

# *The Use of Mapping Knowledge Domains in Analyzing Subject's Research Overview and Hot Spots\**

Liu Yan; Tang Yingjuan  
College of Humanities & Law  
North China Univ. of Tech.  
Beijing, China  
jljaly@163.com

**Abstract**—This paper uses scientometrics method of knowledge mapping domains, on the basis of CAJD database, summarizes and analyses the discipline structure of “Teaching Chinese as a Second Language (TCSL)”. It introduces the method of knowledge mapping analysis to the research area of Teaching Chinese as a Second Language. It shows how to present the research overview and hot spots of a subject intuitively and visually by analyzing keywords frequency and co-words.

**Keywords**— Teaching Chinese as a Second Language (TCSL); Knowledge Mapping Domains; Visualization; Research Overview; Hot Spot

## I. INTRODUCTION

Scientific mapping knowledge domains (hereinafter referred to as Mapping Knowledge Domains) are emerging fields of scientometrics. It is a picture which shows the development process of scientific knowledge and construction relationship. It describes human beings' knowledge resources overtime and its carrier by using visualization technology. It draws, searches deeply, analyzes and presents scientific technology knowledge as well as their mutual connection. It creates an environment of knowledge sharing in the organization to promote the cooperation of science and technology research and in-depth. In 2003, Prof. Liu Zeyuan brought it to China and established a professional research team of mapping knowledge domains. At present, many domestic scholars begin to apply research method of mapping knowledge domains to analyze and research on discipline structure, frontier and development. The related research is becoming more and more. However, the research objects are mainly on science and engineering. It has been applied quite few in humanities and social science and has not been used in the TCSL area. Research method of mapping knowledge domain is based on big data and can provide a powerful tool for TCSL to analyze and research on literature systematically.

As “China Popularity” is increasing in the world, the international promotion of Chinese is developing vigorously.

The depth and breadth of TCSL study is also expanding. TCSL discipline theory construction began in the early 1980s. In 1998, “Teaching Chinese as a foreign language” as a formal branch was listed in the secondary discipline of “linguistics and applied linguistics” by the degree committee of the state council, and for the first time to set up “Teaching Chinese as a foreign language” as the main concentration of “linguistics and applied linguistics” Doctoral (master) program. This suggested that “Teaching Chinese as a foreign language” as a discipline began to enter a new stage of development.

This paper, as a discipline of “Teaching Chinese as a foreign language” as the research object, by using modern scientometrics method, draw the discipline knowledge mapping, describe and interpret the subject structure of the discipline, the research hotspot and frontier problems.

## II. THE RESOURCE AND PROCESSING OF DATA

This study uses China Academic Journal Network Publishing Database (CAJD) as the basic literature database and sets up research literature database of “Teaching Chinese as Foreign Language”. It searches all the Chinese literature which are 23808 papers in total on “Teaching Chinese as a foreign language” by using CAJD database. By eliminating the non-research papers manually, there are 11196 valid papers. It forms a research literature database of “Teaching Chinese as a foreign language” by obtaining basic information like titles, authors, units, summaries, keywords, references from these Chinese and English literature and standardizing the data. This is the foundation of making mapping knowledge domain next step. This paper mainly analyzes key words, uses modern scientific metrology method comprehensively, present research field, subject structure, hot issue of “Teaching Chinese as a foreign language” discipline.

We use Bibexcel software to do the first round of word frequency analysis and get the keywords list. Before the analysis of key words, we standardize the keywords firstly which include several steps, such as stop words list filtering, key words' similarity and so on.

---

Supported by Beijing Higher Education Young Elite Teacher Project  
YETP1432

### III. RESEARCH HOT SPOTS AND OVERVIEW ANALYSIS OF TEACHING CHINESE AS A SECOND LANGUAGE

#### A. The result of keyword frequency analysis

This study used CAJD as basic literature database. It analyzed the keywords from 11195 valid articles by Bibexcel software and got the frequency table of keywords. It can roughly describe research hot spots in recent decades of teaching Chinese as a foreign language by selecting the keywords which have the highest frequency.

TABLE I. TOP20 KEYWORDS IN FREQUENCY

| No. | Keywords              | Frequency |
|-----|-----------------------|-----------|
| 1   | Error                 | 710       |
| 2   | Chinese Character     | 684       |
| 3   | Textbook              | 626       |
| 4   | Vocabulary            | 536       |
| 5   | Culture               | 536       |
| 6   | Teacher               | 516       |
| 7   | Acquisition           | 389       |
| 8   | Grammar               | 373       |
| 9   | Spoken language       | 325       |
| 10  | Pronunciation         | 270       |
| 11  | Teaching method       | 245       |
| 12  | classroom teaching    | 220       |
| 13  | Inter-language        | 211       |
| 14  | Communication Ability | 210       |
| 15  | Chinese education     | 193       |
| 16  | HSK                   | 186       |
| 17  | Comparison            | 184       |
| 18  | Pragmatics            | 180       |
| 19  | compilation           | 177       |
| 20  | Culture Teaching      | 173       |

From the statistics of keyword frequency, we find out the hot spot keywords in the field of TCSL in 40 years. These hot spots' keywords reflect the research hot spot s' fields. There are very close connections among some keywords which reflect the relevant field of study. After the analysis of these keywords, we can get the following hot research fields. These hot spots' areas are the second language acquisition, Chinese character, textbook, vocabulary, culture, grammar, speaking, teaching method, communication, HSK, and international promotion Chinese language and so on. In addition, these keywords refer to listening, reading and writing in classified teaching. Therefore, classified teaching of Teaching Chinese as a foreign language is the teaching research hotspot. Scholars advocate teaching by levels and teaching by countries which are also reflected. The research is mainly on primary level and Japanese students.

#### B. The result of keyword co-occurrence analysis

The co-occurrence relationship of keywords reflects some keywords which often appear together. However, some keywords of high frequency don't have co-occurrence relationship of high frequency. Some keywords which usually co-occur reflect fixed research fields, research paradigm, and also the subject structure in these research filed. Top 20 keyword co-occurrence frequency analysis was showed by table 2.

TABLE II. TOP 10 KEYWORDS IN CO-OCCURRENCE

| No. | Co-occurrence term 1 | Co-occurrence term 2     | Frequency |
|-----|----------------------|--------------------------|-----------|
| 1   | Compilation          | Textbook                 | 149       |
| 2   | Culture              | Culture Teaching         | 59        |
| 3   | Teacher              | Cultivation              | 54        |
| 4   | Textbook             | Teaching                 | 50        |
| 5   | Teacher              | Classroom Teaching       | 49        |
| 6   | Error                | Acquisition              | 48        |
| 7   | Error                | Inter-language           | 46        |
| 8   | Error                | Pronunciation            | 45        |
| 9   | Teacher              | Student                  | 41        |
| 10  | Chinese Character    | Pictophonetic Characters | 40        |

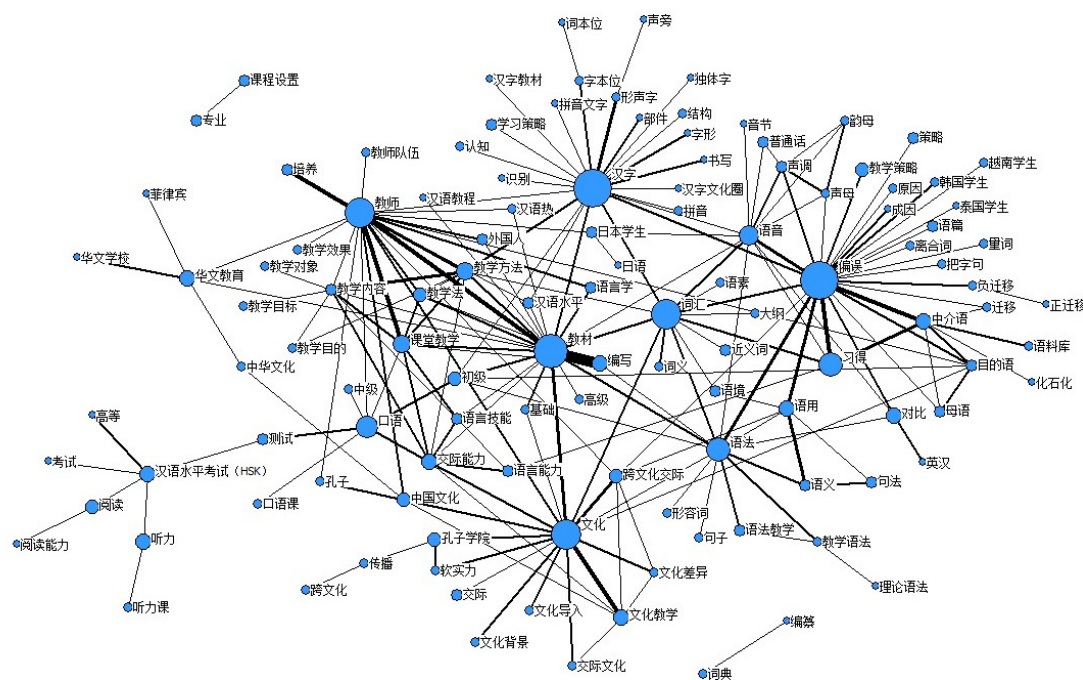
According to table 2, we can find out some comparatively mature research field, such as error, acquisition, inter-language which usually occur together and have close relationship. In the study of errors, the most researches are on pronunciation, pragmatics, and grammar which become hot areas of research. This analysis method can make us understand the research hot spot and blank, find out the keywords in the subject that have close relationship, understand the research paradigm, and clarify the concept.

Only from the co-occurrence relationship, we can't tell the subject structure. With the help of keyword frequency and number of co-occurrence, we can draw the knowledge mapping domains of TCSL discipline structure.

It extracted the keywords from 11195 Chinese literature, standardized the keywords, analyzed keyword frequency statistically, calculate the frequency of keywords of co-occurrence (appear in the same article is as co-occurrence once), choose high frequency keywords for the construction of high frequency word co-occurrence matrix, and then draw the co-occurrence map of key words by visualization software which is shown in the above picture (to make the may clear, the picture choose to present the keywords whose co-occurrence frequency is more than 10).

In the picture, the size of the circle represents the frequency of keywords. The bigger of the circle is, the higher frequency of the keywords is. The line shows the co-occurrence frequency of keywords. The line is thicker, the co-occurrence frequency is bigger. It reflects that the keywords that are connected by the line have stronger connection. We can find out that the thick lines areas are mostly in the center of the map.

Fig. 1. Map that shows Co-occurrence Frequency is more than 10



### C. Overview analysis of TCFL research hot spots

The biggest circles in figure 1 represent the hot issues in the development of this discipline. There are a lot of research's areas formed around these hot issues which include Chinese character, error, pronunciation, vocabulary, pragmatic, grammar, textbook, teacher, culture and so on. The research areas which are shown by knowledge mapping domains have both similarities and differences from the description by previous scholars.

Subject overview of the existing research are based on logical clues and form the subject framework according to pronunciation, vocabulary, grammar and other language elements to summarize or according to listening, speaking, reading and writing skills to summarize. Based on knowledge mapping domains analysis of big data, this paper tries to describe the research map of teaching Chinese as foreign language.

With the combination of modern linguistics study achievement, we divide these research areas into fourteen major groups. Each group is research area of related contents. These groups reflect study overview of teaching Chinese as a foreign language over 30 years.

## IV. CONCLUSION

The analysis of the mapping knowledge domains completed the mission that was impossible in the past. Based on huge amounts of data, it analyzed the discipline's overview. Although scientometrics has become a mature discipline and knowledge mapping domains is a mature method, we use mapping knowledge domains for the first time to study the research scope and hot spots of teaching Chinese as a foreign

language. It meets the goal and verifies some classic literature review. Through keyword frequency and co-occurrence analysis and knowledge mapping domains, it researches on overview and hot spots of teaching Chinese as a foreign language which is intuitive and clear. Mapping Knowledge domains reflect the existing research. We can find the blank area of subject research based on the logic analysis of subject structure.

This research method can be applied more. For example, each group can continue to have in-depth analysis on how Chinese character teaching, inter-language, errors analysis change with time and what their status are. From the analysis, we can present and clarify the discipline overview. This paper is an attempt to use mapping knowledge domains to analyze TCFL teaching. And we need to do more work.

## References :

- [1] Liu Xun, Introduction to Teaching Chinese as a Second Language Education. Beijing:Beijing Language and Culture University Press,2000.
- [2] Zhu Zhiping, "TCFL as a Branch of Applied Linguistics", Journal of Beijing Normal University(Humanities and Social Science Version), vol. A162, pp. 105-109, Dcember 2000.
- [3] Cui Yonghua, "Review of Research Hotspots of Teaching Chinese as a Second Language in Twenty Years", Applied Linguistics, pp. 63-70, February 2005.
- [4] Sun Dejin, "More than 10 years of Teaching Chinese as a Second Language Research Overview---'Index of Teaching Chinese as a Second Langue Works' Afterwords", Language Teaching and Research, pp. 45-53, April 2009.
- [5] Chen Yue, Liu Zeyuan, "Emerging Scientific Mapping Knowledge Domains", Studies in Science of Science, A23, pp. 149-154, February 2005.
- [6] Katy Borner, Chen C., Kevin W. Boyack Visualizing knowledge domains. Annual Review of Information Science & Technology, A37, pp179-255, 2003.