

# 基于 Bibliometrix 软件的国外祖传语文献计量分析

姜占好<sup>1</sup> 秦永丽<sup>1,2</sup>

1. 西安外国语大学; 2. 江苏科技大学

**提 要:**借助基于 R 语言的 Bibliometrix 文献分析软件,以 Web of Science 为文献来源,从文献计量分析描述和研究领域可视化两个层面上对国外祖传语文献计量分析,回答了该领域三个问题:哪些作者和国家相关研究突出?哪些文献比较重要?哪些研究主题已经探讨?文献梳理发现主要领域有祖传语习得、祖传语教学、祖传语语种和身份认同,但缺乏祖传语历时研究和心理语言学的介入。

**关键词:**祖传语;文献计量分析;Bibliometrix 软件

## 引言

祖传语对应的英文是 heritage language<sup>①</sup>。该英文术语最早出现在 1977 年加拿大“安大略祖传语项目”的描述中。到了 20 世纪 90 年代,该术语在国外,尤其是北美地区语言政策和语言教育领域广泛使用。

国外围绕该术语研究开展的活动不胜枚举,其中最早的是 1996 年美国外语教学学会制定的《21 世纪外语学习标准》高频率提及 heritage speaker;其次是 1999 年开始召开的多届全美祖传语大会和 2010 年开始的国际祖传语大会;第三个标志是围绕该研究,2002 年创建了学术期刊 *Heritage Language Journal*,国际知名期刊开设专栏或专刊(如《现代语言学刊》、《第二语言习得研究》、《国际双语教育和双语教学制》和《应用语言学年度评论》),以及权威语言学研究手册(综览)都辟出专栏或词条讨论“祖传语”(如 *Routledge Handbook of Second Language Acquisition*, *The Handbook of Language Teaching*, *Encyclopedia of Language and Education*),甚至出现了有关祖传语研究的专门手册和语言政策书籍,如

*Handbook of Research and Practice in Heritage Language Education* (2018) 和 *Heritage Language Policies around the World* (2018);2003 年在加州大学洛杉矶分校成立了美国祖传语中心。

国外祖传语研究如火如荼,吸引了语言教学、语言规划和政策众多研究者。本文旨在使用基于 R 语言的 Bibliometrix 软件,快速梳理 Web of Science 已有国外研究论文,回答上述问题,为国内相关研究人员提供参考。

## 一 Bibliometrix 软件简介

Bibliometrix 是一款基于 R 语言的文献科学计量软件。目前该软件处理的文献信息主要源于数据库 Scopus 和 WOS(Web of Science)。该软件基于脚本运行,相对于其他基于图形用户界面操作的文献计量软件,具有两大优势:(1)使用脚本运行,不需要在图形界面进行多次重复操作,提高文献处理

<sup>①</sup> heritage language 国内译名有“遗产语言”、“传承语”、“继承语”、“族裔语”、“祖裔语”、“族裔传承语”、“祖籍传承语”和“族语”,本文使用“祖传语”,相关论述见方小兵(2017)。

效率,减少了出错概率,更适合数量大、重复性、多步骤的文献计量分析工作;(2)基于 R 语言环境,开源免费,而且用户可以根据实际需要,自己(他人)重新编写或修正 R 语言程序代码,不需要跨平台进行操作和处理,不需要任何费用,方便经济。

Bibliometrix 软件主要功能有:数据导入和文献格式转换、文献计量分析描述、文献和属性矩阵构建、文献归一化处理、文献降维分析、网络矩阵构建和构图(李昊等,2018)。我们选取文献计量分析描述和研究领域可视化两大模块中的分项,在 Web of Science 的“主题”模块中输入 heritage language 进行查询,并对命令运行结果进行解

释。启动 Bibliometrix 软件,对检索到的国外祖传语文献进行描述分析,得到如下相关信息:文献总体基本信息、年发文量、以国家为单位论文总被引量、与研究主题关系最为密切的期刊、最高被引文献信息(作者、发表年份、杂志名称)、发文量最多作者、相关性最高的关键词列表。

二 国外祖传语文献计量分析

本文选取其中三个项目进行解释:发文量最多的作者和国家;最高被引文献信息;话题最关联关键词。

(一) 发文量最多的作者和国家

软件对文献作者的分析结果见表 1。

表 1 发文量最多的作者

| Most Productive Authors |                         |          |                   |                         |
|-------------------------|-------------------------|----------|-------------------|-------------------------|
| Authors                 |                         | Articles | Authors           | Articles Fractionalized |
| 1                       | Montrul S               | 13       | Montrul S         | 8,75                    |
| 2                       | Polinsky M              | 6        | Polinsky M        | 4,67                    |
| 3                       | Llombart Huesca A       | 5        | Llombart Huesca A | 4,50                    |
| 4                       | Mu G M                  | 5        | Mu G M            | 4,50                    |
| 5                       | Albirini A              | 4        | Burgo C           | 4,00                    |
| 6                       | Burgo C                 | 4        | Lee J S           | 3,50                    |
| 7                       | Curdts Christiansen X L | 4        | Melo Pfeifer S    | 3,33                    |
| 8                       | Elola I                 | 4        | Albirini A        | 3,00                    |
| 9                       | Guardado M              | 4        | Atoofi S          | 3,00                    |
| 10                      | Jegerski J              | 4        | Beaudrie S M      | 3,00                    |

从表 1 中可知,围绕 heritage language 研究发文量最多的前 10 位作者中,Silvina Montrul 以 13 篇文章居榜首,其次是 Maria Polinsky 和 Amàlia Llombart-Huesca 两位。这三位作者都是美国高校的语言学教授,研究领域都涉及 heritage language。2016 年,Montrul 教授出版了《祖传语习得》(*The Acquisition of Heritage Languages*)一书,对过去 20 年祖传语使用者和祖传语

研究进行了权威性整合与批判性概述。Maria Polinsky 是马里兰大学帕克分校的语言学教授,她关注的是祖传语及其对语言学理论的影响。Amàlia Llombart-Huesca 是加州州立理工大学英语与现代语言系的西班牙语语言学教授,她关注西班牙祖传语学习者。

软件对文献作者所属国家的分析结果见表 2。

表 2 发文量最多的国家

| Corresponding Author's Articles |                |          |        |     |     |           |
|---------------------------------|----------------|----------|--------|-----|-----|-----------|
| Country                         |                | Articles | Freq   | SCP | MCP | MCP_Ratio |
| 1                               | USA            | 259      | 0.5307 | 241 | 18  | 0.0695    |
| 2                               | Canada         | 52       | 0.1066 | 35  | 17  | 0.3269    |
| 3                               | United Kingdom | 26       | 0.0533 | 19  | 7   | 0.2692    |
| 4                               | Germany        | 20       | 0.0410 | 14  | 6   | 0.3000    |
| 5                               | China          | 14       | 0.0287 | 7   | 7   | 0.5000    |
| 6                               | Australia      | 11       | 0.0225 | 7   | 4   | 0.3636    |
| 7                               | New Zealand    | 11       | 0.0225 | 10  | 1   | 0.0909    |
| 8                               | Spain          | 9        | 0.0184 | 8   | 1   | 0.1111    |
| 9                               | Netherlands    | 5        | 0.0102 | 4   | 1   | 0.2000    |
| 10                              | Sweden         | 5        | 0.0102 | 4   | 1   | 0.2000    |

SCP: Single Country Publications

MCP: Multiple Country Publications

由表 2 可知,发文量最多的国家中,前五位分别是美国、加拿大、英国、德国和中国。发文量多少是一个国家对祖传语研究重视程度的标杆。祖传语通常与移民语言和濒危语言(endangered language)有关(张广勇,2013)。美国和加拿大都是移民国家,两个国家发文量最高正说明北美是世界祖传语研究中心,祖传语研究是语言政策和语言教育领域广受关注的热点。英国、德国和

中国都是由多个民族组成的单语制国家,在国家通用语言和文字普及过程中,少数民族语言如何保护和传承成为多民族文化国家亟待解决的问题,而濒危语言的保护正是祖传语研究的范畴。

(二) 最高被引文献信息

祖传语研究领域哪些文献比较重要?我们以软件给出的最高被引文献信息来考察(见表 3 和表 4)。

表 3 前 10 个高被引文献信息

| Top manuscripts per citations |   |     |           |
|-------------------------------|---|-----|-----------|
| Paper                         |   | TC  | TCperYear |
| 1                             | Valdes G, 2005, Mod Lang J                | 159 | 11.36     |
| 2                             | Cummins J, 2005, Mod Lang J               | 156 | 11.14     |
| 3                             | Montrul S, 2008, Lang Learn               | 138 | 12.55     |
| 4                             | Ricento T, 2005, J Socioling              | 94  | 6.71      |
| 5                             | Hornberger N H, 2005, Mod Lang J          | 89  | 6.36      |
| 6                             | Benmamoun E, 2013, Theor Linguist         | 88  | 14.67     |
| 7                             | Polinsky M, 2011, Stud Second Lang Acquis | 88  | 11.00     |
| 8                             | Rothman J, 2009, Int J Biling             | 86  | 8.60      |
| 9                             | Rothman J, 2007, Int J Biling             | 82  | 6.83      |
| 10                            | Potowski K, 2004, Mod Lang J              | 81  | 5.40      |

表 3 中最高被引文献是 Guadalupe Valdés 于 2005 年发表在 *The Modern Language Journal* 上的论文: Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized? 作者在讨论第二语言习得(SLA)和语言教学交集的时候提出若干扩大 SLA 和语言教育界面研究的途径;基于双语主义思想,作者提出教授祖传语面临的挑战,特别讨论了祖传语学习者的双语现象,进而作者宏观上列举了祖传语者研究中出现的问题;最后,作者从专业实践和学科领域描述了祖传语研究,呼吁第二语言习得领域的研究要借鉴其他领域的成果,为世界范围内少数民族语言儿童教育添砖加瓦,为更深一步了解人类语言能力作出贡献。该篇文章整体上是基于文献梳理的质性研究,因而对其他学者从事祖传语研究具有理论借鉴和启示意义。笔者对 10 篇高被引文献的研究内容和研究方法进行简单梳理发现,10 篇文章中只有 3 篇是实证研究,通过基于个案的实验方法或课堂观察调查祖传语习得的问题(文献 3、9 和 10),其他 7 篇都是基于文献梳理,从社会学或语言学视角分析祖传语资源保留的策略(文献 2)、美国祖传语保护方法(文献 4)、祖传语教育的发展空间(文献 5)、祖传语研究与语言学理论建构的联系(文献 7)及开展祖传语实证研究和认识论的相关问题(文献

7)。这些高被引论文研究内容和方法可以折射出现有祖传语研究大都是关于祖传语的教育和保护策略方法的理论探讨,而从实证角度验证对于祖传语教学和习得的实效性有待增强,心理语言学对祖传语习得研究不可或缺。

同时,读者和研究人员还可以通过表 4 了解哪些刊物经常刊发祖传语话题的论文,进而关注这些期刊,了解该话题的重要文献,跟踪国外该话题的研究趋势。我们发现这些期刊有一个共同特点,即刊物主题都是关于外语/二语习得、教学,语言身份和教育相关的领域,所列的 10 种期刊名中有二语、外语字样的就有 5 个。刊发祖传语相关文章最多的 *Modern Language Journal* 是国际二语习得领域的重要期刊,主要刊登关于外语/二语习得与教学的研究和论述,刊发量最多佐证了祖传语习得与教学是二语习得研究的热点。排名第二的期刊 *Hispania* 是美国西班牙语和葡萄牙语教师联合会(AATSP)官方刊物,主要刊登西班牙语和葡萄牙语教学论文。在美国,这两种祖传语都有很大的学习者群体,发文量三甲之一的 Amàlia Llombart-Huesca 教授关注西班牙祖传语学习者,Silvina Montrul 教授被引用最多的两篇文章都是关于在美祖传语西班牙语学习者的文章。

表 4 最相关文献来源期刊

| Most Relevant Sources |   |          |
|-----------------------|---|----------|
|                       | Sources   | Articles |
| 1                     | Modern Language Journal   | 33       |
| 2                     | Hispania—A Journal Devoted to the Teaching of Spanish and Portuguese  | 30       |
| 3                     | Canadian Modern Language Review—Revue Canadienne Des Langues Vivantes | 29       |
| 4                     | Foreign Language Annals   | 29       |
| 5                     | Journal of Multilingual and Multicultural Development                 | 19       |

续表

|    |   |    |
|----|---|----|
| 6  | International Journal of Bilingualism                         | 17 |
| 7  | International Journal of Bilingual Education and Bilingualism | 15 |
| 8  | Language Culture and Curriculum                               | 15 |
| 9  | Journal of Language Identity and Education                    | 11 |
| 10 | Bilingualism—Language and Cognition                           | 10 |

（三）话题最关联关键词

国外祖传语研究的重点从软件的“话题最关联关键词”分析结果中可以观察到（见表 5）。

表 5 话题最关联关键词

| Most Relevant Keywords |                             |          |                    |          |
|------------------------|-----------------------------|----------|--------------------|----------|
| Keywords (DE)          |                             | Articles | Keywords-Plus (ID) | Articles |
| 1                      | Heritage Language           | 68       | Acquisition        | 45       |
| 2                      | Spanish                     | 28       | Spanish            | 42       |
| 3                      | Bilingualism                | 27       | Speakers           | 42       |
| 4                      | Identity                    | 23       | Learners           | 35       |
| 5                      | Heritage Speakers           | 20       | English            | 31       |
| 6                      | Heritage Languages          | 14       | Identity           | 29       |
| 7                      | Language Maintenance        | 13       | Maintenance        | 27       |
| 8                      | Heritage Language Education | 12       | Children           | 26       |
| 9                      | Incomplete Acquisition      | 12       | Bilingualism       | 23       |
| 10                     | Motivation                  | 12       | 2ND Language       | 22       |

从表 5 我们得知,以“作者关键词”进行检索查询出来的研究话题和对应的论文数量,论文数量大于 20 的话题有 Heritage Language (68)、Spanish (28)、Bilingualism (27)、Identity (23)、Heritage Speakers (20)。以“关键词拓展”进行检索查询出来的研究话题和对应的论文数量,其中前 5 位的是 Acquisition (45)、Spanish (42)、Speakers (42)、Learners (35) 和 English (31)。综合两种检索结果,我们发现国外祖传语研究的重点如下:

1. 祖传语的习得问题

这方面的研究基于第二语言习得研究

理论和研究范式,探讨祖传语习得研究与二语习得和一语习得之间的异同,突出对比研究的重要性(Valdés, 2005),关注的具体语言现象有祖传语学习者的发音和音位特征(Chang et al., 2011),语言迁移和语言石化问题(Albirini et al., 2013),以及影响祖传语习得的因素(Noels, 2005; Comanaru & Noels, 2009)。Montrul 综述了围绕祖传语习得者的语法知识和词汇知识发展研究(如语音、音位、词汇、句法、词汇句法界面),从习得模式、语言知识类别、任务类型、课堂教学反应类别和最终收获,对比了一语、二语和祖传语习得者的异同,探讨了祖传语不完

全习得的特性,总结了三种理论框架(形式语言学、认知语言学和涌现理论)。这方面的研究发现,祖传语的习得与母语习得和二语习得既有共同点,也有差异,祖传语习得有母语和二语习得的共性特征,是母语和二语的交叉点。

习得问题研究中另外一个聚焦点是祖传语习得者的语言潜能和认知潜能。由于习得者所处的语言输入环境好于二语习得者,其语音能力、语法认知和语用意识要优于二语习得者,在语言潜能和认知潜能方面具有相对优势(Benmamoun, Montrul & Polinsky, 2013)。除了语言环境和年龄外,语言输入和输出的质、量也会影响到语言潜能的形成。

但祖传语习得者相对优越的语言潜能和认知潜能并不意味着其完整的祖传语的习得,这种不完全性体现在祖传语习得者书面语形态(如动词、名词)判断错误多于二语者(Montrul, 2011),语言使用的简化和语言使用的泛化(Polinsky, 2011)。导致不完全习得的原因有狭窄的语言输入渠道(相对于母语习得)、语言磨蚀(Polinsky, 2011)、学校祖传语教育机会少(Hornberger, 2015)。

## 2. 祖传语习得者的身份认同

表 5 中 identity, bilingualism, 以及 speaker, learner 反映出祖传语习得研究中的身份认同问题。受 20 世纪 90 年代应用语言学领域社会学转向的影响,身份认同成为热点问题,体现在祖传语研究中,有祖传语习得者的身份定义问题,如身份认同不是一个固定的一成不变的状态,而是一个发展变化、流动性的过程;身份认同受制因素众多,如习得者的家庭环境、学校环境和国家的语言政策及规划,受制于习得者的成长经历、语言实践;身份认同的研究方法多样,如问卷、访谈、课堂观摩等(Leeman, 2015)。

## 3. 语种问题

通过表 5,我们发现祖传语相关研究中,西班牙语和英语成为研究者关注的主体语种,发表的论文分别为 42 篇和 31 篇。这同祖传语的缘起和相关活动相关。正如引言中所提及的,祖传语缘起 1977 年加拿大“安大略祖传语项目”,关注的对象主要是北美移民、土著和殖民者后裔。其中美国的祖传语研究对象大多数为西班牙语的拉丁裔,所以西班牙语成为首要研究语种。其次,祖传语相关会议和活动也以北美为主要阵地,也促使研究者聚焦英语作为主导语言下的西班牙语等传承语的学习问题。这从一个侧面表明,祖传语研究的语言对象过于狭窄,没有“真正实现对弱势语言、濒危语言的关照”(方小兵, 2017)。

## 4. 教学问题

表 5 中以“关键词拓展”搜索出 35 篇以 learners 为关键词的论文,围绕学习者开展的祖传语教学是有一个研究热点。首先是祖传语在外语教学中的定位问题。目前,祖传语较多的国家如美、加、澳等将外语教学和祖传语教学等同在一起,忽略了祖传语学习者的特别需求,进而 Kagan 和 Dillon (2001)针对两类不同的学习者,提出宏观和微观的途径,进行有针对性的教学(详见曹贤文, 2017)。Kondo-Brown (2010)以翔实的文献综述了在美国实施的旨在提高祖传语学习者祖传语语言能力的研究,在祖传语教学大纲设计和教学法维度方面提出具体建议:如制定以学习者为中心的大纲,以教学内容为中心的课堂教学活动,以教学结果为中心的课程设计,双语双向“浸入式”教学项目的实施,强调学校在学习者文化和种族认同方面所起的作用,协调学习者祖传语口语和书面语不平衡发展的问题,区别对待不同祖传语水平的学习者,以及祖传语教师的培训问题。

上述研究话题和热点还可以通过图 1

的“关键词共现”得以验证。在图 1 中,我们看到关键词贡献分为三大领域(圆圈的大小表示该话题关注频率,连线表示各话题之间的关系):习得、二语和学生(课堂)。这三个区域分别对应祖传语习得(包含身份认同 identity)、祖传语研究中的语种问题和祖传

语的教学问题。当然,它们之间的关系不是截然分开的,而是相互叠加,具有交集。例如:祖传语习得的区域包含了 English, Chinese 和 Spanish 的语种问题。另外,从图 1 中我们还可以推断出其他的研究话题,例如祖传语的保护、青少年对祖传语的态度。

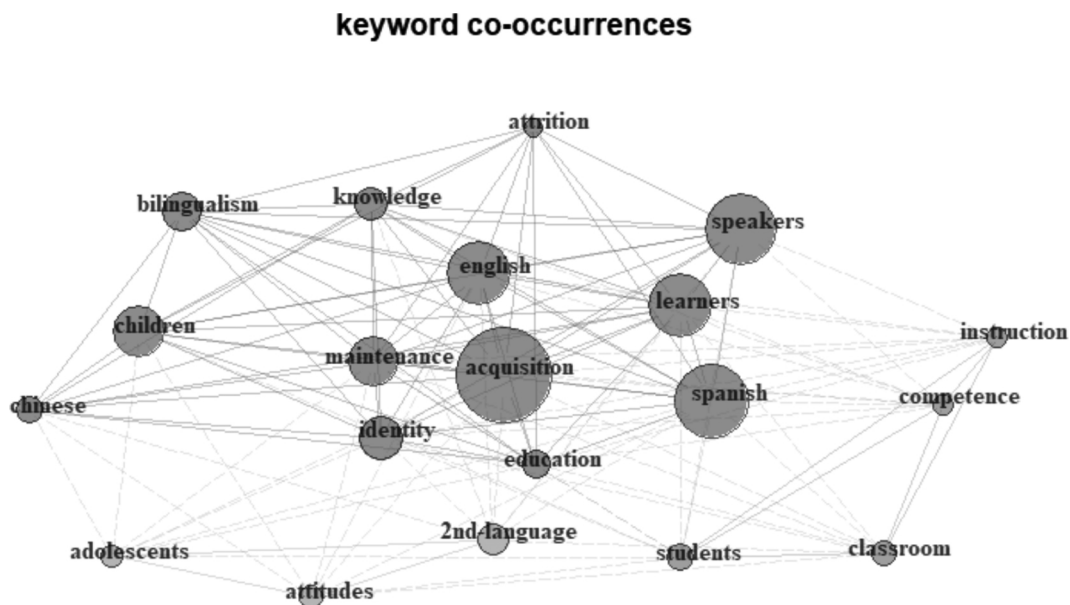


图 1 关键词共现

### 三 结论

基于 R 语言的 Bibliometrix 软件,本文对 WOS 上主题为 heritage language 的论文文献进行梳理发现:该软件能够快速高效地回答围绕某主题词的国外研究的三个研究问题:该领域哪些作者和国家研究成果突出?哪些论文比较具有代表性?现有哪些研究问题和热点话题?就本文而言论文数量前三名的作者有 Silvina Montrul、Maria Polinsky 和 Amàlia Llombart-Huesca,前五位的国家是美国、加拿大、英国、德国和中国。代表性的论文共计 10 篇。研究的热点和主题主要有祖传语的习得、教学、身份认同和语种问题,具体体现为祖传语的习得与

母语(二语)习得既有相同点,也有不同点;祖传语的习得和教学受各种因素的制约、影响(学习中因素有性别、态度、身份、年龄等;社会因素有家庭、学校和社会的环境、语言政策和规划等)。研究方法多为基于文献耙梳和问卷调查的质化研究,也有通过课堂观察和个案研究的实证研究,总体来讲,缺乏历时研究和心理语言学视角的介入。由此,祖传语的研究有助于二语习得理论的验证、研究领域的拓展,同时也提出了新的挑战,为语言习得和教学提供了新的研究领域。

### 参考文献

- Albirini A, Benmamoun E, Chakrani B. 2013. Gender and Number Agreement in the Oral Production of Arabic Heritage Speakers. *Bilingualism: Language and Cog-*

- inition (16).
- Benmamoun E, Montrul S, Polinsky M. 2013. Heritage Languages and Their Speakers: Opportunities and Challenges for Linguistics. *Theoretical Linguistics* (39).
- Chang C B, Yao Y, Haynes E, et al. R. 2011. Production of Phonetic and Phonological Contrast by Heritage Speakers of Mandarin. *Journal of the Acoustical Society of America* (129).
- Comanaru R, Noels K A. 2009. Self-determination, Motivation, and the Learning of Chinese as a Heritage Language. *Canadian Modern Language Review* (66).
- Hornberger N H. 2015. Opening and Filling up Implementation and Ideological Spaces in Heritage Language Education. *Modern Language Journal* 89(4).
- Kondo-Brown K. 2010. Curriculum Development for Advancing Heritage Language Competence: Recent Research, Current Practices, and a Future Agenda. *Annual Review of Applied Linguistics* (30).
- Leeman J. 2015. Heritage Language Education and Identity in the United States. *Annual Review of Applied Linguistics* (35).
- Montrul S. 2011. Morphological Errors in Spanish Second Language Learners and Heritage Speakers. *Studies in Second Language Acquisition* (33).
- Noels K A. 2005. Orientations to Learning German: Heritage Language Learning and Motivational Substrates. *The Canadian Modern Language Review* (62).
- Polinsky M. 2011. Reanalysis in Adult Heritage Language: New Evidence in Support of Attrition. *Studies in Second Language Acquisition* 33(2).
- Valdés G. 2005. Bilingualism, Heritage Language Learners, and SLA Research: Opportunities Lost or Seized? *The Modern Language Journal* 89(3).
- 曹贤文. 2017. 海外传承语教育研究综述. 语言战略研究 (3).
- 方小兵. 2017. 国际祖传语研究焦点分析——基于《祖传语期刊》历年文献. 语言战略研究 (4).
- 李昊, 潘宇光, 王磊. 2018. Bibliometrix: 一款新的基于 R 语言的文献计量软件介绍与评价. 大学图书馆情报学刊 (36).
- 张广勇. 2013. 美国继承语教育对我国少数民族语言保护的启示. 贵州民族大学学报(哲学社会科学版) (3).
- 张广勇. 2014. 国外继承语习得研究新进展. 现代外语 (1).

#### 作者简介

姜占好,男,西安外国语大学英文学院教授,博士生导师。研究方向为二语习得、语用学、语言政策和规划。电子邮箱:447939126@qq.com。

秦永丽,西安外国语大学英语教育学院博士研究生,江苏科技大学外国语学院讲师。研究方向为应用语言学。电子邮箱:ylqin81@163.com。



# Bibliometrix-based Literature Analysis: A Case of Heritage Language Studies Abroad

Jiang Zhanhao<sup>1</sup>, Qin Yongli<sup>1,2</sup>

1. *Xi'an International Studies University*;

2. *Jiangsu University of Science and Technology*

**Abstract:** This paper conducted research on heritage language literature from Web of Science with the help of R-based software Bibliometrix in terms of quantitative description of the literature and literature visualization to answer three questions concerning heritage language studies: 1) main researchers and main countries in this field; 2) important and typical papers in this field, and 3) canonical research topics and foci that include heritage language acquisition, teaching, language categories and identity. However, longitudinal and psycholinguistic explorations into this field remain scanty.

**Key words:** heritage language; bibliometrics and literature analysis; Bibliometrix software