

The Contemporary Chinese Dictionary's New Entries: A Review of the CCD7's Inclusion Principles

Weiping Ren¹ and Dechun Qin²

¹School of Foreign Studies

Beijing Information Science & Technology University, Beijing, China

No. 12, Xiaoying East Road, Haidian District, Beijing, China

renweiping@bistu.edu.cn

²School of Foreign Studies

Beijing Information Science & Technology University, Beijing, China

No. 12, Xiaoying East Road, Haidian District, Beijing, China

derekqin@126.com

Selected paper from CLSW2019

ABSTRACT. *Since its first edition in 1978, the Contemporary Chinese Dictionary (CCD) published by the Commercial Press has been an authoritative source of reference for the use of standard modern Chinese. As of 2018, the CCD has undergone 7 updates, with the latest edition featuring more than 400 new entries. This paper conducts an empirical study into a number of theoretical considerations for selecting new CCD7 entries, namely the common use principle, the standardization principle, the systemization principle, and the consultation frequency principle. Besides assessing the merits of the four fundamental principles, the author also examines some problematic aspects of the CCD7's new entry inclusion process and puts forward three suggestions for the dictionary's future revision.*

Keywords: Syntactic Analysis, Software Natural Language Processing, Automatic Generation

1. Introduction. As a phenomenon unique to human society, language consists of three elements: pronunciation, vocabulary and grammar, among which the transformation of vocabulary is the most noticeable. The meaning of a character (phrase) is closely related to how a society changes. In line with rapidly developing political and economic situations, together with changing cultural, scientific and technological landscapes, new things and

phenomena keep coming and going. All this necessitates adding corresponding phrases in Chinese spoken and/or written discourse.

As China's first authoritative Chinese-Chinese dictionary on standard modern Chinese, the Contemporary Chinese Dictionary (hereinafter referred to as CCD), has been keeping pace with the times since its first print edition in 1960 [1]. Up until now, the continuous revision of the CCD has witnessed a comprehensive and accurate reflection of the characteristics of modern Chinese vocabulary. The compilation of each new edition of the CCD, which has been based on its previous version, includes additions, deletions and changes in the meanings of the phrases, which are closely related to the development and changes of Chinese society. The newly included phrases and meanings can reflect social phenomena and highlight the characteristics of the times.

By 2016, great changes had taken place not only in modern Chinese, but also in the fields of linguistic studies, lexicographical studies, and the regulations stipulated by government departments and other public bodies administering the use of the modern Chinese language. Therefore, a more scientific, modern, and improved version of the CCD7 was published. With academic research as the guiding spirit and following the principle of standardizing modern Chinese, the CCD7 is "a comprehensive implementation of the China's guidelines and policies regarding the proper use of standard modern Chinese, which draws on and reflects the latest achievements by the State Language Working Commission in the drafting and revision of relevant glyphs and phonetics [1]."

This paper carries out some research on the newly added 400 entries against the commonly adopted inclusion principles. Samples of the new entries are classified according to their deriving fields. Through an in-depth analysis, the author finds that the CCD7's newly included phrases generally follow the four fundamental principles of standardization, common use, systemization, and consultation frequency. A systematic study, however, also reveals some unscientific aspects of entry selection, which are inconsistent, as in the CCD6, to the compilers' principles [2].

2. A Brief Review of New Entries in CCD7. The CCD7 features 403 new phrase (expression) entries, which belong to different syllable structures, parts of speech, and various semantic fields. From the perspective of parts of speech, 14 out of the 395 newly included characters are nouns, 174 are verbs; 16 are adjectives [1].

Regarding syllable structures, 38 out the 403 new entries are monosyllabic phrases (characters), 233 are disyllabic phrases (expressions), 54 are trisyllabic, and 79 have four syllables. It is clear that disyllabic phrases (expressions) account for largest percentage, which reflects the overall trend of Chinese vocabulary's development [1].

From a semantic point of view, the CCD7's new entries mainly derive from politics, the economy and trade, laws and regulations, social life, recreation and sports, cyberspace, human emotions, the military, diplomacy, career development, school education, transportation, environmental protection, linguistics, bioengineering, infrastructure, and medical care. In particular, with the development and popularization of social media, neologisms in the Chinese language come into existence also. Many have become catchphrases widely used by the country's netizens. This is also in line with the basic

principle of "adapting to the development of the society and better reflecting the new features of modern Chinese vocabulary" [3].

Generally speaking, the newly included phrases (expressions) in the CCD7 are mainly nouns, verbs, adjectives. Phonetically, these new phrases are mostly disyllabic. As for the semantics, the entries cover almost every aspect of social life and are very much updated.

2.1. CCD7's Principles of New Entry Inclusion. Historically, the CCD has attached great importance to a scientific and systematic approach to the compilation work [4]. In order to reflect the development of modern Chinese vocabulary in recent years, the CCD researchers and compilers have made an exhaustive effort to gather new phrases and delete outdated ones. The corpus-based compilation work is able to comprehensively reflect the development of modern Chinese vocabulary [5]. Attention is also paid to the new achievements on the form and pronunciation of Chinese characters as formulated and revised by experts and scholars organized by the State Language Commission.

Like its predecessors, the CCD7 has a few fundamental considerations when including new entries. And these are the common use principle, the standardization principle, the systemization principle, and the consultation frequency principle [6].

2.1.1 The Common Use principle. The common use of phrases is reflected in the high frequency of their use in society, the wide range of their use, and the fact that they are mostly recognized by the mainstream media. As Jiang Lansheng [3] puts it, "the inclusion of new entries and new meanings in the CCD is based on the principles of common use and stability. The point is to reflect the development and changes in the Chinese vocabulary system and to promote the standardization of language use. The so-called common-use principle refers to a wide range of commonly used phrases (expressions) in the society, most of which have been approved by the mainstream media. The so-called stability refers to the form of a word, the meaning of a word and its usage has basically stabilized. Some new phrases are not long, but they fit grammar, and can fill the gaps in the Chinese vocabulary system in meaning and usage, so that the more colorful expressions can also be included. In addition, factors such as how to guide the use of society and the norms of the language itself are taken into consideration and are not new to them. "

Some of the phrases that have been widely used because of political and economic activities are also included in the CCD7, such as "中国梦" (the Chinese dream), "新常态" (the new normal), "四风" (Four unhealthy trends), "供给侧" (Supply-side), "双创" (mass entrepreneurship and innovation), "一带一路" (the Belt and Road Initiative), "互联网+" (Internet Plus, which is similar to "Information Superhighway and Industry 4.0"). Obviously, the CCD7 has been well aware of socio-economic development in the process of entry inclusion. It has basically been able to keep pace with the times [1].

The CCD can play a "standardization" role in Chinese vocabulary. It can not only provide entry definitions, but also promote the standardization of the use of Chinese phrases [6]. The purpose of compiling the CCD is to promote the standardization of modern Chinese vocabulary, and the new phrases highlighted in the CCD are usually used more

frequently than others. New phrases used with a low frequency, or within a small group of people, or not accepted by mainstream media are rarely included in it, while phrases that have been deleted are usually phrases that have little or no value to the public.

2.1.2 The Systemization Principle. As a standardized dictionary, the CCD also pays attention to the system and comprehensiveness of the phrases included [7]. Vocabulary of a language is not an isolated element, but a rather systematic one. Supposedly, the CCD should reflect the overall features of the standard Mandarin vocabulary. In the process of including phrases, dictionary compilers should not only consider whether the phrases conform to the standard requirements, but also examine the relationship between the phrases and other phrases in the vocabulary system. The systemization principle requires that phrases in the same group should be included in a uniform manner [8]. At the same time, phrases with similar semantic features and generic relationships may also be included in the dictionary.

The systemization principle emphasizes the connection between the collected phrases and other phrases, and the CCD pays attention to the relevance between the collected phrases in its collection. In the process of including new phrases and deleting old phrases in the CCD7, it is necessary to examine such antonyms as “正能量” (positive energy) and “负能量” (negative energy), which are typical of coinages.

Semantically related phrases are also included in the CCD7, such as such as “裸捐” (an all-out donation), “裸考” (taking an exam unprepared), “裸官” (Naked official or officials who stay in mainland China while their spouses and children reside abroad), and “裸退” (Naked retirement or Assuming no government duty after retirement. There are a myriad of similar phrases in the Chinese language. When compiling a dictionary, it is likely that individual phrases are omitted even with extensively collected data. Following the systemization principle, therefore, can cover new entries as comprehensively as possible.

2.1.3 The Standardization Principle. The aim of compiling the CCD is to popularize Putonghua and promote the standardization of Chinese, besides reflecting the development of this language [9]. As the fundamental principle of the CCD, the standardization principle plays a major role in making sure that the included phrases meet the primary criteria for using Chinese. Any existing language has been going through change and transformation, a continuous process that impacts the development of vocabulary most considerably. Therefore, the standardization of Chinese vocabulary is of particular importance. There are approximately 65,000 phrases in the CCD7 and the standardization work is relatively complicated.

In a face-to-face interview on the CCD6’s inclusion and exclusion of some phrases, Jiang Lansheng [10] opined that “We have included ‘宅男’ (indoorsman) and ‘宅女’ (indoors woman), but we refuse to incorporate ‘剩男’ (literally, leftover men, men at a mature age unable to find an ideal spouse) and ‘宅女’ (literally, leftover women, woman at a mature age with a good education and high salary but unable to find an ideal spouse), because the

former two are neutral in meaning, but the latter two show disrespect for the persons in question, as the reasons for their being single vary. [10]” Jiang added that “as a standard dictionary, we refuse to include this category of vocabulary, though we do not care how people use it in private. [10]”

Apparently, since the CCD6, the use of some phrases has prevailed and become established in both spoken and written discourse, but they are still not included in the CCD7 because they fail to conform to the CCD’s standardization principle [11]. In addition to “leftover man” and “leftover woman”, there are also “同志” (homosexual) and “学渣” (D student; duffer; underachiever or loser). Interestingly enough, the phrase “学霸” (straight-A student) has just been included in the CCD7.

2.1.4 The Consultation Frequency Principle. As an authoritative and widely used dictionary, the CCD not only considers the practicability, versatility and standardization of phrases, but also satisfies the need of a wide range of readers to consult for usage [2]. Therefore it is essential that the CCD follow the principle of “consultation convenience”. In other words, readers should find it easy to look up phrases (including some obscure ones) from the core vocabulary of the Chinese language. For instance, “𧈧” (offend; violate), pronounced as “[gān]”, is a monosyllabic phrase with an added meaning (sense) of 𧈧 “(mosquito larvae)”, pronounced as “[hán]”. Another example is “猻” (golden monkey mentioned in ancient books), pronounced as [róng] [1].

Such phrases have derived from ancient Chinese or dialects and are seldom used in people’s daily work and life today. Consequently, their pronunciation and meanings may have to be looked up before they can be understood. The inclusion of phrases of this type can give readers assistance with any problems that may arise in reading ancient Chinese.

On the other hand, the CCD does not deem it necessary to include every phrase that is associated to the same root character. For example, both “肠炎” (enteritis) and “阑尾炎” (appendicitis) are “old phrases” in the CCD, but phrases like “胃炎” (gastritis), “皮炎” (dermatitis), despite their high frequency uses in real life, have not yet included in the CCD7 mainly because they are formed by certain combination that can be understood with relative ease [11].

3. Some Drawbacks to CCD7 and Suggestions for Improvement.

3.1. Drawbacks. As an authoritative dictionary, the CCD has always neglected the importance of including certain types of proper nouns. In some cases, such neglect is beyond belief. A case in point is “阿司匹林” (aspirin). As a well-known and common medicine, its exclusion forms a sharp contrast to the CCD7’s entry of “阿伏伽德罗常量” (the Avogadro constant), which is by no means a household name [1]. Moreover, like the previous edition, the CCD7’s entry inclusion is inconsistent from place to place [12]. On the one hand, the dictionary does not include geographic names such as “中国” (China), “美国” (the U.S.), “韩国” (South Korea) and “长江” (the Yangtze River), which are also important information for dictionary users. On the other hand, under the entries of “中” and “美”, the definitions show “指中国” (relating to China) and “指美国” (relating to the U.S.)

as both countries' abbreviations. Nevertheless, under the entry of “韩”, there is no such definition as “指韩国” (relating to South Korea).

Another example is the phrase “朋友圈” (circle of friends or moments), which is a function of a social networking app. This term has become an integral part of the lives of the vast majority of ordinary people. Referring to a "relationship circle" made up of acquaintances and semi-acquaintances, it is actually an important channel for individuals to exchange greetings and information. It is worthy of note that Chinese President Xi Jinping started to use this phrase in his formal speeches back in 2015 or even earlier. A similar case is “表情包” (emoticon or sticker). After several years of “trial” use, both phrases have become well-understood and accepted by the public, so they should be considered for inclusion in the CCD. There are many other similar words that are not listed here [12].

3.2. Suggestions. The notion of "standardization" deserves some rational understanding. Being standard does not mean "a rigid adherence to conventions". Language is evolving and so are its standardization requirements. As Chao Jizhou et al. [13] pointed out, "The inclusion of standard terms does not mean rejecting all non-standard entries, nor it means adopting them at will. Rather, it means deciding which phrases to choose while keeping a sense of proportion. In emphasizing “standardization”, we should take into account the objective existence of linguistic facts, and we should not overlook the growth trend in the emergence of particular phrases. When including commonly-used phrases, we should relax the rules for including certain phrases...A proper amount of relaxation in the inclusion criteria is not contradictory to the purpose of standardization. On the contrary, it is a necessary complement to the standardization principle."

In the first place, special attention should be paid to how phrases come and go in modern Chinese society. A contemporary dictionary is indispensable to the general public, so its compilation should not be done behind closed doors. In this regard, a complete corpus should be established for the convenience of data analysis and dictionary compilation.

Secondly, the inclusion principles should be applicable so as to avoid the influence of subjective factors. When it comes to the establishment of the principles, a scientific approach has a vital part to play. To a large extent, the CCD shoulders the responsibility for promoting the standardization of the Chinese language. Therefore, its role of guiding language use may backfire, thus undermining its reliability and authority. It is wise to maintain that lexical norms should adhere to the principle of flexibility. That is to say, flexibility of linguistic norms, rather than the rigidity of those rules, should favor a certain degree of “leniency”.

Finally, the CCD team can turn their eyes outward and learn useful lessons from their counterparts across the globe. For example, the Longman Dictionary of Contemporary English (LDOCE), the Cambridge Advanced Learner's Dictionary, the Oxford Advanced Learners Dictionary (OALD), the Collins English Dictionary, and the Webster's Dictionary...have all done a fantastic job in popularizing and standardizing the English languages.

4. **Conclusions.** Language is a constantly evolving system. The development of a society can have a major impact on the evolution of language components. How a language develops hinges on changes in the cultural, industrial, and economic landscape. Some social phenomena appear and disappear quickly, while others are here to stay. The emergence of new phrases, the disappearance of old expressions as well as their changing senses are always a fair reflection of societal norms that are continuously undergoing transformation. As a very standard and authoritative dictionary, the CCD has accurately reflected the features of modern Chinese vocabulary. Meanwhile, The CCD should also be increasingly mindful of the evolution of the modern Chinese language and take a more open-minded and pragmatic approach to the inclusion of new entries.

REFERENCES

- [1] Dictionary Compilation Team, Institute of Linguistics, Chinese Academy of Social Sciences, *The Contemporary Chinese Dictionary*, The Commercial Press, 2016.
- [2] G. Zhao, The Contemporary Chinese Dictionary, *International Journal of Lexicography*, 2015.
- [3] L. S. Jiang, An Overview of the 6th Edition of the Contemporary Chinese Dictionary, *Lexicographical Studies*, no.2, pp. 1-19, 2013.
- [4] S. X. Lv and S. Hu, *Proc. of the Symposium on the Contemporary Chinese Dictionary*, The Commercial Press, 2004.
- [5] Z. R. Peng, Modernization of the Contemporary Chinese Dictionary and Dictionary Compilation. *Proc. of the 2003 SMCL Conf. on Decision & Control*, The Society for the Modernization of Chinese Language (SMCL). no.11, pp. 87-97, 2013.
- [6] Z. T. Chen, Inclusion of New Entries by the Contemporary Chinese Dictionary (5th Ed.), *Applied Linguistics*, no.1, pp. 12-14, 2006.
- [7] J. T. Han, *An Anthology of Academic Papers on the Compilation of the Contemporary Chinese Dictionary*, The Commercial Press, pp. 129-131, 2004.
- [8] S. X. Lv, On New Meanings of New Words, *Lexicographical Studies*, no.1, pp. 8-14, 1984.
- [9] J. G. Li, A Research into Neologisms and Dictionary Compilation, *Lexicographical Studies*, no.03, pp. 30-38, 1996.
- [10] L. S. Jiang, *Going on in My Own Way [TV Interview]*, accessed on January 25, 2019. http://tv.cntv.cn/video/C14128/20121029100477_CN26, 2012.
- [11] X. J. Li, On the Standardization of Some Phrases from the Perspective of Language Development and Social Psychology, *Language Planning*. no.05, pp. 31-43, 1987.
- [12] L. Su and C. A. Wu, The Impact of Chinese New Phrases on the Development of the Chinese Language, *Journal of Northeast Normal University (Philosophy and Social Sciences)*, no.01, pp. 53-57, 2016.
- [13] J. Z. Chao, Y. H. Shan and J. B. Han, On the Inclusion of Phrases by Standardization-type Dictionaries, *Applied Linguistics*, no.2, pp. 89-93, 1995.