

# 汉字部件学

沈克成 沈迦 著



机械工业出版社

# 汉字部件学

沈克成 沈 迦 著

机械工业出版社

为了增强全社会的语言文字规范意识,实现新时期识字教学的科学性;为了改变中文信息处理技术领域的“万马奔腾”局面,实现编码方案的优化,本书作者参照国家语言文字工作委员会最近颁布的有关规范,对汉字部件进行了科学的量化考察和研究,分析了汉字的笔画、笔顺、结构;指出了汉字部件学在识字、写字、检字、打字等方面的应用;介绍了本书作者发明的颇具规范性、教育性、普及性的计算机汉字编码——沈码。

本书适用于中文爱好者,文字编辑工作者,语文教学工作者,汉字研究工作者,中文信息处理工作者及计算机录入员和操作人员。

## 图书在版编目(CIP)数据

汉字部件学 / 沈克成, 沈迦著 — 北京: 机械工业出版社, 1998. 12  
ISBN 7-111-06777-0

I. 汉… II. ①沈…②沈… III. 汉字信息处理参考资料 IV. TB 391

中国版本图书馆 CIP 数据核字 (98) 第 30238 号

出版人: 马九荣 (北京市百万庄大街 22 号 邮政编码 100037)

责任编辑: 盛君豪 李万宇 封面设计: 姚学峰

北京机工印刷厂印刷 · 新华书店北京发行所发行

1998 年 12 月第 1 版第 1 次印刷

787mm×1092mm<sup>1</sup>/32·5.5 印张·114 千字

0 001-3 700 册

定价: 9.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

# 序

张 普

沈克成和沈迦两位先生的《汉字部件学》是一部关于汉字部件的研究与应用方面的专著。我拜读了这部专著，仅就此浅述几点个人的见解。

一、什么是汉字部件？1994年12月7日由国家技术监督局公布的《汉语信息处理词汇 02 部分：汉语和汉字》(GB/T 12200.2-94)中对“汉字部件”的定义是：“由笔画组成的具有组配汉字功能的构字单位。现代汉字部件按是否成字可分为成字部件与非成字部件。例：成字部件：木心口也；非成字部件：彳亻卩礻。最近20多年以来，汉字部件的研究与讨论十分活跃。无论是对部件进行定量分析的规模，还是从事汉字部件教学实验的范围，或者是进行汉字部件分析研究的人员及所发表的论文数量，从哪一方面看都是历史上最繁荣的时期，当然也是最聚讼纷纭莫衷一是的时期。汉字的部件教学（包括中小学语文教学和对外汉语教学），特别是汉语信息处理中的汉字部件编码，都迫切需要有一个标准或规范，而标准或规范的制定，是一件极其困难的事情。教学与信息处理都需要部件，但是两种部件分析有同有异。部件分析并不是始于今日，在部件分析之前有字根分析，部件与字根又有同有异。部件与偏旁、部件与部首、部件与独体也都互有异同。此外，部件与汉字笔画的关系、部件与汉字结构的关系、部件的成字或非成字问题、部件的读音或称说的

问题等等都没有统一的认识。即使对上述问题的认识统一了，还有对汉字部件拆分的原则、拆下限、拆分方法以及汉字部件的结构理据与现代汉字字形的矛盾等一系列问题的认识需要解决。所以沈氏父子认为：“近年来，随着汉字理论研究的不断深入和中文信息处理技术的迅猛发展，在汉字学领域逐渐形成了一门新学科——汉字部件学”；“汉字部件学逐渐成了现代汉字学讨论的一个热点”；“汉字部件学的核心，就是提供一种科学、规范的现代汉字字形拆分法，即部件拆分法”；“汉字分解的目的是使识字教学和汉字信息处理融为一体，实现识字、写字、查字和打字四者合一”。这些认识来自近 20 多年部件分析研究的实践，正是这些实践为“汉字部件学”的提出和建立准备了必要条件。汉字部件学已经到了呼之欲出的前夜。

二、本书的上篇第一、第二两章讲“汉字的笔画”和“汉字的笔顺”，为讲“汉字的部件”做了铺垫。从第三章到第五章，一共花了 16 节的篇幅，详细论述了有关汉字部件的各个方面。诸如：汉字部件学的研究范畴、内容、意义、目的以及汉字部件的定义、类型、定名、汉字的结构、汉字部件的组配、汉字部件的拆分方法、原则、下限、汉字部件的规范、汉字部件编码的设计等均有详细论述。论述中作者客观公允地介绍了学术界主要的有代表性的观点和一些不同看法，作者书后还附有大量的参考文献和参考书目。对于没有疑义的观点，作者常常指出出处，予以肯定而不掠美，对于不同看法，作者一般不妄加褒贬，而是仔细介绍其由来，指出区别所在，当然作者自己也必定有所取舍，若有取舍，往往认真地心平气和地摆事实、讲道理。或者作者并不囿于成见，另有独到见解。其中，作者书中提到的许多观点我都

是十分同意的。例如：对于单笔画部件争论的热点的介绍(p32-34)、作者选择承认“单笔画部件”的理由、以及为“单笔画部件”确立合法地位的三点分析，我认为都是十分得当的，我本人也历来赞同这种看法。学术争论的公开性、平等性和心平气和是文明与进步的表现。在学术争论方面过于偏激，甚至盛气凌人、辱骂对手，往往掩盖着理论上的苍白、内心里的空虚或学术外的焦急。我以为本书所讲述的汉字部件学的内容是适宜的，其探讨比较敏感的学术问题的态度也是应该提倡的。

三、作者极力主张部件规范、拥护部件规范、并且使沈码处处符合部件规范，这在本书中是随处可见的。他们说：

“研究汉字部件学的主要目标是实现汉字电脑化和电脑汉字化。在汉字部件学领域，我们应该紧密结合传统汉字学的特点，在‘以形为主，兼顾音义’的前提下，遵循‘以部件为主，以笔画为辅的拆分方法，加强对汉字的结构分析，并实现对汉字部件的规范化和标准化。”(p90)。他们甚至把《汉字部件规范》的《汉字基础部件表》收为该书的附录。1998年5月1日《信息处理用GB13000.1字符集汉字部件规范》(以下简称《汉字部件规范》)开始实施以后，他们对“沈码”不符合规范的“容错”之处(参见本书下篇第七章第六节“容错的措施”)立即进行了修订，认为“由国家语委主持制定的《汉字部件规范》的发布实施，将对我国语言文字应用的规范化、标准化，尤其是中文信息处理中语言文字规范化、标准化起到积极的推动作用。”(p65)。作者沈克成先生在1998年4月27日的《计算机世界》上还专门发表文章《在规范面前不能有“容错”》，说“在修订过程中，不再设容错码。也就是说，在规范面前我们不应有容错的余地”。但我

倒是主张不要把规范和为了方便用户使用的“容错技术”绝对地对立起来，既要维护规范的严肃性，又要让还没掌握规范的用户得到方便，可以在“容错技术”的支持下用不规范的编码输入汉字，那么，有什么办法可以两全呢？我认为只要发挥电脑的智能化特点，再增加“纠错技术”就可以了。这样，输入法不但可以继续以容错技术方便用户，还使计算机具备了辅助学习的功能，随时帮助违反规范者学习正确的规定，提高规范的水平。

四、《汉字部件学》尚有不足之处。

1、下篇的“汉字部件应用”与上篇相比，显得比较单薄。在汉字编码应用方面，只写了“沈码”，或只以“沈码”为例，作为一门“部件学”的应用，显然不足，至少要分析到不同类型的有代表性的部件输入法。

2、在汉字部件应用的三个主要方面中，“汉字部件输入法”是谈得比较具体、比较充分的，另外的两个方面——“汉字部件检索法”和“汉字部件识字法”相比之下就欠缺一些。

3、书后虽然开列了一些“参考文献”和“参考书目”，但仍显不足，如果尽量搜集全，附列一个文献索引和书目索引就更好了。另外，历史上曾有的部件分析统计表如能放在附录中，也是会受读者欢迎的。

本书填补了汉字学的一个方面的空缺，也许有修订的机会，我希望看到作者的补充，看到汉字部件学的发展和完善。

张普先生是我国著名中文信息处理专家，北京语言文化大学语言信息处理研究所所长、教授，中国应用语言学会理事，中国语现代化学会理事，中国中文信息学会理事，计算语言学专业委员会副主任。

# 前 言

信息时代，电脑的普及面越来越广，电脑使用者的主体也已发生根本性变化，逐步形成从小学生到老教授，从个体户到公务员等各年龄段、各阶层的社会用户群。

随着电脑应用的日渐广泛及计算机教育的逐渐普及，作为发展现代信息产业关键性技术之一的中文信息处理技术已经成为大家关注的热点。但是，在这一领域数百种编码所形成的万码奔腾局面，给文字使用和信息处理带来了混乱。各个编码方案对汉字的拆分、部件的选取和归类上自成体系，有的甚至对汉字肆意“肢解”。这种混乱局面给计算机教育和识字教育造成很大困难，既不利于计算机应用的发展，也不利于语言文字的统一规范。为此，国家语言文字工作委员会最近相继颁布了《现代汉语通用字笔顺规范》和《信息处理用 GB13000.1 字符集汉字部件规范》，于 1998 年 5 月 1 日起开始实施，这对规范中文信息处理、增强全社会的语言文字规范意识，都具有重大意义。

汉字编码主要分音码、形码和音形码。音码的优势是易学，儿童从小就学拼音，用拼音输入中文自然顺理成章。但音码也有缺点，一是同音字太多，输入效率低；二是对用户的发音要求较高；三是难于处理不识的生字。但是随着一批智能产品和优秀软件的相继问世，随着中文输入跨进了“以词输入为主导”的境界，重码选择已不再成为音码的主要障碍了，所以在中小学校，我们应该不遗余力推广音码。

形码由于重码率低，不受方言干扰，易于实现完备性与快速性兼顾，深受人们的青睐。多数形码采用“部件归类和



拆分”的设计思路，但由于汉字的字数多，部件多，笔数和笔顺都有严格的限制，致使形码方案往往难学难记。首先是部件的键位设计难，把几百个部件归类到 20 多个键位上，绝不是轻而易举的事；其次是输入操作时的拆字难，需要耗费较多的心理操作。如果这些部件的归类和拆分，与我们在学校里所学的背景知识不统一，与源远流长的汉字文化传统缺乏内在的传承，就会使形码的学习难上加难。

要实现形码方案的优化，必须对部件进行科学的量化考察与研究，必须使形码方案符合规范性。在信息时代，规范性是评价汉字编码及其它信息处理系统的重要标准，也是使用者选择输入系统的首要条件。它既关系到使用汉字传播信息的速度与质量，又关系到教育手段的更新及识字教学的科学性。

汉字部件规范是汉字规范的一个有机组成部分，它是指符合汉字构形规律，在部件的信息量和系统性上符合信息处理的需要，在理论上和实际操作中处理好汉字形体与理据的关系、历史与现实的关系、普遍规律与特殊情况的关系。汉字部件规范遵循“从形出发，尊重理据，立足现代，参考历史”的原则，定出符合规律、符合实际、符合优化的操作性条例。国家既然已颁布了两个规范，就要对编码有约束作用，几乎所有的编码都要因此作出相应的调整，有些偏离规范较大的编码方案也可能会因此而消亡。计算机用户应该怎样面对这一形势？答案只有一个：顺应潮流，向规范化、标准化靠拢。汉字键盘输入技术只有通过规范化、标准化，才能与国家的语言文字政策相一致，才有可能进入中小学教育，成为国民教育的一部分。

# 目 录

## 序 前 言

### 上 篇 汉字部件学研究

第一章	汉字的笔画	2
第一节	笔画的组合	2
第二节	笔画的形状	3
第三节	笔画的分类	6
第四节	笔画笔形的排序	9
	参考文献	10
第二章	汉字的笔顺	11
第一节	笔顺的形成	11
第二节	笔顺的运用	13
第三节	笔顺的规则	14
第四节	笔顺的规范	16
第五节	一些容易出错的字例	18
	参考文献	22
第三章	汉字部件学的研究	23
第一节	汉字部件的定义与特征	24
第二节	汉字的结构类型	25
一	独体字与合体字	25
二	单部件字与多部件字	30
第三节	汉字的部件类型	32
一	单笔画部件与多笔画部件	32

二	基础部件和合成部件·····	34
三	成字部件和非成字部件·····	35
第四节	汉字部件的定名·····	36
一	汉字部件的定位·····	36
二	成字部件的定名·····	37
三	非成字部件的定名·····	37
第五节	汉字的间架结构·····	41
第六节	汉字部件组成的空间关系·····	43
一	相离型·····	43
二	相交型·····	45
三	相接型·····	45
四	相聚型·····	46
	参考文献·····	47
第四章	汉字的部件拆分·····	48
第一节	部件拆分的原则·····	49
第二节	部件拆分的下限·····	52
第三节	部件拆分的方法·····	53
第四节	结构理据与部件拆分·····	56
第五节	汉字结构分析的不同层次·····	57
第六节	平面拆分法和层次拆分法·····	59
一	两种拆分法的异同·····	59
二	两种拆分法的比较·····	60
第七节	汉字部件的规范·····	64
	参考文献·····	67
第五章	汉字的编码部件·····	68
第一节	编码部件的定义·····	68
一	部件的归并·····	69

二	部件与偏旁的异同·····	71
三	编码部件与部首的异同·····	73
第二节	编码部件的确定·····	74
一	应以传统部首为核心·····	74
二	应作必要的字频分析·····	75
三	应以形为主兼顾音义·····	77
第三节	编码部件表的设计·····	80
一	数量应以 200 个为适度·····	80
二	以音托方式实现键盘映射·····	81
	参考文献·····	81

## 下 篇 汉字部件学应用

第六章	汉字部件识字法·····	84
第一节	我国识字教学的现状·····	84
第二节	识记单部件字是关键·····	85
第三节	按照部件的构字规律教学·····	86
第四节	部件识字法的特点·····	87
	参考文献·····	88
第七章	汉字部件输入法·····	89
第一节	沈码概说·····	91
第二节	代码的设置·····	94
第三节	单字的取码·····	102
一	独体结构汉字的取码·····	102
二	二分体结构汉字的取码·····	104
三	三分体结构汉字的取码·····	106
四	四分体结构汉字的取码·····	109

五	万能键的检索·····	109
第四节	词组的取码·····	110
第五节	简码的设置·····	111
第六节	容错的措施·····	113
第七节	关于笔画、笔形和笔顺的说明·····	116
第八节	关于层次拆分法的说明·····	119
第八章	汉字部件检索法·····	123
第一节	部件检索的设计·····	124
第二节	部件检索的方法·····	125
第三节	部件检索的优点·····	128
	参考文献·····	129
附录一	《信息处理用 GB13000.1 字符集汉字部件规范》 汉字基础部件表及其称说·····	130
附录二	汉字部件动态、静态统计表·····	152
附录三	沈码 V 4.0 部件及笔画代码表·····	160
	参考书目·····	162

上 篇

**汉字部件学研究**

# 第一章 汉字的笔画

所谓汉字笔画，是指用笔写字时笔头在纸上所画的线条。每写一个汉字，笔头一般要在纸上起落若干次。在每一次起笔、落笔的过程中，不管笔头在纸上所走的路线是横的还是竖的，是长的还是短的，是直的还是弯的，统称为一个笔画、一笔或一画<sup>①</sup>。掌握汉字的笔画数目、组合特点及形状，对识字、查字、打字均具有重要作用。

## 第一节 笔画的组合

笔画是汉字最小的构字单位。笔画的数目、形状、组合方式构成一个严密的系统。现代汉字的笔画系统以楷书为标准。是识字、写字教学的一项重要内容，是字典编纂、情报检索、文档管理中字、词排序的一种依据。

在通用汉字中，只有“一、乙”两字是由单笔画组成的，其余都是由多笔画构成的。每个汉字的笔画数都是一定的，多一笔少一笔就可能成为另一个字，如“鸟”字多一点就成了“乌”字，“目”字少一横就成了“日”字，“目”字多一撇又成了“自”字。计算笔画数的原则是：在同一笔画上，笔头只能走一次，不能走回头路。笔尖离纸，加算一画；写横

只能由左向右，写竖、写撇、写捺只能由上向下；按照由上到下，由左到右的笔顺，不管多长均可算作一画。

每个汉字的笔画组合也是一定的，同样的笔画放错了位置，组合方式不同，就会成为另一个字，如“犬”字的一点放到下面就成了“太”字。汉字笔画之间的不同组合关系，具有区别字形的作用。例如“人、入、八”三个字，虽都由一撇一捺组合而成，但因笔画之间的结构关系不同而构成不同的字。

汉字笔画的组合关系有相离、相交、相接等三种。

相离：如“二、三、八、川、小、心”等字。

相交：如“十、丈、力、七、九”等字。

相接：分两类，即：一个笔画的笔首或笔尾与另一个笔画的身段相接，如“人、入、刀、丁、上、正、乍、己、卫、久”等字；一个笔画的笔首或笔尾与另一个笔画的笔首或笔尾相接，如“厂、了、几、己、口、弓”等。

有些字的笔画间并非只有一种组合关系，如“大”字里有相交、相接，“国”字里有相离、相接，还有相交。掌握汉字的笔画组合关系，对准确计算汉字笔画数有很大帮助。

## 第二节 笔画的形状

所谓汉字笔形，是指汉字笔画的形状，或称汉字笔画的种类。每个汉字的笔画形状是一定的，该撇则撇，该捺则捺，如“千”字的起笔写成横就成了“干”字，“天”字的末笔写成钩就成了“无”字，“刀”字的撇笔写长了就成了



“力”字，“末”字的起笔写短了就成了“未”字，“己”字的竖弯钩上端稍长一点就成了“巳”字、如再长一点封了口就成了“已”字。

把笔画的基本形式和变形加在一起，就有许多不同的笔形。汉字的笔形种类繁多，有二十多种，但汉字的笔形并无确定的数量，无确定的名称<sup>[2]</sup>。所以只能给基本笔画一个较为习惯的称呼，如：

横 heng 竖 shu 撇 pie 点 dian 折 zhe

钩 gou 捺 na 提 ti

如果给笔形以规范的定名、定量，不仅便于称呼分辨，也利于指导书写和排序，是汉字规范的一个重要组成部分。汉字的笔形分为两大类：一类是不曲折的笔形，称为平笔笔形；一类是曲折的笔形，称为折笔笔形。平笔笔形有六种，即“横、竖、撇、点、捺、提”。折笔笔形一般依起笔笔形而有不同的称谓，如起笔是横的称为横折，起笔是竖的称为竖折，起笔是撇的称为撇折等等。

### 一、平笔笔形(基本笔画)

类别	名称	笔形	例 字
横	长横	一	苦、十、喜、册
	短横	一	慧、地、天、日
竖	长竖	丨	上、中、目、丰
	短竖	丨	工、刊、贡
撇	平撇	一	夭、后、千
	斜撇	丿	人、方、徐
	竖撇	丿	周、月、师
点	斜点	丶	主、内
	竖点	丶	空、心