



语言规范 与语言应用探索

史灿方 孙曼均 著



南京大学出版社

H1
115

语言规范 与语言应用探索

史灿方 孙曼均 著

内 容 提 要

本书对语言文字应用、语言与文化、语言规范等方面进行研究,特别针对新时期语言规范过程中出现的城市流行语、网络语言、普通话推广和方言保护、语言的法律和道德问题进行了探索。本书既有理论探究,也有实例剖析,文笔活泼,行文流畅,可作为语文工作者以及大学生和研究生语言研究和语言学习的参考书。

图书在版编目(CIP)数据

语言规范与语言应用探索/史灿方,孙曼均著. —南京:
南京大学出版社,2008.6
ISBN 978-7-305-05401-3

I. 语… II. ①史…②孙… III. ①汉语规范化—研究
②汉语—应用语言学—研究 IV. H1

中国版本图书馆 CIP 数据核字(2008)第 059192 号

出 版 者 南京大学出版社
社 址 南京市汉口路 22 号 邮 编 210093
网 址 <http://press.nju.edu.cn>
出 版 人 左 健

书 名 语言规范与语言应用探索
作 者 史灿方 孙曼均
责任编辑 王 琪

照 排 南京南琳图文制作有限公司
印 刷 丹阳兴华印刷厂
开 本 850×1168 1/32 印张 9.25 字数 262 千
版 次 2008 年 6 月第 1 版 2008 年 6 月第 1 次印刷
ISBN 978-7-305-05401-3
定 价 24.00 元

发行热线 025-83594756
电子邮箱 sales@press.nju.edu.cn(销售部)
nupress1@public1.ptt.js.cn

* 版权所有,侵权必究

* 凡购买南大版图书,如有印装质量问题,请与所购
图书销售部门联系调换

前 言

语言应用和语言规范这些年来受到了前所未有的重视,特别是 90 年代以后,随着市场经济的不断推进,语言交际随之扩大和频繁,语文生活日渐丰富和活跃,语言文字应用中诸多问题也日益凸显,语言应用和语言规范的理论研究显得愈加紧迫,许多学者开始关注并投身这一领域的研究。我们的研究也是从这一时期开始的。本书收录的是我们这些年来写作的一些文章,大多已公开发表过。内容分为五个部分:一、语言文字应用研究;二、语言与社会、文化;三、语言文字的法制与道德;四、语言规范论略;五、语言文字规范和应用实例剖析。

关注语言现实和语言生活,这几乎是我们一以贯之的研究思路。语言是活的语言,用的语言,是与人们的生活最为密切的现象。我们觉得,作为语言工作者,尽管有选择课题的自由,可以在许许多多的语言课题中自由选择研究方向,但从现实需求看,更迫切需要关注现实语言生活,研究活的语言和用的语言,为现实的语言交际服务。所以,我们的不少文章,涉及了现时代语言生活中热点问题和比较敏感的话题,比如新词新语问题、城市流行语问题、语言的道德和法律问题等等,这些文章记录了我们的所思所悟和所感。文章中的语料,也大都采选了当下现实生活中鲜活的语言实例。

语言规范,这是个老生常谈且历久而弥新的话题。说“历久”,是因为语言规范由来已久,从孔子时代的“诗书执礼皆雅言”到秦

始皇统一文字，一直到当今的普通话推广中的语言规范，一直是语言政策和语言生活的重要内容。说“弥新”，是因为语言规范在不同时期有着不同的重点，而且，语言发展过程中不断会出现新的语言规范现象。对语言规范问题进行积极的探究，是我们这些“语言人”义不容辞且责无旁贷的事儿。规范工作很重要，但难度也很大。用英国词典编纂大师约翰逊的话来说，那就是“规范语言，是件吃力不讨好的工作”。因为规范问题不仅涉及理论研讨，而且涉及实践操作。有时候，理论上说得通，而实践中且行不通；也有时候，实践中行得通，而理论上且说不通。我们觉得，不管是理论还是实践，只要不通，就有问题，要么是理论不行，要么是实践不行。在规范问题上，我们认为理论探讨还相对比较滞后，表现在一是现有理论中有的说服力不够，二是对现实语言规范的描写和解释不够。当然，这些问题的解决，需要时间，需要艰苦的探索。本书收录的语言规范方面的几篇文章，只是我们在这方面所做的几点粗浅的思考。是否妥当，只能留待专家同仁的批评了。

目

录

● 前言

语言文字应用研究

- 3 ● 汉字应用水平测试用字的统计与分级 孙曼均
- 16 ● 形声字形旁表义度浅探 费锦昌 孙曼均
- 26 ● 汉字教学的规律和幼儿识字教学 孙曼均
- 37 ● “注·提”实验中的错别字调查 孙曼均
- 47 ● 关于生造词问题的再思考
——兼议词汇规范与词典编纂 孙曼均
- 55 ● 新词新语的规范问题述评
金惠淑 孙曼均 袁 晖
- 69 ● 流行新语词的形义特点分析 史灿方
- 72 ● 动词对宾语褒贬选择的语义分析 孙曼均
- 78 ● 《马氏文通》共时比较方法的运用及其意义
史灿方

89 ● 关于法律语体风格类型研究的几个问题

史灿方

103 ● 语境漫谈

孙曼均

语言与社会、文化

113 ● 论中国当代社会和谐语境的构建

史灿方

123 ● 城市流行词语及其社会文化分析

孙曼均

137 ● 北京的流行词语与当代北京城市文化

孙曼均

148 ● 从京都流行语看当前大众心理的一些倾向

孙曼均

157 ● 漫谈汉语祝福语的变迁与社会的变革

孙曼均

164 ● 广告语言与社会心理

孙曼均

176 ● “撮”“嗑”“开”“练”及其他

孙曼均

180 ● 汉字书写文化向何处去

史灿方

185 ● 重视中国年俗语言文化的保护和研究

史灿方

188 ● 浅谈语言文字的形式力

史灿方

语言文字的法律与道德规范

195 ● 加强语言文字法制建设

史灿方

199 ● 论语言规范解释的现状及其应对策略

史灿方

209 ● 从法律地位看汉语方言的保护

史灿方

218 ● 建设语言文明是实现社会和谐的重要基础

史灿方

223 ● 漫谈语言文字的法律规范

史灿方

 语言规范论略

- | | | |
|-----|-------------------------------|-----|
| 229 | ● 语言规范是把“双刃剑”吗 | 史灿方 |
| 233 | ● “全民病句”问题之管见 | 史灿方 |
| 239 | ● 由“保卫方言”引发的思考 | 孙曼均 |
| 244 | ● 重新认识网络语言规范 | 史灿方 |
| 248 | ● 名从主人与兼顾客人
——对“汉城”易名的几点看法 | 孙曼均 |
| 250 | ● 也说“错字门事件” | 史灿方 |
-

 语言文字规范和应用实例剖析

- | | | |
|-----|------------------|-----|
| 255 | ● 话语要有条理 | 史灿方 |
| 257 | ● 《新华汉语词典》封底文字商榷 | 史灿方 |
| 259 | ● “盛大”能否修饰“期待” | 史灿方 |
| 261 | ● “陆指退干”的简称合法吗 | 史灿方 |
| 263 | ● “推理”还是“推测” | 史灿方 |
| 265 | ● “不仅”的位置该放哪儿 | 史灿方 |
| 267 | ● 此处不该用问号 | 史灿方 |
| 269 | ● “题不对文”例两则 | 史灿方 |
| 271 | ● 标点在标题中的妙用 | 史灿方 |
| 274 | ● 还是用“三八妇女节”的名称好 | 史灿方 |
-

- | | |
|-----|-------|
| 277 | ● 后 记 |
|-----|-------|
-

语言文字应用研究

汉字应用水平测试用字的统计与分级

孙曼均

一、前言

“汉字应用水平测试研究”课题是国家语委“十五”科研规划重点项目,目的是要研制一套科学的标准化汉字应用水平测试体系,以检验社会上与汉字使用有密切关系的一些行业的从业人员在书面交际过程中掌握和应用汉字所达到的程度。不同的人使用汉字的数量不同,掌握汉字的程度及汉字应用的水平也有较大差异。根据我们的调查,一些与汉字应用有关的行业 and 部门早就迫切希望制定一个切实可行的汉字应用水平测试标准,作为考核、录用和培训从业人员的依据之一。因此,制定汉字应用水平等级标准和编制汉字应用水平测试大纲,不仅是对社会上不同人群汉字实际使用情况的客观反映,也是为了满足社会对提高人们语言文字水平的一种迫切需求。

汉字应用水平测试在我国尚属首次,这是一项具有开创性的关于汉字应用水平的量化研究,其研究成果不仅能丰富现代汉字学、语言测试学等学科的内容,而且对提高全社会的语言文字规范意识和应用水平,进而提高人们的文化素养,促进社会应用汉字的规范化和标准化,起到积极的推动和促进作用。这对我国语言文字应用的理论和实践都具有非常重要的意义。

这项课题研究的重点之一就是确定测试用字的数量和范围,以便据此逐步建立汉字应用水平等级标准及测试大纲,并建设相

应的测试题库。

制定《汉字应用水平测试字表》是这项课题的基础性工作,也是今后测试的重要内容。它将为汉字应用水平等级标准的划分、测试大纲的编写,以及题库的建设等提供基本的内容和大致的范围。没有测试字表,测试就无从谈起。鉴于汉字应用水平测试在我国历史上是首次,它不同于以往的语文水平测试或书面语测试,面向社会上使用文字较多的行业和人群,针对人们汉字使用中的各种实际问题,应用性很强。因此,《汉字应用水平测试字表》必然不同于已有的常用字表和通用字表,其字种和字量必须能够真实反映出社会上使用文字较多的行业和人群的汉字应用的实际情况。

二、汉字统计的历史及现状

我国历史上就有对常用汉字和通用汉字的统计。编写常用汉字字表是中国语文教育的传统,也是古代语文教学的基础。古代的蒙学课本就是典型而突出的例子,南北朝的《千字文》共有汉字 1 000 个,宋代的《百家姓》共有汉字 576 个,《三字经》共有汉字 1 248 个。西汉的字书《仓颉篇》有汉字 3 300 个,清代的《文字蒙求》则从《说文解字》收录的 9 353 个汉字中选取 2 044 个常用字,附加字义的解释。这些文字学习教材都是选取当时的常用汉字汇编成册,供儿童学习汉字使用,可以算得上是中国古代的常用字表。历代的字书则反映了不同时代通行的汉字字量和字种,是对当时所使用汉字的统计和汇编。随着时代的推移,字书所收汉字的数量日益庞大,到了现代,大型辞书所收汉字大多有数万之巨。但是除掉其中的异体字、罕用字、方言字、古字甚至废字,古往今来,就社会上应用的汉字而言,一般都在万字之内,人们日常使用的汉字大致在 3 000—6 000 字之间。

我国现代意义的动态性字频统计是教育家陈鹤琴于 1928 年

完成的,统计出常用汉字4 261个。从1928年到1988年的60年间,我国进行的汉字统计资料(重点是对常用汉字的统计)约有十几种。在这些字表中,有两个字表特别值得注意。一个是《信息交换用汉字编码字符集基本集》,这是中国最大的一次字频统计工作。在数以亿计的浩瀚文献资料中,统计出实际通用的汉字数为6 763个,其中3 000多个汉字的累积使用频度达到了99.9%,另外3 000多个汉字的累积频度不到0.1%。这个成果说明了常用汉字与次常用汉字的数量不足7 000个,框定了电脑汉字输入常用字的范围,为国家制定汉字库标准提供了重要的科学依据。另一个是当代汉字整理与统计研究的重要里程碑,即国家语委于1988年完成的《现代汉语常用字表》和《现代汉语通用字表》,这两个字表按照汉字的频率和使用度,确定了2 500个常用字、1 000个次常用字,以及包括上述两类字在内的7 000个通用字。这个成果说明了当代汉字使用的范围在7 000字以内,常用汉字的范围在3 500字之内,由此确定了教学用字以及出版印刷、辞书编纂、信息处理等通用汉字的范围。

据新华社最新统计信息,^①目前汉字的总数虽已超过了8万。而历代日常书面语常用的汉字数量一般都在三四千个。据统计,李白994首7.7万字的诗文用字3 560个;杜甫1 500余首诗用字4 350个;白居易3 000余首诗共18万字,用字4 600个。老舍的代表作《骆驼祥子》共10余万字,单字仅有2 413个;66万字的《毛泽东选集》1—4卷的用字量为2 891个,毛泽东公开出版的全部著作也仅用单字3 136个;《孙中山全集》的用字量为2 673个。

从上述统计资料可以看出,无论在中国古代还是当代,对常用汉字的选择大都在2 000—4 000字之间,即3 000字上下;而通用汉字则一般在7 000字以内,基本为6 000字左右。这些材料和数据是我们编制《汉字应用水平测试字表》的重要参考依据。

^① 参见2001年9月26日《北京晚报》。

在编制测试字表的过程中,我们遇到了一些问题。主要是目前我们对现代汉字的研究还缺乏一些最基础的定量研究,对现代汉字的实际应用情况缺乏宏观掌握和微观分析。比如,对不同文化程度、不同行业等不同人群的识字量,人们往往凭语感来估算,没有源于实际调查的科学、准确的统计数据;又如,对汉字诸如构词能力等诸因素尚无全面而精细的定量分析,等等。这就为测试字表的编制带来了很大困难,也使进一步提高全社会的汉字应用水平无所依据。而现有的常用字表和通用字表,一个是反映受过基本教育、具有中等文化程度的人的常用识字量,一个是反映远远大于一般人识字量的现代汉字流通、应用的范围,它们都不能代表社会人群的实际识字量,更反映不出具有不同文化水平、分属不同行业的人们对汉字实际掌握的字量和字种。因此,现有的各种字表都不能作为汉字应用水平测试所需要的测试用字表。

根据以上情况,课题组经过认真分析研究、多方征求意见,认为应为汉字应用水平测试编制专用的汉字测试字表。《汉字应用水平测试字表》的字量和字种应以常用汉字为基础,但要高于常用汉字;应在通用字的范围之内,但应低于通用字;测试用字必须要从实际中来,要能够精确反映受测人群的汉字实际使用情况。课题组决定,以3500个常用字为测试用字的基本内容,在7000个通用字的范围内,对社会上经常使用文字的人群进行实际测查,根据掌握的第一手数据来选取测试用字;同时,依据国家语委语料库的汉字频率统计数据,对经测查而筛选出来的汉字进行必要的人工干预。测试用字字量和字种的大致框架由此确定。

三、《汉字应用水平测试字表》的编制

(一) 测试字表编制的依据和原则

1. 编制的主要依据

测试字表编制的主要依据是:《现代汉语常用字表》和《现代汉语通用字表》,实际测查统计结果,国家语委两个大规模语料库的字频统计数据。

我们首先确定《汉字应用水平测试字表》应以《现代汉语常用字表》和《现代汉语通用字表》为重要依据,即测试字表以《现代汉语常用字表》为基础,测试用字包括 3 500 个常用字;测试字表以《现代汉语通用字表》为选字的基本范围,字量要低于 7 000 字,字种则基本在 7 000 字之内选定。

其次,鉴于上述面临的问题,我们认为确定《汉字应用水平测试字表》的字量和字种必须首先进行实际调查,以掌握社会上与文字使用密切相关的人群的汉字实际使用情况,并以此作为《汉字应用水平测试字表》的重要依据。在研制《汉字应用水平测试字表》的过程中,无论是对测试用字字量的统计和字种的选取,还是对测试用字进行分级,我们都以课题组测查的实际数据作为主要参数。

另外,我们充分利用国家语委两个大规模语料库的字频统计数据,以此对实际测查的汉字的字量和字种进行增删调整。这两个语料库,一个是国家语委于 1990 年开始建设,反映现代汉语全貌的国家级“现代汉语大型通用语料库”,其库容量为 7 000 万汉字,时间跨度从 1919 年至今(其中 1977 年至今的语料选取比例为 50%)。另一个是语用所于 2002 年承接,面向中文信息处理领域,为自然语言处理研究服务的国家高技术研究发展计划(863 计划)《智能化中文信息处理平台》课题的子课题“超大规模通用平衡语料库”,其库容量为 15 亿字(2002 年第一期完成 7 000 万字),时间

跨度以近十年为主。

2. 编制总原则是实际定量统计与专家定性分析相结合

首先进行目标定性,然后根据定量原则为主、定性原则为辅的选字原则,对实际测查的汉字进行筛选,以其中识别率高、在语料库中频率高、使用度大的字作为选字的基本范围。再采用人工联想的定性方法作为定量标准的补充。即从定性到定量统计,从实际测查到字频统计,再到专家干预。如此反复,不断使之更完善,更接近实际。

3. 汉字筛选的具体原则

(1) 规范性原则:从语言学的角度,测试用字应为普通话语词范围内的用字,方言用字一律不收;测试用字应为规范的现代汉字字型,不规范的字形不收。如:“囡、伢、嬷、甬、攢、姪、擗”等,《现代汉语词典》均注明为方言用字,测试字表不收。又如:“垆”,是“壑”的异体字,虽然目前它尚在 7 000 通用字内,也不予收入。

(2) 应用性原则:从应用语言学的角度,测试用字应在现代汉字的范围内选字,要充分考虑到现代汉字的实际使用情况。对于一些常用的口语用字,如“涸、焯、磴、砾、蕻、疥、搽”等,虽然在书面语中较少使用,字频统计不到或频率很低,但在日常生活中却较常用,则适当选取。对于文言用字,除现在还使用的或有构词(包括成语)能力的,一般不收,如“筮、笏、耜、骡、黼”等,这些字指称的古代事物已经消亡,在现代汉语中既不再使用,又无构词能力,则不予选取。

(3) 常用性原则:依据频率统计的原理,测试用字应以实际人群的汉字识别率以及在语料库中的统计频率为重要参数,即选取高识别率和高频率的字。

(4) 均匀性原则:依据频率统计的原理,测试用字还要考虑其在语料库中的覆盖率,即选取分布广泛、覆盖率大的字。

(5) 联想性原则:任何字表的编制,都不能完全依照频率,必要而适当的人工干预必不可少。对一些成系列的字要进行纵横相

关的多种联想,综合考虑其取舍。如:天干地支、节气用字,数字的汉字大写用字,民族名称,省、自治区、直辖市的简称和别称,地、县级以上的地名用字以及连绵字等,我们都是综合考虑,平等对待,成系列收取。其中为保持地县级以上的地名用字的完整性,我们收入了“邛、埽、埴、獠”等4个7000通用字以外的县级地名用字。对动植物名称、金属、化学元素名称以及姓氏、人名用字等则根据常用程度和频率的高低决定取舍。

(6) 等级性原则:测试用字既要能反映出被试的汉字实际使用情况,又要便于测试等级的确定;因此,测试用字要根据识别率和字频的高低、覆盖范围的大小、构词能力的强弱及其构成词语的常用程度分出等级。

以上6条原则不是孤立的,应综合运用以决定取舍。

(二) 测试字表编制经过

从目标定性开始,到定量统计,再回到定性筛选,最后形成限定性等级字表。这就是测试字表产生的基本过程。

1. 首先进行目标定性。确定3500常用字为测试字表的核心内容,同时在7000通用字范围内确定字量,筛选其他字种。

2. 在7000字中除去3500常用字,对余下的3500字进行实际测查。经研究决定,在使用汉字较多的行业中,选定对汉字掌握的广度和深度有着不同层次的三种人群,即把中等教育程度、高等教育程度、具有高等教育程度又从事文字工作的三种人群作为测试对象。经过实际测试,依照实测人群掌握汉字的平均水准,筛选出其中识别率高的2023个字作为备选字。实际测查结果不仅较为准确地反映了使用汉字较多行业的人群实际掌握汉字的数量,即高限为5500字上下,低限为4000字上下(均含3500字);而且基本上筛出了相应的字种。从源于实际的第一手材料筛选备选字,避免了用传统主观方法确定字表时受专家个人文化素养、专业学科、社会地位以及个人用字习惯等影响而造成的缺陷。