Hongloumeng, Honglou Meng, Hong Loumeng, or Hong Lou Meng

Yuanqiong Wu
School of Foreign Languages, Shanghai University of Engineering Science, China

Abstract—The current Pinyin Romanization of Chinese book and journal titles is rich in examples of inconsistencies, and this problem has much more been identified than examined. The current paper traces the problem back to the guiding documents, analyzes their inborn problem. It is argued that the currently dominant practice of aggregating syllables is the source of the inconsistencies, and it results from ambiguous wordings and misconception of “ci” as the basic unit in the guiding documents. Based on this analysis, a practice of Romanizing Chinese on the basis of “zi” is put forward, and the underlying rationale analyzed. The purpose is to contribute to the solution of the issue of inconsistency and offer an approach to standardizing the practice of Pinyin Romanization of Chinese book and journal titles.

Index Terms—Pinyin Romanization, Chinese Book and Journal Titles, Inconsistencies, Aggregation of Syllables, Guiding Documents, Zi, Ci

I. INTRODUCTION

Chinese book and journal titles have a diverse outlook in the English world, as a result of adopting different methods such as translation, and transliterating according to Wade-Giles system or Pinyin system. Even under the same system, Chinese book and journal titles are rich in examples of diversity, and subsequently inconsistency. The inconsistency caused by Pinyin romanization of Chinese materials has mostly been felt in library cataloging and retrieving field (Arsenault, 2001; Diao, 2015; Huang, 2004; Li, 2008). However, much effort has been spent identifying, instead of analyzing and solving the problem itself.

The current paper focuses on the different presentations of the classical Chinese novel title “红楼梦” under Pinyin system: “Hongloumeng”, “Honglou Meng”, “Hong Loumeng”, and “Hong Lou Meng”. The following two tables shed a light on the diversity in both real life and the virtual world.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>A SURVEY ABOUT PINYIN ROMANIZATION OF “红楼梦” IN A GROUP OF 10 NATIVE CHINESE TEACHERS1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pinyin Romanization of “红楼梦”</td>
<td>Results</td>
</tr>
<tr>
<td>Hongloumeng</td>
<td>1</td>
</tr>
<tr>
<td>Honglou Meng</td>
<td>1</td>
</tr>
<tr>
<td>Hong Loumeng</td>
<td>1</td>
</tr>
<tr>
<td>Hong Lou Meng</td>
<td>5</td>
</tr>
<tr>
<td>Others</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>WORLDCAT RETRIEVAL OF THE TITLE “红楼梦” IN PINYIN SYSTEM2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Romanization of “红楼梦”</td>
<td>Retrieval Results</td>
</tr>
<tr>
<td>Hongloumeng</td>
<td>72</td>
</tr>
<tr>
<td>Honglou Meng</td>
<td>64</td>
</tr>
<tr>
<td>Hong Loumeng</td>
<td>5</td>
</tr>
<tr>
<td>Hong Lou Meng</td>
<td>1472</td>
</tr>
</tbody>
</table>

The inconsistencies illustrated here represent the general state of Chinese book and Journal titles in Pinyin Romanization. For instance, another classic Chinese novel “西游记” has a similar existence in English with “红楼梦”. According to information from WorldCat accessed on 11 Oct., 2017, there are 1,518,892 representations of “西游记” in the form of “Xi You Ji”, 201 “Xiyouji”, 452 “Xiyou Ji”, 442 “Xi Youji”. This paper will focus on the Pinyin romanization of “红楼梦” to analyze the inconsistency problem.

1 Grant Information: This paper has been supported by grants from Shanghai University of Engineering Science (Grant Number: K2018180002; K2018180003).
2 On 23 Sep., 2017, a survey was conducted among a group of 10 native Chinese teachers who were instructed to present their Pinyin Romanization of the title of “红楼梦”. Among the results, two were the translation of “红楼梦”, which did not fall into the Pinyin system, so they were indicated in the column of “others”.

WordCat was accessed on 11 Oct., 2017. The retrieval was limited in the fields of “title” and “author Xueqin Cao ”, and results were restricted within formats of print books and e-books.
Superficially, the differences are about semantic segmentation: in “Hongloumeng”, the three Chinese syllables for “红楼梦” were rendered into one semantic unit; in “Honglou Meng”, “Honglou” were rendered into one unit, and “Meng” another one, thus “Honglou Meng” consisting of two semantic units; this is also the case in “Hong Loumeng”; and in “Hong Lou Meng”, the three characters were rendered into three units. Further examination of these renderings reveals an appalling realization that they are actually about an essential issue: the basic unit of Chinese language. Given the nature of this issue, we argue that the inconsistency phenomenon in current Pinyin romanization of Chinese book and journal titles will be analyzed and checked, and a feasible standard be established. To fulfill this end, three research questions will be asked:

1. What is the source of these variations: “Hongloumeng”, “Honglou Meng”, “Hong Loumeng” and “Hong Lou Meng”?
2. What are the problems with the source?
3. What should be the proper way of romanizing Chinese book and journal titles? And why?

To provide an answer to these questions, the current paper begins by reviewing the historical backgrounds of Pinyin romanization of Chinese. Specific attention will be given to documents guiding the practice of romanizing Chinese book titles. Then the problems of these documents, which are also the sources of inconsistencies, will be analyzed. Lastly, a proposal based on both historical and linguistic facts, and a conclusion of the current paper will be followed.

II. HISTORICAL BACKGROUNDS AND GUIDING DOCUMENTS OF PINYIN ROMANIZING CHINESE

Pinyin system is one of the two major systems of romanizing Chinese in currency today, another being Wade-Giles system. Pinyin romanization of Chinese, as well as Wade-Giles, is but one spot in the long continuum of engaging Chinese language to the international world, the source of which could be dated back to early 17th century. Different from Wade-Giles system which was initiated with the desires of foreigners inside China to approach Chinese and materials carried by the language, Pinyin system is a result of efforts initiated at the turning point of 19th century and 20th century by Chinese people to simplify Chinese language and to promote education on a massive scale. After decades of exploration, in September of 1957, out of the 6 proposals, Scheme for a Chinese Phonetic Alphabet (《汉语拼音方案》) (it will be referred to as “SCPA” in this paper), one adopting Roman alphabets to indicate the pronunciation of mandarin Chinese, was passed by the State Council of PRC, and approved by the National People’s Congress in February 1958. Twenty years later, in 1978, the National People’s Congress of China resolved to use SCPA as standard romanization of Chinese personal and geographical names both at home and abroad. SCPA prescribes the 26 Latin alphabets, the initials and endings adopted to indicate mandarin Chinese. Following this resolution, more documents and laws have been issued to consolidate the role of Pinyin system. In 1984, Basic Rules for Hanyu Pinyin Orthography (《汉语拼音正词法基本规则》) was issued, and it became a national standard in 1996, and was revised in 2012 (GB/T 16159 2012). Most recently, Article 18 of Law of the People’s Republic of China on the Standard Spoken and Written Chinese Language (《中华人民共和国通用语言文字法》) (which was established in 2000, and has come into force since January 1st, 2001) restates that “SCPA is to be used as the tool of transliteration and phonetic notation for the standard spoken and written Chinese language” and it is the “unified norm of transliterating names of Chinese people, places, and documents”.

Corresponding to the national popularity of Pinyin system, the international world responded actively to transfer from Wade-Giles to Pinyin, and some guidelines for romanizing Chinese were issued. As early as 1958, the British Library started to use Pinyin for its bibliographical control of its Chinese collection, and UN approved to adopt SCPA for its romanization of Chinese geographical names in 1977. Five years later in 1982, ISO decided that Pinyin system be the international standard in spelling the Chinese language, and so far several specific documents have been issued to regulate the romanization of Chinese (ISO 24615-2; ISO 7098:2015).

The above mentioned documents demonstrate some consistent practice. First of all, aggregation of characters is a typical practice. For instance, SCPA prescribes the basic elements of romanizing Chinese: the letters, tones and apostrophe use, while the notes include such examples as “儿童” to be romanized into “ertong” and “花儿” into “huar”. This practice became dominant in later-on documents including GB/T 16159 2012. Secondly, GB/T 16159 2012 prescribes that “ci” (词) be the basic unit in romanizing Chinese. As “ci” can be monosyllabic, disyllabic, or polysyllabic, this principle actually underlies the practice of aggregation characters. The rationale is twofold. Firstly, according to Feng (2016, p. 6-7), there are only 405 syllables available in Chinese. Therefore, against the 8105 characters included in the List of Standard Chinese Characters for General Use (《通用规范汉字表》), one syllable is to indicate over 20 characters, resulting in ambiguity on a large scale. Then, to reduce ambiguity, “ci” which may contain more than one syllable is adopted as the basic unit of romanization of Chinese. Hence, the aggregation of syllables becomes the norm giving birth to ambiguity.

Once issued, the guiding documents have been received differently, with two schools of thought: the appreciative school and unappreciative schools. The appreciative and accepting school exclaims that Pinyin system adopts Latin alphabets to romanize Chinese, so it is a monumental system to bridge the gap between Chinese language and the international language family, enabling Chinese language to participate in the international dialogue (Jie, 1990; Li, 2008; Tang, 2008; Wang, 2013). At the same time, the appreciative scholars have also pointed out the problems of Pinyin
system, among which over-generalization (Shang, 2008; Cai, 2013) and lack of consistency are most frequently referred to (Gao, 2003; Liu, et al, 2014). On the other hand, the unappreciative school holds that Pinyin system is a system totally turning away from Chinese language tradition, thus not suitable to Chinese language, so it should be eradicated (Duan, 1989; Li, 2000). Despite the problems mentioned above, Pinyin system has not been eradicated. It has been consistently revised and remains to be the dominant system of romanizing Chinese both at home and abroad.

III. AN ANALYSIS OF GB 3259-92----THE DOCUMENT GUIDING ROMANIZING BOOK AND JOURNAL TITLES IN CHINESE

The romanization of book and journal titles is governed by SCPA and other documents mentioned in the above section. However, GB 3259-92 is a document issued for the specific purpose of romanizing Chinese book and journal titles. So, in this section, specific articles in GB 3259-92 will be examined, and the problems with this document analyzed.

The current GB 3259-92 was first issued in 1982, with the title of Transliterating Rules for Basic Chinese Alphabet on Titles for Books and Periodicals in Chinese (《中文书刊名称汉语拼音拼写法》). It was established as a national standard in 1992, thus GB 3259-92. In the following, focus will be given to a close examination of this document, especially to the inborn problems, for these problems have much to do with the confusion and inconsistencies in the current romanization of book and journal titles, which were mentioned at the beginning of this very paper.

A. Ambiguous Wordings in GB 3259-92

GB 3259-92 has five sections: Content and its Application, Terminology, romanizing Principle, Reference, and romanizing Rules. The romanizing Principle is of particular importance, because it governs the following romanizing Rules. In Section 3 of the document, it reads:

3. 拼写原则
“以词为拼写单位，并适当考虑语音、语义等因素，同时考虑词形长短适度。”

This principle is translated into English as:
The romanizing practice is conducted with “ci” being the basic unit, and at the same time phonetic and semantic aspects are to be considered appropriately, together with the proper length of “ci”.

The wording of the principle invites problems, for such ambiguous concepts as “适当考虑” (be considered appropriately) and “长短适度” (proper length) are of subjective perception and open to variation. Similar wordings also appear in the specific rules in Section 5:

5.1 中文书刊名称拼写基本上以词为书写单位，每个词第一个字母要大写，因设计需要，也可以全用大写。
5.2 结合紧密的双音节和三音节的结构（不论词或词组）连写。

The corresponding translation is:
5.1 Transliterating Chinese book and journal titles is basically conducted with “ci” being the basic unit, with the first letter capitalized. For specific purposes, each letter can also be capitalized.
5.2 The closely bound disyllabic or trisyllabic structures (whether they are word or phrases) are to be transliterated aggregates, without space in between.

According to online Oxford English Dictionary, a rule is “one of a set of explicit or understood regulations or principles governing conduct or procedure within a particular area of activity”. Such ambiguous wordings as “也可以” (can also be), “结合紧密” (closely bound), “适当考虑” (be considered appropriately) and “长短适度” (proper length) fail the requirement of being “explicit” and “understood”. When the concepts in guiding documents are ambiguous, they are left in the hands of readers and the prescribing and guiding function of rules falls flat. “红楼梦” is offered in GB 3259-92 as an example of “closely bound disyllabic or trisyllabic structures” in rule 5.2, and it was romanized into “Hongloumeng”. Further, it is also justifiable to ask why “长城恋” is not a “closely bound” trisyllabic structure and then is romanized into “Changchenglian”, instead of “Changcheng Lian”, as was offered in GB 3259-92 as an example for rule 5.1?

B. The Problematic Key Concept of “Ci”

Another source of the problem lies in the key concept of “词” (ci), which is held to be the basic unit of romanizing Chinese named entities, including book and journal titles. But what is “ci”? How to judge whether a chunk of Chinese characters falls into the category of “ci” or not? GB 3259-92 did illustrate different types of “ci” through 16 specific

rules, but unfortunately, it did not offer a response to these two very essential questions, and the concept of “ci” remained to be problematic and elusive. If the concept of “ci” remains to be ambiguous, romanization of Chinese with “ci” as the basic unit is to meet with individualistic idiosyncrasies, and henceforth inconsistencies in the current practice.

IV. A PROPOSAL FOR ROMANIZING CHINESE BOOK AND JOURNAL TITLES

After exploring the source of inconsistencies in current practice, a key issue looms large, and that is “what should be the proper way to romanize Chinese book and journal titles”. To check the current chaotic situation, our proposal is that the romanization of Chinese book and journal titles should be conducted on the basis of “zi”----characters, instead of “ci”. The following points would be adopted to back up our argument:

A. “Zi” Has Been in Chinese Language System since the Ancient Times, While “Ci” Is a Rather Recent Phenomenon

Li Wang (1958) divided the history of Chinese language into four periods (Table 3). From Ancient Chinese to Early Modern Chinese periods, “zi” (zi) instead of “词” (ci) played a central role in

<table>
<thead>
<tr>
<th>Historical periods</th>
<th>Time line</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ancient Chinese</td>
<td>Before 3rd Century</td>
</tr>
<tr>
<td>Middle Chinese</td>
<td>4th Century----12th Century</td>
</tr>
<tr>
<td>Early Modern Chinese</td>
<td>13th Century----19th Century</td>
</tr>
<tr>
<td>Modern Chinese</td>
<td>May 4th Movement----now</td>
</tr>
</tbody>
</table>

Table 3

Chinese language system. Firstly, the monosyllabic “zi” was the basic unit in such classical texts as Sun Zi Jing (A Three-character Canon 三字经), and Qian Zi Wen (A Thousand-character Classic 千字文). Secondly, dictionaries were compiled to explain the meaning of “zi” (Table 4), thus “zi” being an indispensable part of the names of dictionaries. In modern times, “Xin Hua Zi Dian”, a dictionary containing explanation of 13000 “zi”, is a must for school students to learn Chinese. Lastly, “zi” was the starting point of the traditional Chinese language study which included phonology, morphology and

exegesis: Phonology was about the auditory sound of “zi”; morphology about its written form; and exegesis about its semantic aspect.

Compared with “zi”, “ci” (word) is a rather recent phenomenon. In Ma Shi Wen Tong (马氏文通) (1898), a monumental book in Chinese linguistic study, “ci” started to be frequently addressed. However, the entity of “ci dian” (词典) (Dictionary of Ci) did not appear in dictionary names until 1945 when San Zi Jing (A Three-character Canon 三字经), and Qian Zi Wen (A Thousand-character Classic 千字文). Secondly, dictionaries were compiled to explain the meaning of “zi” (Table 4), thus “zi” being an indispensable part of the names of dictionaries. In modern times, “Xin Hua Zi Dian”, a dictionary containing explanation of 13000 “zi”, is a must for school students to learn Chinese. Lastly, “zi” was the starting point of the traditional Chinese language study which included phonology, morphology and

exegesis: Phonology was about the auditory sound of “zi”; morphology about its written form; and exegesis about its semantic aspect.

Compared with “zi”, “ci” (word) is a rather recent phenomenon. In Ma Shi Wen Tong (马氏文通) (1898), a monumental book in Chinese linguistic study, “ci” started to be frequently addressed. However, the entity of “ci dian” (词典) (Dictionary of Ci) did not appear in dictionary names until 1945 when Guo Yu Ci Dian (《国语词典》, A Dictionary of Ci in Chinese Language) was issued. Bi-syllabic or polysyllabic “ci” first outnumbered monosyllabic “zi” in Ci Yuan (《辞源》) published in 1950: 50000 “ci” vs. 13000 “zi”. But once “ci” appeared, it took the dominance. For instance, in Zhong Guo Wen Fa Yao Lue (《中国文法要略》) (1942/2002: 8), Lv (1904-1998) divided the basic unit of Chinese language into “dan ci” (single monosyllabic word) and “fu ci” (复词 bilingual or polysyllabic word). “Dan ci” was what had been traditionally known as “zi”. In other words, the traditional “zi” was assimilated into the concept of “ci”, thus losing its independent status. This dominating role of “ci” reflected in many linguistic works such as History of Chinese Grammar (1983) and History of Chinese Ci Hui (1984), where the contents were based on the central concept of “ci”.

From the above sketch of Chinese language history and, it can be observed that “zi” has been in Chinese language since the ancient times; the modern concept of “ci” is an extension of “zi”, not vice versa. To romanize Chinese on the basis of “zi” is to make the history of Chinese language consistent, while to romanize Chinese on the basis of “ci” is to cut that historical consistency short.

B. “Ci” ≠ “Word”

GB/T 16159 2012 regulates that “ci” is equivalent to “word”, and it is the smallest independent unit of Chinese. This
theory is problematic in two aspects. Firstly, the concept of “ci” being the smallest independent unit of Chinese language invites counter arguments. The smallest linguistic unit, according to Lü (1981, p. 211), means nothing can be inserted in-between. GB 3259-92 regulates that “红楼梦” is a “ci”, so it is romanized as “Hongloumeng”, which means no linguistic units could be inserted between these three characters. While in fact, characters do appear in between “红楼梦”, as in examples like “红楼梦复梦”, “红楼梦正梦”, “红楼梦续梦”. With the smallest unit falling flat, the concept of “being independent” cannot hold itself. Secondly, the correspondence of “ci” with “word” is problematic, and linguists even questioned the very existence of “ci” in Chinese language. Y.R. Chao (1968, p. 136) argued that “Not every language has a kind of unit which behaves in most (not to speak all) respects as does the unit called ‘word’”. Lü stated in “Yu Wen Chang Tan” that “word is ready-made in English...while what is ready-made in Chinese is zi...The reason why there is no satisfactory definition for ‘ci’ in Chinese is that there is no such a ready-made concept of ‘ci’ in Chinese” (1981, p. 45). As to the Chinese correspondence of “word”, Pan (2001) put forward that if a correspondence must be established, the Chinese “zi”, instead of “ci”, is corresponding to the English term “word” for the following reasons:

1. “Zi” is the natural basic unit in Chinese, and it corresponds with “word”, the unit par excellence in English. Using this natural unit as the starting point of language study is always better than using an ambiguous and controversial unit which is a result of personal idiosyncratic analysis.
2. “Zi” is the basic unit for Chinese people to learn about the world, which corresponds to the Bible sentence “In the beginning there was the word”.
3. “Zi” is the interface of phonetic, orthographic, and semantic study of Chinese language, as is the English “word”.
4. “Zi” is the nexus of grammatical study of Chinese language, for it links lexicology on the one hand, and syntax on the other. In English, “word” is such an nexus.

C. Context Helps to Disambiguate

In terms of ambiguity caused by homophony—characters sharing the same pronunciation, it is important to remember that ambiguity is “a feature of all languages” (Wang, 2011, p. 5) and the context in which linguistic chunks occur helps to disambiguate the meaning. The current aggregation of syllables in Pinyin romanization of Chinese book titles was based on the effort to reduce ambiguity—by linking syllables together to indicate that they are an independent unit. However hard they tried to illustrate examples of “ci” of different kinds, such documents as GB/T 16159 2012 and GB 3259-92 failed in explicitly defining the boundary between “ci”, and “ci” continues to remain as a vague concept. Without a clear definition that could clarify the specific purpose of “ci”, romanization of Chinese on the basis of “ci” is just like a castle built on sand: it is fluid and unreliable. According to the guiding documents, “Hongloumeng” is a “closely bound” structure, for it is the title of the classic Chinese novel “红楼梦”, but with the absence of explicit definition of “ci” and inborn ambiguous wordings mentioned above, it is also legitimate for such book titles as “名医谈《红楼梦》与现代心理学” to be romanized as “Mingyitanhongloumengyuxiandaixinlixue”, “Mingyi Tan Hongloumeng Yu Xiandaixinlixue”, or “Mingyi Tan Hongloumeng Yu Xiandai Xinlixue” and get away with the chaotic practice. The prescription in guiding documents is the very source of ambiguity and even chaos. If the title is romanized as “Ming Yi Tan Hong Lou Meng Yu Xian Dai Xin Li Xue”, the designation of each syllable can be determined by the linguistic context in which the syllables are situated. Also, the context beyond linguistic domain (in this case, a psychology learning classroom ) will help to solve the issue of ambiguity, by gradually narrowing down the designation of each syllable. Finally, readers carry with them a certain relevant knowledge, and it cannot be assumed otherwise. If the romanization of Chinese is based on “zi”, the flexibility of word segmentation in different contexts is allowed and consistency and accuracy of “ci” identification is guaranteed.

D. Lessons Can Be Learned from the Other

Lessons can always be learned from the “other”. As was mentioned above, ambiguity is a feature general to all languages, and every language has abundant examples of ambiguous linguistic units. The classical English example “white house” can provide a perfect example of how ambiguity operates in language. In English, “the white house” can refer to any specific house that is painted white, or the only place where the president of the United States resides, or the Russian parliament building. With this ambiguity, English language does not come up with the practice of aggregation of syllables. Instead, English adopts “the White House”, with the first letter in each word capitalized, to indicate the presidential residence or the Russian parliament building, leaving the context to work out its specific designation.

Lessons of similar kind can also be learned from Wade-Giles system, where the practice is to separate syllables, either by space or a hyphen. For instance, “红楼梦” has been romanized as “Hung Lou Meng” or “Hung-lou-meng” in Wade-Giles system. When the world is thinking about transferring from Wade-Giles to Pinyin system, the wisdom in Wade-Giles should not be deserted.

V. Conclusion

The status of “ci” occupies an ambiguous status in Chinese language: it lacks a workable definition, and the boundary between “ci” is elusive. romanizing Chinese on the basis of “ci”, as was prescribed in such guiding documents as GB/T
16159 2012 and GB 3259-92 is in fact “a matter of fiat, not a question of fact” (Chao, 1968, p. 136), and is the very source of the current chaos in romanizing Chinese book and journal titles. This chaotic situation cannot go examined or unchecked. In this paper, we have argued that “zi” should be the basis of romanizing Chinese book titles, for this practice firstly eradicates the chaotic situation in the current practice, and secondly, it is a lesson learned from history and current practice. So, we finally advocate that the Pinyin romanization of “Hongloumeng” be “Hong Lou Meng”, not “Hong loumeng”, “Hong lou Meng”, nor “Hong Loumeng”.

ACKNOWLEDGEMENTS

I should first thank Professor Wenguo Pan from East China Normal University, China, for his encouragement of conducting the current study and sharing with me many important resources in this area. My sincerely felt thanks also extend to Dr. Rhada Iyer at Faculty of Education, Queensland University of Technology, Australia. The meeting with her during my stay at QUT has helped me forge the line of thinking in conducting the current study. And her careful reading of the drafts has contributed a lot to the final image the current paper takes.

REFERENCES

Yuanqiong Wu, an associate professor at Shanghai University of Engineering Science. Her research fields include translation studies, cross-cultural communication, and TESL.