Journal of Social Science Humanities and Literature, Vol.7, Issue 5, (Oct)
ISSN 2277-842X **2024**
DOI: 10.53469/jsshl.2024.07(05).17

Advances in Fuzzy Linguistics Research in the Last Two Decades - CiteSpace-based Visualisation and Analysis

Xiaoli Cheng

Jiangxi Normal University, Nanchang, Jiangxi, China

Abstract: ***This paper analyzes the research progress of fuzzy linguistics in the last two decades based on the visual analysis method of CiteSpace. It introduces the concept and development background of fuzzy linguistics. In this paper, the research progress in the field of fuzzy linguistics is explored in depth through visual analysis using CiteSpace, and the related research directions and development trends are discussed and summarized in detail. The results of these analyses are of certain reference significance for further in-depth research in the field of fuzzy linguistics and for advancing the technological development and application in the field of fuzzy linguistics.***

Keywords: Fuzzy linguistics; Fuzzy language; Fuzziness; Citespace.

1. RESEARCH BACKGROUND AND SIGNIFICANCE

The study of linguistic ambiguity has a long history both in China and abroad. In ancient times, Chinese literary and philosophical scholars noticed this phenomenon. Qu Yuan once wrote in his "Bu Suan Zi": "Sometimes an inch is long, and a foot is short." In ancient Greece, the Megalain school, represented by the philosopher Eubulides, first noticed the phenomenon of ambiguity, known as the "sorites paradox" - although they did not directly propose the concept of ambiguity (Wu Tieping, 1999).

For about 2000 years thereafter, the ambiguity of natural language was severely neglected. It wasn't until 1965 that L. A. Zadeh proposed the theory of fuzziness, which changed the situation. His paper titled "Fuzzy Sets," published in "Information and Control," established the formal status of the theory and then radiated to many different research fields (Sperber, D. & Wilson, 2001). It is widely recognized that Wu Tieping's 1979 article "Basic Methods of Fuzzy Language," published in the fourth issue of "Foreign Languages," was the earliest paper on ambiguity in China. In 1999, Professor Wu Tieping published China's first monograph on fuzzy linguistics, "Fuzzy Linguistics," which is considered the cornerstone of Chinese fuzzy linguistics.

The birth of fuzzy linguistics is an inevitable result of the development of linguistics. In addition to its own characteristics, fuzzy linguistics also has all the features of linguistics and fuzzy theory. Therefore, it has been considered a branch of linguistics from the very beginning. As a new and relatively independent discipline, fuzzy linguistics is generally interpreted as the scientific study of ambiguity in natural language. Since ambiguity is an attribute of natural language, fuzzy linguistics should first answer basic questions about ambiguity, namely, what is ambiguity. Secondly, since precision and ambiguity are inherent attributes of all natural languages, fuzzy linguistics must explore their dialectical relationship, study the causes and functions of ambiguity. Thirdly, since ambiguity is an innate attribute of natural language, it can be reflected in various dimensions of language, such as phonetic ambiguity, semantic ambiguity, pragmatic ambiguity, and rhetorical ambiguity. Lastly, it is also necessary to study the relationship between fuzzy linguistics and other disciplines, such as translation studies, rhetoric, aesthetics, etc. All of these constitute the research scope of fuzzy linguistics. Language, as a fundamental and crucial means of communication, is self-evidently important to study.

CiteSpace is a visualization analysis software based on the Java platform, which can be used in various fields such as literature analysis, knowledge discovery, and scientific evaluation. The general steps for using CiteSpace include: downloading and organizing the literature to be analyzed into a file format that meets the data input requirements of CiteSpace, typically a .txt or .xml file. Then, input the parameters to be used, including the input file path, the calculation method for node relationships, the setting of node size and color, the setting of the time window, the layout algorithm, etc. Network analysis: After completing the parameter settings, network analysis and visualization can be carried out. CiteSpace will analyze the input files to generate literature relationship networks, keyword co-occurrence networks, etc. Based on the parameter settings, CiteSpace can generate various

charts and visualization interfaces, such as funnel charts, timelines, 3D views, clustering maps, etc., to facilitate user analysis and reasoning. CiteSpace provides a more intuitive and clear perspective, allowing one to see the development of fuzzy linguistics in recent years, as well as the missing aspects, to explore the current state of research and research hotspots, to speculate on the development trends of research, and to provide reference for future research in this field.

2. CITESPACE ANALYSIS METHOD

2.1 Literature Data Acquisition

English literature data is retrieved through the Web of Science (WOS) Core Collection platform, using "Fuzzy linguistics," "Vague Linguistics," and "Fuzzy Language" as the search terms, with document types limited to Articles and Reviews (original works, reviews), and language type restricted to English. The search period is from the inception of the database to May 22, 2023. English literature data is exported in plain text format through WOS and deduplicated and selected through CiteSpace software; Chinese literature is searched on China National Knowledge Infrastructure (CNKI), using "模糊语言学" "模糊语言" "语言模糊性" and "语言模糊" as search terms or keywords, with the search period being from the inception of the database to May 22, 2023. The selection is made by reading the titles and abstracts of the literature, and the data is exported in Refworks format, with each record mainly including information such as the author of the literature, the publishing institution, abstract, and keywords.

2.2 Inclusion and Exclusion Criteria

2.2.1 Inclusion criteria:

(1) The content of the literature is related to fuzzy linguistics;

(2) Documents with repeated content or multiple submissions are counted as one.

2.2.2 Exclusion criteria:

(1) Documents that only mention search terms like fuzzy linguistics but do not elaborate in detail;

(2) Conference notices, activity records, experience exchanges, or call for papers.

2.3 Analysis Method

The number of Chinese and English literature published from the inception of the database to May 22, 2023, is statistically analyzed, and the trend is plotted using Excel. CiteSpace 5.8.R3 software is used for visual analysis with institutions, authors, countries, and keywords as analysis nodes. Before visualizing the CNKI literature, data transformation processing is carried out with CiteSpace software, with the time span set from 1993 to 2023, the time slice set to 1 year, and the threshold selected as the system's default parameter (Top 50). Parameters are adjusted based on the number of selected nodes, and the results view is cropped using the pathfinding method to ensure readability. Keyword mutation maps are generated for both Chinese and English literature included in the study to track research hotspots at different stages.

3. RESEARCH CONTENT AND VISUAL ANALYSIS

A total of 76 English documents were retrieved from WOS, and 500 Chinese documents were retrieved from CNKI. After screening out documents that did not meet the criteria, a total of 492 documents were included. There were 55 valid English documents, including 7 reviews; there were 492 valid Chinese documents, including 13 master's theses.

3.1 Publication Volume

The earliest English document on fuzzy linguistics in WOS appeared in 1983, while the first Chinese document was found in 1985 according to CNKI. Overall, the publication volume of Chinese documents on fuzzy linguistics

shows a wave-like growth trend, with English documents showing a larger increase from 2012 to 2013, and from 2016 to 2021, reaching a peak (11 documents) in 2021. The publication volume of Chinese documents on fuzzy linguistics shows a wave-like growth trend, with English documents showing a larger increase from 2006 to 2008, and from 2010 to 2013, reaching a peak (34 documents) in 2008.

Figure 1: WOS Fuzzy Linguistics Chinese Publication Volume

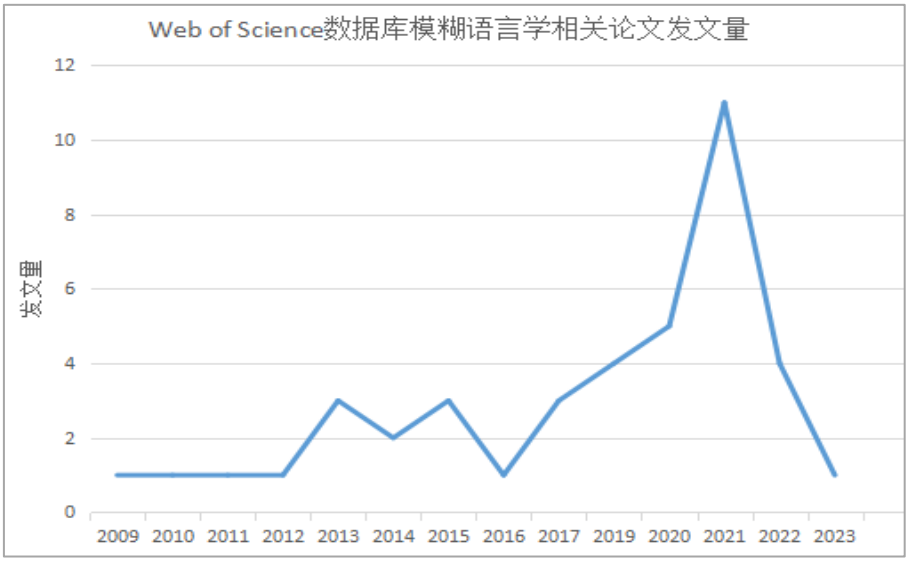

Figure 2: CNKI Fuzzy Linguistics English Publication Volume

3.2 Publishing Institutions

In the knowledge map of publishing institutions, each node represents one institution, the font size represents the number of publications by the institution, and the lines between nodes represent the cooperative relationships between institutions. The thicker the line, the stronger the cooperation between the two.

3.2.1 WOS

The publishing institutions of English literature in the WOS database are shown in Figure 2. The top three institutions by publication volume are RLUK - Research Libraries (5 papers), Systems Research Institute of the Polish Academy (4 papers), and University of Alberta (3 papers). Institutions with named identifiers have published more than 2 papers each, and the knowledge map includes 44 nodes and 26 lines.

Figure 3: Knowledge Map of Institutions Associated with English Literature in WOS

3.2.2 CNKI

The knowledge map of the institutions associated with the Chinese literature retrieved from CNKI is shown in Figure 3. The results include a total of 2 nodes and 3 lines. Institutions that have published one or more papers are identified in the graph. There are 218 institutions that have published literature related to fuzzy linguistics in Chinese. The top four institutions in China by publication volume are Heilongjiang University (13 papers), Anhui University (9 papers), Guangxi Normal University (9 papers), and Sichuan Normal University (8 papers).

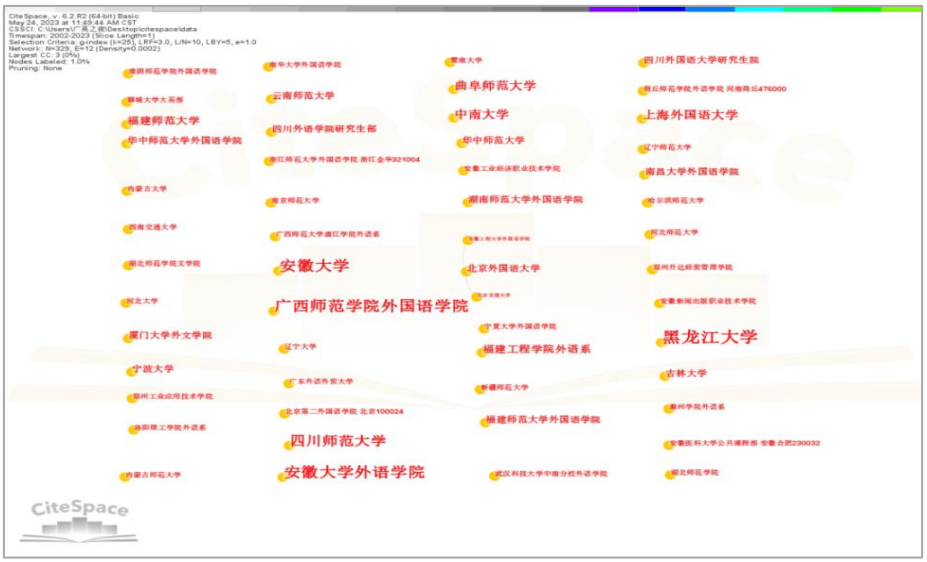

Figure 4: Knowledge Map of Institutions Associated with Chinese Literature in CNKI

4. PUBLISHING AUTHORS AND COUNTRIES

In the knowledge map, each node represents an author or a country, and the lines between them represent the cooperative relationships between authors or countries. The thicker the line, the closer the cooperation between the two.

4.1 WOS

The co-occurrence distribution of authors in the English literature on fuzzy linguistics in WOS is shown in Figure 4. Authors who have published more than 5 papers are identified in the graph, with Pedrycz, Witold (4 papers) and Proell Simon (3 papers) ranking in the top 2 positions. The graph includes a total of 24 nodes and 6 lines, with a network density of 0.68.

4.2 CNKI

The author with the highest number of publications in the Chinese literature on fuzzy linguistics retrieved from CNKI is Xianghua Weng (8 papers), followed by Jinshu Li (5 papers), with all others having 2 papers.

Figure 5: Co-occurrence Distribution of Authors in WOS Literature

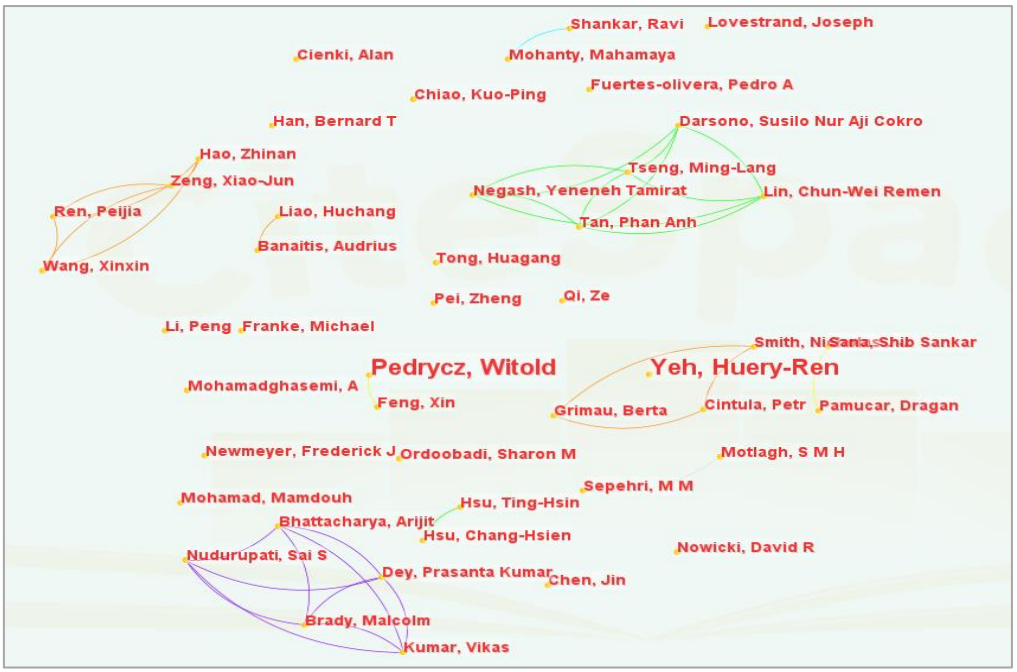

Figure 6: Co-occurrence Distribution of English Authors in WOS

5. KEYWORD CO-OCCURRENCE ANALYSIS

In the keyword clustering knowledge map, each node represents a keyword, and the lines represent the co-occurrence of two keywords. The thickness of the line indicates the frequency of their co-occurrence, and a centrality greater than 0.1 is considered a key node. The keyword emergence function can be used to analyze the research hotspots in the field during different periods.

5.1 WOS

The keyword co-occurrence knowledge map of the English literature data from WOS is shown in Figure 7. Additionally, it can be seen from the figure that there are a total of 12 clusters. After analyzing the keyword groups within the clusters, the top three are decision-making, models, and fuzzy models, indicating that these three are the research hotspots in the field. After removing the initial search terms and counting the keywords in the English literature, the top 10 high-frequency keywords in foreign literature are listed in Table 1. According to the table, the high-frequency keywords in English literature on fuzzy linguistics include "corpus linguistics," "cognitive linguistics," "decision making," "fuzzy sets," "group decision making," "term sets," "consensus model," "aggregation operators," "operators," and "management." Among them, "decision making" is a key node. The research frontier hotspot mutation map of the English literature on fuzzy linguistics is shown in Figure 8. From Figure 8, it can be seen that the latest research hotspot in English literature on fuzzy linguistics is in the field of management.

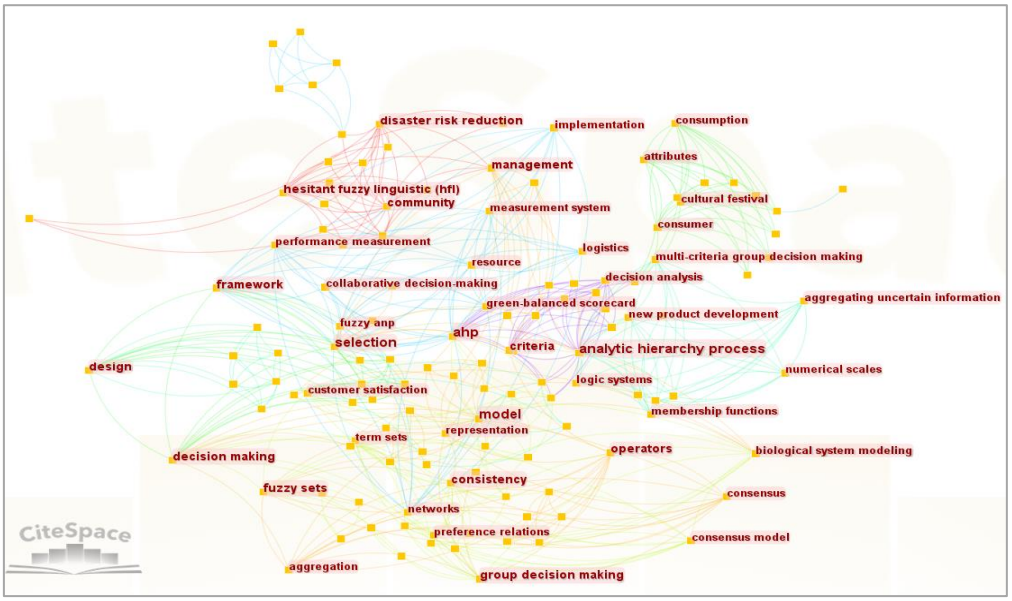

Figure 7: Keyword Co-occurrence Distribution in WOS

Keywords	Year	Strength	Begin	End	2009 - 2023
corpus linguistics	2011	1.29	2011	2012	
cognitive linguistics	2011	1.2	2011	2014	
decision making	2017	2.86	2019	2023	
fuzzy sets	2020	1.61	2020	2023	
group decision making	2019	1.14	2019	2023	
term sets	2019	0.96	2019	2020	
consensus model	2019	0.96	2019	2020	
aggregation operators	2019	0.96	2019	2020	
operators	2021	1.34	2021	2023	
management	2021	1.14	2021	2023	

Figure 8: Keyword Mutation Map of Research Hotspots in English Literature in WOS

5.2 CNKI

The keyword clustering knowledge map of the Chinese literature data retrieved from CNKI is shown in Figure 9. After performing cluster analysis and marking the keywords, 7 clusters were obtained. The modularity index Q value of the cluster view is 0.712, and the silhouette index S value is 0.892, indicating that the structure of each cluster is significant and highly convincing. Analyzing the keyword groups after clustering reveals that the top three search terms are "fuzzy," "fuzzy language," and "pragmatic function," indicating that these three keywords are the research hotspots in this field in China. The top 10 high-frequency keywords in Chinese literature are "fuzzy language," "semantic ambiguity," "pragmatic function," "euphemism," "category," "fuzzy words," "fuzzy rhetoric," "categorization," "cooperative principle," and "literary translation," as seen in Table [Table number]. Among them, "fuzzy language," "fuzzy language," and "pragmatic function" are key nodes. The research hotspot mutation map of Chinese literature on fuzzy linguistics is shown in Figure 10. The results show that the first research hotspot to emerge in this field in China is "translation studies." This indicates that fuzzy linguistics initially attracted the attention of domestic scholars as a translation method and was studied accordingly. "Review" as a research hotspot from 2018 to 2020 indicates that domestic scholars tend to use literature reviews as a vehicle to summarize domestic and international research.

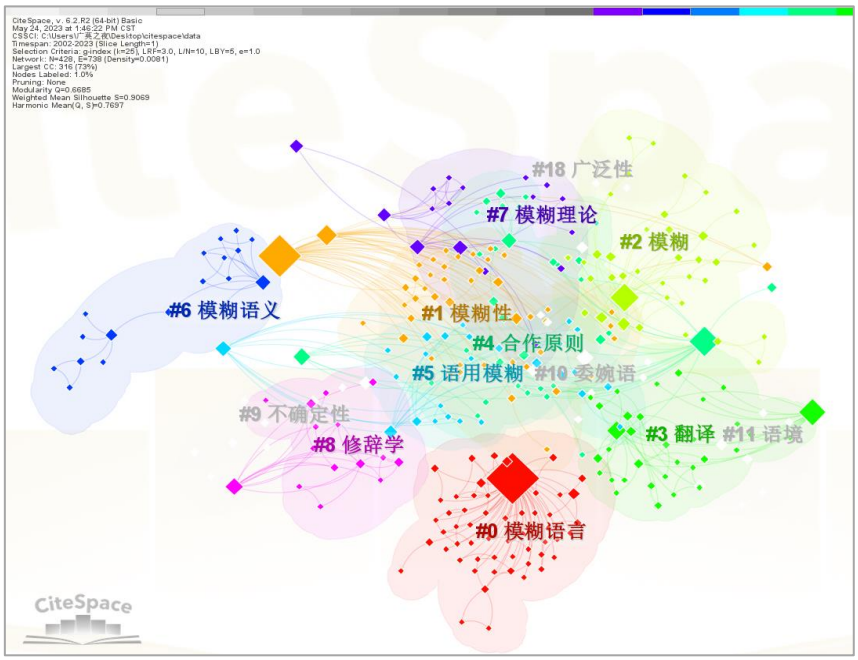

Figure 9: Keyword Co-occurrence Distribution in CNKI

Keywords	Year	Strength	Begin	End	2002 - 2023
修辞学	2002	1.69	2002	2007	
认知	2002	1.6	2002	2003	
语言	2003	1.7	2003	2005	
模糊语义	2003	1.62	2003	2005	
翻译	2002	2.42	2007	2009	
语义模糊	2008	3.08	2008	2009	
广告语	2009	1.92	2009	2011	
英语教学	2012	2.51	2012	2021	
语用模糊	2005	1.99	2012	2014	
模糊理论	2006	1.92	2012	2016	
模糊语	2004	2.78	2013	2018	
口译策略	2013	1.57	2013	2015	
翻译策略	2007	2.35	2014	2015	
翻译教学	2019	1.89	2019	2021	

Figure 10: Keyword Mutation Map of Research Hotspots in Chinese Literature in CNKI

6. DISCUSSION AND SUMMARY OF RESULTS

6.1 Analysis of Research Status Characteristics

From the data statistics, the current state of fuzzy linguistics research has three major characteristics: First, there is a significant gap between domestic and international research, with the main research forces concentrated in American scientific institutions. As can be seen from the publication volume (Figure 1), English literature first appeared in 1983, while the first Chinese literature in the form of a research review appeared in 2003. Looking at the distribution of national research strengths, the United States has a total publication volume of up to 2134 papers, far exceeding China and other countries, and has close exchanges with other countries. Moreover, the top three contributing institutions in foreign research also come from the United States, with frequent cooperation among institutions and a concentrated research force. In contrast, domestic research forces are more dispersed, and there is a lack of cooperation among various research institutions. Second, both domestic and international research is developing towards quantitative research. Third, the field of fuzzy linguistics is developing rapidly and has a promising research prospect. The publication volume of Chinese literature in fuzzy linguistics has been increasing year by year since 2015, reaching a peak in 2017, then slightly declining but has remained stable; English literature has also maintained a trend of increasing year by year, with a single-year publication volume as high as 57 papers in 2019. Overall, the publication volume in the field of expressive writing shows a continuous increasing trend. Fuzzy linguistics has a philosophical perspective on the essential issues of language research.

The theoretical research of fuzzy linguistics requires the support of disciplines such as philosophy, cognitive linguistics, logic, and philosophy of language. The applied research of fuzzy linguistics also needs to be combined with disciplines such as mathematics, electronics, engineering, and mathematical logic. The comparative study of multiple languages is also a relatively weak link at present. Studying language in isolation is of little significance and limited value. It is important to focus on conducting quantitative research. Zadeh's fuzzy sets have made the quantitative research of category semantic ambiguity possible, but statistical data indicate that the quantitative research of fuzzy language in China is still relatively weak. Quantitative research will promote the qualitative research of category semantic ambiguity to develop in depth.

6.2 Analysis of Research Hotspots and Trends

Keywords are the core and theme of a literature, and high-frequency keywords can reflect the hot topics in the field. The keyword mutation knowledge map can detect the intensity and changes of research hotspots in the field over time. In terms of research hotspots, first: theoretical research on fuzzy language. This is one of the main research hotspots in fuzzy linguistics, mainly focusing on the study of fuzzy language characteristics, quantification of fuzzy language, and fuzzy language logic. Second: applied research on fuzzy language. With the continuous development of information processing technology and artificial intelligence technology, the application of fuzzy linguistics in fields such as finance, risk assessment, and natural language processing is becoming more and more extensive. Third: applied research on fuzzy mathematics. Fuzzy linguistics is an important branch of fuzzy mathematics, so research on processing methods and application technologies for fuzzy language based on fuzzy mathematics is also one of the current research hotspots. In terms of research trends,

Fuzzy language processing methods based on machine learning. Machine learning has become an important technology in natural language processing and is widely used in fields such as natural language understanding and knowledge graph construction. In the future, fuzzy language processing technology based on machine learning will become a research focus and be widely applied in various fields. First, formal theoretical research on fuzzy language. With the development of fuzzy linguistics, the formal theory of fuzzy language is also increasingly receiving attention. In the future, we can expect to see more refined and rigorous research results in fuzzy language theory. Second, interdisciplinary research between fuzzy language and cognitive science. Along with fuzzy language is the continuous progress of cognitive science. In the future, the interdisciplinary research between these two fields will further deepen our understanding of the generation and comprehension of fuzzy language and provide more comprehensive and in-depth support for the application of fuzzy language. In general, with the continuous improvement of information technology, the in-depth research and widespread application of fuzzy linguistics will be the trend in the future, and various problems and challenges will also emerge, requiring continuous deepening of theory and expansion of applications to meet future needs and challenges.

7. CONCLUSION

By conducting a visual analysis of the literature related to fuzzy linguistics in the WOS and CNKI databases using CiteSpace software, an objective analysis of the research status and hotspots in the field of fuzzy linguistics has been conducted, and the future research trends have been explored and predicted. In summary, fuzzy linguistics has a broad research prospect. However, there is a significant gap between domestic and international research. There is a lack of prolific authors in China, and no long-term scientific research system has been formed, which is not conducive to improving the depth and breadth of research. Moreover, the distribution of research institutions in China is relatively scattered, and there is a lack of cooperation, which is not conducive to pooling the scientific research strength between institutions. Therefore, it is necessary to establish relevant long-term scientific research support policies in the future to encourage authors to continue to track and explore, and to strengthen the exchange and cooperation between research institutions, further enhancing China's research level in this field. Future research will mainly focus on exploring the different application effects of expressive writing in different research populations and the empirical and quantitative research of different theoretical mechanisms.

REFERENCES

[1] Sperber, D. & Wilson, D. Relevance: Communication and Cognition [M]. Beijing Foreign Language Teaching and Research Press, 2001.
[2] Gutt, E. A. Translation and Relevance: Cognition and Context [M]. Shanghai Foreign Language Education Press, 2004.
[3] Chen Guoquan. Fuzzy Expression in Knowledge Engineering [M]. Science Press, 1989.
[4] Li Xiaoming. Fuzziness: The Enigma of Human Cognition [M]. The People's Publishing House, 1985.
[5] Miao Dongsheng. Introduction to Fuzzy Studies [M]. China Renmin University Press, 1987.
[6] Li Yingjie. A Review of Domestic Research on Fuzzy Linguistics in the Past Forty Years (1979-2019) [J]. Journal of Changchun Normal University, 2021, 40(07): 68-73.
[7] Fu Weifeng, Yi Zhihong. Fuzzy Language Information Integration and Decision Analysis [J]. Systems Science and Mathematics, 2020, 40(08): 1447-1455.
[8] Su Xuelian. A Brief Analysis of the Communicative Function of Fuzzy Language and Translation [J]. Language Construction, 2017(18): 65-66.
[9] Qi Yuanyuan. Exploration of Semantic Fuzziness and Cognition [J]. Journal of Harbin Normal University (Social Sciences), 2016, 7(03): 102-104.
[10] Yi Xinqi. A Review of Domestic Research on Fuzzy Language [J]. Theory Study, 2013, No. 654(12): 183-184.
[11] Liu He. A Study on the Development of Fuzzy Linguistics Abroad [J]. The Economist, 2013, No. 288(02): 57-58.