

DISTRIBUTION AND EXPLANATION OF THE MEANING OF THE MODERN CHINESE POLYSEMIOUS WORD “PIN”

XiaoHan Lin¹, XueFang Zhou^{2*}

¹*School of Chinese Language and Culture, Zhejiang Yuexiu University, Shaoxing 312000, Zhejiang, China.*

²*School of International Education, Zhejiang Yuexiu University, Shaoxing 312000, Zhejiang, China.*

Corresponding Author: XueFang Zhou, Email: zhouxf30@163.com

Abstract: The revision of the Dictionary of Modern Chinese (7th Edition) makes full use of various corpora to select or test new words, new meanings and new usages, showing the results of the combination of dictionary compilation and corpus technology. However, it takes time to update and supplement the dictionary, and it is not possible to accomplish it in one go. Some words in the Modern Chinese Dictionary (7th edition) still have problems such as the lack of accuracy in describing the denotation and the lack of scientificity in dividing the denotation. In this paper, we have chosen the word “pín” as the research object, and with the help of the CCL corpus of Peking University, we have statistically sampled the corpus of the modern Chinese word “pín”, and we have taken the 5,000 collated corpus as an important reference, and on the basis of which we have compared it with the English-Chinese dictionary, and have reinterpreted it using the theory of generative thesaurus. Based on this, we compare the English-Chinese dictionaries and apply the theory of generative thesaurus to reinterpret the distribution of “pín” in modern Chinese, and adjust the order of the items according to the frequency of use. is used to help Chinese learners better understand the word “pín”. It is hoped that it can help Chinese learners to better understand and utilize the character “pín” and provide new ideas for lexicographers.

Keywords: Corpus; Distribution of denotations; pín

1 INTRODUCTION

Corpus refers to a large-scale electronic text database that has been scientifically sampled and processed, which can be used for both corpus linguistic research and empirical language research. The CCL corpus is a large-scale corpus produced by Peking University, which can be utilized in various fields of language research, especially in lexicography. However, it still takes time to update and improve the lexicon, and Li S. C. also said, “The refinement of the meaning of common polysemous words in Chinese medium-sized language dictionaries is still at the level of the 1950s and 1960s when the meaning of words was determined by the sense of language, which led to the sloppy division of the meaning of Chinese medium-sized language dictionaries and the omission of the meaning of the words.” In other words, there is still much room for expansion in the current dictionary. In addition, “Interpretation is the soul and core of a dictionary, which has a great influence on the readers' correct understanding of the meaning of words and the use of words.” The completeness, accuracy and systematicity of the interpretation are important criteria for evaluating the quality of a dictionary; adjectives, as an important part of vocabulary, reflect people's perception of the nature or attributes of things, and studying the interpretation of adjectives clearly is of great value for improving the quality of the interpretation of adjective dictionaries and understanding the way of thinking of the Han Chinese people. On these bases, this paper selects the commonly used polysemous word “pín” in modern Chinese as the object of study.

2 REVIEW OF RELEVANT RESEARCH

Li Qiang [1] conducted a pioneering study on noun definitions in lexicography from the perspective of Generative Lexicon Theory, identifying several systemic issues in current lexicographical definitions and proposing solutions grounded in qualia-based analysis. Their research demonstrated that adopting a “thematic argument + qualia role” definitional schema enhances both the precision and systematicity of lexical entries. This model prioritizes three definitional sub-principles—patterning, uniformity, and identity—thereby facilitating the construction of lexicographic definitions that are more standardized, streamlined, and reader-accessible. Building upon Yuan Yulin's [2] theoretical framework for qualia role classification, this study applies the “thematic argument-qualia role” schema to reanalyze the adjectival senses of the character “pín”, aiming to achieve a comprehensive and semantically precise description of its adjectival semantics.

The CCL Corpus, developed by the Peking University Center for Chinese Linguistics (PKU-CCL), constitutes a large-scale Chinese linguistic database encompassing not only written texts (e.g., newspapers, books, academic journals) but also multimodal resources such as spoken language corpora (e.g., Beijing dialect surveys), film/TV scripts, and web-based materials. As one of the most extensive repositories of contemporary Chinese texts, the CCL Corpus prioritizes granular corpus categorization, thereby enhancing the representativeness of sampled data. This study leverages the PKU-CCL Corpus to extract linguistic instances of the character “pín”. As of September 2023, the corpus contains

2,773,443 modern Chinese instances of “pín”. Following the methodology introduced by Li S.C. [3], a balanced subcorpus specific to “pín” was constructed to ensure analytical rigor.

3 RESEARCH DESIGN

3.1 Distribution of the meaning of “Pín”

3.1.1 Adjective: poor (as opposed to rich)

It is used most frequently in 4521 items, accounting for 90.42% of the total corpus. In the present Han, the meaning of the word “pin” is interpreted in the way of word by word interpretation. In the interpretation of the word by word, it is generally used to explain the meaning of another word by a synonym, a related word or an antonym.

When we searched for the word “pín” in the Modern Chinese Dictionary (7th Edition), we found that its corresponding meaning is as follows: “It means living in poverty and lacking money (as opposed to being rich).” [4] It can be seen that the character “pín” is used in a descriptive sense in the Han dynasty. Descriptive interpretation can better avoid semantic repetitions and conflicts and deepen dictionary users' understanding and application of the interpreted words. Referring to the interpretation pattern of this item, we can find that the interpretation of the word “pín” highlights the subject element ‘life’ and the evaluative role “poverty”. The thesis can find that the interpretation of the word “pín” highlights the subject element “life” and the evaluating character “poverty” and “lack of money (as opposed to ‘rich’)”. Among them, “subject element” refers to the indispensable element that constitutes a complete sentence and is mainly used as the subject; “evaluative role”, as a kind of “object role”, is a structuralization that gives the lexical items relational force. The “evaluative roles”, as a kind of “object roles”, are structured representations that give relational force to lexical items.

According to the Chinese-English Dictionary, the English equivalent of the word “pín” is “poor”. The English equivalent of the word “pín” is ‘poor’ in Oxford Advanced Learner's English-Chinese Dictionary (9th Edition) as follows: “has very little money; not having enough money for basic needs”. In the interpretation of the word “pín” in the book, the evaluative role of “having very little money” and the functional role of “not having enough money for basic needs” are captured separately.

The analysis of the annotated corpus reveals that the adjective “pín” predominantly modifies thematic arguments related to economic domains, including “region”, “society”, “condition” and “population”. Representative examples from the corpus demonstrate this pattern:

- ① China adheres to an independent foreign policy of peace, advocating that all countries worldwide, regardless of size, wealth, poverty, or strength, should be treated equally. (Contemporary/Periodicals/People's Daily/October 1993)
- ② He found an elderly couple living in extreme poverty. The man was nearly blind, and the woman lay bedridden. Their household retained only two days' worth of coal. (Contemporary/Periodicals/Reader/Reader Anthology)
- ③ The anthem Towards Magnificence resonated in the theater: 'We pursue you, joining arms, guiding our nation from destitution to prosperity, making the world witness China's miracles' — this embodies the aspirations of the 'Ten Outstanding Youths'. (Contemporary/Periodicals/People's Daily/October 1993)

The semantic feature of “pín” lies in its characterization of economic deprivation, foregrounding the evaluative role of the modified thematic argument. It is therefore recommended that its lexical entry be redefined through a descriptive approach as follows: Adjective. Economically deprived (as antonymous to affluent).

3.1.2 Lacking; insufficient

There are a total of 312 entries for this term, accounting for 6.24% of the total corpus, and it is used more frequently. The Chinese Journal of Modern Chinese has adopted the interpretation methods of word-for-word and combination-for-explanation for this item. Among them, the combination of interpretation is to use two or more words to make annotations, mainly by supplementing or restricting the meanings of the interpreted words to illustrate the meanings of the interpreted words.[5] This way of interpretation, as a way of interpretation that makes full use of people's existing knowledge and does not directly state the connotation of the word meaning, stays at the surface level of interpretation, and cannot comprehensively and deeply reveal the lexical characteristics of the word being interpreted.

In the Modern Chinese Dictionary (7th Edition), “lack” is interpreted as “lack of movement (mostly referring to the insufficient number of people or objects)” with the subject element pointing to people or objects; ‘insufficiency’ is interpreted as “insufficiency (of a certain number)” 19, with the subject element of this interpretation being “number”. “not enough (of a certain number)”, the subject element of this interpretation is “number”. Combined with the labeled corpus, we can find that the subject thesis element modified by “pín” is often “land”, “person”, “kick”, “figurative thinking”, ‘image’, and “object”. figurative thinking”, etc. These subjective elements can be categorized into two groups according to concrete and abstract perspectives, because when the subjective elements change, the related object roles will also change accordingly. When the subject-matter element modified by “poor” is a concrete person or object, such as “land”, ‘man’, “young woman “sun” and so on, the word ‘poor’ shows the characteristics of people or things at a low level of quantity or quality”, highlighting the formal role of the subject element. The total number of such items in the corpus of the word “pín” is 222, with a usage frequency of about 4.44%. Representative examples from the corpus demonstrate this pattern:

- ① China adheres to an independent foreign policy of peace, advocating that all countries worldwide, regardless of size, wealth, poverty, or strength, should be treated equally. (Contemporary/Periodicals/People's Daily/October 1993)
- ② He found an elderly couple living in extreme poverty. The man was nearly blind, and the woman lay bedridden. Their household retained only two days' worth of coal. (Contemporary/Periodicals/Reader/Reader Anthology)
- ③ The anthem Towards Magnificence resonated in the theater: 'We pursue you, joining arms, guiding our nation from destitution to prosperity, making the world witness China's miracles' — this embodies the aspirations of the 'Ten Outstanding Youths'. (Contemporary/Periodicals/People's Daily/October 1993)

Drawing on the "Thematic Argument + Qualia Role" framework from Generative Lexicon Theory (Pustejovsky 1995), the semantic extension of “pín” in these examples demonstrates its capacity to quantify or qualify material/abstract entities through scalar evaluations of deficiency.

3.1.3 low number or quality of people or objects, etc.

When the subject element of “(pín)” points to abstract things, the total number of items in the corpus of the word “pín” is 90, and the frequency of use is about 1.8%. In these corpora, the subject elements modified by the word “pín” are abstract existences, such as “life”, “kicking”, “knowledge”, etc., which show that thinking and thinking are not the same. “which characterizes thinking or skills as at a low level” and highlights the evaluative role of the subjective element. Representative examples from the corpus demonstrate this pattern:

- ① This concrete imagistic thinking in children is inextricably linked to the impoverishment of their knowledge and experience and the dominance of their first signal system activity. (Contemporary \CWAC\APM0096)
- ② Despite better material conditions abroad, there is a spiritual impoverishment compared to home. Coupled with my lumbar injury (herniated disc), I cannot adapt to the intense competition overseas. (Contemporary Press: Selected Newspaper and Journal Articles (1994))
- ③ For an astute painter or writer, who excels at capturing people's fleeting glances and idiosyncratic speech, the absence of these elements would render their imagistic thinking markedly impoverished. (Contemporary Periodicals/Reader/Reader Bound Volume)

3.2 Distribution of the Meaning of “Pin²”

3.2.1 The rambling is tiresome

This is an adjectival interpretation of “pín”, mostly used in dialects. It is used mostly in dialects. There are 47 entries in the corpus, accounting for 0.94% of the total corpus, and the frequency of its use is second to none. The present Chinese text adopts a word-for-word interpretation for this item. It emphasizes the formal role of “rambling” and the evaluative role of “repulsive”.

The subject thesis elements in the above example sentences are “Cao Xueqin himself,” “the person who writes the various ‘after-dreams,’” “I” When the object of description is a person, such as “I am not poor”, “I am poor”, “I am poor”, “I am poor”, it refers to the fact that Cao Xueqin is not poor. “, meaning a person's language is so garrulous that one gets tired of listening to it. “pín” actually modifies the “language” spoken by the “person”, and ‘mouth’ is used to describe a person's “mouth”. The actual modifier of “pín” is the “language” spoken by the “person”, not the “person”, and the implied ‘language’ is the functional role of the subjective element “person”. The implied “language” is the functional role of the subject element “man”; when the subject element “pín” refers to the organs of the human body such as “mouth” and “tongue”, “pín” actually modifies the ‘language’ of “man” as well. [6] Actually modifies the “language” spoken by the “mouth” and “tongue” rather than the ‘mouth’ and “tongue” themselves. “The implied ‘language’ is the functional role of the subjective element “mouth”. The subjective element of the word “pín” actually points to ‘language’, so the original term is re-described as a new term by the method of “subjective element + object role”.

4 CONCLUSION

Adjectives are one of the most important types of vocabulary in the modern Chinese lexical system and have always been the focus of attention in the academic field. However, the current research on adjectives focuses more on syntactic functions, and the research on individual adjectives has not yet formed a systematic and comprehensive study. The birth of corpus technology has pushed forward the progress of lexicography, but it takes time to update and replenish the lexicon, and it is impossible to do it overnight.[7] Many commonly used polysemous words are included in “Present Chinese”, and comparing with Chinese and English dictionaries, [8]we can see that the fineness of the meaning of some of the words still needs to be improved. In this paper, based on the interpretation of “pín” in *Modern Chinese Dictionary*, we apply the theory of generative thesaurus, and combine it with 5000 specific corpora extracted from the CCL corpus, to try to reclassify and describe the distribution of the meaning of “pín” in modern Chinese, and to propose an optimization strategy. The optimization strategy is proposed.

In terms of the scientific division of terms, according to the rule that when the subject thesis element changes, the physical role will also change accordingly, combined with the specific corpus, and in accordance with the depiction of “subject thesis element + physical role”, the author divides the “pín¹” term 2 into two terms.

In terms of the accuracy of the depiction of the meaning items, the author combines the specific corpus, compares the English-Chinese dictionary's mode of interpretation, according to the depiction mode of "subject thesis + object role", re-describes the meaning items of "pín¹" and "pín²". The meaning of "pín¹" and "pín²" are re-described in accordance with the "subject thesis element + object role".

Due to the limitations of knowledge, the author has only conducted a preliminary study on the interpretation of the word "pín", which is still incomplete in many aspects. For example, the theoretical study is not solid enough, the selected corpus is not sufficient, and the reference dictionaries are not diversified enough, etc. I hope that will be able to make some useful references for the future study of the interpretation of the word "pín". I hope that this study can provide some useful references for future research on the meaning of the word "pín".

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

- [1] Li Q. A Study of Semantic Knowledge Systems Based on Generative Thesaurus Theory and Theory of Theoretical Meta-Structure. *Chinese Language Teaching in the World*, 2022, (03): 24-28.
- [2] Yuan Y L. A Study of Semantic Knowledge Systems Based on Generative Thesaurus Theory and Theory of Theoretical Meta-Structure. *Chinese Journal of Informatics*, 2013, (06): 42-45.
- [3] Li S C. A Corpus-Based Study on the Distribution of Meaningful Items of the Chinese Character "huáng" in Modern Chinese. *Journal of Ocean University of China (Social Sciences)*. 2016, (05): 29-35.
- [4] Zhang Y H. The New Concept of International Rhetoric Modernization Technology: Rhetoric Corpus Dataization. *Rhetoric Research*, 2012, (02): 38-44.
- [5] Pan Y Y. Frame Semantics: Theory and Application. *Foreign Language Studies*, 2003, (05): 25-26.
- [6] Zhang Z Y, Zhang Q Y. *Theoretical lexicography*. Beijing: The Commercial Press, 2015.
- [7] Tan J C. Semantic induction and grammatical analysis in dictionary interpretation--Talking about the revision of entries in the 6th edition of the Modern Chinese Dictionary. *Chinese Language*, 2012, (06): 52-56.
- [8] Gu Y. On the Introduction of Meta-Structural Theory. *Contemporary Linguistics*. 1994, (01): 12-16.