

# Developing an Online Learning Material for Kanji Vocabulary

Midori Indoh, Keisuke Tanatsugu,  
Tomomi Tabuki, Yayoi Ohji

(Midori Indoh, Tomomi Tabuki, Yayoi Ohji: Center for International Education and Exchanges)/(Keisuke Tanatsugu: Honorary professor, The University of Kitakyushu)

## Keywords

Online learning materials, Kanji vocabulary, Text, Reading drills, Writing drills

## Summary

Although foreign students who come to Japan are highly motivated to learn the Japanese language, they often meet with great difficulties that set them back. Mastering the reading and writing of the kanji vocabulary seems to present especially great hardships to them. Therefore, it is very important to provide a learning environment that will enable them to engage in serious learning wherever and whenever they wish. For this reason, we decided to develop an online learning program that will motivate and assist self-learners. In this paper, we will discuss the significance of this software system, describe the system's structure, and summarize the program.

## 1. Introduction

For non-native speakers of Japanese, reading and writing is always more difficult than listening and speaking. This is especially true in the matter of kanji. In addition to the complicated shapes and stroke orders of kanji, in many cases, a single character has multiple readings<sup>1</sup> which are determined by how the character is used in combination with other

---

1 For example, the Chinese characters, “上” and “下” can each be read in eight different ways.

## Developing an Online Learning Material for Kanji Vocabulary

characters or words.<sup>2</sup> Also, since Japanese sentences are written both with kanji that represent the meaning of words and with “kana” (phonetic alphabets), the result is extreme variety and complexity when it comes to reading.

In order to write a certain language, there exists a set of symbols to be used. This set is called the “character set”<sup>3</sup> of this Language. Japanese character set consists of hiragana, katakana, kanji, and various other symbols. The size of the character set will vary greatly based on whether kanji is included.

Currently, approximately 2000 kanji are formally recognized as Jōyō Kanji (regular use kanji), since they are necessary for daily life and must be learned during compulsory education. Furthermore, kanji to be used in the cyber environment are determined by the JIS standard, in which there are approximately 6000 kanji, consisting of First Standard and Second Standard characters. For this reason, there are fifty times more characters in Japanese than in English, which only uses alphabets.

For example, let us assume we are randomly creating a character string of length 2. If there are about 100 English characters, we can generate 10,000 different strings. Assuming there are approximately 5,000 Japanese characters, there will be 25 million possible strings. Therefore, even for such a short string, there are potentially 2500 times more variety in Japanese to express an idea, than in English.

However, in reality, we only use a small portion of potential string combinations, and, therefore, having many characters does not necessarily mean it is better. On the contrary, a language with many characters can be tedious and inefficient to use, and its complexity can be an obstacle for learners.

If that is the case, how can we efficiently teach kanji vocabulary to them?

Foreign students will have already attained a certain level of Japanese when they enter school. They will also be required to attend and understand lectures, and write reports and theses, in Japanese. Furthermore, they would have enough reading and writing abilities to carry on daily lives in Japan, without too much emotional stress. Regardless of their target level of accomplishment, they all need to make conscious efforts to master kanji vocabulary efficiently and in a short period of time.

With the above points in mind, we decided to develop an online program for learning kanji

---

<sup>2</sup> A single word may have multiple meanings.

<sup>3</sup> For some languages, it is difficult to define “character set”.

vocabulary, and with the goals outlined below. Item(1) is especially indispensable if learners are to improve their skills efficiently and quickly.

- (1) To provide an environment where students can learn anytime and anywhere.
- (2) To allow students to learn kanji words used in their daily lives repeatedly.
- (3) To guarantee an interactive learning process that stimulates the learner's interest.
- (4) To create a flexible program that fits a learner's level of interest and ability.

## 2. System Summary

All pages consist of two frames, Frame 1 showing the title and menu, and Frame 2 showing the contents selected or specified by the menu. Image 1 shows the first page.

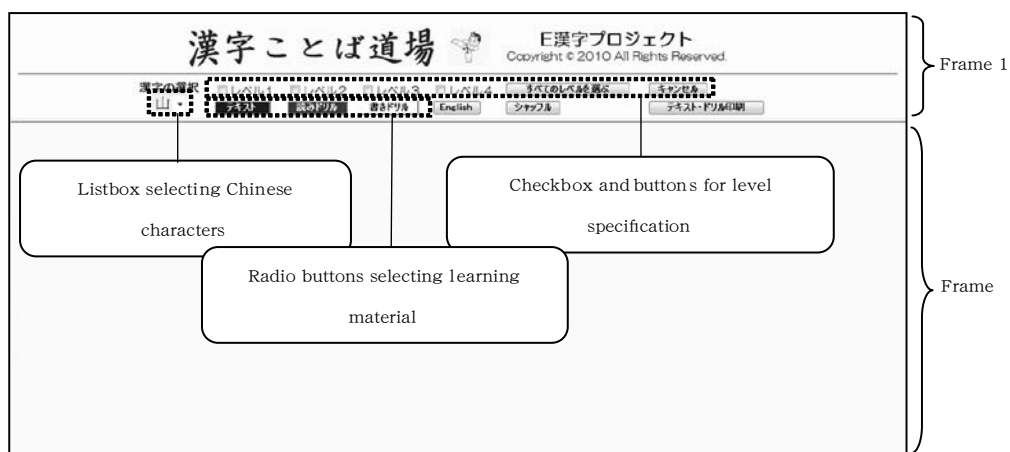


Image 1. Top page (partial view)

### 2.1. Structure of the Menu

So that a learner can easily select his/her learning process according to his/her needs, the following buttons and boxes were created for the menu.

#### Required menu for displaying learning materials

- (1) List box for choosing kanji

The user chooses one kanji from the list.

- (2) Check box for selecting the learning level: [LEVEL1] to [LEVEL4], [ALL SELECT] and [ALL CANCEL].

User can select multiple levels simultaneously.

## Developing an Online Learning Material for Kanji Vocabulary

In order to facilitate the selection process, two buttons were added, [ALL SELECT] and [ALL CANCEL].

(3) Radio button to select the learning materials: [TEXT], [READING DRILL] and [WRITING DRILL]

The selected learning materials will be displayed in Frame 2 in their designated format.

### Option Menu

(4) Button for display in English [ENGLISH]

To aid users who are not proficient enough in Japanese yet, the vocabulary words and examples can be shown in English. However, the function of this button is still at the trial stage and is, therefore, applicable only to a limited number of kanji.

(5) Shuffling button for rearranging the items [SHUFFLE]

By shuffling and rearranging items to be learned, learners can design their own drills.

(6) Button [PRINT TEXT OR DRILL]

By clicking this button, Frame 2 will be selected for printing. A message will also appear indicating how to adjust the page for proper printing.

## 2-2. Displaying the learning materials

Learning materials corresponding to the desired level will be displayed on the screen as follows. The following image is an example when all levels have been selected.

(1) Course materials [TEXT]

Learners read the materials. Up to Level 4 can be displayed.

The maximum number of vocabulary words are six for Levels 1-3, and ten for Level 4.



Image 2-1. Text (upper page)



Image 2-2. Text (lower page)



### **3. XML and DHTML**

#### **3-1. Construction of a Kanji Lexicon Database based on XML**

The XML (Extensible Markup Language) technology conceptually separates the data description part and the look description part of an HTML (Hyper Text Markup Language) document, and denotes a method to describe data with precise structure, and a method to display them efficiently on the Web.

Its features are as follows:

- (1) Can use tags just as in HTML, and allows data to be described in plain text
- (2) Can describe data structures containing multiple layers
- (3) Permits tags to be freely created, allowing the contents of the data to be described by tag names, and their meanings shared
- (4) Allows the description of data to be standardized
- (5) Simplifies the storage, maintenance, and updating of data

Suppose we wanted to display, on the Web,  $n$  number of data whose structure is clearly defined and whose format is the same. If we do not use XML, we must prepare an  $n$  number of HTML documents that include long scripts to realize interactivity with the user.

The greater the number  $n$ , the more complex the task and the more likely that errors will occur. Also, if a change is needed in the data structure, it would become necessary to change and check all  $n$  number of HTML documents.

By using XML, we can create one HTML document including the program script and the visual description. This allows changes to be made more quickly and accurately, while greatly reducing the amount of work required.

#### **3-2. Dynamic content using DHTML**

DHTML (Dynamic HTML) is a collective term for the technology that allows motion and interaction within the static HTML. This is made possible by embedding within the HTML document, programs that use scripts.

More than 90% of the portion of the HTML file (ekANJI.html) that constitutes the core of the Kanji Lexicon Learning System was written using script language.

#### **3-3. The flow of the Program**

The following is a summary of the program flow, from the display of the top page to the

display of individual lessons.

(1) The top page is displayed.

- a. The Kanji List File (kanjilist.xml) is loaded.
- b. The file defining the frame (frame.html) is loaded, and the title and menus are displayed in Frame 1.

(2) Individual events are processed based on the options selected and the buttons clicked in the menu. The following are especially important:

- a. When a kanji character is chosen in the Kanji List, its corresponding XML file (such as 2-1sun.xml) is opened, its contents are read and then pre-processed for display. At the same time, the image file for the stroke order (for example, 2-1sun.jpg) is loaded.
- b. When either the [TEXT], [READING DRILL], or [WRITING DRILL] buttons are clicked, the text corresponding to that event is immediately generated and injected into the “body” of the HTML document, and its contents are then displayed.

Finally, let us describe the role of the program in this system. One kanji character corresponds to one XML file. The reason the display differs between the Text, Reading and Writing Drills is due to the handling by the program. By using programs within a single HTML file (ekanji.html), it is possible to display a wide range of data based on the needs of the user, while maintaining responsiveness and ensuring the autonomy of the XML data.

#### **4. Conclusion**

The level of complexity for a language is often discussed on its level of difficulty to a foreigner attempting to learn it. In such cases, the “level of difficulty,” which should be an objective index, contains within it “how difficult the language is relative to the native tongue,” which is a totally subjective notion. As is the case with formal languages, we believe that objectivity and generality in the discussion of natural languages can be obtained by discussing complexity from a mathematical approach. This will be a topic for further research and discussion.

Under present circumstances, it is presumed that either the learner or the instructor prints out the drills from the Web learning page, and the answers are manually written down. Naturally, it is ideal to equip the system with a program to recognize handwritten script, so that the answers can be input and immediately rated on the Web page. However, this will be a topic for long-term research, since we foresee difficulties and challenges in realizing such a

system.

## References

- Chiu, Hsueh-Chin 2010. The processing mechanism for Japanese kanji words in learners of Japanese: concerning interlingual similarities in orthography, phonology and semantics. *Journal of Japanese Language Teaching* 146:49-60.
- Fujii, R. 1997. Overview: teaching Kanji. *Journal of Japanese Language Teaching* 94:85-90.
- Kawaguchi, Y. 1993. A communicative approach of teaching Kanji. *Journal of Japanese Language Teaching* 80:15-27.
- Matsumoto, K. 2004. Tayoona ninchi nooryoku-ni sasaerareta gainen shuutoku-no sokumen-kara-no gengo kyooiku. Paper presented at JCLA (The Japanese Cognitive Linguistics Association) Conference, Kansai University.
- Matsumoto, Y. and Oyama H. 2009. Perspectives of language processing technologies on writing assistance and vocabulary learning. *Journal of Japanese Language Teaching* 140:37-47.
- Shimizu, Y. et al. 1998. On improving practical ability in Kanji. *Research Bulletin, The International Student Center of Kyushu University* 9: 95-104.
- Suzuki, T. 2003. Psychological prototype from the viewpoint of the semantic network: through a meaning analysis of the Japanese polysemic verb tsuku. *Journal of Japanese Language Teaching* 116: 59-68.
- Takada, A., K. Nishina and I. Kawashima. 1993. Kinoo-kara mita kihon kanji-no bunseki-to gakushuu-hoo. *Humanities Review, Tokyo Institute of Technology* 19: 88-98.