

地理语言学

《汉语方言地图集》前言^{*}

曹志耘

提要：本文介绍《汉语方言地图集》课题的背景、工作过程、调查研究情况以及该地图集的基本内容。本课题实地调查了全国汉语地区 930 个地点，绘制了汉语方言语音、词汇、语法特征地图 510 幅。《汉语方言地图集》是世界上第一部在统一的实地调查的基础上编写的全面反映 20 世纪汉语方言基本面貌的原创性语言特征地图集。

关键词：汉语方言；方言地图；地理语言学

^{*} 《汉语方言地图集》，曹志耘主编，商务印书馆即将出版。

一 背景

汉语方言历史悠久，纷繁复杂，其差异性甚至远远超出欧洲许多语言之间的差异。汉语方言是当今世界上最古老、最丰富的语言文化资源之一，是中国文化和人类文明的宝贵遗产。20 世纪中叶以来，随着我国经济、文化、教育和交通事业的迅速发展，汉语方言已经发生了巨大的变化，各地方言中的特殊现象正在急剧消失，一些弱势方言甚至在走向衰亡。全面、科学地调查描写汉语方言的传统面貌，及时抢救和保存汉语方言资料，保护民族语言文化遗产，已经成为我国语言学界一项迫在眉睫的历史使命。

方言地图是语言学研究中的一种重要手段，在描写、展示和保存大面积地区、众多方言现象的面貌和分布状况方面具有不可替代的作用。自 19 世纪后期以来，许多国家如法国、德国、瑞士、意大利、丹麦、英国、捷克、美国、加拿大、日本、菲律宾、泰国等都投入巨大力量，调查、编写、出版了本国的方言地图集，例如吉叶龙 (J. Gilliéron) 的《法国语言地图集》(1902~1910 年)，芮德 (F. Wrede) 等的《德国语言地图集》(1926~1956 年)，雅伯尔格 (K. Jaberg)、依德 (J. Jud) 的《意大利瑞士语言地图集》(1928~1940 年)，国立国语研究所的《日本言语地图》(1966~1974 年) 和《方言文法全国地图》(1989~2006 年)，等等。

早在 20 世纪 20 年代，林语堂、刘复等先辈已经提出在调查的基础上编绘汉语方言地图的宏伟设想。此后，赵元任先生主持了多次大规模的方言调查，并首次在《湖北方言调查报告》(1938 年成书，1948 年出版) 里收入 66 幅方言地图。20 世纪 40 年代，贺登崧 (W. A. Grootaers) 先生把西方地理语言学引入中国，在大同、宣化等地调查方言，并发表了多篇附有方言地图的论文。不过，由于历史的原因，长期以来始终未有人实现编写全国汉语方言地图的梦想。直到 20 世纪 80 年代，中国社会科学院和澳大利亚人文科学院共同编写出版了《中国语言地图集》(1987 年)，包括总图 5 幅、汉语方言图 16 幅、少数民族语言图 14 幅、

共 35 幅地图。这是第一部汉语方言和中国少数民族语言的分布和分区地图集，具有开创意义。20 世纪 90 年代以来，岩田礼先生等日本学者致力于汉语方言地图的编写工作，从 1992 年至 2007 年，以“研究成果报告书”的形式印行了《汉语方言地图（稿）》等 6 种地图集，每种收有几十幅汉语方言特征分布图，均据已发表的方言材料编写而成。日本版的汉语方言地图集为编写汉语方言特征图作出了有益的探索和重要的贡献，不过由于材料所限，难以全面、系统地反映汉语方言的基本面貌。

方言特征分布图（简称“特征图”）是方言地图的基本形式，同时也是地理语言学、历史语言学、社会语言学、语言类型学等语言学科的研究基础。长期以来，国内外学术界热切期待一部全面的汉语方言特征地图集的出现。在《法国语言地图集》问世一个世纪后的今天，中国已经没有任何理由继续留着这一巨大的空白了。尽管深知《汉语方言地图集》调查、编写工作的艰巨性，但历史的使命感和科学创造的诱惑力，驱使我们义无反顾地投入《汉语方言地图集》的工作。

二 工作过程

自 1997 年以来，我们多次向有关部门提出开展“汉语方言地图集”课题研究的申请计划。2001 年 12 月 7 日，“汉语方言地图集”获得批准成为教育部人文社会科学研究“十五”规划项目。2001 年 12 月 12 日，“汉语方言地图集”获得批准成为北京语言大学“十五”规划项目。研究工作自此正式全面启动，整个过程大致分为以下几个阶段。

2.1 前期准备（2001.12~2002.11）

在此期间，举行了多次研究课题论证会和调查条目论证会，集中精力完成了选定调查点、拟定调查条目、制定工作规范的工作，并分别对全国各大方言进行了 20 次试验调查。在反复调查、反复修改之后，于 2002 年 10 月印制了专门供本课题调查使用的《汉语方言地图集调查手册》（后于 2003 年 2 月由澳门语言学会正式出版），于 2002 年 11 月印制

了《汉语方言地图集工作手册》。

2.2 调查 (2002.12~2006.12)

整个调查过程共持续了4年时间。第一个点于2002年12月5日开始调查,最后一个点于2006年12月19日完成调查。有关调查的一些具体情况参看下文“三 调查和研究”。

2.3 资料整理、数据库建设和图目编写 (2005.3~2006.12)

(1) 录入

有37个调查点的材料是调查时用电脑直接记录在电子文件里的,其余893个点的材料都是先记在《汉语方言地图集调查手册》上,然后再录入电脑。调查材料的录入工作,大约有70%调查点是由调查人自己负责录入的,其余约30%调查点由课题组聘用的录入员录入。每个点的材料按语音、词汇、语法分别录入到统一制定的3个Word表格里面,重点点还把《方言调查字表》的材料录入到另一个Word表格里面。国际音标一律使用IpaPanADD3.0版字体。

(2) 校对

从2006年5月开始,课题组先后组织36人(包括课题组成员和在读研究生),对已录入调查材料的电子文件进行校对。校对实际上还包括对材料的核实以及统一用字等工作。除了录入人自己所作过的校对以外,课题组的校对工作通常分为一校、二校和终校,但也有不少点实际上进行了四校甚至五校。为了规范校对工作,我们编印了《汉语方言地图集校对手册》,其中包括用于统一用字的“用字规范”。

(3) 数据库建设

完成调查材料的终校后,把所有Word表格转换成Excel表,然后再导入数据库。至2006年12月底,建成全部930个调查点的“汉语方言地图集数据库”。

(4) 图目编写

“图目”是指地图条目。编写图目的工作从2005年上半年着手进行,

经过一年多时间的讨论和修改，至 2006 年底，基本上确定了语音、词汇、语法各卷的图目。

2.4 分类和画图 (2007.1~2007.12)

“分类”是指对图目的方言材料进行归类，分类的结果即体现为图例里的各种类型。分类工作的程序包括从数据库提取材料、编写分类表、提交讨论、修改、审定等，所有图分类表都由主编终审。分类表终审后，在 Excel 表里填写分类信息（类码），再导入数据库。由画图人利用地理信息系统绘制方言地图。初图画出后，经过校对、核查和修改等程序，最后定稿。为了规范分类和画图阶段的工作，我们编写了《汉语方言地图集编图手册》，其中包括详细的“分类规范”“画图规范”等。

从 2007 年 1 月开始进入分类阶段，大部分的分类工作集中于上半年完成，但分类表的终审工作一直持续到 2007 年 12 月。试验画图工作从 2005 年就开始了，经过长期的摸索和改进，逐渐确定了样图的形式。语音卷的画图工作主要集中在 2007 年 9~10 月进行，全部画图工作于 2007 年 12 月底完成。

此外，自 2005 年以来，先后举行了 6 次地图集编写工作会议，以及几十次编委会的讨论会，对调查、资料整理、数据库建设、分类、画图等各个环节的工作进行了反复讨论。

三 调查和研究

自 20 世纪初以来，汉语方言的调查研究取得了巨大的成就，积累了非常丰富的资料。但这些资料年代不同，详略各异，发音人身份多种多样，调查人的水平和处理方法也有很大差别。与此同时，还有很多地区的方言没有经过调查。显然，现有的资料无法满足编写一部全面、系统、具有可比性的《汉语方言地图集》的要求。因此，本项研究的所有地点都由课题组成员自己专门调查，所有材料都由课题组统一处理。

整个调查和材料分析处理的工作极其繁复，遇到的困难和问题不计其数，因篇幅所限暂不展开讨论，这里只交代调查研究工作中的一些基本情况。

3.1 调查点

截至 2006 年底，中国大陆县级行政单位（包括县、自治县、旗、自治旗、县级市、市辖区等）共有 2864 个（其中甘肃省嘉峪关市、广东省东莞市和中山市、海南省三亚市 4 个地级行政单位下无县级，暂计为县级），香港、澳门和台湾地区缺相应数据。我们从其中的汉语地区选取了 930 个地点（其中包括少量“一县二点”的情况）作为调查点。

考虑到各地方言复杂程度的不同，东南方言地区原则上每个县（指县级行政单位，下同）设 1 个点，官话和晋语地区约每 3~4 个县设 1 个点。但有以下例外情况：

(1) 近几十年中国县级行政区划变动很大，而明清时代的府县与汉语方言的分布具有密切的关系，因此，在东南方言地区，如近几十年有撤县或设县情况的，设点时一般以清代的行政区划为准。

(2) 位于城区或主要位于城区的所有市辖区合为一个县级行政单位对待，位于城区之外的市辖区视为独立的县级行政单位。

(3) 一个县里有两种区属不同而且势力相当或都很重要的方言，设 2 个点。例如福建省浦城县，设吴语和闽语 2 个点。畲话散布于福建、浙江等地的闽语、吴语区里面，为了反映畲话的情况，在福建省宁德县设闽语和畲话 2 个点，在浙江省景宁畲族自治县设吴语和畲话 2 个点。

(4) 与东南方言“直接接壤”的官话方言县均选作调查点，不受“3~4 县设 1 点”的限制。这是为了观察方言交界地区的相互影响情况。

各省市区的调查点数及其方言区属的情况见表 1，各调查点的具体情况见附录，各调查点在地图上的位置见各卷前面的“调查点图”。

表 1 调查点数

省级	县级 数	调查 点	官话	晋语	吴语	徽语	闽语	粤语	客家 话	赣语	湘语	平话	其他
安徽	105	47	22		10	8				7			
澳门		1						1					
北京	18	3	3										
重庆	40	7	7										
福建	85	69			1		59		8				1 畲话
甘肃	87	15	15										
广东	123	92					16	43	27				6 土话
广西	109	66	3					15	7			36	5 土话
贵州	88	17	17										
海南	21	14					13						1 儋州话
河北	172	31	23	8									
河南	159	26	23	3									
黑龙江	128	13	13										
湖北	102	32	23							9			
湖南	122	92	15						3	15	43	1	4 乡话, 11 土话
吉林	60	10	10										
江苏	106	40	21		19								
江西	99	87	2		3	3			21	58			
辽宁	100	12	12										
内蒙古	101	11	5	6									
宁夏	21	7	7										
青海	43	5	5										
山东	140	30	30										
山西	119	24	5	19									

续表

省级	县级数	调查点	官话	晋语	吴语	徽语	闽语	粤语	客家话	赣语	湘语	平话	其他
陕西	107	21	16	5									
上海	19	11			11								
四川	181	19	19										
台湾		15					12		3				
天津	18	1	1										
香港		2						1	1				
新疆	99	11	11										
云南	129	15	15										
浙江	90	84			77	4	2						1 畲话
小计			323	41	121	15	102	60	70	89	43	37	29
合计	2791	930	364		566								

西藏自治区无调查点，故未列入。

其中有 91 个是重点点，按方言区片选定。重点点的调查内容比其他点多出“两字组连调表”和《方言调查字表》，重点点的材料供编写《方言语料对照集》之用。

所有 33 个省会级城市和 3 个非省会级城市的方言区代表点苏州、厦门、梅州均设为调查点，只调查城里或城郊方言，不调查乡下。其余 894 个调查点一律只调查乡下方言，不调查城里（但有少数调查点调查的是城乡交界地区）。

在官话和晋语地区设调查点没有特别的标准，但尽量使设为调查点的县在地理上互不相连。在一个县里到底调查哪个村，由调查人在当地根据具体情况决定，一般要求是尽量选位于县境中部的村，该村使用的方言在本县内较为通用，不选方言岛。

3.2 发音人

本研究要求以 1931～1945 年之间出生的男性为调查对象。发音

人必须在当地出生和长大，未在外地长住，能说地道的当地方言，有一定文化（小学或中学程度为宜）。一个地点以 1 名发音人为主，不要求调查年龄差异。但由于各种原因，也有少数地点的发音人在年龄、性别方面超出了上述要求。

发音人的简况见表 2，详细情况见附录。

表 2 发音人

	出生年					性别		文化程度			职业			
	1930 年前	1931 ~ 1935 年	1936 ~ 1940 年	1941 ~ 1945 年	1946 年后	男	女	小学 及 以下	中学	大专 及 以上	农民	教师	干部	其他
人数	86	139	232	309	164	921	9	233	578	119	307	397	98	128
百分比	9.2%	14.9%	24.9%	33.2%	17.6%	99%	1%	25.1%	62.2%	12.8%	33.0%	42.7%	10.5%	13.8%

一个点如有多个发音人，以主要发音人为统计对象。1946 年后出生的 164 人中，1946~1950 年出生的有 131 人，1951~1955 年出生的有 20 人，1956 年后出生的有 13 人。“干部”包括除村干部以外的各级官员。“其他”包括工人、商人、医生等。

3.3 调查条目

调查条目使用课题组专门编写的《汉语方言地图集调查手册》。该手册包括调查点和发音人信息，以及音系、字表、“两字组连调表”、词汇、语法等内容。其中单字音、词汇、语法的调查条目为单字 425 个，词汇 14 类 470 条，语法 65 类 110 条，共计 1005 个条目。所有调查点都需要调查记录除“两字组连调表”外的所有项目，重点点还要记录“两字组连调表”和中国社会科学院语言研究所编的《方言调查字表》（但不再单独记录《调查手册》里的“字表”部分）。

设立调查条目的主要原则是：（1）反映重要的地域差异；（2）反映重要的历史演变。具体而言，选定的单字必须能反映重要的古今语音演变关系，即具有“代表字”的性质。选定的词汇应是语言中较重要的概

念，同时具有重要的地域差异。语法条目则涵盖了普通话语法系统中的主要形式，同时还收录了一些比较重要的只见于方言的语法形式，例如前缀“阿”、“数量+头”、量词定指、后置成分“添”“过”“先”等。为了保证调查材料的可比性，《调查手册》给许多条目加了必要的“限定条件”，例如“张”字要求调查做量词时的读音，不调查做姓氏的读音，又如人称代词“我”，要求调查“我姓王”里的“我”，不调查处于宾语或定语位置的“我”。

3.4 调查方法

所有调查点一律赴当地（所要调查的县）进行调查。在传统的书面记录之外，还采用数字录音方式录制全部调查项目的有声语料。

调查方法基本上采用传统的提问法。

重点点的纯调查时间需要 5~7 天，一般点的纯调查时间需要 2 天左右。

3.5 调查人

承担调查工作的人员共计 46 名，来自中国、日本、美国的 34 所高校和研究单位。其中具有汉语方言学专业背景博士学位的达 34 名，具有副高以上职称的也达 34 名。这些人分别负责调查本人比较熟悉的方言地区。不过，各人所负责调查的点数是不同的，有的人调查了五六十个点，多数人调查了二三十个点，也有的人只调查了 1 个点。

调查人名单如下（各人所负责的调查点的情况见附录）：曹志耘、陈晖、陈泽平、范耀斌、甘于恩、高晓虹、顾黔、郭风岚、胡士云、胡松柏、黄晓东、李建校、李蓝、李连进、李咸菊、刘倩、刘淑学、刘新中、罗自群、孟子敏、钱曾怡、秋谷裕幸、瞿建慧、石汝杰、Richard VanNess Simmons（史皓元）、孙宜志、唐伶、汪平、王建设、王莉宁、王临惠、王文胜、吴伟雄、吴永焕、项梦冰、邢向东、徐越、严修鸿、张世方、张双庆、张维佳、张燕芬、张燕来、赵日新、周晨萌、庄初升。

3.6 分类人

承担分类工作的人员共计 20 名，均为课题组成员和在读研究生。分类是对材料的分析和研究，工作难度很大。虽然每条图目的分类都由专人负责，但每张分类表都要经过反复讨论和修改，所以实际上是很多人合作研究的结果。

分类人名单如下：曹志耘、邓楠、范耀斌、高晓虹、郭利霞、黄晓东、贾坤、李咸菊、刘倩、郗远春、石汝杰、王莉宁、夏俐萍、杨璧菀、杨慧君、张世方、张燕芬、赵日新、周晨萌、邹妍。

3.7 数据库和画图

930 个点的原始调查数据多达 100 多万条，靠人工来处理这些数据是不可想象的。在本课题的研究过程中，数据库发挥了巨大的作用。它一方面为方言材料的保存、修改、检索、提取、比较、分析提供了极大的方便，同时也是电脑绘图不可缺少的一个基础。

在画图阶段，我们利用中国国家测绘局国家基础地理信息系统 (NFGIS) 1:400 万的全国地图数据和美国 ESRI 公司的 ArcView9.1 版绘图软件，建立了一个“汉语方言地理信息系统”，在该系统的基础上，进行方言地图的绘制工作。事实表明，该系统简便实用，能够满足绘制方言地图的要求。

四 内容简介

根据调查结果，我们从全部调查条目中归纳出最有价值的 510 个地图条目，绘制成方言特征分布地图，分为语音、词汇、语法 3 卷。其中语音卷 205 幅，词汇卷 203 幅，语法卷 102 幅。每幅地图均包括 930 个调查点的信息。

4.1 图类

各卷地图的类型及其数目如下：

语音卷 205	- 音类图 160 - 特字图 32 - 音值图 13	- 规律图 23 - 代表字图 137
词汇卷 203	- 概念图 188 - 词形图 6 - 分合图 4 - 综合图 5	
语法卷 102	- 结构图 51 - 语法词图 39 - 综合图 12	- 普通话出条图 29 - 方言出条图 22

这是对地图的属性进行归类的结果。但为了便于阅读和使用，各卷里地图的顺序仍按通常的习惯排列。

4.2 语音卷

语音图以反映重要的历史演变为主要目的，同时也反映重要的地域差异。

(1) 音类图：反映音类的古今演变情况，语音图中有 78% 是音类图。音类图包括“规律图”和“代表字图”两类。

规律图：反映音类的古今演变规律，即某个音类保留、分化、合并、变化的类型和条例。例如：平声的分化，入声的演变，浊塞音、塞擦音声母的演变，泥来母的分合，阳声韵尾的演变，入声韵尾的演变。

代表字图：“代表字”是指音类中有代表性的字，代表字图用以表现“规律”的具体读音情况。例如：坐—大浊上去声调的异同，铜定的声母，酒—九尖团声母的异同，猪抽虫知彻澄的声母，大果开—的韵母，心—星深梗韵母的异同。

(2) 特字图：“特字”是指从读音的来源或演变的角度来看比较特殊

的字。例如：鼻特字的声调，桶特字的声母，打特字的韵母。

(3) 音值图：选择部分比较有代表性的、比较重要的音值，以反映汉语方言在音值方面的特点。例如：[tɕ tɕʰ ʂ]声母，[ɬ]声母，舌尖元音。

4.3 词汇卷

词汇图以反映重要的地域差异为主要目的，有些图也可以反映历史演变的情况。

(1) 概念图：反映一个概念在各地方言中的具体说法，词汇图中有93%是概念图。例如：太阳~下山了，今天，吃~饭，热指天气。

(2) 词形图：反映一个词形（或者说一个语素，如“房”）在各地方言中的意义。例如：“手”和“脚”的词义，“房”的词义。

(3) 分合图：反映在各地方言中，不同的概念是否使用同一个词形。例如：“外孙—外甥”说法的异同，“喝~酒—吃~饭”说法的异同，“胖人~—肥猪~”说法的异同。

(4) 综合图：对现有材料进行进一步综合、分析后画的图。例如：“月亮”的拟人称呼，“阴茎”的源自动物名的称呼，“猪舌头”的婉称。

4.4 语法卷

语法图以反映重要的地域差异为主要目的，有些图也可以反映历史演变的情况。

(1) 结构图：反映某个语法结构（包括词法、句法）在各地方言中的相应形式。结构图包括“普通话出条图”和“方言出条图”两类。

普通话出条图，例如：“桌子”的后缀，我吃了一碗饭完成体，他把碗打破了处置句，我比他大比较句。

方言出条图：有些语法现象只见于方言，在普通话里没有对应的形式，这种现象只能从方言的角度设立图目，称为“方言出条图”。例如：只鸡死了〈方〉量词定指，囤〈方〉名词后缀，“你去先”的“先”〈方〉“先”是后置成分，表示领先。

(2) 语法词图：“语法词”是指语法属性比较强的词，主要是代词、

副词和虚词等。例如：我～姓王，很今天～热，不明天我～去，的我～东西，了他来～，被衣服～贼偷走了。

(3) 综合图：反映在构词或语法功能表达方面的一些综合性现象。例如：人称代词复数表示法，否定词的分类，小称形式，动物性别表示法。

4.5 语料

方言地图所反映的是经过提炼、归纳的语言现象。由于版面的限制，地图上只能反映最重要的语言特征，而难以容纳更多的语料，例如语音图往往只反映声、韵、调当中的某一方面，而未能反映整个音节的全貌；词汇图一般只反映方言说法的字形，而未能反映其具体读音；有些语法图则经过了高度抽象化，只反映方言说法的结构，不反映其用词。而对于方言地图的理解和解释，对于语言现象的分析研究来说，全面、详细的语料无疑都是非常重要的。为了弥补地图的不足，我们拟在本地图集出版之后，组织力量编写《汉语方言语料集》和《汉语方言地图集（电子版）》，把调查所得的全部语音、词汇、语法材料以列表对照的形式或以数据库的形式提供给研究者使用，在电子版里，甚至可以提供有声语料，提供多种检索、显示、比较、分析以及互动的功能，使之发挥更强更大的作用。

参考文献

- [1] 北京语言大学语言研究所 2003 《汉语方言地图集调查手册》，澳门语言学会。
- [2] 曹志耘 2004 地理语言学及其在中国的发展，《北京语言大学汉语语言文学萃·方言卷》，北京语言大学出版社。
- [3] 陈章太 詹伯慧 伍 巍 2001 汉语方言地图的绘制，《方言》第3期。
- [4] 国立国语研究所 1966~1974 『日本言語地図』，日本国立国语研究所。
- [5] 贺登崧 2003 《汉语方言地理学》（石汝杰、岩田礼译），上海教育出版社。

- [6] 金有景 主编 1992 《中国拉祜语方言地图集》，天津社会科学院出版社。
- [7] 项梦冰 曹 晖 2005 《汉语方言地理学——入门与实践》，中国文史出版社。
- [8] 中国社会科学院 澳大利亚人文科学院 合编 1987 《中国语言地图集》，香港朗文出版（远东）有限公司。
- [9] 中华人民共和国民政部编 2007 《中华人民共和国行政区划简册 2007》，中国地图出版社。

（原文载于《语言教学与研究》2008 年第 2 期）

作者简介

曹志耘，男，1961 年出生，毕业于山东大学，获文学博士学位，北京语言大学教授，博士生导师。