Using Affordances in Electronic Chinese/English Dictionaries for Non Chinese-Speakers[1]

Karel Svoboda
Department of Psychology
Carleton University
Ottawa, ON
ksvoboda@chat.carleton.ca

Robert West
Department of Psychology
Carleton University
Ottawa, ON
robert_west@carleton.ca

ABSTRACT
Chinese is a difficult language to learn and it doesn't help that virtually all electronic, Chinese/English dictionaries are designed for Chinese-speakers who are studying English. We found that the most common form of this type of dictionary, the root radical indexing system, is very difficult to use for non Chinese-readers. However, further investigation revealed that this system exploits visual affordances that non Chinese-readers are surprisingly sensitive to. Furthermore, the affordances highlight structural aspects of the characters that are used by experienced readers to identify the characters. Difficulties in using this system seem to be due mainly to missing information that non Chinese-readers require to use the system. We argue that a modified, electronic version of this system could be very effective for non Chinese-readers provided the missing information is supplied.

Keywords
Chinese, dictionary, affordances

INTRODUCTION
Organizing a dictionary for a logographic script, such as Chinese, is a complex problem because an alphabetizing strategy cannot be used. However, Chinese characters have an internal structure that does allow them to be indexed. The majority of Chinese characters are compound characters composed of smaller units, referred to as radicals. The radicals themselves are made up of strokes (i.e., originally brush strokes).

The type of indexing system most widely used by students studying the Chinese language is the root radical method (most available Chinese/English dictionaries use this system**).** To look up a character, the *root radical* of the character is identified and the number of its strokes counted. The number of strokes is then used to locate it in the root radical table, which provides the location for another table specific to that radical. To use this table, the rest of the strokes are counted and the total is used to find the subsection of the table that contains the character and a link (e.g., page number for paper versions) to its location in the dictionary (see [2] for a typical example).

Research has shown that experienced readers of Chinese use the radical system to organize their memory for characters (e.g., [1]). Therefore, the root radical method is good for students as it reinforces this structure in their minds. Computerized systems that can read and translate characters from the screen, or even scan them in from the environment, lack this property. Although ideal for quick translations there is no learning embedded in the process. However, the root radical method depends on the user's ability to identify the root radical, which can be a problem for non Chinese-readers as there is nothing directly identifing the root radical. We hypothesized that this could be overcome by using a computerized system that allows users to start the search with any radical in the character. However, this would only work if the user can identify the radicals in the first place. If the user cannot do this then indexing off all of the radicals would not provide any benefit

There is a vast literature on how Chinese-readers read Chinese, but almost nothing about non Chinese-readers. Therefore, we undertook this study to examine the extent to which non Chinese-readers can identify radicals and, by extension, the feasibility of using the root radical system to create a usable electronic Chinese/English dictionary for non-Chinese readers (e.g., in most cases beginner students of the Chinese language).

The strokes of different radicals frequently touch each other within a character. Since non-Chinese readers have little or no knowledge of the radicals, the only way they could parse them is if the characters contain affordances [3] that interact with the visual system to suggest their identity. For example, this could result from Gestalt properties [4] imbedded in the characters. It would not be too surprising if such an intensively visual script had

[1] Carleton University Cognitive Science Technical Report 2002-06
URL http://www.carleton.ca/iis/TechReports

evolved to take advantage of how the visual system operates.

USABILITY TEST

We began with a quick usability test. We gave two well-motivated, intelligent volunteers a Chinese/English dictionary [2] and explained how to use the root radical indexing system. We then gave them a series of Chinese characters to look up. Both users failed to find any of the characters in the dictionary, even though it was clear that they understood the instructions.

RADICAL RECOGNITION STUDY

Although we used only two subjects, the usability test clearly showed that this version of the root radical method, which can be considered typical, is highly problematic for non Chinese-readers. To get a better idea of the extent to which not being able to identify the radicals is a problem we conducted a study on radical recognition.

Seventeen non Chinese-readers with no background in the Chinese language (male = 10, female = 7) participated in the experiment. Twenty-eight characters representing common words were selected from a Chinese-character dictionary. The characters were composed of two to four radicals each (which is true of the majority of Chinese characters in everyday use). Subjects were told that the characters were made of smaller structures and were asked to guess what these were for each character. Subjects identified the structures by drawing them.

There were 73 possible correct responses (i.e., the characters we used contained a total of 73 radicals). Some subjects identified less than 73 radicals and some more. Therefore, instead of calculating correct responses as a function of possible correct responses, the correct responses were calculated as a function of attempted responses.

The ratio between the number of radicals attempted and the number of correct radicals identified was 70.46% with a standard deviation of 8.63%. Of the 17 subjects, 12 scored in the 71 - 80 % range. This is a large number considering that none of the respondents had any previous knowledge of Chinese writing. With the exception of two subjects, all scored above 60%.

An ANOVA determined that there was a significant effect for order of guessing on the mean percentage of radicals guessed correctly ($p = 0.004$). Paired sample t-tests were used to analyze the differences between subjects' first, second and third guesses. Subjects were significantly better at selecting a correct radical on the first guess than on the second guess ($p < 0.001$), and significantly better on the first guess than the third guess ($p = 0.004$). There was no significant difference between the second and third guess ($p = 0.448$).

On their first attempt, subjects selected a correct radical 78.9% of the time. Interestingly, they also correctly identified the root radical 62.1% of the time, which was much higher than would be expected by chance. To evaluate whether there would be a difference between using the traditional root radical method and the proposed (any radical) method, a paired sample t-test was done to compare the percent chance of guessing the root radical correctly on the first trial to the percent chance of guessing any radical correctly on the first trial. The t-test revealed that guessing any radical was significantly more likely to produce a correct answer ($p < 0.001$). These results suggests that being able to select any radical to search for a character would increase the odds of guessing a correct radical on the first trial by approximately 17% over having to select the root radical.

DISCUSSION

The results of this study show that non Chinese-readers have approximately a 70% hit rate for identifying radicals overall and an 80% chance on their first guess. This level of performance was surprisingly high given the results of our usability study and the fact that our subjects had no knowledge of the Chinese language. These results suggest that Chinese characters contain strong affordances that act on the visual system to highlight the differences between radicals. We believe the affordances are based on Gestalt properties and we are currently carrying out research to identify them. The poor results in the usability test appear to have been due mainly to other factors. We have since identified a host of problems that can prevent a non Chinese-reader from going from correctly identifying the root radical to finding the character in the dictionary. These have to do with specialized knowledge about how to count strokes and how to identify different forms of the same radical. The ability to identify the root radical at an above chance level is most likely due to a non-perceptual affordance. It turns out that the root radical is statistically more likely to be located to the left side of the character and/or towards the top. We speculate that our subjects (English-readers) extended their habit of reading from left to right and top to bottom to selecting their first guess.

The results of this study indicate that the root radical method is a viable approach for non Chinese-readers. We believe that the problems not related to identifying radicals can be dealt with by providing instruction and support in an electronic version of the dictionary. For example, an electronic dictionary could easily show the different variants of a radical. Also, an electronic version would make the whole process much faster. The advantage of using this system is that the process can be considered a form of situated learning that highlights the structure of the characters. This may be more important than we originally thought since, as we have seen, the structure contains useful visual affordances.

REFERENCES

1. Chen, Y. P., Allport, D.A., and Marshall, J.C. (1996). What are the functional orthographic units in Chinese word recognition: The stroke or the stroke pattern? *The Quarterly Journal of Experimental Psychology*, *49A*, 1024-1043.
2. Cowie, A. P., & Evison, A. (Eds.) (1986). *Concise Oxford English-Chinese, Chinese English Dictionary*. Hong Kong: Oxford University Press.
3. Gibson, J. J. (1986). *The ecological approach to visual perception*. Hillsdale, NJ: Erlbaum.
4. Koffka, K. (1935). *Principles of Gestalt psychology*. Harcourt, Brace and Company: New York