

珍爱中华语言资源

李宇明

国家语言文字工作委员会、教育部语言文字信息管理司

世界经济正在走向一体化，然而世界的文化应当多元化。文化是各个民族的精神家园，绝不能被经济一体化所破坏。这是世界多数人的认识，也是中国大多数人的共识。中国共产党十七大报告指出：“当今时代，文化越来越成为民族凝聚力和创造力的重要源泉、越来越成为综合国力竞争的重要因素……”“弘扬中华文化，建设中华民族共有精神家园。中华文化是中华民族生生不息、团结奋进的不竭动力。……加强中华优秀传统文化教育，运用现代科技手段开发利用民族文化丰厚资源。加强对各民族文化的挖掘和保护，重视文物和非物质文化遗产保护，做好文化典籍整理工作……”语言是文化的重要组成部分，是文化最为重要的载体与根基，维护文化的多样性，离不开维护语言的多样性。

一、语言是国家的重要资源

人类文化有 3 种载体形态：（1）由实物承载的，如建筑、地下出土的文物等；（2）由文献记载的；（3）由口语涵载的。第一类文化需要语言来阐释，后两类文化都是由语言（包括文字）负载的。由此来看，离开语言来谈文化的多样性，几乎不可能。

百年来的中国语言规划，常把语言作为“问题”看待，重点解决的是语言给社会带来的麻烦。比如语言分歧带来的交际障碍，学习某种较为流行的语言（包括外语）而付出的巨大社会成本和经济成本，怎样和谐处理语言间、方言间的关系，等等。这当然是必要的，但并不全面。对社会来说，语言不仅是“问题”，而且也是资源。由于将语言作为资源看待的意识相对薄弱，对语言资源的保护、开发不够自觉，措施不力，加上现代社会条件的剧烈变化，造成了语言资源的严重流失，并由此引发出一系列语言问题。树立语言资源观念，了解中华语言资源的基本状况，制定切实可行的语言资

源保护、开发措施,已经成为当今中国语言规划的必务之事、当务之急。

语言首先是语言的资源。每种语言都是一种特殊的语言样品,具有其他语言无法代替的语言学上的认识价值。例如,汉语的平上去入的声调系统及其变调特点,世间少见;哈萨克语有关马的毛色的词语达 3500 多个,形容骏马的词语有 100 多个,有关马的其他特征的词语有 600 多个,非马背上的民族难有这样的语言奇观。语言在不断的发展演变之中,语言成分的借用,历史元素的承袭,是语言发展不可忽视的因素。通过不同语言样品的比较分析,可以帮助再现语言的历史,探索语言接触的各种情况。

语言是文化的资源。民族的语言与文字不仅表现着民族智慧,而且常常成为民族图腾般的象征。而且,80%的文化是通过口语和文字传留下来的。世界上没有文字的民族多于有文字的民族,他们的文化主要通过口语流传下来。如我国的鄂伦春族、羌族、东乡族、保安族等。在中国首批公布的 518 项非物质文化遗产保护名录中的传说、故事、号子、歌谣、戏曲等,都无不牵涉到语言文字的问题。

更不可忽视的是,当今时代,一些语言职业和语言产业逐渐形成,语言已经进入经济和高科技领域,成为经济发展的重要资源。如英语教育产业、语言翻译产业、文字速录师职业、计算机字库提供商、语言文字信息处理软件产业等。随着信息时代的发展,语言作为经济资源的性质会体现得越来越明显,其经济意义会越来越显著。

二、急需开展中国语言普查

看待问题的角度往往决定认识问题的深度。当我们具有语言资源意识、回眸审视中华语言资源之时,就会突然发现:

一是对中国语言状况了解不够。中国究竟有多少种语言?各语言又有多少方言?这些语言与方言的分布区域、使用人口、功能活力如何?无人能清晰地回答这些问题。

二是对语言资源的保护和开发利用不够。例如,许多语言职业和语言产业没有得到重视;许多语言的语料库质量不高,没有发挥其应有的作用;对于现实世界和虚拟空间的语言资源保护没有明确的法规制度,等等。

三是语言资源流失严重,许多语言及方言已衰落或濒危。许多儿童已不能熟练掌握方言,母语水平滑坡,一些少数民族青少年甚至放弃了母语。中国的一些语言和方言已经处于濒危状态甚或消亡。当然,语言资源流失、一些语言濒危,是当今世界普遍存在的问题。目前世界上有 5000~6000 种语言,90%的语言只被 10%的人使用,如果不采取有效措施,这 90%的语言在百年或几百年内可能消亡。语言消亡意味着人类将失去不可复得的语言样品,将失去不可再生的文化基因,将永远失去一些历史记忆。有鉴于此,在新世纪我国应当尽快开展语言普查工作,以了解基本的语言国情,保护和开发中华语言资源。

20 世纪 30 年代以来,一直有学者对中国语言和方言的种类、分布、结构特点、谱系关系等进行调查研究,出版了若干汉语方言和民族语言的系列丛书和词典,绘制了中国语言的分布图等。作为国家组织的大型语言普查有两次:

第一次是根据国务院指示于1956年开始的汉语和少数民族语言调查。当时主要调查的是语言结构,旨在为方言区学习普通话服务,为创制民族文字服务,没有重视语言运用调查;而且限于当时条件,调查较为粗糙。当时调查了1849个县市的汉语方言,并组成7个民族调查队,调查了主要民族地区的语言。这次大规模的调查,培养了一支从事方言和民族语言调查的专业队伍,发展了语言学学科,积累了经验,了解了一些语言国情。

第二次是1999年教育部等11个部委联合开展的中国语言文字使用情况调查。当时主要关心的是语言文字的使用现状,但是没有采录语料,不能了解语言及方言的话语实态,也不能直接解决语言保护等问题。那次调查涉及全国1063个县(市、区),被调查对象47万多人,积累了田野调查和问卷调查的丰富经验。

现在一些有识之士正在呼吁开展新世纪的语言普查。果若能成,这次语言普查就应兼取前两次之长,弄清中国语言及其方言的种类、分布区域、运用人群及使用变化状况等,建立起可永久保存的中国语言多媒体语料库及相关数据库,绘制细致的可传至后代的多媒体语言地图。

语言普查是一项工程浩繁、功在当代而惠及千秋的事业,应由国家立项,调动全国的语言学力量、社会各界积极支持配合,这是不言而喻、无须论证的。就普查工作而言,应做到以下几点:

第一,要对国内外已有的相关成果进行梳理、整合。这些成果,包括各种语言和方言的调查报告、辞书和语言地图、记录的音档,也包括地方志中的相关内容。通过这些成果的收集梳理和整合集成,为新的语言普查提供工作基础。当然,这本身就是一项浩瀚的文化工程。

第二,在语言普查、成果整理的一系列过程中,应充分利用现代语言处理手段和各种信息技术与设备。如用于语言普查的录音、转写技术,地理信息系统(GIS)技术和卫星定位(GPS)技术,用于语料存储和检索的多媒体数据库技术,用于语料分析、标注的软件技术,以及录音笔、计算机、服务器、大容量存储器等多种硬件设备。

第三,也是最为重要的,就是要制订科学的普查方案。参考已有的研究成果,根据语言国情、技术条件和工作目标,对普查内容、普查方式、采录标准、人员培训、资料汇总、成果管理、项目监督等一系列问题进行科学论证,形成方案。并选取有典型意义的县(旗)作普查试点,在试点中积累经验、完善方案。语言普查以县(旗)为单位布点为宜,语言情况复杂的地方加大调查点的密度。除调查清楚每个调查点语音、词汇、语法等语言结构的情况之外,还应录取实态语音材料。

第四,重视数据的保存、开发与共享。普查数据,特别是语言、方言的实态语料数据要多备份,异介质、异地方安全保存,重视对这些资源的行政开发和学术开发,为国家的各行各业和学术的发展提供数据支撑。通过网站、开放实验室、语言博物馆等方式,实现数据共享,最大限度地让这些数据造福学界、国家与人类。

第五,重视人员培训。语言普查的专业性很强,也很辛苦,需要对普查人员进行培训,以保证普查使用统一的方式、技术和工作规范。培训的专业内容主要包括听音

训练、用国际音标记音、语言调查字表和词表的使用、计算机辅助田野调查技术、语言数据的建库技术，等等。

三、结 语

对文献文化的保存、整理与利用，中国有着悠久的传统，历来受到重视。国务院正组织开展全国第三次文物普查，这是对实物文化进行收集、整理的新世纪的国家行动。但是，对于由口语涵载的文化，常被忽视。语言普查如果能够实施，并且重视文化因素的采集，如用结构访谈和采风的方式采录各地有特色的文化语料，增加语料的文化含量等，其实也就是对中华口头文化的大规模收集，意义宏大而深远。

语言及其方言是国家不可再生的、弥足珍贵的非物质文化遗产，也是构成文化多样性的前提条件。通过语言普查建立中华语言的语料库，其实就是建立中华文化的知识库、“基因库”，是语言资源保存与开发的重大而有效的举措。中国是世界上语言资源最丰富的国家之一，保护好中国的语言资源，也是中国对人类的贡献。

作者:

李宇明

作者单位:

国家语言文字工作委员会;教育部语言文字信息管理司

本文链接: http://d.g.wanfangdata.com.cn/Conference_7771981.aspx