MEASURING ACCURACY AND SPEED OF L2 KANJI WORD IDENTIFICATION

YOKO MATSUMOTO-STURT
UNIVERSITY OF EDINBURGH
This paper investigates the quality of interlanguage lexical representations in reading kanji (Chinese characters) compound words. Within second language (L2) Japanese pedagogy, it has been well established that learning to read and write Japanese orthography, especially kanji, is an outstandingly difficult task for learners from non-kanji background.

This study includes the first demonstration of frequency effects on naming latencies of kanji compounds for English-speaking (EJ) learners of Japanese at different proficiency levels: literate novice learners (n=18), and skilled advanced learners (n=11).
In order to test accuracy and speed of word pronunciation, we need to consider the frequency of the words, as this is a fundamental factor in word recognition (e.g. high frequency words are pronounced faster and more accurately than low frequency words). This created a methodological problem for the present study. First of all, there is no previous literature in L2 acquisition of Japanese that deals with frequency issues of kanji compound words. Next, it is difficult to be convinced that L1 word frequency databases are relevant to our literate beginners, whose language exposure is limited to classroom learning (i.e. tutored acquisition).

Accordingly, this study employs an original L2 frequency database, from which test words are drawn for two reading tasks (the pronunciation task and the naming task). I found several lines of evidence for the well-documented L1 frequency effect in the L2 naming task.
ABSTRACT

Firstly, a significant negative correlation confirmed that EJ readers' naming latencies decreased as word frequency increased, and vice versa.

Secondly, we found evidence that the rate of correct pronunciation was correlated positively with word frequency in the novice group. This line of evidence suggests that word frequency affects the quality of early interlanguage lexical representations.

Thirdly, I found evidence that a low-frequency ambiguous word caused an L2 reading difficulty, which may not be associated with a particularly high latency (i.e. speed-accuracy tradeoff). Since the Second Language Acquisition (SLA) literature has just begun exploring the frequency effect on L2 reading, this finding suggests a potential interest in the use of psycholinguistic-oriented methodologies for SLA research such as measuring Reaction times of L2 readers.
OVERVIEW

- Introduction to Japanese orthography
- The problem of reading/pronouncing kanji
- Teaching and learning background
- Background to the experiment
- Construction of word-frequency database
- Experimental setup
- Results
INTRODUCTION TO JAPANESE ORTHOGRAPHY

- Phonetic kana (syllabary)
  - Hiragana  e.g. いもうと  i-mo-o-to  younger sister
  - Katakana  e.g. ブリストル  bu-ri-su-to-ru  Bristol

- Kanji (Chinese characters)
  - e.g. 妹  younger sister
Multiple readings for any given kanji
  - Example: 一日

No possibility of sub-lexical reading

Kanji has no (reliable) phonetic cue

Many kanji compounds
  - These are words: vocabulary learning
  - Often whole word pronunciation differs from sum of parts
- University of Edinburgh degree course
- 4 years, with 3rd year spent in Japan
- 1st year students can start from scratch
- All 4 skills covered in all levels
Mora deletion/insertion errors common in L2.
- Long-short vowel distinction is notoriously difficult for L1 English learners of Japanese

Previous studies have not considered word frequency.
- Word frequency is known to affect L1 processing

Experiment part of a larger study examining the relation between spoken and written production.
- Today I will discuss only the spoken part
CONSTRUCTION OF A FREQUENCY DATABASE

- L1 corpora limited;
  - Unrealistic for classroom learners
- I computed raw word counts from 2 textbooks
  - ICU textbook for College Students.
  - Basic Kanji Book
- Textbooks are used by 1st & 2nd year students.
Experimental materials drawn from frequency database
- 30 experimental words plus 40 fillers
- All experimental words included a long /o:/ sound balanced equally between word-initial, word-medial and word-final positions

Tasks:
- **Naming task** (name kanji in isolation)
- **Pronunciation task** (pronounce the same kanji embedded in a frame sentence)
3. 妹

これは_____と読みます。
Stimuli presented on computer
Responses digitally recorded
  - Naming time and pronunciation accuracy were recorded for later analysis
Subjects:
  - 18 2nd-year students (novice group)
  - 11 4th-year students (advanced group)
Subjects tested individually in a quiet room
Mean latency differed between groups (p < .001)
Scatter plot of latency x frequency

Item by item means (Novice data)

$r = -0.62, p < 0.05$
- Significant negative correlation for novice group \( (r = -0.62, \ p < 0.05) \).
- Marginal correlation for skilled group \( (r=-0.17, \ p < 0.09) \).
Positive correlation between frequency and accuracy for novice group \((r= .51, \ p < .05)\)

No frequency/accuracy correlation for advanced group \((r = .17, \ p = .37)\)
Accuracy x Latency scatterplot (all subjects)

Percentage correct pronunciation

(all subjects)

$\text{Correct pronunciations}$

$\text{Reaction times}$

$r = -0.62, p < .05$
Negative correlation between accuracy and naming latency for combined group data \( r = -0.62, p < .05 \)

There seems to be a single outlier in the scatterplot
Accuracy x Latency scatterplot (all subjects)

Percentage correct pronunciation (all subjects)

\[ r = -0.62, \ p < 0.05 \]
Accuracy x Latency scatterplot (all subjects)

\[ r = -0.62, \ p < 0.05 \]
The outlier is *isshoo* (= one’s life).
- This word was very rarely pronounced accurately, but was named relatively quickly.
- Speed/accuracy tradeoff
- The word is also very low frequency
CONCLUSIONS

- Frequency effect played a role in L2 word recognition
- Low frequency exceptional word caused L2 reading difficulty regardless of proficiency levels with relatively short RT
- Shows viability of using frequency and on-line methods in L2 research