

***Chinese Character Dictionary – A new approach
to arranging, explaining and looking up Chinese
characters. Released through Amazon’s self-
publishing service CreateSpace***
Adrian van Amstel

Reviewed by Zhang Xinming

1. Background

Impelled by the dissatisfaction towards today’s most Chinese-English dictionary, committed to solve the confusions encountered in Chinese study and find out more about the structure of Chinese character, the editor decided to compile his own dictionary in 1994 which he believes may provide a new and better approach to arrange, explain and look up Chinese characters. The first version of Chinese Character Dictionary (CCD) was printed in 1998 and now in 2016, the present revised edition is available. Compared with the first version, an approximately 1,500-word supplement is added to the original 6,000 entries. Besides, crucial changes are made to the dictionary to make it more functional and professional.

Any dictionary is an organic whole consisting of vertical macrostructure and horizontal microstructure. Since CCD’s new arrangement and retrieval of Chinese characters are its spotlights, this review would mainly describe and evaluate CCD from macrostructure perspective.

Affiliation

East China Normal University, Shanghai, China.
email: 52160400012@stu.ecnu.edu.cn

2. Microstructure of CCD

2.1 Arrangement of entries

Briefly, the most innovative thing about CCD is its new approach to arrange characters in the dictionary. Different from current commercially available Chinese–English dictionaries whose entries in the main body are arranged according to the official phonetic system ‘Hanyu pinyin’, the arrangement of entries in CCD is organized in terms of the shape of each Chinese character’s phonetic part.

Per the editor, this new arrangement of entries is put forward due to a regret of today’s ordinary dictionary that looking up a character is convenient once the sound of it is known, but it will be useless if only the form of a character is available, and even though a dictionary contains a function searching a character according to its stroke count, it will still be of little help for learners who desire to explore characters with similar shape. Similar Chinese character is known to be one of the major features of Chinese character and the reason for this phenomenon can be traced to its origin. Pictographic is the most basic characteristic of Chinese characters, which leads to the result that certain characters share the same form because of the similarities among objective things. Another noteworthy reason is that the essential composition of characters is stroke and meanwhile basic strokes are nothing but rightward stroke, downward stroke, falling leftwards, falling rightwards, tiny dash and 90 degree turn. These strokes are composed to shape such a large amount of characters that necessarily cause similarity among them.

This dictionary provides a new approach to locate character by its phonetic part and meanwhile this approach groups together characters with similar patterns, which undoubtedly brings certain benefits compared with other dictionaries. First, most foreign learners of Chinese are struggling with memorizing Chinese characters, while grouping together phonetic parts provides them with a better way to remember characters. Mashima ones stripped 214 radicals from 41,000 Chinese characters and got 4,081 phonetic parts, among which 1,689 are in common use (Yang, 2012). He points out we can obtain 70%–80% of Chinese characters if combining these commonly used phonetic parts with 214 radicals. Therefore, remembering the phonetic part which is the main and most prominent part of a character can be considered as one of the most crucial steps to learn Chinese. Second, this arrangement can facilitate foreign learners of Chinese to recall characters with similar patterns, which may provide help in word recognition. Third, learner’s interest towards Chinese study may be aroused through exploring the similarities and differences among characters with similar patterns and meanwhile etymological explanation is available after a series header which may help to better explore characters. Fourth, the discrimination of Chinese characters which are similar

in form is always one essential aspect and even a challenge in Chinese character learning. This arrangement may help learners to form the thought model enabling them to correctly recognize Chinese character from its shape and structure and to further improve their ability to identify easily-mixed compound Chinese characters.

Actually, Chinese researchers never stop working on characters with similar patterns and many reference books or dictionaries aiming at them are compiled to facilitate students and writers. These purely Chinese character dictionaries also classify characters into groups with similar forms, accompanied by phonetic notation, definition, examples and explanations about differences. To say the least, the amount of this kind of dictionary has reached a large number, but the main service subjects for these works are Chinese primary and secondary school students or writers involved in text work. Few dictionaries of this kind aiming at foreign learners of Chinese have been found on the market till now. What's more, even for foreign learners who can understand most Chinese, these dictionaries are of limited usefulness because they focus their attention on discrimination of forms, which means only characters with similar forms are included and could not be used as a dictionary for looking up characters. In this regard, CCD can be considered as an unprecedented innovative practice since it combines a dictionary of characters with similar patterns with a Chinese-English dictionary together. CCD successfully makes up the insufficiency in previous dictionary editing history. A rough research is made to count how many versions of Chinese-English dictionaries are available on JD which is one of the largest online stores in China, and the result shows there are seven versions and all of them are compiled by Chinese specialists. Chinese people may not be able to gain a firsthand experience for the difficulties that foreigners go through during the learning of Chinese. As a consequence, they may neglect the fact that foreign learners, just like underage Chinese children, also need a dictionary of characters with similar form as their reference.

Undeniably, CCD's creative arrangement of characters has its benefits, but its weakness cannot be neglected as well. For a dictionary whose entries are arranged by the Pinyin system, characters can be looked up by either the ABC Pinyin order or stroke count method. Various searching methods are all for user's convenience and make access to information more effective and faster. However, for CCD, since characters are placed per shape of component, the Pinyin system does not work for this arrangement, which indicates this dictionary is of limited usefulness if merely the sound of a character is known.

Another aspect of entry arrangement is the amount of characters in CCD. Although CCD comes nowhere close to dictionary tomes, it already can meet

people's needs. According to the International Program of Chinese characters released by Chinese government in 1981, 6,763 characters are identified as Chinese characters vocabulary. Containing 7,450 characters, CCD can satisfy the quantitative criteria.

2.2 Index structure

Another distinct characteristic of CCD is its unique approach to look up a character. Nowadays, retrieval approaches in existing dictionaries are based on either the Pinyin System or stroke count. The alphabetical order method is one of the most effective ways to construe the macrostructure of a dictionary and is the most ideal information arranged system adopted by basically all kinds of dictionaries (Yong, 2003: 15). Also for a dictionary of characters with a similar form mentioned above, most of them adopt both the Pinyin System and stroke count method, since the number of characters included in those dictionaries is simply 2,000 words or so. This amount of entries indicates all characters can be listed in an appendix by the ABC order or stroke count and are followed with page number, which will not occupy too many pages in the dictionary.

However, CCD aims at compiling a dictionary containing all necessary Chinese characters, which inevitably requires a new system of index structures to locate characters. Referring to Ann's classification of Chinese characters, the editor summarizes 7,450 traditional and 1,450 simplified characters and then clusters them into 395 character tables and further divides components in each character table into different series. Each series is led by a series header (or component/phonetic part) which is the part that all characters in one series have in common. Correspondingly, the editor creates a new system of index parameters to retrieve characters through deciding the right component (non-radical) and locating the series header.

Not employing a long list of strokes, the editor of CCD only summarizes one table called the Main Component Table as the main retrieval clue. The user first strips radicals of characters and chooses the component first appearing in the MCT, and then locates the series header in the header list of that component. The lexicographer always uses a semantic radical to locate a character, but in CCD, it is the first time (as far as I know) in dictionary editing history to utilize the phonetic part as the index parameter. Besides, in order to let users detect the right component conveniently, the editor introduces carefully how to identify the radical part in part 1. CCD's approach must have its own imperfection, yet it indeed groups together characters with similar forms as far as possible and provides a way to locate them. Besides, it saves users time on counting strokes.

It must be admitted that the CCD proposes a completely new and different approach to compile a dictionary, but the new thing is inevitably controversial. For me, the validity of this approach remains open to question.

The first doubt of CCD is about the split of phonetic parts, which is one of the most important steps in looking up a character in CCD. For most foreign learners of Chinese, the radical indexing method has a certain difficulty, because learners must acquire a certain amount of characters, memorize Chinese radicals, and meanwhile detect the right radical of a character. That's the reason why the Pinyin system is more popular. However, in CCD, difficulty is increased because the part needing to be split is the remaining part other than the radical one. It is not easy to split the phonetic part because large amounts of components must be acquired in advance to judge which part is a component and which part is simply a stroke, which requires a high level of Chinese. Even though the phonetic part is successfully split into several components, the component which is the first in order of MCT still needs to be picked out. It seems to me this complicated searching approach might be a challenge for Chinese people, let alone foreign learners.

Then let us talk about the MCT. It can be seen that 395 components are classified into 17 categories with a representative and page number. The component is the part occurring with highest frequency in the structure of Chinese characters. Depending whether it can be used as radical, components can be divided into two categories. In Wang's Book *An Introduction about Chinese Components and Radicals*, he counts 816 components among which 254 can be used as radical and others cannot. Similarly, in CCD, the editor also makes a distinction between two kinds of components, and only components which cannot be used as radicals are included in MCT. However, if checking with Wang's book, it is an obvious component in CCD that is fairly different from the traditional Chinese component.

First there is a difference in the amount of components, but after checking with CCD, it may be found that the editor takes some components as a series header. The reason for this arrangement may be because according to the editor, 400 components are sufficient to identify all phonetics, but it inevitably results in certain problems such as time and energy wasting. For example, 考 is regarded as a component in Wang's book, but it is a series header in CCD and is grouped together with 孝、老、者, etc., in the same CT under the component category 耂. However, 耂 is not a component in most Chinese dictionaries, and it seems strange for most Chinese to consider them as a whole. As a result, when searching characters containing 孝, 老 or 者 components, 耂 is neglected and 土 or 子 or 匕 or 日 are searched.

Second, there are nearly one thousand components in Chinese, but how are these 394 components chosen to be included in MCT and further classified

into 17 groups? There is not an explicit standard of classification for users to follow and it may result in the trouble of locating the component. Sometimes, it seems a little over-classified for the MCT and it is hard to connect the component of target characters with components in MCT. For example, it is difficult for me to understand why 氏、瓦 can be grouped together under 𠃉. Also for 衷, after trying 中 and taking a long time to check components in MCT one by one, I was uncertain whether to try the component 衣 and found it was the right one, but it is still confusing for me to accept the fact that 衷 is included in the component 衣. Compared with 衣, why not choose 衤 as a component in MCT? These subjective classifications make the location of character harder.

Third, different from stroke counts which are listed by the order from small to large and can provide a quick way to locate the accurate place of the component, the order of 17 categories in MCT seems somewhat difficult to follow and memorize for people who are not quite familiar with the structure of character. The editor attempts to arrange similar components that are adjacent to their characteristic features, but the order is still subjective and users have to find the components one by one. All components in MCT are classified first into two groups. The first group is those with slanting strokes and the second group is components consisting of horizontal and vertical strokes. This classification certainly gives users a general direction to locate the component, but the same issue is its uncertainty. For example, why 月 周 丹 are grouped under the second group? With slanting strokes and strokes bending to the left, it is more reasonable for me to classify them under the first group. Besides, a reasonable order of category may help users to find the target component quickly and correctly, but the order of 17 groups in MCT might cause a little trouble when locating the component. For instance, 儿 and 凵 are far away from each other and 弋 is arranged near 冫 rather than 凵 or 人. The same issue exists for the radical table. Users may take some time to search the needed radical, so I am just considering whether it would be more practicable to list them in some sorted order.

Another remarkable feature of CCD is that it avoids counting the number of strokes. The editor sets a series of rules to decide which component should be used to locate the character, but this procedure occasionally takes more time and energy compared with the stroke count. The question is incorrect where a prior component might be chosen, whereas the stroke count is seldom wrong. Besides, for most traditional Chinese characters, counting the number of strokes may be boring and time-wasting, but for most simplified characters, structure of character is greatly simplified and stroke count just needs several seconds.

Other imperfections of CCD's searching approach are listed below. First, for characters with an easy structure, this index approach works perfectly because

users needn't waste time in counting the number of strokes, especially for Chinese characters with many strokes. However, for characters with complicated form and more components, this method might take a longer time, since once users strip incorrect radicals or judge wrong prior components, all steps have to be repeated. Moreover, in previous dictionaries, users may utilize the radical part to locate characters; however the first step of CCD's approach is to strip it. The radical part, the most identifiable element in a word, is stripped, which means once users could not pick out the right component, there are no other ways to find the needed character. Besides, in order to master this procedure, users must be quite familiar with rules in each step and all 394 components in MCT to locate characters quickly and correctly. Once they do not use CCD for a short period of time, they might not be able to use it efficiently. In addition, the layout of components in MCT and series headers in CT is liable to cause illusions at certain moments when picking the target components. This is not only because they all look pretty similar with components nearby but because components or series headers of each category are printed in horizontal lines. This format, whereas saving spaces, is inconvenient to users.

2.3 Traditional and simplified characters

Another feature of CCD is that it deals with both traditional and simplified Chinese characters. It greatly facilitates users of both characters, but also brings some problems for traditional and simplified components. It leads to repetitive explanation. For example, 宗 appears in the category 8 and 宗 appears in the category 384, which causes all characters with component 宗 to appear twice.

3. Microstructure of CCD

Microstructure of a dictionary is mainly about its entry. Consisting of two parts, entry character and its translation, entry is the most basic structure and functional unit of a dictionary (Huang, 2001). In CCD, each series is composed of two parts: series header and entry characters. Series header is explained from etymology. Each entry character is followed with phonetic notation, translation with one illustrative example and occasionally etymological explanations.

The etymology note is added to each series header or some characters which may help to better understand the structure of character and distinguish between similar headers. Many Chinese scholars have already noticed the importance of distinguishing between phonetic parts of characters. He (2005) states great attention should be paid to distinguish characters by component formation in teaching Chinese as a foreign language. He (another scholar) published a book in 1998 and emphasized in the appendix that it is

necessary to observe different parts of two similar characters and analyze the etymology of different part to deepen learns understanding about the structure. However, occasionally etymological explanations are difficult for users to imagine and get the connection between explanation and component.

Illustrative example is the extent of translation. Sometimes, new meaning would be expounded when entry characters combined with other characters. Therefore, richness of illustrative example is proportional to the size and quality of a dictionary (Yang, 2012). However, in CCD, the amount of illustrative examples is not large. This may be due to the size of CCD and its main focus. A dictionary of this size and scope cannot be expected to be complete. However, users may feel disappointed to find that some useful everyday expressions such as 差不多, 近乎 and 场所 are ignored because of limited examples. Besides, syntactic category is not annotated for most characters. Admittedly, the functive class of Chinese is a fairly tricky issue in English translation, but it can serve users better to define a character. What's more, the addition of synonym and antonym may help users to broaden their Chinese vocabulary.

4. Conclusion

On the whole, this work appears to be a careful compilation. Containing two writing systems, introducing etymological explanations, showing how to recognize radicals and creating a unique, almost perfect, but a little complicated way to split elements in a character, all these creations indicate the editor to be a scholar with extensive knowledge about Chinese character culture and history. However, deficiencies of CCD greatly minimize the usefulness of this dictionary. While the editor does not state clearly which level of Chinese is CCD intended, for me, CCD is quite useful to advanced foreign students who has an interest in the Chinese language and wish to know more about the structure of Chinese character. It might not work well for beginners, however, because it requires a high level of users' Chinese ability. Moreover, as the editor mentions in the preface, it is not easy to get people changing from one system to a new one unless there are substantial advantages. Lastly, we have to accept the fact that with the technology development, apps in mobile phones greatly take users' hands off a paper dictionary, such as the apps' word recognition function with the help of a video camera on the back of the phone. It is not necessary to locate characters with such complicated procedures, and just pressing a button will solve everything. All these are challenges confronted by CCD.

References

- 黄建华、陈楚祥 《双语辞典学导论》，商务印书馆，2001。
冉红 《巧记形近字》，北京邮电大学出版社，1998。

王述加 《偏旁部首简说》，湖南人民出版社，1985.

杨慧玲 《19世纪汉英词典传统》，商务印书馆，2012.

雍和明 《交际词典学》，上海外语教育出版社，2003

He, Hongfeng (2005). On 'Distinguishing characters by component formation of strokes' in teaching Chinese as a foreign language. *Chinese Language Learning* 2005 (03)

