

【语言学】

国家语言能力指数体系完善与研究实践

张天伟

【摘要】国家语言能力是一个国家综合国力的重要体现。我们依据客观数据,运用多指标综合评价分析方法,将国家语言能力指数体系中的管理能力、掌控能力、创造能力、开发能力、拓展能力5个一级指标进一步完善为12个二级指标和26个三级指标,将世界各国国家语言能力的差异以量化形式直观呈现;据此体系进行数据分析,发现国家语言能力与国家经济实力紧密相连,世界各国的国家语言能力形成了呈金字塔形分布的3个层级,美国、中国和英国占据塔尖;5个一级指标中掌控能力、创造能力和拓展能力对整体国家语言能力影响更大;而中国在语言研究、语言教育、语言产业和学术话语方面还需继续努力,以提升国家语言能力。

【关键词】国家语言能力;语言政策;指数体系;主成分分析法

【作者简介】张天伟,男,北京外国语大学教授,中国外语与教育研究中心/国家语言能力发展研究中心,主要研究方向为语言政策、句法学,E-mail:zhangtianwei@bfsu.edu.cn(北京 100089)。

【原文出处】《语言战略研究》(京),2021.5.12~24

【基金项目】北京高校高精尖学科“外语教育学”建设项目(2020SYLZDXM011)。

一、引言

国家语言能力是一个国家软实力和综合国力的重要体现之一,它与国家语言实力、国民语言能力以及公民个人语言能力既有所区别又相互联系。美国学者Brecht & Walton(1993)最早提出国家语言能力这一概念,将其定义为“国家应对特定语言需求的能力”;此后,Brecht & Rivers(2000,2005,2012)陆续阐释和完善了这一观点。国内李宇明(2011)、文秋芳等(2011)、赵世举(2015)、魏晖(2015)等也从不同角度对国家语言能力的定义和理论框架进行了讨论。国内外一些学者围绕这一主题,从理论建构、宏观思考、分项能力、具体项目和国别区域等不同维度进行了研究,详见文秋芳、张天伟(2018),文秋芳(2019)的述评。

虽然学者们对国家语言能力进行了多方面的阐释和分析,但现有成果存在两方面的问题。一是大多偏重于宏观理论分析和整体描述,缺少对国家语言能力各分项的微观化、量化研究。文秋芳(2016)虽然构建了国家语言能力的评价指标,但没有进一步进行量化分析。2020年国家社科基金立项了“全球治理视域下国家语言能力评价指标体系研究”,但还未见到相关研究成果发表。二是没有在全球范围内对不同国家的国家语言能力进行量化比较研究,而在全球竞争由经济、军事竞争转移到包括文化、话语权等在内的综合国力竞争背景下,如何评价国家语言能力的建设成效,提高建设效率,提升建设效果,找出中国与世界发达国家的总体差距以及比较各分项能力的长短,制定一套科学的、动态的、行之有效的国家语言能力评价指标体系,就显得尤为迫切和必要。

为此,我们采取多指标综合评价分析方法,搜集各方面数据,力图构建一个可对世界各国国家语言能力进行分析评价的量化指数体系,其核心是统计指数的构建。指数是统计学概念,是一个反映社会现象综合变动情况的相对数,其实质是测定多项内容,表现形式为动态相对数。国家语言能力指数构建是针对国家语言能

力这一研究对象,建立一个动态测评的指标体系,利用具体的统计方法,对搜集的数据资料进行整理,将多项指标比较分析,整合形成一个概括性的指数,将不同国家语言能力具体量化,并根据反馈数据进行校正修改,最终形成一个结构合理、真实可信、动态更新的国家语言能力综合评价体系。

二、国家语言能力指数体系构建和完善的理论依据

国家语言能力指数体系是围绕“国家语言能力”这一核心概念展开的。国家语言能力指“在政府层面,国家运用语言处理一切与国家利益相关事务的能力”(文秋芳2019),它与国家语言实力、国民语言能力、公民个人语言能力的区别主要在于实施主体、处置对象、应用领域、战略目标、研究重点等方面。因此,本文依据文秋芳、张天伟(2018),文秋芳(2019)的前期研究,将国家语言能力重新定义为:国家围绕国家利益、国家安全、国家需求、国家意识等维度,运用语言及其各种手段,处理一切相关事务的能力。文秋芳(2016)构建了国家语言能力的评价指标体系,将国家语言能力分为5个一级指标:管理能力、掌控能力、创造能力、开发能力和拓展能力。本研究以此为基础,结合实际收集的数据进行了完善,添加了二级和三级指标。具体来说,本研究共构建了5个一级指标、12个二级指标、26个三级指标(见表1)。

表1 国家语言能力指数评价指标及数据来源

一级指标	二级指标	三级指标	数据来源
1. 管理能力	1. 组织力	1. 有无语言文字管理机构	各国政府官网
		2. 语言文字相关智库机构排名	相关全球智库排名
	2. 规划力	3. 有无国家层面语言文字事业规划	各国政府官网
		4. 有无国家层面教育规划	各国政府官网
		5. 有无语言文字法律法规	各国政府官网
		6. 有无少数民族语言文字规划	各国政府官网
	3. 执行力	7. 是否设有国家通用语言文字水平测试	文献检索
2. 掌控能力	4. 应急力	8. 应急语言服务的可获得性	国际翻译能力指数数据 ^①
		9. 有无语言普查数据	联合国统计司数据
	5. 通晓力	10. 国家开设语种专业的数量	国家翻译能力指数数据
		11. 语言服务社团开设的数量	CSA Research ^② 数据
		12. 各语种语言人才储备情况	国家翻译能力指数数据
3. 创造能力	6. 支配力	13. 全球主要语言可提供性	国家翻译能力指数数据
		14. 开设语种教学的高校数量和水平	QS ^③ 现代语言高校排名
	7. 实践力	15. 语言学研究能力排名	QS 语言学高校排名
4. 开发能力	8. 科学力	16. 机器翻译能力的典型个案	国际机器翻译大赛(WMT)
	9. 机器翻译力	17. 信息检索能力的典型个案	国际音频检索评测大赛(MIREX)
		18. 语音和语种识别能力的典型个案	国际多通道语音分离和识别大赛(CHiME)
			NIST 说话人识别技术评测(SRE)
5. 拓展能力	10. 信息挖掘力	19. 该国官方语言是否联合国大会官方语言	联合国6种官方语言
		20. 该国官方语言是否G20成员的官方语言	G20成员官方语言
		21. 该国官方语言是否主要国际组织的官方语言	国际组织年鉴在线数据库
		22. 有无国家通用语言国际传播机构	文献检索
		23. 国家通用语言国际传播机构在全球布局的国家数量	各国语言文化传播机构官网
		24. 国家通用语言国际传播机构在全球布局的机构数量	各国语言文化传播机构官网
	11. 影响力	25. 国家通用语言在 SCIE、SSCI 和 A & HCI 中使用的比率、占比和排名	Web of Science 官方数据
		26. 有无国家通用语言的语言文字国际标准	国际标准化组织(ISO)官网及相关成员国数据

国家语言管理能力指国家对涉及战略利益的语言事务进行管理的能力,包括4个二级指标:组织力、规划力、执行力和应急力。组织力指语言文字管理机构和智库的布局及其任务的合理性,在指数体系构建中,有2个三级指标:一是有无语言文字管理机构,二是语言文字相关智库机构排名。规划力指对语言文字未来需求的预判以及制订应对方案的计划性,包括4个三级指标,即有无国家层面语言文字事业规划、有无国家层面教育规划、有无语言文字法律法规、有无少数民族语言文字规划。执行力指落实规划的速度和效率,以及通过何种手段促进规划的实施,以是否设有国家通用语言文字水平测试这个三级指标来衡量。应急力指处理海内外突发事件的语言能力,其三级指标为应急语言服务的可获得性。

国家语言掌控能力指国家对语言人才资源的掌控能力,包括通晓力和支配力2个二级指标。通晓力指国家对所拥有语言资源的种类和质量的熟悉程度,具体包括有无语言普查数据、国家开设语种专业的数量2个三级指标。支配力指调用语言资源服务国家需求的速度和准确性,具体包括语言服务社团开设的数量、各语种语言人才储备情况和全球主要语言可提供性3个三级指标。

国家语言创造能力指国家对语言人才的培养和教育能力,包括实践力和科学力2个二级指标。实践力指落实培养国家短缺语言人才计划的能力,体现在开设语种教学的高校数量和水平这个三级指标上。科学力指培养短缺语言人才的有效能力,体现在语言学研究能力排名这个三级指标上。

国家语言开发能力指国家对语言技术的开发能力,包括机器翻译力和信息挖掘力2个二级指标。机器翻译力指运用机器翻译语言的速度和质量,我们以国际机器翻译大赛(WMT)的个案研究为三级指标进行量化。信息挖掘力指挖掘公开情报的自动化程度和准确性,用信息检索能力和语音、语种识别能力的典型个案2个三级指标来衡量。

国家语言拓展能力指国家对通用语言国际影响的拓展能力,包括影响力和传播力2个二级指标。影响力指国家通用语言在国际交流中使用的广度和深度,具体包括该国官方语言是否联合国大会官方语言、是否二十国集团(G20)成员的官方语言、是否主要国际组织的官方语言、有无国家通用语言国际传播机构、国家通用语言国际传播机构在全球布局的国家数量和机构数量6个三级指标。传播力指国家通用语言创造和推广新知识的能力,包括国家通用语言在SCIE、SSCI和A & HCI中使用的比率、占比和排名,有无国家通用语言的语言文字国际标准2个三级指标。

由于指标的选择关系到整个指数体系的科学性和可靠性,且国家语言能力的理论建构尚无公认标准,因此本研究在确认指标时主要采用定性研究中的德尔菲法,通过与此方面的专家进行反复访谈和函件调查,确立每个分项能力具体对应的二、三级指标。同时需要说明的是,这些指标是我们依据已有的国家语言能力理论研究进行构建的,反映的是目前我们所认识到的国家语言能力的各个维度。而国家语言能力这一概念与社会政治经济现实、国际国内形势密切相关,其内涵肯定会不断发展变化。因此,我们的指标体系和指数体系构建也不会一劳永逸,将在后续研究中不断根据现实进行调整,以尽可能科学、全面地体现世界各国国家语言能力的动态变化和发展。

三、国家语言能力指数体系构建和完善的具体过程

(一)数据搜集

进行指标体系构建,数据是基础和关键。本研究的数据收集遵循如下原则:权威性、可验证性、可操作性、

动态性。权威性和可验证性指数据来源基本为国家、国际组织或公共组织的官网数据和公开数据,如国家层面的语言政策规划、国家法律法规等均从各国官网查询得出并附有具体查询网址。可操作性指本研究的数据均可收集、可复制、可验证。而一些构想中的二级或三级指标,数据很难量化和收集,因此暂时不纳入本指数体系的计算范围内,如原构想中的言语治疗和语言康复指标,因难以收集世界范围内的公开数据而被删除。动态性指本数据的收集均有公开、可靠的信源并可以动态更新,有利于后续研究,如各国语言文化传播机构的数量、语言产业中语言服务社团的数量都可以根据数据来源持续更新。

我们通过各种方法,从不同渠道收集数据,力求客观公正地展现国家语言能力的各个方面。首先,研究根据联合国官方网站所列成员国,选定我们的研究对象为193个具体国家;然后,通过查询各国政府官网,各种数据资料,得出这193个国家在不同指标方面的分值。具体数据来源如表1所示。

国家语言管理能力包括4个二级指标,8个三级指标。有无语言文字管理机构这一指标的数据主要通过查阅各国政府的官方网站,找出各国语言文字的专门管理机构或归属部门来衡量,如法国境内语言总司、美国国防语言与国家安全教育办公室、马来西亚国家语文局等专门机构,也有部分国家将其归属于教育部、科技部等。语言文字智库机构的排名采用美国宾夕法尼亚大学“智库与公民社会计划”(Think Tanks and Civil Societies Program, TTCSPP)研究机构编写的《2020年度全球智库报告》(2020 Global Go to Think Tank Index Report)作为参考数据。《2020年度全球智库报告》统计列出了全球11175家智库,并按总体排名、大洲、主题等方式分类排序。本指标采用其中与语言文字相关的教育政策和国际关系智库作为数据收集指标,这些智库的相关研究内容都包括语言文字内容,如《2020年教育政策领先智库》(2020 Top Education Policy Think Tanks)和《2020年外交政策和国际事务领先智库》(2020 Top Foreign Policy and International Affairs Think Tanks)。研究统计了各国在上面两个榜单中的上榜机构数量,并根据机构排名高低,赋予分值。规划力中的4个三级指标的数据,主要通过政府在网站上检索而得。如有无国家层面语言文字事业规划,中国有《国家语言文字事业“十三五”发展规划》、美国有《国防语言变革路线图》《国防部语言技能、区域知识、文化能力战略规划:2011-2016》等国家层面的语言文字事业规划。国外教育规划中通常包括语言文字部分,如印度《国家教育政策2020》,因此本研究在材料中收集了包含语言文字内容的国别教育政策文件。语言文字法律法规既包括专门的语言文字法,如《中华人民共和国国家通用语言文字法》,也包括宪法或其他法律中涉及的语言文字部分,如《中华人民共和国宪法》规定“国家推广全国通用的普通话”。保护和发展少数民族语言文字能够促进语言生态和谐,提升国家语言能力,因此本研究将少数民族语言规划也纳入规划力中,如《欧洲区域或少数民族语言宪章》(European Charter for Regional or Minority Languages)等。根据实际数据收集情况,本研究将有无国家通用语言文字水平测试作为执行力的量化指标,如雅思、托福、普通话水平测试、汉语水平考试等都是各国执行力的体现。应急力指国家对海内外突发事件的语言处置力,即应急语言服务的可获取性,主要统计能够提供应急语言服务的国家,如美国有国家语言服务团、土耳其有灾害应急语言服务项目,中国在新冠肺炎疫情期间推出《抗击疫情湖北方言通》《疫情防控外语通》《疫情防控“简明汉语”》等应急语言服务。

国家语言掌控能力中通晓力的数据收集主要有2个维度:一是有无语言普查数据,主要依据联合国统计司的公开数据,目前有86个国家和地区向联合国上报了通过人口普查获得的语言数据;二是国家开设语种专业

的数量,如美国开设了310种语种教学(Looney & Lusin 2019),中国开设了101种外语语种教学。支配力的数据收集主要有3个维度:一是语言服务社团开设的数量,语言服务社团的量化研究主要采用了CSA的研究结果进行再分析。在语言领域,CSA提供全球最佳语言服务商排名。2020年,CSA对589家语言服务商进行了比较和排名,通过对各公司的财务、市场、公司总人数、业绩等数据的比较,形成了地区排名和整个行业的大排名。二是各语种语言人才储备情况,主要指世界各国开设外语翻译专业的高校数量。三是全球主要语言的可提供性,该指标考察全球30种主要语言的可提供数量。

国家语言创造能力即国家语言教育能力,它的2个三级指标——开设语种教学的高校数量和水平、语言学研究能力排名分别依据的是QS关于现代语言和语言学的高校排名,并归属于国别统计。

国家语言技术开发能力包括机器翻译力和信息挖掘力2个二级指标。在机器翻译力中,本指标体系收集机器翻译领域重要国际交流赛事的相关数据,并对结果进行排名和分析。其中典型个案代表以2020年国际机器翻译大赛的数据为基础,该赛事是全球学术界公认的国际顶级机器翻译比赛之一,也是各大科技公司、顶尖院校与学术机构展示自身机器翻译实力的较量舞台。在信息挖掘力中,信息检索能力数据以2020国际音频信息检索大赛的结果和排名为个案,对各国的信息检索能力做出评价。语音和语种识别是自然语言信息处理的重要内容,也是当前极具竞争性的新兴产业。做好语音和语种识别也是一个国家具备良好的信息挖掘力的体现。语音相关评测的种类很多,目前国际上在此领域具有代表性的高端赛事有国际语音识别大赛、国际语音合成大会(Blizzard Challenge),以及由美国国家标准技术研究院(NIST)举办的说话人识别评测(SRE)等。由于国际语音合成大会是没有官方成绩排名的交流大会,因此本指标体系选择以国际语音识别大赛和说话人识别评测2020年度的比赛结果为参考来衡量各国的语音和语种识别能力。

国家语言拓展能力中影响力的数据收集包括对各国语言在主要国际组织中使用情况的考察。如联合国大会的6种工作语言,即汉、英、法、俄、阿拉伯、西班牙语,是一个重要的衡量指标。G20成员都是工业化的发达国家或新兴市场国家,在全球化和世界经济一体化中占有重要地位,因此是否G20成员的官方语言也纳入数据收集范围。影响力还体现在国际组织的语言使用中,如果一个国家的官方语言是这些主要国际组织所使用的官方或工作语言,能够在一定程度上证明该国语言在国际交流中发挥的重要作用,如中文正式成为联合国世界旅游组织官方语言,表明中文的国际影响力持续扩大,综合国力不断增强。本研究主要依据国际组织年鉴在线数据库所发布的语言统计结果进行分类量化分析。有无国家通用语言国际传播机构、国家通用语言国际传播机构在全球布局的国家数量和机构数量也是影响力的重要衡量指标,本研究依据世界上主要语言传播机构的官网和年报的动态信息,对上述数据进行了量化统计,这些主要的语言文化传播机构包括孔子学院、歌德学院、法语联盟、英国文化教育协会、塞万提斯学院、卡蒙斯学院、日本国际交流基金会、世宗学堂等。国家语言拓展能力中传播力的第一个三级指标是国家通用语言在SCIE、SSCI和A & HCI中使用的比率、占比和排名,主要依据Web of Science数据库进行检索分析;传播力的第二个三级指标是有无国家通用语言的语言文字国际标准,依据国际标准化组织成员国官网及相关数据来考察,如中国有国际标准ISO 7098《信息与文献——中文罗马字母拼写法》,表明中文在这方面的传播力较高。^④

(二) 指数构建方法

在评价指标确立以及具体数据采集完成后,各指标的权重问题是影响评价结果的重要因素。近年来,国内外综合评价研究中所用到的赋权方法主要可分为定性的主观赋权法和定量的客观赋权法。主观赋权法是指专业的研究人员依靠自身的知识和经验,通过其主观判断确定指标,对指标进行综合评价。其优势在于操作简单,在对复杂的概念进行评定时可以集思广益,结合专家的建议从多角度考虑各种因素,并具有较强的时效性。然而,人为的主观判断在某些情况下与客观事实存在偏差。主观赋权法对专家的能力要求较高,不同专家对评价指标认识的不完全一致可能影响分析结果的准确性,并且收集专家观点所耗的时间成本和人力成本较高。主观赋权法主要以德尔非法为代表,包括选择专家、设计问卷调查、循环调查和统计量化等步骤,此外主观赋权法还包括层次分析法等。客观赋权法是指从原始数据出发,从样本中提取信息,通过一定的数学方法来计算权重。其优点在于其分析过程完全基于客观数据,有较强的统计学理论依据,可避免人为因素造成的结果偏差,能反映出众多评价指标真实信息含量差异。客观分析法包括主成分分析法和熵值法等。

根据研究目的和实际情况,本研究在确定指标体系时,主要采取主观的德尔非法;在进行指标权重计算时,主要采取客观的定量评价方法。研究以量化评价为基础,通过可获得和可核查的数据资料,采用客观分析法中的主成分分析法计算权重,经过具体核算和检验,最终得出一个能有效区分各测评对象国家语言能力的客观可靠的指数体系。

(三) 指数构建过程

国家语言能力指数构建的具体计算步骤如下。

1. 数据处理

在进行指数计算之前,我们需要将原始的指标数据进行标准化和正向/负向处理。在多指标评价体系中,由于各指标具有不同量纲和不同数量级,无法放在一起直接比较,因此通过标准化处理,将各指标值转换到同一个数量级别上,从而在保证数据之间可比性的同时方便后续的计算和分析。如本研究中有某些指标通过赋分“1”和“0”表示其工作的有无,有的指标按照具体成绩进行数值计算,分值是0~1之间的小数,因此,计算之前,我们需要将其统一转化为标准化数值。数据的正向/负向处理指通过对与指数呈反向变动的指标进行负向处理,以确保在计算指数时所有指标都正向相关,但本指数所用数据均为正相关(即数据数值越大,对指数的正向变动影响越大),因此未进行负向处理。

2. 检验现有数据是否适用主成分分析法

主成分分析法是国内外目前广泛使用的一种多元统计综合评价分析方法,但并非对所有的数据分析都适用,只有各指标(变量)之间线性相关,才可通过检验。因此,我们将各指标数据通过计算进行检验。以掌控能力中的支配力为例,首先查看本研究变量之间的相关系数矩阵,支配力的三级指标包括语言服务社团开设的数量、各语种语言人才储备情况、全球主要语言可提供性,它们的相关系数均大于0.6,表明各变量之间存在较好的线性相关性。然后,我们将变量进行KMO检验,比较变量间的简单相关系数和偏相关系数,得到的KMO值为0.723,表明各变量之间的相关性较强。最后,我们对研究数据进行Bartlett's球状检验,其p值为0。以上检验结果表明,该数据适合采用主成分分析法进行分析。

3. 提取主成分

接下来,我们在通过检验的数据中提取尽可能少的主成分,并使提取的主成分包含尽可能多的信息(数据变异解释力)。一般来说,有3种提取主成分的判断方法:(1)在因子分析中,提取特征值大于1的数;(2)提取解释总方差比例大于80%的数;(3)以碎石图中趋缓的主成分点为转折点,提取之前的点作为主成分。本研究提取满足特征值大于1的因子作为主成分,如在对规划力包含的三级指标做因子分析时,提取特征值为2.105的数作为主成分。此外,结果显示,本研究各指标提取的主成分在碎石图中均位于趋缓的转折点,并且大部分满足解释总方差比例大于80%。

4. 计算权重

最后,我们根据以上数据计算主成分的权重。之所以可以利用主成分分析法进行赋权,其原理在于各个指标的权重可以借助于各主成分的权重(方差解释率占比)和各指标与主成分的相关性计算得出。简言之,主成分分析法是根据各个指标的信息含量赋权的:若一个指标的方差解释率越高,则信息含量越高,进而赋予的权重越高。所以,本研究采取一级指标非等权重的方式计算最终结果,以尽可能反映各个指标的实际信息含量,使指数体系更加科学有效。以管理能力中的组织力为例,通过以上步骤,我们最终得出管理能力的权重是18.845%,组织力是28.473%,有无语言文字管理机构和语言文字智库机构排名各占比50%。下页表2为经过计算后的国家语言能力各级指标权重比例。从权重数据来看,掌控能力、创造能力、拓展能力的权重都超过了21%,对总体的国家语言能力影响较大。这几种能力具体对应着国家语言资源、语言教育、语言文化传播等方面的发展水平。

(四)指数体系结果

经过以上构建过程,我们得出了数据标准化处理后的193个国家的国家语言能力指数,下页表3列出了世界排名前20名和后20名的国家。^⑤

从排名情况来看,各个国家的语言能力并不相同。美国、中国和英国居于第一集团行列,他们的语言能力指数得分都大于0.5,美国更是达到了0.855,处于世界领先地位。中国排名第二,得分0.767,虽然和美国还有一定差距,但比其他国家仍然超出不少。德国、俄罗斯、法国等国家的得分在0.2~0.5之间,这些国家共40个,属于第二集团。其余150个国家,大多数是第三世界国家,得分都在0.2以下,这是国家语言能力的第三集团。从3个集团的划分可以看出国家语言能力在全世界范围内也呈现金字塔式的分布(如图1所示)。

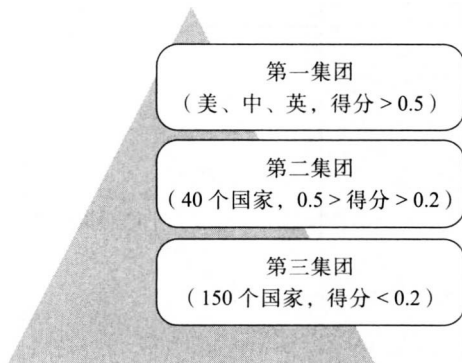


图1 世界范围内国家语言能力的分层情况

表2 国家语言能力各级指标权重

一级指标	一级指标权重	二级指标	二级指标权重	三级指标	三级指标权重
1.管理能力	18.845%	1.组织力	28.473%	1.有无语言文字管理机构	50.000%
				2.语言文字相关智库机构排名	50.000%
		2.规划力	24.278%	3.有无国家层面语言文字事业规划	32.821%
				4.有无国家层面教育规划	25.430%
				5.有无语言文字法律法规	31.431%
				6.有无少数民族语言文字规划	10.318%
		3.执行力	24.278%	7.是否设有国家通用语言文字水平测试	100.000%
2.掌控能力	21.359%	5.通晓力	50.000%	8.应急语言服务的可获得性	100.000%
				9.有无语言普查数据	50.000%
		6.支配力	50.000%	10.国家开设语种专业的数量	50.000%
				11.语言服务社团开设的数量	33.065%
				12.各语种语言人才储备情况	32.989%
				13.全球主要语言可提供性	33.946%
		7.实践力	50.000%	14.开设语种教学的高校数量和水平	100.000%
3.创造能力	21.246%	8.科学力	50.000%	15.语言学研究能力排名	100.000%
		9.机器翻译力	50.000%	16.机器翻译能力的典型个案	100.000%
4.开发能力	17.463%	10.信息挖掘力	50.000%	17.信息检索能力的典型个案	50.000%
				18.语音和语种识别能力的典型个案	50.000%
5.拓展能力	21.087%	11.影响力	50.000%	19.该国官方语言是否联合国官方语言	15.672%
				20.该国官方语言是否G20成员的官方语言	17.797%
				21.该国官方语言是否主要国际组织的官方语言	5.246%
				22.有无国家通用语言国际传播机构	19.561%
				23.国家通用语言国际传播机构在全球布局的国家数量	21.212%
				24.国家通用语言国际传播机构在全球布局的机构数量	20.511%
		12.传播力	50.000%	25.国家通用语言在SCIE、SSCI和A & HCI中使用的比率、占比和排名	50.000%
				26.有无国家通用语言的语言文字国际标准	50.000%

表3 各国国家语言能力指数排名(部分)

世界前20名国家			世界后20名国家		
排名	国家	国家语言能力指数	排名	国家	国家语言能力指数
1	美国	0.855	174	圣文森特和格林纳丁斯	0.061
2	中国	0.767	175	莱索托	0.055
3	英国	0.570	176	安道尔	0.052
4	德国	0.439	177	圣马力诺	0.051
5	俄罗斯	0.434	178	基里巴斯	0.048
6	法国	0.401	179	瑙鲁	0.047
7	澳大利亚	0.393	180	所罗门群岛	0.027
8	日本	0.388	181	尼日利亚	0.022
9	加拿大	0.383	182	赤道几内亚	0.019
10	新加坡	0.366	183	摩纳哥	0.018

续表 3

世界前 20 名国家			世界后 20 名国家		
排名	国家	国家语言能力指数	排名	国家	国家语言能力指数
11	意大利	0.360	184	索马里	0.018
12	西班牙	0.357	185	刚果(金)	0.018
13	荷兰	0.349	186	格林纳达	0.018
14	韩国	0.342	187	利比里亚	0.018
15	葡萄牙	0.332	188	几内亚	0.016
16	南非	0.314	189	也门	0.016
17	爱尔兰	0.305	190	中非	0.015
18	土耳其	0.292	191	尼日尔	0.014
19	瑞士	0.278	192	佛得角	0.013
20	墨西哥	0.269	193	图瓦卢	0.008

四、世界各国国家语言能力分析

本研究采用统计分析的基本方法,对世界上 193 个国家的各项数据进行分析,得出了国家语言能力排名总表。但本研究不是为了建构指数体系而建构,而是要通过这个体系,将国家语言能力这一概念以数据化形式直观呈现,并从数据中找出对我国国家语言能力提升有实际帮助的具体措施和建议。因此,我们对指数体系的各项数据进行分析,提出以下几点思考,希望助力我国的语言文字工作。

(一)国家语言能力与国家经济实力紧密相连

不同国家的语言能力与自身的经济发展水平有很大的关系。在第一集团的 3 个国家里,美国和中国是世界前两大经济体,而英国得益于英语在全世界范围内的普及,得分也达到了 0.57。第二集团的 40 个国家大部分是发达国家,GDP 总量和人均 GDP 排名在世界上均处于前列。而第三集团的 150 个国家,大多经济不够发达,没有足够的实力支撑国家的语言规划和不同类别语言能力的提升,国家语言能力也相对较差。

从地域分布来看,欧洲国家的语言能力普遍较强,非洲国家的语言能力整体偏弱。在排名前 50 的国家中,欧洲国家占了 27 个,超过半数,非洲国家只有 2 个;在排名后 50 的国家中,非洲国家有 26 个,而欧洲只有 2 个。欧洲和非洲的国家数量都很多,但是其各自语言能力差别很大。除了欧洲整体经济发达这一因素之外,还与其国家历史有关。非洲大部分国家是在二战之后世界范围内的民族解放运动和国家独立运动中才独立的,之前它们大多处于欧美国家的殖民之下,本身的语言和文化遭受了严重的破坏,而独立之后的语言自主并不能完全实现,因而国家语言能力难以提升。

从语种分布来看,以印欧语系语言为主的国家,其语言能力偏强。印欧语系包含了 400 多种语言,其中既有英语、法语等世界范围内有强大影响力的语种,也有阿尔巴尼亚语、亚美尼亚语等非通用语种。它是当代世界分布区域最广的一个语系,使用者几乎遍及整个欧洲、美洲、大洋洲,还有非洲和亚洲的部分地区。从 15 世纪开始,随着欧洲殖民势力不断扩张,欧洲语言陆续传播到世界其他许多地区,这是印欧语系分布如此之广的直接原因。在三四个世纪的殖民侵略过程中,被征服地的语言文化往往受到严重破坏,只能接受殖民者所带来的语言,因此,以印欧语系语言为官方语言的国家,其语言能力在世界范围内整体上处于领先状态。可见,要提高一个国家的国家语言能力,首先要保证自身历史文化的传承和发扬,重视本国语言的使用、保护、推广。

(二)中国的国家语言能力尚存许多短板

中国的国家语言能力指数为0.767,在世界上仅次于美国,居第二位。从整体指数来看,中国国家语言能力与中国的综合国力基本匹配。但是分析各分项能力的数据,可发现中国和美国、英国这两个以英语为主导语言的国家相比,还有不少方面存在差距,需要在语言研究、语言教育、语言产业和学术话语方面加强建设。

中国相比于美国、英国的优势在于语言人才的储备、语言技术竞赛的名次和国际中文教育传播机构的数量。这些优势有的是由中国国情决定的,如中国是世界上人口最多的国家,也是少数民族语言、跨境语言资源丰富的国家,而且,中国从晚清提倡“睁眼看世界”以来,就非常重视外语教育,因而语言人才的储备量很大,目前要继续提升的是非通用语人才、高端语言人才的培养;有的是由于中国政府和社会对语言教育的重视而形成的,如语言技术竞赛,中国在信息技术、人工智能发展的过程中,努力创新,攻克技术堡垒,因而在国际相关赛事中成绩名列前茅。这些优势既有助于提升中国国家语言能力,更是中国占领信息技术高地、提高科技水平的重要环节,需持续加强。

但从进一步加强国家语言能力的角度来看,中国要提高自身的国家语言能力,在语言研究、语言教育、语言产业、语言传播方面还需要持续追赶。语言研究和语言教育是创造能力的重要体现,这一方面美国 and 英国排名世界前2名,中国排第3。美国国力强大,其国家主导语言也在世界范围内普遍使用,境内又由于移民众多而拥有丰富的多语种资源,同时在全球化背景下,美英两国在科技、教育方面的传统优势确保了其语言研究、语言教育能力在世界上遥遥领先。2020年QS语言学高校排名前10名中,美国占了7所,其余3所是英国高校。中国排名最高的北京大学只在第16位,前50名中中国高校只有5所,且除了北大之外都是港澳台地区高校。因此在语言学科发展方面我们和世界前列的国家还有很大差距,要通过教育改革、资金投入加大追赶力度。

语言产业也是中国相对于美国的短板。美国的第三产业非常发达,语言服务产业也同样如此。从CSA Research对全球语言服务商的排名来看,前100名中,美国的语言企业最多,所占市场份额也最大,英国、德国、日本紧随其后,而中国只有1家企业,排在第19位。世界一体化进程中,我们越来越意识到,语言不仅是工具,是文化,也是信息,是资源,发展语言服务产业是未来发展经济的必然需求。所以,我们要重视语言产业方面的差距,尽快追赶。

在拓展能力方面,中国的国际中文教育事业迅速发展,大大提高了中国的影响力。但是它的另一个指标传播力,中国相比美国还落后较多,即中国的国家通用语言中文在世界学术语言体系中占比太小,2020年Web of Science数据库中美国的发文数量为474 012篇,中国虽然排名第2,但是数量仅为281 596,只有美国的60%左右。而且中国学者发表的文章很多还是用英文写作的,因此,中文在世界学术发表体系中的实际占比更小。据学者统计(李宇明2020),2010-2019年在自然科学领域,中文排名第4,但使用比例只有0.28%;在人文社科领域,中文排名更加靠后,到了10名开外。近些年来,学界逐渐意识到,学术语言的使用关乎知识创造体系的话语权,关乎一种语言表达知识创造知识的能力,我们不应助长英语一家独大的势头。所以,重视学术领域的语言使用问题,采取措施切实提高中文的科学表达能力也是语言文字工作的重要方面。

(三)在资源有限的情况下,要优先发展掌控能力、创造能力和拓展能力

我国正处于世界百年未有之大变局,在国际上面临政治、经济、文化等各方面的压力,因此,提高自身实力

是应对一切内外问题的根本。在资源有限的情况下,可以优先发展国家语言掌控能力、创造能力、拓展能力。因为从权重指标体系来看,这几个方面对国家语言能力的整体提升有较大帮助,可以快速增强我国的国家实力,更好地维护国家利益。

掌控能力包括语言普查数据、语言人才库建设等指标,实际对应着国家对自身语言资源的掌握和运用能力。邓小平说过“科学技术是第一生产力”,^⑥而习近平进一步指出“人才是第一资源”。^⑦语言是由人掌握、运用的,语言能力必定和一个个具体的人联系在一起。对于国家来说,语言资源的掌握和控制,有助于国家在经济、政治、外交、社会管理等各方面的能力提升。

创造能力包括高校语言教学能力和研究能力两个指标,这是一个国家语言教育水平的集中体现。语言教育既是个人语言能力提升的重要途径,也是国家语言能力水平的实现方式,其指数权重在国家语言能力中占比较大。

拓展能力包括影响力和传播力,其二级指标主要面向国际使用和拓展。在世界进入全球化时代之后,不同语言在世界融合交流过程中的发展速度并不相同,有的越发强势,有的逐渐萎缩甚至濒临消亡。而从指数来看,拓展能力对一个国家的国家语言能力来说至关重要,因此国家在制定语言规划和战略时,一定要平衡好国家内部不同语言之间的关系,提高国家主导语言在国际社会的影响力。

如表4所示,国家语言能力的具体排名情况也佐证了我们的推断。德国和澳大利亚都是发达国家,同属于第二集团,德国排名第4,澳大利亚排名第7。但这两个国家的管理能力和开发能力排名都不高,都在10名之外。使得他们整体能力提升的正是上述掌控能力、创造能力和拓展能力。德国这三项能力分别排名6、4、7,澳大利亚是4、5、9。这些能力方面的优势使德国和澳大利亚整体居于前列。韩国正好相反,其管理能力排名第5,但是掌控能力和拓展能力分别是78和12,最终其整体国家语言能力也掉到了14。由此可见,要提升国家语言能力,就要在这3个方面不断加大投入,因此,中国在语言资源建设、语言学科建设、国际中文教育、语言产业方面应当加大投入,继续努力,以提升国家语言能力。

表4 部分国家分项能力排名对比

国家	总排名	管理能力排名	掌控能力排名	创造能力排名	开发能力排名	拓展能力排名
美国	1	1	1	1	2	2
中国	2	3	2	3	1	1
英国	3	2	3	2	6	4
德国	4	12	6	4	15	7
俄罗斯	5	7	7	10	8	5
法国	6	4	71	12	11	3
澳大利亚	7	43	4	5	21	9
加拿大	8	25	9	11	4	8
荷兰	13	10	22	13	18	19
韩国	14	5	78	7	14	12

总而言之,国家语言能力是一个国家整体实力的体现,是国家软实力的重要组成部分。我们对世界各国的国家语言能力进行指数分析,希望为发现不同国家之间的能力差异及原因、找出我国国家语言能力的优势和短板、明确我国国家语言能力的发展方向和路径提供有价值的参考。

五、结语

国家语言能力指数体系是提升国家语言能力建设效率和有效性的重要一环。目前还未有文献对国家语言能力进行专门量化研究并进行排名。本研究具有一定的开创性,在学术研究上,指数体系的建立能进一步完善国家语言能力的理论建构;在对策性研究上,指数体系的建立能够比较出我国与世界其他国家在国家语言能力建设上的优势和不足,为提升我国的国家语言能力提供更多的参考数据和资政建议。

本研究虽然已经达到了预期目的,得出了世界大多数国家的国家语言能力指数体系,但是在具体研究过程和方法上还有可继续加强和改进的空间。比如,本研究虽采用定量和定性相结合的方法,但定性研究主要是应用德尔菲法进行指标的确定,而指标的权重计算是依据客观赋权的主成分分析法进行赋权,如能在指标权重方面也用德尔菲法进行主观赋权计算,并与客观赋权的结果相比较,依据两者结果进行指标或数据调整,指数体系的科学性会更高。但在疫情的大背景下,由于此方法对选择专家的广度、数量、国别背景、学术积累和调查频度都有较高要求,这种操作在当前现实中很难实现,因此本研究在权重分析时还是采用主成分分析法。此外,本研究采用主成分分析法,属于客观方法,但多种客观方法之间的量化结果有无差异,如果差异不大,说明指标体系设计合理,否则说明指标体系需要调整。此方面工作需要在后续研究中落实和验证。

有鉴于此,本研究将在指数指标的优化、数据搜集的完善、研究方法的多元和实践服务功能的加强等方面继续努力,争取将国家语言能力指数体系打造得更科学、更完善。

本文的部分内容修改自即将出版的《全球指数2021》(外语教学与研究出版社,2021)“国家语言能力指数报告”一节。感谢北京外国语大学区域与全球治理高等研究院“全球指数项目组”的资助。感谢孙莎博士在研究方法和数据处理上提供帮助,感谢卜凡尧老师等提供团队支持。文中错误之处,概由本人负责。

注释:

①参见外语教学与研究出版社即将出版的《全球指数2021》中国国家翻译能力指数课题组的相关数据,下同。

②CSA(Common Sense Advisory, Inc.)指美国的卡门森斯咨询公司,它在很多领域提供全球市场分析,其报告经常被各个领域采纳。

③QS(Quacquarelli Symonds)指夸夸雷利·西蒙兹公司,是英国一家国际高等教育咨询公司。

④限于篇幅,本研究的具体数据来源和分析参见《全球指数2021》。

⑤因本文篇幅有限,所有国家的详细排名可参见《全球指数2021》一书第四部分“国家语言能力指数报告”。

⑥1988年9月5日,邓小平在会见捷克斯洛伐克总统胡萨克时,提出了“科学技术是第一生产力”的重要论断。

⑦2018年3月7日,习近平在两会期间参加广东代表团审议时强调“发展是第一要务,人才是第一资源,创新是第一动力”。

参考文献:

[1]李宇明 2011《提升国家语言能力的若干思考》,《南开语言学刊》第1期。

- [2]李宇明 2020《中文怎样才能成为世界通用第二语言》,《光明日报》1月4日第10版.
- [3]魏晖 2015《国家语言能力有关问题探讨》,《语言文字应用》第4期.
- [4]文秋芳 2016《国家语言能力的内涵及其评价指标》,《云南师范大学学报(哲学社会科学版)》第2期.
- [5]文秋芳 2019《对“国家语言能力”的再解读——兼述我国国家语言能力建设70年的成就及不足》,《新疆师范大学学报(哲学社会科学版)》第5期.
- [6]文秋芳,苏静,监艳红 2011《国家外语能力的理论构建与应用尝试》,《中国外语》第3期.
- [7]文秋芳,张天伟 2018《国家语言能力体系构建研究》,北京:北京大学出版社.
- [8]赵世举 2015《全球竞争中的国家语言能力》,《中国社会科学》第3期.
- [9]Brecht, R. D. & A. R. Walton. 1993. National strategic planning in the less commonly taught languages. NFLC Occasional paper (ED367184).
- [10]Brecht, R. D. & W. P. Rivers. 2000. Language and National Security for the 21st Century: The Role of Title VI/Fulbright-Hays in Supporting National Language Capacity. Dubuque, IA: Kendall/Hunt Publishing Company.
- [11]Brecht, R. D. & W. P. Rivers. 2005. Language needs analysis at the societal level. In M. H. Long(Ed.), Second Language Needs Analysis, 79–104. Cambridge: Cambridge University Press.
- [12]Brecht, R. D. & W. P. Rivers. 2012. US language policy in defence and attack. In B. Spolsky(Ed.), The Cambridge Handbook of Language Policy, 262–277. Cambridge: Cambridge University Press.
- [13]Looney, D. & N. Lusin. 2019. Enrollments in languages other than English in United States institutions of higher education, summer 2016 and fall 2016: Final report. <https://www.mla.org/content/download/110154/2406932/2016-Enrollments-FinalReport.pdf> [2020-12-01].

National Language Capacity Evaluation Indexes and Their Applications

Zhang Tianwei

Abstract: National language capacity is an integral component and important manifestation of a country's comprehensive national power. To evaluate the national language capacity of a country, this study proposes a three-level evaluation index system, developed with reference to objective data and a multi-index comprehensive evaluation approach. The evaluation system consists of 5 first-level abilities, 12 second-level abilities, and 26 third-level specifications, whereby the first-level indexes, including management ability, control ability, creation ability, development ability, and expansion ability, perform paramount roles. Using these indicators, the language capacities in various countries can be quantified and compared. With further data analysis and verification concerning the system, this study reveals that the language capacity of a country is closely related to its economic strength. The values of national language capacity in the individual countries over the world form a pyramid structure, with the United States, China, and the United Kingdom occupying the top. Among the five first-level indexes, the control ability, creation ability and expansion ability have greater impacts on the overall national language capacity of a country. When it comes to the national language capacity of China, this paper proposes that efforts need to be made in aspects such as language research, language education, language industry, and academic discourse in order to enhance its overall national language capacity.

Key words: national language capacity; language policy; index system; Principal Component Analysis (PCA) method