

基于定量分析的 对外汉字学

王汉卫
苏印霞 著

清华大学出版社



基于定量分析的 对外汉字学

王汉卫 著
苏印霞

清华大学出版社
北 京

内 容 简 介

本书基于对《汉语国际教育用音节汉字词汇等级划分》的定量研究,以《等级划分》3000字为主要对象,研究了汉字的笔画、部件、结构、理据、语义、多音字、同音字,以及外向型检字法等相关问题。本书数据翔实、观点新颖、直接面向汉语国际教育的需要,适合于汉语国际教育从业者及相关专业的学生阅读。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。
版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

基于定量分析的对外汉字学 / 王汉卫, 苏印霞著. —北京:
清华大学出版社, 2015
ISBN 978-7-302-39698-7

I. ①基… II. ①王…②苏… III. ①汉字 - 对外汉语教学 -
教学研究 IV. ①H195

中国版本图书馆CIP数据核字(2015)第059399号

责任编辑: 张莹
封面设计: 傅瑞学
责任校对: 王凤芝
责任印制: 刘海龙

出版发行: 清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址: 北京清华大学学研大厦A座 邮 编: 100084

社总机: 010-62770175 邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者: 三河市金元印装有限公司

经 销: 全国新华书店

开 本: 170mm×240mm 印 张: 12.75 字 数: 225千字

版 次: 2015年5月第1版 印 次: 2015年5月第1次印刷

定 价: 42.00元

产品编号: 058123-01

序

王汉卫教授把他与其妻苏印霞合作的《基于定量分析的对外汉字学》专著打印稿送来，并嘱我为之作序。面对汉卫兄的请求，难以推辞，我想原因有二：

首先是我们的交往甚深，情不容辞。记得是在2000年夏天，他刚从南开大学研究生毕业，来暨南大学华文学院求职，我当时担任培训部负责人，经过简短交谈，发现此君思维敏捷，基础扎实，是难得人才。立即跟院方推荐，院长也欣然同意。孰料，王君狡兔三窟，多处求职，左右逢源，在其他因素影响之下，最终去了天津新蕾出版社当编辑。我闻之叹息，并料定编辑小儿读物非其志也。果然，一年之后，他给我来信，抱怨工作无聊，学非所用。我鼓励他跳槽并再次向院方推荐，终于引进成功。汉卫兄来后心无旁骛，一心钻研教学科研，羽翼渐丰，成绩斐然。我们一起出差期间，常常联床夜话，切磋学术，二人见解相同、惺惺相惜的时候居多，但也不乏各有所据，互不相让之时。不过每次聚谈皆有收获，他的奇思妙想给了我不少启发，深感此兄外显粗犷却心细如发。他后来跟随郭熙教授，在职攻读了博士学位。近年来又主持了教育部和国侨办的重大项目，在华文测试和汉字教学研究方面，都取得不俗的成就。第二个原因是我自己对汉字研究也饶有兴趣，也出版过对外汉字研究和教材数种。故得到新作，迫不及待，一睹为快。

读毕此书，最突出的感受是，王苏伉俪的基础研究，做得全面、深入、细致。此前汉卫兄已发表了多篇汉字研究的论文，有了先期研究作为基础，著书时就得心应手了。汉字是国际汉语教学所面临的最大难点，也是外国人掌握汉语最突出的薄弱环节。过去，汉字教学作为汉语教学的附庸，没有独立的地位，教学效果不如人意。人们逐渐意识到，汉语教学必须依照汉语汉字的特点展开，才有可能提高效率。近年来，涌现了一批汉字和汉字教学的研究成果。跟前贤的研究相比，此书独辟蹊径，从汉字教学的实际需要出发，以教育部和国家语委发布的《汉语国际教育用音节汉字词汇等级划分》3000汉字表为范围，依据详尽可靠的统计分析数据来展示汉字的笔画，部件，结构，理据，多音与同音，字、语素、词的关系，语义类等特点，甚至还研制了3000汉字的外向型检字法，为汉字教学及研究提供一个坚实的理论基础。

笔画是汉字的基础，过去学界重视不够。汉字的笔画到底有多少？从各家的主张来看，张志公 26 种，傅永和 29 种，林祥楣 30 种，张静贤 31 种，张斌 32 种，胡裕树 33 种，邢福义 35 种，黄伯荣、廖旭东 36 种，众说纷纭，难成一统。本书从 3000 基本汉字出发，以宋体为宗，突出笔位的观念，分析统计了对外汉语教学用字（即三千字）的笔画种（32 种），并进而提出对外汉字教学笔画种（24 种）和基础笔画种（18 种）。他们还进一步统计了 18 个基础笔画的综合价值，发现高达 97.4%，这就科学地确立了对外汉字的笔画教学的重点。

再比如笔顺，虽然关于笔顺问题的研究和见解没有在本书中专门阐述，有一次我听他的课，也发现汉卫兄也是见解独到，不妨一述。汉语教学中广泛使用的一种方法是汉字分笔画抄写，不仅教材这样编排，有些教师上课时也喜欢带领学生口唱笔画名称，一笔一画地书写或书空。海内外都对这种做法习以为常，沿袭至今。胡适先生说，“做学问要在不疑处有疑”，汉卫兄发现古人书写同一汉字的笔顺往往各不相同，因此质疑所谓的笔顺规范教学到底有多大价值，对分笔画抄写的做法明确提出了反对意见：“不少汉字笔画繁多，分笔画抄写不胜其烦，直接导致了学习者的厌倦。”事实上，汉字的形音义结构才是汉字认知的核心，对笔画序的过度关注，可谓因小失大。

汉字的理据分析也是一个热点问题，近年来对外汉字教学界涌现了不少汉字溯源、联系古文字分析汉字理据的书，它们通常印刷精美，图文并茂，通过古字形（主要是象形字）讲述先民造字的道理和故事，中国学生读来或有兴味，但问题是，古今文字形态变化很大，讲解文字理据的办法在对外汉字教学中究竟有多大作用？是不是最有效的教学方法？万业馨教授曾在一次学术研讨会上不无嘲讽地指出，“如今越是没有研究过文字学的人，越喜欢讲字源、说理据。”其实讲字源、说理据、谈文化，常能活跃课堂，引发兴趣，也是可行的手段，但我们需要把握一个合适的度。为了掌握好这个度，作者认为，首先需要研究的就是 3000 字的理据性到底有多高。他们研究了 3000 现代汉字的构字方式，主张分为形义字、会意字、形声字、会意兼形声、唯形字、唯声字、无理据字这七种。综合计算，常用汉字的表义表音理据度仅各有 0.41，而且初级汉字表义度很低，仅有 0.33，表音度更低，只有 0.27。如果考虑到古今字形的巨大差异，现代汉字的理据就更不明显了。显然，如果在汉字课上一味侈谈造字理据，追本溯源，甚至牵强附会，对学生掌握现代汉字本身作用并不大，还占用了宝贵的课堂教学时间。其实，汉字的构造无论多么精巧，充满了多么有趣的文化，从宏观来看，依然只是记录语言的一种符号，如果过度关注符号本身，而忽视了符

号的运用，那就有些得不偿失了。前辈学者苏培成先生明确提出“不是每个字都能说出字理来，说不出字理的字就不要勉强谈字理”；“更不能强作解人”。本书通过理据度的统计，顺理成章地提出“谈理据要谨慎，要有选择”。为了帮助汉字教师进行选择，教材还细心地梳理出可用于理据教学的 155 个形义字，并呈现了古字形。这种踏踏实实为教学实践服务的精神，值得赞许。

书中创见颇多，再举一例：在汉语教学界，曾经有过字本位、语素本位、词本位的争论，作者指出：“争论的根本原因在于字、语素、词的不重叠：有些字没有意义，只能跟其他字连绵在一起表示一个意义；有些字有意义，然而却不能单说单用，是语素而不是词；有些字同时既是语素也是词。这样就产生了字、语素、词之间的差异，进而带出学术界关于本位的争论。”可谓条分缕析，切中肯綮。

书中也有百密一疏之处，比如计算表音部件的表音度时说：“第四级，不表音，表音部件的声母及韵母跟整字完全不相关的，计 0 分。”道理似乎完全正确，但如果细致考察，有些字的表音部件虽跟整字的声韵母都不一样。如我（wo）跟饿、鹅、俄、哦、娥、蛾、峨（e）等字声韵母虽无相同之处，却能明显提示发音。汉字古今形音变化很大，有些隐藏的规律还需细察。如果一定要说此书的缺憾，窃以为是遗漏了“偏旁”一章。汉字的主体是形声字，偏旁在其中凸显表义、表音功能，甚至说它是解开汉字的一把钥匙也不为过。实际上，作者也清醒地意识到偏旁的重要性，并在“部件”一章中指出：“部件是观察汉字构成单位表音表义属性的一个角度，但跟偏旁的角度相比，也许不是一个好的角度。”我们期待未来修订再版时，重新梳理部件与偏旁的关系，补上偏旁这个“好的角度”，以成完璧。

大作即将付梓面世，可喜可贺！又当与汉卫兄相约小酒馆，浮一大白了。

周健

2014 年 11 月 18 日于香港中文大学

目录

第一章 3000 字的笔画 / 1

第一节 3000 字的笔画种 / 2

第二节 笔画的排序 / 7

第三节 对外汉字教学的笔画 / 9

第四节 基础笔画的遴选 / 11

第五节 不同笔画数的拥字量 / 12

第二章 3000 字的部件 / 16

第一节 部件拆分的原则和结果 / 17

第二节 表义部件 / 18

第三节 表音部件 / 23

第四节 唯形部件 / 25

第五节 部件的形序 / 25

第三章 3000 字的结构 / 30

第一节 3000 字中的独体字 / 30

第二节 按部件数量的分类 / 33

第三节 按结构方式的分类 / 37

第四章 3000 字的理据 / 44

第一节 3000 字的造字法分析 / 46

第二节 3000 字的构字法分析 / 48

第三节 3000 字的理据总结 / 52

第五章 3000 字的多音与同音 / 65

第一节 多音字的定义和统计原则 / 66

第二节 多音字的统计结果和分析 / 68

第三节 简化多音字的教学 / 73

第四节 同音字问题 / 86

第六章 3000 字范围内字、语素与词的关系 / 107

第一节 连绵字、语素字与词字 / 108

第二节 字与语素的对应关系 / 113

第三节 字与词的对应关系 / 117

第七章 3000 字的语法语义类 / 126

第八章 以 3000 字为总体的外向型检字法 / 136

第一节 外向型检字法的特点 / 136

第二节 检字元素的选取和检字方案的确定 / 137

第三节 “四码方案”的效果 / 142

参考文献 / 153

附录 1 《汉语国际教育用音节汉字词汇等级划分》汉字表 / 158

附录 2 《汉语国际教育用音节汉字词汇等级划分》词汇表 / 162

后记 / 192

第一章 3000 字的笔画^{*}

笔画的研究和教学应该是整个对外汉语教学中至关重要的一个环节,是汉字教学,乃至汉语教学成功的起点和保证。遗憾的是,也许学术界认为“笔画”太简单了,为二语教学的笔画本体研究并不多见,而从留学生的汉字习得情况来看,笔画的教学算不上成功,这从现有大量关于笔画习得偏误的研究上可以得到证明。

笔画是汉字的基础,学习汉语汉字之初,在很短的一段时间内,就应该系统地学习到,并且准确、熟练地掌握常用笔画的形态。从理论上说,这不是什么“高标准,严要求”,而从实践上看这近乎天方夜谭(汉字圈的学生除外),母语是拼音文字的学生,在学习汉字很长一段时间后,笔画错误仍是一个不容忽视的问题。

施正宇(2000)曾呼吁汉字教学应该重视笔画。江新、柳燕梅(2004)的研究表明,拼音文字背景的留学生基于字形的错误在各类汉字错误中最多。孙文访(2006)的研究进一步显示,在增加笔画、减少笔画、笔画形错三种笔画层面的错误中,“笔画形错”占的比例最高,占全部笔画层面错误的一半左右。尽管笔画层面的错误仅占笔画、部件、整字三种错误类型的 1/4 左右,但基于笔画是汉字构形上的基本元素,可以合乎逻辑地认为,对笔画的认识不清会导致更多其他方面的汉字错误。

重视笔画教学首先需要重视笔画本体的研究工作。一直以来,汉字研究主要集中在习得及偏误分析上,本体方面的研究虽有张静贤(1987、1990)、刘永山(1990)、

^{*} 本章主要内容发表于《世界汉语教学》2012 年第 2 期,本次成书有扩充修改。

陈子骄(2011)等,总的来说为数不多,有些基本问题从现有的学术成果内还找不到答案。其次,重视笔画教学应该致力于寻求笔画教学的合理性,应该“以简驭繁”,既能使笔画教学为汉字教学打下坚实的基础,又能使笔画教学本身不令学习者厌烦。有鉴于此,我们需要明确三个不同层次上的概念:

1. 对外汉语教学用字笔画种——这个概念指的是依据对外汉语教学用字的相关“大纲”得到的笔画种类。

2. 对外汉字教学笔画种——这个概念指的是在“对外汉语教学用字笔画种”的基础上,经过筛选得到的具有教学意义的笔画。

3. 对外汉字教学基础笔画种——这个概念是对“教学笔画”的深化,是更优先、更重点教学的笔画。所谓“基础”,包括三层意思:首先,笔画主要是初级阶段的教学内容,基础笔画应该来自基础字;其次,基础笔画是整个笔画系统的基础,对非基础笔画具有笔画形态和组合方式两方面的基础意义;最后,常用性也是基础笔画选取的一个重要考量。

下文拟以上面的三个概念为线索,以对外汉语教学为背景,以对外汉字教学为服务目标,主要讨论这三个概念下的相关问题。本书将尽可能详细地提供一些笔画方面的基本数据,希望能为对外汉字教学和研究提供一些参考。

第一节 3000 字的笔画种

笔画是“书写工具从和书写材料接触到离开,这期间书写工具的移动在书写材料上留下的痕迹就是笔画”(苏培成,2001:65),简单地说,就是笔的一次起落留下的痕迹。笔形是指“笔画的形状”(《现代汉语词典》第72页;《汉语大词典》卷八第1161页)。本书严格区别使用这两个概念。

1.1.1 前人及时贤的笔画分类

汉字笔画有多少,这是一个十分基础的问题,也是一个非常复杂的问题,历来的说法不尽相同。我们统计了12家的笔画情况,具体如下:

26种(张志公,1982:74)、29种(傅永和,1992)、30种(林祥楣,1995:485)、31种(张静贤,1987:496)、32种(北京大学中文系,1995:157;张斌,2004:83)、33种(胡裕树,1987:196;武占坤,1985:136)、35种(邢福义,1986:19)、36种(黄伯荣、廖序东,1997:197;张威,1998;孙汝建,2003:22)。其中,

张斌、孙汝建分别跟较早出版的北大版和黄伯荣版的笔画数量和具体笔形完全相同，下文分析有差异的其余 10 家的笔画构成。

统计显示，10 家累计涉及笔画 46 种，如果 46 种是必须的，10 家总计应该有 460 种次，而实际上 10 家累计是 321 种次。4 种笔画为一家独有，而 10 家共有的笔画也仅有 17 种，就是说各家笔画的广义共同度为 69.78% (321/460)，狭义共同度仅为 36.96% (17/46)，具体如表 1-1 所示。

表 1-1 10 家笔画的同异

独 / 共有	笔画种	笔 画
一家	4	丨 (悬针竖) 乚 (用于乙亿艺等) 冂 (月的第二笔)
两家	2	丶 (撇点) 一 (短横)
三家	3	丿 (斜撇) 丨 (短竖) 丶 (挑点)
四家	1	ノ
五家	3	㇏ 丶 ㇀
六家	7	㇏ ㇀ ㇁ ㇂ ㇃ ㇄ ㇅
七家	3	㇏ ㇀ ㇁
八家	2	㇏ ㇁
九家	4	㇏ ㇏ ㇏ ㇏
十家	17	㇏ ㇏ 丨 丶 ㇏ ㇏ ㇏ ㇏ ㇏ ㇏ ㇏ ㇏ ㇏ ㇏ ㇏ ㇏ ㇏

除却可能的统计失误，笔画差异的主要原因有以下几点。

第一，统计范围的差异。上文列举的多家数据，只有张静贤 (1987) 明确说明统计对象是《印刷通用汉字字形表》的 6196 字，其他多家大都没有明确的统计来源，有的表明只是举例或从俗性质的。例如，“楷书笔画系统举例”(武占坤,1985:136)，“现代汉字的笔画有多少呢，一般认为……”(林祥楣：1995:485)。”理论上说，统计对象不同，得到的结果自然会有差异，例如，GB13000.1 字符集中有横捺 (㇏)、撇钩 (ノ)，这两个笔画不要说常用字范围，即便一般中型语文词典上也找不到它们的踪影。但根据我们的统计，从常用字范围，一直到中型语文词典万字左右的收字范围，笔画种类都是稳定的。而事实上，有的笔画表甚至不包括例如 “㇏、㇏、㇏、㇏、㇏、㇏” 等常见笔画，这有可能是统计范围太小，但更有可能是举例，或其他原因。

第二，字体的差异。尽管宋、楷两种最常用字体的笔画系统相差不多，但还是会带来一些差异，例如 “㇏、丨、㇏”，依据楷体则有，依据宋体则无。有的笔画表

叠加了宋、楷两种字体，笔画种自然会比较多。

第三，笔画的归属不同。例如，“丶”（‘当’的第三笔），有人认定为撇、有人认定为点，“乚”，也有点和挑（提）两种归属。但归属的不同不会产生笔画种的差异。

第四，对笔画位（以下简称“笔位”）的把握不同，这是带来笔画种不同的主要原因。笔位的概念虽然并不见于以往的论述，但实际上已经运用在以往的笔画研究中，《现代汉语词典》、《新华词典》等语文工具书上的笔画定义也秉持了笔位的理念，例如，“撇”在《现代汉语词典》（第5版）的解释是“向左斜下，形状是‘丿’”（第1047页）。这样的定义就是规定了“撇”这个笔画位，是各种“撇”的总括。有些学者实践了笔位这个概念，有些学者则没有，或仅是部分地实践了这个概念。

笔位这个概念的提出来自于音位的启发。辨别音位的三原则是辨义功能、互补分布、音感差异，笔位同样可以依据这样的思路来界定，只不过需要把“音感差异”改为“笔形差异”。例如，长横短横^①、长竖短竖悬针竖，再例如折笔画的“乚、乚”等，这些成对成组的笔画既无辨义功能，与此同时，也有互补分布和笔形近似的特点，是“横、竖、横折钩”三个笔位。至于长横、短横等只是笔画变体的差别，而不是笔位的不同。突破笔位走向笔画变体，笔画的种类就难有共识也不胜其繁。例如，单单一个“派”字，认真看起来，四个撇四个模样，大小长短斜度各不相同，直撇、平撇、斜撇仍不够描写具体的“撇”。

1.1.2 对外汉语教学的笔画分类原则

延续上文的讨论，对外汉语教学的笔画分类原则表述如下。

第一，明确3000字的汉字范围。尽管在一个相当大的弹性空间内笔画种是稳定的，但不限定范围最终将产生不同（见上文），另外，不同的范围得到的笔画频率等数据也可能会有差别。面向汉语二语教学，需要以业内的“大纲”为依据，《汉语国际教育用音节汉字词汇等级划分》（以下简称“等级划分”）是最新的音节、汉字、词汇表，本研究暂以该大纲的汉字表3000字为笔画拆分的对象。

第二，明确字体范围。基于下面的两个原因，本书以宋体为宗：一是宋体为最常用字体，各种报刊书籍，包括教材、词典，都以宋体为最常用。二是笔画的讨论为书写服务，也为阅读、检索汉字服务，书写有较大的灵活性，而阅读、检索则仍

① 长横短横只在“土、士”这一个情况下有辨义功能，只能算是个例外。

以宋体为主。

第三,明确笔位的观念。笔画种的确定即笔位种的确定,由于折笔形笔画以直笔形笔画为基础,而以往笔画种的差异又主要集中在直笔形笔画上,所以我们有必要明确6个直笔形笔位的界定。

横:直线状笔画,水平或有一定斜度。

竖:直线状笔画,垂直或有一定斜度。

提:楔子状笔画,左下粗右上细。

点:雨点状笔画,上尖下圆。

撇:尾巴状或楔子状笔画,右上粗左下细。

捺:扫把状笔画,左上细右下变粗再变细。

上面的界定不同于《现代汉语词典》等,没有给出“代表笔形”,只是提取了具有共性的笔形特征。例如“点”,不管是左点、右点,还是长点都符合上面的定义。结合“撇”的定义,所谓“撇点(丷)”也就可以明确地归为“撇”。至于“挑点(㇀)”,实在是一个笔画综合了“挑”和“点”两种笔形,上半部是“挑”、下半部是“点”。参考大多数专家的看法,本文将“㇀”归为“提”。

明确了直笔形笔位,就明确了每一个直笔形笔画变体的归属,进一步也就明确了折笔形笔位——不同直笔形笔位的首尾组合(含不是独立笔位的“钩”和“弧”),或相同直笔形笔位的不同方式组合(例如弯和折不同、左向右向不同等),就构成了不同的折笔形笔位。

1.1.3 3000字范围内的笔画种类

根据上面的界定,可以统计出3000字范围内的笔画种。

统计程序:①3000字按笔画排序;②标注每个字的笔画数;③分拆并记录每个字的笔画;④以笔画数来核对分拆结果,每个字分拆出的笔画数等于按现有相关规范规定的笔画数,3000字分拆结果应该等于3000字的总笔画数;⑤核对无误后得出3000字的笔画种及其按使用频次的排序。示例如表1-2所示。

用这种方法,我们就可以将3000字准确分解成笔画,并进一步统计出每一种笔画的使用情况,如下表1-3所示。

统计显示,3000等级字总笔画为28358,笔画种32,平均每字9.45画。除了笔画种,上表的统计也明确了对外汉语教学用字范围内每一个笔画的使用状态,为进一步整理教学笔画奠定了基础。

表 1-2 拆分统计方法示例

等级字	横	竖	撇	点	捺	提	横折	横钩	横撇	横折折	横折钩	横折提	横捺钩	横折折折	横撇弯钩	横折折折钩	横折折折撇	竖钩	竖提	竖折	竖弯	竖钩	竖折折	竖弯钩	竖折撇	竖折钩	撇横	撇提	撇点	捺钩	笔画数量
一	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
七	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2
九	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
了	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2
二	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
人	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
儿	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2
入	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
八	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
几	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
刀	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
力	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
十	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
厂	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
又	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
万	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
三	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
上	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
下	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
个	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
久	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
么	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3
义	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
也	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3
习	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
乡	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	3
于	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	3
亿	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
千	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
笔画数量	16	7	17	4	7	1	0	1	2	0	5	0	0	0	0	3	0	0	2	0	0	0	0	3	0	0	0	3	0	0	71

表 1-3 按使用频次的笔画排序

排序	笔画	名称 ^[1]	使用次数	使用频率	排序	笔画	名称	使用次数	使用频率
1	一	横	7687	27.11%	17	乚	撇点	82	0.29%
2	丨	竖	4813	16.97%	18	㇏	捺钩	78	0.28%
3	ノ	撇	4661	16.44%	19	㇀	横折提	77	0.27%
4	丶	点	3936	13.88%	20	㇚	撇横	63	0.22%
5	㇇	横折	1852	6.53%	21	㇄	横撇弯钩	62	0.22%
6	㇏	捺	939	3.31%	22	㇆	横折弯钩	47	0.17%
7	㇀	提	798	2.81%	23	㇄	弯钩	41	0.14%
8	㇇	横折钩	678	2.39%	24	㇚	竖弯	28	0.10%
9	㇚	竖钩	550	1.94%	25	㇆	横捺钩	22	0.08%
10	㇀	横钩	441	1.56%	26	㇏	横折弯	21	0.07%
11	㇚	竖弯钩	349	1.23%	27	㇆	横折折折钩	20	0.07%
12	㇀	撇提	316	1.11%	28	㇆	横折折撇	14	0.05%
13	㇇	横撇	314	1.11%	29	㇚	竖折撇	4	0.01%
14	㇚	竖提	212	0.75%	30	㇏	横折折	1	0.00%
15	㇚	竖折	161	0.57%	31	㇆	横折折折	1	0.00%
16	㇚	竖折折钩	89	0.31%	32	㇏	竖折折	1	0.00%

第二节 笔画的排序

笔画排序应该体现出笔画的价值。表 1-3 是严格按照使用次数的排序。使用频率最高的是横, 占了总笔画的 27.11%; 其次是竖、撇、点, 仅仅这四个笔画的使用频率就占到了总笔画的 74.4%。排在前十位的笔画使用频率占总笔画的 92.94%, 而排在最后的 10 个笔画的使用率合起来也不足 1%。可见, 频率序对认识各笔画种的价值很有好处。

然而, 仅仅按频率序认识各笔画种的价值还不够严谨, 除了使用次数外, 笔画的构字次数也是认识笔画价值的一个重要指标。使用次数跟构字次数高度相关, 但毕竟不是一回事。表 1-4 是反映构字率^[2]的数据。

[1] 笔画名称采用张静贤(1987)的命名,撇折例外,分化为撇提和撇横。撇折可以说是过去的统一认识,但“红”的第一笔、第四笔跟“东”的第二笔不同,前者是撇提,后者是撇横。

[2] 构字率的算法是每个笔画种的构字数除以所有笔画种的构字总数 (16310)。

表 1-4 笔画的构字频次

排序	笔画	构字数	构字率	排序	笔画	构字数	构字率	排序	笔画	构字数	构字率
1	一	2742	16.81%	12	ㄣ	292	1.79%	23	ㄣ	41	0.25%
2	ノ	2511	15.40%	13	ㄥ	209	1.28%	24	ㄥ	28	0.17%
3	丨	2407	14.76%	14	ㄤ	200	1.23%	25	ㄥ	22	0.13%
4	丶	2038	12.50%	15	ㄥ	152	0.93%	26	ㄥ	21	0.13%
5	ㄣ	1504	9.22%	16	ㄥ	87	0.53%	27	ㄥ	20	0.12%
6	ㄥ	929	5.70%	17	ㄥ	78	0.48%	28	ㄥ	14	0.09%
7	ㄥ	765	4.69%	18	ㄥ	78	0.48%	29	ㄥ	4	0.02%
8	ㄥ	635	3.89%	19	ㄥ	77	0.47%	30	ㄥ	1	0.01%
9	ㄥ	531	3.26%	20	ㄥ	63	0.39%	31	ㄥ	1	0.01%
10	ㄥ	411	2.52%	21	ㄥ	62	0.38%	32	ㄥ	1	0.01%
11	ㄥ	339	2.08%	22	ㄥ	47	0.29%	-	-	-	-

对比表 1-3 和表 1-4 发现，大部分的笔画不管按哪种标准排序位置都不会改变，但少数笔画还是有不同。就构字能力来看，“撇”比“竖”优先，“横撇、竖提”比“撇提”优先。

上面的情况说明需要将笔画的使用频率和构字频率这两个指标综合起来，这样才能更有效地反映一个笔画的使用价值。将同一笔画的使用率和构字率数据相加再除以 2，便可得到一个反映笔画综合价值的序列，如表 1-5 所示。

观察按综合价值的排序，有下面几点值得我们注意。

表 1-5 按照综合价值的笔画排序

排序	笔画	综合价值	排序	笔画	综合价值	排序	笔画	综合价值	排序	笔画	综合价值
1	一	21.96%	9	丨	2.60%	17	ㄥ	0.39%	25	ㄥ	0.11%
2	ノ	15.92%	10	ㄣ	2.04%	18	ㄥ	0.38%	26	ㄥ	0.10%
3	丨	15.87%	11	ㄥ	1.66%	19	ㄥ	0.37%	27	ㄥ	0.10%
4	丶	13.19%	12	ㄣ	1.45%	20	ㄥ	0.31%	28	ㄥ	0.07%
5	ㄣ	7.88%	13	ㄥ	1.02%	21	ㄥ	0.30%	29	ㄥ	0.02%
6	ㄥ	4.51%	14	ㄥ	1.17%	22	ㄥ	0.23%	30	ㄥ	0.01%
7	ㄥ	3.75%	15	ㄥ	0.75%	23	ㄣ	0.20%	31	ㄥ	0.01%
8	ㄥ	3.14%	16	ㄥ	0.42%	24	ㄥ	0.14%	32	ㄥ	0.01%

第一，尽管通常“横、竖”并提，但“撇”以其自身的综合实力插入其中。

第二，尽管通常以平笔画为基础笔画，但“横折”的价值远在“捺、提”之上。

第三，尽管有个别笔画远远超出其同类的价值，但总体而言，从0折笔画到4折笔画，随着折点的增加，笔画的种类及综合价值都有十分巨大的变化和差异。就载荷来说，0折笔画是主体，就笔画数量来说，1折笔画是主体，二者加起来，综合价值和笔画数量都构成了笔画体系的主要组成部分。特别是价值载荷，0折和1折的18个笔画，载荷高达93%以上，2~4折的14个笔画，载荷尚不足7%。如表1-6及图1-1（同折点的笔画按综合价值排序）所示。

表 1-6 按折点的笔画分类

折点	笔画种	笔 画	综合价值
0 折	6	一ノ丨丶丶丶	75.20%
1 折	12	㇏㇏㇏㇏㇏㇏㇏㇏㇏㇏㇏㇏	18.33%
2 折	8	㇏㇏㇏㇏㇏㇏㇏㇏	5.42%
3 折	5	㇏㇏㇏㇏㇏	1.12%
4 折	1	㇏	0.01%

以表1-6为基础，得到图1-1。

第三节 对外汉字教学的笔画

一般操作上，笔画教学发生在汉语学习的初级阶段，上文得出的32个笔画不都具有教学价值。“教学用字的笔画”跟“教学笔画”是两个不同的概念，前者是对教学用字的笔画描写，后者则要进一步考虑教学阶段、教学策略、笔画价值等因素，从描写出的32个笔画中筛选出具有教学价值的笔画。

第一层瘦身是删除非基础汉字的笔画。所谓教学笔画，其前提必须是基础汉字的笔画，如果笔画栖身的汉字不是基础字，这样的笔画显然不具备初级阶段汉语教学的价值，否则，笔画教学就脱离了汉字教学的需要，就成了为笔画而笔画。

不管按照“等级划分”的一级（初级）900字，还是按照《汉语水平词汇与汉字等级大纲》的甲乙级1604字，初级阶段都没有使用到“㇏、㇏、㇏”，因此完全可以排除。删除这三个笔画后，笔画只剩下29种。

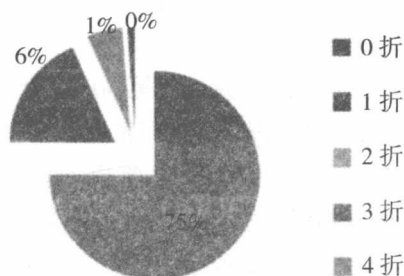


图 1-1 不同折点笔画的综合价值