

刘俐李

侯超

等著

江阴方言新探

汉语方言学研究系列

刘俐李 主编

世界图书出版公司

江阴方言新探

刘俐李 侯 超 等著

世界图书出版公司

北京·广州·上海·西安

图书在版编目 (CIP) 数据

江阴方言新探/刘俐李, 侯超等著. —北京: 世界图书出版公司北京公司, 2012. 12

ISBN 978-7-5100-5041-1

I. ①江… II. ①刘… ②侯… III. ①吴语—研究—江阴市 ②江淮方言—研究—江阴市 IV. ①H17

中国版本图书馆 CIP 数据核字 (2012) 第 311009 号

江阴方言新探

著 者: 刘俐李 侯 超 等

责任编辑: 陈晓辉

出 版: 世界图书出版公司北京公司

出 版 人: 张跃明

发 行: 世界图书出版公司北京公司

(地址: 北京朝内大街 137 号 邮编: 100010 电话: 64077922)

销 售: 各地新华书店和外文书店

印 刷: 三河市国英印务有限公司

开 本: 711 mm × 1245 mm 1/24

印 张: 8.1

字 数: 186 千

版 次: 2013 年 1 月第 1 版 2013 年 1 月第 1 次印刷

ISBN 978-7-5100-5041-1

定 价: 29.00 元 (含一张 CD)

版权所有 翻印必究

《语言科技文库》总序

李葆嘉

当代语言学已经进入了一个科学与技术的互补时代，信息处理水平成为衡量国家现代化水平的重要标志之一。知识世界的载体是语符系统，信息处理的根本对象是语言信息处理。与计算机的出现使得语言符号有可能成为数据处理对象相似，神经科学实验仪器设备的应用，使得在大脑神经层面探讨语言机制成为可能。这些无疑都引导语言研究走向科技化，“语言科技新思维”（李葆嘉 2001）应运而生。所谓“语言科学”包括理论语言学、描写语言学、历史语言学、应用语言学等分支学科，所谓“语言技术”指语言研究的现代技术手段，包括语言信息处理、语音实验分析，以及语言的神经、心理和行为实验分析的技术手段等。就语言信息处理而言，又可以分为语料库研制技术、知识库研制技术、知识挖掘和抽取技术、句法信息处理技术、词汇信息处理技术、语音信息处理技术、语义信息处理技术、语用信息处理技术等。

2001 年 5 月，南京师范大学文学院创办了史无前例的“语言科学及技术系”，率先迈出了从传统文科教育范型向现代科技教育范型转变的步伐。“十五”期间，南京师大“211 工程”重点学科建设项目“语言信息处理与分领域语言研究的现代化”（陈小荷教授主持），以基础平台建设、资源建设和理论探索等为主，迈出了语言科技研究的一大步。

“十一五”期间，南京师大文学院、外国语学院和国际文化教育学院联袂申报“211 工程”三期重点学科建设项目。该项目以“语言科技”为引导，以“多学科交叉、跨院系整合、开放型营运”为理念，建设具有前瞻性、原创性、成长性的语言科技高级工作平台。以典型课题的工作原理为核心，进行资源开发和系统研制，拓展语

音科技、二语习得的神经机制研究、言语能力受损儿童的语言能力研究等新方向。同时造就新一代学术领军人物和培养一批高层次复合型人才，以期形成一支高水平的交叉学科团队。该项目设计，体现了工作平台建设、理论创新、应用研究、人才培养、团队建设的学科发展一体化思路。其旨趣在于，加速语言研究从传统文科范型向现代科技范型的转变，以引领 21 世纪语言科技的新潮流。

作为新兴交叉学科项目，通过教育部组织的专家匿名评审，“语言科技创新及工作平台建设”（2008 ~ 2011）获批，总投入 1000 万元。总体而言，这一“语言科技创新”团队，分支学科齐全，专业知识互补。涵盖了理论语言学、计算语言学、语义科技、语音科技、实验方言学、历史语言学、神经语言学、二语习得研究、话语行为语言学等领域。这一期间，项目组成员获批的国家级基金项目达 20 多项。该项目理念之前瞻、实力之雄厚、工程之浩大、经费之保障，为学界瞩目。

2008 年秋，本项目以南京师范大学语言科技研究所为实施单位正式启动。主要有三大任务：建设一个领先性的语言信息科技实验室、建立一个独创性的语言科技工作平台、撰著一套有特色的语言科技文库。

从实验室方案设计到设备招标采购，再到实验室用房改造，经过 8 个月的努力，2009 年 12 月，语言信息科技实验室建成，为语言研究从传统范型向科技范型的转变提供了基本保障。该实验室划分为实验工作区、科研工作区和管理服务区。实验工作区建有语音实验与计算室、神经认知实验与计算室、课堂话语实录室三个专门实验室。科研工作区建有语义科技工作室、语音科技工作室、方言实验工作室、知识工程工作室 I（先秦词汇）、知识工程工作室 II（中古词汇）、知识工程工作室 III（敦煌俗语言文字）、语言习得神经机制工作室、语言习得中介机制工作室，以及参研工作室。管理区服务包括办公室、管理室、编辑室和交流室。出席“语言科技高层论坛暨语言信息科技实验室落成仪式”（2009 年 12 月 14 日）的专家认为，该实验室体现了语言学跨学科研究的当代性和先进性，具有整体性、科技型、开放型三个特点，处于全国领先地位，是“语言科技新思维”的又一体现。同时认为，该实验室的科研工作涵

盖了四个二级学科、四个博士学位点，有稳定明确的研究方向，有合理的设计规划和很好的科研基础；整体设计合理，功能齐备。以教育部重点实验室建设标准衡量，很多方面超过了指标。

语言科技工作平台是基于工作原理（课题定位—理论方法—技术路线—关键技术—评估方式）而建设的高级平台。一方面，从语言信息、语言知识和语言机制三个层面，围绕典型课题进行设备配置、资源建设和软件开发；一方面，将典型课题研究和工作平台建设融为一体，依据典型课题建设的子平台应具有解决同类课题的功能。

建设语言科技工作平台的目标是要实现语言研究手段的技术化和模型化，总体设计包括三个二级平台和八个子系统。

一、语言信息工作平台 1. 语义科技工作系统（李葆嘉教授主持）：基于词汇语义—句法语义的一体化研究思路，开发“人—机交互语义标注工具”，研制“深度语义标注信息库”；研制“幼儿（2~6）日常话语跟踪语料库”，完成幼儿语义系统和话语行为分析研究。2. 语音科技工作系统（顾文涛教授主持）：研制“多语言、多语境、多语用的语音语料库”，基于声学信号分析、感知实验和数学建模，完善语音韵律理论与相关技术应用。3. 方言实验工作系统（刘俐李教授主持）：完成“网络版汉语方言有声语料库”，拟定系统的可操作性语音、词汇、语法实验模型和研究方法，进一步完善新兴交叉学科“实验方言学”。

二、语言知识工作平台 1. 先秦词汇统计与知识检索系统（陈小荷教授主持）：研制“先秦文献语料库”、“专名知识库”、“汉语词汇档案库”等，开发先秦文献自动分词算法、古籍版本异文自动发现算法、同指专名检索软件工具等，完成“先秦汉语词汇统计与知识检索”。2. 中古词汇统计与知识检索系统（董志翘教授主持）：研制“中古文献语料库”、“专名知识库”、“中古汉语词汇档案库”等，开发中古文献自动分词和标注工具等，完成“中古汉语词汇统计与知识检索”。3. 敦煌俗语言文字统计与检索系统（黄征教授主持）：研制“敦煌文献资料库”、“敦煌文献俗语词语档案库”，开发相应工具，完成“敦煌文献资料与知识检索”。

三、语言机制工作平台 1. 二语习得的神经机制研究系统（倪

传斌教授主持):研制“英语受蚀词汇库”等,基于行为学、脑成像和脑电三维度模型,进行中国人英语习得与磨蚀的神经机制研究,完成“基于神经机制的英语个性化学习分析系统”。2. 二语习得的中介机制研究系统(肖奚强教授主持):研制“留学生汉语口语中介语语料库”,基于中介语理论、对比分析理论、偏误分析理论以及二语习得影响因素等,完成“留学生汉语习得的中介机制研究”。

这一工作平台,既是科技研究平台,也是人才培养平台,即一个现代化的科学研究和人才培养工作体系。

作为本项目的文本成果,《语言科技文库》包括计算语言学研究、语义语法学研究、汉语方言学研究、古代汉语学研究、语言教学与研究、语言新专题研究六个系列。其总体特征为:领域的开拓性、理论的原创性、选题的新颖性、方法的交叉性、考据的精确性、成果的应用性。在研究过程中,除了数据采集分析、资源建设和软件开发,更重要的还是要有新思路、新理论和新材料。陈小荷提出的先秦文献信息处理新方法,从先秦典籍注疏文献中挖掘出用于自动分词和词义消歧的知识,再注入已开发的古汉语分词和词性标注工具中去,所取得的先秦古籍版本异文自动发现、先秦词汇知识自动挖掘等成果均具开拓性。李葆嘉提出的语义语法学理论和话语行为理论,基于研制专用语料库或语义信息库和技术手段,开拓了语义网络建构、深度语义分析和话语行为研究等新的领域。刘俐李建构的实验方言学理论和方法,为方言学向现代科技方法的转型研究提供了新路,并取得了一系列新成果。黄征多年来从事敦煌文献及其俗词语文字研究,古代汉语学研究系列中的敦煌文献校录整理,以及敦煌写本字词考释、以古佚和疑伪经为中心的敦煌佛典词语和俗字研究、两汉声母系统研究等新见迭出。肖奚强基于汉语中介语语料库的二语习得研究,在对外汉语教学研究界已经产生了影响。钱玉莲的汉语介词与相应英语形式比较研究等专著各有亮色。倪传斌依据语言测试和认知实验等数据,从行为学、生理学和语言学三个层面分析影响中国英语学习者外语磨蚀的相关因素。刘宇红基于隐喻的理论探讨,对各类隐喻形式的结构、特性和解读规律进行了多视角的深入探讨。

《语言科技文库》所收论著,由作者在2008年12月申报选题,

2011 年始逐步完稿。系列主编审读了书稿，主要就其学术价值、章节安排、内容关联、行文表述、图表绘制等方面，提出审阅意见。此后，作者们对书稿又进行了修改和润色。《语言科技文库》的作者，大多数是具有博士学位的年轻教师。对于我们这些 20 世纪 80 年代走进语言学研究领域的而言，出版论著可能已不足为道。然而，对于年轻学者而言，其论著的出版既是几年来研究的结晶，也是对其继续探索的促进。换言之，“211 工程”重点学科建设的目的之一，就是为年轻教师搭建一个可持续发展的科研和教学平台。学科带头人的主要任务之一就是提携后进。

尽管从根本上来说，科学或学术研究是一种个人的探索行为，然而复杂问题的研究，无疑需要群体协作。“学科建设”或团队合作模式，是 20 世纪 90 年代后期出现的一个新概念。这种模式涉及总体规划、多方协调，是需要付出精力和心血的。2008 年，通过投票方式推举我担任该项目总负责时，就意识到自己成了一个“劳动班委”。2009 年，前往安徽大学拜访黄德宽教授时，曾谈到“学科负责人的任务就是规划设计，争取项目经费和提供科研设备设施”，得到黄教授的赞许。2010 年，申报江苏省高校哲学社会科学重点研究基地时，评审专家柳士镇教授提问的“作为一个交叉学科项目，各学科之间的协调是怎么考虑的，有什么做法”，可谓一语中的。作为后学，深知交叉研究之艰、学科整合之难。相关学科之间的整合协调需要借助行政机制，但凭借行政方式并非就能完成。当时的回答是，目前做到的是建成了一个可以合作研究的场所，至于学科之间的进一步沟通合作应有较长过程。有一点很明确，只有通过交叉项目，相应学科才能渗透，合作者才能逐步磨合。我们只是在一步步探索。

十一五期间的“211 工程”建设项目即将完成，但是学科建设的任务并没有结束。2010 年，“语言信息科技研究中心”被评审为江苏省高等学校哲学社会科学重点研究基地，为“语言科技”这一交叉领域注入了新的建设活力。重点研究基地建设，除了“跨院系整合、多学科交叉、开放型运行”理念，需要凸显“合作性攻关”。围绕交叉性项目，实施计算语言学、语音科技、神经语言学、语义科技等力量的联合攻关计划。只有通过全面开放以及和与国内外同

行的合作交流，才有望建成具有影响的语言科技研究、人才培养和学术交流基地。

十年前，我（2001）曾写道：“语言科技”的内涵是以理论研究为指导，以描写研究为基础，以应用研究为枢纽，促使语言研究向计算机应用、认知科学和现代教育技术领域等延伸，沟通文理工相关学科以实现语言研究过程及其成果的技术化。“语言科技”的外延为语言工程科技、语言教育科技和语言研究科技。其中，“语言研究科技”是将语言研究活动与资源建设、软件开发相结合，其目标是实现语言学自身的科技化。还应包含语言实验、数据处理这些实验语音学、神经语言学研究的科技手段。

虽然语言学家不可能也不必要都转向语言计算或实验研究，尽管描写、考据和内省始终是最基本的方法，但是具有一定的语言科技意识却非常必要。语言学家只有了解有哪些可供利用的资源、软件或仪器，才能提高其研究深度、精度和效率。语言学家也只有了解到信息处理的语言研究需求，才有可能为之提供可资应用或参考的基础成果。“语言科技”是21世纪语言学研究的潮流。

此为出版缘起。是为总序。

2011年8月谨识于南都

序

2007 年底接到江阴市史志办公室李阳主任电话，约编江阴方言志，即应允。有两个原因促成此事：第一，2008 年恰好是我连续主持的两个国家社科基金课题的空挡，“同源异境三方言百年演变研究”（2004—2007）刚结项，而“基于有声语料库的实验方言学构建”（2009—2012）尚未立项，时间尚可安排；第二，与研究旨趣合拍，当时正想就一个县市级行政单位的方言点进行实验方言学的探究。

实验方言学是综合方言学、实验语音学、语料库语言学以及当代信息科学理念及技术的一门新学科。研究对象为方言，研究手段多样，研究目标不限于方言的客观描写，还在于反映方言的变化，并探讨变化原因。自 2003 年创建文学院语音实验室（现名“南京师范大学综合语音实验室”）以来，我们一直在进行实验方言学的探索。就一个方言点的全面研究而言，“同源异境三方言百年演变研究”课题中的哈萨克斯坦东干语调查是实验方言学的第一次系统操练，江阴方言研究是第二次。

实验方言学植根于传统方言学与实验语音学，但又与二者不同。其方法论要点是：用语料库语言学方法收集、储存方言语料，使所记方言事实可验证、可重复使用，包括他人同类研究的再应用，避免重复调查；用实验语音学方法分析审订方言语料，提高语料描写的客观度和精准度；用社会语言学方法调查、分析方言应用的动态现状，包括该方言使用社团的方言种类及其使用频度，语言态度与语言心理的统计分析，并在科学抽样的基础上剖析方言变化的趋势与原因。

江阴方言的调查研究按照实验方言学的方法论原则设计与实施。

首先是确定方言点。江阴是个县级市，辖 15 镇，另加市政府所在地澄江街道，有 16 个行政单位，设为 16 个调查点。江阴的主方

言是吴语，但沿江的利港镇有江淮方言，再设一个调查点。此外，为系统描写江阴吴语近 20 年的发展变化，除应用 20 年前的方言语料对比分析外，我们还用年龄级差形成的方言差异反映和佐证这一变化，即在澄江街道对新派与老派进行同等调查。16 个行政单位，另加澄江街道新派和利港镇江淮方言两个点，共计 18 个方言调查点。根据资料与调查，江阴方言可分作 5 片：吴语 4 片，江淮方言 1 片。

其次是确定发音合作人。按照方言调查惯例，各镇调查老派方言，澄江街道另加新派。18 方言点各选 1 人，其中 5 方言片的代表点与新派发音人各再备选 1 人，共 24 人。此外还有社会语言学的调查合作者 120 人（可与发音合作人重叠）。确定和物色这 100 多人是调查成功的关键。在实地调查前，特拟定两类人选的入选条件（详见书后附录一），送请江阴市史志办公室的纪晓华科长事先帮助物色。课题组到现场后，经过预调查，筛选符合条件者。纪晓华科长鼎力协助，这项工作进展顺利。发音合作人信息见书后附录二。

最后是确定调查内容。调查内容取决于研究目标。此项调查目标有三：1. 江阴方言的现状，包括方言种类、结构及其社会应用；2. 江阴方言近 20 年来的变化；3. 变化原因。这三项研究有弹性，详略可伸缩。一般而言，以详尽为上，但有尺度。在时间、精力、经费允许的前提下，我们取详略结合的原则：代表点详，非代表点略。基本原则是：6 方言代表点作全面调查，其余 12 方言点作专题调查。全面调查含音系与词汇的系统调查，其中澄江老派另含同音字汇及语法、语料；另 12 方言点进行了声调的系统调查（这是方言分片的依据），以及反映近 20 年方言变化的特征词、特征字的调查。根据这些调查项目，我们分别设计了相应的语音、词汇、语法及语料的各类专用调查表。

对调查表中的词汇表和江阴话问卷表要作些说明。词汇表分基础词（748 词）、特征词（163 词）、核心词（100）三类（实际调查到的三类词各方言点不等）。我们认为，基础词是常用词，反映方言词汇的主干；核心词是词汇中最稳固的部分，反映汉语词汇的共性；特征词反映方言的个性。这样分类收词是顾及词汇系统的各结构层面，能在词汇系统的结构层避免疏漏，有助于全面反映词汇面貌。

同时词汇的调查量可得到控制，即用可控词量最大限度地反映方言词汇的真实状态。核心词表采用斯瓦迪士的前 100 核心词，特征词根据钱乃荣（2003）《北部吴语的特征词》修订而成，基础词表选用了最常用词，按词义分为 24 类。

江阴话问卷表是为调查江阴方言现状与变化原因而设，分家庭环境、工作环境、其他环境以及语言态度、语言心理等 5 类，共计 25 项。其中前 3 类是调查江阴人使用语言的三种环境，而后 2 类是为调查江阴方言变化原因所设。该表见书后附录三。

2008 年 3 月 31 日（周一）课题组进入江阴市，按照既定程序调查并录音。两套录音设备同步采录语料，并加班加点核实调查材料。4 月 4 日（周五）完成调查。返校后分析整理。其间两次再赴江阴补充调查，还有多次电话询问复核，最后记音成文，修改校对。终稿于 2009 年 1 月 21 日完成，正文 13.4 万字。

数据库的构建颇费精力，主要是切分音档。为方便检索，单字大体一字一个音档，词汇则一词一个音档，都是人工切分。多亏有一个工作努力、技术娴熟的团队。

分析整理的第一步是整理音系，其中声调，尤其调值的拟定是传统方言研究的难点，我们采用声学分析的方法，这项工作由侯超实验师统一完成。依照我们实验室拟定的声调实验模式，他对 18 方言点声调音档逐一标注、提取基频数据、进行五度值转换，归纳出单字调系统。我们根据单字调古今调类的对应并合，将江阴方言分作 5 片，与 20 年前江阴地方志中的方言志结论一致，但 20 年前没有说明分区理由。本书关于吴语与江淮方言的区分标准没有沿用传统的声母是否分清浊，而是根据中古去声的分合：吴语分，江淮方言合。粗查更多资料，发现这一分类标准大体符合江阴境外的吴语与江淮方言。这一问题需要专题深入研究，已超出本书范围，不赘言。声韵母记音以及语句语篇记音均采用语音分析软件 praat 辅助分析与审订。

当 13.4 万字的终稿交付江阴市史志办公室后，却没有这么多版面，江阴地方志留给方言的版面只有 5 万字，需要删节 8.4 万。李阳主任说：“你们可以出书嘛”，于是有了出书打算。但书稿一直静卧于电脑。直至南京师范大学“211 工程”三期重点学科建设项目

“语言科技创新及工作平台建设”（2007—2012）策划出版方言学丛书。我在这一重点学科建设项目中主持子课题“方言实验工作系统”，本书正是这一工作系统的实践，算是名至实归。

书名中的“新探”有两层含义，一是采用实验方言学这一新的研究模式，二是相对江阴方言志 1988 年成稿、1992 年出版的老版本，本书为新。老版方言志的作者未知姓名，但调查细致、全面，信息量大，而表述则言简意赅，是方言学训练有素者所为，对他心怀敬意。

本书虽起因于 2008 年《江阴志·方言卷》的撰写，但与它很不相同。除字数悬殊外，成书时对江阴方言进行了再研究。我们依据语图分析并结合听辨感知，对江阴六方言点的五、六千条词汇进行逐音逐字的过滤式校验，修订错误，补正疏漏，并再归纳音系；修改数据库程序使界面更方便使用者。书稿虽即付梓，但疏漏及偏误一定不少，期待读者赐教。

本项目研究者除笔者与侯超外，主要还有 2008 年在读的博士生顾劲松、宋益丹、彭玉康和硕士生冯青青、王丽，以及目前正在读的唐志强、滕菲，对他们的努力与付出深表谢意。

该项研究得到江阴市史志办公室的资助与协助，合作很愉快，并致谢忱。

刘俐李

2012 年 3 月 15 日

本书符号使用说明

① “□”表示难以写出的字。

② 字下加波浪线“~”表示同音替代字，如表13“喊呢（咱们）”中的“喊”。

③ 字下加“_”表示白读，加“__”表示文读。如江阴新派韵母表中的a：介（白读），ia：介（文读）。

④ “h”为送气符号。

⑤ 声调五度数字下的横线表示声调短，如23。

⑥ 本书注音使用国际音标，一般省去惯用的方括号。

⑦ 多音、多义以及冷僻字一般用下标加以说明。下标“又”表示同一个字另外的读法。

⑧ 因软件功能所限，以上③的下划线以及⑦的下标字在光盘所附“江阴方言有声数据库”中均不能显现，改用圆括号加注的方式表达，如“介（白读）”“介（文读）”分别表示“介”的白读音和文读音，“诬（又）”表“诬”的又读音。②的下划线也不能显示，数据库未作标注，可参看本书。

目 录

引 言	1
第一章 江阴方言的发展变化和地域差异	2
第一节 江阴的语言现状	2
第二节 江阴方言的发展变化	5
一 近 20 年来江阴声调的变化	5
二 江阴声调的古今演变	7
三 近 20 年来江阴声韵的变化	8
第三节 江阴方言新老两派的差异	9
一 语音差异	9
二 词汇差异	11
第四节 江阴方言的地域差异	12
一 江阴方言的内部分区	12
二 江阴 18 方言点字音对照	14
三 江阴 18 方言点特征词语对照	19
第二章 语音	33
第一节 声韵调系统	33
一 江阴老派音系	33
二 江阴新派音系	35
三 澄南片音系	36
四 澄东片音系	38
五 澄西片音系	39
六 江淮片音系	41
七 江阴方言六代表点音系总表	42
第二节 音变	45
一 连读变调	45
二 文白异读	46

三 合音现象	48
第三节 同音字汇	48
第三章 词汇	68
第一节 基础词表	68
第二节 核心词表	119
第三节 特征词表	127
第四节 江阴方言词汇分析	139
第四章 语法	143
第一节 词法	143
一 构词法	143
二 特殊词类	145
第二节 句法	149
一 被动句、给字句、否定句	149
二 疑问句	150
三 语序例证	150
四 语法例句	150
第五章 方言语料	154
第一节 熟语	154
第二节 民间故事	158
第三节 地方戏曲	165
第六章 江阴方言有声数据库	166
第一节 数据库结构和功能	166
一 首界面	166
二 概况	166
三 语音	166
四 词汇	167
五 语法	167
六 语料	167
第二节 数据库的操作	167
一 首界面操作	168
二 概况部分操作	169
三 语音部分操作	169

四	词汇部分操作	171
五	语法部分操作	171
六	语料部分操作	172
参考文献		173
附录一	发音合作人条件	174
附录二	发音合作人信息	175
附录三	江阴话问卷调查表	176