

Research on the Construction of Chinese Song Hierarchical Database Oriented to Chinese Teaching*

Yuqi Zhang

Beijing International Studies University, China

Xiujuan Han

Beijing International Studies University, China

Qingping Hu

Beijing International Studies University, China

Fei Song¹

Beijing International Studies University, China

Abstract—As a critical auxiliary method in teaching Chinese as a foreign language, Chinese songs have the characteristics of a strong sense of rhythm, extensive vocabulary, and rich cultural knowledge. Previous studies have mostly discussed the use of Chinese songs in Chinese teaching from the perspective of teaching methods, teaching experiment design and textbooks. There is no specific research on the construction of a database of Chinese song ratings for Chinese learners of different levels. Based on Internet resources, this study collected 9098 Chinese song lyrics and built a database. Based on the analysis of the characteristic data in the database and the method of building a vector space model, four songs of different difficulty levels were divided, which was respectively targeted at Chinese learners of different levels. This research can reduce Chinese teachers' lesson preparation time and improve teachers' teaching efficiency.

Index Terms—Chinese songs, Database, Chinese teaching

I. THE ORIGIN OF RESEARCH

Today, with the vigorous development of Chinese teaching, first-line teachers gradually introduce Chinese songs into Chinese class, thereby arousing the interest of learners to achieve their goals of spreading Chinese. However, at different stages of Chinese language learning, the role of Chinese songs in the Chinese class is very different. In the primary stage, the use of Chinese songs can effectively alleviate the students' fear and anxiety about Chinese learning. In the intermediate stage, the interesting lyrics in the Chinese songs can help the learner to expand their vocabulary. In the advanced stage, the Chinese songs can help the learner to gain a deeper understanding of its connotation and enhances their own cultural knowledge.

In China, Zhao Yuanren (1980) is the first expert who conducted a research in linking music and language which is aimed to study Chinese. The "Five Degree Tone Marking Method" he founded opened the door of music in linguistics. From the perspective of song genre analysis, Zhao Shouhui and Luo Qingsong (1994) first proposed the application of folk songs in teaching Chinese as a foreign language. Then, scholars such as Zhao Fangyu (1998), Zhang Jie (2007), and Sun Ke (2013) also believed that ethnic music was an excellent carrier which could promote Chinese teaching as a foreign language. From the perspective of Chinese song textbooks, Wang Yu (2017) used singing, learning, listening, singing, Chinese, and China as keywords to search on major sales platforms. A total of 18 textbooks of this type were found, and this type of textbook is mainly suitable for Children and beginners who just started to learn Chinese.

A database is a warehouse with specific rules for storing data. In language teaching, the database is mainly embodied as a corpus. Zhang Baolin (2010) and Zheng Yanqun (2013) emphasize that the construction of a corpus can help to summarize the usage and meaning of a certain language form in the context. In addition, some scholars such as Xing Hongbing (2005), Shi Zhengyu (2008), Liu Xiang (2010), and Liu Shantao (2011) took Chinese characters and vocabulary as the primary research objects and established hierarchical Chinese character splitting and Chinese character basic component databases, teaching fonts, and Chinese characters, the representative ontology knowledge bases such as the teaching resource database of word source literacy and the teaching information database of Chinese

* The paper is supported by the graduate scientific research program of Beijing International Studies University

¹ Corresponding author: Fei Song, bestsongfei@126.com, Beijing International Studies University, China

new words to foreigners also which provides great help for the development of the syllabus of Chinese characters and vocabulary.

Throughout previous studies, the author found three problems: first, when scholars introduced Chinese songs into the teaching of Chinese as a foreign language, there was no standard for Chinese songs in terms of song search and selection. Second, the research of database construction is versatile. But currently, there is no database of Chinese song grades that can be directly used for Chinese language teaching. Third, in the teaching of Chinese language, the division of the text difficulty level mainly depends on the Chinese characters and word frequency in the statistical text. The method of dividing the difficulty level of the text from the perspective of the similarity calculation of the vector space model and typical individual cases has not been widely used.

Therefore, this study will focus on the application of Chinese songs in Chinese language teaching, plan to establish a database of Chinese song grades, and draw the criteria for the lyrics of song lyrics in order to better serve the teaching of Chinese as a foreign language.

II. PREPARATION OF THE DATABASE CONSTRUCTION

Before building a "Chinese song Hierarchical database", a lot of preparation work needs to be done. It mainly includes collecting Chinese song lyrics corpus and performing word segmentation processing on song lyrics, selecting vocabulary level outline and vocabulary connection with lyrics text and other tasks. It will be described in detail below.

A. Collection of Database Materials

1. Collection of Lyrics Data

Based on Internet resources, the author uses web crawler technology to randomly obtain 10,000 initial Chinese lyrics data from the Chinese song music platform. After effective secondary data cleaning, a total of 9,098 correct song lyrics data are retained. After a general analysis of these Chinese songs, it was found that theme type and creation time, the collected lyrics of 9098 songs all met the needs of this research considering the song style, and they were stored in the database as "ID", "Song title" and "Lyrics".

2. Vocabulary Hierarchical Outline Selection

This research needs an authoritative vocabulary level outline, which is used as the standard for determining the word level after the song lyrics are segmented. Zhang Baolin (2005) pointed out that the current grade syllabus in China is divided into three types, namely syllabus, level syllabus and vocabulary. Considering the needs of this research, the author is comparing "Chinese Proficiency Vocabulary and Chinese Character Grade Outline" (1992) (hereinafter referred to as "Old HSK Vocabulary Outline"), "New Chinese Proficiency Test (HSK) Vocabulary Outline" (2009) (hereinafter referred to as "New HSK Vocabulary Outline", "Chinese International Education Syllable Chinese Character Vocabulary Classification" (2010) (hereinafter referred to as "Vocabulary Classification"), it is planned to select "Vocabulary Classification" as the judgment of the lyrics of Chinese songs in this article. Regarding the standard, the main reasons are considered from the status of the grade outline and the vocabulary included.

The "Old HSK Vocabulary Outline" (1992) was published earlier. The 8822 words included in it are divided into four levels: A, B, C, and D, and they cannot adequately meet the needs of society. The "New HSK Vocabulary Outline" was launched in 2009. It was developed by researchers after drawing on the advantages of the "Old HSK Vocabulary Outline" and drawing on the latest achievements of the International Chinese Language Test. It received 5,000 words and was divided into one to six levels which is for HSK candidates. "Vocabulary Classification" was launched in October 2010. It is the first national standard of Chinese international education and a representative of the times and authority. A total of 11,092 words were received and divided into three levels in detail to serve Chinese learners of different levels.

After collecting the corpus of the database, the author will then perform word segmentation processing on the song lyrics and associate the processed new words to the corresponding vocabulary level which also provides a reference for the establishment of the later hierarchical database.

B. Construct a Song Vocabulary Hierarchical Database

1. Song Lyrics Participle Processing

After obtaining the lyrics of 9098 Chinese songs based on data mining, the author plans to use Corpus Word Parser to perform part-of-speech tagging on the lyrics of each song, and statistically obtain the length of lyrics, the number of the participle, participle, and part of speech. According to statistics, the total word count of 9098 Chinese song lyrics totals 2575389 words. Afterword segmentation, a total of 1785273 words were obtained, and the average number of words per Chinese song was 283.07. Select Access as a library building tool and import all statistical data into the database. The structure of the library is: ID, song title, length of lyrics, number of participles, participle, part of speech (as shown in Figure 2-1). Name the database "Lyrics Vocabulary database" refers to a vocabulary database that includes word segmentation processing. The "lyric vocabulary database" is the basis of the "lyric vocabulary rating database".

ID	song title	length of lyrics	number of parts	participles	part of speech
1	Angel's wing	256	182	zhe	u
1	Angel's wing	256	182	wo	r
1	Angel's wing	256	182	ti	v
1	Angel's wing	256	182	tianshi	n
1	Angel's wing	256	182	ge	q
1	Angel's wing	256	182	shouhu	v
1	Angel's wing	256	182	hui	vu
1	Angel's wing	256	182	wo	r
1	Angel's wing	256	182	congci	d
1	Angel's wing	256	182	meiyou	v
1	Angel's wing	256	182	cibang	n
1	Angel's wing	256	182	huaguo	v

Figure 2-1. Example of lyrics vocabulary database

2. Related Song Vocabulary Level

After the song vocabulary database is completed, the words in it need to be related to the same words in "Vocabulary Classification" to find the level corresponding to each word.

In "Vocabulary Classification", some words have some part-of-speech tags, and the words contain two or more parts of speech (see Figure 2-2), which need to be properly resolved.

ID	number	level	word	pinyin	part of speech
9 8		附	挨家挨户	āi jiā āi hù	
10 9		三	挨着	āizhe	
12 11		二	挨打	ái dǎ	
38 37		—③	安排	ānpái	v/n
40 39		二	安慰	ānwèi	v/n
49 48		—③	按	àn	v/p
83 82		—③	把握	bǎwò	v/n

Figure 2-2. Examples of special words in "Vocabulary Classification"

First of all, a preliminary analysis of the "no part of speech" words selected in the "Vocabulary Classification" found that these words are mostly phrases, proper nouns, idioms, etc. The level corresponding to each word is unique (as shown in the figure 2-2). Therefore, it is vital to associate these "no part of speech" words with the "Lyric Vocabulary Database" to match the corresponding phrases, idioms, proper nouns in the database to the corresponding levels. Then, the words "two or more parts of speech" were screened from "Classification of Vocabulary" for analysis. It was found that the two parts of speech in the part of words corresponded to a level, so when the part of words was associated with the "Lyrics Vocabulary Database". If the words in the "Lyrics Vocabulary Database" can be matched with a part of speech in this type of word, it will be considered that the word level can be found. Finally, it is difficult for the massive vocabulary in the "Song Vocabulary Database" to be related to the corresponding level in "Vocabulary Classification", so the author named the words that are not related to the corresponding level as "expansion-level words". According to statistics, this part of the total number of words is 844122, accounting for 47.28% of the total number of words in the database. The word types are 524558 in total, accounting for 53.58% of the database word types.

So far, the author has matched all the words in the "song vocabulary database" to a level. There are four levels of words, namely level one, level two, level three, and extended level.

3. Calculate the Proportion of Song Vocabulary Level

Next, we need to count the number of word types and word frequencies contained in the songs in the "song vocabulary database" at different levels, and then calculate the proportion of word frequencies at that level. The specific calculation formula of the level ratio is:

$$\text{proportion} = \frac{\text{word frequency}}{\text{total word frequency}}$$

Through word segmentation processing, vocabulary level association, and level ratio statistics, important data such as word frequency, number of word types, and level ratio in a song are obtained. They are mainly "first-level word frequency, first-level word number, first-level word" Proportion, secondary word frequency, secondary word types, secondary word ratio, tertiary word frequency, tertiary word type, tertiary word ratio, tertiary word frequency, tertiary word number, tertiary word, Word Proportion, Expansion Level Word Frequency, Expansion Level Kind of Words, Expansion Level Proportion. Import it into the database to establish a "song vocabulary rating database", reflecting the proportion of each song vocabulary rating. The fields included in the library are: ID, song title, lyrics, total word count,

first-level word ratio, first-level word frequency, first-level word type, second-level ratio, second-level word frequency, second-level word type, third-level Proportion, tertiary word frequency, tertiary word types, tertiary appended ratio, tertiary attached word frequency, tertiary attached word number, extended level proportion, extended level word frequency, extended level word number.

III. CHINESE SONG HIERARCHICAL DATABASE CONSTRUCTION

When using the vector space model to classify Chinese lyrics, this paper plans to let the computer automatically calculate the similarity of the lyrics, and merge automatically above a certain threshold to form different classification categories.

A. The Specificity of Song Lyrics Suitable for Teaching

1. Typical Song Lyrics Data Analysis

After reading the previous research results and conducting a four-month classroom practice, the author found that "Meet", "Blue Lotus", "Where did all the time go", "The Moon Represents My Heart" and "Lake Baikal" these 5 Chinese songs are more suitable for teaching Chinese as a foreign language. Among them, the teaching effect of "The Moon Represents My Heart" and "Where Does Time Go" is obviously better than the other three. After analyzing the data of these five songs, it is found that the higher the level of the song's lyrics, the better the teaching effect of the song in the class; the higher the proportion of expansion-level words, the poorer the teaching effect of the song in the classroom. That is, the proportion of the first-level vocabulary in song lyrics has a crucial position.

2. Analysis of all song data summary

If the first-level words in a song's lyrics occupy a relatively high level, it reflects that the song's use of simple words is not difficult for students to understand, and the teaching effect may be better. On the contrary, if the expansion of Chinese songs is relatively high, it means that the difficulty of using words in a song is not conducive to students' understanding. Therefore, based on the proportion of first-level words in all songs in the "Song Vocabulary Rating Database and Song Vocabulary Rating Database" established in Chapter 2, the author arranges them in descending order to make a scatterplot 3-1) for analysis. In the figure, the abscissa represents the sequence number of each song, and the ordinate represents the proportion of first-level words.

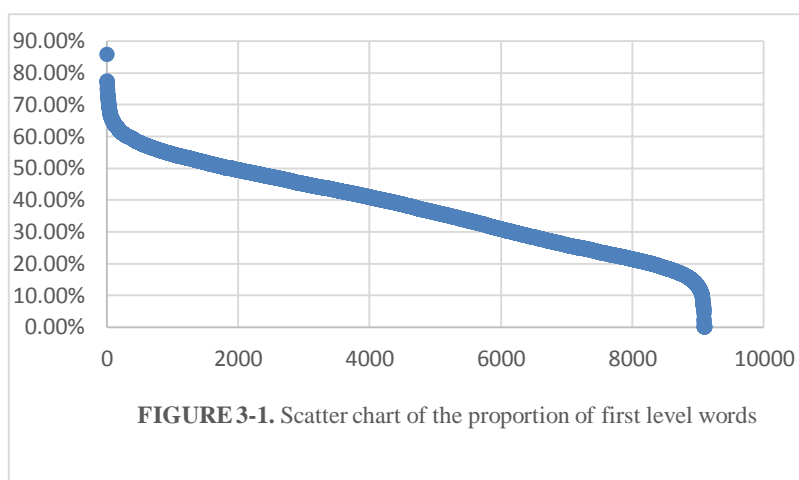


FIGURE 3-1. Scatter chart of the proportion of first level words

According to the data change trend, two inflection points can be found in the figure: one is (174,62.28%), that is, at the data number 174, the first-level vocabulary accounts for 62.28%. The changing trend of the data on both sides shows a trend from fast to slow. Another place is (8698, 17.11%), that is, when the data serial number is 8698, and the first-level share reaches 17.11%, the changing trend of the data on both sides of the data shows a change from slow to fast. Therefore, according to these two inflection points, songs can be roughly divided into three categories: first-class words occupy a relatively high class, first-class words account for a gradual change in class, and first-class words occupy a relatively low class. Next, the author will analyze these three types of data in detail in order to discover the characteristics of the data.

a. Analysis of the Data at Higher Proportion

Screening out Chinese songs with "the first-level vocabulary accounted for more than 62%", a total of 181 lyrics data were obtained. A random sampling of 10% was conducted to analyze one by one, and it was found that such songs sing clearly when singing, most of the lyrics are common daily vocabulary, and the theme is mostly describing love, mood or a specific thing. In addition, the author also found that the second-level and third-level words in this category account for a relatively small amount, and the rest are extended-level words. Obviously, this part of the Chinese song is suitable for teaching Chinese as a foreign language, especially for elementary learners with a low level of Chinese.

b. Analysis of Data at Lower Proportion

Filter out the Chinese songs whose "the first-level vocabulary accounts for less than 17%" and get a total of 388 lyrics data. A random sampling of 10% was conducted to analyze one by one, and it was found that many words of this type of song have phonation during the singing process. Besides, the lyrics are mostly ancient words while the meaning is obscure. The theme of the song is mostly to describe love, but there are many expressions. The tendency is to use lyrical sceneries, lyrical sceneries, etc. Besides, in the extracted sample data, three-level vocabulary and extended-level vocabulary usually occupy an important semantic component in the sentence. Therefore, when the proportion of first-level words in a song is less than 17%, the lyrics of the song are too tricky, and the learning of such words is not very helpful for students to improve their oral ability and pronunciation. Therefore, the author believes that such songs. It is not recommended to be used in Chinese as a foreign language class. It can be used as an extra-curriculum song of interest in teaching Chinese as a foreign language to arouse students' interest in Chinese song culture.

c. Data Analysis of First- level Accounted for Relatively Flat Areas

By observing the pictures, we can know that the lyrics data of these songs change evenly, so the author decided to explore the rules further. The author conducts stratified sampling of the data in this area and performs a 1% sampling every 15% with a first-level ratio. After the division, the data is divided into three layers: the first layer extracts 22 data, the second layer extracts 23 data, and the third layer extracts 30 data.

Through the analysis of the word segmentation of these sampled data, it is found that in addition to the evenly decreasing proportion of the first-level vocabulary, the distribution of the second-level, third-level words and extended levels is irregular. Therefore, through manual screening, the author conducted a one-by-one investigation on the 75 Chinese songs, and finally successfully found 29 Chinese songs suitable for intermediate-level Chinese students to learn (as shown in Figure 3-2).

1	歌名	16	有点甜
2	偏爱	17	情侣装
3	微行循	18	素颜
4	K歌之王	19	感恩的心
5	花房姑娘	20	亲爱的，你的棉裤真的不在我家
6	后会无期	21	小蝌蚪找妈妈
7	我的好兄弟	22	爸爸
8	当你孤单你会想起谁	23	沈阳北站
9	痴心绝对	24	放生
10	当爱已成往事(霸王别姬 电影插曲)	25	春天里
11	父亲	26	冰雨
12	王妃	27	昨天是找不到的永远
13	过火	28	蓝莲花
14	童话	29	铁窗泪
15	会呼吸的痛	30	有多少爱可以重来

Figure 3-2. 29 songs suitable for intermediate Chinese teaching

The author decided to use the vector space model to analyze this part of the data, and the 29 Chinese songs manually selected as the feature data modelling analysis. In the comparison song vocabulary rating database, the word segmentation results of each level of the song are counted, the number of words and the proportion of word frequency of each lyric is recorded, and this is used as the dimension of a Chinese song rating. There are three main reasons: first, the frequency and scope of different levels of words are different, which can reflect the difficulty of the lyrics; second, the number of words of different levels of words can reflect the difficulty of the lyrics and the impact on teaching; Third, the proportion of grades and the number of words in a song can be considered as independent features, independent of the total number of words in the song.

Next, the writer will use the vector space model method to divide the song lyrics data suitable for the intermediate level of teaching Chinese as a foreign language in the gentle area. The calculation formula of the angle cosine value is as follows: where A represents the proportion of each level, and B represents the number of words.

$$\frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

The vector space model has a total of ten dimensions, namely: first-level word-level ratio, first-level word types, second-level word-level ratio, second-level word type ratio, third-level word-level ratio, third-level word words The number of species, the proportion of the three-level attached words, the number of the three-level attached words, the proportion of the expanded word grades, and the number of expanded word words.

B. Construct a Vector Space Model

1. Data Normalization

Before constructing the vector space model, it is necessary to normalize the data of different units. The specific conversion function of normalization is as follows: x represents the current value, \min represents the minimum value in this dimension, and \max represents the maximum value in this dimension.

After normalizing 29 songs, the 29 sets of data are calculated by weighted average, and the calculation result is used as a vector suitable for the vector space model of intermediate Chinese teaching.

2. Iterative Calculation to Determine the Threshold

By calculating the angle cosine of the space vector of the data in the gentle area and the vector suitable for the intermediate teaching model, the threshold for dividing songs of medium and high difficulty is found in the process of iterative calculation, and the data above this threshold are classified as a class, low The threshold data is classified into another category.

After four manual iterations, sample analysis and manual selection of songs suitable for intermediate Chinese teaching, the author found that when the vector space angle cosine is 0.95, after random analysis of 5%, 88.4% of the Chinese songs It can be applied to intermediate Chinese teaching classes for foreigners. Therefore, the author uses the cosine value of 0.95 as a threshold because manual detection found that 88.4% of the Chinese songs in the songs divided by this threshold are suitable for intermediate Chinese teaching.

C. Song Difficulty Level Division Results

1. Beginner-difficulty Songs

By observing the "scatter plot of first-level word proportion", we can find that the first-level proportion has a great influence on the song. Therefore, the lyrics data of more than 62% of the first-level vocabulary is classified as "primary difficulty songs", a total of 181 songs. The following is a list of ten representative songs of primary difficulty:

TABLE 3-1.
REPRESENTATIVES OF SONGS OF PRIMARY DIFFICULTY

1	Too much love	6	When you
2	Confession of love	7	Little train
3	Just follow me if love me	8	Suspect
4	The day will never know how dark the night is	9	The moon represents my heart
5	Tea world	10	You used to be a young boy

2. Intermediate-difficulty Songs

By constructing a vector space model suitable for intermediate-level Chinese teaching songs, the threshold between finding intermediate-level difficulty songs and advanced-level difficulty songs is 0.95. Therefore, when the angle cosine of a Chinese song is calculated to be higher than 0.95, the song is classified as "intermediate difficulty song". Through calculation, a total of 3085 Chinese songs were obtained as intermediate difficulty songs. Here are ten representative songs of intermediate difficulty:

$$x^* = \frac{x - \min}{\max - \min}$$

TABLE 3-2.
REPRESENTATIVES OF SONGS OF INTERMEDIATE DIFFICULTY

1	Don't love you	6	The life in Beijing East Road
2	The winter of love	7	How do you want me to be
3	Zebra, zebra	8	Fish looking for water
4	Aries	9	How much love can be repeated
5	Flower Girl	10	Meet

3. Advanced-difficulty Songs

When the angle cosine value of a Chinese song is less than 0.95, the song is classified as "advanced difficulty song". Through calculation, a total of 5444 Chinese songs were obtained as advanced difficulty songs. The following is a list of ten high-level difficulty representative songs:

TABLE 3-3.
REPRESENTATIVES OF SONGS OF INTERMEDIATE DIFFICULTY

1	Ambiguous	6	Life in full bloom
2	Hug back to back	7	On July
3	Secret	8	Normal road
4	No contact any more	9	Love you
5	Once you	10	Finally you are here

4. Interest Expansion Songs

By observing the "scatter plot of first-level word proportion", it is found that the first-level proportion has a great influence on the song. When the first-level vocabulary of song lyrics is less than 17%, analysis of this part of the data reveals that this part contains a lot of ancient songs, which makes it difficult for students to understand, and the effect of promoting student learning is not obvious. For "Interest Expansion Songs", the following lists ten representative songs of interest expansion categories:

TABLE 3-4.
REPRESENTATIVE SONGS OF INTEREST EXTENSION

1	The moon in border town	6	The trouble beauty brings
2	Picking Osmund	7	How much sorrow
3	Spring morning	8	The reeds and rushes
4	The elegance and talent in the world	9	Peer rain
5	Ah, homeward bound I go	10	Black robe alley

IV. RESEARCH RESULTS

After statistics, the Chinese song ranking database has a total of 9098 Chinese songs. The distribution of songs in each level is shown in Figure (4-1)

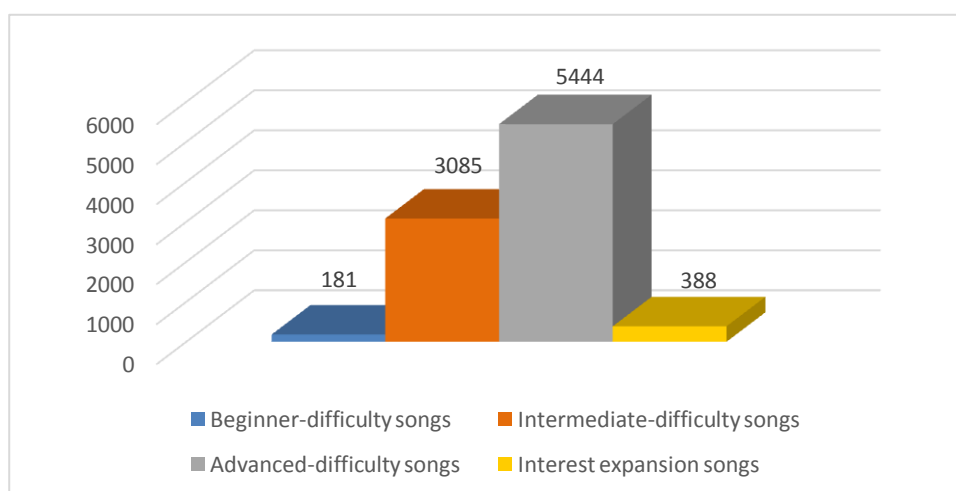


FIGURE 4-1. The distribution of songs in the Chinese song ranking database

It can be seen from the chart that the database divides Chinese songs into four difficulty levels, namely: primary difficulty, intermediate difficulty, advanced difficulty, and interest expansion. The author believes that the Chinese songs in the database can be applied to teaching Chinese as a foreign language, but when faced with different teaching objects, the use of songs is very different.

First, for learners of elementary Chinese level, at the beginning of Chinese learning, the Chinese characters and vocabulary they master cannot be fully applied in daily life. It is difficult for them to choose the appropriate Chinese song teaching. Therefore, when the first-level vocabulary accounts for more than 62%, it means that other types of vocabulary are lower, which is more suitable for elementary Chinese learners.

Second, for intermediate-level Chinese learners, they already have a particular vocabulary and master some commonly used Chinese syntactic structures and grammatical functions. Therefore, in the database, they can not only learn songs of intermediate difficulty but also harness the teachers' Guide to learn.

Third, for learners with a Chinese major or a higher level of Chinese, who have taken Chinese as their major. Then based on mastering commonly used Chinese vocabulary, they need to improve their Chinese proficiency further and continuously expand their vocabulary. Therefore, Chinese songs at this level can help them achieve these goals, and at the same time, they can understand the meaning of the song and the cultural content contained in the lyrics.

Fourth, in the Chinese song ranking database, the content of the first-level vocabulary in some song lyrics is shallow, and the second-level, the third-level and extended-level vocabulary account for more, which makes the lyrics of such songs obscure and challenging to learn. Therefore, it is recommended that such songs be used as interest-expansion songs and used as expansion songs after Chinese as a foreign language, in order to increase the interest of learners and increase the enthusiasm for Chinese songs.

V. RESEARCH PROSPECTS

From an application point of view, first, the database can help teachers reduce the pressure of lesson preparation,

improve teaching effects, and help students increase their enthusiasm and interest in Chinese learning. Second, the research of the database mainly solves the problem of determining the difficulty level of song lyrics text while filling the gaps of previous research. Besides, it has strong innovation and educational reference value. Third, the construction of the database is in line with the current trend of corpus big data and reflects the innovative research concept of interdisciplinary and fusion. The application of data mining, data analysis, natural language processing and other related technologies in the research process plays a vital role in achieving correct personalization and autonomous learning and is also an essential reflection of the frontier research in this discipline.

ACKNOWLEDGMENTS

The authors wish to thank Beijing International Studies University. The paper is supported by the graduate scientific research program of Beijing International Studies University.

REFERENCES

- [1] Chinese Proficiency Grade Standards and Grade Outline. (1988). Edited by the Research Group of Chinese Proficiency Grade Standards of the Chinese Society of Teaching Chinese as a Foreign Language. Beijing: Beijing Language and Culture University Press.
- [2] Chinese vocabulary classification of syllables for international education in China. (2010). China National Foreign Language Teaching Leading Group Office and other editors. Beijing: Beijing Language and Culture University Press.
- [3] Liu, S. Li, M. Kang S. (2011). Research and Implementation of Teaching Information Database for Chinese New Words for Foreigners. *Application of Language and Characters* (01):129134. doi: 10.16499/j.cnki.1003-5397.2011.01.019.
- [4] Liu, X. (2011). Exploring the construction of Chinese character generation system. Southwest University, p79-83
- [5] Shi, Z. (2011). Interpretation of the teaching mode of words, morphemes and Chinese characters. Confucius Institute at Sapporo University, World Chinese Rhetoric Society, International Rhetoric Communication Society. 2011 International Forum of Rhetoric and Communication Studies—Language and Cultural Education and Intercultural Communication Proceedings. Confucius Institute at Sapporo University, World Chinese Rhetoric Society, International Rhetoric Communication Society: Global Rhetoric Society: 138.
- [6] Sun, K. (2013). Research on Teaching Chinese as a Foreign Language with National Music as the Carrier. *Modern Chinese (Language Research Edition)* (07):102-105.
- [7] Wang, Y. (2017). Research and practice of Chinese songs assisting Korean and Chinese teaching. Shandong Normal University. p16-23
- [8] Xing, H. (2005). Statistical analysis of Chinese character components in "(Chinese Proficiency) Chinese Character Grade Outline". *World Chinese Language Teaching* (02):49-55+118. doi: 10.13724/j.cnki.ctiw.2005.02.006.
- [9] Zhang, B. (2005). Some thoughts on the syllabus of Chinese as a foreign language. World Chinese Language Teaching Society. Selected papers of the 8th International Conference on Chinese Language Teaching. World Chinese Language Teaching Society: World Chinese Language Teaching Society: 28-38.
- [10] Zhang, J. (2007). Learning to sing folk songs and learning Chinese. *Overseas Chinese Education* (01):35-41. doi: 10.14095/j.cnki.oce.2007.01.008.
- [11] Zhao, F. (1998). Introducing Chinese folk songs into teaching Chinese as a foreign language. *Journal of Beijing Second Foreign Studies University* (02):55-62.
- [12] Zhao, S. & Luo Q. (1994). Chinese Class Introducing Singing. *Chinese Learning* (04):47-51.
- [13] Zhao, Y. (1980). A set of standard letters (English). *Dialect* (02): 81-83.
- [14] Zheng, Y. (2013). Perspective of the application of corpus technology in Chinese teaching. *Application of Language and Characters* (01):131-138. DOI: 10.16499/j.cnki.1003-5397.2013.01.026.

Yuqi Zhang was born in Mudanjiang City in 1996. She graduated from Harbin Normal University in 2018 and graduated with a master's degree from the Chinese Academy of Beijing Second Foreign Language College in 2020, majoring in Chinese international education. As the principal person in charge, she undertook the critical scientific research project of graduate students in the school and was awarded the titles of excellent graduates of the Chinese Language Institute of Beijing Second Foreign Studies University and excellent graduates of Beijing.

Xiujuan Han is an associate professor and a professor from Beijing Second Foreign Language College. From 1996 to 2000, she studied at Yantai Teachers College with a bachelor's degree. From 2000 to 2007, she studied at Beijing Language and Culture University with a master's degree and a doctoral degree. From 2007 to present, she has worked in the Chinese Language School of Beijing Second Foreign Language College.

Qingping Hu is a female lecturer. From December 1998 to July 2000, she taught at East China Normal University. From July 2000 to present, she has taught at Beijing Second Foreign Language College. From September 2007 to July 2009, she studied in the postgraduate class of linguistics and applied linguistics of Beijing Normal University. In July 2014, she participated in the Hanban Hosted the "Advanced Workshop on the Sustainable Development of Culture and International Chinese Education" and the "Advanced Workshop on Chinese International Education" sponsored by Beijing Language and Culture University.

Fei Song is a PhD, associate professor and master tutor of the Central University of Nationalities and Columbia University who mainly engaged in research on language intelligence, international Chinese teaching measurement research, and cultural and technological integration. Currently, he presides over many scientific research projects such as the Beijing Social Science Fund Project. As a key member, he participates in the National Social Science/National Self-Study Fund.