

## ТЕОРЕТИЧЕСКИЕ ПРОБЛЕМЫ СОЦИОЛИНГВИСТИКИ

## THEORETICAL ASPECTS OF SOCIOLINGUISTICS

УДК 81 272

DOI: 10.37892/2713-2951-2020-3-3-10-29

THEORIES AND PRACTICES OF CHINA'S LANGUAGE RESOURCES<sup>1</sup>

Li Yuming

Beijing Language and Culture University, People's Republic of China

*Forty years have passed since the concept of language resources was proposed. In the 1980s, Australia formulated the National Language Policy based on this concept. Since the beginning of the 21st century, China has started to put this concept into practice. National Language Resources Monitoring and Research Center was established, and projects such as the Chinese Language Audio Database Resources and the National Project of Chinese Language Resource Preservation were carried out. China has successfully held International Conference on Role of Linguistic Diversity in Building a Global Community with Shared Future in cooperation with the UNESCO and has adopted the Yuelu Proclamation. China has been outstanding in advocating the concept of language resources and implementing language protection. Research on China's language resources is carried out on the basis of the practices of language planning. The focus of research has moved from the nature of language as resources and their classification to the understanding of the functions of language resources. This article offers a detailed historical survey of the theories and practices of China's language resources. It classifies language resources into three categories: oral, written, and derivative resources, and analyzes three functional domains of language resources: language reservation, language information processing, and language learning. In addition, this article proposes the view of language knowledge and regards language as not only a semiotic system, but also a bank for*

<sup>1</sup> This paper was read at the first "World Conference on Language Resources Protection" (世界语言资源保护大会) (Changsha, September 19–20, 2018) under the title "Theories and Practices of China's Language Resources Protection". The relevant content was published in the CPPCC Daily (10th edition, January 14, 2019) with the title "Concept of China's Language Resource". Previously, this paper was read in the title of "The Needs of The Times for Language Resources" (《语言资源的时代需求》), "Language Resources of the Intelligent Age" (《智能时代的语言资源问题》), "Digitalization of Language Resources and Localization of World Knowledge" (《语言资源数据化及世界知识中文化》) respectively at School of Education, University of Hong Kong (December 27, 2017), the Workshop of "One Belt, One Road" Language Resources and Evaluation (B&R LRE) of the "11th International Language Resources and Evaluation Conference" (LREC 2018) (Miyazaki, Japan, May 7, 2018), the launching ceremony and expert review meeting of "A Multilingual Dictionary Database of Chinese and Foreign Languages" project of Foreign Language Teaching and Research Press (January 13, 2018), the "2nd International Symposium on Information and Intelligence of Language and the 11th Forum of Young Language Scholars in Shanghai (July 8, 2018, Fudan University), the "Founding Meeting of the Experts' Committee of the "One Belt, One Road" Ethnic Culture Big Data Center"(March 30, 2018, Minzu University of China), and the "Frontier Lecture on Linguistics for Postgraduates" (June 14, 2018). Thanks to Guo Xi (郭熙), Li Zhijiang (李志江), Wang Lining (王莉宁), Yang Erhong (杨尔弘), Sun Le (孙乐), Rao Gaoqi (饶高琦) and Liang Jingtao (梁京涛) for providing relevant data for this paper.

*storing human language knowledge system and cultural system.*

**Keywords:** *language planning, language resources, types of language resources, functions of language resources, the view of language knowledge*

## СОХРАНЕНИЕ ЯЗЫКОВЫХ РЕСУРСОВ КИТАЯ: ТЕОРИЯ И ПРАКТИКА

Ли Юймин

Пекинский университет языка и культуры, Китайская Народная Республика

*Концепция языковых ресурсов насчитывает уже более сорока лет с момента своего возникновения. На этой концепции основана сформулированная в 1980-х годах национальная языковая политика Австралии. С начала XXI века эта концепция начала применяться в Китае. Был создан Национальный центр мониторинга и исследований языковых ресурсов, а также осуществлены такие проекты, как «Ресурсы голосовой базы данных китайского языка» и «Национальный проект сохранения ресурсов китайского языка». Китай успешно провел Международную конференцию, посвященную роли языкового разнообразия в построении глобального сообщества с общим будущим в сотрудничестве с ЮНЕСКО и принял «Декларацию Юэлу». Китай добился выдающихся успехов в пропаганде концепции языковых ресурсов и защиты языков. Исследование языковых ресурсов Китая проводится на основе практики языкового планирования. Фокус исследования переместился от признания ресурсной природы языка и его классификации к пониманию функций языковых ресурсов. В данной статье предлагается подробный исторический обзор теорий и практик использования языковых ресурсов Китая. В ней языковые ресурсы классифицируются по трем категориям: устные, письменные и производные, а также анализируются три функциональные области языковых ресурсов: сохранение языков, обработка языковой информации и изучение языка. Кроме того, в данной статье предлагается взгляд на языковое знание и рассматривается язык не только как семиотическая система, но и как банк для хранения языковых и культурных знаний человека в системном виде.*

**Ключевые слова:** *языковое планирование, языковые ресурсы, типы языковых ресурсов, функции языковых ресурсов, представление о языковых знаниях*

There are over 100 languages in China. An immense number of works written in Chinese has accumulated for thousands of years. China is a country with rich language resources.<sup>2</sup> At the

---

<sup>2</sup> How many languages there are in China is still an academic issue to be discussed. Zhou Qingsheng [2015: 43–50] discussed “landscape of ethnic minority languages”, arguing that there are 70 to 80 languages in China, and pointed out in Note 2 that “there are more than 70 according to the data released by the government; more than 80 according to *The Encyclopedia of China*; some domestic experts and scholars think that there are more than 100 kinds, and some foreign scholars think that there are more than 200.” In fact, scholars at home and abroad now think that there are more than 300. For example, Huang Xing [2018: 113] thinks that: “China is one of the countries with the richest linguistic diversity in the world. There are more than 300 languages belonging to the Oriental Sino-Tibetan language family, the Western Indo-European language family, the northern Altaic language family, and the Southern-Asian language family and South-Island language family in the south. The geographical and family distribution of languages in Chinese is probably unique in the world.

same time, language database is an important foundation of language information processing. The development of language information processing in China has been aiming to catch up with the world's most advanced levels. Some technologies and their applications have pioneered in the world, which urges Chinese scholars to understand language resources from the perspective of the information age. Throughout history, China has dedicated great efforts to the protection and development of language resources. For example, the "Xiping Stone Scripture" ("熹平石经") of the Xiping period of Emperor Ling of Han Dynasty and the "Three Body Stone Classics" ("三体石经") of the Zhengshi period of the reign of the King Qi of Wei in the Three Kingdoms period are models of protecting and displaying the scriptures and characters by erecting steles. Rhyme books and dictionaries of the past dynasties are more direct products of language resources protection. It has been a Chinese tradition to pay great attention to written language, and a good job has been done in respecting and protecting the classics and characters; however, the oral language was relatively ignored. The record protection of oral resources in the past dynasties was not done well, despite the existence of *Dialects* written by Yang Xiong (扬雄) in the Han Dynasty and many words and proverbs in local chronicles. There are many lessons and experiences worth summing up and inheriting in the awareness, protection, and utilization of language resources in the history of China.

However, it has been only since the last decade or so that we realized the significance of language resources rationally, took language resources as an important concept of language planning, and took actions to protect language resources at the national level and across the country. As a social resource, language has only been recognized and valued by the society in recent years.

## 1. Language is resource

It is not easy to understand the nature of language resources. Language does not exert influence directly on human material life as coal, oil, electricity, or waterpower does. This is also in line with the law of human cognition of resource.

### 1.1. Cognitive path of resources: from natural resources to social resources

Resources are the materials on which human production and life depend. Whether a certain kind of material can become a socially recognized resource depends on two preconditions:

First, the "Usefulness" of resources.

With the development of science and technology, materials that could not be used in the past and could not be used in this or that way can be used for production and life nowadays. For example, electricity and oil have existed since ancient times. However, electricity and oil only became resources for humankind when science and technology reached a certain stage of development. In this way, resources are closely related to the progress of science and technology. We should not only fully understand and make full use of the "available resources" under the current state of modern science and technology, but also pay attention to the "potential resources" that may be used in the future with the development of science and technology, especially the resource benefits that may be produced by the frontier science and technology, so as to make the resource construction predictable.

Second, the recognition of the "Usefulness" of resources.

Some materials are indispensable for production and life, but for some reason, people do not realize or ignore their nature as resources. For example, when the haze is not that severe, people cannot recognize fresh air as a resource; when the pollution is not serious, people cannot

recognize clean water as a resource. In this way, resources include “recognized resources” and “unrecognized resources”. We should pay special attention to those “unrecognized resources”, including “available resources” that have been used in our production and life and “potential resources” that may enter our production and life in the future.

Whether the “Usefulness” of resources can be recognized can be seen from four aspects: first, whether the society has a scientific view of resources and whether resource science as a discipline is developed; second, the level of research on the relationship between a certain resource and human beings; third, the sensitivity to science and technology of the resource; fourth, scarcity of the resource. The most easily perceived resources are scarce resources. Useful and scarce resources are bound to have high prices. Therefore, “scarcity” has almost become an additional attribute of resources.

There are two types of resources: natural resources and social resources. In the human consciousness of resources, natural resources are first recognized, and then social resources. *Contemporary Chinese Dictionary* (《现代汉语词典》) is a dictionary reflecting the “common sense of citizens”. From the 1st edition to the 6th edition, the explanation of the term “resource” has always emphasized that it is the “natural source” of means of production and means of living. The examples cited involve four kinds: “underground resources”, “waterpower resources”, “human resources” and “tourism resources”. The interpretation in the 7th edition in 2016 has made great progress:

*[Resource] As noun: Resources are the sources of means of production or means of living, including natural resources and social resources: underground~; waterpower~; tourism~; human ~; information~ [Contemporary Chinese Dictionary, 2016: 1732].*

The interpretation in the 7th Edition no longer emphasizes the “naturalness” of resources. It clearly divides resources into natural resources and social resources, and adds “information resources” to the examples, which reflects the new understanding of resources in the information age.

*Cihai* (《辞海》) is an encyclopedic reference book. From the 1st edition to the 4th edition, “resource” was interpreted as “the source of wealth, generally refers to natural financial resources”, which is obviously narrower than the definition in *Contemporary Chinese Dictionary*. By the 5th edition in 1999, *Cihai* added a new meaning to “resource”:

*The general term of material resources, financial resources, human resources and other material elements owned by a country or a certain region. It is divided into natural resources and social resources. The former includes sunlight, air, water, land, forest, grassland, animals and mineral resources; the latter includes human resources, information resources and material wealth created by labor [Cihai, 1999: 3881].*

This understanding can be considered accurate, as it not only no longer emphasized the “naturalness” of resources and divided resources into natural resources and social resources, it also considered social resources to include “material wealth created by labor”. Of course, if we look further, social resources include not only “material wealth created by labor”, but also “spiritual wealth created by labor”.

## 1.2. The establishment of language resource awareness

Looking back at the definitions of resources in *Contemporary Chinese Dictionary* and *Cihai*, we can find that they do not mention “language resources” in their examples. It is unlikely that “language resources” was not listed because of limited number of examples, but rather that the society and lexicographers at that time did not realize the nature of language as resources.



The understanding of language resources by humankind is indeed relatively recent. Language planning scholars from abroad have also mentioned the issue of language resources, but the research was not systematic. Fishman (1973) began to mention “language resources”: “Language is different from agriculture, industry, labor, waterpower and other resources... It is obvious that language is a kind of resource only in the sense of its value. In any case, language is a special resource, which is difficult to manage with the existing cost-benefit theory. The reason is that it is hard to measure language and separate it from other resources. However, there is still ample reason to explore the similarities and differences between language and other resources in resource planning.” (see Zhou Qingsheng, 2001: 422–423).

Wang Hui (2007) pointed out that Ruiz put forward three orientations affecting language planning in 1984: language-as-problem, language-as-right, and language-as-resource. Language resource orientation can alleviate the conflict between the first two orientations and help reestablish people’s attitude towards language and language communities. Language is a resource that needs to be managed, developed and protected. Bilingual and multilingual abilities are language resources. Languages of ethnic minorities are special and important resources. Kaplan (1997) and Grin (2003) regard language as an important “human resource” and an important “human capital”, and language planning should be an aspect of national planning.

In the practice of national language planning, Australia advocated the policy of cultural pluralism from the 1970s to the end of 1980s. On the basis of the “four reports” including *Grassby Report* (1973), *Galbally Report* (1978), the *Toward a National Language Policy* (1982), and *A National Language Policy* (1984), the *National Policy on Languages* was published in 1987. The *National Policy on Languages* proposed four strategies to guide Australian Language Policy: first, protect Australian language resources; second, develop and expand these language resources; third, integrate language teaching and language usage initiatives in Australia with national economic, social and cultural policies; fourth, provide information and services in a language that customers can understand. This is one of the few cases in which the concept of language resources has been transformed into a national language policy. However, it is a pity that in the 1990s, with the publication of the *Green Book* (1990)<sup>3</sup> and the *White Book* (1991)<sup>4</sup>, language policies in Australia had evolved to give priority to English literacy, which has relatively weakened the proportion of language resources in language policy.<sup>5</sup>

The concept of “language resources” was first used in China in the 1980s. Qiu Zhipu (1981) discussed the development of language resources from the perspectives of informatization, language teaching and language planning, especially in the fields of “language project” and the global spread of Chinese. In 1988, the Chinese journal of *Construction of Languages and Characters* (《语文建设》) published the *Australian National Language Policy* by Lo Bianco, who is a famous Australian expert in language planning and a drafter of *National Policy on Languages* (1987). The article of Lo Bianco (1988: 55–57) published by the journal is just an abstract of the *National Policy on Languages* (1987), but its introduction of the concept of language resources in Australia and related measures was relatively detailed, such as: “but there are still many people whose mother tongue is other than English. This is a very valuable language resource, and it is of great significance to make full use of these language resources,”

<sup>3</sup> *The Language of Australia: Discussion Paper on an Australian Literacy and Language Policy for the 1990s* (《澳大利亚的语言:20 世纪 90 年代澳大利亚读写能力与语言政策讨论》).

<sup>4</sup> *Australia’s Language: The Australian Language and Literacy Policy* (《澳大利亚的语言:澳大利亚语言与读写能力政策》).

<sup>5</sup> For details of Australian language policy, see Wang Hui (2010).

“the main objective of the national language policy is to maximize the benefits of Australia’s access to its rich language resources,” “all in all, only careful planning can allow for the most use of Australian language resources,” “language resource planning requires the cooperation and coordination of... government at all levels.”

In the 1990s, the concept of language resources proposed by Qiu Zhipu and Lo Bianco did not seem to have any repercussions in the circle of linguists. Only Tao Yuanke (1996) proposed that attention should be paid to the development and utilization of the social language resources in Macao. Qiu Zhipu (2000) once again stated his viewpoint in 1981, and Zhang Zhengbiao (2000) cited the value of western dialects in language research. However, according to Wang Shikai (2009: 25–28), Chinese literary circles were discussing how to explore and utilize language resources (old vernacular, dialect, folk language, etc.) to do a good job in writing and the relationship between western language resources and Chinese literary creation. In China, although language and literature fall under the same discipline of “Language and Literature” and there are two sub-disciplines, “Chinese Language and Literature” and “Foreign Language and Literature”, there is little academic interaction between language and literature circles, as if “the sound of chickens and dogs can be heard, but people never come into contact” (鸡犬之声相闻，老死不相往来). On the issue of language resources, there is no dialogue between the language and literature circles in China to this day.

In the 21st century, the discussion of language resources is gradually increasing, and the awareness of language resources has begun to be established in China. Around 2004, Professor Zhang Pu often discussed the issue of language resources with colleagues from the Bureau of the Management of Language, Characters and Information of the Ministry of Education (教育部语言文字信息管理司), such as Li Yuming and Wang Tiekun. It was well recognized that language is a resource; language resource is the most important resource in the information society, and it is a national resource alongside of mineral resources, land resources, marine resources, water resources, and forest resources. The state should legislate on the management of language resources to monitor, protect, develop and utilize them. The academic achievements of these discussions are reflected in Zhang Pu’s important paper *On National Language Resources* (2007). The practical achievement is that the State Language Commission has set up the “National Center for Language Resources Monitoring and Research”, which was officially launched in June 2004. The Center gradually established the China Language Resources Development and Application Center (The Commercial Press) and 6 branch centers, including print media (Beijing Language and Culture University), audio media (Communication University of China), network media (Central China Normal University), languages of ethnic minorities (Central University for Nationalities), educational and teaching materials (Xiamen University), and overseas Chinese (Jinan University).

In July 2005, the Ministry of Education and the State Language Commission held the “Conference on the Construction and Informatization of National Language Standards” in Urumqi. Yuan Guiren, then director of the State Language Commission, made the written speech “Establishing a Scientific Outlook of Development and Initiating a New Situation on the Construction of Ethnic Language Standards and Information Work”, in which he proposed to “protect cultural diversity and develop ethnic language resources”, and held that “language resources are important information resources and cultural resources”. The establishment of the National Center for Language Resources Monitoring and Research and the speech made by Yuan indicate that China’s awareness of language resources has gradually become clear and the

concept of language resources has been established in national language planning.<sup>6</sup>

## 2. Practical activities on language resources in China

The concept of language resources and its academic development in China are driven by the practice of language planning. Understanding the practical activities of language resources in China in the 21st century can help us better understand China's language planning and the research on language resources.

### 2.1. National Center for Language Resources Monitoring and Research

The mission of the National Center for Language Resources Monitoring and Research is completed by its many branch centers. According to the common concept and technical specifications, each branch center collects and operates corpora of “print media, audio media, network media, ethnic minority languages, educational and teaching materials and overseas Chinese”. Collecting and maintaining these corpora is the preservation of language resources. Through the corpora, the Center can also find out characteristics of language use, new language phenomena and some important “real” statistical data, such as the frequency of use of Chinese characters and vocabularies, new words, catchwords, language situation on the Internet, etc. Many data are often released to the public through the press conference of the Ministry of Education and “The Inventory of Chinese”(汉语盘点) activities, and become important chapters in the annual *Report on the language Situation in China* (Green Book) (《中国语言生活状况报告》绿皮书). These data have played a certain role in the development of the *General Standard Chinese Character Table* (《通用规范汉字表》), the formulation of Chinese curriculum standards in primary and secondary schools, the relevant standards of Chinese international education, and the compilation of Chinese teaching materials.<sup>7</sup>

In September 2007, the Bureau of the Management of Language, Characters and Information, and Beijing Language and Culture University jointly hosted the Summit Forum on “National Language Resources and Applied Linguistics”, which is also an academic conference where domestic and foreign scholars were invited by several branch centers of the Center for Language Resources Monitoring and Research. This is the first ever large-scale academic conference on language resources held in China. Besides “Construction, Monitoring and Research of National Language Resources”, the topics discussed in the forum also included: national language resources and language services, language planning, normalization of languages and characters, social application of language and characters, mother tongue teaching, global spread of Chinese, teaching Chinese in ethnic areas, and the relationship between language information processing. The scope of research is relatively wide and 50 papers were received at the conference, 28 of

---

<sup>6</sup> After director Yuan put forward the concept of “language resources” in his speech, the relevant departments of the State Language Commission have constantly affirmed and actively explored the concept of language resources. For example, in May 2006, Li Yuming, director of the Bureau of the Management of Language, Characters and Information of the Ministry of Education, introduced the language situation of China in 2005 at the 11th press conference of the Ministry of Education in 2006, and Wang Tiekun, deputy director of the Bureau, made further elaboration on “language resources” in his speech. In July 2007, at the conference of “Industrialization in Europe and China: Language Contact and Identity” held at Leiden University in the Netherlands, Li Yuming made an academic report entitled *Language Planning in China*, which clearly pointed out that we should “establish the concept of language resources and cherish language resources”. In August 2007, at the press conference of “the language situation of China”, Li Yuming expounded again the view of language resources. For details, please refer to Wang Shikai (2009: 19–22).

<sup>7</sup> Details can be found in the *Report on the language Situation in China* since 2005 and Wang Tiekun (2009).

which were selected to form a collection *On Language Resources in China* (《中国语言资源论丛》) published by The Commercial Press (商务印书馆). It was the first collection of papers on language resources in China, which still exerts academic influence today.

## 2.2. China Audio Database for Language Resources

Around 2006, many scholars believe that to protect national language resources, we need to understand the actual national conditions of language resources. Language census is the best measure to achieve that goal. In 1956, according to the instructions of the State Council, a survey of Chinese and languages of ethnic minorities was conducted. In 1999, the Ministry of Education and other 10 Ministries or Commissions in China jointly carried out a survey on the use of Chinese language and characters. These two surveys have played an important role in understanding the national conditions of the language, but they have not reached the level of “Census”. In particular, they did not collect the corpus and failed to understand the language and dialect discourse. During this period, there were also scholars who especially understood the situation of language census in some foreign countries and began to design the program of language census in China.<sup>8</sup>

After consulting with the relevant departments, the State Language Commission was told that the time for the national language census was not yet ripe, so it took a new approach: a research group was organized to study the construction of the “China Audio Database for Language Resources”. The name of “Audio Database” aims to emphasize oral language, audio data, and database technology; to embed “language resources” in the name is to highlight and spread the concept of language resources; some contents of language census are reflected in the survey items.

The main experts who participated in the study were Cao Zhiyun (曹志耘), Dai Qingxia (戴庆厦), Guo Longsheng (郭龙生), He Rui (何瑞), Huang Xing (黄行), Li Rulong (李如龙), Liu Danqing (刘丹青), Pan Wuyun (潘悟云), Qiao Quansheng (乔全生), Wei Hui (魏晖), Xie Junying (谢俊英), Xu Daming (徐大明), Zhang Zhenxing (张振兴), etc. According to the research results, the Office of the Leading Group for the Construction of China Audio Database for Language Resources (中国语言资源有声数据库建设领导小组办公室) compiled the Part of Chinese Dialects of the *Survey Manual of China Audio Database for Language Resources* (《中国语言资源有声数据库调查手册》). Cao Zhiyun did the actual writing, and Gu Qian (顾黔), Hou Jingyi (侯精一), Liu Lili (刘俐李), Sun Maosong (孙茂松), Wang Ping (汪平), Yang Erhong (杨尔弘), and Zhao Xiaoqun (赵晓群) also participated in the discussion. These research results are reflected in the handbook and also in Li Yuming’s paper *On the Construction of China Audio Database for Language Resources* (2010) 《论中国语言资源有声数据库建设》.

According to the management standards and technical requirements stipulated in the *Survey Manual of China Audio Database for Language Resources*, the State Language Commission officially launched the construction of China Audio Database for Language Resources in 2008. Jiangsu had undertaken the pilot work for the construction of audio database, which took over a year. In 2009, Jiangsu officially started the construction work. After that, Shanghai, Beijing, Liaoning, Guangxi, Shandong, Hebei, Fujian, Hubei and other places also carried out investigation, research and database construction. The construction of China Audio Database for Language Resources has popularized the concept of language resources in academia and society, cultivated a backbone team, formed a set of management regulations and technical

<sup>8</sup> See Li Yuming (2008a).



standards, and accumulated a number of language resource data.

### 2.3. Language Resources Protection Project of China<sup>9</sup>

On the basis of the construction of China Audio Database for Language Resources, in 2015, the Ministry of Education and the State Language Commission issued the *Notice on Launching the Chinese Language Resources Protection Project* (《关于启动中国语言资源保护工程的通知》), which started the largest language protection project in China and in the world (hereinafter referred to as “the Language Protection Project”). By the end of 2018, the project has been fully launched in 34 provinces and autonomous regions of China, involving more than 350 universities and scientific research institutes, involving more than 4500 professionals. A survey was conducted in 1495 survey sites, and the data collected included all Chinese dialects and more than 120 languages.

On the basis of the Part of Chinese Dialects of the *Survey Manual of China Audio Database for Language Resources*, the Language Protection Project has formulated the *Survey Manual of China Language Resources-Dialects of Chinese* (《中国语言资源调查手册·汉语方言》), and the *Survey Manual of China Language Resources-Languages of Ethnic Minorities* (《中国语言资源调查手册·民族语言》) (divided by language family, totaling 8 volumes), designed language survey and processing software such as “language preservation video recorder” and “language preservation annotating software”, etc., and successively released landmark achievements such as “*Selected Collection of Languages and Cultures in China*” (20 volumes), “*Records of Endangered Languages of China*” (30 volumes, to be published by the end of 2019).

In September 2018, the Chinese government and UNESCO jointly held the first “World Conference on the Protection of Language Resources” in Changsha. At the meeting, the UNESCO, governments, representatives from relevant academic institutions, experts and scholars attending the meeting discussed and adopted *Protection and Promotion of Linguistic Diversity of the World: Yuelu Proclamation* (《岳麓宣言》) (Draft). After the meeting, UNESCO widely solicited opinions in accordance with the procedure and formed the final text of the proclamation, which was officially announced on its official website on January 18, 2019. On February 21, 2019, the Ministry of Education of China, the Representative Office of UNESCO in China, the China National Committee of UNESCO, and the State Language Commission of China jointly held a press conference in China to officially issue the *Yuelu Proclamation*. The first “World Conference on the Protection of Language Resources” was held in Changsha because China’s language protection work was recognized by the international community; the formulation of the *Yuelu Proclamation* was supported by the concept and experience of China Audio Database for Language Resources and China Language Protection Project.<sup>10</sup>

---

<sup>9</sup> Thanks to Professor Wang Lining (王莉宁) for providing some data on the Language Protection Project.

<sup>10</sup> In June 2014, the Chinese government and UNESCO jointly held the “World Language Conference” in Suzhou. Officials and scholars from more than 100 countries and regions discussed issues such as language capability and social sustainable development, language capability and educational innovation, language capability and international exchange cooperation, and reached the *Suzhou Consensus* (《苏州共识》), which integrates the research results and practical experience of China language planning circles in recent years. If *Suzhou Consensus* is the first international consensus promoted by language planning achievements of China, then *Yuelu Proclamation* is the second one.

## 2.4. Chinese Linguistic Data Consortium<sup>11</sup>

Large-scale language resources are the basis of computer language information processing. According to the introduction of Xu Bo, Sun Maosong and Jin Guangjin (2003: 218–224), in 2003, with the support of the special project “Chinese Corpus Construction under the “973 Plan”, the Institute of Automation of the Chinese Academy of Sciences, Tsinghua University, Peking University, the Institute of Computing of the Chinese Academy of Sciences, Shanxi University, Shanghai Jiaotong University, the Application Institute of Languages and Characters of the Ministry of Education and various other organizations jointly undertook the initial construction of resources for the “Chinese Linguistic Data Consortium” (CLDC). After that, the Chinese Information Society of China specially set up a “Language Resources Construction and Management Working Committee” that is responsible for the resources construction and database operation of CLDC.

The establishment of CLDC is based on the experience of American “Linguistic Data Consortium” (LDC)<sup>12</sup>. The purpose is to build and share Chinese resources and promote the technical progress of language information processing. There are 36 providers of resources. The establishment and development of CLDC is a practice of gathering Language resources for language information processing and of language resources co-construction and sharing mode in China, reflecting the scientific needs of language informatization.

## 2.5. Beijing Advanced Innovation Center for Language Resources

In October 2015, the “Construction Plan of High-end Innovation Center of Beijing Universities” was officially launched. In May 2016, the Advanced Innovation Center for Language Resources of Beijing Language and Culture University (Beijing Advanced Innovation Center for Language Resources) was established. It is the first research institute of its kind in China, with the purpose of collecting, developing and utilizing the language resources of the world. Its main academic goal is to keep a copy of the world’s languages (more than 7000) in China, so as to help machines understand human languages.

The Center focuses on the collection of Class A and Class B language resources: Class A is language resources for language library and display, mainly natural language resources; Class B is language resources for language intelligence, including various raw corpus, processed and annotated idioms, language knowledge, social common sense, etc. For language resources of Class A, the Center is implementing the “Circle 3” strategy, that is, on the basis of the China Audio Database for Language Resources and the Language Protection Project, after completing the corpus collection of dialect circle (Circle 1) and languages of ethnic minorities circle (Circle 2) in China, the Center is focusing on collecting corpus of cross-border languages and border languages of China (Circle 3). It is expected that, in the near future, 200 main languages in 60 countries of the “One Belt, One Road” Initiative will be collected.

---

<sup>11</sup> Thanks to Sun Le, Yang Erhong and Rao Gaoqi for providing relevant information.

<sup>12</sup> LDC is funded by the Advanced Research Projects Agency and the National Natural Science Foundation of the United States hosted by the University of Pennsylvania. Its purpose is to construct, collect and distribute language resources for research, teaching and development in the field of language information processing. According to Xu Bo, Sun Maosong and Jin Guangjin [2003: 218], in 2003, more than 100 universities, companies and government departments had joined LDC, with 220 kinds of resources in English, German, French, Spanish, Chinese, Japanese, Arabic and other languages, and provided resources to more than 700 organizations. LDC provides a new mechanism for the large-scale construction and wide sharing of language resources, promoting the research and development level of related fields.

In addition to the five major language resource practice activities described above, there are many other construction projects of language resource databases in China. For example, the *Phonetic Database of Modern Chinese Dialects* (《现代汉语方言音库》) presided over by Hou Jingyi at the end of the 20th century contained 40 kinds of modern Chinese dialects pronunciation files<sup>13</sup> and was published by Shanghai Education Press from 1995 to 1999. Qian Nairong (1995) commented that “the establishment of the modern Chinese dialect phonetic database has ended the history of Chinese dialect publications which only described dialect phenomena in written form.” Another example is that Inner Mongolia has also developed language resource databases of Mongolian, Ewenki, Oroqen and Daur languages (蒙古语、鄂温克语、鄂伦春语、达斡尔语). It is difficult to list all examples one by one. It can be said that as of today, China has become the world leader in the promotion the construction of language resources.

### 3. Academic research on language resources in China

The academic research on language resources has been mentioned above in many aspects. Here are some focused discussions on several issues.

#### 3.1. Practice of language planning and research on language resources

CNKI (China National Knowledge Infrastructure) is a very good database of scientific and technological literature. Currently, CNKI is one of the available tools for literature analysis, though there may be defects in the literature retrieval process of CNKI, for example, some literatures may not be included, and some may not be suitable for retrieval due to keyword annotation. In CNKI, using “language resources” as topic and keyword matching retrieval, 403 articles<sup>14</sup> were retrieved by March 31, 2019, involving more than 370 authors.

As shown in Figure 1, the earliest literature appeared in 1981. By 2003, the annual number of papers published was still less than 5. In the past 22 years, only 17 papers were published, which means that it is a place where scholars seldom set foot in. There was a small peak of research in 2004, and the number of published papers reached 7. In 2007, the research heat increased significantly, with 16 papers published annually; this trend continued until 2011, with 26 papers published annually; the second peak area formed from 2007 to 2011; the third peak area appeared from 2015 to 2017, with the peak in 2016 where the annual number of published papers reached 55.

---

<sup>13</sup> These 40 dialects are representative: Beijing, Tianjin, Jinan, Qingdao, Nanjing, Hefei, Zhengzhou, Wuhan, Chengdu, Guiyang, Kunming, Harbin, Xi'an, Yinchuan, Lanzhou, Xining, Urumqi, Taiyuan, Pingyao, Hohhot, Shanghai, Suzhou, Hangzhou, Wenzhou, Shexian, Tunxi, Changsha, Xiangtan, Nanchang Xiamen, Fuzhou, Jianou, Shantou, Haikou, Taipei, Guangzhou, Nanning, Hong Kong, Meixian, Taoyuan. Each sound file contains more than 100 pages (small 32-divided) of texts and about 60-minute long recording on matching tapes.

<sup>14</sup> This retrieval is operated by Liang Jingtao on CNKI. Liang also helped with data analysis and provided some literature for this paper. The results show that the double matching of subject and keyword retrieval is more scientific than that of keyword-only retrieval, and the number of articles obtained may be more precise and less than that of similar reports. At the beginning of this query, 405 literatures were retrieved, but when the search report was generated, the total number of literatures was reduced to 403, and the reason is unknown. It may be that the repeated counting of some two articles is excluded.

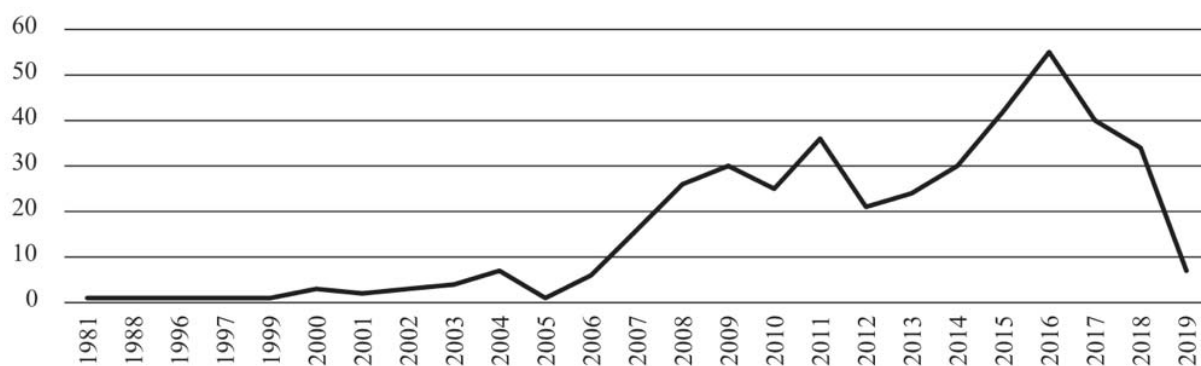


Figure 1. Trend of Literature on Language Resources Research in China

This shows that: first, in the 20th century, the field of “language resources” was almost a virgin land in China, and it was an academic field that only began to be developed in the 21st century. Second, the study of language resources is closely related to the practice of language planning. The small peak in 2004 corresponds to the establishment of the National Center for Language Resources Monitoring and Research; the period between 2007 and 2011 was when the “China Audio Database for Language Resources” brewing and being formal constructed; from 2015 to 2017, the Language Protection Project began to be constructed, and it was also the establishment period of Beijing Advanced Innovation Center for Language Resources. The reason why the peak of research corresponds to the joint point of practical activities is that we always have to do some research before the beginning of practical activities, and then the practical activities can also drive the research.

In China, it is always the practice of language planning that promotes the research of language resources, providing demand, materials and application for research; the research of language resources also provides academic support for practice; the mutual promotion of academic learning and practice is due to the construction of a modern “intellectual revolving door” between government, society and academic circles. 2019 was also an important year for language resources research, due to the first “World Conference on the Protection of Language Resources” held in Changsha in 2018, and the *Yuelu Proclamation* which was officially issued in the spring of 2019.

### 3.2. The nature and types of language resources

The early study of language resources in China mainly demonstrated that language has the nature of resources and established the rationality of language resources. Zhang Pu’s *On National Language Resources* (2007) (《论国家语言资源》) dedicated a significant amount of time and effort to demonstrate that language is a resource. From resources to resource science, and then to natural resources and social resources, the purpose of making such a lengthy detour is to show that language is also a resource, and specifically a social resource. We can see how difficult it was to say that “language is a resource” at that time. Chen Zhangtai’s *On Language Resources* (2008) (《论语言资源》) also used a lot of ink to illustrate that “language is a special kind of social resources”.

After demonstrating that language has the nature of resources, researchers focused on enumerating and classifying language resources. Chen Zhangtai (2008) views language resources from narrow and broad perspectives respectively: in the narrow sense, language



resources refers to “various corpora and language databases used in language information processing, as well as various language dictionaries”; in a broad sense, language resources refer to “language ontology and its social and cultural equivalent value”. What he wants to discuss is the broad sense of language resources, which may explain two issues: first, the concept of “language resources” was more widely used in the academic circle of language information processing at that time; second, the extension of language resources to language ontology was a great innovation.

Zhang Pu (2007) divides language resources into three categories<sup>15</sup>: language resources, speech resources and language learning resources. The classification of language learning resources into one category takes into account the use of language resources. Since scholars have begun to pay attention to language resources, such as Qiu Zhipu (1981), they have put more emphasis on the application of language resources in teaching. Mr. Zhang Pu is also a researcher and advocator for digital teaching, so language learning resources are easier to enter his research field. However, his classification of language resources into “language resources and speech resources” superficially conforms to the “abstract-concrete” thinking habit of the academic circle of “language and speech”, particularly the keen emphasis on language resources generated by the use of language; while in fact, as far as the language database construction is concerned, it is impossible to build a database with “language”, as the text or audio materials in it can only be specific “speech”.

Wang Shikai (2009) wrote the first monograph on language resources in China. While proposing the establishment of the discipline “language resources”, he focused on exploring the various features of language resources. Meanwhile, he also regarded language resources as a system composed of low-level resources (phonetic resources, lexical resources, grammatical resources, semantic resources and text resources) and high-level resources (rhetoric, register and style). This is basically a division of language resources based on the structural elements, while taking into account the text and rhetoric, register, style and so on.

According to Wei Hui (2015) language resources include four categories: (1) language ontology (knowledge) resources, which are composed of phonetics, vocabulary, grammar and semantics; (2) language application resources, including all kinds of general, specific, static, dynamic, multilingual and parallel corpora, as well as knowledge bases, databases and standards related to corpus processing; (3) language learning resources; (4) human resources, that is, talents who master different languages (including foreign languages). Wei Hui (2016) reiterated this view and once again stressed that “human resources are the most essential language resources and also the most dynamic language resources”.

Obviously, so far there is no consensus on the identification and classification of language resources in the academic community, and there is neither a standard for the identification of language resources nor a more appropriate classification system. This paper holds that the basic attribute of language resources is its “usefulness”. What functions does language and its related things have, that is, what can become language resources? This is a question needs to be explored. Especially with the rapid development of language intelligence, many things that we are not aware of may enter the ranks of language resources. According to the above research and the current level of understanding, language resources can be divided

---

<sup>15</sup> Zhang Pu (2007) said that “language resources can be divided into the following four categories”, while there are only three types below. “Four categories” may be a clerical error.

into three categories<sup>16</sup>: (1) oral language resources; (2) written language resources; (3) language derived resources, including language knowledge, language technology, language art, language talents, etc.

Natural language mainly exists in the form of spoken language and written language, which are the most basic language resources. As far as the collection, collation, database construction and protection of resources are concerned, there is no opposition between language and speech, and all that we are exposed to, are language resources of speech. “Speech” exists in reality and daily life while “language” exists in linguistics, in the brains of linguists and in the operation of academic stream. In this regard, there is no need to distinguish “language resources” and “speech resources”; or, there is only “speech resources”, but no “language resources”.

“Oral language resources” and “written language resources” are the reclassification of natural language resources. In many documents, language resources refer to these two types of resources. In fact, “language derived resources” are also very important: first, language knowledge, language technology, language talents, etc., play an important role in the collection and collation of language resources, labeling and storage, analysis and research, development and application of language resources; second, language art itself (including calligraphy, literature, and art mainly relying on language, such as drama, crosstalk, sketch, joke, etc.) is a valuable language resource. Therefore, language resources should include language derived resources.

### 3.3. The functional perspective of language resources

The social significance of language resources lies in its function. To treat language resources, we must form a “functional perspective”. We should, from the perspective of function, recognize language resources, evaluate the construction of language resources, and develop and utilize language resources. The function of language resources is gradually developed and recognized with the progress of society. In our cultural tradition, characters and written language are more valued than spoken language. However, in the recent research and practice of language resources with language protection as the primary task, more attention has been paid to the significance of spoken language resources, while less attention has been paid to the meaning of written language resources. As for the derivative language resources, such as language knowledge, language technology, language art, language talents, etc., though involved in varying degrees, are still lacking conscientious sorting.

According to the current practice of language resources utilization, the functional domains of language resources mainly include three aspects: language protection, language information processing, and language learning.

---

<sup>16</sup> Li Yuming (2009) once divided language resources into three categories: the first category is natural languages and their characters, including Chinese characters and Chinese dialects, minority languages and their dialects, foreign languages and characters, etc.; The second category is the derivatives of natural languages, such as dictionaries, various proofreading methods, information retrieval methods using language and characters, language textbooks, various standards of language and characters, corpus, language knowledge base, computer font, various keyboard input methods, various software technologies for processing language (including language translation), etc.; The third category is language capability, including native language capability and foreign language capability. People with excellent language capability means a variety of language talents. This paper is to optimize the “three categories” in 2009, that is, the first type of language resources in 2009 is divided into two categories, and the latter two categories are integrated into “language derived resources”. “Language capability” can be listed separately in the third category or can be classified into a subcategory with “language talents”.

### 3.3.1. Language protection

Language protection is a hot topic in the world. According to experts' prediction, 90% of the world languages will be endangered or even extinct by the end of this century. From the perspective of communication, the endangerment or extinction of these languages may not be a serious problem; but from the perspective of culture, it is a cultural disaster. Each language records the history, experience and world outlook of a certain nation (tribe), and most of the spiritual wealth have not been understood by modern human society and have not become a part of modern human knowledge. Language protection is a race against time to save the lost and irrecoverable human spiritual assets. There are more than 100 languages and a large number of dialects in China, and many of them are already in an endangered state or showing signs of endangerment. The task of language protection is urgent and difficult.

There are three levels of language protection: the first level is "language preservation". In other words, language (including dialects) can be recorded by means of written records, audio and video recordings, and databases and museums should be built to preserve these "language specimens". At present, most of the work done by scholars is at the level of language preservation. The language resources involved in language preservation are mainly oral resources, especially those of endangered languages. In fact, ancient written literature, folk language art and so on also have important language preservation value. Further research will find that the language resources as "language samples" can be further divided into two types, one is "exhibition samples" ("展示本"), the other is "holographic samples" ("全息本"). Exhibition samples are mainly used for language display, language teaching, basic research, etc., which require that language resources can reflect the basic characteristics of a language, such as the basic structural features, communicative features, and cultural features. Holographic samples require that language information and cultural information of a language to be preserved as much as possible. If necessary, the language can be "revived" with the help of virtual technology. This is the most romantic ideal of language protection.

The second level is the "living language protection," to extend the life of language and maintain the vitality of language through various measures. Because the protection of living language will inevitably interfere, in different degrees, with the lifestyle and environment of the language's users, the ethical requirements are very high, which makes the work very difficult to carry out. At home and abroad, some explorations have been made and experiences accumulated, such as the living protection of Indian language in North America, Maori language in Oceania, Yugu language(裕固语) and Miao language in Guizhou province, China. In general, however, the results are far from notable and the prospects are not clear.

The third level is the "development and utilization of language resources", to further develop the achievements of language preservation and language living protection, thus acquiring the social "bonus" of language protection. The consciousness of "bonus" is of importance. It can stimulate the power of language protection, release for the full effectiveness of language protection in time, constantly increase the strength of language protection, and ensure the sustainable development of language protection. It should be noted that the recipients of the "bonus" of language protection should first consider the providers of language resources, including pronunciation partners and their affiliated communities.

### 3.3.2. Language information processing

Language information processing is the main user of language resources at present, and also an active advocate and builder of language resources. In China, the establishment of the Chinese Language Resources Alliance is the best endorsement of the above-mentioned understanding. Language information processing is not only a matter of science and technology, but also an important force to promote social progress. It is its development and progress that makes today's society stride forward to the "intelligent era". Artificial intelligence is the most important technical force in the era of intelligence, and its core is language intelligence. The machine's acquisition of language intelligence mainly depends on the training of language big data. Language big data is also a form of language resources. From the perspective of language intelligence, language resources are within the category of means of production, and their significance to human beings is even more significant.

The language resources involved in language protection mainly preserve the past human world, human history, human experience and world outlook; language resources involved in language information processing are used to create new knowledge and discover new world. Language protection is for the past, and language information processing is for the future, although understanding the existing world would help discovering new worlds.

The language resources used in language information processing can be further divided into two types: one is to help achieve "language proficiency" ("语言通"), the other is to help achieve "information proficiency" ("信息通"). "Language proficiency" is to enable the machine to understand and use language, including machine translation, machine writing, etc. "Information proficiency" enables machines to process information in various fields, including "social computing" and secondary processing of scientific and technological documents. "Language proficiency" also requires the machine to have certain professional knowledge and social common sense, that is to say, it needs the support of "information proficiency", while "information proficiency" needs to be based on the development "language proficiency". The language resources used in "language proficiency" and "information proficiency" have different requirements. Therefore, the language resources for language information processing can be further divided.

### 3.3.3. Language learning

Language learning is one of the most important forms of language life for humans. With the development of society, the task of language learning is becoming ever more demanding. People need not only to learn their mother tongue (especially written language), but also to learn foreign languages, even several foreign languages. Even if the development of language intelligence reaches a high level and machine translation can meet the basic needs of translation, foreign language learning is still needed, because it is a quality of well-rounded human development.

Language learning can only be carried out by relying on language resources. Language learning resource is an academic field that language resource researchers have entered relatively early. Traditional language learning resources include textbooks, reference books, extracurricular books, language teachers, and necessary language environment. In the era of information technology, new concepts such as Intelligent Learning, Smart Learning, MOOC and Micro-lectures are dazzling. Constructing language learning resources suitable for "Internet+" has become the theme of education reform. Language knowledge, language technology, language talents, language courses, and language learning environment are important components of



language learning resources in the new era. There have been many practices in this field, but real research and applications are still limited.

### 3.4. The view of language knowledge

There are many more functions of language resources, but the three discusses above are enough to show the great value of language resources, the significance and urgency of language resources construction, and the significance of conducting language resources research and establishing the discipline of language resources.

Since Saussure established modern linguistics, “language is a symbolic system” has become an axiomatic academic common sense. However, from the perspective of language resources, especially the function of language resources, language is not only a “symbolic system”, but also a “knowledge system”. For example, language protection should not only record and protect language symbols, but also human language knowledge systems and cultural systems. Traditionally, language investigations conducted by linguists aim to reveal the semiotic system of language, which is obviously insufficient in revealing the knowledge system and cultural system of language. From the standpoint of “the view of language knowledge”, vocabulary, grammar, pragmatics and discourse are most important. Human knowledge and thinking are mainly stored in vocabulary, grammar, pragmatics and discourse. From this point of view, language protection must be greatly improved in the methods and content of language investigation.

If we look at language from the perspective of language information processing, it is not possible for a computer to understand human language only by understanding the symbolic system of language; what computers learn from language big data is not only the symbolic system of language, but also the knowledge system and even the cultural system of language. This may be the reason why computers cannot communicate by learning only the “rules” given by linguists. Language information processing can also be described as “machine language learning”, and human language learning works in similar ways as well. Now people emphasize that foreign language learning should change from “learning a foreign language” to “learning with a foreign language”, which shares a similar truth.

## 4. Conclusion

It was only in the 1970s that human beings began to realize the nature of language as social resources. Australia took the lead in materializing the awareness of language resources by implementing relative policies into its national language planning in the 1980s. China’s language information processing academia began to prepare for the establishment of “Chinese Linguistic Data Consortium” in 2003, pioneering in the construction of language resources. In 2004, the State Language Commission of China established the “National Center for Language Resources Monitoring and Research”, and the concept of language resources began to enter national language planning. After that, with the successive development of “China Audio Database for Language Resources” and “Language Resources Protection Project of China”, China has become the most prominent country in the world to vigorously promote the concept of language resources and carry out language protection. In 2018, China and UNESCO jointly held the first “World Conference on the Protection of Language Resources” in Changsha and formed the “*Yuelu Proclamation* (Draft)”, both provided by China’s language planning and shared with the international community.

Chinese language resources research is carried out under the promotion of language planning practice and has formed a positive interaction with the practice of language planning.

The period from 1981 to 2003 can be regarded as the academic brewing period of language resources. There are three academic peaks or peak areas in 2004, 2007–2011 and 2015–2017. Language resources have developed into an important and attractive academic field, and the discipline construction of language resources has been put on the agenda.

The early research mainly determined the resource attributes of language and focused more on the categorization of language resources. The academic circles have different views on the types of language resources. This paper attempts to divide language resources into three categories: oral language resources, written language resources and derivative language resources (language knowledge, language technology, language art, language talents, etc.), and holds that to better solve the classification of language resources, a functional view of language resources should be established. At present, the most important functional domains of language resources are language protection, language information processing, and language learning. The classification, construction specifications, and evaluation criteria of language resources should be considered in line with these functions.

From the perspective of language resources, especially the function of language resources, language is not only a symbolic system, but also a “knowledge base” in which human language knowledge systems and cultural systems are stored. Only by establishing the “view of language knowledge”, can the construction of language resources meet the needs of language protection, machine language learning and human language learning.

Finally, it should be pointed out that the protection and aggregation of language resources require the cooperation among human societies, including the cooperation of different regions, countries, international organizations, social sectors and disciplines. In order to ensure the smooth development of such cooperation, a series of international standards need to be formulated, including technical standards, working standards and social ethics standards. Language resources are related to human knowledge inventory, spiritual home and the centralized management of means of production. Academic research should be strengthened, and social action should be accelerated.

### References

- Chen, Zhanqiang (2008) Lun yuyan ziyuan [On language resources], *Yuyan wenzi yingyong*, no. 1 (In Chin.)
- Fan, Junjun (2006) Lianheguo jiaokewenzuzhi guanyu baohu yuyan yu wenhua duoyangxing wenjianhuibian [UNESCO documents on the protection of language and cultural diversity], Beijing: The Ethnic Publishing House (In Chin.)
- Fan, Junjun & Xiao, Zihui (2008) Yuyan ziyuan lungang [An outline of language resources], *Nanjing shehui kexue*, no. 4 (In Chin.)
- Fishman, Joshua A. (1973) Yuyan xiandaihua he guihua yu guojia xiandaihua he guihua de bijiao [Language modernization and planning in comparison with other types of national modernization and planning] (Translated by Gao Jianping), in Zhou Qingsheng (Eds.) 2001 (In Chin.)
- Guojia minwei minzu wenti yanjiu zhongxin [Center for Ethnic Studies, National Ethnic Affairs Commission of the People’s Republic of China] (2007) Ouanzuzhi minzu wenti ziliao huibian (Unpublished data), Beijing: The Ethnic Publishing House (In Chin.)
- Huang, Xing (2018) Zhongguo yuyan ziyuan duoyangxing ji qi chuanguan yu baohu guihua [Diversity of language resources in China and their modernization and protection plan] in Yan Guohua (Eds.) *Yuyan keji yu renlei fuzhi – Shoujie Zhongguo Beijing guoji yuyan wenhua bolanhui lun wenji* [Language science and technology and human well-being –

- Proceedings of the first Beijing International Language and Culture Expo, China], Beijing: Foreign Language Teaching and Research Press (In Chin.)
- Li, Rulong (2008) Hanyu fangyan ziyuan ji qi kaifa liyong [Chinese dialect resources and their development and utilization], Zhengzhou daxue xuebao (Zhexue shehui kexue ban), no. 1 (In Chin.)
- Li, Yuming (2006) Guanzhu yuyan shenghuo (Yuyan shenghuo redian wenti kalian jiyu [Attention to language life (Language life hot issues column)], Zhangjiang xueshu, no. 1 (In Chin.)
- Li, Yuming (2008a) Yuyan ziyuan guanji Zhongguo yuyan pucha [Language resources and Chinese language Census], Zhengzhou daxue xuebao (Zhexue shehui kexueban), no. 1 (In Chin.)
- Li, Yuming (2008b) Dang jinrenlei sanda yuyan huati [Three major language topics of modern society], Yunnan shifan daxue xuebao (Zhexue shehui kexueban), no. 4 (In Chin.)
- Li, Yuming (2009) Gongmin yuyan nengli shi guojia yuyan ziyuan – xu (muyu, wenzhang, jiaoyu) [Citizens' language ability is a national language resource-preface (Mother tongue, article, education)], Zhongguo daxue jiaoxue, no. 2 (In Chin.)
- Li, Yuming (2010) Lun Zhongguo yuyan ziyuan yousheng shujuku de jianshe [On the construction of Chinese language resources audio database], Zhongguo yuwen, no. 4 (In Chin.)
- Lo Bianco, Joseph (1988) Aodaliya de guojia yuyan zhengce [National language policy of Australia], Yuwen jianshe, no. 5 (In Chin.)
- Qian, Nairong (1995) Hanyu fangyan yanjiuzhong de xin shouhuo – zhuhe xiandai hanyu fangyan yinku faxing jianping “Shanghaihua yindang” [New achievements in the study of Chinese dialects – congratulations on the release of the Shanghai Dialect Audio Database within the Database of Modern Chinese dialects], Yuwen yanjiu, no. 4 (In Chin.)
- Qiu, Zhipu (1981) Shilun yuyan ziyuan de kaifa – jianlun hanyu mianxiang shijie wenti [On the development of language resources – concurrently on the issue of Chinese language facing the world], Yuyan jiaoxue yu yanjiu, no. 3 (In Chin.)
- Qiu, Zhipu (2000) Yingyong yuyanxue de xin gainian [New concepts of applied linguistics], Zhenjiangshi zhuanxuebao (shehuikexueban), no. 3 (In Chin.)
- Tao, Yuanke (1996) Ying zhuyi kaifa liyong Aomen shehui de yuyan ziyuan [Attention should be paid to the development and utilization of the language resources of Macao society], Xueshu yanjiu, no. 4 (In Chin.)
- Wang, Hui (2007) Yuyan guihua de ziyuan guan [The resource view of language planning], Beihua daxue xuebao (Shehui kexueban), no. 4 (In Chin.)
- Wang, Hui (2010) Aodaliya yuyan zhengce yanjiu [Australian language policy research], Beijing: China Social Sciences Press (In Chin.)
- Wang, Shikai (2009) Yuyan ziyuan yu yuyan yanjiu [Language resources and language research], Shanghai: Xuelin Press (In Chin.)
- Wei, Hui (2015) Guojia yuyan nengli youguan wenti tantao [Discussion on issues related to national language ability], Yuyan wenzi yingyong, no. 4 (In Chin.)
- Wei, Hui (2016) Wenhua qiang guo shijiao de guojia yuyan zhanlüe tantao [A discussion on national language strategy from the perspective of cultural power], Wenhua ruanshili yanjiu, no. 3 (In Chin.)
- Xu, Bo, Sun, Maosong, Jin, Guangjin (2003). Zhongwen xinxi chuli ruogan zhongyao wenti [Some important issues in Chinese information processing], Beijing: Science Press (In Chin.)
- Xu, Daming (2008) Yuyan ziyuan guanli guihua ji yuyan ziyuan yiti [Language resource management planning and language resource issues], Zhengzhou daxue xuebao (Zhexue shehui kexue ban), no. 1 (In Chin.)
- Ruan, Guiren (2005) Shuli kexue fazhangan, kaichuang minzu yuyan wenzi guifan biao zhun

- jianshe ji xinxihua gongzuo de xin jumian [Establish a scientific outlook on the development and create a new situation in the construction of national language standards and informatization], Jiaoyu butong bao, no. 17, General Office of the Ministry of Education (In Chin.)
- Zhang, Pu (2007) Lun guojia yuyan ziyuan [On national language resources], Minzu yuyan wenzi xinxi jishu yanjiu – dishiyi jie quanguo minzu yuyan wenzi xinxi xueshu yantaohui lunwen ji [Research on National Language Information Technology – Proceedings of the Eleventh National Seminar on National Language Information] (In Chin.)
- Zhang, Pu (2012) Zhang Pu yingyong yuyanxue lun wenji [Collection of Essays by Zhang Pu on Applied Linguistics], Beijing: Beijing Language and Culture University Press (In Chin.)
- Zhang, Pu & Wang, Tiekun (2009) Zhongguo yuyan ziyuan luncong [Collection of Theses on Chinese language resources], Beijing: The Commercial Press (In Chin.)
- Zhang, Zhengbiao (2000) Xibu yuyan ziyuan zhongyaoxing liyi [Examples of the Importance of Western Language Resources], Zhongguo yuwen, no. 1 (In Chin.)
- Zhao, Jun, Xu, Bo, Sun, Maosong et. al. (2003) Zhongwen yuyan ziyuan lianmeng de jianshe he fazhan [Establishment and development of the Chinese language resource alliance], in Xu Bo, Sun Maosong, Jin Guangjin (Eds.) Zhongwen xinxi chuli ruogan zhongyao wenti [Important Issues of Chinese Information Processing], Beijing: Science Press (In Chin.)
- Zhongguo yuyan ziyuan yousheng shujuku jianshe lingdao xiaozu bangongshi [Office of the leading group for the construction of China Audio Database for Language Resources] (2010). Zhongguo yuyan ziyuan yousheng shujuku diaocha shouce hanyu fangyan [Chinese Language Resources Audio Database Survey Manual – Chinese Dialects], Beijing: The Commercial Press (In Chin.)
- Zhou, Qingsheng (2001) Guowai yuyan zhengce yu yuyan guihua Jincheng [Foreign language policy and language planning process], Beijing: Language & Culture Press (In Chin.)
- Zhou, Qingsheng (2015) Yuyan shenghuo yu yuyan zhengce [Language life and language policy]. Beijing: Social Sciences Academic Press. (In Chin.)
- Grin, F. (2003). Language planning and economics. Current issues in language planning, 4(1), pp. 1–66.
- Kaplan, R. B. & Baldauf, R.B. (1997). Language Planning from Practice to Theory. Clevedon: Multilingual Matters Ltd.

---

**Li Yuming** – Professor and doctoral advisor of Beijing Language and Culture University, People's Republic of China. Address: 100083, China, Beijing, 15 Xueyuan Road, Haidian District.

E-mail: [liyum@263.net](mailto:liyum@263.net)

---

*Для цитирования:* Ли Юймин. Сохранение языковых ресурсов Китая: теория и практика. [Электронный ресурс]. Социоллингвистика. 2020, № 3(3). С. 10–29. DOI: 10.37892/2713-2951-2020-3-3-10-29

*For citation:* Li Yuming. Theories and practices of China's language resources. Sociolinguistics, 2020, no. 3(3) [online], pp. 10–29. DOI: 10.37892/2713-2951-2020-3-3-10-29