Systematization and improving efficiency in Kanji learning, based on quantitative analysis involving structural decomposition and coding

Galina Vorobeva

Abstract

This study categorized problems in learning kanji encountered by Japanese language learners (non-kanji background) into three groups: (a) problems within the kanji itself, (b) problems in the teaching method, and (c) problems with language learners' attitude and approach to kanji study. This categorization was based on evidence from Japanese language learners' difficulties with kanji study, evidence regarding Japanese language education in Kyrgyz Republic, as well as previous studies related to Japanese language education. In order to help solve these problems, this study has analyzed the structure of kanji, aimed to produce a more systematic and efficient method of remembering kanji, and used this method to develop kanji teaching materials and a kanji teaching method.

The method of this study includes the structural decomposition of kanji, the coding of kanji and their elements, the creation of a kanji code database, development of an algorithm to use in the structural decomposition of kanji, analysis of kanji in groups, the development of a method of kanji education based on a hierarchical approach, and the production of kanji teaching materials focused on memorization by associating related kanji together.

This study introduced previously unused concepts when analyzing the structure of kanji and kanji groups. These new concepts are as follows: the alphabetical codes of strokes in kanji; constituent elements of kanji that are not radicals as sub-radicals; symbolic and semantic codes for kanji constituent elements; displaying the form of kanji using these codes; indexes by the alphabetic, symbolic and semantic codes; a numerical indicator for the structural complexity of kanji; categorizing kanji based on their complexity, selectivity of kanji indexes and efficiency of kanji lookup indexes (coefficient of selectivity); and the hierarchical structural decomposition of kanji.

This study can be summarized as follows. First, based on the idea that a deeper understanding of the meanings of kanji can be brought about in language learners by kanji decomposition and by considering the meaning of a kanji as a whole relative to its constituent elements, the author introduces the coding of strokes (the smallest visual element of any kanji) and constituent elements (the smallest elements within a kanji that hold their own meaning). Following this coding, analysis was carried out on applications for the stroke and composite element codes, as well as the frequency of each stroke and composite element, etc.

Additionally, on the basis of coding of kanji the author also analyzed efficiency of kanji lookup in kanji dictionaries. For this purpose, a new concept was defined – a coefficient of selectivity of a kanji index – and on the basis of this calculation compared the efficiency of existing kanji indexes. This study also introduced new types of indexes based on the kanji coding mentioned above, and compared their efficiency to previous indexes.

Also, quantifiable characteristics of kanji were analyzed, and criteria for the structural complexity of a kanji were determined based on this quantitative analysis, and from this a numerical indicator of kanji complexity was defined. The 2136 Kanji for Regular Use (*Shin Jōyō Kanji*) were categorized based on this complexity indicator. This is thought to assist the development of a rational order of presenting kanji to language learners.

Based on analysis of the structure of kanji and their groups as described above, possible applications to kanji learning and teaching methods were considered, with discussion on the selection of which kanji to be taught and the order in which to teach them. Analysis of the order in which current kanji teaching materials introduce kanji was carried out based on the kanji codes mentioned

above. This was done by creating a database of kanji codes for each teaching material, and by developing an algorithm and computer program to calculate the number of cases where a more complicated kanji was introduced in teaching materials earlier than one of its simpler components (that is used by itself as a character). Results from this analysis showed that the percentage of such cases was 42.6%.

This study developed a rational order of teaching kanji to language learners using the structural complexity indicator as well as data on usage frequency. This development was founded on the idea that teaching materials based on the principle of introducing structurally simple kanji first followed by more complicated kanji would be easier to use.

Based on the analysis above, a kanji teaching method was proposed for non-kanji background learners. Following a discussion on the characteristic of character recognition peculiar to non-kanji background learners, as well as their motives for learning kanji, a hierarchical approach to teaching kanji was defined. A teaching method using this approach was proposed whereby writing would be taught in the order of (a) strokes, (b) katakana, (c) kanji, using the similarities between the stroke shapes and stroke order of katakana and kanji. Also, teaching methods that develop learners' ability to learn kanji autonomously, as well as study methods using ICT (Information and Communication Technology).

Finally, this paper described plans of writing *One Thousand and One Kanji Stories*, a textbook using an associative memorizing technique, including 1006 kanji.

This study builds upon the results of previous research on the structure of kanji and their teaching methods, and aims to improve their systemization and improve efficiency, as well as produce teaching materials that could be used in classrooms for non-kanji background learners.

Specifics of this study when compared to previous research can be summarized as follows:

- The definition of new concepts regarding kanji.
- A new method for the structural analysis of kanji.
- Improved efficiency in looking up kanji in dictionaries.
- Determination of criteria for the structural complexity of a kanji and categorization of kanji by complexity.
- Analysis of the order of displaying kanji in current textbooks, etc.

This study makes a significant contribution to Japanese education in Kyrgyz Republic with the publishing of the kanji teaching materials *Kanji Stories I, II* (*Kanji Monogatari I, II*) (in Japanese and Russian) and *Kanji Teaching Handbook* (in Russian).

The significance of this study in terms of language policy in Japan can be summarized as follows:

- The constructing of a system of constituent elements *Shin Jōyō Kanji* (the smallest parts within that have their own meaning) is important for a deeper understanding of the structure and meaning of kanji, as well as the development of a systematic memorization method. The system developed in this study, along with previously developed systems, can be used is a reference point in the creation of a standard system of kanji constituent elements.
- Kanji pose a significant difficulty when considering the globalization of Japanese. This study demonstrates a new system for categorizing the 2136 Kanji for Regular Use based on quantitative criteria of their structural complexity. This new system is expected to contribute towards future language policy.

Keywords: Structural decomposition of kanji, Coding, Quantitative analysis, Systemization of kanji study, Numerical indicator of the structural complexity of kanji, Numeral indicator of the efficiency of kanji lookup indexes