

# A visualized study of literature review on analysis of language and nationalism based on CiteSpace (2015-2024)

Yunfeng LIU

Shandong University of Finance and Economics, Jinan 250300, China

---

**Abstract:** Using CiteSpace as a scientometrics method, this paper makes a visual analysis of 500 Web of Science papers on international language and nationalism in the past 10 years (2015-2024). The number of published papers, authors and institutions are formed into a knowledge network; the key words, research hotspots and emergent words are sorted out; the future research trend is predicted, and the hottest research field--language policy is analyzed and discussed from the perspective of linguistic economics. The study found that, language policy, identity, national culture and other related studies are still hot topics. There are still some problems in the field of language and nationalism that are worthy of further exploration. For example, most of the relevant research results on language and nationalism focus on international politics and other disciplines, and there are relatively few cross-cooperation studies between different disciplines. Western scholars have done more research, but Chinese scholars have done less localization research.

**Key words:** language and nationalism; Citespace; language policy; linguistic economics

---

## 1 Introduction

Language is not only a tool for daily communication, but also a vehicle of culture and an important symbol of national identity, as well as an important content in national political demands. The same is true of the Marxist theory of the nation, and Stalin's famous first definition of the nation is "having a common language". There is a complex and close relationship between language and nationalism [1]. In order to understand and analyze the current situation and development trend of language and nationalism studies, this paper uses CiteSpace, an analytical tool of scientometrics, to outline the knowledge map of language and nationalism studies in the past 10 years (2015-2024), analyze research hotspots, and visually present the above contents [2].

## 2 Research method and data sources

### 2.1 Method

CiteSpace, the tool used in this study, is a second generation information visualization technology, which is multi-component, time-sharing and dynamic. First, the online analysis platform was used to conduct a preliminary analysis of the number of posts and other content in the field of language and nationalism. These data were then analyzed statistically and visually by Microsoft Excel software. After that, the selected domestic and foreign literature is exported in Plain Text format, and saved as "txt" format with "download\_01" as the filename suffix. Finally, CiteSpace software is used to build

the mapping knowledge of author, institution and keywords domains.

## 2.2 Data sources

Using Web of Science (WOS) database as the English literature source, the data retrieval period was from 2015 to 2024, and the keywords were language and nationalism. The less relevant papers were deleted and the remaining 500 papers were used as English data sources [3].

## 3 Results and analysis

### 3.1 Author and cooperation network analysis

#### 3.1.1 Author distribution

After setting the node category as author, run CiteSpace and set threshold to 1, that is, the author of one published paper is shown in the result. We can adjust the node size to get the author cooperation network diagram.

The result shows the number of nodes  $N = 244$  in the upper left corner. In the figure, it can be seen that the distance between nodes is relatively long, the distribution is uniform and dispersed, and only a few line segments are connected between authors. Density represents the correlation strength of each node, shown in the figure. The density is 0.0011, indicating that the cooperation between researchers in the field of language and nationalism is loose, and no effective cooperative relationship has been formed. In addition, small-scale cooperation also exists, such as Barrenechea, Ignacio, Beech, Jason; you can also see the collaboration of Echavarren, Jose MColler, Xavier, Alvarez-galvez, Javier, but this collaboration is not close or frequent, and it does not form a network of cooperation [4].

#### 3.1.2 Distribution of research institutions

After the node is set as an institution, the distribution map of research institutions can be obtained after the software is operated, which lists the research institutions with a number of publications greater than or equal to 3.

There are 9 institutions that have published more than 4 papers from 2016 to 2021, and most of them are internationally renowned institutions. Only the University of Hong Kong in China has done a lot of research in the field of language and nationalism, while mainland institutions have done little research. This shows that Chinese scholars need to carry out deep cultivation in this field.

### 3.2 Language and nationalism research topics and research hotspots

#### 3.2.1 Keyword co-occurrence and clustering

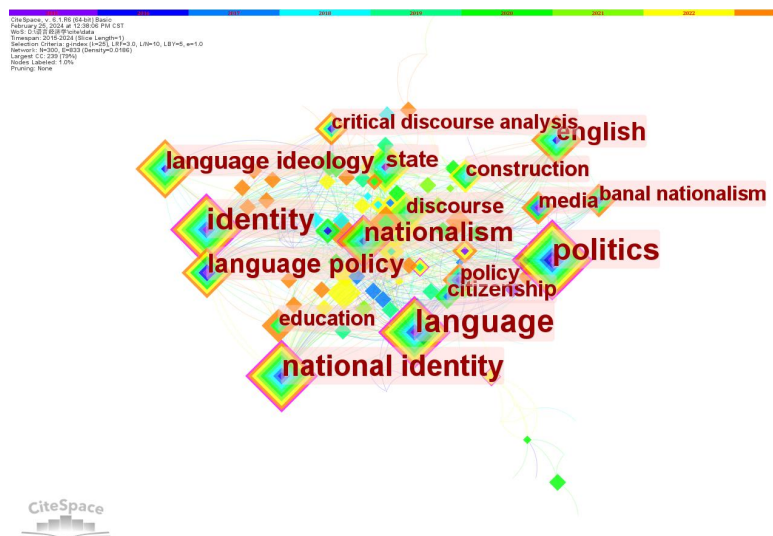


Figure 1. The keyword co-occurrence map

Keywords reflect the research topics. We import the keyword data into the software, set the nodes as keywords, adjust

the display of keywords with a frequency of more than 10, and draw the keyword co-occurrence map according to 500 papers screened in the knowledge network (see Figure 1).

Figure 1 shows that there are 300 keywords in the upper left corner, and these keywords are connected by 833 connections, with dense nodes and complex connections. Node size represents the frequency of keyword occurrence, and node size proves this topic as the research focus. In the figure, a network is formed around 10 key words: language, policy, nationalism, language policy, identity, etc.

On the basis of Figure 1, LLR algorithm is selected to cluster according to keywords. The software will cluster nodes that meet the rules, and set the display cluster range, as shown in Figure 2. If the Q value is within the interval of [0,1] and  $Q > 0.3$ , then obvious clustering can occur; if the S value is above 0.5, clustering is reasonable. The module value  $Q = 0.7303$ , the average contour value  $S = 0.9121$ , and the values of Q and S in the figure are all within a reasonable range [5], which proves that the clustering formed in Figure 2 is effective.

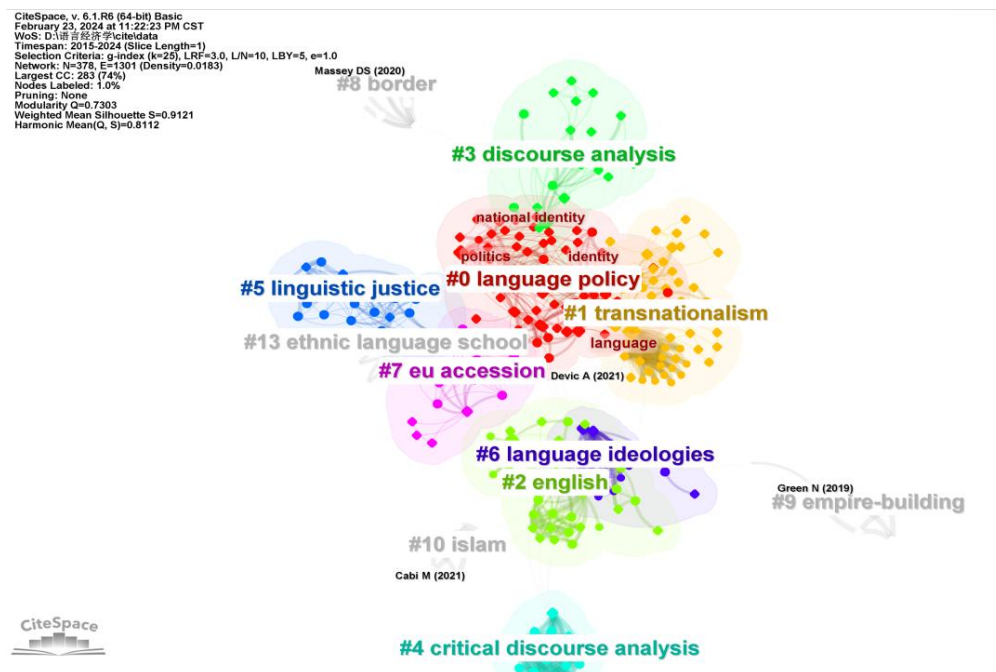


Figure 2. The display of cluster range

Figure 2 presents 14 cluster labels, each of which forms a cluster. The density of distributed nodes and connections in the cluster represents whether the research results of the cluster are rich. It can be seen from the figure that nodes are densely distributed in the clusters of language policy, trans-nationalism and discourse analysis. There are a large number of studies under these three clusters, and they are the focus of current research.

### 3.2.2 High-frequency keywords

Make a keyword emergence map, in which you can see the emergence changes of the top 18 keywords over time, reflecting the concerns of researchers and academic circles. From the result, we can find that "national identity, diversity, border and rise" have become research hotspots in this field in recent years, indicating that the number of people studying such issues has increased. As can be seen from the figure, the red line of the keyword "rise" continues until 2024, indicating that the research heat of this topic will continue.

## 4 Discussion

Through the above research, we can find that "language policy" has always been a hotspot in the study of language

and nationalism. Language policy is also closely related to national economic development [6]. Language is not only communicative, but also capital [7]. In a sense, language is a part of individual worker's human capital, and more precisely, language ability is a kind of human capital. Because language meets the three elements of human capital: its acquisition needs cost; it can bring benefits to people; it is attached to the human body. As human capital, language ability is not fundamentally different from other skills acquired in education and training [8]. To sum up, the relationship between language policy and economy is very close. A country needs to formulate a reasonable language policy, only in this way can it promote national development, reduce social conflicts and improve people's culture.

## 5 Conclusion

Through the visual analysis and review of key literature in the research field of language and nationalism in the past decade, it can be found that the research field has gradually deepened and expanded, and considerable research results have been achieved, but there are still some problems worthy of further in-depth discussion. Most of the research results related to language and nationalism are concentrated in the field of international politics and other disciplines, and there are relatively few cross-cooperation studies among different disciplines, which may cause the application scope of language and nationalism to be limited to some extent. Future research can further strengthen the cooperation between researchers and research institutions in different disciplines to better promote the in-depth development of the research in terms of theory and application. For example, scholars can discuss the impact of language policy on national economy and the impact of nationalism on national economy from the perspective of linguistic economics.

## Conflicts of interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

- [1] Bonikowski B, Luo Y, Stuhler O. 2022. Politics as usual? Measuring populism, nationalism, and authoritarianism in U.S. presidential campaigns (1952-2020) with neural language models. *Sociological Methods & Research*, 51(4): 1721-1787.
- [2] Ho JC. 2023. Hong Kong anti-colonial nationalism during the Chinese language campaign. *The China Quarterly* (Published online), 1-16.
- [3] Laitin DD. 2015. Language conflict and violence: The straw that strengthens the camel's back. *European Journal of Sociology/ Archives Européennes de Sociologie*, 41: 97-137. (Reprint of "Language conflict and violence: The straw that strengthens the camel's back". *International Conflict Resolution after The Cold War*. Edited by P.C. Stern and DDrukman. Washington D.C.: National Academy Press, 1999.)
- [4] Law A, Mooney G. 2012. Competitive nationalism: state, class, and the forms of capital in devolved Scotland. *Environment and Planning C-government and Policy*, 30 (1): 62-77.
- [5] McManus C. 2016. Irish language education and the national ideal: the dynamics of nationalism in Northern Ireland. *Nations and Nationalism*, 22(1): 42-62.
- [6] Migué L. 1970. Le nationalisme, l'unité nationale et la théorie économique del information. *Revue Canadienne d'Économique*, 3: 183-198.
- [7] He QX, Zhong L, Liu XM. 2022. Visualization map analysis of semen cassiae research progress based on CiteSpace. *World Science and Technology - Modernization of Chinese Medicine*, 24(1): 398-407.
- [8] Weng J. 2020. End of an era: transforming language and society in Japan, Korea, and Vietnam, c. 1870-1950. *European Journal of Sociology*, 61(2):269-299.