Developing a competency-based training tool to introduce word reading drills to novice learners

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Citable link
https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37367429

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Developing a competency-based training tool to introduce word reading drills to novice learners

A thesis presented by

Lien Tran

to

The Department of Computer Science and the Department of Linguistics
in partial fulfillment of the requirements for the degree of Bachelor of Arts with Honors
in the subjects of
Computer Science and Linguistics

Advised by

Gordon McKay Professor of Computer Science Krzysztof Gajos
Professor of the Practice of Japanese Language Wesley Jacobsen

Harvard University
Cambridge, Massachusetts

November 2020
Abstract

There are two major goals in this study: (1) propose a novel use of technology for learning Japanese katakana that enables the use of early word reading drills for beginners to help them learn new characters through finishing the drills, and (2) evaluate the effects of introducing assisted word reading drills to beginners compared to character recognition drills.

To achieve these goals, we designed and implemented a training tool for word reading skills that was built upon the framework of intelligent tutor systems and mastery learning. We evaluated this tool by comparing it with its character recognition counterpart via (1) 8 qualitative interviews with beginner learners and those who had never studied Japanese before, (2) a quantitative experiment where 57 participants were randomly assigned one of word-based version or character-based version and were asked to finish the pre-training and post-training surveys and quizzes, and (3) indirect observations of quantitative participants’ interactions with the tool. In the quantitative experiment, the two versions (word-based vs. character-based) were compared based on: participants’ change in self-efficacy, excitement towards the learning task, subjective perception of workload estimation, objective measurement of performance, and additional optional comments.

Participants in the quantitative experiment, 77% of whom had never learned Japanese before, were able to practice with word-reading drills as their first step to learn katakana with approximately similar ratings in perceived workload and technology acceptance compared to the tool’s counterpart that targets character recognition. Qualitative interviews and indirect observations further revealed qualitative feedback on participants’ perception of the nature of learning new characters through practicing to read words, their interactions with the versions, and further suggestions for improvements.
These results suggest that the tool as currently designed or similar computer-assisted training tools could be further developed and explored to assist beginners with learning katakana effectively. Further study would also be desirable regarding the possibility of building upon the concepts of module-based mastery learning proposed in this tool to propose other designs for computer-assisted word reading drills that may be less intimidating for beginners.
Acknowledgements

First and foremost, I would like to thank my thesis advisers. Thank you for your immense patience, and dedicated involvement in every step throughout the process even when my thesis coincided with your paternity leaves. Thank you, Professor Krzysztof Gajos, for your guidance in my journey of learning to becoming a better researcher, from shaping a research topic and thinking critically about the problem, to putting the pieces together into words. Thank you, Professor Wesley Jacobsen, for helping me connect the research topic to its practicality in teaching and learning Japanese, and for your incessant encouragement and discipline that helped me stay on track with my thesis progress.

I would also like to thank the members of the Human-Computer Interaction at Harvard SEAS for your comments and suggestions on this study, especially with my research direction and experiment design.

I would like to thank my family for your effort to provide me with a decent work-from-home environment in our shared space when I was working remotely, and for your supportive phone calls from half a globe away when I had moved to the US.

I would like to thank Tuan Le, for your immense support, and for your contagious uplifting spirit.

I would like to thank my friends for helping me recruit participants for my study, and for your warm support in life.
Contents

1. Introduction
   1.1. Background
   1.2. Objectives of the current study
   1.3. Our contributions

2. Related work
   2.1. Self-efficacy and Mastery Experience
   2.2. Module-based content: Mastery Learning
   2.3. Cognitive load can be heightened when introducing word reading drills
   2.4. Reducing visual cognitive load
      2.4.1. Asynchronous display of romaji
      2.4.2. Visual mnemonics to help with character recognition

3. Implementation
   3.1. Overview
   3.2. Iterative process to improve the design and concept of the tool
      3.2.1. Testing with intended components and potential practice materials
      3.2.2. Linear display of words
      3.2.3. Prioritizing words with the fewest unfamiliar characters
      3.2.4. Criteria for implementation
   3.3. Designing the Module Builder
      3.3.1. The big picture
      3.3.2. Proposed solution
   3.4. Designing the Tutor Model and User Interface

4. Evaluation
   4.1. Qualitative study
   4.2. Quantitative study
      4.2.1. Participants
      4.2.2. Materials and procedure
      4.2.3. Survey design
   4.3. Qualitative indirect observation of users’ interaction

5. Results
   5.1. Results from qualitative interviews
   5.2. Results from quantitative experiment
      5.2.1. Self-efficacy
5.2.2. Learner’s excitement 43
5.2.3. Workload measurement 44
5.2.4. Technology Acceptance Model 46
5.2.5. Objective measurement of participants’ performance 47
5.3. Results from qualitative indirect observation of users’ interaction 49

6. Discussion 50
  6.1. Evaluating the hypotheses 50
  6.2. Further implementation considerations 52
  6.3. Limitation of the study and future development 53

7. Appendix 55
  7.1. Pre-training survey 55
  7.2. Post-training survey 67

8. References 76
1. Introduction

1.1. Background

Learners have to memorize a large amount of distinct symbol-sound correspondence when they start to learn Japanese. Beginners are required to memorize the Hiragana and katakana (the Japanese kana phonetic alphabet), each containing 46 basic characters, or 71 including diacritics. Each character represents a syllable in Japanese. Within each kana system, there is no explicit similarity between the visual representation of a character and its corresponding pronunciation, and between characters that have similar pronunciation. As learners progress through the curriculum, their acquisition of other linguistic processes can be inhibited by the "cognitive overload" caused by unfamiliarity with the writing system and the fact that they do not have a "bootstrapping" mechanism to process the visual cues of the new language (Brockett 2017).

Although learning an unfamiliar orthography can be difficult, adult learners are usually expected to finish learning at least one of the kana systems in their early lessons, and reading drills are typically asynchronous with character recognition training. One approach is to quickly help learners gain exposure to all the characters, and let them memorize the characters by themselves before diving into reading drills (Yone and Makiko 1998). Another approach is to explicitly instruct learners to practice with sets of characters in alphabetical order (Saito-Abbott et al. 2000). For this approach, only after they have mastered every character in the alphabetical set do learners begin to practice with reading drills.

One advantage of the first method is that it helps to introduce learners to real reading context early in the curriculum. It has been proposed that recognition of whole orthographic
words should be at an early stage of learning how to read Japanese (Alberta Learning, Edmonton (Canada). Curriculum Standards Branch. 1999; Brockett 2003), and word recognition has been shown to be a good predictor of students’ reading performance (Chapman and Tunmer 2003).

However, this approach introduces a large number of new characters to learners, while only approximately seven meaningful items can be stored in short-term memory at a time (Miller, 1956). In the information processing model proposed by Schmidt (1991) (Figure 1), unrehearsed information is forgotten from short-term memory. Since there are a large number of characters and the frequency and order of each character is not controlled in reading drills, they run the risk of forgetting the less frequently used characters.

![Image of the information processing model](image)

Figure 1: The information processing model (Schmidt 1991)

Regarding this aspect, learning unfamiliar characters in manageable sets, i.e. using the second approach, offers the opportunity to reduce cognitive load and help learners advance in their memory cycle from short-term memory to long-term memory. Learners are given enough time to master every character in a set. As a result, they are expected to have gained character recognition proficiency before beginning word reading drills. However, character recognition and word recognition require different cognitive processes (Howard 1991; Rapcsak, Rubens, and Laguna 1990; Scaltritti and Balota 2013). As a result, skills might not directly transfer from character recognition to word recognition. Indeed, some reading difficulties faced by
second-script readers may be due to pedagogical strategies that employ decoding and assembling words character by character (Brockett 2003). For example, focusing excessive attention on individual characters may divert the learners from improving their word recognition skills and slow down the process of learning to read Japanese.

A possible solution to combine the advantages of the two aforementioned methods is to introduce learners to word reading drills while being able to control the characters that appear in the drills. Learners will be able to practice word recognition skills at an early stage in their learning, which is the advantage of the first method. They will also be introduced to a small number of unfamiliar characters, which can reduce their cognitive load and help with their memory cycle, which is the advantage of the second method. However, it can be difficult and time-consuming to come up with such drills for Japanese kana.

A possible solution to combine the advantages of the two aforementioned methods is to introduce learners to word reading drills while being able to control the characters that appear in the drills. Learners will be able to practice word recognition skills at an early stage in their learning, which is the advantage of the first method. They will also be introduced to a small number of unfamiliar characters, which can reduce their cognitive load and help with their memory cycle, which is the advantage of the second method. However, it can be difficult and time consuming to come up with such drills for Japanese kana.

This issue can be resolved if approached computationally. However, past technological solutions have been focused on developing and implementing strategies to help learners memorize characters as separate units: flash-card tools (citation), quizzes (citation), mnemonic incorporation (citation), etc. The question remains whether and how technology can be used to develop drills for developing beginners’ reading skills beyond the character level.
A supposed disadvantage of early word reading drills approach is that introducing beginners, and even novice learners to word reading drills can do more harm than good compared to learning separate characters. For example, when offered opportunities to practice reading skills at the word level, learners can feel intimidated by the task due to a higher cognitive load: more unfamiliar characters will appear at the same time. On the other hand, as will be discussed in the section on Related Work from past research, being able to accomplish more difficult tasks improves self-efficacy, i.e. one’s beliefs and confidence in their capabilities to achieve good performance. Yet, if the tasks are too difficult, failure may reduce self-efficacy.

Would learning through practicing with words help or hurt students’ performance compared to practicing with separate characters? How might we design a tool for word reading drills that have low cognitive load so that learners can gain the benefit of practicing with their word recognition skills without feeling overwhelmed?

1.2. Objectives of the current study

Given the issues discussed in the previous section, there are two major goals in this study: (1) propose a novel use of technology for learning Japanese katakana that will enable the use of early word reading drills for beginners that combine the advantages of the two current methods and (2) evaluate the effects of introducing assisted word reading drills to beginners compared to character recognition drills.

Regarding the first major goal, the tool should be able to:

**Design goal A.** Maintain, if not improve, students’ belief and confidence about their character and word recognition skills compared to practicing with separate characters
**Design goal B.** Provide cognitive support to help learners get used to unfamiliar aspects of the orthography to be learned

**Design goal C.** Maintain, if not improve, students’ performance in character and word recognition skills compared to practicing with separate characters

Regarding the second major goal, through evaluating the design of the tool against its version for character recognition training, I seek to answer the following questions:

**Research question 1.** Would learning new characters through practicing with words help or hurt beginners’ performance compared to practicing with separate characters in terms of:

a. Self-efficacy
b. Cognitive load
c. Their experience with using the tool itself

**Research question 2.** What are the considerations needed for the design and implementation context of a word-reading drill technology at an early learning stage?

My approach is to design and build a computer-assisted training system that rolls out appropriate word-based reading drills based on its detection of the learner’s familiarity with individual characters. For the learning material, I use Japanese words that sound similar to English words. Katakana characters are primarily used for representing borrowed words from foreign languages, primarily English. The learners will therefore have a level of cognitive support and potentially find it interesting to compare and contrast Japanese and English (design goals A and C). The tool targets learners’ attention to the visual and auditory aspects of the characters by providing learners with visual and auditory aids before revealing the romanized version of the characters (design goal B).
While building the tool, I achieve my second major goal by comparing the tool with its counterpart for character recognition training, and gather qualitative and quantitative opinions from users.

Following are the hypotheses of the current study that I have:

**Hypothesis 1.** Learner’s self-efficacy will generally increase after practicing with word reading drills compared to a counterpart version of character recognition drill.

Word reading is generally viewed as a more difficult task to accomplish. As will be discussed in the next section, success in doing difficult tasks can enhance learners' self-efficacy. As a result, if the level of difficulty of word reading is kept within a reasonable range for learners, their self-efficacy can increase.

**Hypothesis 2.** Word-based reading drills using words that sound similar to English words will be at least as effective as learning individual characters in terms of helping learners to read Japanese words rapidly and accurately.

**Hypothesis 3.** Learning real Japanese words helps learners feel more engaged in learning Japanese characters compared to learning individual characters.

Because learners can compare and contrast how words they are familiar with in English are pronounced in Japanese, learning words will give them a better sense of progress when learning a language, and they will feel interested in exploring and learning more.

**Hypothesis 4.** Learners perceive learning characters at the word level as more cognitive demanding than learning separate characters in terms of the task itself, and of their experience using the tool.

Regarding the task, whereas character recognition drills focus users’ attention to only one character, word recognition drills display multiple new characters at the same time, and introduce
not only the new characters but also the meaning of the words. Regarding the experience using
the tool, if each character is built as an interactive component, then word reading drills give users
multiple interactive components at the same time and increases the possible interactions that
users have to explore to understand the tool.

1.3. Our contributions

Given the objectives of this study, there are two main areas of contribution:

1. We propose a novel use of technology for learning Japanese katakana that will
   enable the use of early word reading drills for beginners that combine the
   advantages of the two current methods. This contribution is elaborated in the
   Implementation section.

2. We evaluate the effects of introducing computer-assisted word reading drills (our
   proposed solution) to beginners compared to computer-assisted character
   recognition drills (an approach that current computer-assisted tools for beginners
   have been focusing on).

2. Related work

2.1. Self-efficacy and Mastery Experience

In research question 1, I choose to evaluate the tool on the basis of learners’
self-efficacy for two main reasons: (1) it is a good indicator for students’ performance, and (2)
introducing word reading drills can affect learner’s self-efficacy in both good and bad directions
compared to the current popular character recognition drills. This section elaborates on those two reasons.

Regarding the first reason, past research from several areas indicate that self-efficacy “is a key factor that affects learners’ interest, persistence, extent of effort students invest in learning” (van Dinther, Dochy, and Segers 2011), which are important for language learning, especially for the early stage where they have to persist through the initially high level of unfamiliarity. Learners’ beliefs can be a better indicator for performance prediction than their measured ability (Schunk 1991).

Among the factors that affect learners’ self-efficacy, mastery experiences are the most influential source of efficacy information. Mastery experience produces stronger efficacy beliefs than with either of the other modes of influence that rely solely on vicarious experiences, cognitive simulations, or verbal instruction (Albert Bandura, Freeman, and Lightsey 1999). Mastery experiences occur when people have mastered something, i.e. attempted and succeeded at doing something. Mastery experiences are the most effective way to boost self-efficacy because people are more likely to believe they can conquer new challenges if the challenges are similar to something they have done well (A. Bandura 1994). While success strengthens self-efficacy, early easy successes can exaggerate disappointment from failure; difficulties are learning opportunities to become more resilient to future setbacks (A. Bandura 1989).

Considering this notion of mastery experience, the choice for practicing with words or characters at the early learning stage can affect learners’ self-efficacy in both desired and undesired ways. Given the current teaching methods and available technology, learners gain their mastery experience and self-efficacy at the early learning stage through being able to recognize individual characters. This mastery can strengthen their confidence to proceed with reading
drills. However, as discussed in the introduction, character recognition and word recognition require different cognitive processes and learners might struggle with transferring character recognition skills to word reading skills, or even worse, struggle with both character recognition and word reading. This can cause discouragement in learners.

On the other hand, when beginners face early difficulty with word reading, if they are given enough support to overcome the discouragement from the difficulty of the task, they can improve self-efficacy due to their success in accomplishing a more difficult task. However, if they fail to finish the word reading drills due to a high level of difficulty, they lack the success needed to improve self-efficacy.

As a result, explored whether and how learners’ self-efficacy and learning experience can be affected by computer-assisted word reading drills and character recognition drills. I achieved this goal by gaining insights about learners’ attitude and behavior during qualitative interviews with research subjects in this study, and through a quantitative study that took cognitive load, measurement of self-efficacy, and objective measurements of learners’ real performance into consideration. Design decisions were made such that any new cognitive barriers are reduced as much as possible, making it possible to focus in the study on the distinction between word reading drills and character recognition drills.

2.2. Module-based content: Mastery Learning

In order to meet Design Goal A (section 3.2.4) and facilitate master experiences, this study adopts the framework of mastery learning to propose a design for a competency-based training tool where learners progress at their own pace. Learning content is divided into digestible modules and content will only be rolled out at the point that a satisfactory level of
mastery is demonstrated. Enactive Mastery Experience can be reinforced in that a restricted amount of new materials is presented so that learners can challenge themselves with unlearned characters, yet do not feel too challenged and discouraged in doing so.

Mastery learning (ML) is often brought up as an effective method to promote such competency-based training. In mastery learning, students are required to demonstrate their mastery of a unit of knowledge before moving on to the next unit. Although mastery programs can vary according to course content and procedures used, they have been shown to have positive effects on student attitudes toward course content (Kulik, Kulik, and Bangert-Drowns 1990). Keller’s Personalized System of Instruction (PSI) is one influential approach in ML teaching methodologies (Kulik, Kulik, and Bangert-Drowns 1990). In this system, students progress at their own pace through course materials that are largely in written form. If they fail to demonstrate mastery of a particular unit, students will restudy the material, usually with tutorial help, and retake tests until they receive satisfactory scores for the unit. (Kulik, Kulik, and Bangert-Drowns 1990)’s meta-analysis of 103 studies that examined effects of PSI and ML revealed that this competency-based approach was more effective than the traditional approach by an effect size of 0.52 on average. An effect size in educational research helps to indicate the effectiveness and impact of an intervention (Schagen and Elliot 2004) by investigating the “magnitude of the difference between groups” (Sullivan and Feinn 2012). An effective size of 0.52 is considered to have a medium impact (Cohen 1992).

In order to adopt the mastery learning framework, I will discuss the process of coming up with and redefining “satisfactory level of mastery” in the Implementation section. I will also discuss how the tool can roll out content that can help learners gain an enactive experience based on my definition of mastery in the context of the tool.
2.3. Cognitive load can be heightened when introducing word reading drills

Also regarding research question 1, I chose to evaluate the tool based on cognitive load and the experience with the tool itself because even though the approach has the potential to reduce cognitive load by limiting the number of new items being introduced, this approach can incur further cognitive load not only because it introduces a more difficult task to learner, but also because it can introduce an interface that demands more cognitive load.

In the introduction, I discussed that an advantage of character based introduction is the fact that learners can focus on a limited number of new items at the same time. This small number of new items that learners have to learn facilitates the advancement of new materials in the memory cycle. I argued that introducing word reading drills while controlling the number of new characters introduced to learners can produce a similar benefit.

However, learners might, in such a case, have to endure heavier cognitive load. Whereas character recognition drills focus users’ attention to only one character, word recognition drills display multiple new characters at the same time, and introduce not only the new characters but also the meaning of the words.

Regarding the experience using the tool, if each character is built as an interactive component, then word reading drills give users multiple interactive components at the same time and increases the possible interactions that users have to explore to understand the tool. Due to the split-attention effect, where information in the same visual or auditory channel can interfere with the learning process (Mayer and Moreno 1998), visual aspects of the interface can interfere with learners’ effort to get used to unfamiliar visual aspects of the new orthography. As a result, I strived to reduce the visual cognitive load while designing the tool.
In this study, I proposed to reduce the visual cognitive load from the learning task itself by using the methods discussed in the next section.

2.4. Reducing visual cognitive load

2.4.1. Asynchronous display of romaji

A popular method to overcome the initial unfamiliarity with the writing system is to display a romanized version of Japanese characters (romaji) in early reading exercises, either in place of or together with Japanese characters. This method lessens the perceived level of difficulty by using a familiar orthography system to introduce learners to the instruction materials. This familiarity helps learners to learn more materials at the early learning stage while they are still getting used to the new Japanese orthography system, which helps them gain a sense of competence.

However, learners’ familiarity with the romaji orthography is likely to divert the learner’s attention from the unfamiliar non-cognate orthography, which can induce the blocking effect discussed in (Kamin 1967) and prevents learners from adopting Japanese characters effectively. Blocking effects may occur during the conditioning process if an association has already been formed between a conditioned stimulus and an unconditioned stimulus, so that introducing another conditioned stimulus together with the conditioned stimulus will impair the association between the new conditioned stimulus and the unconditioned stimulus. Learners have, that is, already formed strong associations between the romanized script, and character pronunciation, and introducing an unfamiliar writing system together with the romanization can “block” the association between the new characters and their pronunciation.
In an experiment reported on in Chung (2003), displaying the romanization of Chinese characters as feedback (i.e. after learners have seen the Chinese characters) yielded better learning outcome in terms of the number of correct recalls and number of trials to recall given Chinese characters compared to displaying the romanization together with the characters.

Figure 2: Percentage of correct pronunciation responses during immediate post tests up to the criterion trial (top) and mean number of trials to reach the criterion (bottom)

The tool developed in this study was designed so that learners have to focus on the visual features of the characters before seeing the character’s romaji. The asynchronous display of romaji also helps to reduce the visual items that are shown to users at a time.
2.4.2. Visual mnemonics to help with character recognition

As discussed in the introduction, Japanese characters have a high degree of unfamiliarity for learners and there is no explicit similarity between the visual representation of a character and its corresponding pronunciation, and between characters that have similar pronunciation.

As discussed earlier in section 2.4, this unfamiliarity, together with the competing potential visual distraction from the tool can interfere with learners’ progress due to split attention effects. To lessen this effect, I propose using visual mnemonics as an aid to focus learners’ attention to the visual characteristics of the characters.

Mnemonics are used in some Japanese curriculums to help learners familiarize themselves with Hiragana and katakana (Alberta Learning, Edmonton (Canada). Curriculum Standards Branch. 1999). Among the types of association that can be strengthened with visual mnemonics for character recognition, an effective mnemonic device is the one that emphasizes both the shapes of the characters and its pronunciation.

In this study, I used the embedded mnemonic system available on Tofugu’s website (Francisco and Koichi 2014)

3. Implementation

3.1. Overview

This study explores different training components using the architecture components of Intelligent Tutoring Systems.
The Domain Model sets a standard of domain expertise: the rules, concepts, and knowledge related to the domain to be learned. This domain helps to evaluate students' performance and detect their errors. In this study, this domain is used to store the katakana-romaji pairs, and the learning modules.

The Student Model contains data from and about the learner that reflects their learning state and progress. In this project, the state and progress is determined by their familiarity with each character in the alphabets to be learned.

The Tutoring Model makes decisions about tutoring strategies based on the data gained from the domain model and student model. In this tool, it is the tutoring model that determines what instructions to deliver and when for different characters based on the student’s learning progress.
3.2. Iterative process to improve the design and concept of the tool

3.2.1. Testing with intended components and potential practice materials

For the tool designed in this study, I created digital prototypes to test different learning strategies that are possible using the tool and the potential user interface that the tool can have.

![Early sketches of the user interface of the tool](image)

Figure 4: Early sketches of the user interface of the tool

After that, finding criteria for choosing words to present was the most unclear aspect of creating the tools. There had not been previous research that discussed similar methods, and there can be many ways to generate and prioritize the content of the modules. In proposing a set of criteria for a satisfactory tutoring model, I let my decision be guided by the goal of maximizing interactive and incremental development in the learner.
3.2.2. Linear display of words

At the early stage of design exploration, I asked both experienced and novice Japanese learners to interact with an early web prototype and express their thoughts as they used the tool. The prototype merely displayed the words linearly according to incremental word length. I also asked them to compare using the tool and using an online tool called Tiny Cards to study katakana using flashcards. Even though there were different design components of both tools that could affect participants’ experience, this activity helped me form a better understanding of where learners might have trouble, and what areas word reading drills can be different from character reading drills.

3.2.3. Prioritizing words with the fewest unfamiliar characters

In section 2.2, I discussed mastery learning, a framework of learning where learners can progress through a curriculum by demonstrating a satisfactory level of mastery of prior units of the material. In this phase of the study, I defined the units of materials as characters themselves, and I defined mastery as the ability to type in the romaji for a character correctly without viewing mnemonic hints. Because characters appeared in word context, choice of words affected how learners progressed through the units of mastery. This version prioritized displaying words with the fewest unfamiliar characters.

By default, all characters are considered “unfamiliar”. Users’ mastery of a character is detected once a user has correctly typed in a character X times without having to click on the hint card. X was determined to be 2 after testing with the database. A character is converted to “unfamiliar” again in 3 situations: (1) the learner types in that character incorrectly, (2) the
learner has to view the hint card for the character, (3) the learner incorrectly types in the romaji corresponding to the character for another character.

In order to help learners consolidate a new character quickly while it is still stored in short-term memory, the tool prioritized words that contain that character. I expected that this implementation would allow learners to gradually extend the number of units mastered because more characters will become familiar as learners use the tool, and learners are exposed to a small number of new characters at a time.

Figure 5: Simple states for each character
3.2.4. Criteria for implementation

Based on the qualitative feedback gained in the aforementioned versions, the following criteria were adopted for the design’s implementation:

1. There should be a mechanism for gradually extending the characters that learners practice with, until all the characters in the database have been included.
2. Newly learned characters and incorrect characters should be prioritized so that those characters are still in the user’s short-term memory.
3. A character should only be considered as familiar for the mastery unit if learners have gotten that character correct multiple times.
4. Learners should be introduced to a small set of new characters at a time.
5. Even when a character has been detected as familiar, it should be presented to learners in the future to take advantage of interleaved practice and in-session spaced repetition.
6. Learners should have a visual sense of progress and have the bird’s eyes view of the katakana chart.

The design discussed in the previous section satisfied goal 1, 2, and 3, but failed in goal 4 and sometimes failed in goal 5. Even though users could interact with a small number of new characters at a time, they could be introduced to a series of small numbers of new characters, which eventually resulted in introducing too many characters. This made the design fail in goal 4. In addition, characters that were already considered familiar would not appear again unless users ran into the same character in words, or mistakenly typed in the character for another character. These drawbacks call for a higher level control that is beyond keeping track of single characters. This gives rise to the idea of implementing a module builder that will be discussed in the next section.
3.3. Designing the Module Builder

3.3.1. The big picture

The Module Builder receives a training database that contains words and characters and, based on that, generates training modules such that the modules satisfy the criteria discussed in the previous section, and follow the definition adopted in the tool of mastery units.

In implementing this, mastery units are defined in terms of sets of characters that learners are introduced to at their current level. This new definition allows the tool to achieve goals 4 and 5 compared to a definition of mastery that is based on the familiarity of single characters.

Figure 6: The Module Builder in the context of the system
3.3.2. Proposed solution

**Module Builder overview**

Each module consists of a set of characters (module’s characters) and a set of words (module’s words) that are formed only from the module’s characters. The subsequent module is built up from the previous module by adding to the module’s characters the least number of new characters possible.

This idea is based on mastery learning and interleaved practice. Learners progress to a new module once they have shown satisfactory mastery of the current module, and in every module, they practice with materials combining content from all the learned materials.

![Diagram of module builder](image)

Figure 7: How the module builder works. The circles on the left represent individual characters
To allow users to practice with new characters more frequently, each character will be assigned a repetition time based on how new the character is. The repetition time is decreased for a user if as the character in consideration is input correctly. If not, the score of the character is reset to a new character’s repetition time and the user will have to practice more times with that character. A module is considered done when all the character repetition times have been decreased to 0. Interleaved practice is reinforced when the repetition times for current characters are reset to 1 so that learners will revisit learned characters when they have progressed to a new module.

To present characters that need to be revisited early, a focus queue is created to store characters that need to be prioritized. This queue is maintained by the Tutor Model. The Tutor Model also chooses the word to practice with. To do so, the tool randomly chooses a character that has not been repeated enough, and randomly chooses a word among the module's words that contain that character. Once the current module is done, the module attempts to extend the current module before choosing a word.

**Module Builder implementation**

The task given to the Module Builder is modeled as a search problem. At each iteration, it looks for the character and corresponding new words to add to the training modules. To guarantee that the Module Builder can find a new word to add to the module and extend the module’s characters, it looks for words that introduce the least number of new characters and adds the characters to the module accordingly to update the module’s words. If there is a tie between different characters it can add, the Module Builder uses a greedy approach and chooses the one that has a higher frequency in the database. To speed up looking for new words that it
can add to the current module, the module builder looks at “word union,” a set of words that contain some characters that are already in the module’s word.

Figure 8: big picture of the relationships between module’s words, word union, and words in the database and different sets of characters

3.4. Designing the Tutor Model and User Interface

**Feedback for incorrect answers**

A common difficulty that students mentioned during the interviews was that characters looked alike. When giving feedback for the incorrect answers, the tool displays the current character and the character that a user types in so that they can compare the difference between the two.
Visual sense of progress

Several people who had not learned Japanese before asked how many characters there were in the katakana chart, and how much of the chart they would learn using the tool. A katakana progress chart was added to give them the big picture and to encourage them to fill up the chart.

Figure 9: a screenshot from when a real user got an incorrect answer.
Figure 10: visualization of learners’ progress. This chart displays all the basic katakana. Characters users have already seen are highlighted in color. They can hover over each card to review the romaji of that character.

4. Evaluation

4.1. Qualitative study

Participants

I recruited 8 participants who were either taking introductory Japanese or had not taken any Japanese classes to conduct a qualitative study of the tool over Zoom. Among them, 3 were taking the introductory Japanese course either at Harvard College or the Harvard Extension school.

Materials

Two versions of the tool were shown to the participants. They have almost identical visual design, except for those that are unique to the word reading drills. The content of the character-based version is taken from the modules of the word version. Users progressed through the modules of the character version similarly to how they progressed through the modules of the word version.
Figure 11: The training content for character-based version (left) vs. word-based version (right)

Figure 12: How mnemonic cards are displayed for each character: for character-based version vs. word-based version

Procedure
For people who were already learning Japanese, I interviewed them about their experiences with learning the language: what motivated them to learn Japanese and what methods they had used to learn Japanese kana. For people who were not taking Japanese courses at the time, I asked how familiar they were with Japanese language, whether and how much they knew about the Japanese alphabets, and how they would attempt to learn the alphabets.

Then I asked the participants to share their screens when opening a link to one of the two versions. The order of the tools were balanced out among the participants. I asked them how to describe how they would interact with the tool given the screen they saw, and then use the tool until a 7 minute timer popped up. Participants were encouraged to comment on the tool as they interacted with it. After the 7 minutes had passed, participants were asked to reflect on their experience with the tool before repeating the same procedure with the other tool.

After using both tools, participants were then asked to compare the pros and cons of the two tools, whether and how they would use the tool to learn katakana, and which tool they would prefer to use and why. Interviews ranged from 40 minutes to an hour. Upon completion of an interview, an interviewee received a $10 Amazon gift card.

Some design aspects of the tool were changed in between the qualitative interviews, which will be discussed in the Results section.

4.2. Quantitative study

4.2.1. Participants

57 participants finished the quantitative experiment. These were recruited through sending emails to mailing lists of Harvard College and posting on online Japanese learning forums.
Among these participants, 44 participants had not learned Japanese before, 8 had been learning Japanese at introductory level, and 5 had been teaching themselves the basics. 34 identified themselves as female and 23 as male. Their ages ranged from 18-25, with the median age being 20.

4.2.2. Materials and procedure

The two versions used for qualitative interviews were also used for the quantitative study. The participants finished the experiment on their own computer at their own convenient time. They were given a pre-training survey to fill in their demographics, finish a short quiz, and a self-efficacy questionnaire. At the end of the form, they were randomly given the link to one of the two tools. After using the tool for 20 minutes, they would see a pop up that led them to the post-training survey. The whole experiment typically took 30 minutes. Upon completion of the post-training form, a participant received $5 in the form of an Amazon gift card, cash, or donation to charity.

Their interactions with the tool were also used for qualitative indirect observation, which will be discussed in section 4.3.

4.2.3. Survey design

Self-efficacy

In order to compare how participants’ self-efficacy changed after using the tool, a same set of questions was included in both the pre-training and post-training survey. In this set of questions, I adopted the sub-section of self-efficacy from the Motivated Strategies for Learning Questionnaire (MSLQ) (Paul R. Pintrich, David A. F. Smith, Teresa Garcia, and Wilbert J.
McKeachie 1991). Participants rated themselves on a 7-point Likert scale, from 1 (not at all true of me) to 7 (very true of me).”

In addition, because assessment of self-efficacy is task-specific (A. Bandura 1989), especially for learning tasks (Smith and Fouad 1999; Pajares and Miller 1995), and there had not been previous study that focused on self-efficacy for character or word recognition tasks, I also came up with two questions regarding participants’ confidence in their ability to recognize characters and words written in katakana.

This part of the surveys helps to answer research question 1a: “Would learning new characters through practicing with words help or hurt beginners’ performance compared to practicing with separate characters in terms of efficacy?”, and explore hypothesis 1: “Learner’s self-efficacy will generally increase after practicing with word reading drills.”

**Workload estimation**

I included the questions from the subscales from NASA-Task Load Index in the post-training survey to measure: (1) mental demand, (2) temporal demand (how hurried or rushed the pace of the task was), (3) performance, (4) effort (how hard they had to work to accomplish their level of performance), and (5) frustration (how insecure, discouraged, irritated, stressed, and annoyed they were).

The questions were used to identify the factors that affect the perceived workload of the training tools. This part helps me to explore hypothesis 4: “Learners perceive learning characters at the word level as more cognitively demanding than learning separate characters in terms of the task itself, and of their experience using the tool.”
If hypothesis 4 is true, analyzing this part of the survey should also help to identify the area that affects learners’ perceived cognitive load.

**Technology Acceptance Model (TAM)**

This measurement is used to further investigate a potential cause of heightened cognitive barriers (if any) and any differences in user experience when using the two different versions of the training tool. For example, because learners can be more familiar with tools that help them learn separate characters, a tool that introduces them to word reading drills can be less well-accepted by users. I used the questions from the TAM fast form approach (Chin, Johnson, and Schwarz 2008) to compare the usefulness and ease of use of the two versions.

**Objective measurement of participants’ performance**

One quiz was included in each of the pre-training survey and post-training survey. The quizzes had the same pre-training and post-training content for character recognition questions, and different pre-training and post-training content for word recognition questions for words with length of 2, 3, 4, and 5. Pseudo-words were chosen instead of real words to make it possible to show users novel words that only contain a limited set of characters. There are only a few numbers of such words and are already included in the training material. These pseudo-words were generated from the characters that were in module 13.

Learners also report the module level they were at when using the tools.

4.3. Qualitative indirect observation of users’ interaction

LogRocket, an online session-screen-recreation is used to simulate a user's screen and interaction during their sessions based on event logs. I used this to assess qualitatively how users
interacted with the tool to further understand the meaning of the responses collected from the surveys.

Using this tool also allowed me to reduce non-response bias. Only responses from the users who have used the tool for 20 minutes were recorded in the post-training survey. As a result, looking at the quantitative data overlooks the troubles or differences that non-respondents might have experienced. For example, their experience might have been too confusing that they discontinued using the tool. Evaluating these situations qualitatively helps me to further consider the answer to research question 2.
5. Results

5.1. Results from qualitative interviews

The interaction with the word-based version

When participants interacted with the character-based training, they could quickly proceed with the tool, especially for participants who were more familiar with katakana characters. When participants interacted with the word-based version, even though they could describe that they would have to read the word, they needed some time making sense of how to use the tool, but all participants proceeded to finish the task without asking any questions.

Some participants were confused when the word they just typed redisplayed on the tool. They were confused about whether there was a bug that prevented the tool from recognizing their correct answer. From the backend, this happened because there were not many words to satisfy the character restriction. Even though the same thing happened with the character-based version when the character-based version was shown either before or after the word-based version, the participants did not, in that case, show signs of confusion.

I made a major change to the visual design of the word display throughout the qualitative interviews. Initially, characters were placed inside card-like boxes so that learners could better focus on one character at a time and did not feel overwhelmed by the amount. However, two participants tried to press enter after inputting every romaji for a character in the word version: “Oh, I thought that I have to input one character at a time”. Another participant preferred the simplicity of the character version over the word version: “This word version has multiple instances of the other tool on the screen”. These events prompted me to remove the card-like design and display the words so that they look similar to a typical word reading drill. In addition,
it mimics the look of a typing game so that it is more apparent that learners are asked to type in the whole word.

Figure 14: How the design of the word display was changed. Left: the card-like design. Right: the new bare word design.

**Auditory vs. Visual support**

Regarding goal B: “provide cognitive support to help learners get used to unfamiliar aspects of the orthography to be learned”, visual mnemonics and audio were expected to aid with visual and auditory aspects respectively. The tool played the audio automatically and displayed mnemonic images to aid learners with auditory and visual aspects of the language. However, autoplaying the audio was a source of distraction for 4 participants. In addition, 4 participants gave explicit comments about the effectiveness of mnemonic devices for their learning, while all positive comments on the audio were about its naturalness.

<table>
<thead>
<tr>
<th>Audio</th>
<th>Mnemonics</th>
</tr>
</thead>
</table>
| Sample positive comments | “I like that the audio doesn’t feel... canned?.”  
“I like that the pronunciation has correct ups and downs like a real” | “It’s funny and interesting.”  
“I really like these mnemonics. They are helpful to remember the katakana.” |

38
Since audio could cause frustration due to different personal preferences and it could introduce noises to the quantitative experiment, audio was turned off by default for the quantitative experiment, and participants could turn it on using a toggle button on the screen.

**Practicing with words that sound like English words**

There were no comments from interviewees about the fact that the words sounded like English words or whether it was helpful for them. 5 out of 8 participants (3 novice learners and 2 current learners) preferred the word version over the character version. They felt that learning with real words was more interesting and challenging than learning with single characters. One current learner pointed out that they were already practicing reading katakana in real life context, e.g. reading cereal brand names, and therefore would definitely prefer the word-based version.

1 novice learner and 1 current student preferred the character version over the word version. One stated that the character version had more “visual density” than the other version. They got to see the mnemonic of every character that they interacted with, whereas at the end of the word, they only saw the word card that did not have visualizations. The other participant stated that they liked the simplicity of the character over the word based version. Both of them

<table>
<thead>
<tr>
<th>Sample negative comments</th>
<th>“I was annoyed… not annoyed but surprised when I heard the audio”</th>
<th>“The image is not relevant to the character and doesn’t make sense.”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“When I practice with the characters, I just want to focus on typing in the characters. The audio is kind of distracting.”</td>
<td>“This mnemonic compares the character to Hiragana, but I don’t know Hiragana.”</td>
</tr>
</tbody>
</table>
suggested that there can be another version that mixed both versions together: users could practice with character firsts, and then move on to practicing with words.

1 novice learner liked both versions equally and had a similar suggestion to combine the 2 versions.

The need for a more comprehensive introduction and explicit structure of the modules

When asked to compare the two tools, or when seeing the second tool, more than half of the participants commented that the word-training tool seemed to be the “next step” that followed the character-training tool. Two learners were surprised to see the screen of the word version and had to ask whether they missed any screens. These participants interacted with the word version first, but they expected to learn separate characters when they started.

Three participants who had not taken Japanese before suggested that the characters displayed in the tool should follow alphabetical order so that the tool would be more beginner-friendly. One suggested that they should be able to see what new characters they were going to learn.

5.2. Results from quantitative experiment

5.2.1. Self-efficacy

Even though the self-efficacy score for the questions based on the MLSQ q increases less for word version, there is no statistically significant difference between the change in self-efficacy in the two versions. The difference in self-efficacy change is smaller to none for the questions that were derived directly from the definition of self-efficacy.
Figure 15: Change in self-efficacy based on the MLSQ questions. Higher score change is desired. The error bars show the standard errors.

T-statistic: 1.595, p-value: 0.116

Figure 16: Change in self-efficacy based on the participants’ confidence in their ability to recognize katakana characters. Higher score change is desired. The error bars show the standard errors.

T-statistic: -0.059, p-value: 0.953
Figure 17: Change in self-efficacy based on the participants’ confidence in their ability to recognize katakana words. Higher score change is desired. The error bars show the standard errors.

5.2.2. Learner’s excitement

Learners’ excitement to learn katakana decreased for both versions, but decreased more for the word version. However, the difference is not statistically significant.
Figure 18: Change in learners’ excitement to learn Japanese katakana. Higher positive score change is desired. The error bars show the standard errors.

5.2.3. Workload measurement

There is no statistically significant difference between the perceived workload in the two versions.

Figure 19: Participants’ rating for mental demand. Lower score is desired. The error bars show the standard errors.
Figure 20: Participants’ rating for temporal demand. Lower score is desired. The error bars show the standard errors.

T-statistic: -1.105, p-value: 0.274

Word-based

Character-based

Figure 21: Participants’ rating for their frustration. Lower score is desired. The error bars show the standard errors.

T-statistic: -0.474, p-value: 0.637

Word-based

Character-based
Figure 22: Participants’ rating for their task performance. Higher score is desired. The error bars show the standard errors.

5.2.4. Technology Acceptance Model

The two versions had similar technology acceptance by the participants.

Figure 23: Participants’ rating for helpfulness. Lower score is desired (due to the rating system of the questionnaire). The error bars show the standard errors.
Figure 24: Participants’ rating for ease of use. Lower score is desired. The error bars show the standard errors.

5.2.5. Objective measurement of participants’ performance

Figure 25: Participants’ levels after using the app for 20 minutes as measured by the level they have reached in the app. The error bars show the standard errors.
Figure 26: Number of new characters learned based on the difference between pre-training and post-training. The error bars show the standard errors.

T-statistic: 1.993, p-value: 0.051

Figure 27: Score difference regarding word recognition for pseudo-words of length 2 and 3. The error bars show the standard errors.

T-statistic: 0.93, p-value: 0.357
Figure 28: Score difference regarding word recognition for pseudo-words of length 4 and 5. The error bars show the standard errors.

5.3. Results from qualitative indirect observation of users’ interaction

General observations

Participants that received the word version reviewed the mnemonics at the end of each word while participants of character version sped through the tool without reviewing the cards.

It took a while for some participants to figure out what to do with the word based tool: they clicked on some other clickable components of the user interface before starting to type.

Observing the interactions of non-respondents

2 participants in both versions experienced a bug that prevented them from proceeding to the next level. Others quit after less than a minute interacting with the app. Here are some sample observations that could have potentially affected the quantitative results
6. Discussion

6.1. Discussion

6.1.1. Hypothesis 1 and learners’ self-efficacy

**Hypothesis 1.** Learner’s self-efficacy will generally increase after practicing with word reading drills.

As shown in section 5.2.1, there is no statistically significant difference between the change in self-efficacy in the two versions. This result does not support hypothesis 1.

A factor that could have affected the outcome of self-efficacy measurement is the significant difference in the mastery experiences participants gain in each version. Specifically, from the result of objective measurements of participants, those that practiced with the character-based version managed to practice with more characters and as a result, recognized more characters and gained more mastery experiences. In section 2.1, I discussed how mastery experiences could affect learners’ self-efficacy. As a result, it remains to be explored whether balancing the mastery experiences that learners have gained through using the tool can further

<table>
<thead>
<tr>
<th>Version</th>
<th>Observed behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>character</td>
<td>They interacted normally with the tool but quit after 10 minutes</td>
</tr>
<tr>
<td>word</td>
<td>They did not seem to understand the task (typed in the mnemonic)</td>
</tr>
<tr>
<td>word</td>
<td>They interacted normally with the tool but quit after 16 minutes</td>
</tr>
<tr>
<td>word</td>
<td>They used the app for 20mins but didn't open the form</td>
</tr>
</tbody>
</table>
affect learners’ self-efficacy. I proposed this exploration because in the real learning context, learners will eventually have to learn all the characters at an early stage of their study to move on to the next learning materials. Balancing the mastery experience can help with investigating whether and how learners’ self-efficacy differ when given the same quantitative expectation.

6.1.2. Hypothesis 2 and the effectiveness of different learning approaches

**Hypothesis 2.** Word-based reading drills using words that sound similar to English words will be at least as effective as learning individual characters in terms of helping learners to read Japanese words rapidly and accurately.

This hypothesis is weakly supported regarding helping learners with accuracy. In terms of word recognition, the difference regarding word recognition scores measured in the quizzes’ results is not statistically significant. However, on average, participants who received the word-based version scored less than the other group.

This hypothesis is not supported regarding helping learners learn rapidly. Participants in the quantitative experiment who received the word-based version progressed slower than those who received the character-based version in terms of module level. The difference in the levels participants were able to reach is statistically significant (t-value = 2.517, p-value = 0.015). In addition, the difference in the number of new characters learned between the character-based version and the word-based version is statistically significant (t-value = 1.993, p-value = 0.051). In general, when given the character-based version, learners progressed through significantly more levels, as a result, learned more characters, and recognized significantly more characters.

From the result of qualitative interviews and indirect observation, one factor that might have created the difference between participants’ progress is how they interacted with the tools. Users usually had to spend more time to learn how to interact with the word-based version.
compared to the character version. Users of word-based versions were also more likely to review
the mnemonic cards at the end of each word compared to the other group, who sped through the
levels. These interactions might have slowed down users’ learning progress. Regarding
participants’ interaction with the tools, task definition might have also impacted their behavior.
Participants were given a general task - to learn katakana using the tool. A more specified task
with explicit constraint, such as “learn as much katakana as possible in the given time using the
tool” could have resulted in participants’ behavior and attitude towards the two versions.

These results show initial indications that participants who practiced with character
recognition drills showed less effective skill translation from character recognition skills to word
recognition skills. Even though they managed to learn significantly more new characters, their
word recognition score (using the same scoring scale for character recognition) was not
significantly higher. However, a more extensive quantitative experiment should be done to
validate this aspect.

6.1.3. Hypothesis 3 and students’ engagement

**Hypothesis 3.** Learning real Japanese words helps learners feel more engaged in learning
Japanese characters compared to learning individual characters.

This hypothesis is not supported. Based on quantitative results, participants in both
groups showed decreased excitement to learn katakana. Even though the decrease is higher for
word training, there is no statistically significant difference between the two versions regarding
this aspect.

As discussed in section 5.1, results from qualitative interviews suggested that learners
might not notice or feel that they benefit from learning words that sound similar to English
words. However, they did comment on the fact that practicing with words can be more interesting and challenging.

6.1.4. Hypothesis 4 and learners’ subjective perception of the tasks

**Hypothesis 4.** Learners perceive learning characters at the word level as more cognitively demanding than learning separate characters in terms of the task itself, and of their experience using the tool.

The hypothesis is not supported. There is no significant difference between participants’ subjective ratings in terms of workload and technology acceptance between the two versions. Most areas received similar subjective ratings, except for “temporal demand”, where the average rating suggests that participants perceived their task to have a more “hurried or rush” pace when given the word-based version. This difference in temporal demand is relevant to the result discussed in the evaluation of hypothesis 2 where I discussed the fact that participants that received the word-based version progress significantly slower than the other group.

However, this result can be subject to non-response bias. Indirect observation of non-respondents’ interactions with the tools showed that some participants with the word version might have been confused by the tool, or were disengaged, but did not proceed to finish the quantitative experiment. In addition, among the people that did not finish the experiment, the majority received the word version. Also regarding potential confusion and frustration, results from qualitative interviews suggested that learners faced initial confusion when attempting to use the word-based version because it was not immediately clear how they should interact with the tool to finish the reading tasks. Yet, the interviewees were all able to interact with the tool without asking how to use the tool and expressing frustration.
Another factor that arose from the results of qualitative interviews and from the feedback section of the quantitative is how participants view the tool as more appropriate for a later phase of beginning studying a language. Qualitative interviewees believed that the word version was the sequel of the character version, and gave suggestions to add character recognition drills to the word version. In the quantitative experiment, some comments from the participants that received the word version mentioned that the tool seemed to benefit “intermediate students” more.

6.2. Limitation of the study and future development

The results of this study only provide insights into learners’ impression and performance after a short exposure to the tool developed. The constant variable in the quantitative experiments was how long the users were exposed to a tool, and the qualitative interviewees also only used the tool for a few minutes. On one hand, this helps provide a baseline to compare a snapshot of users’ impressions and acceptance of the tools. On the other hand, as discussed in the previous sections, different interactions can interfere with participants’ interactions with the tool within such a short time, and this evaluation alone is not enough to conclude the evaluation of introducing computer-assisted word reading drills to beginners. It remains to be seen how longer exposure to the tool could affect learners’ performance and self-efficacy. If limited to a shorter time constraint, module level in the tool can be used as the constant variable: learners have to master a satisfactory number of modules in order to finish the experiment. This can help to further explore how different forms of mastery experiences can affect learners’ self-efficacy.

The majority of the participants were not learning or attempting to learn Japanese. This helped to reduce interference from other factors from the course experience that might have affected the participant’s opinions. However, as discussed in the previous section, learners can
potentially benefit more from using the tool if they have already gained a basic foundation for the writing system. It remains to be seen whether collecting opinions from more learners who have started to study Japanese would yield different results. For example, while people who are taking a course have the motivation to memorize katakana in order to be able to continue with their course, people who have no intention of studying Japanese used the tool lack such realistic motivation. The differences, if any, can be used to investigate when and how different computer-assisted reading drills can benefit learners at different levels.

The results of this study are mainly specific to the implementation of the proposed design. Other designs and implementations can be proposed to explore the effects of using computer-assisted training tools to expose learners to reading drills with different degrees of complexity at early learning stages.

6.3. Conclusion

In this study, I have proposed, implemented, and evaluated a novel use of technology for learning Japanese katakana that enables the use of early word reading drills for beginners that combine the advantages of the two methods for teaching early reading skills that are commonly adopted in Japanese language education: (1) quickly gaining an overview of all the characters and start early with reading drills and (2) learning subsets of characters to guarantee mastery of character recognition before moving on to the next learning materials.

Participants in the quantitative experiment, 77% of whom had never learned Japanese before, were able to practice with word-reading drills as their first step to learn katakana with approximately similar ratings in perceived workload and technology acceptance compared to the tool’s counterpart that targets character recognition. In addition, results from the quantitative
study suggest that word recognition skills might have been gained more effectively when learners practiced with word reading drills. These results suggest that the tool as currently designed or similar computer-assisted training tools could be further developed and explored to assist beginners with learning katakana effectively.

The proposed learning drill can be used along with a curriculum where learners are expected to memorize the characters by themselves after receiving basic instructions about the Japanese writing system. Results from the qualitative interviews and answers by participants to optional feedback questions at the end of the survey indicated that participants felt that they would benefit more from the tool if they could receive a more formal and systematic introduction to the Japanese writing system prior to using the word-based training tool. A comprehensive introduction in their course curriculum could help learners gain the foundation and, as a result, confidence to conquer more challenging practice drills, while the competency-based training offered by the tool could help learners consolidate the large number of new characters introduced in their course more effectively. Further study could be conducted to explore these proposed benefits.

If used as a standalone learning tool, the tool should include a more comprehensive walkthrough of the larger picture of the Japanese writing system so that learners feel they have a more solid starting level. The tool could also be redesigned to introduce character recognition drills before showing word reading drills, as some participants suggested.

Further study would also be desirable regarding the possibility of building upon the concepts of module-based mastery learning proposed in this tool to propose other designs for computer-assisted word reading drills that may be less intimidating for beginners.
7. Appendix

7.1. Pre-training survey
Why we are doing this research? The primary goal of this project is to develop and evaluate a tool that helps with learning Japanese characters.

What you will have to do? You will be asked to participate in a training session on a web app and fill out two forms before and after the training. In the forms, you will also be asked to answer demographic and survey questions, which include your age, gender identity, and your experience with learning Japanese.

What you will get out of it? You will gain early access to a new learning tool. You may learn new knowledge about Japanese and/or reinforce what you have already known about the language. Participants who have finished all parts of the experiment will receive an Amazon gift card.

Privacy and Data Collection: Your responses will be kept anonymous. We do not store any information that could be used to directly establish your identity. We may, however, collect browser data, the city and country you are located in, mouse and click data, social media sharing activity, scrolling behavior, scores earned, timestamp, time spend on each page, or other data that may be derived from your behavior on our page. This data will be stored securely on our servers.

Are there any risks? We do not expect there to be any risks. You are also free to end the experiment at any time by closing your browser window.

Expected Duration: 30 minutes.

Contact Information: If you have questions about this research or if you think you have been harmed by participating, you may contact Lien Tran at Harvard College at lien_tran@college.harvard.edu

Funding: This research is supported by the Harvard College Research Program.

I agree to participate in this research. Participation in this research is voluntary, and I can stop at any time without penalty. I feel that I understand what I am getting into, and I know I am free to leave the experiment at any time by simply closing the web browser.
Please check off **both** the to-do items that are required for this experiment

- I will finish all parts of this experiment in one sitting (from filling out this form to receiving a message that the experiment is over)
- I will finish this experiment on my computer
How old are you? (Please type a number)

How long have you learned Japanese?

- I haven't learned Japanese before
- I've been teaching myself the basics
- <1 year attending Japanese courses
- At least a year

Gender

- Male
- Female
- Non-binary
Thank you for your interest in participating! We are currently recruiting participants who are at least 18 years old and are at beginner’s level. Please still submit the form if you are interested in learning more about the progress of the tool in the future!

Please also forward this form to the people you know so that they can participate in the research. Thank you!

https://harvard.az1.qualtrics.com/jfe/form/SV_d45C1nbeVeHbF3
The next section of the form is timed.

Please enter the romaji/ English syllable of the Katakana characters that you can recognize as quickly and accurately as possible.

If you don't know a Katakana, just leave the answer blank. You can click on the next arrow to view the next questions.
End of timed section. The next section of the form is also timed. Please enter the romaji/English syllables of the words written in Katakana as quickly and accurately as possible. The characters will be shown on several question blocks. You can click on the next arrow to view the next questions.
マク

デス

ミキ

セロト

フクリ

イラニ
キロラム

ムイラス

キムハラ

トウタル

モノタロ

カラテシロ
Please rate whether the following statements apply to you (1=absolutely incorrect, 7=absolutely correct). For your reference, the Katakana chart is displayed at the end of the page.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>I believe I correctly</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>answered all of the</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Katakana words in the last</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I’m certain I can read the</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>most difficult Katakana</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>words in the app</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I’m confident I can read</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the basic Katakana words</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in the app</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel confident that I can</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>learn to read Katakana</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>characters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel confident that I can</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>learn to read words</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>written in Katakana</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel excited to learn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japanese Katakana</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**KATAKANA**

<table>
<thead>
<tr>
<th>ノ</th>
<th>ソ</th>
<th>ッ</th>
<th>ニ</th>
<th>ユ</th>
<th>エ</th>
</tr>
</thead>
<tbody>
<tr>
<td>コ</td>
<td>リ</td>
<td>フ</td>
<td>ラ</td>
<td>マ</td>
<td></td>
</tr>
<tr>
<td>ヨ</td>
<td>ル</td>
<td>ワ</td>
<td>フ</td>
<td>ム</td>
<td></td>
</tr>
<tr>
<td>ロ</td>
<td>レ</td>
<td>ム</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ニ</td>
<td>ク</td>
<td>ナ</td>
<td>テ</td>
<td>メ</td>
<td>サ</td>
</tr>
<tr>
<td>ヤ</td>
<td>タ</td>
<td>モ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ポ</td>
<td>カ</td>
<td>ス</td>
<td>イ</td>
<td>ヘ</td>
<td></td>
</tr>
<tr>
<td>ハ</td>
<td>オ</td>
<td>ネ</td>
<td>ア</td>
<td>ト</td>
<td></td>
</tr>
</tbody>
</table>
CLICKING NEXT WILL FINISH THIS FORM - MAKE SURE THAT YOU HAVE OPENED THE LINK BEFORE YOU CLICK NEXT

You will have to use the tool on a computer and sign up using the same email address that you use in this form.

After you have used the tool continuously for 20 minutes, a popup will appear in the tool with a URL to another form. You will have to finish the second form that pops up in the app in order to finish this experiment.

If you run into issues while running the app, refresh the page and sign into your account. If the problem persists, please click “Help” and fill out a form so that we can assist you with the experiment.

Here’s the link to access a version of the training tool: https://tale97.github.io/testRepo/

I have read and understood the instruction. I will open the training tool in another window and submit this form.

You have finished form 1. Please remember to use the tool for 20 minutes on your computer and fill out form 2 that pops up in the app. You will only receive compensation after finishing form 2.
7.2. Post-training survey

Welcome back! What email did you provide in the previous form?

What level were you at in the app?
To aid me in learning how to read Katakana, overall, I feel that using the web app as an e-learning tool is:

<table>
<thead>
<tr>
<th>Feature</th>
<th>Rating</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficient</td>
<td>★★★★★</td>
<td>Inefficient</td>
</tr>
<tr>
<td>Performance enhancing</td>
<td>★★★★★</td>
<td>Performance degrading</td>
</tr>
<tr>
<td>Effective</td>
<td>★★★★★</td>
<td>Ineffective</td>
</tr>
<tr>
<td>Helpful</td>
<td>★★★★★</td>
<td>Unhelpful</td>
</tr>
<tr>
<td>Quite useful</td>
<td>★★★★★</td>
<td>Quite useless</td>
</tr>
<tr>
<td>Easy to learn</td>
<td>★★★★★</td>
<td>Difficult to learn</td>
</tr>
<tr>
<td>Easy to manipulate</td>
<td>★★★★★</td>
<td>Difficult to manipulate</td>
</tr>
<tr>
<td>Clear to interact with</td>
<td>★★★★★</td>
<td>Obscure to interact with</td>
</tr>
<tr>
<td>Easy to master</td>
<td>★★★★★</td>
<td>Difficult to master</td>
</tr>
<tr>
<td>Very Usable</td>
<td>★★★★★</td>
<td>Very cumbersome</td>
</tr>
</tbody>
</table>
Given the task to learn Katakana using the tool...

How mentally demanding was the task?

How hurried or rushed was the pace of the task?

How successful were you in accomplishing what you were asked to do?

How insecure, discouraged, irritated, stressed, and annoyed were you?
The next section of the form is timed. Please enter the romaji/English syllable of the Katakana characters that you can recognize as quickly and accurately as possible. The characters will be shown on several question blocks. If you don’t know a Katakana, just leave the answer blank. You can click on the next arrow to view the next questions.
End of timed section. The next section of the form is also timed. Please enter the romaji/English syllables of the words written in Katakana as quickly and accurately as possible. The characters will be shown on several question blocks. You can click on the next arrow to view the next questions.
マス

テラ

ミア

ロセト

コリア

ニクリ
ハルラム

クラレス

トムハラ

ホシテル

カラロテス
Please rate whether the following statements apply to you (1=absolutely incorrect, 7=absolutely correct). For your reference, the Katakana chart is displayed at the end of the page.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>I believe I correctly answered all of the Katakana words in the last test</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>I’m certain I can read the most difficult Katakana words in the app</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>I’m confident I can read the basic Katakana words in the app</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>I feel confident that I can learn to read characters written in Katakana</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>I feel confident that I can learn to read words written in Katakana</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>I feel excited to learn Japanese Katakana</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

KATAKANA

<table>
<thead>
<tr>
<th>ノ</th>
<th>ソ</th>
<th>ッ</th>
<th>ニ</th>
<th>ユ</th>
<th>ニ</th>
<th>エ</th>
</tr>
</thead>
<tbody>
<tr>
<td>コ</td>
<td>リ</td>
<td>フ</td>
<td>ラ</td>
<td>マ</td>
<td>マ</td>
<td>マ</td>
</tr>
<tr>
<td>ヨ</td>
<td>ル</td>
<td>ワ</td>
<td>メ</td>
<td>ム</td>
<td>ム</td>
<td>ム</td>
</tr>
<tr>
<td>ロ</td>
<td>レ</td>
<td>ラ</td>
<td>ヌ</td>
<td>ヌ</td>
<td>ヌ</td>
<td>ヌ</td>
</tr>
<tr>
<td>ヒ</td>
<td>ク</td>
<td>ナ</td>
<td>テ</td>
<td>メ</td>
<td>メ</td>
<td>メ</td>
</tr>
<tr>
<td>セ</td>
<td>ケ</td>
<td>チ</td>
<td>サ</td>
<td>サ</td>
<td>サ</td>
<td>サ</td>
</tr>
<tr>
<td>サ</td>
<td>タ</td>
<td>ム</td>
<td>キ</td>
<td>キ</td>
<td>キ</td>
<td>キ</td>
</tr>
<tr>
<td>ボ</td>
<td>カ</td>
<td>ス</td>
<td>イ</td>
<td>ヘ</td>
<td>ヘ</td>
<td>ヘ</td>
</tr>
<tr>
<td>ハ</td>
<td>オ</td>
<td>ネ</td>
<td>ア</td>
<td>ア</td>
<td>ア</td>
<td>ア</td>
</tr>
</tbody>
</table>
The questions on this page are optional and will help us improve the design of the learning tool and the experiment.

Do you have any other comments on the tool?

What are some other things that you’d like to see in the tool?

What other comments/suggestions do you have?

Clicking next will submit this form and finish the experiment.


