
L2	Second language.
SPOT90	The 'Simple Performance-Orientated Test' is an L2 Japanese proficiency test which measures the automaticity of Japanese linguistic processing, forming part of a larger battery of tests called the Tsukuba Test Battery of Japanese (TTBJ).
SRCKanji	Adapted from Tseng et. al's (2006) SRVoc, the 'Self-Regulatory Capacity for learning Kanji' was developed by Rose (2017) to assess the self-regulatory capacity and strategy use in L2 Japanese kanji learning.
SRCVoc	The 'Self-Regulatory Capacity for learning Vocabulary' was developed by Tseng et. al., (2006), is a 20-point survey used to measure learners' underlying self-regulation in L2 vocabulary learning.
SRS	SRS is an acronym for Spaced Repetition System; a pedagogical approach to learning underpinned by an algorithm which systematically presents content to learners at intervals based prior recall ability.
TTBJ	The Tsukuba Test Battery of Japanese consists of several L2 Japanese proficiency tests developed by researchers at the University of Tsukuba, Japan.

[1] For more information on the JLPT please see https://www.jlpt.jp/e/

GLOSSARY

Alphabetic script	A type of orthographic writing system based on alphabetic characters, where graphic units represent phonemes
Furigana (振り仮名)	Furigana are small syllabic characters written above kanji to aid word pronunciation usually written in syllabic hiragana. Furigana are a pedagogical tool used in native early learning texts.
Grapheme	Smallest unit of meaning, typically a kanji may contain one (being a grapheme itself) or several smaller component graphemes.
Hànzi (汉字)	Mandarin Chinese logographic characters.
Hiragana (平仮名)	*Hiragana* forms one of the two syllabic Japanese writing systems (along with *katakana*).
Jukugo (熟語)	Jukugo（熟語）are words which are typically formed by combining two or more kanji characters to make a kanji compound, often forming nouns
Jouyou kanji (常用漢字)	2,136 official 'regular-use *kanji'* as set-out by the Ministry of Education (MEXT), forming the total number of kanji studied by Japanese children after 12 years of formal schooling.
Kana (仮名)	*Kana* is the name used to combine both of the two Japanese syllabic scripts, *hiragana* (ひらがな) and *katakana* (カタカナ). *Kana* can be used to decode Japanese phonological units
Katakana (カタカナ)	A fully syllabic Japanese writing system typically reserved for foreign or loan words
Kanji (漢字)	Japanese logographic characters which were imported from Mandarin Chinese *hànzi* (汉字).
Kango (漢語)	A Japanese word of Chinese Mandarin origin, sharing semantic and/or phonological similarities
Kun-yomi (訓読み)	Japanese reading of *kanji* character
Logographic script	A type of orthographic writing system that uses symbols to represent individual words or concepts.
Morphographic script	A writing system which uses symbols to represent semantic and/or phonological meanings.

Phonological reading aid	A phonological clue, usually written above or next to kanji character to assist in deriving a word's pronunciation. In Japanese *furigana* uses syllabic *hiragana*, in a similar way that Chinese *pinyin* uses the Roman alphabet letters to aid word pronunciation.
Pinyin	A pedagogical tool which to aid word pronunciation. *Pinyin* uses the Latin alphabet, which each letter corresponding to a character's sound.
On-yomi (音読読み)	Chinese reading of Japanese kanji character
Romaji (ローマジ)	A Romanised version of Japanese words using the Latin or Roman alphabetic script.
Syllabic writing system	A writing system where each character represents a character. Japanese *hiragana* and *katakana* are both fully syllabic writing systems of Japanese. For example, the characters さ・サ both represent the syllable 'sa' in hiragana and katakana respectively
Wago (和語)	Words of native Japanese origin, often written in the two Japanese syllabic scripts hiragana and katakana.

1. INTRODUCTION

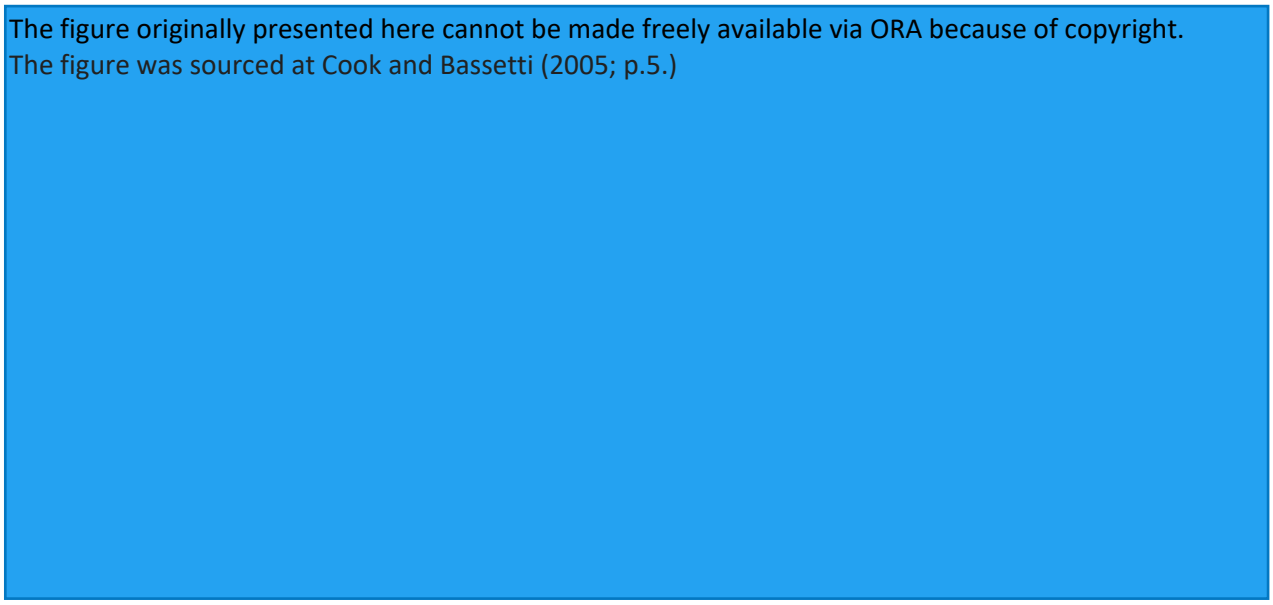
1.1. Statement of the problem

Japanese is considered one of the most difficult languages in the world to learn (Everson, 2011). In particular, leaning of thousands of kanji (Chinese characters), has been identified as one of the most challenging aspects of L2 Japanese (Gamage, 2003; Mori and Shimizu, 2007). As a result, L2 Japanese rate of progression is much slower compared to students studying European languages (Everson, 2011). A particularly challenging problem in learning Japanese kanji is being unable to successfully decode phonological representations of kanji characters and kanji word compounds. This phonological inaccessibility is largely due to the complex and opaque nature of the orthographic writing system (Mori, 1998, Joyce, 2011). For Japanese second language (L2) learners, deficiencies in kanji word decoding (i.e., word recognition), manifests into reduced automaticity required for important higher-level word processing, and L2 Japanese reading comprehension (Koda, 2017). The immense cognitive challenges presented by the orthographically complex writing system inevitably leads to learner frustrations, and has been linked to high attrition rates on Japanese language programmes (Samimy & Tabuse, 1992). Indeed, staying motivated is a one of the major challenges for L2 Japanese learners (Rose, 2017), yet understanding the nature of *how* various individual Japanese learners self-regulate and monitor their own behaviour during the kanji learning task is far less known.

1.2. The Japanese writing system

Combined, the Japanese writing system comprises of four scripts; one morphographic script with uses kanji, or Chinese characters where characters represent both sound and meaning, two syllabic hiragana and katakana (collectively called kana), where each character represents one syllable, and one alphabetic, a Romanised version of Japanese called romaji. Japanese natives are able to effortlessly switch between reading phonetic and morphographic writing systems at equal speeds (Tamaoka, 2014). Paradoxically, despite the complexities of the Japanese writing system, literacy rates among natives are extremely high (Paxton & Svetanant, 2013). Figure 1.1 highlights the orthographic differences between *syllabic* kana and *morphographic* kanji compared with the *phonemic* sound-based Roman alphabetic script.

Figure 1.1 Differences between major writing systems (with examples written in their given native script)

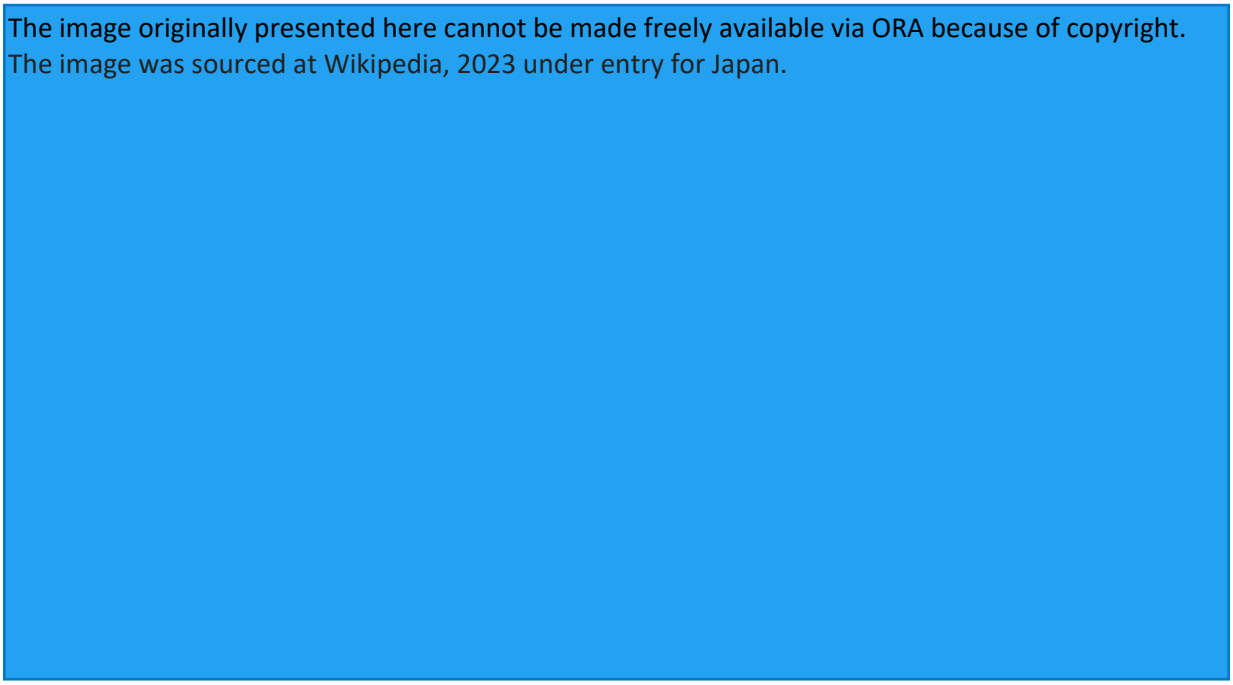


1.3. Challenges in L2 Japanese kanji learning

Japanese kanji form the backbone of the Japanese writing system, and are used to convey meaning in modern Japanese written texts (Tamaoka et. al., 2017), with kanji forming up to 70% of the words in current usage in the Japanese dictionary (Yokosawa & Umeda, 1988). Figure 1.2 shows the dominance of kanji in a typical passage of Japanese written text. For native speakers, after nine years of formal schooling, Japanese children will have command of at least 2,276 kanji characters required to function in modern day Japan (Rose, 2017). Even beyond formal schooling, the route to kanji literacy is 'often a lifelong process' (Akamatsu, 2006; p.486). Thus, for the L2 learner, kanji is one of the most difficult hurdles to overcome, as learners must know the semantic and phonological representations of thousands of kanji characters which places an immense cognitive burden and results in delaying the time to reach proficiency (Paxton & Svetanant, 2013).

Kirwan (2005) estimates knowledge of 300-400 kanji characters is required to gain literacy. Unfortunately, learning hundreds of characters is typically outside the scope of many formal Japanese second language programmes. Ono (1973) points out that the failures of formal

Japanese programmes which fail to teach such minimum number of characters required to reach anywhere near fluency as 'educationally immoral and criminal' (Ono, 1973, p.4). There is often not enough time to cover course material *and* teach kanji, so often kanji learning in done individually, outside of the classroom. Thus, problems may arise when the cognitive burden of learning thousands of kanji characters falls upon the learner him or herself. Learners must learn to regulate and control their behaviour to achieve their desired learning goals.



1.4. Phonological inaccessibility in L2 kanji learning

Compared to Chinese hànzi, Japanese kanji are significantly more phonologically opaque (Joyce, 2011), as Japanese kanji 'do not lend themselves to phonological access by their readers' (Everson, 2011; p.253). Gaining full phonological understanding of kanji is multifaceted, and requires not only phonological decoding, but also knowledge of complex 'compositional structures' (Mori, 2014; p.414). In terms of phonological organisation; kanji are orthographically deep (i.e. not congruent), can differ both at the character or word level, and can have several possible phonological readings depending on word order or context (Cook & Bassetti, 2005; p.8). For example, the kanji character for moon [月] , can be read in as '*tsuki'* (つき), '*gatsu'* (がつ), or '*getsu'* (げつ). Similarly, when the kanji '*moon*' (月) appears in jukugo

compounds, kanji word pronunciations for '*every month*', '*August*' and '*next month*' are read; *mai-tsuki* (毎月), *hatchi-gatsu* (八月), and *rai-gestu* (来月) respectively. A lack of phonological transparency, or the 'correspondence between symbols and their corresponding sounds', can result in a 'phonological deficiency' for L2 learners (Cook & Bessetti, 2005; p.8). More simply, when Japanese learners encounter a kanji word which they do not know how to decode, they may become 'stuck' in being able to successfully identify its correct phonological form. The degree of L2 phonological access may be influenced by the orthographic similarity (or distance). In terms of phonological access for western learners with alphabetic L1 backgrounds, learning orthographically deep lexical forms without prior knowledge of Chinese logograms is particularly challenging (Dwyer, 1997; Kato, 2005; Rose & Harbon, 2013). On the other hand, for learners from L1 logographic script backgrounds such as Mandarin Chinese, although Chinese hànzi, share more than half of semantic similarities with Japanese kanji (see Chen, 2002), their script advantage is significantly reduced for Japanese kanji phonology due to complex phonological forms of Japanese kanji (Tamaoka, 2014).

1.5. Phonological reading aids

For early L2 learners of Japanese, there are two key phonological reading aids which can be employed to overcome the challenge phonological inaccessibility arising from complex kanji and kanji words. Romaji, a Romanised form of the Japanese script (forming one of the four Japanese scripts[2]), and furigana, small hiragana characters which are typically written in small superscript, written above the kanji characters. Both of these reading aids can be used to assist in L2 (and L1) kanji word pronunciation. Table 1.1 provides a visual example of how romaji are typically used as kanji word pronunciation cues in early-learner L2 Japanese textbooks, and how furigana appears in L1 children's textbooks (or manga comic books intended for adolescents).

[2] Romaji forms part of the writing system, although romaji is *not* formally recognised in the Japanese orthographic system (Kess & Miyamoto, 2000)

Table 1.1 Types of reading aids in L2 Japanese pedagogy with example texts		
L1 native text (no aid)	**L2 text** (romaji aid)	**L1/2 text** (furigana aid)
突然	totsuzen	とつぜん 突然
突然、雨が降り出して私たちは傘をささなかった。	**totsuzen**, ame ga furidashite watashitachi wa kasa o sasanakatta.	とつぜん　あめ　ふ　だ **突 然**、雨 が降り出して わたし　かさ 私 たちは 傘 をささなかった。

1.6. Rationale for the study

In Japanese foreign language (JFL) contexts, there is much debate from practitioners whether to use romaji or kana from the outset of Japanese second language learning (see Emmerson, 2011). To date, there is extremely scarce empirical evidence to support using either as a pedagogical aid, but those who support using romaji, assert that is no detriment in learning outcomes in delaying the onset of more authentic scripts such as hiragana. In English as second language (ESL) contexts, Milton & Meara, (1995), point out that successful learners acquire 2,500 words per year. In JFL instructed contexts, learning the required number of kanji to be become proficiency is often well-below the required amount (Ono, 1973). Therefore, understanding more efficient ways to learn L2 Japanese words (kana and kanji) is a relevant area of for Japanese second language acquisition research (SLA) to explore. The rate of L2 vocabulary gains plays a pivotal role in early second language acquisition (Ellis, 1995), but little evidence is known about the acquisition of kanji and kanji words in non-alphabetic languages such as Japanese (Okuyama, 2007). Therefore, the rationale for this study is to provide necessary empirical evidence in the field which is significant (and relevant) for both SLA theory and for L2 Japanese pedagogy.

A further reason for the study is that much of the Japanese SLA literature is heavily skewed toward English L2 learners of Japanese (Mori et. al., 2021), despite only accounting for a small fraction of the total Japanese learner population (JEES[3], 2019). Over 80% of L2 Japanese learners are based in East and Southeast Asia region, with Chinese-script learners (e.g., China,

[3] Japan Educational Exchanges and Services (JEES)

Hong Kong, Taiwan) forming the overwhelming majority of Japanese learners. UK applicants account for just 0.1% of total JLPT test takers, compared to mainland Chinese, who represent just under a quarter (24%) (see Table 1.2). Therefore, an additional aim of this study is to address this research gap by sampling Japanese learners from non-English L1 backgrounds such as China, Hong Kong and Taiwan providing much needed empirical data from more representative population of Japanese learners, which are relevant for the generalisability of the results of the study.

Table 1.2. JLPT examinees by country, 2019

JLPT test site country	**Number of examinees**	**% of total examinees**
Japan*	405,862	37%
China	267,290	24%
Korea	81,016	7%
Taiwan	77, 304	7%
Vietnam	74,334	7%
Myanmar	44,022	4%
India	26,604	2%
Thailand	25,134	2%
Indonesia	20, 984	2%
Philippines	16,298	1%
Hong Kong	13,010	1%
United Kingdom	1,332	0.1%
Total	**1,100,896**	100%

*Examinees by country not available (JEES)

1.7. Aims of the study

This aim of this study is to address the problem of phonological inaccessibility in L2 Japanese kanji reading. It will do this by examining the effectiveness of two phonological reading aids (romaji and furigana) on kanji compound word pronunciation at two different stages of Japanese language development (beginner and intermediate), and by learners from alphabetic and logographic L1 backgrounds. The study also aims to explore the role of cross-linguistic influence between alphabetic and non-alphabetic learners on the efficiency of learning novel

kanji words using orthographically dissimilar reading aids. This thesis further aims to explore the impact of the cognitive demands of L2 kanji learning on Japanese learner strategic behaviours, underpinned by the theoretical framework of self-regulation. Through the lens of self-regulation, this study seeks to understand the impact of L2 kanji learning on learner behaviour as proficiency level changes. The study hopes to develop on works by Rose & Harbon (2013), and Rose (2017), in exploring qualitative differences between self-regulated behaviour in learners of Japanese in similar (logographic), and dissimilar (alphabetic) writing systems.

1.8. Research questions

Given the aims outlined above, the research questions of this study are centred around the effectiveness of phonological reading aid type on kanji word pronunciation. Sub-research questions further aim to explore the effectiveness of different reading aids across proficiency level and L1 background. The following research questions (RQ), and sub-questions are outlined below;

RQ1. To what extent do different phonological reading aids (romaji and furigana) help in the learning and retention of L2 Japanese kanji word pronunciation?

1.1 What are the relative effects of different reading aids and L2 Japanese proficiency level on L2 kanji word pronunciation?

1.2 What are the relative effects of different reading aids and L1 background on L2 kanji word pronunciation?

A secondary research question aims to explore the relationship between self-regulation and the strategic behaviour in kanji word learning in learners across different proficiency levels, and L1 backgrounds. Therefore, the following research question is raised;

RQ2. What is the relationship between self-regulation, L1 background and L2 proficiency in Japanese kanji word learning?

2. LITERATURE REVIEW

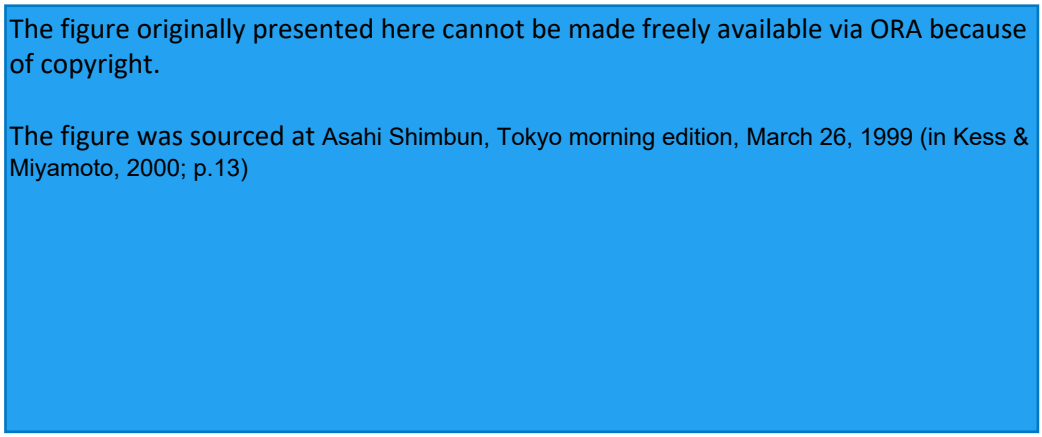
2.1. Phonological reading aids in L2 Japanese word learning

In order to navigate the phonological barrier of unknown kanji words, there are two key phonological reading aids used in L2 Japanese pedagogy, romaji and furigana. Romaji, which literally means "Roman letters" (Tamaoka, 2014), are primarily used for transcribing L2 Japanese words (Hadamitzky & Spahn, 1981), and are a simply a transcribed version of Japanese written in the Latin alphabet. Furigana, are small hiragana characters typically written above difficult, or ambiguous readings of kanji words (Kirwan, 2003). Unlike romaji, furigana form part of native L1 reading materials, particularly for younger school children, or used as an aid to decode difficult kanji words in adolescent comic books. L1 texts containing furigana can be used directly for L2 reading, but learners' must be familiar with hiragana. Hence, there is an argument for early learners to learn kana immediately from the outset, which may be extremely important in developing interactions with unknown kanji words, in turn may develop learners' automaticity in L2 kanji character and word recognition.

2.11 Romaji

In modern Japanese, romaji is typically used in Japanese advertising (e.g., Tokyu Hands department store) (see Figure 2.1), where the Roman script is perceived to being more 'stylish' (Akamatsu, 2006; p.485), for acronyms (e.g., JR for Japan Railways), or used in specific corporation names (e.g., Sony). Used alongside the three Japanese scripts, romaji contrastive. Figure 2.1 shows an example of romaji in Japanese media advertising, often used as a contrast to regular Japanese text (Kess & Miyamoto, 2000). Unlike kanji and kana, romaji and are not used to write strings of continuous text in the same way as the other Japanese scripts (Neustupny, 1984; Seeley, 1991). Thus, one of the most obvious problems with using romaji as a pedagogical tool as that romaji is not used in the same manner as other Japanese scripts. Moreover, romaji is not as widespread compared to kanji and kana (Akamatsu, 2006), perhaps as the romaji script is not formally recognised as part of the Japanese writing system officially (Kess & Miyamoto, 2000). Thus, for the L2 learner, while learning Japanese using romaji may be beneficial during the early stages, there are clear limitations of romaji's benefit as learners progress through the language using more authentic scripts.

Figure 2.1. Media advertisement which shows the use of the Roman alphabet in modern Japanese



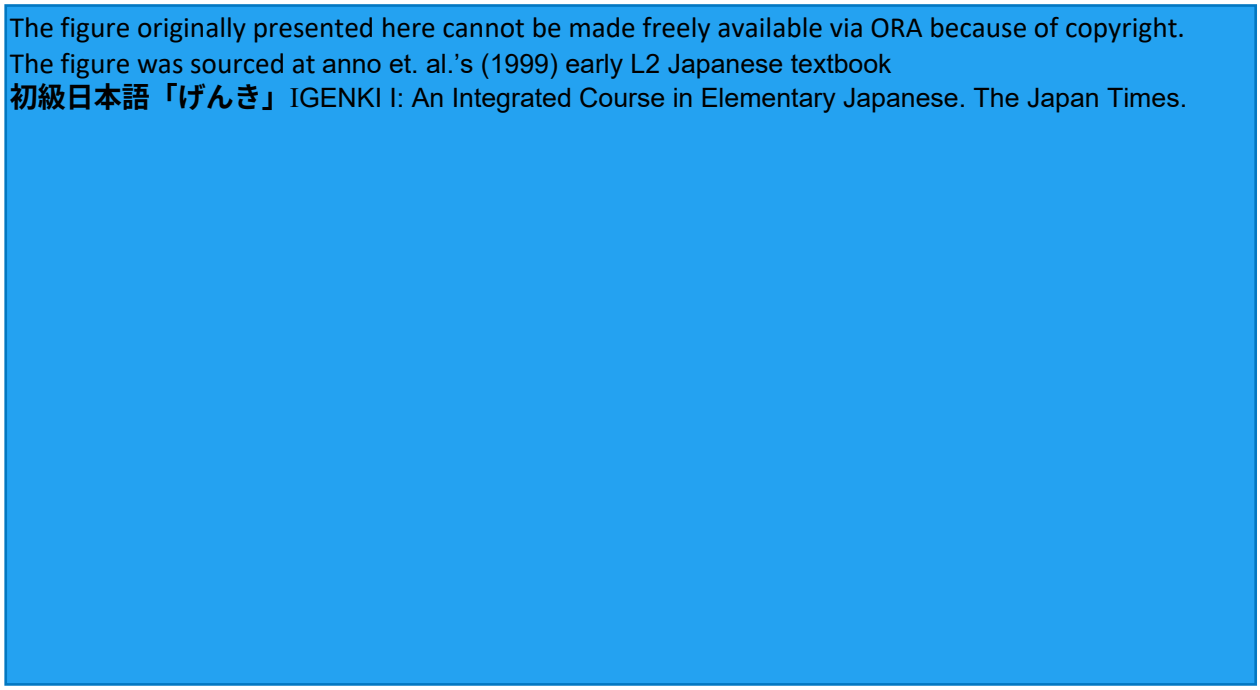
Reprinted from the *Asahi Shimbun*, Tokyo morning edition, March 26, 1999 (in Kess & Miyamoto, 2000; p.13)

Despite romaji's long-term limitations, for beginner-level L2 Japanese learners, romaji is widely used as an early pedagogical tool in L2 textbooks as a method to access and decode the phonologically complex kanji (Dwyer, 1997). The prevalence of romaji in L2 textbooks is particularly prevalent in the United States (Okuyama, 2007). Many early-learner self-study Japanese textbooks[4] are written entirely in romaji (Okuyama, 2007). In a review of 355 L2 textbooks over 127 years, 52.1% solely used a form of Romanisation. However, in recent years, the percentage of romaji in L2 Japanese texts have been steadily been reducing in favour of more authentic texts (Kirwan, 2003; p.155). Figure 2.2 shows an excerpt from an early-level beginner textbook called GENKI. A benefit of using romaji is that beginner-level learners can immediately access Japanese kanji without having to memorise kana (46 hiragana and 46 katakana characters) beforehand. However, some theorists argue that kana can be learned with relative ease, and delaying the onset of authentic kana can hinder language progression (Koda, 1992, Rose, 2017). Nevertheless, from an educational psychology perspective, allowing early learners to use romaji as a pedagogical tool has been found to reduce some L2 Japanese language anxiety associated with learning such an orthographically complex script (Okuyama, 2007). Romaji may be more 'beneficial' for learners who originate from L1 alphabetic scripts (Kess & Miyamoto, 2000; p.114), suggesting that using romaji may make learning Japanese kana more easily due to employing a script which is already firmly cemented

[4] Self-study romaji texts include; *Japanese in 10 minutes a day* (Kershul, 1992), and *Master the Basics: Japanese* (Akiyama & Akiyama, 1995)

as part of their first language orthographic system (Kess & Miyamoto, 2000). Similarly, Chinese Mandarin learners may also benefit from romaji in Japanese kanji decoding having knowledge of the phonological system pinyin.

Figure 2.2 Excerpt from early L2 Japanese textbook 'GENKI I' 「げんき」which utilises the romaji script alongside Japanese kana



Excerpt from Banno et. al.'s (1999) early L2 Japanese textbook *初級日本語「げんき」Ⅰ GENKI I: An Integrated Course in Elementary Japanese*. The Japan Times.

There is much debate from L2 Japanese practitioners surrounding the use of romaji. Romaji is often criticised as being a 'crutch' for learning development (Rose, 2017; p.134). Indeed, early proponents of using romaji as a pedagogical tool by Crowley (1965), have been continuously challenged by more recent thought (Everson, 1994; Pye, 1984), suggesting that romaji may cause more complex problems in the long run (Koda,1992). One particular issue surrounding romaji is that there is no consensus in terms of regulating the Romanised form in the way in which they phonetically describe word sounds (Kess & Miyamoto, 2000). There are

currently at least two main systems of romaji; the Hepburn[5] system (1867) and the Kunrei-shik[6] (訓令式) system (1937), which both simultaneously used in Japan (Hannas, 1997).

Each system can represent word sounds in phonologically different ways. Unlike regular Chinese pinyin, Romanised Japanese words do not always match exactly with grapheme-phoneme correspondences. For example, the hiragana character [ふ] may be written as [fu] versus [hu], which could result in Mt Fuji being mis-read as 'Mt Huji' (Okuyama, 2013). The mixing of two different romaji systems could lead to a confusion and potentially be hazardous, negatively affecting kanji word pronunciation when read phonetically. For example, the word 'arigatoo' ('*thanks*'), could be pronounced wrongly as "*arigatwo*" (Kess & Miyamoto, 2000). In addition, romaji can also be similarly be problematic when being used to transcribe foreign loan words, which often have a large 'orthographic gap', resulting in a source of confusion for learners (e.g., computer vs kompyuutaa) (Kess & Miyamoto, 2000; p.114).

However, paradoxically, despite grapheme-phoneme inconsistencies in using romaji, there is exploratory evidence that romaji can have a facilitative effect on L2 Japanese word pronunciation. Tamaoka (n.d.) as cited in Kess & Miyamoto 2000, found that learners of Japanese using romaji were rated as sounding more 'nativelike' as rated by Japanese native speakers, and that using romaji helped learners reduce their error rate in word pronunciation by as much as 300%. Tamaoka's study is one of the only studies which has studied L2 Japanese comprehensibility using romaji, so we must be cautious in interpreting the results until more robust supporting empirical evidence is uncovered in L2 spoken Japanese.

Because using romaji can directly delay the introduction of kana and kanji, there is considerable debate in amongst theorists and practitioners surrounding whether to delay the introduction of authentic scripts (Hatasa, 2002). Miura (1987) asserts that romaji delays the acquisition of L2 Japanese word acquisition as learners may be reluctant to make the switch to authentic texts after learning with romaji. Indeed, many Japanese instructors see romaji as a hindrance to L2 Japanese learning, arguing that learners should start learning Japanese with authentic texts such as kana immediately from the outset (Rose, 2017). Reliance on romaji has also been criticised for lengthening the amount of time required for kana to become an automatic process in reading and writing (Rose, 2017). Koda (1992) agrees, pointing out that

[5] The Hepburn form of romanisation was proposed as early 1885, an adopted by missionary James Curtis Hepburn in 1987 (Kess & Miyamoto, 2000)
[6] The Kunrei-Shiki (訓令式), or 'Cabinet Style' was formally recognised by the Cabinet Order as standard version of romaji in Japan in 1937 (and again 1954) (Kess & Miyamoto, 2000)

from a pedagogical perspective, although romaji may initially help in learning Japanese sounds, once authentic texts are introduced, pre-existing models of the language would have to be remapped reconstructed once again, delaying the language process. Thus, Koda (1992), argues that authentic writing systems (e.g. kana and kanji) should be presented *together* with non-authentic systems (e.g. romaji) to help forge stronger links between the sounds and images of Japanese kanji words through increased exposure to native texts.

In light of the controversy surrounding the use of romaji, surprisingly little empirical evidence exists to disprove its perceived pedagogical limitations. A study by Hatasa (2002) stands as one of the first (and currently only) studies in the field which provides supporting evidence for the use of romaji in a real JFL classroom setting. Her study investigated whether or not delaying the introduction of kana over a period of one 16-week semester made any difference on oral-aural (speaking, listening and pronunciation) L2 Japanese abilities. 62 English-speaking participants at an American university were assigned to either an 'lag' Roman-script group, or 'non-lag' syllabic kana-script group (acting as a control). In each group, participants were assigned to either a romaji or kana version of Japanese text. The experimental lag group, used romaji for eight weeks, and were then introduced to hiragana for the second half of the semester. By week 10, the experimental treatment was completed, and all instructions were subsequently conducted in kana. Participants were tested on their L2 conversation ability, as well as for L2 listening, reading and grammar. The results found that delaying the introduction of hiragana syllabary characters by eight weeks had *no* significant facilitative effect on the development of oral-aural abilities. In other words, compared with the non-delay kana group, the delay romaji group did not differ significantly in L2 Japanese speaking, listening and pronunciation. The results also highlight that although the delay group were effectively 46 characters behind the kana group after eight weeks, they *were* able to catch-up from week 10, after which kana (and kanji) were used as the sole method of instruction. Hatasa's (2002) study highlights that compared with the using romaji at the beginner stages of Japanese learning, the early introduction of kana from the outset does *not* result in any differences in L2 learning outcomes. Hatasa (2002) reports that the performance of the delay group may be partly explained due to the fact that syllabic hiragana is very closely related to phonetic romaji. While this may be partially true, due to romaji's inconsistent grapheme-phoneme correspondences, caution should be exercised when making phonological links between romaji and hiragana. A replication of Hatasa's (2002) study delaying the onset of *kanji* (rather than syllabic kana) presents an area for future research. From a cognitive perspective, this thesis hopes to contribute further to the findings of Hatasa (2002) by identifying if Japanese kanji words can be learned more efficiently using romaji (compared with kana).

Perhaps only second to Hatasa's (2002) study, is Okuyama's (2007) experimental research which set-out to find empirical evidence for using romaji as an effective pedagogical tool to help improve the efficiency of L2 Japanese kana word learning. Underpinned by Katz & Frost's, (1992) Orthographic Depth Hypothesis - where transparent scripts are easier to read than opaque ones - Okuyama's (2007) study filled a research gap in attempting to provide a link between romaji and alphabetic background learners familiarity with their own L1 script compared with learning in hiragana, a non-alphabetic syllabic script. In his study, 61 English-speaking students from two American universities were divided into either a romaji or hiragana learning groups. Using a specialist computer assisted learning (CALL) programme, participants learnt 40 novel Japanese nouns (kanji compound words) over four lessons. During the learning task, word illustrations and phonological sound glosses were added to aid word recall. The results of the post-test found that romaji did *not* significantly aid beginner-level alphabetic script learners' Japanese kana word gain scores. The significance of this study is that despite romaji and Latin script sharing similar phonetic features, there appears to be no facilitative effect on improving word learning. However, Okuyama's (2007) study found that the use of audio recordings were found to strongly correlate with word recall, regardless of the presence of romaji or not. Thus, for beginner-level JFL learners, with existing phonetic knowledge, having audio assistance appears to help the learning of highly syllabic kana. This provides further support that Japanese hiragana word recognition (decoding) is largely a phonological process. While Okuyama's (2007) results are significant for the field, the nature of kana word learning using CALL does limit the study's results which may not be direct applicability to JFL instructed classroom settings. It should be noted that Okuyama's (2007) study only measured the effect of romaji assistance on kana words (hiragana), rather than for kanji, the more salient writing system in Japanese. Far less is known about using romaji to assist with kanji word learning. This study aims to follow-up Okuyama's study to examine the effects of romaji on Japanese kanji (rather than kana) word learning. Studies have found that Chinese characters are processed more visually (Chen et. al., 2002), thus the use of romaji may not be as effective in learning kanji compared with kana. Furthermore, the addition of sound assistance to learn kanji may also be qualitatively different compared to learning kana.

2.12. Furigana

Just as romaji can help decode unknown kanji (and kana), furigana – small hiragana characters typically written above Japanese kanji - are a type of phonological reading aid primarily used

for clarifying the pronunciations of difficult or low-frequency kanji (Koda, 1992). As Japanese kanji often have multiple (and irregular) *on* and *kun* readings, furigana regularly used in early L1 texts (but rarely for adult material). Similarly, for L2 learners, furigana are used as a pedagogical tool to access the correct phonological information of kanji words (Kirwan, 2003; Okuyama, 2007). Since furigana use fully syllabic hiragana characters, unlike romaji, phonological decoding of kanji using furigana is precise. In terms of using furigana in L2 contexts, furigana remains equally divisive among teachers and learners.

Generally, there are many benefits for L2 learners using furigana. Once the hiragana script is mastered – which can be learned in a little as a month (Otake, 1987) - furigana can be used independently as a phonological device to greatly facilitate L2 reading comprehension of complex texts which contain a high number of low-frequency kanji or 'unfamiliar characters' (Koda, 1992; p.503), increasing their exposure to kanji and kanji words. From a cognitive perspective, using furigana can help develop more fluent and automatic processing of kanji word recognition and help to free up valuable cognitive resources (see LaBerge & Samuels, 1974). In addition, furigana can also be used in L2 JFL settings by teachers as a way to introduce more complex characters to learners early on outside of the syllabus, potentially increasing learner self-efficacy and attitudes toward learning kanji (Kirwan, 2005). Furthermore, compared to learners who delay authentic scripts, the introduction of kanji with furigana has been found to reduce emotional shock for learners when introducing kanji Japanese language learners (Kirwan (2005). Moreover, there may be benefits for learners' self-efficacy using furigana as they may perceive greater progress in the language using more authentic scripts, compared with using a Romanised script which could be negatively perceived as childish by learners.

In the face of furigana's perceived benefits, like romaji, the use of furigana is also thought of as a crutch to learning. There are often highly divergent negative perceptions of furigana as being overused or over-relied upon in L2 Japanese learning (Kirwan, 2005). To illustrate the bias toward the perceived negative effects of furigana on a popular L2 Japanese kanji learning website called WaniKani[7], claims that using furigana will negatively impact L2 kanji learning.

> *"..[using furigana] will actively hinder your kanji [and kanji word] learning…[because] when we read Japanese with furigana, your eyes automatically go to it first…we don't teach kanji with furigana….as we don't take kindly to anything that might hold you back*" (WaniKani website)

[7] WaniKani (WK) is a popular online web application which uses spaced-repetition (SRS) and mnemonics to learn Japanese kanji and kanji words. Available at https://www.wanikani.com.

Despite the debate surrounding of using furigana in L2 learning circles as 'holding back' learners (WaniKani), the impact of using furigana has scarcely been investigated empirically in JFL settings. Perhaps, the only empirical evidence within JFL instructional settings comes from Kirwan's (2003) unpublished doctoral thesis. In his study, a series of mini-studies were carried out with L2 Japanese language learners at an Australian university over 14-weeks. The results of the experiments were mixed. In one of Kirwan's (2003) mini studies, he found for lower-level learners of Japanese being exposed to kanji with furigana facilitated implicit learning of unknown kanji (beyond those taught in class). The participants were given supplementary textbooks with furigana written above kanji words, but were *not* explicitly taught any kanji. At the end of the semester, the students were tested on the completely unknown kanji words, scoring on average 7.4% of kanji word pronunciations. In other words, implicitly, students were able to recognise more than 1 in 20 unknown kanji words implicitly when furigana was present. This provides some weak, but exploratory evidence that the use of furigana may be a useful tool for implicit kanji learning. However, as Kirwan (2003) acknowledges himself, it is hard to say whether or not this incidental learning took place outside of the classroom or not (Kirwan, 2003; p.314). The study also lacked a sufficient control group to make direct comparisons between learners. Thus, remains an area to be fully explored in in future research.

In a subsequent follow-up, Kirwan (2003) investigated the role of furigana and explicit kanji instruction with 31 intermediate-level learners. Participants were given two-sets of reading comprehension passages. One set contained a passage with furigana written above the kanji words through the text, whereas the other only contained the reading sample with kanji words presented without furigana. To aid in semantic and phonological kanji word comprehension, each of the reading passages were accompanied with a vocabulary word list which provided the English meanings of the kanji. The kanji word lists which accompanied the non-furigana reading task were provided with furigana, whereas the furigana condition reading condition only provided the kanji words themselves (without furigana) along with the English meanings The results found that Japanese kanji reading and English meaning test scores were higher in the non-furigana passages but these differences were *not* significant. The results are contrary to what we might expect based on the implicit learning gains from furigana in the previous study. It should be noted that these were intermediate level learners of Japanese and that perhaps they learn kanji words more effectively when presented a list of vocabulary items with their pronunciations and meanings separately rather than given the phonological

pronunciations in-text using furigana (Kirwan, 2005). Therefore, based on these preliminary results, there is some evidence to suggest that furigana may *not* be the most effective way for intermediate-level learners to learn new words compared to beginner-level learners who performed better in the implicit learning of kanji with furigana. The results may suggest that for advanced learners, furigana may be a distractor rather than a pedagogical aid. Kirwan's (2005) study would benefit from replication with lower-level learners to provide further support for this hypothesis. In sum, evidence to support kanji word learning using furigana is weak, and requires further investigation. Results from Kirwan's 2005 interventions paint a mixed picture of the effectiveness of furigana in learning Japanese kanji words. Therefore, there is scope for this thesis to explore effect of furigana and kanji word learning across different L2 proficiencies.

2.2. Japanese L2 kanji word processing

Lower-level L2 word recognition processing involves 'extracting linguistic information directly from print' (Koda, 2007; p.4), is widely considered by the field as an important factor in reading comprehension (Grabe, 2009; Koda, 2017; Walter, 2008). Simply defined, L2 word recognition (e.g. decoding) is the ability of a reader to recognise printed letters as something meaningful (Chikamatsu, 2006). For Japanese jukugo (kanji compound) words, on-yomi variations typically do not contain kana characters between them making the decoding process more difficult (Kayamoto, Yamada & Takashima, 1998; Yamada, 1992). This is exacerbated by the opaque grapheme-phonology correspondences contained in the graphemes of on-yomi kanji compounds (Koda, 2017). For example, the first sound in the following on-yomi kanji compounds all contain [ka] 可能、家庭、火事、科目, which makes it difficult for readers to create strong phonological representations within their mental schemata (Koda, 2017).

In L2 Japanese learning, Koda (2017) who defines Japanese kanji recognition as 'decoding', argues that L2 kanji recognition is highly correlated with overall Japanese reading comprehension and stresses the importance of visual skills in extracting phonological (and morphological) information from Japanese kanji words. Logically, unlike the fully transparent syllabic kana characters, if a learner is unable to decode the meanings and pronunciations of highly opaque kanji, then reading comprehension would be severely affected. To illustrate this point, Japanese grapheme-to-phonology relationships are complex due to having both on-yomi (Chinese reading) and kun-yomi (Japanese reading) kanji pronunciations, each requiring

learners to build independent grapheme-phonology correspondences, making the L2 decoding process highly opaque (Koda, 2017). Adding further complexity to L2 Japanese script decoding, kanji characters often embed incongruent or irregular semantic and phonological information (Mori, 2014). Even for natives, decoding a kanji word's phonological representation from its written form, is highly complex (Coumas, 1989; Taylor & Taylor, 1995; Yamada, 1992). Native readers are able to simultaneously switch between syllabic kana and logographic kanji fluidly and effortlessly (Rose, 2017). From a theoretical standpoint, the Orthographic Depth Hypothesis (see Katz and Frost, 1992) asserts that compared to more shallow or transparent orthographic forms deeper or more opaque writing systems generate more errors and require longer cognitive processing times for learners. Thus, for alphabetic learners of Japanese without prior knowledge of Chinese logographic characters, the decoding process becomes a much more difficult cognitive challenge.

There has been much theoretical debate surrounding whether Japanese kanji are activated through *visually* or *phonologically* through sounds (see Chen et al 2007; Mori, 2014). Japanese morphographic kanji contain both semantic and phonological components, and understanding kanji word activation is one of the most researched areas of L1 Japanese word processing (Kondo-Brown, 2006). From a theoretical point of view, the dual route framework is useful theoretical tool understand how kanji are activated at the lexical or phonological level. The dual route theory explains that words can be decoded either visually or phonologically, yet evidence to support when these activations occur is less clear. In L1 kanji decoding, Matsunaga, (1995) found supporting evidence that native speakers of Japanese process kanji not only through their semantic orthographical forms, but also through phonology. A study by Horodeck (1987), found that kanji errors by natives in university entrance examinations were found to be largely homophonic, suggesting the importance of phonology in kanji decoding. Chikamatsu (2005; p.86) also found in L1 and L2 kanji word production, phonological errors formed the largest type of error. Thus, for learning logographic kanji, evidence suggests that visual processing is very important.

One study which investigated how L2 kanji are decoded was Koda's (1992) research which assessed 58 alphabetic JFL students in the United States. Participants were tested on a range of Japanese reading comprehension tasks. Koda's study found strong a strong relationship between efficient lower-level word decoding and L2 Japanese reading comprehension, arguing that for learners without knowledge of logographic writing systems, are more reliant on *visual* features as lower-level processing. The results highlight that early learners of Japanese from alphabetic script backgrounds show a preference for *visual* decoding of the characters due to

insufficient lower-level kanji character decoding skills, which uses valuable resources for higher level-processes such as lexical inferencing (Koda, 1992). A later study by Hagiwara (2016) challenged earlier assumptions that alphabetic learners use phonological information to decode Japanese words, discovering that intermediate level-learners of Japanese were able to learn novel kanji words without the mediation of kanji word pronunciation. In her study, she found there were *no* significant differences, even for learners were did *not* receive any audio stimulus.

Tanaka (2015) investigated the effects of *sound* and L2 kanji word processing of learners from Chinese Mandarin logographic backgrounds. Her results found that Chinese do *not* activate phonological representations of kanji when reading phonological similar cognates, and these results were not significant. In other words, even though Chinese speakers their phonological knowledge of hànzi does not seem to result any phonological advantage in L2 kanji decoding, and may even lead to inhibitory effects (eg H. Chiu, 2003; Y. Chiu, 2006; Hayakaya, 2010; Hayakawa & Tamaoka, 2012). The significance of Tanaka's (2015) findings for this thesis is that although L1 Chinese learners may have a semantic script advantage, their phonological knowledge of hànzi may *not* facilitate kanji learning. In fact, it may make the kanji learning task actually more cumbersome. Thus, the implications for this study is that the presence of audio stimulus during the kanji word learning may *not* necessarily enhance kanji learning for logographic script learners.

2.3. Cross-linguistic transfer

Nation, (2001) asserts that 'for learners whose first language is not related to the second language, the learning burden will be heavy' (Nation, 2001; p.24). This view is echoed by Emmerson (2011), who asserts that learning logographic scripts present special challenges for learners whose first language is alphabetic. These differences (or similarities) in L2 word processing can be described as cross-linguistic influence or transfer. Odlin (1989), defines cross-linguistic influence as being 'the influence resulting in similarities and differences between the target language and any other language that was previously (and perhaps imperfectly) acquired' (Odlin, 1989; p.27). In other words, features of the first language can creep into aspects of L2 language production and word recognition (including semantic or phonological decoding (Hayashi, 2012). However, evidence from cross-linguistic transfer has largely been carried out in ESL or Indo-European language contexts, so it is difficult to make generalisations to all languages. Far less is known about L1 transfer in morphographic scripts

such as Japanese (Chikamatsu, 2006). This study aims to fill this research gap by recruiting L2 Japanese learners from logographic script backgrounds.

When orthographic knowledge is shared between L1 and L2, lower-level processing skills (e.g., decoding) which in turn, can enhance L2 reading performance (Koda, 1996; p.455). However, differences in L1-L2 orthographic distance can impact lower-level processing skills as the cognitive load increases (Akamatsu, 1999; Koda, 1997). The lexical knowledge gap between L1 and L2 reading means that L2 readers often lack the automaticity in lexical processing, which drains resources requires for higher-level processing (e.g. word inferencing) (Chikamatsu, 2006; Stanovich, 1980). Koda (1996) argues that these differences occur from cross-linguistic differences, and can therefore affect phonological decoding. The challenge is exacerbated when attempting to decode L2 words whose writing systems are typologically distant (Woore, 2022). Koda (2005) explains these differences as the orthographic distance effect. The universal phonological principal explains that word activation is an automatic process which happens directly at the word level.

Chikamatsu (2006) for example, was one of the first researchers to address the research gap in L1-L2 linguistic transfer among Japanese L2 learners from alphabetic script learners at different proficiency levels. She, examined the transfer of phonological encoding of familiar and unfamiliar syllabic hiragana and katakana words to see if orthographic influence diminishes as proficiency increases. The results found learners' strategy use and mechanisms in memory and L2 retrieval change as proficiency improves. Chikamatsu (1996) conducted a similar study to explore the effect of divergent L1 background learners of Japanese kana processing. The results found that Chinese learners relied more on visual information compared with English speakers who relied more on their phonological forms. The implications of Chikamatsu's study is that there is some empirical evidence suggesting differences in L1 script characteristics may transfer to L2 word recognition (Chikamatsu, 1996). A limitation of this study was that Chikamatsu only used syllabic kana for the study rather than morphographic kanji. The relevance of these finding for this thesis is that it is reasonable to expect some L1 transfer, particularly for lower-level learners. This transfer may be seen in alphabet to romaji or even romaji to kana which share phonetic similarities. As proficiency improves, the effect of script familiarity may influence cognitive processing preferences toward more authentic scripts such as kana.

2.4. L2 Japanese kanji word flashcard learning

Flashcard learning is a useful way for L2 learners to build-up their knowledge of vocabulary with repeated exposure, and is widely considered a valid pedagogical tool in L2 word learning (Everson, 2011). In optimising flashcard learning, theorists stress the importance of rote learning (Nesbit, 2009), spaced-repetition (Nation, 2001), and retrieval practice (see Webb, 2007b; Baddeley, 1990). Elgort, (2010) asserts that flashcards are one of the most efficient methods of intentional vocabulary study. This view is shared by Nation (2022) who argues that continued spaced-repetition is most effective way to make sure a learner encounters an item which they have previously seen. There is some evidence that compared with L2 incidental learning conditions, intentional learning using flashcards, results in *higher* vocabulary recall gains (Hulstijn, 2003). This may be because when learners use flashcards, they deliberately make associative links, reinforcing their memory trace. Nation (1980) claims that using bilingual word pair flashcards, it is possible to intentionally learn between 30 and 100 new words per hour, whereas Webb (2007b), argues that between 10 and 20 repetitions are required for meaningful learning to take place.

Indeed, for Japanese learners, learning thousands of logographic characters through rote memorisation is a popular method by Japanese L2, both for self-study, and in JFL pedagogical settings (Shimizu & Green, 2002). However, in recent years, Nakata's (2011) review of the field, revealed that computerised programmes have largely replaced manual paper-based flashcard learning techniques. In online language learning apps, vocabulary is the most popular area for L2 learners (Huang, 2020). Ashcroft et. al., (2017) carried out a study using the online learning tool Quizlet, verses a paper flashcards. Their results found that compared with traditional flashcard learning, online flashcard apps were *more* effective in vocabulary gains for beginner-level learners in Japanese EFL contexts. This may be due to more interactive online learning task compared with paper-based learning. Although there is evidence for the benefits of flashcard learning, there are a dearth of L2 studies which have examined flashcard learning in Japanese foreign language contexts. This study hopes to provide some evidence for learning kanji using flashcards through word repetitions, supplemented with audio pronunciations, participants should be able to effectively recall and build awareness kanji orthographic forms in this experimental study.

2.5. Memory and L2 kanji word learning

Memory is important in L2 learning (particularly in reading comprehension) as it helps to interpret incoming information from text (or aurally), and aids in learning novel phonological forms to in acquiring new words (Baddeley, Gathercole & Papagno, 1998). From a theoretical perspective, memory is divided into long-term and short-term memory, where short-term memory, or working memory, is seen as having limited capacity. In this limited short-term store system, Baddeley & Hitch, (1974) theorised the that working memory has a large phonological component known as the phonological loop which (briefly) holds aural input for around 2 seconds (Baddeley, 2012). In addition to the phonological loop, an area used to manipulate visual and spatial lexical information was coined the visuo-spatial sketchpad (Baddeley & Hitch, (1974). In Japanese learning, visual memory is particularly important for learning highly visual and complex kanji characters. Koyama et al., (2008) carried out a cognitive study with L1 natives and found that visual long-term memory was a *strong* predictor of kanji reading comprehension. Researchers have been interested in understanding how differences in L2 lexical processing affects working memory and L2 kanji word learning. For example, L2 learners who with script backgrounds with close orthographic script similarities, the expected cognitive demand would be expected to be reduced. For example, Tamaoka (1997) found that Chinese background learners read Japanese kanji compounds *faster* than English words. The relevance to this study is that, compared with alphabetic script learners, Chinese learners may free-up more valuable working memory and direct their attention to the novel kanji phonological forms. If this is true, then we should see logographic script learners perform better in the kanji pronunciation recall task, regardless of which phonological aid group they are randomised to.

2.6. Self-regulation and L2 kanji word learning

Within the field of language learner strategy (LLS) research, the introduction of Dörnyei's (2005) self-regulation theory was highly influential. Born out of the field of educational psychology, Dörnyei (2005) proposed a theoretical shift by conceptualising language learner strategies underpinned by the concept of self-regulation. Self-regulation was seen as a way to broaden the scope of strategy research (Rose, 2017). However, self-regulation was also highly controversial, and many scholars saw self-regulation as a direct replacement of, and end to strategy research itself (Rose, 2012). Self-regulation sees as a shift away from 'thoughts' and 'actions' of learner behaviours, to more of the 'glue and the engine' of L2 strategic learning (Weinstein et al., 2011, p.47). A limitation of examining learner behaviours through the lens of self-regulation is that it only identifies the underlying self-regulatory capacity

for strategic behaviour, rather than understanding strategy use itself (Dörnyei, 2005). Some scholars fully embraced self-regulation (e.g., Mizumoto & Takeuchi, 2012), while others saw self-regulation as an 'outsider' construct and not theoretically valid for assessing strategy use (Ellis, 1994, Rose et. al., 2018). However, despite its explosive arrival within the field of L2 strategy research, many studies have continued with the status quo, failing to fully adopt self-regulation (Rose et. al, 2018). While some scholars have embraced self-regulation in LLS research (e.g., Mizumoto, 2012; Mizumoto & Takeuchi, 2012; Ranalli,2012), others argue that self-regulation is more of an add-on construct and not valid in attempting to measure complex and multifaceted strategic behaviours (Gao, 2007; Grenfell & Macaro, 2007; Gu, 2012). Dörnyei's (2005) self-regulation model, with its roots in educational psychology, is a motivational control taxonomy based motivational control strategies (Dörnyei, (2001). Dörnyei (2005) explains that self-regulation is a 'multidimensional construct, including cognitive, metacognitive, motivational, behavioural, and environmental processes that learners can apply to enhance academic achievement' (p.101). Under the framework of self-regulation, we can examine what kind of strategies students use through a motivational control taxonomy, defined by Dörnyei as; *commitment control*, *metacognitive control, satiation control, emotional control*, and *environmental control* (Dörnyei, 2005; p.113). Thus, examining learner behaviour underpinned by self-regulation, we can understand the underlying drivers of strategic behaviour itself. Although self-regulation it is only useful for examining motivational behaviour itself which is necessary in language learning (Graham, 2022), it does *not* explain *how* successful learner behaviour will be, as the relationship between motivation and success is not linear (Csizér and Dörnyei, 2005; p.20).

Self-regulated learners have been described as resourceful learners who are able to assertively navigate obstacles and take control of their own learning (Zimmerman, 1990; p.4). In other words, if they are faced with difficulties, they are able to take the necessary steps and find ways to overcome these challenges. In terms of the highly complex and cognitively demanding processes such as Japanese kanji learning task, one's ability to cope and manage their own learning is particularly pertinent, as failing to overcome these major hurdles may lead to learner frustrations in language progression (Rose & Harbon, 2013). One would argue that lack of adequate self-regulation in L2 Japanese learning, could lead to a complete breakdown in learning. Indeed, evidence has found that learners of Japanese often set unrealistic or unachievable goals due to underestimating the cognitive challenges of the Japanese writing system (Rose, 2017). As a result, failing to see a sense of achievement in learning Japanese, can take an immense physical toll on L2 learners (Rose, 2017). Scholars point out that the learning of the Japanese kanji requires a conscious and systematic approach in monitoring

one's own progress and learning outcomes (Kubota & Toyoda, 2001; Mori, 2002). Therefore, it could be argued that a failure in one's self-regulatory behaviour would be catastrophic for the long-term acquisition of Japanese. To overcome the enormous cognitive barriers in learning Japanese kanji, a learners' ability to control and self-regulate one's own learning is absolutely essential in achieving successful learning outcomes in L2 Japanese (Rose, 2017). In other words, in order to control and manage the kanji learning task, Japanese learners must develop sufficient self-regulatory capacity.

As far as validated instruments which incorporate self-regulation in Japanese L2 learning, most self-regulatory instruments have been heavily used (and tested) in English second language contexts. For example, Tseng et. al's (2006) Self-Regulatory Capacity for learning Vocabulary (SRCVoc) aimed at measuring self-regulatory capacity in Taiwanese ESL learners. The SRCVoc was a highly influential tool, and one of the first tools which was specifically situated to an L2 learning task (i.e., L2 vocabulary learning) underpinned self-regulation. Similarly, Mizumoto and Takeuchi (2012) adapted and tested the SRCVoc, and applied it to Japanese EFL contexts, where the instrument was found to be reliable with a Cronbach's alpha of α = 0.84. However, Mizumoto and Takeuchi (2012) also found 'definitional fuzziness' in *environmental control* sub-components when retesting the SRCVoc in EFL contexts in Japan. 'Definitional fuzziness' had plagued the development of learner strategy research (Rose et al, 2018), largely as a result of using taxonomies and Likert scales to measure attitudinal behaviours which while has brought widespread acclaim, but has also drawn sharp criticism from several scholars in the field (e.g. Dörnyei, 2005; Woodrow, 2005), who argue for the use of more qualitative studies in self-regulation research.

One study which answered the calls from the field was Rose and Harbon (2013), who used a qualitative approach to investigate how self-regulation changes in L2 kanji learning as proficiency increases with exchange students in a year-long programme in Japan. Although alphabetic Japanese learners face the most difficulty in when first starting out in Japanese, Rose and Harbon (2013) found that as learners advance in the language, they suffer from a breakdown of self-regulatory control due to 'frustrations caused by a lack of progress in learning, or due to self-criticism over an inability to reach goals' (p.96).

In terms of the development of specific self-regulation instruments in Japanese L2 kanji learning, one of the first scholars to examine kanji learning strategies through the theoretical concept of self-regulation was Rose (2010). Rose used the SRCVoc as a template for identifying and categorising kanji learning strategies in qualitative research, which he coined

the 'SRCKanji'. Although, some argue that questionnaire instruments such as the SRCVoc still suffer from the same definitional fuzziness with hampered SILL, in-roads are being made to validate such tools (see Ardasheva and Tretter, 2013). In addition, in an attempt to bridge self-regulation and writing strategy research, Teng & Zhang's (2016) Writing Strategies for Self-Regulated Learning (SRL) is an example of high-quality research which is moving the field forward with context-specific, and well tested and validated self-regulatory control measurement instruments. In a similar manner, this study sees an opportunity to quantitatively validate the SRCKanji for the first time, with a medium-sized sample of L2 Japanese learners.

2.7. Summary and relevance to the study

Based upon the limited empirical evidence from our literature review, it is hypothesised that for beginner-level learners, one might expect that Roman-script based romaji would enable alphabetic script learners an advantage at the beginner level learner stage due to their reliance on their L1. In a similar way, as *pinying* is used in Mandarin-Chinese as a pedagogical tool, romaji may enable beginner Chinese-background learners to learner Japanese vocabulary more effectively. Furigana on the other hand should be as just as effective for each L1 group (alphabetic or logographic) in decoding and learning new words, and possibly more effective for more advanced learners as they are more familiar with more authentic Japanese kana.

To the best of the author's knowledge, *no* prior experimental research has investigated the effects of romaji and furigana reading aids in an experimental comparison of L2 Japanese kanji word learning. We found that based upon the academic literature, the role of both L2 proficiency development, L1 learner background is vastly unknown in L2 Japanese kanji learning. Through the learning of novel kanji words using different phonological aids, this study may help us understand whether or not differences in proficiency and/or differences of L1 writing systems result in more efficient kanji word learning.

From our review of the field in terms of self-regulatory control in L2 kanji learning, we might expect to find higher-level Japanese learners suffering from a loss of self-regulation as they frustrations with perceived language development. While Rose and Harbon (2013) were one of the first studies to explore the relationship between self-regulation and different L2 proficiencies, the relationship between self-regulation and non-alphabetic script learners of Japanese has scarcely been investigated. This study aims to explore this research gap by investigating self-regulation across different L1 writing system backgrounds.

3. SAMPLE

3.1. The Sample

In total, 146 beginner and intermediate-level adult L2 Japanese speaking participants were recruited online for this study. All participants were proficiency profiled, and 38 participants were removed from the sample as their proficiency scores fell outside the predetermined 'beginner' or 'intermediate' cut-offs[8] scores (TTBJ). Eight participants were excluded either because they had completed the proficiency test but not the online experiment (or vice versa). In addition, two participants requested to pull-out of the study and were removed. Of the remaining 98 participants, it was found that 23 had prior knowledge of more than 50% of the 16 target words. The researcher felt the cognitive burden of learning the pronunciation of less than half of the target kanji words was significantly reduced, and the inclusion of these participants may have distorted the post and delayed post-test scores, thus were excluded from the study. Table 3.1 shows a breakdown of the resulting *final* sample of N=75, (36 beginners, and 39 intermediate learners).

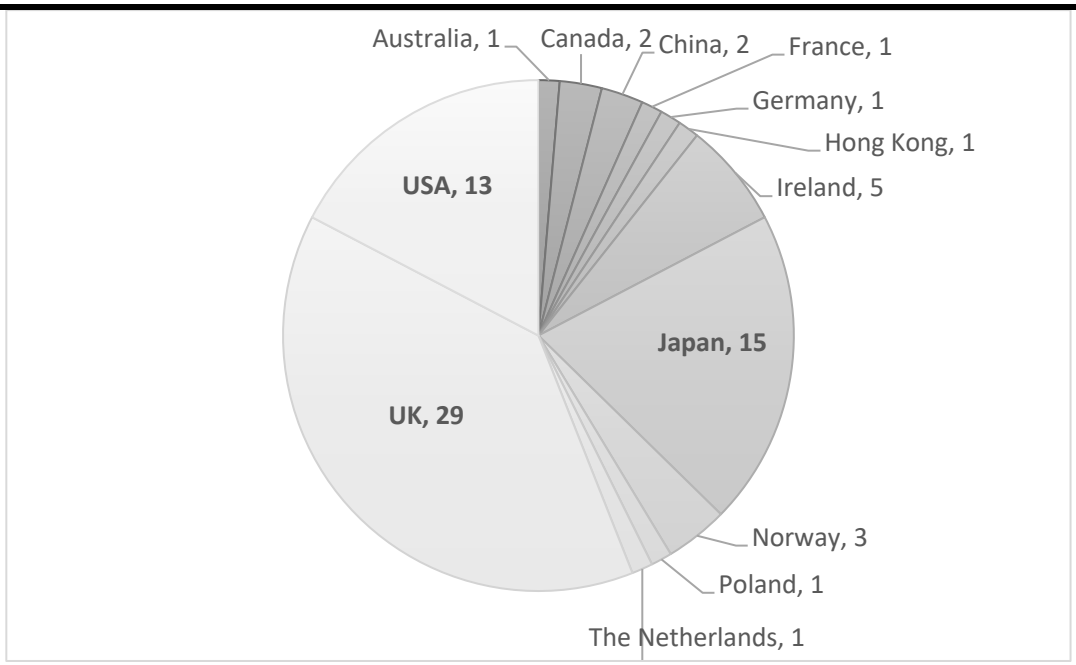
Table 3.1 Breakdown of final sample (N=75) by proficiency and treatment group

L1 background	L2 proficiency	Treatment group	Number of participants
Alphabetic	Beginner	Romaji	11
		Furigana	18
	Intermediate	Romaji	17
		Furigana	19
Logographic	Beginner	Romaji	5
		Furigana	2
	Intermediate	Romaji	2
		Furigana	1
TOTAL			75

[8] A score on the University of Tsukuba's SPOT90 test defines a score between 31-55 as 'beginner', and 56-80 for 'intermediate'. The SPOT90 forms part of a larger TTBJ test 'battery' of Japanese L2 proficiency. Available at https://ttbj.cegloc.tsukuba.ac.jp/en/p1.html

14 countries were represented in the final sample, with the UK, Japan and the United States comprising of 76% of the total sample (see Figure 3.1). In terms L1 orthographic background, alphabetic script learners (N=65) accounted for 86.6% of the final sample, with 13.3% (N=10) logographic script learners. Appendix B lists a full breakdown of participants by geographical location, level of L2 proficiency, and L1 script background.

Figure 3.1. Chart showing participant country of residence of the final sample (N=75)



The mean age of the sample (N=75) was M=26.53, (SD=6.42, SE=0.74), with a slight bias toward the number of females of who took part in the study (Female, N=40, Male, N=31, Prefer not to say, N=4). The mean number of years of Japanese language study for the sample was M=6.03 (SD=5.8, SE=0.68), with participants mainly reporting self-study *or* formal Japanese language programmes as their main method of instruction. The mean proficiency score of the final sample was M=55.76, (SD=13.34, SE=1.54) placing the overall sample at the upper-cusp of 'beginner'-level (TTBJ). Within the sample, the long length of mean of Japanese L2 study (M=6.03), together with the low proficiency mean score of 'beginner' (M=55.76), is consistent with Everson's (2011) view, who highlights that due to the complexities of the script, a longer amount of time is needed to acquire proficiency in Japanese compared to European languages. Table 3.2 shows descriptive statistics by proficiency level, age and mean number of years of Japanese language study.

Table 3.2. Descriptive statistics of the sample (N=75)

	Number (N)	Minimum (min)	Maximum (max)	Mean (M)	Standard deviation (SD)	Standard error (SE)
Proficiency score	75	30	80	55.76	13.341	1.542
Beginner	36	30	55	44.28	7.700	1.283
Intermediate	39	56	80	66.36	7.180	1.150
Age	-	18	45	26.53	6.419	0.741
Japanese study (years)	-	0	26	6.03	5.882	0.679

Participants were recruited for the study mainly from universities and Japanese language schools via a combination of convenience and snowball sampling. The justification of convenience sampling was due to limited resources and time-constraints. Furthermore, the number of L2 Japanese learners in the UK is low, with just 1,332 students taking the JLPT in England and Scotland in 2019 (refer to Table 1.1). As Bryman (2022) states 'convenience sampling is appropriate for studies that examine hard to reach populations' (p.83). Thus, convenience (and snowball sampling) were used to draw L2 Japanese learners from outside of the UK. The study also aimed to recruit a diverse range of Japanese learners outside of college and university-level learners, answering calls from Mori (2017) to include Japanese learners from much more diverse contexts. Out of the final sample, 58.7% were recruited outside of Japanese programmes at universities.

Table 3.3 shows list of universities who formally assisted with the research project comprising the remainder of the recruited sample (41.3%). To recruit students on Japanese language programmes, several hundred dual-language 'letters of invitation' were emailed to Japanese language schools and university language programmes in the UK, the United States, China, Hong Kong, Taiwan, and Japan, along with a detailed recruitment poster (Appendix W), and necessary CUREC ethics approval letter (Appendix T). A further limitation of convenience sampling was that L2 learning background (e.g., formal language school versus self-study) could *not* be fully controlled for. Therefore, this study must acknowledge a limitation of using convenience and snowball sampling which carried the risk of 'non-probability' (Bryman, 2021). In other words, bias arising as a result of not selecting participants randomly from the target population. Thus, we must be cautious about generalising our results to the broader population of Japanese learners which may not be fully 'representative of the sample' (Bryman, 2021;

p.146). Outside of formal instructed settings, the researcher's personal network of former work colleagues from the Japan Exchange and Teaching (JET) Programme[9] was a highly successful snowball sampling recruitment method. A recruitment post[10] within JET alumni forum on social news outlet Reddit (https://reddit.com) with more than 22,000 registered members, received over 5,200 views over the span of five days (see Appendix F).

Table 3.3 List of universities and institutions where participants were recruited who were enrolled on formal Japanese language programmes

Institution	Location	Type of language programme	Number of participants recruited
University of California San Diego	California, USA	Japanese language programmes	5
Slippery Rock University of Pennsylvania	Pennsylvania, USA	Undergraduate Japanese programmes	2
University of Wisconsin-Whitewater	Wisconsin, USA	Undergraduate Japanese language programmes	2
Trinity College Dublin	Dublin, Ireland	Introductory-level Japanese course	2
University of Oxford	Oxford, UK	BA Oriental Studies (Japanese) and MSc/MPhil Japanese studies programmes	5
University of Oxford Language Centre	Oxford, UK	All fast-track and elective Japanese language programmes	4
University of Sheffield	Sheffield, UK	All undergraduate Japanese language programmes	6
School of Oriental and African Studies, University of London	London, UK	Graduate and undergraduate Japanese language programmes	2

[9] The Japan Exchange and Teaching (JET) Programme is a Japanese government initiative which brings graduates from over 50 countries to teach English in primary, middle and senior high schools. https://jetprogramme.org/en/

[10] URL link to Reddit recruitment post available at https://www.reddit.com/r/JETProgramme/comments/14jhl6o/looking_for_participants_for_l2_japanese_study/

Japan University of Economics (日本経済大学)	Fukuoka, Japan	Japanese second language course for exchange students	1
Saga University (佐賀大学)	Saga, Japan	Japanese second language course for exchange students	2
TOTAL			31

In determining the required sample for this study, an a-priory calculation was performed using statistical analysis programme G*Power[11] (see Faul et. at., 2007). G*Power is a powerful tool which allows researchers to calculate the required sample size given predetermined effect sizes, alpha error probabilities, and desired power levels (Faul et. at., 2007). Field (2018) explains that G*Power can be used in a similar fashion to Cohen's (1988) statistical calculation tables. The rationale for predetermining the sample size before collecting data is that it avoids the issue of unethicality in research (see Bacchetti et. al., 2005), and limits the number of underpowered studies in the field, which significantly hamper their ability to make useful and accurate conclusions (Faul et. at., 2007). To calculate the sample for this study, the appropriate statistical analysis were selected in G*Power; *Analysis of Covariance* (*ANCOVA) -> Fixed effects -> main effects and interactions*', -> '*numerator (df)*' set as "1" (2 x treatment groups - 1), and the '*number of groups*' as "4" (2 x treatment groups x 2 proficiency levels x 1 covariate (pre-test scores)). Input parameters were set for detecting a medium effect size (f = 0.25), powered to 80%, with alpha significance level set to p <0.05. These parameters are generally accepted for social science research (Bhandari, 2023). G*Power estimated a resulting sample size of N=128. Thus, 128 was set as the target sample during the recruitment phase of the study.

3.2. L2 Japanese proficiency assessment

The University of Tsukuba's SPOT90 Japanese proficiency test, is a highly effective and reliable proficiency profiler for sorting learners into Japanese language programmes at universities (Ford-Niwa & Kobayashi, 1999; Hatasa & Tohsaku, 1997; Kobayashi, 2016). The SPOT90 is also effective for profiling learners for Japanese SLA research (Kobayashi, 2016; p.186). The test is divided into 3 x 30 blocks of questions and takes about 8-10 minutes to complete online. Cut off scores of between 31-55, and 56-80, places learners as 'beginner'

[11] G*Power is a free application which allows researchers to calculate pre and post statistical analyses for an array of statistical tests. Available at; https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower

and 'intermediate' respectively. Table 3.4 shows some examples of questions used in the test.

Table 3.4 Examples of TTBJ SPOT90 questions

(1)	ニュースを聞い(　)ビックリしました。 [I was surprised when I heard the news]	Answer: ('te') [て]
(2)	東京行きのバスはど(　)ですか。 [Which bus is the one to Tokyo?]	Answer: ('re') [れ]
(3)	私は毎日、新聞(　)読みます。 [I read the newspaper everyday]	Answer: ('wo') [を]

Participants have to listen to (and read) 90 short passages of Japanese text read at near native speed, then have three seconds to select the correct answer from four hiragana multiple-choice responses (TTBJ). As the test progresses, the difficulty increases as the reading passages become more difficult and abstract. The audio sound quality also becomes more difficult to hear as the speech is manipulated (TTBJ)[12]. Fitzpatrick, (2012; p.83), argues that vocabulary should be measured in more spontaneous way as 'vocabulary elicitation process should happen with a minimal contextual inference or scaffolding'. Thus, as there is limited time to retrieve information due to the ephemeral nature of the SPOT90 test design, this type of test taps into learners' procedural knowledge. Figure 3.2 which shows the TTBJ response screen interface, the target sentence, the four multiple choice hiragana responses, and the timer screen.

Figure 3.2 SPOT90 test user interface showing close target item, responses, and timer

[12] For more information about the TTBJ which explains how the test is conducted please visit https://ttbj.cegloc.tsukuba.ac.jp/doc/teacher/doc_en.pdf



After hearing the sentence, participants have three seconds to identify the correct hiragana (in the parentheses) from a choice of four hiragana response items (TTBJ)

As learners from different L1 writing backgrounds were participating in this study, the SPOT90 was piloted with 2 students; one from an alphabetic and one from a logographic script background. The students were matched on proficiency, being both enrolled in the same class in a first-year BA Japanese language programme at the University of Oxford. The results of the pilot found test scores to be comparable, despite the Chinese Mandarin background learner's superior knowledge of Chinese hànzi, scoring 34% (31/90) on the SPOT90 test. The result is consistent with Tamaoka's (2015) findings that existing knowledge of Chinese hànzi is not advantageous in Japanese kanji decoding. As Kobayashi (2016) notes, the SPOT90 requires an understanding and recognition of the Japanese phonological system. Based on the small pilot, the SPOT90 was deemed a suitable instrument to measure L2 proficiency across different L1 backgrounds.

One unexpected side-effect of using the SPOT90 was that it was a poor instrument for profiling lower-level learners, as it may have been too difficult for them. This is because without a fully developed knowledge of hiragana, they were unable to read while listening to near native spoken Japanese, preventing them from decoding the answers written in hiragana within the 3-second time limit. This is in line with Kobayashi, (2016), who found that lower-level learners performed worse in the audio SPOT90 compared to an older paper version with no audio, as lower-proficiency learners have less capacity to perform the dual actions of reading and listening while drawing upon their linguistic knowledge (Kobayashi, 2016; p.188). Although

there are more superior methods to measure Japanese L2 proficiency[13], Eda et. al. (2008) argue that they are time consuming, with the Japanese Skills Test (JSKIT) lasting around 90 minutes. Therefore, the SPOT90 was chosen for its relatively quick assessment, being completed in just 8-10 minutes. Including all of the 121 participants who undertook the SPOT90 in this study, the mean proficiency score was M = 58.77 (SD=17.24, SE=1.57), at the lower cusp of intermediate (see Appendix C).

[13] Eda et. al. (2008) recommends The Japanese Skills Test (JSKIT) as a superior L2 Japanese proficiency test.

4. METHODOLOGY

4.1. Study design

Revisiting our main research questions;

> **RQ1** "***To what extent do different phonological reading aids (romaji and furigana) help in the learning and retention of L2 Japanese kanji word learning",***

This study addresses this research question by measuring the effect of learning kanji word pronunciations under the two different learning conditions at two time points in a quantitative manner. A quasi-experimental design was used, where participants were randomly assigned to one of the two treatment conditions. When designing between-subject experiments, the process of randomisation into groups helps to reduce the effects of unsystematic variation[14] (Field, 2018). Data from existing word knowledge of the 16-target items were controlled for by using a pre-test against post and delayed-post kanji word scores. Any variance between learning groups were then analysed using an Analysis of Covariance (ANCOVA) on Microsoft's SPSS (ver.29).

To address to the second research question in this study

RQ2. ***What is the relationship between self-regulation, L1 background and L2 proficiency in Japanese kanji word learning***

To address the research question, Rose's (2010) 20-point self-regulation instrument SRCKanji was used to measure self-regulation across L1 background and L2 proficiency to explore any significant relationships. It should be noted that the analysis is *not* to find a relationship between self-regulation and learning outcomes, as the relationship between these two variables is not a linear (Dörnyei and Csizer, 2005; p.20). Rather, the aims of this analysis will examine the dynamics between self-regulation and strategic behaviour between L1 backgrounds at the different stages of L2 Japanese learning. As the SRCKanji, derived from Tseng et. al.'s. (2006), the instrument was first checked for internal reliability using Cronbach's alpha across each of the self-regulation coefficients.

[14] Unsystematic variation is known as the 'variation from random factors that exist between the experimental conditions. For example, differences in ability, age, time of day and so forth (Field, 2018).

4.11. Quasi-experimental design

The primary aim of this study is to investigate the role of different phonological reading aids on novel kanji word pronunciation learning. To prevent the risk of bias from external factors and to control the learning environment as much as possible, the online experiment builder Gorilla'[15] was chosen as an appropriate data collection tool (see Anwyl-Irvine et. al., 2020). In the case of this study, Gorilla appropriately controlled for timing of the pre and post-tests, the learning phase, as well as ensuring that each participant was automatically randomised to one of the experimental learning conditions. Data were collected via the Gorilla at five data points; from two surveys, a background survey and self-regulation questionnaire, and from three kanji pronunciation words tests (pre, post and delayed-post). A pilot version of the experiment was tested with two Japanese learners; one beginner and one intermediate learners. 10 pilot kanji words selected from Kondo-Brown's, (2006) study (see Appendix H). The aim of the pilot study was to check if the pronunciation of 10 words were learnable given the 10–12-minute flashcard learning cycle. A secondary aim of the pilot study was to check for ceiling effects in the immediate post-test.

4.12. Data analysis tools

To examine the effect of the two independent predictor variables (romaji and furigana) on post and delayed-post kanji word test scores all participants were randomised into one of the kanji word-learning independent treatment groups. The study used a *within-subject* Analysis of Covariance (ANCOVA), collecting data from *within* each of the participants' respective learning group (romaji or furigana). The two independent variables were each of the reading aid conditions, and the dependent variables were the post and delayed-post test scores. Pre-test pronunciation scores were used as the covariate, acting as a control for existing prior knowledge. The covariate (pre-test score) and the independent variables (reading aid condition) were completely independent, reducing any spurious effects of each treatment group (Field, 2018). Using ANCOVA is a powerful tool to identify if group membership is expected to change (more or less) compared to members in the other group, ensuring that participants have the same baseline starting point Fitzmaurice et. al., (2004). Post-hoc analysis and t-tests were performed to examine any differences between treatment group and learner L1 script background (e.g., alphabetic or logographic).

[15] Gorilla is an online experiment builder and data collection tool (Anwyl-Irvine et. al., 2020), which has been authorised as a tool for collecting data within the Department of Education, University of Oxford. For further details, please visit (https://gorilla.sc)

4.2. Data collection methods

The following section explains what instruments were used, and how they were used to collect the qualitative data in this study. This section will also justify the methodological considerations, and provide rationale for why particular methods and instruments were used instead of others.

4.21. Selecting target vocabulary

In choosing the target kanji words for the study, a small pilot was carried out with the same students used for the SPOT pilot study (one alphabetic and one logographic learner). The aim of the kanji pilot study was two-fold. First it was necessary to assess prior lexical knowledge (particularly for logographic script learners) of kanji pronunciation, and secondly to see how long it would take to learn both the meanings and pronunciations of the 10 kanji target words. The pilot was carried out using flashcards on the online word learning tool Quizlet (https://quizlet.com). The participants in the pilot study were pre-tested to check their existing knowledge of the kanji words (semantic and phonological meanings), and were then instructed to learn the new words on Quizlet. The participants were asked to report how long it took, and reported an estimated mean time of 10-12 minutes. Thus, for the main experiment itself, it was estimated that for the same amount of time, learning phonological forms of 16 novel kanji words would be possible (as the cognitive load learning *just* kanji pronunciations would be significantly reduced).

The pilot study revealed that logographic script background learner had superior semantic knowledge of words despite being a relatively new learner of Japanese, consistent with Chen (2002) who identified that Chinese speakers share as many as 54.5% semantic meanings in Chinese Mandarin. Nevertheless, phonological scores were roughly comparable between both learners, again in-line with Takaoka's (2015) findings. When deciding on the final 16 target items, care was taken to choose not only on-yomi nouns, but also verbs, adverbs, and adjectives as these tend to share less phonological cognates with Mandarin Chinese. Target kanji words were selected from each of Tono et. al.'s (2013) 1,000 – 5,000-word bands. A larger percentage of words were extracted from the upper 3,000 – 5,000 bands as it was reasoned that lower-frequency words would be less likely to be known than more common higher frequency words. This study must acknowledge that a limitation of using word band frequencies assumes that word frequency corresponds the relative amount of exposure of a kanji word (Nation, 2001). So, although using kanji frequencies is a useful tool for selecting

target words, it may not be fully representative of true kanji exposure and cannot be fully controlled for when choosing target items (Kirwan, 2003; p.254). For example the compound 何と [nanto] has a frequency of 2419 compared with 何 [nani], which has the same character but with a much higher word frequency of only 40 (Tono et. al, 2013). Table 4.1 shows a summary of the final 16 target words and the number selected in their corresponding word frequency bands.

Table 4.1 Summary of target 16 words and their corresponding word frequency bands.

Word frequency band	**Number of words**
1000	1
2000	1
3000	4
4000	4
5000	6
TOTAL	**16**

From Tono, Yamazaki, & Maekawa (2013).

The final 16 target items (see Appendix I) were chosen from Tono et. al's. (2013) word frequency dictionary. This frequency dictionary is perhaps the most up-to-date Japanese word frequency resource available for Japanese L2 researchers. The resource draws upon more than 100 million corpora; covering spoken, written, fiction and non-fiction, and modern Japanese texts, organising Japanese kanji words into frequency word-bands ranging from 0 to 5,000.

A limitation of this reference book is that it lacks lower frequency words above the 5,000-word band threshold. An updated version, ranging to 10,000, or 15,000 and above may be prove useful for future L2 Japanese studies. Table 4.2 provides a sample of the final 16 target words with their corresponding word frequencies.

Table 4.2. Sample of target Japanese kanji words used in the experiment by word type and word frequency.

Target kanji word	Reading	Meaning	Type of word	Word frequency (0 – 5000)
感染	かんせん (kansen)	infection	noun	3623
率いる	ひきいる (hikiiru)	to lead	verb	4248
度々	たびたび (tabitabi)	often	adverb	3825
著しい	いちじるしい (ichijirushi)	significant	adjective	2468

Target vocabulary words selected based on their word frequencies in *A Frequency Dictionary of Japanese, Core vocabulary for learners* (Tono, Yamazaki, & Maekawa, 2013).

4.22. Measuring L2 kanji word knowledge

In order to gain a full picture of an L2 learners' vocabulary size, it is necessary to quantify in some manner, an estimate of their lexical knowledge. Nation, (2013), asserts that 'vocabulary size tests are useful for comparing groups of learners' (Nation, 2013; p.522). One such vocabulary size test is Nation's (2013) Vocabulary Size Test (VST); a receptive measure of a learners' total vocabulary tested via known words of the most frequent 5,000-word *families*. A Japanese version of the VST exists called the New Vocabulary Levels Test (NVLT)[16] (see McLean & Kramer, 2015), however, for the purposes of this study, the six-part test, consisting of 150 multiple choice kanji word questions was deemed too long.

Further limitations of the NVLT may arise due to Japanese kanji words being presented as multiple-choice responses. Hence, the NVLT may have favoured participants from L1 logographic backgrounds who may easily be able to make semantic links between the Japanese kanji words and Chinese hànzi. For example, you would expect a Chinese

[16] The Japanese version of Nation's Vocabulary Size Test is available at https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-tests

background script learner to make the link between the kanji word 期間（きかん）with the abstract concept of time, period, or interval in Chinese Mandarin (e.g., "It was a difficult period"). Moreover, we felt that as the NVLT was based on simply translation of the main VST which uses the 2012 British National Corpus and the Corpus of Contemporary American English, that perhaps the NVLT did not represent a true reflection of modern Japanese text as the NVLT is not grounded upon an up-to-date Japanese corpus. Thus, given the lack of a simple, quick and robust instrument to measure Japanese L2 vocabulary, and further given the limited time constraints of the project, a simple *pre-test* was decided upon to measure participants' existing knowledge the 16 target items.

During the online pre-test, participants were given 40 seconds to answer each target item. In terms of test scoring, 1 point was awarded for a fully correct answer, and a half a point for a partially correct response, with a maximum possible score of 32 points (16 for meaning, and 16 for pronunciation). The scores were hand-checked by the researcher to ensure consistent scoring. As the main aim of the study was to measure the effect of phonological reading aids on kanji word pronunciation, only pronunciation scores were used to control for pre-existing phonological knowledge. Finally, a further rationale for using a simple target word pre-test, was that it allowed the researcher to remove participants who scored too highly test (particularly those with high phonological decoding scores).

4.23. Measuring L2 self-regulation

To measure the underlying self-regulatory capacity in L2 Japanese kanji word learning across language backgrounds, a dual-language (Mandarin-English), an online version of Rose's (2010) SRCKanji was administered on Gorilla (Appendix J). Example questions include "*When learning kanji, I have special techniques to achieve my learning goals*.". Translations of the questions were also provided in English and Mandarin Chinese as it was deemed important to collect accurate responses all participants. Data were collected from self-reported responses were measured using a 6-point Likert scale from '*strongly disagree'* to '*strongly agree'.* The 'SRCKanji', coined by Rose, is an adapted version of Tseng et. al's (2006) SRCVoc, which incorporated factors from Bourke's (1996) kanji learning taxonomy. Rose felt that the rationale for using the SRCVoc as a basis was that that the kanji learning task was 'not too dissimilar' from the vocabulary learning task (Rose, 2017; p.12). Tseng et. al.'s (2006) SRCVoc has been found to be a reliable instrument, with each sub-section Cronbach's alpha (α) coefficient score above 0.7, with an overall internal consistency of α = .78 (Tseng et. al., 2006). It was later

validated in Japanese EFL contexts with similar results (Mizumoto & Takeuchi, 2012). Although the SRCKanji was one of the first instruments to incorporate the theory of self-regulation and Japanese L2 kanji learning and L2 Japanese learning contexts (Rose & Harbon, 2013), it has yet to be widely tested or validated. Thus, an internal consistency calculation was performed on SPSS with the full sample (N=98). Overall, the internal consistency for the SRCKanji was high, with Cronbach's alpha α = .78, considered as acceptable in the field (see DeVellis, 1991). However, the sub-component reliability scores were found to be less reliable, with emotional, environmental and satiation control components scoring below Cronbach's alpha of α = .316, (see Table 4.3).

Table 4.3 Cronbach's alpha internal reliability of the SRCKanji (N=98)

Self-regulatory behaviour coefficients	**Cronbach's alpha (α)**	**Number of items**
Commitment control	.649	4
Satiation control	.316	4
Emotional control	.140	4
Environmental control	.203	4
Metacognitive strategies	.689	4
Overall	.774*	20
Commitment control	.649	4
Metacognitive strategies	.689	4
Overall (revised)	**.794***	8

*Cronbach's alpha scores of $\alpha > .7$ are generally accepted as having valid internal consistency (DeVellis, 1991; Field, 2018).

The low internal consistency of the SRCKanji suggest a high probability that the scale items may *not* be accurately measuring the underlying construct (Cronbach,1951). These results are consistent with Mizumoto and Takeuchi (2012), who found in low internal validity of environmental control factors (α = .52) when performing CFA analysis[17] on the SRCVoc. Moreover, this highlights some of the past-issues with 'definitional fuzziness' found in strategy

[17] Confirmatory factor analysis (CFA) is a statistical method used to check the structural validity of a measurement model.

research, where psychometric instruments often lack reliability in measuring the intended construct (Dörnyei, 2005). A limitation of using Likert-type scales to measure highly individual and contextual learner strategy behaviours is in-line with calls for 'richer, more sample-specific data techniques' from qualitative approaches (Woodrow, 2005; p.97). The lack of internal reliability casts doubt over the SRCKanji's general validity. The instrument would benefit from an exploratory factor analysis (EFA), as the reliability of the instrument cannot be interpreted without a quantifiable estimate rate of error (Cronbach,1951). This study is likely to have been the first to empirically test the SRCKanji since Rose (2010) first coined the instrument. Although the overall reliability of the SRCKanji of the instrument is acceptable, due to the unreliable sub-component factors, the instrument must be used with caution. For data analysis in this study, low-scoring sub-components were first removed, and the SRCKanji was *re-checked* again with the remaining sub-components (commitment control and metacognitive strategies) to ensure that the removal of items had not affected the internal structure of the instrument (Field, 2018). The resulting internal reliability of the instrument using was α = .794 (see Table 4.3). Subsequently, the reduced two sub-component SRCKanji was used to measure L2 self-regulation in L2 kanji word learning.

4.24. Materials

Custom flashcards for the 16 target Japanese kanji words were created by using furigana and a modified version of romaji (using romaji as furigana above text). As there no consensus in modern Japanese about which Romanised system should be used (see Appendix J), the Hepburn System was chosen due to its popularity and prevalence (Kess & Miyamoto, 2000). Typically, romaji are used to in early pedagogical texts as a direct replacement for kana or kanji. However, as the aim of the study of this study was to investigate reading aids kanji decoding, romaji letters were used as a phonological cue in a similar way to furigana. Table 4.4 provides an example of flashcards in each condition.

Table 4.4 Example of how novel L2 Japanese vocabulary were presented to participants under each treatment condition

Target kanji word (no phonological aid)	**Word type**	**Kanji word flashcards in each learning condition**	
		Romaji (group A)	**Furigana** (group B)
感染	noun	kansen 感染	かんせん 感染
率いる	verb	hiki 率いる	ひき 率いる
度々	adverb	tabitabi 度々	たびたび 度々
著しい	adjective	ichijiru 著しい	いちじる 著しい

4.3. Experiment procedure

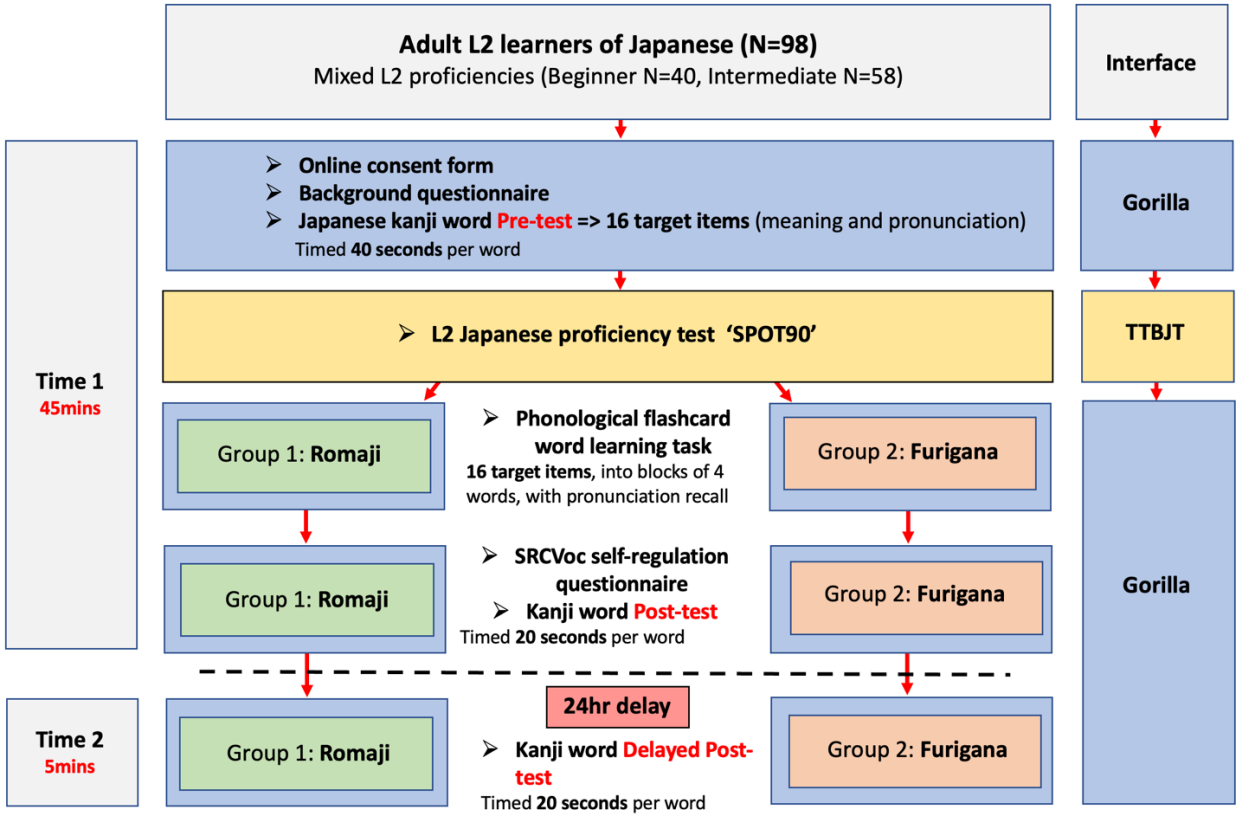
Participants were instructed to join the experiment using the QR code, or URL link which were included on the marketing poster (see Appendix E). The experiment was carried out via two online sessions via Gorilla, each lasting between 40-45 minutes and 5-7 minutes respectively. Figure 4.1, shows a flow diagram of the study, highlighting each stage of the experiment. At the beginning of the online study consent was confirmed via ticking a check-box (Appendix T), followed by a background information survey (Appendix A).

4.31 Pre-test

Next, a 6–7-minute pre-test measured participants' existing lexical knowledge of the 16 target items. Each item had a 40-second time limit, and the order of the target words in each of the

words tests was also fully randomised. Rest-breaks were built-in after each block of 4 words. Breaks were followed by a 'click-to-proceed' button to move onto to the next block of words.

Figure 4.1. Flow diagram of the quantitative design of the study



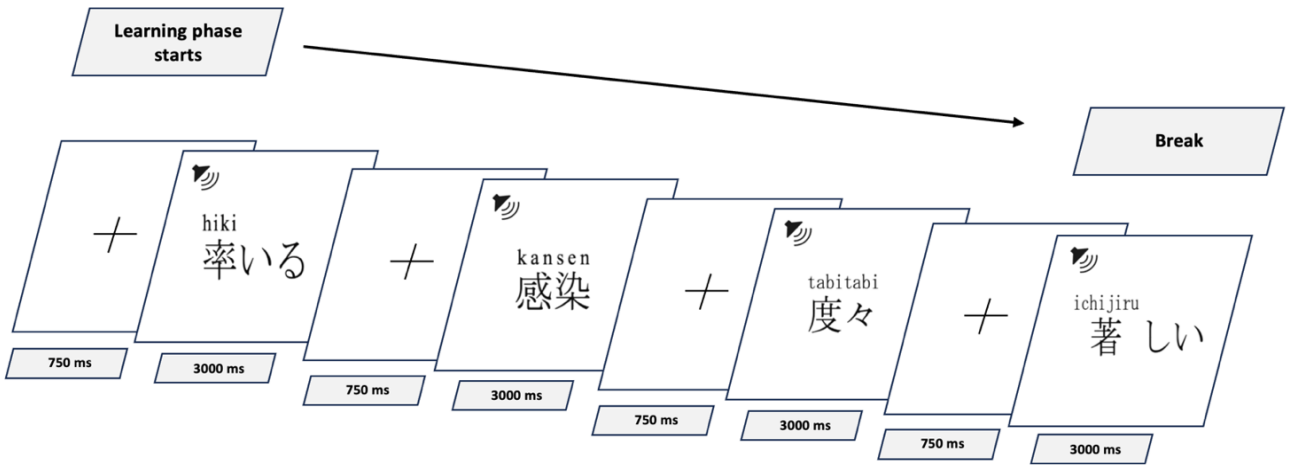
4.32 Flashcard learning cycle

After the pre-test, participants were redirected to take the 8–10-minute SPOT90 proficiency test hosted on the TTBJ webpage. Users were given a unique 'SPOT ID' and password (emailed to participants in advance). As there was no way to redirect participants at the end of the proficiency test, prior to being redirected to the TTBJ website, clear instructions were given to *return back* to Gorilla. Nevertheless, this caused some confusion, and as a result, several participants failed to return back to Gorilla to complete the study. Data from these participants were removed from the final data analysis. The SPOT90 scores were automatically emailed to the researcher upon completion of the proficiency test. The participants' unique SPOT ID numbers were used to match-up SPOT90 scores with their respective data from Gorilla.

Next, participants were randomised on a on a 1:1 allocation to either treatment group A (romaji) or treatment group B (furigana). The randomisation process was fully automated using Gorilla's 'randomiser node' (see Appendix G for a flow map of the study). Original flashcards were created by the researcher for all of the 16 target words for both learning conditions (36 flashcards). Care was taken to ensure in the furigana condition, that the hiragana symbols were placed directly over the top of corresponding kanji character, rather than being stretched across the whole word. This is to avoid confusion about which kanji word the furigana relates to, as noted by Kirwan (2005).

During the flashcard learning sequence, kanji words were presented in blocks (4 words x 4 blocks). Along with each word, an audio file automatically played the word's pronunciation. Audio for the Japanese audio files was recorded using Microsoft's Azure AI Cognitive Speech Studio[18]. The audio files were recorded in both male and female voices. Male and female voices were used to ensure that the audio was balanced during the kanji word learning task. This is because of known variation in the spoken form of Japanese between genders (Okamoto, 1995). The rationale for playing audio with the flashcard sequence was that in pedagogical settings, practitioners generally introduce new words on flashcards together with audio, or by speaking out aloud the word's pronunciation when presenting the novel items. Figure 4.2 shows an illustrative example of how the flashcards were presented to learners. Each randomised block commenced with a 750ms fixation point, followed by the target flashcard with audio for 3000ms.

Figure 4.2 Illustrative representation of the learning task on Gorilla (group B – romaji condition)



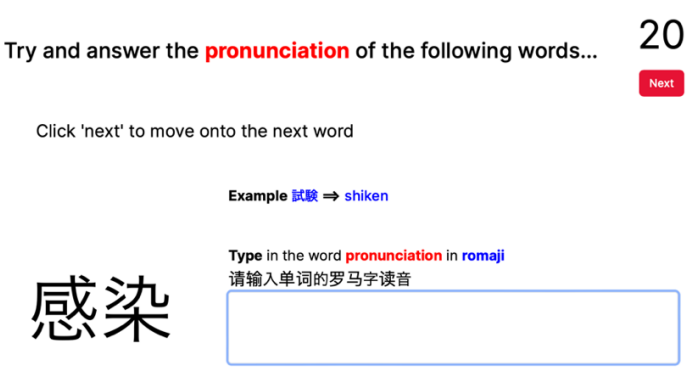
[18] Microsoft's Azure AI Cognitive Speech Studio is an artificial intelligence-based software tool where "text-to-speech" can be generated using a wide-range of different languages, speaking styles, accents, and regional dialects. Available at; https://azure.microsoft.com/en-gb.

After the automatic flashcard cycles, there were short breaks, followed by a recall task, where participants were asked to type in the pronunciation of the word on their machine (using either romaji or furigana depending on the condition). As there is strong evidence of the importance of repetition for word learning to occur (Baddeley, 1990; p.156), each word was presented the maximum of 7 times within the limited 10-minute learning cycle; 5 instances with flashcards, and 2 during pronunciation recall tasks. Baddeley (1990; p.156), stresses that not only is a word a more likely to be recalled based on the number of repetitions, but also having opportunities to recall the word which is being learned. Thus, a word recall task was included into the flashcard learning design. Although, Webb (2007b), argues that between 10 and 20 repetitions are required for meaningful learning, participants in this study were only required to learn the phonological form of the words, so it was felt that seven encounters with each word was sufficient to learn only word pronunciations.

4.33 Immediate-post test

For the post-test, participants had 20 seconds per word to recall kanji word pronunciation using either *romaji* or *furigana* (depending on condition). Instructions for the post-test (and delayed post-test) were given in both Chinese Mandarin and English to avoid confusion between participants with different L1 backgrounds. A score of 1 was given for full phonological access, and 0.5 for partial knowledge (16 maximum). Figure 4.3 shows the post-test Gorilla interface, with automatic timer and dual language instructions.

Figure 4.3 User interface on Gorilla of the immediate *post-test* (group B – romaji condition)



4.34 Delayed-post test

After a 24-hour delay, participants were automatically emailed a unique URL link directly from Gorilla inviting them to the 5-7-minute delayed post-test. The justification for choosing 24-hour delay between post and delayed post-test was that typically in classroom settings students are given a list of vocabulary to learn during their studies, often followed up by a vocabulary quiz the next day. After accounting for learners who knew more than half of the target words, of the final sample (N=75), 57 completed the delayed post-test, resulting in a final delayed post-test attrition rate of 24%[19].

4.4. Research ethics

The University of Oxford has strict measures in place to ensure that user data is used and handled appropriately. On the 5th of April, 2023, an ethics approval application for this dissertation was submitted to the University of Oxford's Central University Research Ethics Committee (CUREC). On the 19th of May, the relevant CUREC ethics approval was granted with ethics approval number EDUC_C1A_23_181 (see Appendix T). All participants were emailed consent forms to participate in the online experiment (see Appendix U). The informed consent form explained how quantitative data from the experiment would be collected, used and stored as recommended by university's CUREC guidelines. In addition, informed online consent was also sought from all participants immediately prior to starting the experiment on Gorilla (Appendix V), giving participants the option to withdraw immediately prior to taking part in the study.

[19] Attrition rate calculation = percentage of loss of participants / initial number of participants = (75-57)/75*100

5. RESULTS

5.1. Analysis and interpretation of data

The dataset of the final sample (N=75), were first checked for outliers, and tests of normality were performed on the SPOT90 proficiency scores and kanji pronunciation test scores to ensure that parametric assumptions were met. The importance of normality is essential, as it is an underlying assumption of performing reliable and valid statistical inferences (Razali & Wah, 2011). Both test scores and proficiency scores were checked for normality because both independent and dependent variables should be considered when performing inferential analysis (Orcan, 2020). Tests of normality were also checked independently at the group level, as normality for each group is an underlying assumption in parametric tests, and because normality should be checked at 'each level of the predictor variable' (Field, 2018; p.232).

5.11. Outlier detection and treatment

First the SPOT90 proficiency and kanji word test scores were checked for outliers using box-whisker plots by proficiency category and treatment condition (see Appendix L & M). No outliers or extreme cases were found in either group greater than between 1.5 to 3 times the upper quartile (Field, 2018). This is an important step, as outliers or extreme cases can affect the normality of a dataset (Baffoe-Djan & Smith, 2019). Table 5.1 shows the descriptive statistics for the overall SPOT90 scores.

Table 5.1 SPOT90 descriptive statistics across L2 proficiency levels (N=75)

	N	Minimum	Maximum	Mean		Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Std. Error	Statistic	Std. Error
All SPOT90 scores	75	30	55	55.76	1.540	13.336	-.143	.277	-.829	.548
Beginner* SPOT90 score	36	30	55	44.28	1.283	7.700	-.561	.393	-.813	.768
Intermediate** SPOT90 score	39	56	80	66.36	1.150	7.180	.095	.398	-1.170	.741

*Beginner 30-55, **Intermediate 56-80. Out of a possible maximum score 90 (TTBJ, University of Tsukuba).

Next, raw pre, post and delayed-post kanji test score data were checked for outliers. This was performed by proficiency level and treatment group. At the treatment learning group level, no outliers or extreme scores were found in the pre-test, however, 3 outliers were found in the post and delayed-post-test data within the romaji learning condition (see Appendix M & N). One of the outliers was in both the post and delayed post-test test data. Upon inspection, it was found that this participant's scores were very low (less than 3 marks in both). Taking into considering their low SPOT90 test score of 30 (on the cusp of beginner-level (30-55), they were removed from the dataset as we felt that the experiment may have been too difficult for the participant. The remaining 2 outliers found in the post and delayed post-tests were adjusted by the process of windsorization. Windsorizing involves 'replacing outliers with the next highest score that is *not* an outlier (Field, 2018; p.264). The 2 outlier scores in the post and delayed-post test score data were windsorized from 5 to 6, and 3 to 5.5 respectively. By proficiency level, 1 outlier was discovered in the intermediate delayed post-test data. The researcher felt that this may have been down to a slight performance anomaly compared to their respective group proficiency test scores. Thus, the decision to also windsorize this score was performed, increasing the low score by 3 marks from 4 to 7. Mean test scores by proficiency are reported in Table 5.2.

Table 5.2 Descriptive statistics showing Japanese kanji word pre, post and delayed post-test raw test scores by proficiency (N=75)

	Pre-test mean scores*		**Immediate post-test** mean scores*		**Delayed post-test** mean scores*	
	Mean (M)	Std. Deviation (SD)	Mean (M)	Std. Deviation (SD)	Mean (M)	Std. Deviation (SD)
Beginner (N=36)	2.800	2.515	10.314	4.218	8.696	3.945
Intermediate (N=39)	3.577	2.352	13.692	2.520	12.071	2.824

*Out of a possible 16

5.12. Normality testing and analysis

To determine whether parametric assumptions had been violated, tests of normality were calculated using Z-score values of skewness and kurtosis. Z-scored are considered a more robust approach compared to the commonly-used Kolmogorov-Smirnov Test of Normality which often lacks the power to detect whether or not assumptions of normality are violated in smaller sample sizes (Field, 2018). The Z-score method is calculated by dividing the distribution's skewness and kurtosis values by their standard errors, and checked to see if the values significantly differ from zero (i.e., different from a normal distribution) (Field, 2018; p.248). If the Z-score skewness and kurtosis results are greater than 1.96, then the dataset is significantly different from a normal distribution at $p =< .05$ (Field, 2018).

Z-score checks were performed for both proficiency groups, and importantly, as our predictor variable is categorical (i.e., romaji and furigana), checks were carried out on test score data for each independent variable (Field, 2018). Firstly, for the SPOT90 proficiency scores, no deviation from normality was found in skewness and kurtosis Z-scores in either group (see Appendix O). An independent paired samples t-test revealed that there were *was* a significant difference between proficiency SPOT90 scores and proficiency categories (t (73) = -12.852, p < .001), suggesting that the SPOT90 is an accurate L2 proficiency profiling instrument for beginner-intermediate level learners of Japanese. This is consistent with findings from Ford-Niwa & Kobayashi (1999). A t-test also confirmed that there were *no* significant differences in proficiency between both treatment groups (t (73) = -.114, p = .455). In other words, there were no statistically significant differences in proficiency between each randomly allocated learning condition.

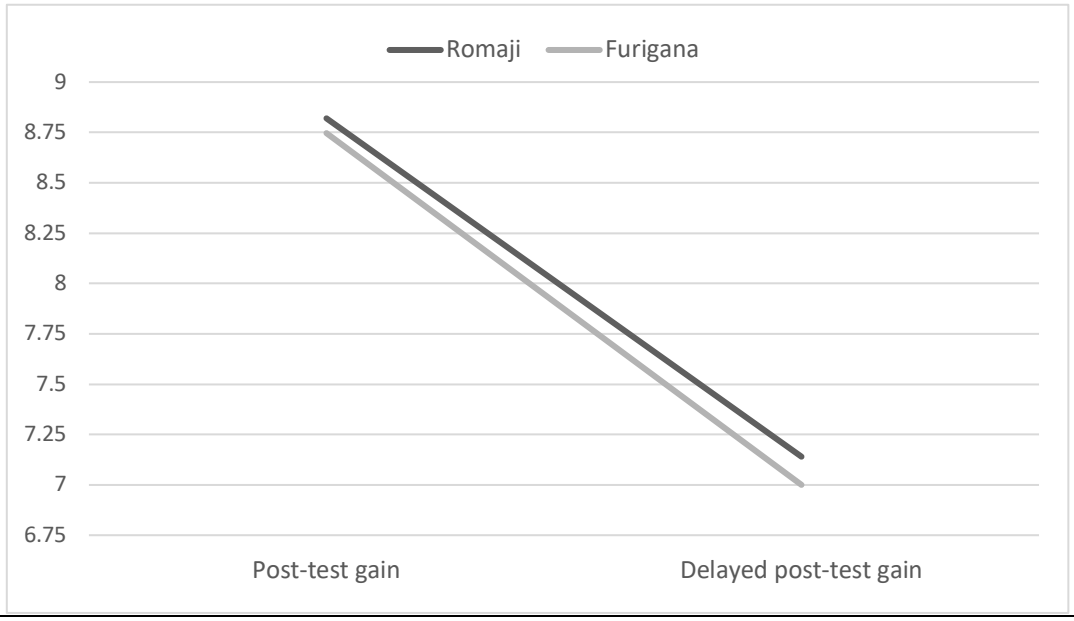
Next test score data was checked for normality by each independent treatment group. Overall, the pre, post and delayed-post Z-scores did *not* deviate significantly from normality (see Appendices P & Q). By treatment group, all 16 Z-score calculations also fell under the accepted Z-score cut-off of parameters of 1.96, except *one* Z-score in the romaji post-test data; with a negatively skewed Z-score of -2.717. The researcher felt that as 15 out of 16 Z-scores had all met the more robust Z-score parametric assumptions, (i.e., falling within the acceptable limits of ±1.96, thus being a non-substantiative diversion from normality (Cameron, 2004; Hair et al., 2022)), then including this data would be acceptable for parametric analysis. Nevertheless, as a violation was found, we felt it prudent to also perform a non-parametric analysis to confirm any significant parametric results. This is echoed by Field (2018), who

states that in applying the *F*-statistic only, then it is advisable to check the 'sensitivity' of parametric assumptions (Field, 2018; p.537).

5.13. Descriptive statistics

Overall, when observing the kanji word gains[20] (change in difference scores) between each treatment group (romaji, N=35, furigana, N=40), marginal gains are seen favour of the romaji condition. Figure 5.2 illustrates the kanji gain scores for each group in the post and delayed post-test.

Figure 5.1 Graph showing gain scores* in post-test and delayed post-test by treatment group (N=75)

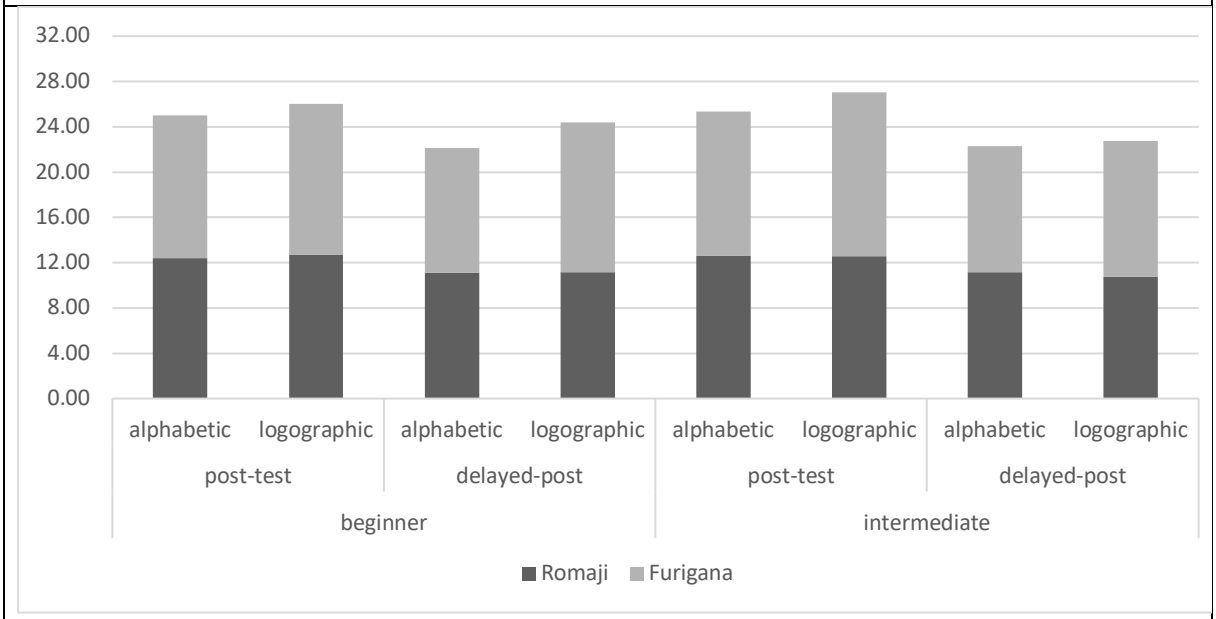


*Gain scores calculated by subtracting pre-test scores from post and delayed-post test results (Smolkowski, 2019)

Figure 5.2 shows for post-test and delayed post-test by proficiency level and L1 background. Observing the gains made for logographic learners, mean test score gains are higher in the authentic furigana condition for all tests. However, these descriptive results must be interpreted with extreme caution due to the small number of logographic learners (N=10) in the final sample across both learning conditions.

[20] Raw gain scores (change in difference scores) were calculated by subtracting pre-test score from post and delayed post-test scores, and then the mean score of those gains were calculated for each group (Smolkowski, 2019)

Figure 5.2 Stacked bar-chart showing average gain scores* by group and proficiency level (N=75)



*Gain score calculated by subtracting pre-test score from post and delayed post-test scores.

5.14. Inferential statistics

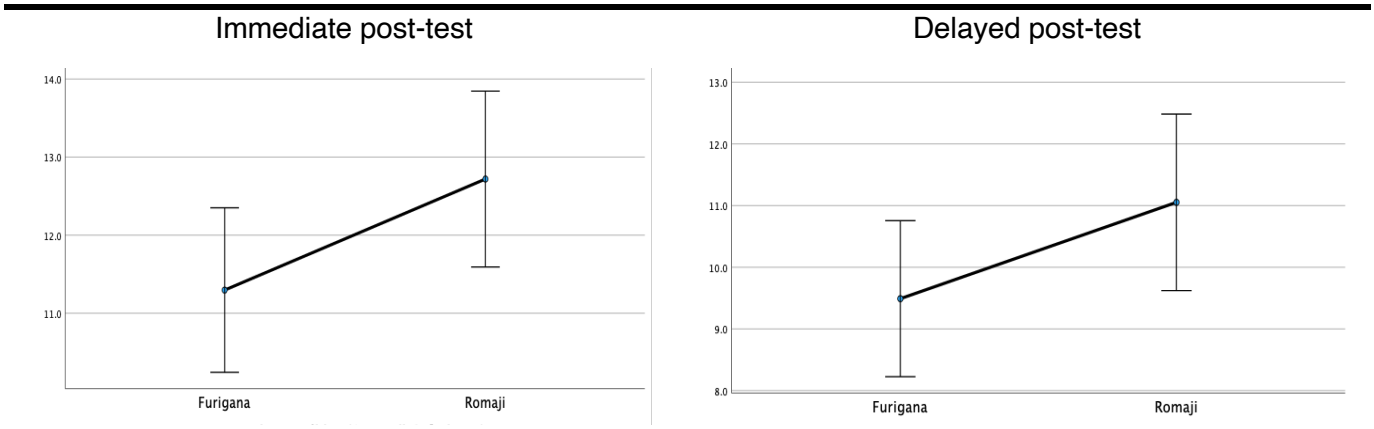
Before the ANVOCA was performed, checks were made to ensure there were *no* significant differences between baseline pre-test scores between groups. No significant differences were found (p = .751). Next, the assumption of homogeneity of regression (homoscedasticity) was checked to ensure that the relationship between test scores (outcome variable) and covariate (pre-test scores) are the same in each learning condition (Field, 2018; p.586). A one-way ANOVA confirmed the effect of each learning condition was *not* significant, meaning the assumption was met as test scores were linear (i.e., not significantly different) in each group (see Appendix R). The homoscedasticity check was performed on both post-test (p = .749), and delayed post-test data (p = .823).

To investigate the overall relative effects of treatment group on kanji recall (RQ1), an ANCOVA analysis was performed on SPSS with test scores (post and delayed post) as the dependent variables, treatment groups as the independent variables controlling for pre-test scores as a covariate. The results of the ANCOVA showed an overall *significant* main effect between treatment groups in post-test scores F (1, 73) = 5.036, p = .028, with reported eta squared

effect size of $\eta^2 = .066$. In other words, 6.6% of the post-test score variance is attributable to romaji. Eta squared (η^2) over 0.06 are considered medium[21] (Cohen 1988).

The ANCOVA was repeated for the delayed post-test scores. Again, a significant difference between groups *was* found $F\,(1, 55) = 4.911$, $p = .031$, with effect size $\eta^2 = .085$ explaining 8.5% of the test score variance. As the *F*-test only shows that there a difference between groups exists (but not in which direction). Figure 5.3 shows a visual representation of the estimated marginal means between treatment groups. While the 95% confidence intervals do overlap, the higher mean scores for group B (the romaji condition) *are* significantly different. Nevertheless, at the group level, the researcher felt that due to parametric violations of skewness in post-test data (see Appendix P), a non-parametric Kruskal-Wallis test should be performed to confirm accuracy of the statistical significance. The Kruskal-Wallis test is used to compare independent scores across different groups (Field, 2018; p.306). The results of the non-parametric test found romaji to be very close to, but *not* significantly different in post-test $H\,(1) = 3.374$, $p = .066$, and delayed post-test scores $H\,(1) = 3.598$, $p = .058$ (see Appendix S).

Figure 5.3 Interaction plot showing the estimated marginal means of the post and delayed post-test scores by treatment group controlling for pre-test scores.



*Covariates in the model estimated at 3.140, Error bars: 95% CI

*Covariates in the model estimated at 3.035, Error bars: 95% CI

Next interaction effects between proficiency level and treatment group were examined (RQ1.1), in both post and delayed post-tests. *No* significant differences were detected for post ($p = .124$,

[21] For Cohen's *d,* effect sizes over 0.5 are considered medium in behavioural science research (Cohen, 1988)

and delayed post-tests (p = .107) respectively). Finally, we looked at the effects of L1 background and treatment group on outcome measures (RQ1.2). Due to a low sample of logographic learners (N=10), unfortunately it was not possible to examine this research question using inferential statistics.

In addressing the relationship between Japanese L2 proficiency and L1 script background on self-regulation, we carried out an ANOVA analysis on the full sample (N=98), with the dependent variables as 'commitment' and 'metacognitive control', and fixed factors as 'proficiency level' and 'L1 background'. The results found the *only* statistically significant difference between L2 proficiency and metacognitive control (survey item 1), $F(1) = 4.793$, $p = .031$. *No* other statistically significant differences were detected, except a large (but non-significant) difference between L2 proficiency and commitment control (survey item 1) $F(1) = 3.693$, $p = .056$. Descriptive statistics from the self-report SRCKanji are shown below in Table 5.3.

Table 5.3 Descriptive statistics for the SRCKanji sub-scale factors mean scores by L2 proficiency and L1 background (N=98).

		Modified SRCKanji* self-regulation survey							
Sub-scale coefficient		**Commitment control** Mean (M) scores				**Metacognitive control** Mean (M) scores			
Sub-scale survey item number (#)		#1	#2	#3	#4	#1	#2	#3	#4
L2 proficiency	Beginner (N=42)	3.12	3.02	3.50	3.86	**3.00****	3.02	2.76	2.57
	Intermediate (N=56)	2.67	3.27	3.31	3.64	**2.51****	2.89	2.78	2.42
L1 background	Alphabetic (N=81)	2.77	3.07	3.39	3.64	2.64	2.93	2.72	2.49
	Logographic (N=17)	3.29	3.59	3.41	4.18	3.12	3.06	3.00	2.47

The modified SRCKanji Cronbach alpha score α = .794* (individual sub-components*; commitment control* α = .649, and *metacognitive control* α = .689). ***p = > .05*

6. DISCUSSION

6.1. Interpretation of results

This study set-out to uncover any qualitative differences in learning novel Japanese kanji words using two different phonological decoding tools. To interpret the results, we will revisit our initial research questions and sub-questions.

RQ1. To what extent do different phonological reading aids (romaji and furigana) help in the learning and retention of L2 Japanese kanji word pronunciation?

Overall, the results found that romaji did in fact facilitate significantly higher Japanese kanji word pronunciation scores, with mean scores in the post-test being 14.5 versus 12 for furigana (see Appendix M). Not only were test scores higher in the romaji group, but box-whisker plots show that there was also less variation in the spread of scores compared with the furigana condition (Appendix M). In other words, the furigana group displayed much larger variation in kanji naming accuracy (i.e., broader range of extreme scores). The broader implication of these results provides some evidence that using non-authentic texts such as romaji may result in more efficient (and less varied) L2 Japanese kanji word learning. The combined main effect also shows that regardless of L1 background, *both* alphabetic and logographic learners may benefit overall using romaji to improve their kanji pronunciation accuracy. This may be due to first language transfer in L2 phonological processing, from knowledge of alphabetic and pinying writing systems. In terms of short-term 24-hour retention, significant differences were also found in favour of the romaji group. However, any longer-term benefits of using romaji for kanji word pronunciation beyond 24 hours is unknown. Thus, in light of the results, we should be cautious in making large generalisations about the benefits of using romaji. Results from the non-parametric analysis, although close to being significant, found *no* significant difference between learning groups.

1.1 What are the relative effects of different reading aids and L2 Japanese proficiency level on L2 word pronunciation?

In terms of the interaction between reading aid treatment and proficiency, although significant differences were found overall between main *groups*, the results by group and L2 proficiency level were *not* found to be significantly different. The results suggest that despite increases of

proficiency level, preference for either learning aid appears to remain stable. This is contrary to evidence found by Chikamatsu's (2006) 'developmental' proficiency influence, which found that cross-linguistic effects may weaken as L2 Japanese proficiency increases.

1.2 What are the relative effects of different reading aids and L1 background on L2 word pronunciation?

Unfortunately, this study was unable to fully explore this research question with inferential statistics based on the limited number of logographic learners recruited for the study. Nevertheless, preliminary descriptive data does suggest a preference for furigana (refer to Figure 5.2). We must be careful in interpreting these findings based on very few logographic learners within the sample. Despite this, even though the logographic learners (N=10) were included in each treatment group, a significant difference between groups in favour of using romaji was still detected, which may negate any hypothesised benefit of furigana for logographic background learners.

RQ2. What is the relationship between self-regulation, L1 background and L2 proficiency?

The results from the ANOVA data analysis from the self-regulation survey found only one significant difference when inspecting metacognitive behaviour and L2 proficiency level. The significant difference was detected between beginner and intermediate level learners (in favour of beginners) with the following metacognitive survey item;

'*When learning Japanese vocabulary (kanji words), I have special techniques to keep my concentration focussed'*

This difference reveals that when compared with more advanced learners, beginner-level learners of Japanese seem to have greater metacognitive control in their kanji word learning. This is perhaps not surprising, as higher-level learners of Japanese often lack self-direction or become overwhelmed by a perceived lack of progress. This would potentially then result in a loss of metacognitive control in keeping concentration levels focussed and their learning on track. These results are consistent with qualitative findings of Rose and Harbon (2013) who found that more advanced learners suffer from self-regulation breakdowns due to learner frustrations due to lack of progress, or self-criticism as a result of not being able to reach language goals.

Further, when examining Japanese learners' commitment control by proficiency level, a large (but non-significant) difference was found in favour of beginners compared to intermediate Japanese learners. The difference was found in the following survey item;

> '*When learning vocabulary, I have special techniques to achieve my learning goals.'*

Again, this suggesting that beginner learners use more commitment control techniques in autonomous goal setting and goal achievement compared to more proficient Japanese learners. Although the result was non-significant, it does hint at the possibility that self-regulated goal-setting behaviour and techniques may deteriorate as learners advance through stages of Japanese language learning. Often, when learners start out learning Japanese, are very proactive in self-regulating their self-study behaviours to achieve their goals. For example, they may keep records, or track their progress. However, as Japanese learners advance, they may find it harder to stay committed learning kanji words as they begin to feel overwhelmed by the sheer number of kanji compounds and the perceived impossibility of learning thousands of complex kanji compound readings.

When examining self-regulation by different writing system backgrounds, *no* other statistically significant differences between groups were found. The null result provides some evidence of little qualitative differences between self-regulation across diverse writing systems (and cultural backgrounds). Given logographic background learners have to spend thousands of hours learning Chinese hànzi in during schooling, it is perhaps surprising that no differences were found in their self-regulatory behaviour compared to non-morphographic script learners. This is certainly a ripe area for future self-regulation L2 kanji research. A limitation of exploring kanji learning under the scope of self-regulation is that the scale response does not reveal more information about the nature of the qualitative differences between these groups of learning in terms of what type of 'special' techniques. Therefore, to build on early work in the field by Rose and Harbon (2013), future studies in self-regulation and Japanese kanji learning would greatly benefit from more depth and qualitative approaches, as pointed out by Woodrow (2005).

6.2. Implications for pedagogy

At the core of the debate between L2 Japanese instructors are two main questions. Firstly, whether using romaji (rather than more authentic scripts) in pedagogical settings is effective in L2 Japanese learning, and second, if we do introduce romaji, when should we introduce furigana (and kanji). In addressing the first question, evidence from this study indicates that while there is benefit to improving the efficiency of Japanese kanji learning with the aid of romaji, the benefit may be reduced longer-term as learners have to eventually move onto more difficult scripts. Some theorists note (e.g., Koda, 1992), that even if romaji helps to facilitate learning, mental representations in a learner's schemata would need to be rebuilt later on when moving toward more authentic scripts. In turn, this may create more a complex and cumbersome task in building the L2 phonological (and semantic) lexical representations. Therefore, in-line with Koda (1992), our view, is to use romaji as an aid in very early L2 Japanese learning, presenting where possible romaji alongside more authentic scripts such as kana and kanji. This would improve exposure and encounters with authentic material to build stronger associative links (Nation, 2022).

In relation to the second argument concerning when to introduce more authentic scripts in the classroom. On the balance of the empirical evidence, romaji appears to do no harm in the short-term, and does appear to provide marginal gains in learning phonological information of kanji word pronunciation. However, the evidence should be interpreted with care by practitioners as the longer-term effects of delaying furigana (kanji) after learning with romaji is far less known. To date, there is weak evidence to suggest that romaji significantly impedes L2 Japanese linguistic development. Nevertheless, the overuse of romaji may hamper linguistic development if overused or over-relied on. Like training wheels attached to a bicycle, romaji may be extremely useful at the beginning of a learner's language journey, but the timing of removing the training wheels and moving onto authentic material is critical in development. In sum, at best romaji can be a useful tool, and at worst appears to do no harm at least in the short-term. However, without further empirical evidence, the use of romaji in Japanese L2 classrooms is likely to remain a controversial topic.

Finally, in terms of self-regulated control, practitioners should be aware of the significant cognitive challenges in L2 Japanese learning. Importantly, these challenges may not only affect early and beginner-level learners, but also disproportionately affect higher-level learners. Results from the self-regulation analysis suggest that more advanced learners may suffer from

higher levels of metacognitive breakdowns. Subsequently learners may face breakdowns in essential goal-setting and commitment control strategies. This is unsurprising, as progress can be negatively impacted without achievable short-term goals (Bandura & Schunk, 1981). Practitioners can help by increasing awareness of self-regulatory and metacognitive strategies (particularly for advanced learners), thus reducing the demotivational effects of learning an extremely difficult and linguistically complex language such as Japanese.

6.3. Limitations of the study

6.31 Methodological limitations

This study must acknowledge several methodological limitations. Where possible justification of these limitations will be explained. First, the quasi-experimental design resulted in the lack of a full control group. The addition of an 'audio-only' control group would have been useful to compare post-test gains as a baseline against the two reading aid conditions. However, the addition of a third group acting as a control, would have resulted in a significantly larger sample required for the ANCOVA analysis, which the researcher felt was beyond the limited resources in this study.

Second, a more robust vocabulary test was also absent in this study. Although, knowledge of the target words was measured using a simple pre-test, we were unable to gain a full understanding participants' vocabulary breadth and depth. A lack of more robust vocabulary measure was primarily due to limited time constraints. Third, in terms of learning novel kanji words, only learning the phonological representation of the kanji could be viewed as a major limitation in this study. As phonological representations of a character or word is only one facet of a word's components of word knowledge (i.e., its form, meaning and use) Nation (2001). The rationale for only measuring phonological recall was because the study's aims was to simply to examine the ability of phonological aids to enhance word 'learning' (i.e., its phonological form). However, the study must acknowledge that only learning kanji pronunciation is a severely gross over-simplification of Nation's (2001) complete knowledge of a word's 'form'. Therefore, the implications for the study's generalisability for reading Japanese reading aids assisting kanji word learning is significantly reduced as these superficial word encounters reduces a learner's ability to forge stronger links with their mental lexicon (Nation, 2022; Perfetti & Hart 2002).

Investigating the effectiveness of pedagogical reading aids beyond phonological access (i.e., semantic meaning or grammatical forms), is an area for future research. Moreover, the use of flashcards in short-term kanji pronunciation is a rather 'shallow' method of learning. As such, it could be argued that this study is merely a test of short-term memory recall, rather than deeper lexical development. A longer delayed post-test (greater than 24-hours) may have also been useful to fully understand the long-term effects of each reading aid condition phonological recall over a longer time frame and may have provided different data between learning groups.

One unexpected result was that lower-level learners found the SPOT90 proficiency test too difficult. Although the SPOT90 worked as an effective proficiency profiler, the nature of the SPOT90 requiring participants listen to excerpts of Japanese at near native speed while reading hiragana was too difficult for lower-level learners. Therefore, in the future if recruiting early or beginner-level learners in future studies, we would not recommend the SPOT90 due to the difficulty and pressured nature of the test.

In terms of using the SRCKanji to measure self-regulation, one major limitation was that the instrument was not methodologically tested and validated as an L2 research tool for kanji word learning. Although this study may have been one of the first to empirically test the SRCKanji, the field could benefit from testing JFL context specific tools such as the SRCKanji in a similar fashion to Ardasheva and Tretter's (2013) work adapting and validating a version of Oxford's (1990) Strategy Inventory for Language Learning (SILL) for younger primary aged English language learners. Nevertheless, a further limitation of using psychometric methods is that instruments such as the SILL and the SRCVoc have been severely criticised for not effectively measuring mental processes which are not always captured through these self-report style questionnaires. Indeed, the field of self-regulated L2 kanji learning would greatly benefit from more in-depth qualitative approaches (such as stimulated recall and structured interviews) (Dörnyei, 2005; Tseng et. al., 2006; Woodrow, 2005). Qualitative methods would provide much richer data and move the field further in understanding *what* type strategic behaviour and 'special techniques' are used by Japanese kanji learners in order to control and regulate their individual learning. This is a particularly relevant in understanding L2 strategic behaviour in challenging languages such as Japanese.

6.32 Pedagogical limitations

Perhaps one of the largest pedagogical limitations is that study was *not* carried out in real JFL classrooms. Burns (2000) asserts that for results to be useful in research, 'they must be generalisable beyond the confines of the particular experiment' (p.143). In other words, for results to be useful they must also be 'externally valid' (Burns, 2000; p.143). Generally, learning L2 vocabulary through flashcards is considered 'general purpose' and typically carried out by learners outside of the classroom (Nation, 2022; p.271). The learning of kanji words in this study is underpinned by a rather shallow flashcard learning technique, where creating meaningful associations is unlikely. In typical Japanese language classrooms flashcard learning may be complimented with additional information from language instructors such as introducing the epistemological history of kanji characters, or by explicitly pointing out a kanji's component parts. This may lead to richer learning, allowing the learner to form deeper and more meaningful lexical representations of kanji words, leading facilitate more efficient retrieval. While there is some merit in repetitive flashcard learning, the simplification of kanji word learning task may not fully represent learning of new lexical items, flashcard practice should be a supplementary word learning activity. Flashcard learning lacks the explicit learning which occurs in a classroom environment, which can create richer interactions with novel words (Sonbul & Schmitt, 2009). In turn, incidental and explicit learning may also be facilitated through conscious 'noticing' (Schmidt,1990). Thus, given the limitations of learning kanji words with flashcards, we must acknowledge the limitations of the results for JFL learning within classroom settings.

A second pedagogical limitation is that the way romaji were used in this study (i.e., as a kanji decoding tool). The romaji script is *not* typically used as a reading aid in L2 Japanese pedagogical contexts, but rather are used as a direct replacement for authentic texts. While there may be a link between romaji and Japanese kanji word learning, the jump is a significant diversion from how romaji are used pedagogically. If this study were to be replicated, romaji could be presented in a more pedagogically authentic manner (i.e., without the use of kana or kanji at all) (refer to Figure 1.4 which highlights typical romaji usage). Thirdly, it should be noted that learners recruited in this study were already familiar with hiragana. This was to ensure that fair learning could occur if a participant was randomised to the hiragana treatment group. A potential limitation of this requirement, is that the effectiveness of using romaji to learn new kanji words (typically used in early stage of Japanese learning), may have been

significantly reduced as subjects had already acquired knowledge of authentic kana. Thus, if replicating this study from a pedagogical standpoint, it would be useful to re-examine the work of Kirwan (2003) exploring the relative effects of using romaji to learn Japanese with learners with *no* prior knowledge of hiragana. This is a core area of study for further investigation in future L2 Japanese research.

7. CONCLUSION

7.1. Summary of findings

This study addressed the phenomenon faced by L2 learners learning Japanese kanji, one of the most phonological inaccessible and orthographically opaque writing systems in the world. In particular, this study set-out to find empirical evidence of the effects of kanji word learning using two different L2 phonological reading aids (romaji and furigana). An overall main effect *was* found between each the two different of reading aids regardless of L2 proficiency with medium effect size (η^2=.066). While the results suggest that using romaji can result in more efficient (and less variability) in phonological decoding of novel kanji words, the long-term effects may be marginal as learners progress through stages of development. The significance of these results, is that there is some evidence to suggest the role of cross-linguistic transfer having a facilitative effect of Japanese kanji word learning for at least for alphabetic (and possibly Chinese script background logographic learners with knowledge of the pinyin system). In sum, while romaji appears to be a relatively effective method in facilitating novel kanji learning which seems to do no harm. The results of different reading aids by L1 logographic learners are still relatively unknown, thus this is a area for future studies to further investigate.

7.2. Directions for future research

If this study were to be replicated, we feel that using phonological aids to learn kanji words beyond phonology (i.e., semantic meaning) would be useful to gain a fuller picture of the effectiveness of these pedagogical tools. Next, as around 25% of participants knew more than 50% of the target kanji words (phonological form), we recommend selecting more difficult target kanji words from higher word band thresholds. In terms of selecting kanji, there is a methodological opportunity to examine learner strategy use by different kanji type[22] argued by Rose (2017). Further, while romaji appears to be effective in the short-term, an investigation into the longer-term effects of using romaji would be highly beneficial for L2 pedagogy. Finally, empirical investigation of Mandarin Chinese learners with logographic backgrounds is paramount, as these learners make up the overwhelming majority of Japanese learners. In addition, learners from the Philippines, Vietnam and Thailand, who have witnessed tremendous growth in L2 Japanese learning (Mori et. al., 2021) and would also benefit from future studies in the field.

[22] Kanji types include; shokeimoji (pictographs), shijimoji (logograms), kaiimoji (ideographs), keiseimoji (semasio-phonetic ideographs), tenchuumoji (derivative characters), kashamoji (phonetic loan characters) (Rose (2017).

REFERENCES

Akamatsu, N. (1999). The effects of first language orthographic features on word recognition processing in English as a second language. *Reading and Writing: An Interdisciplinary Journal*, *11,* 381–403.

Akamatsu, N. (2002). A similarity in word-recognition procedures among second language readers with different first language backgrounds. *Applied Psycholinguistics* 23, 117–133.

Akamatsu, N. (2003). The Effects of First Language Orthographic Features on Second Language Reading in Text. *Language Learning*, 53(2), 207–231.

Akamatsu, N. (2006). Literacy acquisition in Japanese-English bilinguals. In R. M. Joshi & P. G. Aaron (Eds.*), Handbook of orthography and literacy* (pp. 481–496). Mahwah, NJ: Lawrence Erlbaum Associates.

Akiyama, N., & Akiyama, C. (1995). *Master the basics: Japanese*. Hauppauge, NY: Barrons Educational Series.

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, 52(1), 388–407.

Ardasheva, Y. (2016). A structural equation modelling investigation of relationships among school-aged ELs' individual difference characteristics and academic and second language outcomes*. Learning and Individual Differences*, 47, 194–206.

Ardasheva, Y., & Tretter, T. R. (2013). Strategy Inventory for Language Learning–ELL Student Form: Testing for Factorial Validity. *The Modern Language Journal*, 97(2), 474–489.

Ashcroft, R. J., Cvitkovic, R., & Praver, M. (2017). Digital flashcard L2 Vocabulary learning out-performs traditional flashcards at lower proficiency levels: A mixed-methods study of 139 Japanese university students. The EuroCALL Review, 26(1), 14–28.

Azure AI Cognitive Speech Studio by Microsoft (n.d.). Retrieved July 23, 2023, from https://azure.microsoft.com/en-gb.

Bacchetti, P., Wolf, L. E., Segal, M. R., & McCulloch, C. E. (2005). Ethics and sample size. *American Journal of Epidemiology*, 161(2), 105-110.

Baddeley, A. (1990). *Human Memory*. London: Lawrence Erlbaum Associates.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual review of psychology*, *63*, 1-29.

Baffoe-Djan, J. B., & Smith, S. A. (2019). Descriptive statistics in data analysis. In J. McKinley & H. Rose (Eds.), *The Routledge Handbook of Research Methods in Applied Linguistics* (pp. 398–414). Abingdon: Routledge.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), 586–598.

Banno, E., Ikeda, Y., Ohno, Y., Shinagawa, C., Tokashiki, K., (1999). *[初級日本語げんき] GENKI: An Integrated Course in Elementary Japanese.* Tokyo: The Japan Times Ltd.

Benson, P. (2007). Autonomy in language teaching and learning *Language Teaching*, 40, 21–40.

Bhandari, P. (2023). *Statistical Power and Why It Matters | A Simple Introduction.* Scribbr. Retrieved July 25, 2023, from https://www.scribbr.com/statistics/statistical-power/.

Bourke, B. (1996). *Maximising the kanji learning task*. Unpublished doctoral dissertation. The University of Queensland.

Bryman, A. (2021) *Social research measures* (6th ed.), Oxford: Oxford University Press.

Cameron, A. C. (2004). Kurtosis. In M. S. Lewis-Beck, A. Bryman, & T. F. Liao (Eds.), *The SAGE Encyclopedia of Social Science Research Methods* (pp. 544–546). Thousand Oaks, CA: Sage.

Chen, Y-M. (2002). Nihongo niji kanji goi to sore ni tai-suru chugokugo niji kanji goi wa onaji ka: Taiwan oyobi chugoku no chugokugo to no hikaku [Do Japanese kanji compound words and Chinese words share the same meaning?: Drawing a comparison between Mandarin of Taiwan and China]. Gengo bunka to nihongo kyoiku [*Journal of Language Culture and Japanese Language Teaching*] 24:40-53.

Chen, Y., Fu, S. D., Smith S. M., & Matthews, P. M. (2002). Testing for dual brain processing routes in reading. A direct contrast of Chinese characters and Pinyin reading using fMRI. *Journal of Cognitive Neuroscience*, 14 (7), 1088-1098.

Chen, H., Yamauchi, T., Tamaoka, K., & Vaid, J. (2007). Homophonic and semantic priming of Japanese kanji words: A time course study. *Psychonomic Bulletin & Review*, *14*, 64–69.

Chikamatsu, N. (1996). The effects of L1 orthography on L2 word recognition: A study of American and Chinese learners of Japanese. *Studies in Second Language Acquisition*, 18(4), 403–432.

Chikamatsu, N. (2005). L2 Japanese Kanji Memory and Retrieval: An Experiment on the Tip-of-the-pen (TOP) Phenomenon. In V. Cook & B. Bassetti (Eds.), *Second Language Writing Systems* (pp. 71–96). Blue Ridge Summit, PA: Multilingual Matters.

Chikamatsu, N. (2006). Developmental Word Recognition: A Study of L1 English Readers of L2 Japanese. *The Modern Language Journal*, 90(1), 67–85.

Chiu, H, C. (2002). Taiwanjin nihongo gakushusha ni okeru nihongo kanji jukugo no shori katei – Nitchū 2-gengo kan no dōkongo to hi-dōkongo no hikaku [The processing routes of Japanese kanji words for the Taiwanese learners of Japanese: The comparison of cognates and non-cognates between Japanese and Chinese]. Hiroshima Daigaku Daigakuin Kyōikugaku Kenkyūka Kiyō – Dai 2-bu [*Bulletin of Graduate Studies of Education at Hiroshima University* – Part 2] 51. 357–365.

Chiu, H, C. (2003). Taiwanjin nihongo gakushūsha no nihongo kanji jukugo no on'in shori nitsuite – tango taipu/tango no shūtoku nenrei/shūzyukudo no kanten kara no kōsatsu [The phonological processing of Japanese kanji words by Taiwanese learners of Japanese: An analysis of word type, age of acquisition, and fluency level]. Nihongo Kyōiku [*Journal of Japanese Language Teaching*] 116. 89–98.

Chiu, Yu Yuan. 2006. Taiwanjin nihongo gakushūsha ni okeru nihongo tango no chōkakuteki ninchi – Nihongobun no imihandan kadai o mochiita kentō [*Auditory recognition of Japanese words in Taiwanese learners of Japanese: Discussion on meaning decision task of Japanese sentences*]. Hiroshima Daigaku Daigakuin Kyōikugaku Kenkyūka Kiyō – Dai 2-bu [Bulletin of Graduate Studies of Education at Hiroshima University – Part 2] 55. 267–273.

Cohen, J. (1988). The Analysis of Variance. In Statistical Power Analysis for the Behavioral Sciencess (pp. 270–335). Hillside, New Jersey: Lawrence Erlbaum Associates.
In V. Cook & B. Bassetti, *Second Language Writing Systems* (pp. 71–96). Blue Ridge Summit, PA: Multilingual Matters.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The Cognate Facilitation Effect: Implications for Models of Lexical Access. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26(5), 1283.

Coulmas, F. (1989). *The writing systems of the world*. Cambridge, MA: Blackwell.

Cronbach, L, J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.

Crowley, D, P. (1965). *Linguistics and Japanese Reading*. Hilo, Hawaii: Pacific Linguistics.

Csizér, K., & Dörnyei, Z. (2005). The Internal Structure of Language Learning Motivation and Its Relationship with Language Choice and Learning Effort. *The Modern Language Journal*, 89(1), 19–36.

Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystock (Ed.), *Language processing in bilingual children* (pp. 70–89). Cambridge: Cambridge University Press.

Dewey, D. P. (2008). Japanese Vocabulary Acquisition by Learners in Three Contexts. Frontiers: *The Interdisciplinary Journal of Study Abroad*, 15(1), 127–148.

Dewey, D. P. (2004). Connections Between Teacher and Student Attitudes Regarding Script Choice in First-Year Japanese Language Classrooms. *Foreign Language Annals*, 37(4), 567–577.

DeVellis, R. F. (1991). *Scale Development: Theory and Applications*. London: Sage Publications, Inc.

Dizon, G. (2016). Quizlet in the EFL Classroom: Enhancing Academic Vocabulary Acquisition of Japanese University Students. *Teaching English with Technology*, 16(2), 40–56.

Dörnyei, Z. 2001. Motivational Strategies in the Language Classroom. Cambridge: Cambridge University Press.

Dörnyei, Z (2005) *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Mahwah: Routledge

Dwyer, E, S. (1997) *Getting started the right way: An investigation into the introduction of kanji study to neophyte Japanese learners*. PhD thesis. The University of Texas at Austin.

Eda, S., Itomitsu, M., & Noda, M. (2008). The Japanese Skills Test as an On-Demand Placement Test: Validity Comparisons and Reliability. *Foreign Language Annals*, 41(2), 218–236.

Elgort, I. (2010). Deliberate Learning and Vocabulary Acquisition in a Second Language. *Language Learning,61*(2), 367-413.

Ellis, N. C. (1995). The Psychology of Foreign Language Vocabulary Acquisition: Implications for CALL. *Computer Assisted Language Learning*, 8(2–3), 103–128.

Everson, M. E. (2009). Literacy in Chinese as a foreign language. In M. E. Everson & Y. Xiao (Eds.), *Teaching Chinese as a foreign language: Theories and applications* (pp. 97–112). Boston, MA:T

Everson, M. E. (1994). Toward a Process View of Teaching Reading in the Second Language Chinese Curriculum. *Theory Into Practice, 33*(1), 4-9.

Everson, M. E. (2011). Best Practices in Teaching Logographic and Non-Roman Writing Systems to L2 Learners. *Annual Review of Applied Linguistics*, 31, 249–274.

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, 175-191.

Field, A, P. (2018) *Discovering statistics using IBM SPSS statistics*. (5th ed.) Sage Publications: London

Fitzpatrick, T. (2012). Tracking the changes: vocabulary acquisition in the study abroad context. *The Language Learning Journal*, 40(1), 81–98.

Fitzmaurice, G.M., Laird, N.M. and Ware J.H. (2004) *Applied longitudinal analysis*. John Wiley & Sons, Hoboken.

Ford-Niwa, J., & Kobayashi, N. (1999). SPOT: A test measuring “control” exercised by learners of Japanese. In N. Risto & K. Kanno (Eds.), *The acquisition of Japanese as a second language* (pp. 53–69). Amsterdam: John Benjamins.

Fujita, (1999) Nihongo no kanji kana majiri go no ninchi yunitto ni kansuru kentou [A study on the recognition unit for a Japanese Kanji-Kana-combined-word] *The Japanese Journal of Psychology* 1999, Vol. 70, No. 1, 38-44

Fukuda, M. (2004). Listening comprehension in Japanese as a second language and the capacity of working memory by adult-Malay-language-speaking learners. *Acquisition of Japanese as a Second Language* 7. 45–59.

Fukuda, M. (2005). Listening comprehension in Japanese as a second language and working memory span: A study on differences in proficiency, targeting native speakers of Chinese. *Bulletin of the Graduate School of Education, Hiroshima University, Arts and Science Education* 2(53). 299–309.

Gao, X. (2007). Has language learning strategy research come to an end? A response to Tseng et al. (2006). *Applied Linguistics*, 28, 615-620.

Gamage, G. H. (2003). Perceptions of kanji learning strategies: Do they differ among Chinese character and alphabetic background learners? *Australian Review of Applied Linguistics*, 26, 17–30.

Geva, E., & Wang, M. (2001). The development of basic reading skills in children: a cross-language perspective. *Annual Review of Applied Linguistics*, 21, 182–204.

Grabe, W. (2009) *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.

Graham, S. (2022). Aptitude or motivation: Which is the better predictor of successful languge learning? In E. Macaro & R. Woore (Eds.), *Debates in Second Language Education* (pp. 106–121). London: Routledge.

Grainger, P. (2005). Second language learning strategies and Japanese: Does orthography make a difference? *System*, 33(2), 327–339.

Grenfell, M. and Macaro, E. (2007) Language learner strategies: claims and critiques. In, Macaro, E. & Cohen, A. (Eds.). *Language Learner Strategies: 30 Years of Research and Practice.* (Oxford Applied Linguistics) Oxford, UK. Oxford University Press, pp. 9-28.

Gu, Y. (2012). Learning Strategies: Prototypical core and dimensions of variation. *Studies in Self-Access Learning Journal*, 3(4), 330-356.

Hadamitzky W, & M. Spahn, (1981) *Kanji & kana: A handbook and dictionary of the Japanese writing system*. Rutland: Tuttle.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (3rd ed.). Sage: Thousand Oaks, California

Hagiwara, A. (2016). The Role of Phonology and Phonetics in L2 Kanji Learning. *The Modern Language Journal*, 100(4), 880–897.

Hannas, W, C. (1997). *Asia's Orthographic Dilemma.* Honolulu: University of Hawai'i Press

Hassan, X., Macaro, E., Mason, D., Nye, G., Smith, P., & Vanderplank, R. (2005). Strategy training in language learning – a systematic review of available research. Retrieved 2005 from *University of London: PPI-Centre, Social Science Research Unit, Institute of Education*

Hatasa, Y. A. (2002). The Effects of Differential Timing in the Introduction of Japanese Syllabaries on Early Second Language Development in Japanese. *The Modern Language Journal*, 86(3), 349–367.

Hatasa, Y, A., & Tohsaku. Y-H. (1997). "SPOT as a Placement Test." In H. M. Cook, et. al. (Eds.), *New Trends and Issues in Teaching Japanese language and Culture.* Honolulu: University of Hawaii, Second Language Teaching and Curriculum Center, 77–98.

Hayakawa, K. (2010). Chugokugo o bogo to suru nihongo gakushusha ni yoru kanjigo no on'in shori [Phonological processing of kanji compounds by native Chinese speakers learning Japanese]. Nitchuu Gengo Kenkyu to Nihongo Kyoiku [*Journal of Japanese and Chinese Linguistics and Japanese Language Education*]. 3. 100–110.

Hayakawa, K., & Tamaoka, K., (2012). Chugokujin/Kankokujin nihongo gakushusha ni yoru chokaku/shikaku-teiji no gengokan-dokeidogi/gengokan-ikeidogi no niji kanjigo no shori [Processing of auditory-and visually-presented orthographically/semantically similar and different two-kanji compound words by native Chinese and Korean speakers learning Japanese]. Koide kinen nihongo kyoiku kenkyukai ronbunshu [*Koide Memorial Journal of Japanese Language Education*]. 20. 17–32.

Hino, Y., Lupker, S., & Taylor, T. (2012). The role of orthography in the semantic activation of neighbors. *Journal of Experimental Psychology*, *38*, 1259–1273.

Hishinuma, T.(1983). Nihongo to chugokugo no joyo jitai [Commonly-used kanji in Japanese and Chinese]. Chugoku Kenkyu Geppo [*Monthly Journal of Chinese Studies*] 428. 1-20.

Hishinuma, T.(1984). Chugokugo no hyojun jitai to nihon no joyo jitai [The standard orthography in Chinese and the common orthography in Japanese]. Nihongogaku [*Japanese Linguistics*] 3. 32-40.

Horodeck. R. A., (1987) *The Role of Sound in Reading and Writing Kanji*. Unpublished PhD thesis. Cornell University.

Huang, Z. J. (2020) Thirteen years since the first iPhone: A systematic review on the effectiveness of language learning apps on smart devices. Unpublished MSc thesis. University of Oxford.

Kobayashi, N. (2016). Japanese language proficiency assessment with SPOT as primary focus. In M. Minami (Ed.), *Handbook of Japanese Applied Linguistics* (pp. 175–198).

Japan Foundation. (2021) Retrieved August 1, 2023, from https://www.jpf.org.uk/language/

Japan Educational Exchanges and Services (JEES). (2019) Retrieved July 23, 2023, from https://www.jlpt.jp/e/statistics/index.html.

Jeon, E. H., & Yamashita, J. (2014). L2 Reading Comprehension and Its Correlates: A Meta-Analysis. *Language Learning*, 64(1), 160–212.

Jeong, H., Sugiura, M., Sassa, Y., Haji, T., Usui, N., Taira, M., Kawashima, R. (2007). Effect of syntactic similarity on cortical activation during second language processing: A comparison of English and Japanese among native Korean trilinguals. *Human Brain Mapping*, 28(3), 194–204.

Jorden, E. H., & Chaplin, H. I. (1976). *Reading Japanese*. New Haven, CT: Yale University Press.

Joyce, T. (2011). The significance of the morphographic principle for the classification of writing systems. *Written Language & Literacy*, 14(1), 58–81.

Kashiwagi, A. (2011). *Processing relative clauses in first and second language*. PhD thesis. The Ohio Sate University.

Kato, T. (2005). Learning strategies employed by Chinese-background learners in learning Japanese vocabulary. *Japanese Studies*, 25(3), 271–286.

Katz, L., & Frost, R., (1992). Reading in different orthographies. In R. Frost and L. Ka (Eds) *The orthographic depth hypothesis. Orthography, phonology, morphology, and meaning*. Amsterdam: New Holland.

Kayamoto, Y., Yamada, J., & Takashima, H. (1998). The Consistency of Multiple-Pronunciation Effects in Reading: The Case of Japanese Logographs. *Journal of Psycholinguistic Research*, 27(6), 619–637.

Kess, J. F., & Miyamoto, T. (2000). *The Japanese Mental Lexicon: Psycholinguistics Studies of Kana and Kanji Processing*. Amsterdam: John Benjamins.

Kershul, K. (1992). *Japanese in 10 minutes a day*. Seattle, WA: Bilingual Books Inc.

Kirk, B. (2022). *Exploring the relationships between Japanese L2 acquisition rates, study environments and kanji learning strategies*. Unpublished MSc thesis. University of Oxford.

Kirwan, L. (2003). *The role of furigana in Japanese script for second language learners of Japanese.* Unpublished PhD thesis. University of Queensland.

Kirwan, L. (2005). Harnessing furigana to improve Japanese learners' ability to read kanji. *Babel* 40(1) 18-25.

Koda, K. (1989). Effects of L1 orthographic representation on L2 phonological coding strategies. *Journal of Psycholinguistic Research*, 18(2), 201–222.

Koda, K. (1996). L2 Word Recognition Research: A Critical Review. *The Modern Language Journal*, 80(4), 450–460.

Koda, K. (2007). Reading and Language Learning: Crosslinguistic Constraints on Second Language Reading Development. *Language Learning*, 57, 1–44.

Koda, K. (1992). The Effects of Lower-Level Processing Skills on FL Reading Performance: Implications for Instruction. *The Modern Language Journal*, 76(4), 502–512.

Koda, K. (2017). *Learning to Read Japanese*. Cambridge: Cambridge University Press.

Kondo-Brown, K. (2006). How Do English L1 Learners of Advanced Japanese Infer Unknown Kanji Words in Authentic Texts*? Language Learning*, 56(1), 109–153.

Koyama, M. S., Hansen, P. C., & Stein, J. F. (2008). Logographic Kanji versus Phonographic Kana in Literacy Acquisition. *Annals of the New York Academy of Sciences*, 1145(1), 41–55.

Kubota, M., & Toyoda, E. (2001). Learning strategies employed for learning words written in kanji versus kana. *Australian Review of Applied Linguistics*, 24, 1–16.

Li, S. (2019). *The effects of phonics instruction on L2 phonological decoding and vocabulary learning: An experimental trial on Chinese University EFL learners*. Unpublished DPhil Thesis, University of Oxford.

Li, S., & Woore, R. (2021). The effects of phonics instruction on L2 phonological decoding and vocabulary learning: An experimental study of Chinese EFL learners. *System*, 103, 1–12.

Macaro, E. (2006). Strategies for Language Learning and for Language Use: Revising the Theoretical Framework. *The Modern Language Journal*, 90(3), 320–337.

Makino, S., Hatasa, Y., & Hatasa, K. (1998). *Nakama: Japanese communication, context, culture*, *Vol. 1*. Boston: Houghton Mifflin Company.

Matsunaga, S. (1995). *The role of phonological coding in reading kanji: A research report and some pedagogical implications* (Tech. Rep. No. 6). Honolulu: University of Hawaii, Second Language Teaching and Curriculum Center

Matsunaga, S. (1996). The linguistic nature of kanji re-examined: Do kanji represent only meanings? *Journal of the Association of Teachers of Japanese, 30*, 1–22.

Matsunaga, S. (2001). Subvocalization in reading kanji: Can Japanese text be comprehended without it? In H. Nara (Ed.), *Advances in Japanese language pedagogy* (pp. 30–46). Columbus, OH: National East Asian Languages Resource Center.

McLean, S., & Kramer, B. (2015). The creation of a New Vocabulary Levels Test. *Shiken*, 19(2), 1–30.

Milton, J., & Meara, P. M. (1995). How periods abroad affect vocabulary growth in a foreign language. ITL - *International Journal of Applied Linguistics*, 107–108, 17–34.

Miura, A. (1987). America ni okeru nihongo kyooiku no shomondai [Problems in Japanese language teaching in the USA]. *Nihongo Kyoiku*, *63*, 6–19.

Mizumoto, A. (2012). Exploring the Effects of Self-efficacy on Vocabulary Learning strategies. *Studies in Self-Access Learning Journal*, 423–437.

Mizumoto, A., & Takeuchi, O. (2009). Examining the effectiveness of explicit instruction of vocabulary learning strategies with Japanese EFL university students. *Language Teaching Research,* 13(4), 425–449.

Mizumoto, A., & Takeuchi, O. (2012). Adaptation and Validation of Self-regulating Capacity in Vocabulary Learning Scale. *Applied Linguistics*, 33(1), 83–91.

Morita, A., & Tamaoka, K. (2002). Phonological involvement in the processing of Japanese at the lexical and sentence levels. *Reading and Writing*, 15(7–8), 633–651.

Mori, Y. (1998). Effects of First Language and Phonological Accessibility on Kanji Recognition. *The Modern Language Journal*, 82(1), 69–82.

Mori, Y. (2014). Review of Recent Research on Kanji Processing, Learning, and Instruction. *Japanese Language and Literature*, 48(2), 403–430.

Mori, Y., & Nagy, W. (1999). Integration of Information From Context and Word Elements in Interpreting Novel Kanji Compounds. *Reading Research Quarterly*, 34(1), 80–101.

Mori, Y., Sato, K., & Shimizu, H. (2007). Japanese Language Students' Perceptions on Kanji Learning and Their Relationship to Novel Kanji Word Learning Ability. *Language Learning*, 57(1), 57–85.

Mori, Y., Hasegawa, A., & Mori, J. (2021). The trends and developments of L2 Japanese research in the 2010s. *Language Teaching*, 54(1), 90–127.

Nakagawa, A. (1994). Visual and Semantic Processing in Reading Kanji. *Journal of Experimental Psychology: Human Perception and Performance*, 20(4), 864–875.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, I. S. P. (2013). *Teaching and learning vocabulary* Boston: Heinle Cengage Learning

Nation, P., & Meara, P. (2015). Vocabulary. In Schmitt (Ed.), *An Overview of Applied Linguistics* (3rd ed.), pp. 35–54). London: Arnold.

Nesbitt, D. (2009). Achieving unconscious recall of *kanji*: Can rote learning help? *New Zealand Studies in Applied Linguistics, 15*, 61–173.

Neustupný, J, V.(1984). *Introduction to Japanese writing*. Melbourne: Japanese Studies Centre.

Okamoto, S. (1995). “Tasteless” Japanese: Less “Feminine” Speech Among Young Japanese Women. In K. Hall & M. Bucholtz (Eds.), *Gender articulated: language and the socially constructed self.* New York: Routledge.

Okuyama, Y. (2007). CALL Vocabulary Learning in Japanese: Does Romaji Help Beginners Learn More Words? *CALICO Journal*, 24(2), 355–379.

Ono., Hi. (1973) *Japanese grammar.* Tokyo: Hokuseido

Orcan, F. (2020). Parametric or Non-parametric: Skewness to Test Normality for Mean Comparison. *International Journal of Assessment Tools in Education*, 7(2), 255–265.

Otake, Y. (1987). *Japanese word processing as in instructional activity for university students learning Japanese as a foreign language*. Unpublished PhD thesis: The University of Texas at Austin.

Oxford R. L. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House.

Packard, J. L. (1990). Effects of Time Lag in the Introduction of Characters into the Chinese Language Curriculum. *The Modern Language Journal*, 74(2), 167–175.

Paxton, S., & Svetanant, C. (2013). Tackling the Kanji hurdle: Investigation of Kanji learning in Non-Kanji background learners. *International Journal of Research Studies in Language Learning*, 3(3), 89–104.

Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(4), 633–643.

Plonsky, L. (2011). The Effectiveness of Second Language Strategy Instruction: A Meta-analysis. *Language Learning*, 61(4), 993–1038.

Pye, M. (1984). *The Study of Kanji: A Handbook of Japanese Characters*. Tokyo: Hokuseido Press.

Ranalli, J. (2012). Alternative Models of Self-regulation and Implications for L2 Strategy Research. *Studies in Self-Access Learning Journal*, 357–376.

Razali, & Wah. (2011). Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21–33.

Reddit. (n.d.). Retrieved July 23, 2023, from https://www.reddit.com.

Rose, H. (2017). *The Japanese Writing System Challenges, Strategies and Self-Regulation for Learning Kanji*. Bristol: Multilingual Matters.

Rose, H. (2010). *Kanji learning of Japanese language learners on a year-long study exchange program at a Japanese university: An investigation of strategy use, motivational control and self-regulation.* Unpublished PhD thesis. University of Sydney

Rose, H. (2012). Reconceptualizing Strategic Learning in the Face of Self-Regulation: Throwing Language Learning Strategies out with the Bathwater. *Applied Linguistics*, 33(1), 92–98.

Rose, H., & Harbon, L. (2013). Self-Regulation in Second Language Learning: An Investigation of the Kanji-Learning Task. *Foreign Language Annals*, 46(1), 96–107.

Rose, H., Briggs, J. G., Boggs, J. A., Sergio, L., & Ivanova-Slavianskaia, N. (2018). A systematic review of language learner strategy research in the face of self-regulation. *System*, 72, 151–163.

Sakamoto, T. (2015). Processing of syntactic and semantic information in the human brain: Evidence from ERP studies in Japanese. In M. Nakayama (Ed.), *Handbook of Japanese Psycholinguistics* (pp. 457–510). Berlin: De Gruyter.

Samimy, K. K., & Tabuse, M. (1992). Affective Variables and a Less Commonly Taught Language: A Study in Beginning Japanese Classes. *Language Learning*, 42(3), 377–398.

Sawasaki, K., & Kashiwagi-Wood, A. (2015). Issues in L2 Japanese sentences processing: Similarities/differences with L1 and individual differences in working memory. In M. Nakayama (Ed.), *Handbook of Japanese Psycholinguistics*. Berlin: De Gruyter.

Schmidt, R. W. (1990). The Role of Consciousness in Second Language Learning. *Applied Linguistics*, 11(2), 129–158.

Seeley, C. (1991*) A History of Writing in Japan*, Leiden: E. J. Brill.

Shimizu, H., & Green, K. E. (2002). Japanese Language Educators' Strategies for and Attitudes toward Teaching Kanji. *The Modern Language Journal*, 86(2), 227–241

Smolkowski, K. (2019). *Gain score analysis*. Retrieved July 22, 2023, from http://homes.ori.org/keiths/Tips/Stats_GainScores.html.

Sonbul, S., & Schmitt, N. (2009). Direct teaching of vocabulary after reading: is it worth the effort? *ELT Journal*, 64(3), 253–260.

Stanovich, K. E. (1980). Toward an Interactive-Compensatory Model of Individual Differences in the Development of Reading Fluency. *Reading Research Quarterly*, 16(1), 32.

Tan, L. H., Spinks, J. A., Feng, C., Siok, W. T., Perfetti, C. A., Xiong, J., Fox, P., T., Gao, J. (2003). Neural systems of second language reading are shaped by native language. *Human Brain Mapping*. 18(1), 158-166.

Tanaka, M. (2015). Japanese Kanji Word Processing for Chinese Learners of Japanese: A Study of Homophonic and Semantic Primed Lexical Decision Tasks. *Theory and Practice in Language Studies*, 5(5), 900–905.

Tamaoka, K. (n.d.). Romaji wa Yuzai ka? Romaji o Baikai to shita Eigokei Nihongo Gakushusha no bogo kara no onin kannsho nozofuku ni kansuru kennsho [*Is Romaji Guilty? An Examination of Phonological Interference in English-speaking Learners of Japanese through the Romaji Script*]

Tamaoka, K. (1997). Chugokugo to eigo o bogo to suru nihongo gakushusha no kanji hyoki goi no shori horyaku [The processing strategy of words presented in kanji and kana by Chinese and English speakers learning Japanese] Gengo Bunka Kenkyu [*Studies in Language and Literature*] 17(1): 65-77.

Tamaoka, K. (2000). The effects of stroke order and radicals on the knowledge of Japanese kanji orthography, phonology and semantics. *Psychologia* : *An International Journal of Psychology in the Orient* (プシコロギア, 東洋国際心理学会誌) 43. 199-210.

Tamaoka, K. (2014). The Japanese Writing System and Lexical Understanding. Japanese Language and Literature; Special Section: New Perspectives on the Japanese Writing System and Reading Japanese as L1 and L2. *Journal of the American Association of Teachers of Japanese* 48(2), 431–471.

Tamaoka, K. (2015). Processing of the Japanese language by native Chinese speakers. In M. Nakayama (Ed.), *Handbook of Japanese psycholinguistics* (pp. 583–632). Berlin: Walter De Gruyter.

Tamaoka, K., Makioka, S., Sanders, S., & Verdonschot, R. G. (2017). www.kanjidatabase.com: a new interactive online database for psychological and linguistic research on Japanese kanji and their compound words. *Psychological Research*, 81(3), 696–708.

Taylor, I., & Taylor, M. (1995). *Writing and literacy in Chinese, Korean, and Japanese*. Amsterdam: John Benjamins Publishing.

Teng, L. S., & Zhang, L. J. (2016). A Questionnaire-Based Validation of Multidimensional Models of Self-Regulated Learning Strategies. *The Modern Language Journal*, 100(3), 674–701.

Tono, Y., Yamazaki, M., & Maekawa, K. (2013). *A Frequency Dictionary of Japanese, Core vocabulary for learners.* Abingdon: Routledge.

Tseng, W., & Schmitt, N. (2008). Toward a Model of Motivated Vocabulary Learning: A Structural Equation Modelling Approach. *Language Learning*, 58(2), 357–400.

TTBJ The Tsukuba Test Battery, University of Tsukuba (n.d.). Retrieved July 23, 2023, from https://ttbj.cegloc.tsukuba.ac.jp/en/p1.html.

Walton, A, R. (1992). *Expanding the Vision of Foreign Language Education: Enter the Less Commonly Taught Languages.* NFLC Occasional Papers. Washington, DC: National Foreign Language Center.

Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English semantic processing: A comparison of Korean and Chinese English L2 learners. *Cognition, 87,* 129– 149.

WaniKani.(n.d.). Retrieved July 23, 2023, from https://www.wanikani.com.

Webb, S. (2007b). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, *28*(1), 46–65.

Weinstein, C. E., Acee, T. W., & Jung, J. (2011). Self-regulation and Learning Strategies. *New Directions for Teaching and Learning*, 126, 45-53.

Wikipedia [Japan 日本]. Retrieved July 28, 2023, May, 13. From [https://ja.wikipedia.org/wiki/日本#交通.

Woodrow, L. (2005). The Challenge of Measuring Language Learning Strategies. *Foreign Language Annals*, 38(1), 90–98.

Woore, R. (2011). *Investigating and developing beginner learners' decoding proficiency in second language French*: Unpublished DPhil thesis. University of Oxford.

Yamada, J. (1992). Asymmetries of reading and writing Kanji by Japanese children. *Journal of Psycholinguistic Research*, 21, 563–580.

Yamato, Y, & K. Tamaoka. (2009). Chugokujin Nihongi gakushusha no Nihongi kanji no short ni okeru boo no eikyo [The effects of mother-tongue on processing of Japanese kanji-compound words by native Chinese speakers learning Japanese]. Kotoba no Kagaku [*Studia Linguistica]* 22, 117-135

Yano, M. (2012). Chugoku kanji bunka to nihon [*Kanji culture in China and Japan*]. Surugadai Press.

Yokosawa, K., & Umeda, M. (1988). *Processes In Human Kanji-word Recognition.* Proceedings of the 1988 IEEE International Conference on Systems, Man, and Cybernetics, 1, 377–380.

Ziegler, N. (2015). The predictive value of the self-regulating capacity in vocabulary learning scale. *Applied Linguistics*, 36(5), 1-9.

Zimmerman, B. J. (1990). Self-Regulated Learning and Academic Achievement: An Overview. *Educational Psychologist*, 25(1), 3–17.

APPENDICIES

CONTENTS OF APPENDICIES

Background screening survey…….. A

Participants by location, proficiency and L1 background……..B

SPOT90 scores of all subjects tested…… C

SPOT90 score conversion table (TTBJ)…… D

Online marketing poster……..E

Marketing post on Reddit…… F

Gorilla experiment flowchart……G

Pilot vocabulary items (10)……..H

Target vocabulary items (16)…… I

SRCKanji survey (adapted from Tseng et. al. (2006) …… J

Examples of romaji transliteration systems..…… K

Boxplot of SPOT90 scores by proficiency level and treatment group……..L

Boxplot of Test scores by proficiency level……M

Boxplot of test scores ……N

Z-score calculations of SPOT90 ……O

Z-score calculations of test scores, by proficiency…… P

Z-score calculations of test scores, by treatment group…… Q

ANOVA and ANCOVA output from SPSS……R

Kruskal-Wallis output from SPSS……S

CUREC approval…… T

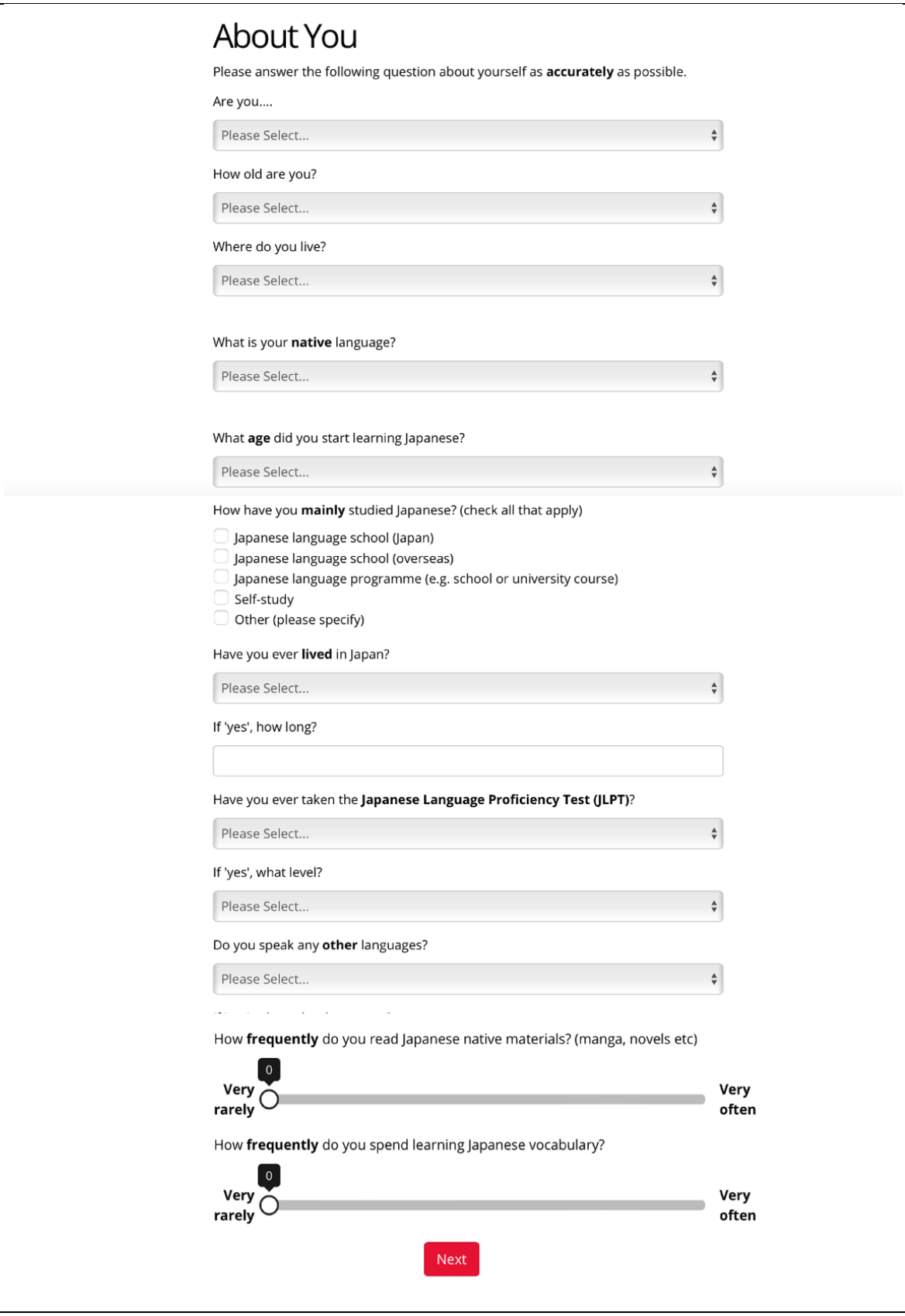
Participant online consent form……U

Gorilla online survey consent form……V

Online recruitment poster……W

Dual language business cards…… X

Appendix A: Background screening survey



Participant background survey on Gorilla (https://app.gorilla.sc).

Appendix B: Participants by location, proficiency and L1 background

Country of Residence	Number of Participants	L2 Proficiency		L1 Background	
		Beginner	Intermediate	Alphabetic	Logographic
UK	30	15	15	23	7
USA	13	6	7	13	-
Canada	2	1	1	1	1
Australia	1	-	1	1	-
Ireland	5	4	1	5	-
Norway	3	1	2	3	-
France	1	-	1	1	-
Germany	1	-	1	1	-
The Netherlands	1	1	-	1	-
Poland	1	-	1	1	-
China	2	2	-	-	2
Japan	15	6	9	15	-
TOTAL	**75**	**36**	**39**	**65**	**10**

Appendix C: SPOT90 scores of all subjects tested

Descriptive Statistics

	N	Minimum	Maximum	Mean		Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Std. Error	Statistic	Std. Error
SPOT90 scores (all)	121	13	85	58.77	1.567	17.237	-.572	.220	-.346	.437

*Beginner 30-55, Intermediate 56-80. Out of a possible maximum score 90 (TTBJ, University of Tsukuba). Available at https://ttbj.cegloc.tsukuba.ac.jp/en/p1.html

Appendix D: SPOT90 score conversion table (TTBJ)

■SPOT90

Total score	Ability assessment*	Explanation	Correspondence with the Japanese Language Proficiency Test**
0–30	Introductory level	Learner has studied virtually no Japanese.	None
31–55	Beginner level	Can understand basic Japanese in everyday situations if spoken to slowly.	N4, N5
56–80	Intermediate level	Can understand Japanese to a certain degree in everyday situations when spoken to at a natural speed.	N3, N2
81–90	Advanced level	Can understand Japanese in a broad range of situations when spoken to at a natural speed.	N1

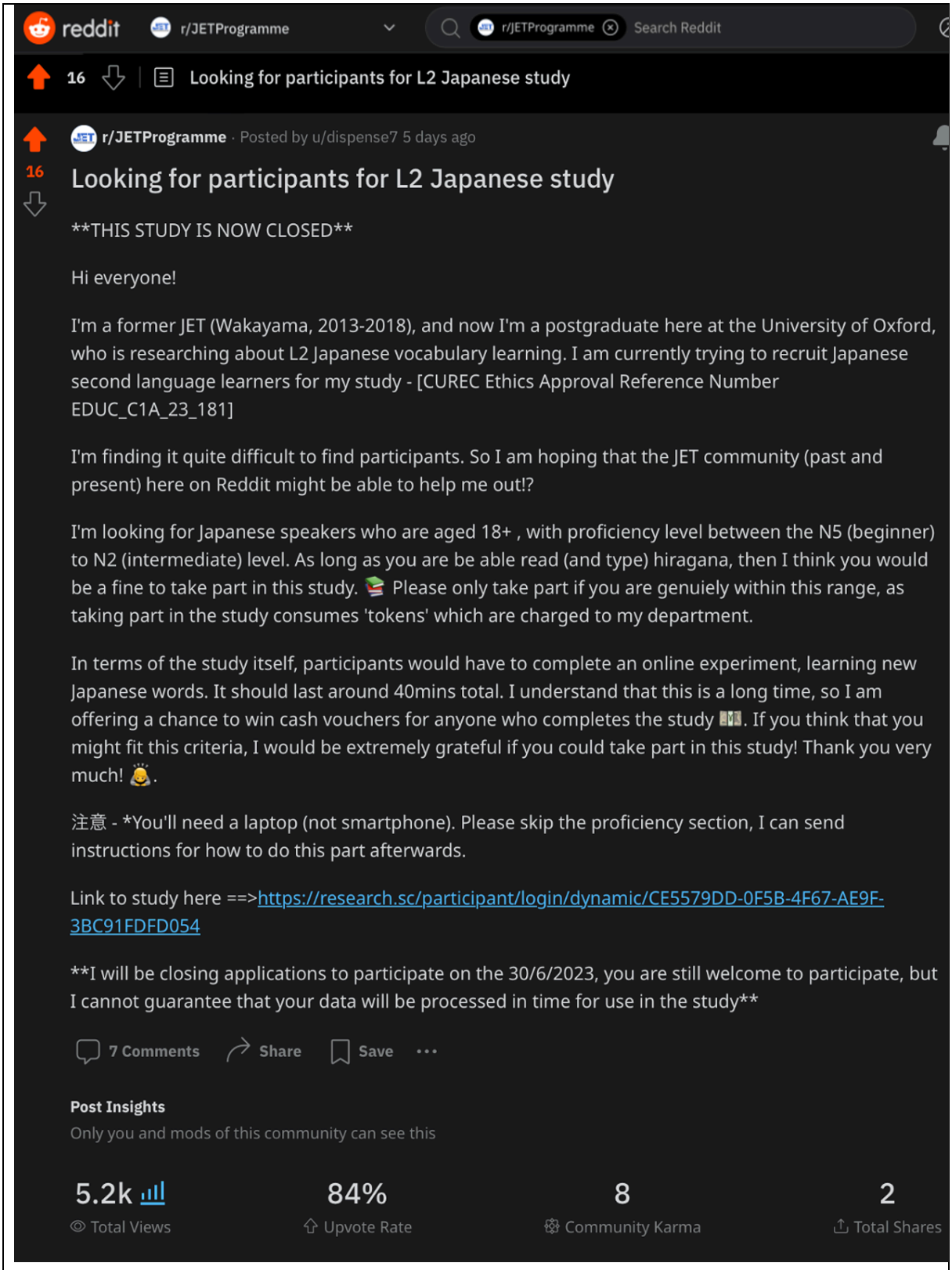
University of Tsukuba SPOT90 test scores interpretation chart and correspondence to the Japanese Language Proficiency Test (JLPT). (https://ttbj.cegloc.tsukuba.ac.jp/en/p1.html).

Appendix E Online marketing poster



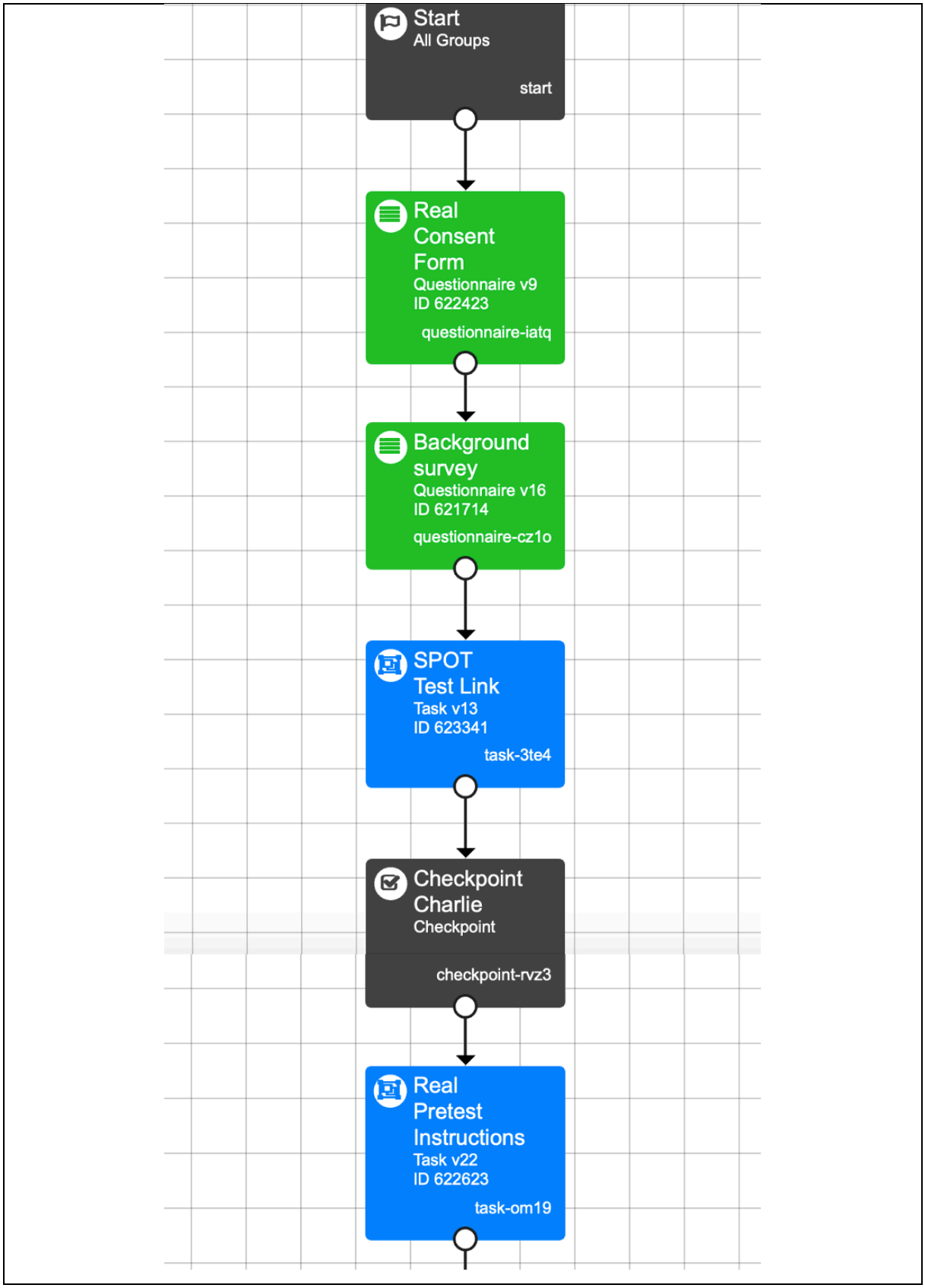
Online marketing poster used for participant recruitment. QR code designed online with 'QR Code Monkey', free QR Code generator. Available at [https://www.qrcode-monkey.com].

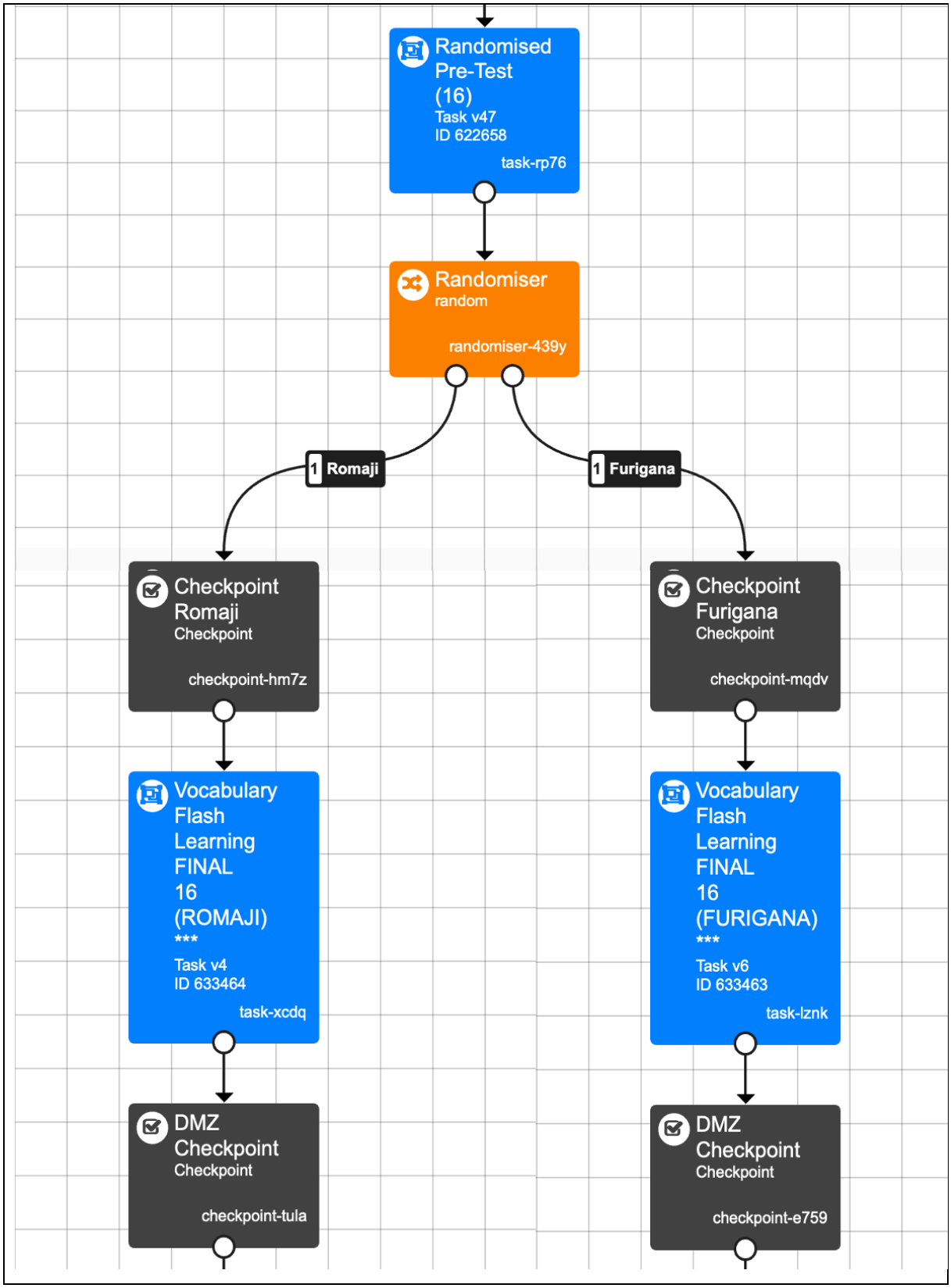
Appendix F Marketing post on Reddit

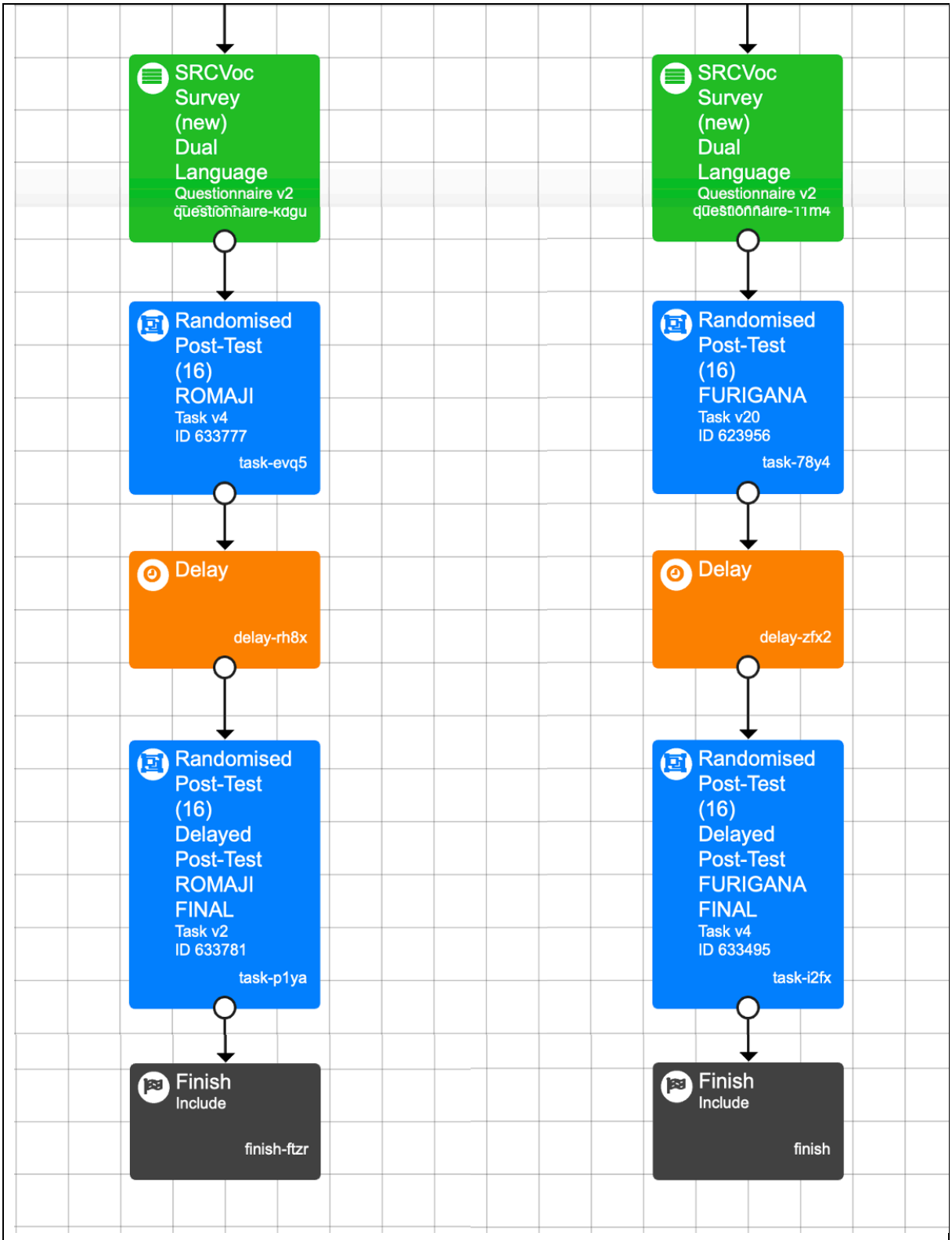


Marketing post on Reddit r/JETProgramme, 5.2k total views in five days. Posted on on 26/6/2023, available at [https://www.reddit.com/r/JETProgramme/comments/14jhl6o/looking_for_participants_for_l2_japanese_study/].

Appendix G Gorilla experiment flowchart




Randomised Pre-Test (16)
Task v47
ID 622658
task-rp76
Randomiser
random
randomiser-439y
1 Romaji
1 Furigana
Checkpoint Romaji
Checkpoint
checkpoint-hm7z
Checkpoint Furigana
Checkpoint
checkpoint-mqdv
Vocabulary Flash Learning FINAL 16 (ROMAJI)
***
Task v4
ID 633464
task-xcdq
Vocabulary Flash Learning FINAL 16 (FURIGANA)
***
Task v6
ID 633463
task-lznk
DMZ Checkpoint
Checkpoint
checkpoint-tula
DMZ Checkpoint
Checkpoint
checkpoint-e759



Experiment built online using Gorilla (www.gorilla.sc), showing 'between-subject' experiment design, with built in automatic 50:50 randomiser, two treatment and assessment pathways, and 24hr delayed post-test. (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020).

Appendix H Pilot vocabulary items (10)

Target kanji compounds	Word frequency (0-5000) and Word type	Reading (hiragana & romaji)	Meaning in English	Example sentence
突然	852 adverb	とつぜん totsuzen	sudden	電車(でんしゃ)か突然(とつぜん)止まった。 The train stopped suddenly.
旧跡	n/a noun	きゅうせき kyuuseki	historic spot	大勢(おおぜい)の人がその旧跡(きゅうせき)を訪(おとず)れる。 A large number of people visit the historic site.
責任	951 noun	せきにん sekinin	responsibility	失敗の責任は私が撮ります。 I take responsibility for my mistake.
他界	n/a noun	たかい takai	death	祖父は５年前に他界した。My grandfather died five years ago.
巡歴	n/a noun	じゅんれき junreki	trip	彼女は巡歴中に写真を撮りました。 She took pictures during her trip.
葬式	n/a noun	そうしき soushiki	funeral	その夫人の葬式は地元の教会行われた。 The lad's funeral was held at the local church.
動揺	4999 noun	どうよう douyou	agitated	彼女は激しく動揺した。 She was violently agitated.
混乱	3036 noun	こんらん konran	chaos	混乱して、何も考えられない。I'm confused, so I can't think about anything.
記憶	652 noun	きおく kioku	record	それは古い過去の記憶だった。 It was an old memory of the past.
派手	3702	はで	showy	彼女の服装はいつも派手だ。

	noun	hade		Her outfits are always flashy
Example Japanese/English sentences from *A Frequency Dictionary of Japanese, Core vocabulary for learners* (Tono, Yamazaki, & Maekawa, 2013)				

Appendix I Target vocabulary items (16)

Target kanji compounds	Word frequency (0-5000) and Word type	Reading (hiragana & romaji)	Meaning in English	Example sentence
飛行機.	876 noun	ひこうき hikouki	airplane	飛行機(ひこうき)が離陸(りりく)した。 The plane took off.
感染	3623 noun	かんせん kansen	infection	インフルエンザに感染(かんせん)した。 I got infected with flu.
倒産	4203 noun	とうさん tousan	bankruptcy	その会社(かいしゃ)は倒産(とうさん)した。 The company went bankrupt.
寄与	4725 noun	きよ kiyo	contribution	彼(かれ)は文化(ぶんか)の発展(はってん)に大(おお)いに寄与(きよ)した。 He made a large contribution to cultural development.
除く	1031 verb	のぞく nozuku	to remove, to exclude	彼(かれ)の作文(さくぶん)は多少(たしょう)の間違(まちが)いを除(のぞ)いてすばらしかった。 His composition was excellent except for a few errors.
蹴る	3823 verb	ける keru	to kick	思(おも)い切りボールを蹴(け)った。 He kicked the ball hard.
漏らす	4000 verb	もらす morasu	to betray	彼女(かのじょ)は彼(かれ)の秘密(ひみつ)を漏(も)らした。 She betrayed his secret.
率いる	4248 verb	ひきいる hikiiru	to lead	私(わたし)は学生(がくせい)を率(ひき)いて博物館(はくぶつかん)へ行(い)った。

				I led my students to the museum.
偶然	2192 adverb	ぐうぜん guuzen	by accident	偶然(ぐうぜん)、駅(えき)で彼女(かのじょ)と会(あ)った。 I met her by chance at the station.
何と	2419 adverb	なんと nanto	How, what	なんと素晴(すば)らしい歌(うた)だろう。 What a beautiful song!
度々	3825 adverb	たびたび tabitabi	often	度々(たびたび)お電話(でんわ)してすみません。 I am sorry to call you so often.
散々	4733 adverb	さんざん sanzan	severely, repeatedly	私(わたし)は遅(おそ)く帰(かえ)ってきて親(おや)に散々叱(さんざんしか)られた。 I came home late and was scolded severely by my parents.
素直	2210 na-adjective	すなお sunao	obedient	彼女(かのじょ)はとても素直(すなお)です。 She is very obedient.
著しい	2468 i-adjective	いちじるしい ichijirushii	significant, remarkable, striking,	あの会社(かいしゃ)の成長(せいちょう)は著(いちじる)しい。 The improvement of that company is tremendous.
貧しい	3516 i-adjective	まずしい mazushii	poor	彼(かれ)は貧(まず)しい人々(ひとびと)を救済(きゅうさい)する活動(かつどう)をしている。 He is active in helping people in poverty.
緩やか	4783 i-adjective	ゆるやか yuruyaka	gentle, soft, mild	私(わたし)の家(いえ)の前(まえ)は緩(ゆる)やかな上(のぼ)り坂(ざか)になっている。 There is a gentle uphill slope in front of my house.
Target Japanese kanji word items and Japanese/English example sentences from: *A Frequency Dictionary of Japanese, Core vocabulary for learners* (Tono, Yamazaki, & Maekawa, 2013)				

Appendix J SRCKanji survey (adapted from Tseng et. al. (2006)

Self-regulation and L2 Japanese kanji learning questionnaire

自我调节与第二语言日语词汇学习。

Please reflect on your own experiences about studying Japanese kanji and answer the following questions.

请回顾你自己学习日语词汇的经验，并回答以下问题

Commitment control

承诺控制

• When learning kanji, I have special techniques to achieve my learning goals.
学习汉字时，我有特殊的技巧来实现我的学习目标。

strongly disagree	1	2	3	4	5	6	strongly agree

• When learning kanji, I believe I can achieve my goals more quickly than expected.
在学习汉字时，我相信我能比预期更快地实现我的目标。

strongly disagree	1	2	3	4	5	6	strongly agree

• When learning kanji, I persist until I reach the goals that I make for myself.
在学习汉字时，我会坚持不懈，直到达到我为自己设定的目标。

strongly disagree	1	2	3	4	5	6	strongly agree

• I believe that I can overcome all the difficulties relating to learning new kanji learning goals.
我相信我可以克服与学习新汉字目标相关的所有困难。

strongly disagree	1	2	3	4	5	6	strongly agree

Satiation control

饱食控制

• Once the novelty of leaning kanji is gone, I easily become impatient with it.
一旦学习汉字的新鲜感消失，我很容易对此变得不耐烦

strongly disagree	1	2	3	4	5	6	strongly agree

• During the process of learning kanji, I am confident that I can overcome any sense of boredom.
在学习汉字的过程中，我有信心能克服任何无聊的感觉。

strongly disagree	1	2	3	4	5	6	strongly agree

• During the process of learning kanji, I feel satisfied with the ways I eliminate boredom.
在学习汉字的过程中，我对自己消除无聊的方式感到满意。

strongly disagree	1	2	3	4	5	6	strongly agree

• When feeling bored with learning kanji, I know how to regulate my mood in order to invigorate the learning process.

当感到学汉字汇变得无聊时，我知道如何调节心情以激发学习过程。

strongly disagree 1 2 3 4 5 6 strongly agree

Emotion control

情绪控制

• When I feel stressed about learning kanji, I know how to reduce this stress.

当我感到学习汉字的压力时，我知道如何减轻这种压力

strongly disagree 1 2 3 4 5 6 strongly agree

• I feel satisfied with the methods I use to reduce the stress of kanji learning.

我对我用来减轻汉字学习压力的方法感到满意

strongly disagree 1 2 3 4 5 6 strongly agree

• When I feel stressed about kanji learning, I simply want to give up.

当我感到汉字学习的压力时，我只是想要放弃。

strongly disagree 1 2 3 4 5 6 strongly agree

• When I feel stressed about my kanji learning, I cope with this problem immediately.

当我感到汉字学习的压力时，我会立即应对这个问题

strongly disagree 1 2 3 4 5 6 strongly agree

Environment control

环境控制策略

• When I am studying kanji and the learning environment becomes unsuitable, I try to sort out the problem

当我在学习汉字时，如果学习环境变得不适合，我会尝试解决这个问题

strongly disagree 1 2 3 4 5 6 strongly agree

• When learning kanji, I know how to arrange the environment to make learning more efficient.

在学习汉字时，我知道如何调整学习环境，使学习更加高效

strongly disagree 1 2 3 4 5 6 strongly agree

• When learning kanji, I am aware that the learning environment affects my learning.

在学习汉字时，我意识到学习环境会影响我的学习。

strongly disagree 1 2 3 4 5 6 strongly agree

• When I study kanji, I look for a good learning environment.

在我学习汉字的时候，我会寻找一个好的学习环境。

strongly disagree 1 2 3 4 5 6 strongly agree

Metacognitive strategies

元认知策略

• When learning kanji, I have special techniques to keep my concentration focussed
当学习汉字时，我有特殊的技巧来保持注意力集中

strongly disagree 1 2 3 4 5 6 strongly agree

• When learning kanji, I think my methods of controlling my concentration are effective.
在学习汉字时，我认为我控制注意力的方法很有效。

strongly disagree 1 2 3 4 5 6 strongly agree

• When it comes to learning kanji, I think my methods of procrastination are effective.
谈到学习汉字时，我认为我拖延的方法很有效。

strongly disagree 1 2 3 4 5 6 strongly agree

• When it comes to learning kanji, I have special techniques to prevent procrastinating.
谈到学习汉字时，我有特殊的技巧来防止拖延。

strongly disagree 1 2 3 4 5 6 strongly agree

End of survey. Thank you

调查结束，谢谢你

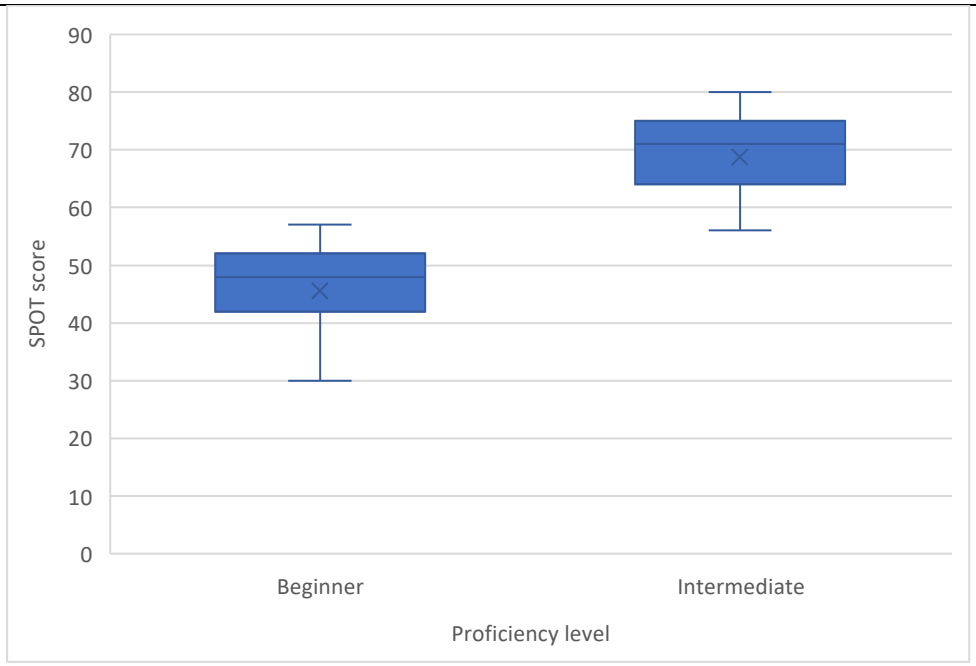
Appendix K Examples of romaji transliteration systems

Hepburn	Kunreishiki	Nihonshiki
shi	si	si
sha	sya	sya
shu	syu	syu
sho	syo	syo
ji	zi	zi
ja	zya	zya
ju	zyu	zyu
jo	zyo	zyo
chi	ti	ti
cha	tya	tya
chu	tyu	tyu
cho	tyo	tyo
ji	zi	di
ja	zya	dya
ju	zyu	dyu
jo	zyo	dyo
tsu	tu	tu
zu	zu	du
fu	hu	hu
o	o	wo

Appendix J Shows the transcription differences between Hepburn, Kunreishiki, and Nihonshiki romanisation systems of Romaji. 9from Kess & Miyamoto, 2000; p.116).

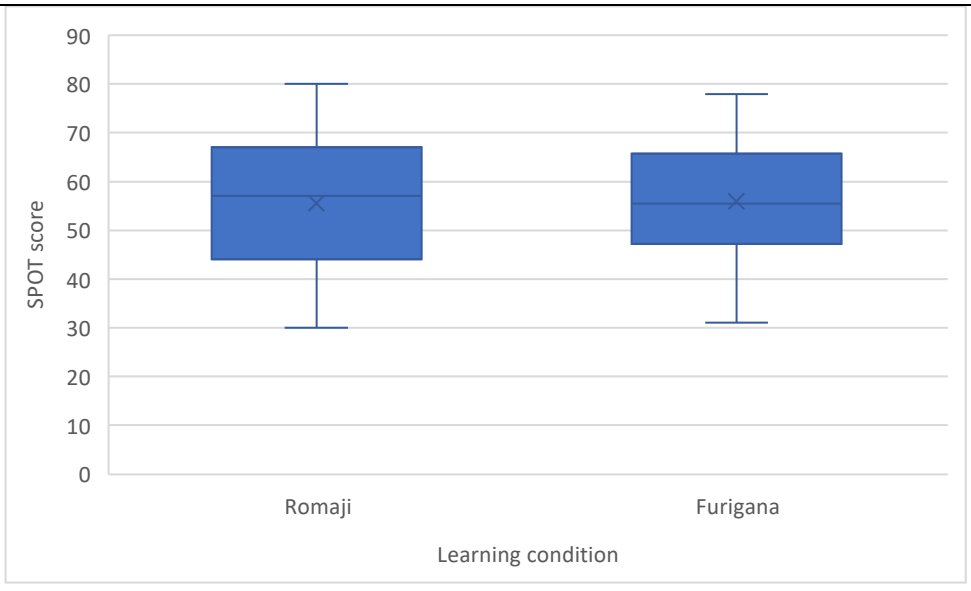
Appendix L Boxplot of SPOT90 scores by proficiency and treatment group

Boxplot of SPOT90 scores by proficiency level



Box-whiskers plot showing no outliers or extreme cases found within the full SPOT90 dataset (N=75)

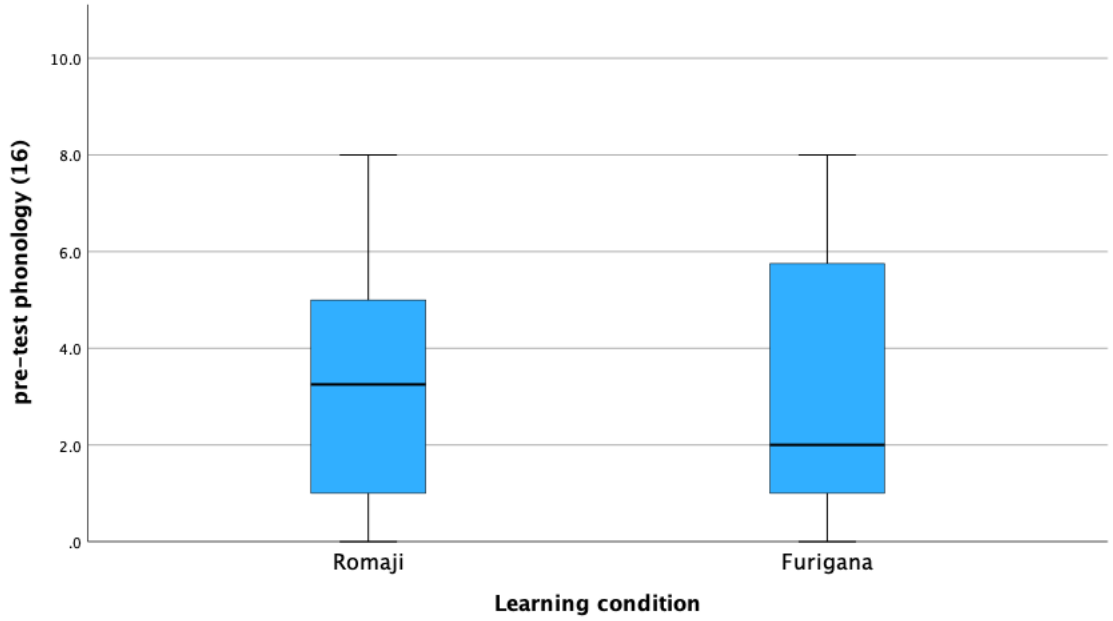
Boxplot of SPOT90 scores by treatment group



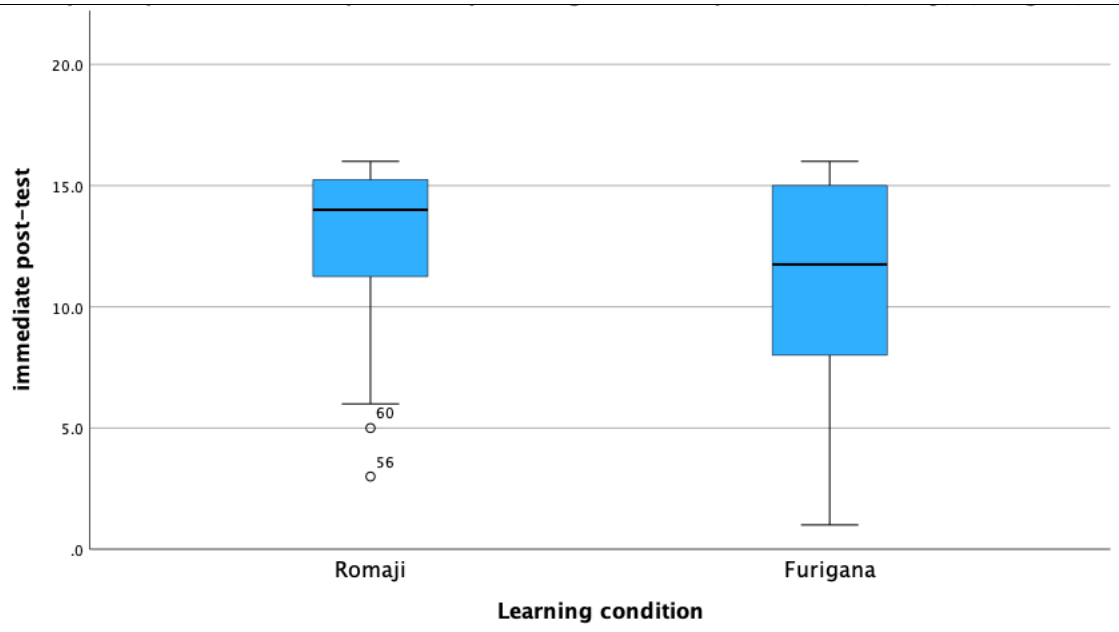
Box-whiskers plot showing no outliers or extreme cases found within SPOT90 scores in each treatment group; Romaji (A) N=35, Furigana (B) N=40

Appendix M Boxplot of test scores by treatment group

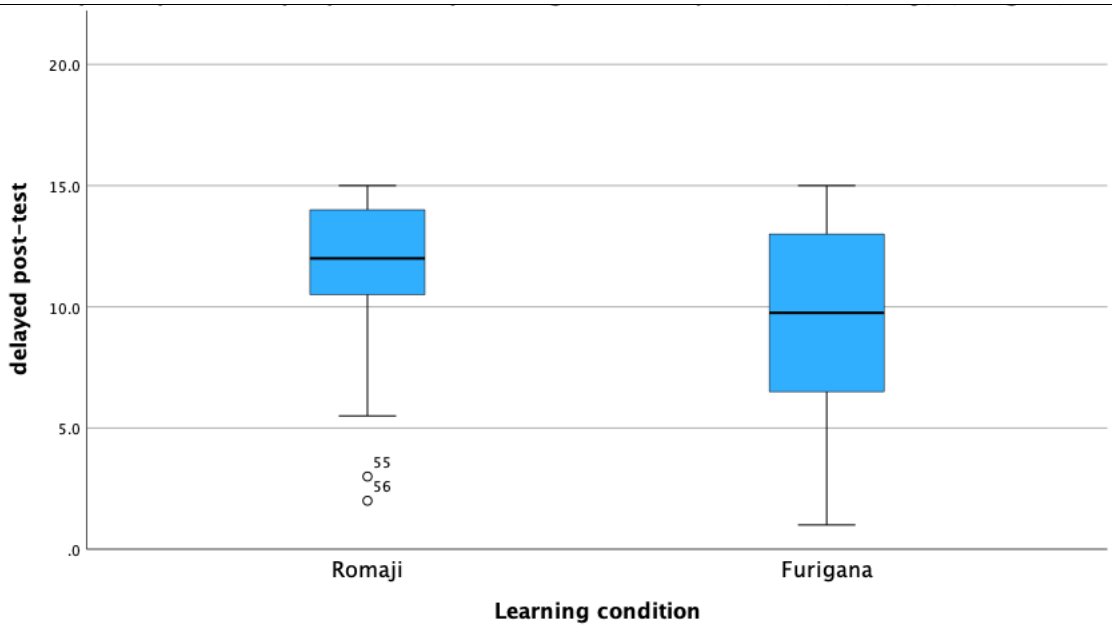
Pre-test



Immediate post-test*



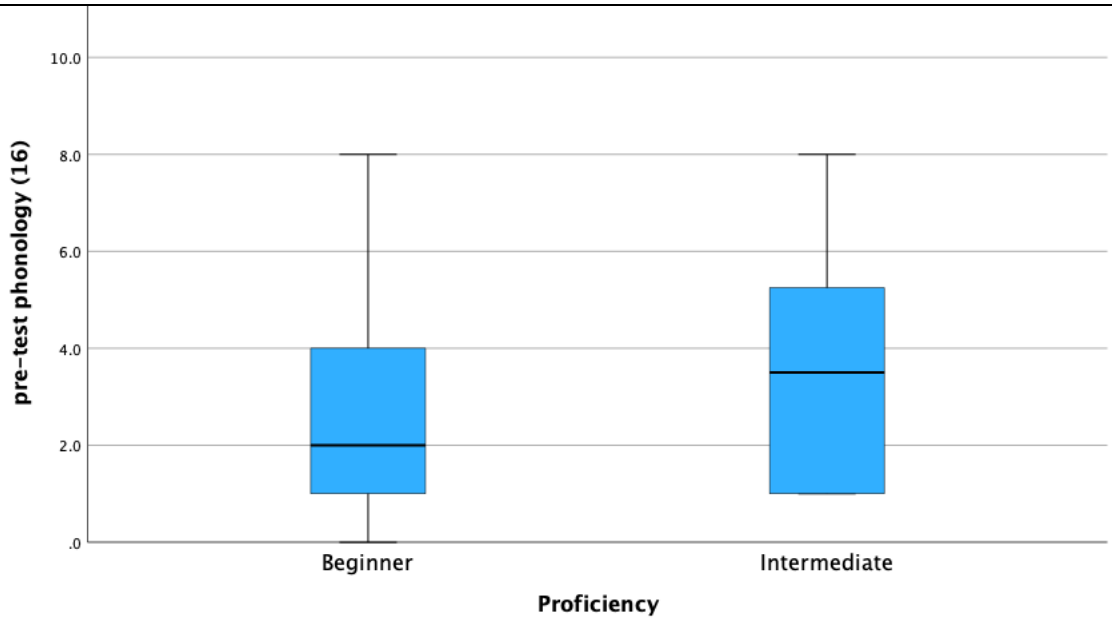
Delayed post-test*



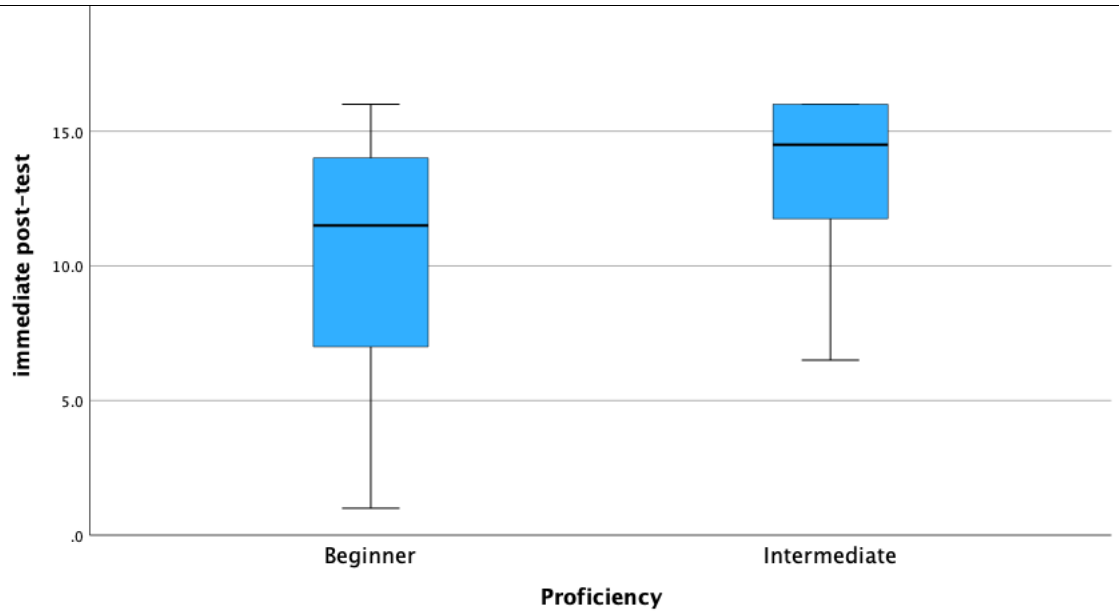
Boxplots for the pre, post, and delayed post-test scores by learning condition. *Three outliers in the romaji learning condition were discovered. Outlier #56 was removed from the dataset. Outliers #55 and #60 were windsorized, to the next non-outlier value (Field, 2018).

Appendix N Boxplot of test scores by proficiency level

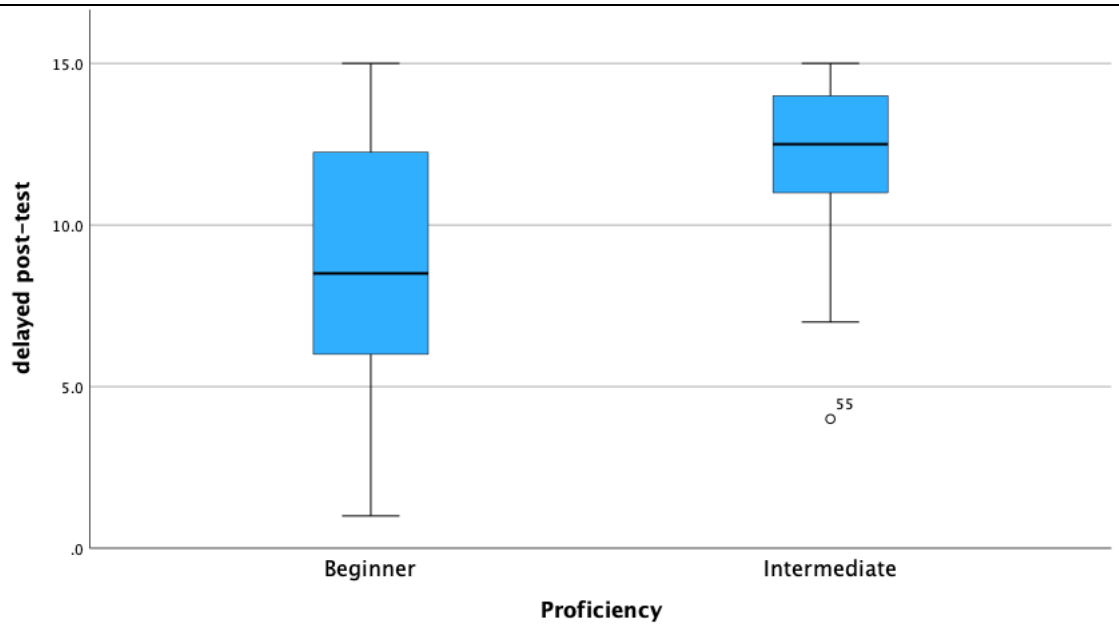
Pre-test



Immediate post-test



Delayed post-test*



Boxplots for the pre, post, and delayed post-test scores by L2 proficiency level (beginner & intermediate). *One outlier #56 in the intermediate group post-test was windsorized, to the next non-outlier value (Field, 2018).

Appendix O SPOT90, Z-score calculations

Z-score calculation formula (Field, 2018)	
Z-skew = (Skew-0)/Standard Error (SE) of skew	Z-kurtosis = (Kurtosis-0)/Standard Error (SE) of kurtosis

Values greater than ±1.96, 2.58, and 3.29 are significant at $p < 0.05$, $p < 0.01$, and $p < 0.001$ respectively. If absolute value is higher than these scores, then violations of normality is significant at the corresponding values. (Field, 2018)

Overall SPOT90 Z-scores

L2 proficiency	Z-score of skewness calculation	Z-score of kurtosis calculation
Beginner (N=36)	Z-skew = (-.561-0)/ .393	Z-kurtosis = (-.813-0)/ .768
Z-score	**1.427**	**1.082**
Intermediate (N=39)	Z-skew = (.095-0)/ .378	Z-kurtosis = (-1.170-0)/ .741
Z-score	**0.251**	**1.578**

SPOT90 Z-scores by treatment group

Treatment	Z-score of skewness calculation	Z-score of kurtosis calculation
Romaji (A) (N=35)	Z-skew = (-.126-0)/ .398	Z-kurtosis = (-.996-0)/ .778
Z-score	**0.316**	**1.280**
Furigana (B) (N=40)	Z-skew = (-.152-0)/ .374	Z-kurtosis = (-.701-0)/ .733
Z-score	**0.406**	**0.956**

Appendix P Test scores, Z-score calculations by treatment group

Z-score calculation formula (Field, 2018)	
Z-skew = (Skew-0)/Standard Error (SE) of skew	Z-kurtosis = (Kurtosis-0)/Standard Error (SE) of kurtosis
Values greater than ±1.96, 2.58, and 3.29 are significant at $p < 0.05$, $p < 0.01$, and $p < 0.001$ respectively. If absolute value is higher than these scores, then violations of normality is significant at the corresponding values. (Field, 2018)	

Pre-test

Pre-test Z-score of Skewness and Kurtosis calculation

Romaji (A) (N=34)			Furigana (B) (N=40)	
Z-skew = (.524-0)/ .404	Z-kurtosis = (-.644.-0)/ .788		Z-skew = (.670-0)/ .374	Z-kurtosis = (-1.072-0)/ .733
1.297	**-0.817**		**1.799**	**-1.462**

Post-test

Immediate post-test Z-score of Skewness and Kurtosis calculation				
Romaji (A) (N=34)			**Furigana (B) (N=40)**	
Z-skew = (-1.095-0)/ .403	Z-kurtosis = (.099)/ .788		Z-skew = (-.731)/.374	Z-kurtosis = (-.381)/ .733
-2.717*	**0.125**		**-1.954**	**-0.519**

Delayed post-test

Delayed post-test Z-score of Skewness and Kurtosis calculation				
Romaji (A) (N=24)			**Furigana (B) (N=32)**	
Z-skew = (-.913-0)/ .472	Z-kurtosis = (-.400-0)/ .918		Z-skew = (-.346-0)/ .414	Z-kurtosis = (-.902-0)/ .809
-1.934	**-0.435**		**-0.836**	**-1.114**

Appendix Q Test scores, Z-score calculations by proficiency

Z-score calculation formula (Field, 2018)	
Z-skew = (Skew-0)/Standard Error (SE) of skew	Z-kurtosis = (Kurtosis-0)/Standard Error (SE) of kurtosis
Values greater than ±1.96, 2.58, and 3.29 are significant at $p < 0.05$, $p < 0.01$, and $p < 0.001$ respectively. If absolute value is higher than these scores, then violations of normality is significant at the corresponding values. (Field, 2018)	

Pre-test

Pre-test Z-score of Skewness and Kurtosis calculation				
Beginner (N=35)			**Intermediate (N=39)**	
Z-skew = (.828-0)/ .398	Z-kurtosis = (-.603.-0)/ .778		Z-skew = (.405-0)/ .378	Z-kurtosis = (-1.147-0)/ .741

-2.080*	**-0.775**		**1.0714**	**-1.547**

Post-test

Immediate post-test Z-score of Skewness and Kurtosis calculation Beginner (N=36), Intermediate (N-39)				
Beginner (N=35)			**Intermediate (N=39)**	
Z-skew = (-.412-0)/ .398	Z-kurtosis = (-.966)/ .778		Z-skew = (-1.056/ .378	Z-kurtosis = (-.544)/ .741
1.035	**-1.242**		**-2.793***	**-0.734**

Delayed post-test

Delayed post-test Z-score of Skewness and Kurtosis calculation Beginner (N=36), Intermediate (N-39)				
Beginner (N=28)			**Intermediate (N=28)**	
Z-skew = (-.099.-0)/ .441	Z-kurtosis = (-1.017-0)/ .858		Z-skew = (-1.194-0)/ .441)	Z-kurtosis = (1.231-0)/ .858
0.224	**1.185**			
-0.224	**-1.185**		**-2.707***	**1.435**

Appendix R ANOVA and ANCOVA output from SPSS

ANOVA calculation to check for differences in pre-test phonology scores in each treatment group.

Tests of Between-Subjects Effects

Dependent Variable: pre-test phonology (/16)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.137[a]	1	.137	.023	.879
Intercept	737.524	1	737.524	125.458	<.001
Learningcondition1romaji2furigana	.137	1	.137	.023	.879
Error	429.143	73	5.879		
Total	1168.750	75			
Corrected Total	429.280	74			

a. R Squared = .000 (Adjusted R Squared = -.013)

Test for homogeneity of regression to satisfy assumptions for ANCOVA calculation that there is a linear relationship between dependent variable (**post-test** scores) and covariate (pre-test scores) between groups (i.e., a non-significant difference).

Tests of Between-Subjects Effects

Dependent Variable: immediate post-test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	499.599[a]	28	17.843	1.456	.128	.475
Intercept	7204.719	1	7204.719	587.990	<.001	.929
Learningconditionquantised1romaji2furigana	27.396	1	27.396	2.236	.142	.047
pretestsemantic16	374.679	17	22.040	1.799	.059	.405
Learningconditionquantised1romaji2furigana * pretestsemantic16	81.590	10	8.159	.666	.749	.129
Error	551.391	45	12.253			
Total	11887.750	74				
Corrected Total	1050.990	73				

a. R Squared = .475 (Adjusted R Squared = .149)

Test for homogeneity of regression to satisfy assumptions for ANCOVA calculation that there is a linear relationship between dependent variable (**delayed post-test** scores) and covariate (pre-test scores) between groups (i.e., a non-significant difference).

Tests of Between-Subjects Effects

Dependent Variable: delayed post-test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	371.343[a]	23	16.145	1.217	.299	.467
Intercept	4409.517	1	4409.517	332.283	<.001	.912
Learningconditionquantisised1romaji2furigana	36.164	1	36.164	2.725	.109	.078
pretestsemantic16	227.099	14	16.221	1.222	.308	.348
Learningconditionquantisised1romaji2furigana * pretestsemantic16	56.577	8	7.072	.533	.823	.118
Error	424.652	32	13.270			
Total	6834.250	56				
Corrected Total	795.996	55				

a. R Squared = .467 (Adjusted R Squared = .083)

ANCOVA SPSS output for *main effect* difference in **post-test** by treatment group, controlling for pre-test scores.

Tests of Between-Subjects Effects

Dependent Variable: immediate post-test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	315.651[a]	2	157.825	15.239	<.001	.300
Intercept	2476.819	1	2476.819	239.147	<.001	.771
pretestphonology16	254.390	1	254.390	24.562	<.001	.257
Learningconditionquantisised1romaji2furigana	52.162	1	52.162	5.036	.028	.066
Error	735.339	71	10.357			
Total	11887.750	74				
Corrected Total	1050.990	73				

a. R Squared = .300 (Adjusted R Squared = .281)

ANCOVA SPSS output for *main effect* difference in **delayed post-test** by treatment group, controlling for pre-test scores.

Tests of Between-Subjects Effects

Dependent Variable: delayed post-test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	197.236[a]	2	98.618	8.729	<.001	.248
Intercept	1501.836	1	1501.836	132.937	<.001	.715
pretestphonology16	132.545	1	132.545	11.732	.001	.181
Learningconditionquantisised1romaji2furigana	55.486	1	55.486	4.911	.031	.085
Error	598.760	53	11.297			
Total	6834.250	56				
Corrected Total	795.996	55				

a. R Squared = .248 (Adjusted R Squared = .219)

ANCOVA SPSS output for *interaction effects* of treatment group by proficiency level in **post-test**, controlling for pre-test scores

Tests of Between-Subjects Effects

Dependent Variable: immediate post-test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	469.309[a]	4	117.327	13.918	<.001	.447
Intercept	2576.756	1	2576.756	305.659	<.001	.816
pretestphonology16	191.471	1	191.471	22.713	<.001	.248
Learning_condition (1romaji,2furigana)	47.242	1	47.242	5.604	.021	.075
Proficiency_quant	122.695	1	122.695	14.554	<.001	.174
Error	581.681	69	8.430			
Learning_condition (1romaji,2furigana) * Proficiency_quant	20.473	1	20.473	2.429	.124	.034
Total	11887.750	74				
Corrected Total	1050.990	73				

a. R Squared = .447 (Adjusted R Squared = .414)

ANCOVA SPSS output for *interaction effects* of treatment group by proficiency level in **delayed post-test**, controlling for pre-test scores.

Tests of Between-Subjects Effects

Dependent Variable: delayed post-test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	338.169[a]	4	84.542	9.418	<.001	.425
Intercept	1605.799	1	1605.799	178.879	<.001	.778
pretestphonology16	83.366	1	83.366	9.287	.004	.154
Learningconditionquantisised1romaji2furigana	57.173	1	57.173	6.369	.015	.111
Proficiency_quant	101.568	1	101.568	11.314	.001	.182
Learningconditionquantisised1romaji2furigana * Proficiency_quant	24.132	1	24.132	2.688	.107	.050
Error	457.827	51	8.977			
Total	6834.250	56				
Corrected Total	795.996	55				

a. R Squared = .425 (Adjusted R Squared = .380)

Appendix S Kruskal-Wallis output from SPSS

Non parametric Kruskal-Wallis calculation

Independent-Samples Kruskal-Wallis Test Summary

Total N	74
Test Statistic	3.374[a,b]
Degree Of Freedom	1
Asymptotic Sig.(2-sided test)	.066

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because there are less than three test fields.

Appendix T CUREC ethics approval

SOCIAL SCIENCES & HUMANITIES
INTERDIVISIONAL RESEARCH ETHICS COMMITTEE
DEPARTMENTAL RESEARCH ETHICS COMMITTEE

Department of Education
15 Norham Gardens, Oxford OX2 6PY
student.curec@education.ox.ac.uk; staff.curec@education.ox.ac.uk



Brendan Allen
Department of Education, Social Sciences Division
University of Oxford

19 May 2023

Dear Brendan,

Research ethics approval

Research title: Self-regulation and learner strategies in Japanese L2 vocabulary learning, and the effects of romaji and furigana phonological reading aids in word learning.

Research ethics reference: EDUC_C1A_23_181

The above application has been considered on behalf of the Education Departmental Research Ethics Committee (DREC) in accordance with the University's procedures for ethical approval of all research involving human participants.

I am pleased to confirm that, on the basis of the information provided to the DREC, ethics approval has now been granted for this study.

Please note the following:

Personal data: It is the responsibility of the PI to ensure that all personal data collected during the project is managed in accordance with the University's guidance and legal requirements.

In-person activities: Any data collection involving in-person interactions with participants must have an up-to-date fieldwork risk assessment in place; further guidance is available from the Safety Office's website.

Amendments: Please notify the committee if you intend to make any amendments to the information in your ethics application as submitted at date of this approval, as all changes must receive ethical approval prior to implementation. The amendment form is available on the SSH IDREC webpage.

We welcome feedback on your experience of the ethical review process and suggestions for improvement. Please email any comments to staff.curec@education.ox.ac.uk / student.curec@education.ox.ac.uk or ethics@socsci.ox.ac.uk.

Yours sincerely

Pinar Kolancali

DREC member

Appendix U Participant online survey consent form

DEPARTMENT OF EDUCATION
15 Norham Gardens
Oxford
OX2 6PY

Professor Heath Rose
Deputy Director
heath.rose@education.ox.ac.uk
+44(0) 1865 274 033

Brendan Allen
MSc Research Student
MSc Applied Linguistics & 2nd Lang Acquisition
brendan.allen@education.ox.ac.uk

PARTICIPANT INFORMATION FOR ONLINE SURVEYS OR TASKS

"Self-regulation and L2 Japanese vocabulary learning"

CUREC Ethics Approval Reference: EDUC_C1A_23_181

General information and research aim

You have been invited to participate as you are 18 and over, with L2 Japanese language ability ranging from beginner-to-intermediate level. You will be asked to complete an online proficiency test, self-regulation survey, followed by a vocabulary learning task and short vocabulary test. The first session should take no longer than 45 minutes to complete. Approximately 24 hours after the first task, you will be asked to join a follow-up session lasting around 3-5 minutes. Based on your performance in the tests, you may be asked to participate in a short follow-up interview lasting no longer than 15mins. Participation in the interview is voluntary, and you will not be required to participate if contacted. To compensate you for your time, all participants who take part in the online study will automatically be entered into a prize draw for a chance to win one of five cash eGift vouchers.

The aim of this research is to help understand the most effective ways to present new words to Japanese language learners and to explore the role of self-regulation and language learning strategies in Japanese L2 word learning. We appreciate your interest in participating in this online study. Please read through this following information before agreeing to participate. If you wish to take part, please indicate your intent by ticking the 'yes' box at the bottom of this form.

The principal researcher for this study is Brendan Allen, Department of Education, University of Oxford. The research is being completed under the supervision of Professor Heath Rose, Deputy Director, Department of Education, University of Oxford. The data is needed for statistical analysis, and will be used to help complete an MSc dissertation project. The data will only be used by myself, (Brendan Allen), and accessible by my academic supervisor, Professor Heath Rose. You are free to ask any questions before deciding to take part in the study by contacting myself or Professor Rose.

Do I have to take part?

No. Please note that participation in this study is 100% voluntary. If you do decide to take part and change your mind, you may withdraw at any point leading up to and during the study. If you are in the middle of the study, you may withdraw before submitting your answers simply by escaping from the programme and closing the browser. In addition, all survey questions in the questionnaire include a 'prefer not to say' option should you prefer not to answer.

How will my data be used?

The quantitative and qualitative data will be used as the basis for primary analysis, the results of which will be used to write my MSc thesis. The submission of this thesis is requirement for the MSc Applied Linguistics & Second Language Acquisition programme, Department of Education, University of Oxford.

What type of data will be collected?

The data we collect that could identify you will be; age, gender, L1 background, length of L2 study, country of residence, audio recording, (if participating in the interview portion of the study), and your Email address. **Your IP address will not be stored**[1]. We will take all reasonable measures to ensure that data remain confidential by using university approved data collection software.

The responses you provide will be stored in a password-protected electronic file on University of Oxford secure servers and may be used in academic publications and conference presentations (if applicable). Hard copies of identifiable information will be securely destroyed after pseudonymisation. Data may be indirectly identifiable after digital pseudonymisation, thus it will be stored separately from the raw data in a password-protected electronic file on University of Oxford secure servers. Pseudonymised data will be held on university secure servers for a total of three years after the submission of the MSc dissertation.

Who will have access to my data[2]?

The University of Oxford is the data controller with respect to your personal data and, as such, will determine how your personal data is used in the research. The University will process your personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest. Further information about your rights with respect to your personal data is available from https://compliance.admin.ox.ac.uk/individual-rights.

Who has reviewed this research?

This research has been reviewed by, and received ethics clearance through, a subcommittee of the University of Oxford Central University Research Ethics Committee (CUREC), approved May 19th, 2023. CUREC Ethics Approval Reference: EDUC_C1A_23_181.

Who do I contact if I have a concern or I wish to complain?

If you have a concern about any aspect of this research, please scan the QR code to contact Brendan Allen, Email: brendan.allen@education.ox.ac.uk or alternatively, speak with their supervisor Professor Heath Rose, Deputy Director, Department of Education, Email: heath.rose@education.ox.ac.uk Tel: +44(0) 1865 274 033, and we will do our best to answer your query. We will acknowledge your concern

[1] Guidance is available within CUREC's Best Practice Guidance (06) on Internet-mediated research
[2] Guidance is available within CUREC's Best Practice Guidance (09) on Data collection, protection and management

within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:

The Chair, Social Sciences & Humanities Interdivisional Research Ethics Committee; ethics@socsci.ox.ac.uk; Research Services, University of Oxford, Boundary Brook House, Churchill Drive, Headington, Oxford OX3 7GB.

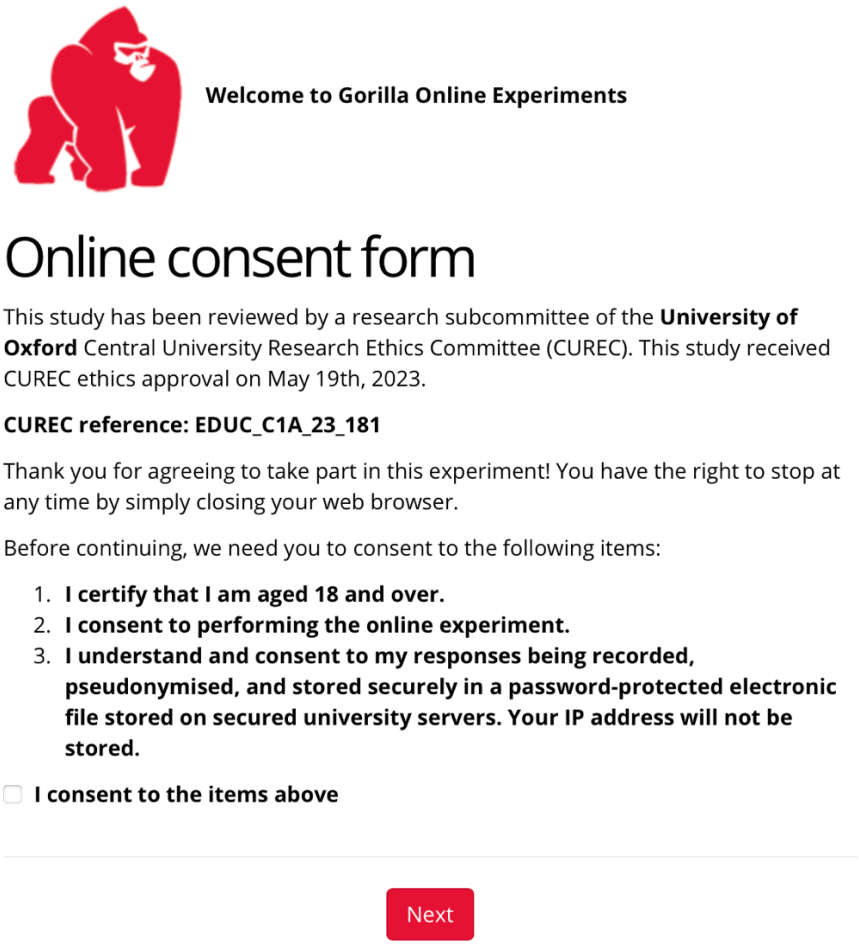
Please note that you may only participate in this survey if you are 18 years of age or over.

☐ I certify that I am 18 years of age or over

If you have read the information above and agree to participate with the understanding that the data (including any personal data) you submit will be processed accordingly, please tick the box below to start.

☐ Yes, I agree to take part

Appendix V Gorilla online consent form



Welcome to Gorilla Online Experiments

Online consent form

This study has been reviewed by a research subcommittee of the **University of Oxford** Central University Research Ethics Committee (CUREC). This study received CUREC ethics approval on May 19th, 2023.

CUREC reference: EDUC_C1A_23_181

Thank you for agreeing to take part in this experiment! You have the right to stop at any time by simply closing your web browser.

Before continuing, we need you to consent to the following items:

1. **I certify that I am aged 18 and over.**
2. **I consent to performing the online experiment.**
3. **I understand and consent to my responses being recorded, pseudonymised, and stored securely in a password-protected electronic file stored on secured university servers. Your IP address will not be stored.**

☐ **I consent to the items above**

Next

Participant online experiment consent form on Gorilla (https://app.gorilla.sc). Participants must consent by checking the box to proceed with the online study.

Appendix W Online recruitment poster

DEPARTMENT OF EDUCATION
15 Norham Gardens
OXFORD
OX2 6PY

Brendan Allen
MSc Applied Linguistics & 2nd Lang Acquisition
brendan.allen@education.ox.ac.uk

UNIVERSITY OF OXFORD

VOLUNTEERS NEEDED FOR L2 JAPANESE VOCABULARY STUDY

"Self-regulation and L2 Japanese vocabulary learning"

自己調整と第二言語(L2)の日本語語彙学習

Ethics Approval Reference: EDUC_C1A_23_181

This research project aims to investigate the learning of Japanese vocabulary for beginner-to-intermediate Japanese second language (L2) learners. We are looking for volunteers, aged 18 and over, who are currently studying Japanese, and/or enrolled on a Japanese language programme.

For those wishing to participate in the study, participants will be invited to complete a short proficiency test, self-regulation survey, and vocabulary training task. The session will be conducted online, and will last no longer than 45 minutes. Participants will then be invited to a follow-up session lasting no longer than 10 minutes.

We may invite a few participants for a follow-up interview regarding their self-regulation and vocabulary learning. The interview is completely voluntary, and there is no obligation to take part. We anticipate that the interview would take no longer than 15 minutes.

If you are interested in taking part in the study, please email Brendan Allen brendan.allen@education.ox.ac.uk, Department of Education, University of Oxford, 15 Norham Gardens, Oxford, OX2 6PY, Alternatively, contact Professor Heath Rose, Deputy Director of Department, Email: heath.rose@education.ox.ac.uk.

Thank you.

"Self-regulation and L2 Japanese vocabulary learning" ・CUREC Ethics Approval Reference: EDUC_C1A_23_181

Appendix X Dual language business cards





Dual language English and Japanese business card used for marketing purposes.
Printed by the University of Oxford Print Studio (printstudio@admin.ox.ac.uk)