

通用规范汉字

构形属性研究

侯冬梅 / 著



 科学出版社

国家语委“十二五”科研规划项目（项目编号：YB125-161）

通用规范汉字构形属性研究

侯冬梅 著

科学出版社

北京

内 容 简 介

本书以《通用规范汉字表》的 8105 字为对象,对通用规范汉字的字形属性进行了全面、系统的考察,提出现代汉字的直观构形理论,构建现代汉字直观构形分析体系,基于该体系构建通用规范汉字构形属性数据库,并在此基础上对通用规范汉字三级构形单位的属性进行量化分析,从统计角度明确了通用规范汉字三级构形单位的系统性和规律性。进而基于上述研究成果,进一步明确了本书的研究意义和应用前景。本书运用理论与实证研究相结合、定量与定性研究相结合的研究方法,提供了大量汉字构形属性分析、描写和统计的图表,为汉字教学和研究提供了方便快捷的资源。

本书可供高等院校中文专业本科生和研究生、中小学语文教师、对外汉语教师和从事中文教育的科研工作者阅读。

图书在版编目(CIP)数据

通用规范汉字构形属性研究 /侯冬梅著. —北京:科学出版社, 2017.12

ISBN 978-7-03-055100-9

I. ①通… II. ①侯… III. ①汉字-构造-属性-研究 IV. ①H122

中国版本图书馆 CIP 数据核字(2017)第 269371 号

责任编辑:张 达 / 责任校对:邹慧卿

责任印制:张欣秀 / 封面设计:黄华斌

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

北京九州迅驰传媒文化有限公司 印刷

科学出版社发行 各地新华书店经销

*

2017 年 12 月第 一 版 开本: 720×1000 B5

2017 年 12 月第一次印刷 印张: 25

字数: 489000

定价: 118.00 元

(如有印装质量问题, 我社负责调换)

序

冬梅的《通用规范汉字构形属性研究》即将出版，我很高兴。

2013年6月5日，国务院公布了《通用规范汉字表》（以下简称《字表》）。《字表》既是贯彻《中华人民共和国国家通用语言文字法》、适应新形势下社会各领域汉字应用需要的重要汉字规范，又是50多年来汉字规范整合优化的最新成果。它的公布，对提升国家通用语言文字的规范化、标准化水平，促进教育事业的发展，具有重要的意义。《字表》8105字，这些字在构形上有些什么特点？能否通过构形的分析，为汉字的教和信息处理提供有益的帮助和启示？这就是该书所要探讨的问题和现实意义所在。

翻阅书稿，留下三点突出的印象，概括起来，就是三个“讲求”。

第一，讲求创新。冬梅的这部著作，以《字表》的收字为对象，对通用规范汉字的字形属性进行了全面的考察，提出了现代汉字的直观构形理论；并在此基础上，提出了现代汉字的直观构形分析体系。根据这一体系，建构了通用规范汉字构形属性数据库；并基于数据库，对通用规范汉字三级构形单位（整字、部件、笔画）的属性进行了量化分析；基于量化分析的结果，总结出通用规范汉字三级构形单位的系统性和规律性。层层深入，环环相扣，这一系列的工作，无不透露出作者的创新意识。

第二，讲求方法。该书不仅提出了富有创新性的认识，而且在研究方法上也显示出鲜明的特点。研究中，注重理论与实证相结合，定量统计与定性分析相结合。理论认识建立在统计分析的基础之上，这样的认识才富有启发性，才具有实用价值。

第三，讲求务实。该书是冬梅刚刚通过答辩并获优秀的博士论文。冬梅读博三年，但她对这一问题的思考和研究远远不止三年。她学习刻苦，勤于思考，思路清晰，视野开阔，尤其是“板凳甘坐十年冷”的求学精神，在学风日下的今天，显得难能可贵。书中展示了大量的图表，比如：通用规范汉字合体字直观构形树图，通用规范汉字部

件表，基础部件的直接组合部件表，合体字基础部件的形变情况表，基础部件笔画、笔顺特征表，合体整字初、高级构形模式统计表，基础部件构字能力表，合体字基础部件分布位置统计表等，这些图表不仅为汉字的教学和研究提供了有用的信息，也反映出作者务实求真的学风。

该书只是冬梅关于汉字构形问题研究的初步结果。我知道冬梅是个一心向学、力求上进的人。我想她不会就此止步，会继续沿着这一路径，开始新的探索。

汪国胜

2017年11月26日

目 录

序	i
第1章 引论	1
1.1 研究缘起	1
1.1.1 汉字问题仍困扰着我们	2
1.1.2 成年人汉字书写水平呈断崖式下降	8
1.2 研究范围	9
1.2.1 学科范畴——现代汉字学	9
1.2.2 研究内容——汉字构形学	10
1.2.3 研究对象——《通用规范汉字表》	11
第2章 现代汉字直观构形体系的构建	15
2.1 现有国家规范对汉字字形认识的分歧	15
2.1.1 认识的分歧	15
2.1.2 认识存在分歧的原因	19
2.2 现有汉字构形理论与构形体系概观	24
2.2.1 “六书”构形理论及其构形体系	24
2.2.2 “形素”构形理论及其构形体系	26
2.2.3 其他汉字构形理论	27
2.3 现代汉字直观构形理论的提出	29
2.3.1 理论假设	29
2.3.2 理论内涵	31
2.3.3 理论基础	32
2.3.4 拆分依据	34
2.3.5 描写方法	35

2.4 现代汉字直观构形理论提出的动机	37
2.4.1 汉字构形理论研究的新视角	37
2.4.2 从字理角度分析现代汉字字形的缺点	39
2.4.3 现代汉字的特殊性	42
2.4.4 流俗文字学纯字形的现代汉字解构	45
2.5 现代汉字直观构形分析体系的构建	47
2.5.1 构建的目的	47
2.5.2 体系的内容	48
第3章 通用规范汉字构形属性的描写	56
3.1 现有汉字字形分析方法概观	56
3.1.1 “六书”分析法	56
3.1.2 “结构—功能”分析法	58
3.1.3 部件分析法	61
3.1.4 其他分析法	65
3.2 现代汉字部件的拆分	66
3.2.1 部件的界定	66
3.2.2 部件和偏旁的区分	67
3.2.3 部件拆分的依据	70
3.2.4 部件拆分的原则	74
3.2.5 部件拆分的下限	83
3.2.6 部件的判定	84
3.2.7 部件的分类	85
3.3 现代汉字笔画的拆分	88
3.3.1 笔画的界定	88
3.3.2 基本笔形与从属笔形	90
3.3.3 笔画拆分的依据	90
3.3.4 笔画拆分的方法	91
3.3.5 笔画拆分的步骤	97
3.4 《通用规范汉字表》整字构形属性的描写	98

3.4.1	《字表》的独体字和合体字	98
3.4.2	《字表》的基础部件	99
3.4.3	《字表》合体字的部件组合特点	101
3.4.4	《字表》合体字的层级特点	101
3.4.5	《字表》合体字的生成过程	101
3.4.6	《字表》合体字的构形模式	101
3.5	《通用规范汉字表》部件构形属性的描写	103
3.5.1	《字表》基础部件的构字	103
3.5.2	《字表》基础部件的功能	104
3.5.3	《字表》基础部件的分布	105
3.6	《通用规范汉字表》笔画构形属性的描写	109
3.6.1	《字表》基础部件的笔画	109
3.6.2	《字表》笔画的组合特点	111
3.6.3	《字表》笔画的笔形变化	113
第4章	通用规范汉字构形属性数据库的建设	115
4.1	现有汉字属性数据库研究概观	115
4.1.1	大陆汉字属性数据库研究	115
4.1.2	台湾地区汉字属性数据库研究	116
4.2	通用规范汉字构形属性数据库的类型	118
4.2.1	数据库的类型	118
4.2.2	ACCESS通用规范汉字构形属性数据库	119
4.3	通用规范汉字构形属性数据库的设计	119
4.3.1	关系式数据库的数据结构	119
4.3.2	关系式数据库数据结构的相关概念	119
4.3.3	关系的性质	120
4.3.4	通用规范汉字构形属性数据库的表设计	121
4.4	通用规范汉字构形属性数据的入库	127
4.4.1	汉字字符库设计	127
4.4.2	整字生成过程信息的抽取	130

4.4.3	部件构字信息的抽取	132
4.5	通用规范汉字构形属性数据库说明	138
4.5.1	数据库“表”属性的补充与修正	138
4.5.2	“Chaifen”表及其分量说明	139
4.5.3	“Zhengzigouxing”表及其分量说明	140
4.5.4	“Bujiangouxing”表及其分量说明	141
4.5.5	“Strokegouxing”表及其分量说明	143
第5章	通用规范汉字构形属性的计量分析	144
5.1	现有汉字属性定量分析研究概观	144
5.1.1	汉字属性定量分析研究的萌芽	144
5.1.2	汉字属性定量分析研究的探索与成熟	144
5.2	通用规范汉字整字构形属性的定量分析	146
5.2.1	通用规范汉字的独体字与合体字	146
5.2.2	通用规范汉字整字组合特点的定量分析	147
5.2.3	通用规范汉字整字层级特点的定量分析	147
5.2.4	通用规范汉字整字构形模式的定量分析	149
5.2.5	通用规范汉字整字构形属性的统计分布特征	173
5.3	通用规范汉字部件构形属性的定量分析	174
5.3.1	通用规范汉字基础部件构字的定量分析	174
5.3.2	通用规范汉字基础部件功能的定量分析	176
5.3.3	通用规范汉字直接组合部件的定量分析	178
5.3.4	通用规范汉字基础部件分布层级的定量分析	179
5.3.5	通用规范汉字基础部件分布位置的定量分析	180
5.3.6	通用规范汉字形变基础部件的定量分析	187
5.3.7	通用规范汉字合体字基础部件构形属性的综合计量分析	191
5.3.8	通用规范汉字部件构形属性的统计分布特征	200
5.4	通用规范汉字笔画构形属性的定量分析	204
5.4.1	通用规范汉字基础部件笔形分布的定量分析	204
5.4.2	通用规范汉字重码笔顺的定量分析	207

5.4.3 通用规范汉字笔画构形属性的统计分布特征	216
5.5 通用规范汉字构形属性的统计分布规律	217
第6章 通用规范汉字构形属性研究的理论意义和应用价值	219
6.1 研究的理论意义	219
6.1.1 提出并实践了现代汉字直观构形分析理论	219
6.1.2 提出并实践了现代汉字直观构形分析体系	221
6.1.3 明确了现代汉字三级构形单位的统计分布特征	222
6.1.4 分析了现代汉字的系统性程度	222
6.2 研究的应用价值	224
6.2.1 为汉字的构形属性提供了描写方式	224
6.2.2 为非成字部件的命名提供了依据	224
6.2.3 为汉字印刷字形的规范提供素材	228
6.2.4 为汉字的整形提供借鉴	232
6.2.5 为简化字的评价提供依据	243
6.2.6 为现代汉字的教学提供帮助	246
6.2.7 为现代汉字手写规范的制定提供参考	262
参考文献	278
附录	289
后记	386

表 目 录

表 3.1	汉字结构与部件形体的关系	71
表 4.1	“拆分基础信息表”结构	122
表 4.2	“整字构形属性表”结构	123
表 4.3	“部件构形属性表”结构	125
表 4.4	“笔画构形属性表”结构	126
表 4.5	整字及部件拆分数数据存储方式	131
表 5.1	整字部件统计	148
表 5.2	合体整字层级特点统计	148
表 5.3	合体整字构形模式统计	171
表 5.4	基础部件构字能力统计	176
表 5.5	基础部件功能统计	177
表 5.6	合体字基础部件直接组合部件统计	179
表 5.7	合体字基础部件分布层级统计	181
表 5.8	合体字基础部件分布位置统计	182
表 5.9	基础部件倾向性分布统计	187
表 5.10	形变基础部件分类统计类型	189
表 5.11	形变基础部件形变规律统计	191
表 5.12	合体字基础部件综合构形属性统计特征异同	192
表 5.13	构字能力极强的合体字基础部件综合构形属性统计特征	193
表 5.14	综合构形属性统计特征相同基础部件的区分	195
表 5.15	基础部件的基本笔形统计	205
表 5.16	基础部件的笔形变化统计	206
表 5.17	笔顺重码的基础部件	208

表 5.18	以笔形为区别特征的重码笔顺	209
表 5.19	以笔画拓扑关系为区别特征的重码笔顺	210
表 5.20	以笔形和笔画拓扑关系为区别特征的重码笔顺	211
表 5.21	基础部件笔画区别特征统计	215
表 6.1	《常用字部件规范》中“代表字+部位”类非成字部件命名的修正	226
表 6.2	《常用字部件规范》中“字源+省”类非成字部件命名的修正	226
表 6.3	《常用字部件规范》中“传统文字学称谓”类非成字部件命名的修正	227
表 6.4	《常用字部件规范》中包含“部件读音”类非成字部件命名的修正	228
表 6.5	基础部件构字能力分级	246
表 6.6	基础部件组合能力分级	247
表 6.7	基础部件分布位置类型	247
表 6.8	合体整字的优势构形模式分类	249

图 目 录

图 1.1	“亲、爱”二字的繁体字和异体字	6
图 1.2	颜真卿楷书作品中不同字形的“爱”字	7
图 1.3	颜真卿楷书作品中的简化字字形	7
图 2.1	《说文解字》中“月、甘、休、枝”的小篆字形	25
图 2.2	徐冰的书法作品《析世鉴——天书》	30
图 2.3	现代汉字直观构形体系	49
图 2.4	“藻”的树图	51
图 2.5	“藻”的初级构形模式	51
图 2.6	“藻”的高级构形模式	52
图 3.1	“结构—功能”分析法分析汉字字形实例	60
图 3.2	“藻、嫉”的树图	84
图 3.3	“嫉、病”的树图	87
图 3.4	“跑、舔”的树图	88
图 3.5	基础部件类型关系图	88
图 3.6	笔顺的三种表示方式	91
图 3.7	“乞亿天元云艺支厅区”的树图	102
图 3.8	“乞亿天元云艺支厅区”的初级构形模式	102
图 3.9	“乞亿天元云艺支厅区”的高级构形模式	103
图 4.1	台湾“ACCESS 全汉字检索系统”截图	117
图 4.2	新造字符代码的选择	128
图 4.3	新造字符参照字的选择	129
图 4.4	新造字符的编辑	129
图 4.5	新造字符的存储	130

图 4.6	构建二叉树结构的流程	132
图 4.7	构建二叉树的搜索示意	133
图 4.8	部件构字信息的抽取步骤-1	133
图 4.9	部件构字信息的抽取步骤-2	134
图 4.10	部件构字信息的抽取步骤-3	135
图 4.11	部件构字信息的抽取步骤-4	135
图 4.12	部件构字信息的抽取步骤-5	136
图 4.13	部件构字信息的抽取步骤-6	136
图 4.14	部件构字信息的抽取步骤-7	137
图 4.15	部件构字信息的抽取步骤-8	137
图 4.16	部件构字信息的抽取步骤-9	137
图 4.17	“Chaifen”表的分量截图	139
图 4.18	“Zhengzigouxing”表的分量截图	141
图 4.19	“Bujiangouxing”表的分量截图	142
图 4.20	“Strokegouxing”表的分量截图	143
图 6.1	大陆印刷楷体、《康熙字典》和台湾“台湾标准字体”	237
图 6.2	台湾标准字体	238
图 6.3	《说文》和《康熙字典》中“辱”的字形	241
图 6.4	现有汉字字形书写规律-1	264
图 6.5	现有汉字字形书写规律-2	264
图 6.6	现有汉字字形书写规律-3	265
图 6.7	字帖中不同风格的手写体	266
图 6.8	印刷楷体对照字形	266
图 6.9	“辶”印刷宋体与印刷楷体的差异	270
图 6.10	字帖手写体与印刷体字形在笔画空间关系方面的差异	272
图 6.11	字帖手写体与印刷体字形在带钩笔形书写方面的差异	274
图 6.12	字帖手写体与印刷体字形在“横”笔、“点”笔笔形书写方面的差异	274
图 6.13	字帖中“错误”的手写字形	275
图 6.14	手写汉字字形评价标准	277

第1章 引 论

1.1 研究缘起

近年来有这样几起事件，再次把汉字问题推入了公众视野。

事件一：2008年全国两会期间，21位文艺界政协委员联名提出“关于小学增设繁体字教育”的提案，希望借繁体字把中国文化的根传下去。

事件二：2009年十一届全国人大二次会议期间，全国政协委员潘庆林明确提出“全国用10年时间，分批废除简体汉字，恢复使用繁体字”的议案。原因是简化汉字违背了汉字的艺术和科学性；利用电脑处理繁体字已经解决了繁体字的繁难问题；恢复使用繁体字有利于祖国统一。

事件三：网络上风传嘲笑讽刺简化汉字的段子，认为简化汉字“亲（親）不见，爱（愛）无心，产（產）不生，厂（廠）空空，面（麵）无麦，运（運）无车，导（導）无道，儿（兒）无首，飞（飛）单翼，有云（雲）无雨，开关（開關）无门，乡（鄉）里无郎，圣（聖）不能听也不能说，买（買）成钩刀下有人头，仑（侖）成人下有匕首，进（進）不是越来越佳而往井里走……”^①这些简化字把汉字原有的象征意义也减掉了，许多会意字的意思不再，汉字的基本内涵和魅力一扫而光。

事件四：2014年3月9日下午，十二届全国人大二次会议宁夏代表团审议会上，全国人大代表、国家民委副主任吴仕民提议，国家应取消简体字，恢复繁体字，“这样有利于传承传统文化，更有利于国际间文化交流”。

事件五：2015年3月全国两会期间，在政协会议分组讨论中，著名导演冯小刚呼吁恢复部分有文化含义的繁体字。冯小刚举例，如“亲爱的”这几个字非常有含义，“亲”的繁体是左边一个“亲”，右边一个“见”，组成了“親”，“愛”是在“爱”中

^① 括号内的字为简化字对应的繁体字。

加了一个“心”，这两字的含义是“亲要相见，爱要有心”。结果两个字简化以后却变成了“亲不见、爱无心”。同时他还提出能不能选择 50 个或更多有含义的繁体字，增加到小学课本里，让小学生感受传统文化。

事件六：2005 年 6 月，上海某大学举办汉语言文字大赛，来自韩国、俄罗斯及新加坡的留学生组成的队伍与中国学生同场竞技，结果留学生组竟力拔头筹取得冠军。这件事为汉语文教学敲响了警钟，也引起了国内的高度关注。

事件七：2013 年 CCTV10 科教频道推出大型汉字听写节目《中国汉字听写大会》，立刻掀起了全民汉字听写自检的高潮。在复赛前五期节目播出后，我们发现成年观众连“癞蛤蟆、蜈蚣、蜥蜴、电饭煲、烙饼、芥菜、纽扣、吝啬、扭捏、摩羯座、妖孽”等生活中常见的词都非常容易写错。

事件八：2014 年 9 月 25 日，《光明日报》第七版《“失写症”影响的是文化遗产》一文记载了这样一则消息：400 字简历写出 24 个错别字，重庆大学生面试遭拒。文章指出，当今“社会上不少群体都患上了或轻或重‘失写症’”。

归纳起来，这八个事件代表了三种观点。事件一、事件二、事件四、事件五认为：在现代汉字中要恢复或部分恢复繁体字。我们把它称作“繁体字（部分）恢复论”。事件三认为：繁体字优于简体字。我们把它称作“繁体字优越论”。事件六、事件七、事件八认为：现代成年人，包括大学生在内，汉字书写水平严重下降。我们把它称作“成年人失写症”。

进入 21 世纪以来，一方面，繁体字优越论、繁体字（部分）恢复论的声音愈来愈强；另一方面，成年人的汉字书写水平不断下降。这迫使我们不得不去思考：汉字怎么了？成年人的汉字书写水平怎么了？

1.1.1 汉字问题仍困扰着我们

1.1.1.1 21 世纪汉字繁简问题争论愈加激烈

繁体字（部分）恢复论和繁体字优越论并非新论，而是汉字繁简之争的延续。这里的繁体字是未经简化的笔画繁多、结构复杂的楷书汉字，简体字则是与繁体字相对，经过简化的笔画相对较少、结构简单的楷书汉字。

汉字繁简之争始于 20 世纪 30 年代。1935 年 8 月中华民国教育部 11400 号令公布了第一批《简体字表》，在全国推行简体字。但是，由于时任中山大学校长、考试院

院长的戴季陶下跪“为字请命”，中华民国教育部不得不于1936年2月收回了推行简化字的命令。在汉字繁简之争的第一回合中繁体字轻松胜出。

中华人民共和国成立后，从1952年2月开始成立中国文字改革研究委员会，着手推进汉字简化工作，历经数年的努力，于1964年5月根据国务院指示发布了《简化字总表》，在全国推广简化字。这次简化汉字运动同样遇到了来自高级知识分子方面的阻力，但是本着从人民大众立场考虑汉字问题的原则，我国政府至今坚决推行简化字。在汉字繁简之争的第二回合中简体字冲破重重阻力取得胜利。

1978年以后，社会用字中繁简字的混用、乱用现象增多。针对这种现象，20世纪80年代后期以及整个90年代，我国文化界开展了一场有关汉字繁简问题的激烈讨论。其中袁晓园“识繁写简”、任继愈“识繁用简”的观点引起了不小的反响。“识繁写简”“识繁用简”都是主张海内外华人既要认识繁体字也要认识简化字，印刷品也用繁体和简化字印刷，而书面语中则可以书写简化字。这一观点一经提出就形成两大旗帜鲜明的阵营，以《汉字文化》杂志为阵营的一批人强烈支持“识繁写简”“识繁用简”的观点。他们的理由是：繁体字能更好地传承中华文化，繁体字表意性强，繁体字有利于海内外同胞交流和祖国统一。一批以《语文建设》、“中国语文现代化学会”为阵营的人坚决反对“识繁写简”“识繁用简”的观点。他们的理由是：“识繁写简、识繁用简”是“崇洋”“崇古”、破坏语文政策、开历史倒车、搞和平演变。双方阵营虽展开了十余年激烈论战却并未真正形成统一意见。为了推动国家通用语言文字的规范化、标准化及其健康发展，保持国家文字政策的稳定性，让汉字更好地发挥其交际工具作用，并解决这次汉字繁简之争中争论最激烈的繁简字使用范围的问题，2000年10月31日《中华人民共和国国家通用语言文字法》（以下简称《通用语言文字法》）颁布，以法律形式认定“国家通用的语言文字是普通话和规范汉字”（规范汉字包含所有的“一简字”）、“国家推广普通话，推行规范汉字”，同时也对繁体字、异体字的六种使用场合和范围进行了专门认定：文物古迹用字，姓氏中的用字，书法、篆刻等艺术作品用字，题词和招牌的手书字，出版、教学、研究中需要使用的字，经国务院有关部门批准的特殊情况用字可以使用繁体字和异体字。《通用语言文字法》的颁布给这场汉字繁简之争的论战画上了暂时的句号。《通用语言文字法》从国家层面上再次肯定了简化字的正字地位。它规定在社会日常通用层面的书面交际中必须以简化字为规范汉字，而在社会特用层面的书面交际中仍可以使用繁体字。这明确表达出国家对繁简字的态度：在我国现代语言生活中，繁简字并非敌对关系，各有其使用范围与生存空间，简化字是正字，繁体字是特殊用字，它们是并存共用的。因此，“识繁写简”“识繁用简”的观点暂时离开了大众视野。汉字繁简之争的第三回合中，虽然在国家层面上简化字最终胜出，但支持“识繁写简”“识繁用简”的观