

现代汉字部首法所用单字切分的研究

陈 燕

(天津师范大学 文学院, 天津 300387)

摘 要:现代汉字部首法所用的切分主要用于合体结构字中,对选取部首具有重要的意义。专门的研究仅有20多年的历史,表现出现代汉字部首法体系趋向缜密并向深处发展的趋势。在切分的依据、原则和难点这些重要问题的研究中,我们发现按照汉字外部结构切分可能会最大限度地覆盖汉字,同时使内部系统性与社会应用性得到统一。

关键词:现代汉字部首法的切分;切分依据;切分原则;切分难点

中图分类号:H124.5 **文献标识码:**A **文章编号:**1671-1106(2006)04-0071-06

自《说文解字》建立部首法开始,切分的内容就隐含在归部之中,并由归部结果体现。《说文》是据义归部,当然其切分的结果也是据义的,表义就是它的理据性所在。20世纪80年代切分问题被明确的提出[1],这是据形定部的取部原则发展到一定阶段的产物。把单字切分独立进行研究,表现出现代汉字部首法日益趋向缜密、所涉及内容正向深处拓展的态势,将更加方便应用。

切分主要用于合体字,对选取部首有重要的意义。它是取部之前难以回避的先行工作,是定部的基础。有关这方面的研究历史较短,有一些文章涉及到切分问题^①,但是很少见到专门讨论的文章。在语言文字应用领域所定的

具体目标,除了有很好的实践支撑外,还应当有相应的理论支持,否则这个体系就会缺少深度,从而缺乏说服力。我们在这方面有所贡献。笔者经过三年多的研究,围绕现代汉字部首法所用单字^②的切分,就以下三个问题进行讨论:切分依据;切分原则;切分难点。期冀在现代汉字部首法切分的理论探讨上抛砖引玉。

一、切分的依据

目前对这个问题的见解大致有四种类型:按照部首切分;依照习惯按照较大部位切分;按照上下或左右大体平衡切分;依照汉字结构切分。它们的共同点都是根据汉字的字形而不是

收稿日期:2006-02-20

基金项目:国家语言文字应用“十五”科研项目“汉字部首规范”(YB105-30)阶段性成果。

作者简介:陈燕(1953-),女,河北河间人,天津师范大学文学院教授。

① 参见曹乃木《统一部首查字法需要解决的主要问题》,载《文字改革》1985年第1期;杭苇《部首查字法的几个问题》,载《辞书论集》,知识出版社,1987年;程养之《再谈实施统一部首的步骤》,载《语文建设》1993年第2期;汤宋岐《统一汉字部首表(征求意见稿)读后》,载《语文建设》1994年第6期;苏宝荣《汉字部首排检法规范化试探》,载《辞书研究》1995年第5期;李青梅《从康熙字典的归部看汉字的归部》,载《语文建设》1997年第2期等。辞书检索系统也有一些研究成果,例如《现代汉语规范词典》“部首检字表”和《学生规范字典》“部首检字表”等。

② 本文所用单字限于20,902个汉字的范围内。字符依据国家语言文字工作委员会《GB13000.1字符集汉字笔顺规范》,上海教育出版社,1999年。

字义切分,据形切分已经成为现在应用者的共识。主要分歧是根据汉字字形的哪部分内容切分,不同的切分可能会造成相异的归部结果,因此,选取切分的依据很重要。下面我们对各种见解作一番梳理,以确定较为合理的切分依据。

按照部首切分要求在规定的取部范围内,见到部首就切分,新版《辞海》^①已经有所实践,有创新之举,对于推广和深入研部新部首法发挥了重要的作用。其不足主要是不考虑汉字的结构,较多顾及到文化水平较低的学习者选取部首的方便而割裂偏旁。以《辞海》为例:将怎归入亻部,把不应该割裂的偏旁“亻”切开;将奔、拿、龕、盒、愈等字归入人部,把偏旁“合”割开;剪、孳归入八(丿)部,把偏旁“前兹”割裂;劣,归入小部,把偏旁“少”割开;壅、饗归入宀部,割裂了偏旁“雍”;翁、贫归入八部,割裂了偏旁“公分”等。以上例字皆将汉字硬性地拆分,破坏了汉字的结构,势必丧失部首法的优势。以上切分也不符合使用汉字者的习惯。现在,几乎每个读者在小学阶段都养成了分析字形的习惯,这种析字方法影响到汉字使用者拆字习惯的形成。在部首法切分时,应当将其作为重要的参考。另外,不论使用哪种切分都要考虑国家语言文字的实情,现在的情况是:“教育逐渐普及,人民文化水平有了很大提高。2000年全国已普及九年义务教育,青壮年文盲率降至4.8%,成人文盲率降至8.72%。”^[2]现在的文化水平与1949年相比有了极大的提高,对文字规范的要求也相应提高。通过以上分析,我们认为以部首为中心切分的原则现在已经不完全适用。

依照习惯按照较大部位切分,主要指“贫案”类字从“分安”而不是从“八宀”切分,前者就是较大部位。这种切分结果有道理。但是我们注意到此说将“依照习惯”放在首位,遣词比较模糊,在操作上容易与别的取部条例发生矛盾。明言实践这个主张的辞书举例为:想,查心部;

凯,查几部;桑,查木部等,主要在右和下的位置取部。依此而论,洋、样、佯、们等字因为右边占的面积较大,应当做部首;但是左边也是部首,其书的查字方法又有“左右有部首的,查左不查右”。在这种情况下,应当取左还是依据较大部位取右为部首?因此,我们认为这种切分方法语意模糊而前后矛盾,较难把握。

按照上下或左右大体平衡切分,也是以单字的面积来衡量的。提出这个主张的辞书举例为:暮,从“𡗗 日”处三分;喜,切分为“吉 𡗗”;壽,切分为“𡗗 𡗗”。既不考虑汉字结构,又与习惯相差较远。

后两种说法出现的时间较晚,都是以单字的面积为切分依据。所不同的是,前者是较大面积,后者为大体平衡,都属于硬性切分,较难把握。虽然各有长处,但是都不完善。

部首法符合汉字结构的特点^②,所以能够长久流传。使用汉字的人掌握了汉字结构,可以从偏旁的组合中提取部首,认识汉字的全貌。我们由此悟出汉字部首法的切分也应以汉字结构为依据。

一些学者在他们的相关文章中已经涉及到汉字结构问题。杜定友将汉字分为四种类型,分别是单体字、左右字、上下字、杂式字,然后按照类型组合检索。^[3]苏宝荣提出两种切分的类型,主要是二分定位归部和三(多)分定位归部。^[4]李青梅也认为,有些合体字的取部应该符合汉字的结构规律。^[5]因此,学者们对于按照汉字结构切分具有一定的认同感。

关于汉字结构的具体内容,目前的研究有两种类型:外部结构和内部结构。^[6]外部结构又叫构形法,由汉字构形角度出发有多种划分。^③我们采用傅永和的分类。他分为:独体字、左右结构、上下结构、全包围结构、半包围结构、特殊结构。^④内部结构又叫构字法,从汉字

① 特指20世纪六七十年代出版的《辞海》,与旧版《辞海》相对而言,以下简称《辞海》。

② 史存直在《试论汉字排检问题》(华东师大学报1960年第1期)一文中曾指出:“旧部首法的最大长处在于它是从汉字的形体结构中归纳出来的,它能符合汉字结构的特点。”

③ 除了杜定友、苏宝荣的论述外,梁东汉分为3式:上下组合式、左右组合式、内外组合式(见《汉字的结构及其流变》,上海教育出版社,1981年)。史存直分作3类:左右并列关系、上下叠置关系、内外包孕关系(见《试论汉字排检问题》,《华东师大学报》1960年第1期)。

④ 见傅永和《中文信息处理》,广东教育出版社,1999年。他在该书中提出粗略和较细的两种分类,粗略的分类如以上所说,还有较细的分类,共14类。他讨论汉字结构的论文还有:《汉字结构及其构成成分的分析 and 统计》,载《中国语文》1985年第4期;《汉字的结构》,载《语文建设》1991年第9期等。

构成的理据出发,传统分为形声、会意、象形、指事。两者相比较,前者更加符合楷书汉字的特点,适用于现代汉字部首法。现代汉字如果从汉字构成的理据性出发,常常需要溯源,这为切分带来一定的困难。现行汉字中有相当一批字难以溯源,譬如简体字,历代积淀的简体字有相当数量。现代“我们的规范汉字,是指简体字系统”[2]。还有偏旁混同等其他因素,所以“用溯源分析代替现状分析是不科学的”[6](P64)。因此,构形法更适用现代汉字。我们认为,现代部首法以检索为主要目的,重在实用,也应当以汉字的外部结构作为切分的主要依据,换种说法就是根据字形切分。它不但可以覆盖现代汉字,而且也能覆盖传承汉字。

依据汉字结构切分的现实意义在于:与笔形笔画和四角号码式的切分相区别;与按照部首或左右大体平衡的切分相区别;按照层次切分,凸显汉字的结构特点。

二、切分的原则

汉字部首法的切分与汉字部件的拆分,其对象都是汉字,两者之间尽可能地保持一致非常必要。汉字部首法切分应当充分借鉴汉字部件拆分的原则。部件拆分顺序基本是从整字到偏旁,再到部件,属于穷尽性的拆分。拆分出的部件有合成部件和基础部件之分,合成部件大致与复合偏旁相当。汉字部首法切分的顺序也是从整字到偏旁,但是取到部首就终止切分。切分的目的就是合理地提取部首。

部首法切分规则与部件拆分的规则相同,相离或某些相接关系可以切(拆)分,相交关系不可以切(拆)分。⊙⊙则不同,部件拆分可以从相接处折出笔画,但是部首法一般不能从相接处拆出笔画。如:天,不能切分出一和大,而取起笔单笔部首时虽然取到笔画,但并非切分所得。两者之间的差异完全由它们的目的所决定,部首法切分以合理而科学地选取部首为目的,带有一定的特殊性,而部件拆分为提取部件。在部首法初期和之后的千百年,相交关系

是可以切分的。例如《辞海》,本、末、未、朮取木部,夫、夷、央取大部,内、冉、册、内取门部,独体字一般穿过交叉的笔画进行切分,《辞海》之前的历代辞书基本都是如此。我们采纳了交叉不拆的部件拆分原则,对传统的切分内容有所更新,不单为了与部件拆分保持一致,而是注重将现代科学的拆分理念注入到古老的汉字部首法中,使之更加科学实用。

部首法切分的规则为避免混乱还要求:相离和相接关系并存时,首先切分相离关系,例如急,先从心处切分;有多条相离的分隔沟^①,先从最长的分隔沟切分,例如颖,在页处切分;左中右和上中下结构字要多分,首先在最左、最上的分隔沟切分,例如:颞,在最左的男处切分;曼,在最上的日处切分。

部首法依据汉字结构切分与上述切分规则并不矛盾。规则要求相交或某些相接的笔画团不拆,相交的笔画团必然没有分隔沟,从汉字结构角度分析都是独体字,不可以切分;相离关系的可拆,合体字的偏旁之间基本都有分隔沟;相离的先于相接的切分,需要切分偏旁时,先切分相离的,没有相离的再切分相接的。

切分的目的、依据和规则构成了切分的原则。依据和规则为目的服务,如果其间发生不可避免的矛盾,应有合理的解释。

三、切分的难点

切分的难点集中在以下3类字:剪、案、贫、鹏类;太、勺、凡、刃类;复合偏旁。这些字难以顺利地切分,影响到选取部首,很有讨论的必要。

(一)关于剪、案、贫、鹏类字的切分

这类字还有盒、贫、壅、慈、冤等,从20世纪80年代被专门提出来[1],一直是研究的难点。有两种切分的方法可以选择:在八、宀、人、宀(非成字偏旁)处切分;在前、安、分、合、今、雍、兹(成字偏旁)处切分。总的倾向是按照成字偏旁切分,一些学者解释为有理有据切分。

有理有据切分与按照汉字结构切分能否画等

^① 分隔沟的说法采自苏培成,见参考文献[6]。

号,应当有所辨识。它不仅关涉到切分依据的取舍,而且还影响到能否设立统一的标准。

汉字理据性可以参考部件拆分术语“结构理据”的解释：“根据字源或参考字源，从汉字的部件组合中分析出造的字意图，称结构理据”[7](P2)。汉字的理据性由结构理据获得，需要根据或参考字源。而据此对楷书汉字进行切分常常需要溯源，以体现造字意图。例如戌，追溯到小篆，《说文》字形分析为“从人持戈”，是会意字，据此进行的切分才是有理据的切分。但是楷书字形已经难以看出这样的结构理据，很难进行有理据切分。因此，有理据的切分方法不能覆盖全部汉字。根据我们研究大约从《龙龕手鏡》开始，全部汉字就已经不能完全依据汉字理据进行切分了。但是，在部分字中还存在着有理据切分的现象，例如《康熙字典》中颖、颍从顷处切分，腾、滕、𨔵从朕处切分，褰、𦐇、𦐈从衣处切分等。现在这种现象在现代较为著名的辞书中已经非常少见。^①

而据形切分的结果与上述不同:戊,是独体字,不能切分;颍、颍、腾、滕、滕从左右最长的分隔沟处切分;褒、衷、褒从最上的分隔沟处切分,基本能与现代辞书的取部结果保持一致。这虽然不是有理据切分,但是不等于无根据的切分。我们认为应当禁止无根据的随意切分。

这就是两者的区别。虽然绝大多数汉字在汉字结构切分与有理据切分之间可以画等号,因为它们切分的结果是一样的,例如:励、台、倘、外、彰、婴、潘、辆、眉、牦、扣等,但是某些汉字不能完全画等号。从一致性出发,我们不赞成在部首法中选择有理据的切分。

我们把剪、盒、案、贫、贪、壅、慈、冤、鹏类字从刀、皿、木、贝、贝、土、心、龙、鸟处切分，而没有在最左最上的分离位置切分，并非不考虑从最上和最左的分离位置切分的规则，当然也绝不是从理据性出发。我们认为必须考虑使用汉字人们的切分习惯。他们熟悉常见字，从识字教学中养成了较为固定的切分习惯，习惯于从成字偏旁切分。从顺应文字约定俗成的规律出

发,对此不能漠然,在这个问题上原则性与灵活性应当统一。

我们认为,常见字的成字偏旁在首次切分时一般不宜切开,为便于使用可以酌情类推。例如:贫从分和贝处切分,盆、忿也从相同位置切分。但一定要用“常见字”限制,表示与据义切分相区别。若遇到僻字而且可以用两种方法切分的字,难以判断应当从何处切分,只好另行处理。例如:苾、莧、茱、蕙、蕝、藟、益、洩、湑、淑、灏、漱、缬、櫚、榭、纈、壩、噉、瞭等皆为不常见字,而蕙、噸是较常用字,他们都有两种以上的切分方法。这类字似乎没有固定的切分习惯,宜遵从一般的切分原则,从最左、最上处切分。

(二)关于太、勺、凡、刃类字的切分

这类字还有氐、丹、叉、术、戍、尤、书、甫、卞、良、永、办、必和正、下、元、开、百、千、壬、𠂔、𠂔、𠂔、𠂔等字，简称做“太类”和“正类”字。

太类字有 20 余个,特点是有单独的点笔依附在字符上。我们曾将点笔看做汉字构成的一部分,把太、勺、凡、刃、氏、丹、叉按照包围结构,术、尤、书、甫按照左右结构,卞、良、永按照上下结构处理,感到不理想。

我们将视角转向现代汉字的研究,希望从他们的研究成果中获得启迪。一些研究汉字部件拆分的学者主张:单独的点笔虽是相离关系,但一般应附属于它临近的结构块。^② 我们认为很有道理。单独的点笔虽然呈分离的状态,但切分之后难以成为一个独立的部分。也难以称说其位置,作为附属成分倒是适宜的。

正类字的首笔是笔画，其下面多是复笔部首。最初我们采纳了《辞海》的切分模式，皆分析为上下结构，从首笔单笔笔画处切分。但是后来改变了看法。

范可育等人认为,如果单独的笔画可以独立成字,则独立作部件(如旦、画、乱等字),否则一律附属于临近的结构块。[8](P102)陈爱文、陈朱鹤说:相粘连的笔画一般属于一个结构块。[8](P117)我们认为,由单笔笔画与一个笔画结构块(有的是成字部件)相接不再切分的说法比较

① 只有《汉语大字典》、《汉语大词典》仍将褰、衷、裊等字从衣切分。

② 持此说的有：陈爱文、陈朱鹤，范可育，陶晓东等。

适合汉字部首法的切分,决定采用。例如千、再、午、亚、下、乏、丙等字,不再切分。而相离关系的则应切分,旦、画、乱等字从分隔沟处切分。

上面两类字都是独体结构字^①不能切分。因此,在切分中确定独体结构字非常重要。我们把具有以下特点的字判断为独体结构字:单笔字,如乙、一;单独的点笔附属于它临近的结构块,如书、义、凡、叉;相交的笔画组合,如由、来、长、东;单笔笔画与一个笔画结构块(有的是成字部件)相接,如平、垂、尺、亚;平行的或大致平行的相离笔画,如二、川、州、乡。

从部首法角度看,不能切分的范围可以扩大到以下四个方面:部首不拆,如宀、羊、鼻、香;与部首字形极为接近的异体字不切分,如尸、黄;镶嵌结构和某些特殊结构的字不切分,如疆、爽、巫、乖、承;独体结构偏旁与独体单字的切分一致,如豫的左旁,歪的上头下底都不切分。

按照汉字结构切分,一般情况下必须坚持在合体结构字而不在独体结构字切分的原则。从构形角度所说的独体结构字与传统的独体字不尽相同。例如:向、兆,古代是象形字;习、发,现代是简化字。它们在《基础教学用现代汉语常用字部件规范》^②中都是部件而不拆分。但是从现代汉字角度看不宜作独体结构字,《学生规范字典》一律定为合体字。^[9]退一步说,如果向、兆可以作独体结构字,但是习、发明显不能作独体结构字。习与匀结构相同,发与灰、友结构相同。拿相类的字形比较就看出能否切分,能切分就不宜看做独体结构字。

(三)关于复合偏旁的切分

复合偏旁的术语由傅永和提出^[10],1999年出版的《中文信息处理》还在使用,我们没有找到他对这个术语的解释。陶晓东解释为:由两个以上分离的笔画结构块组合的成字或不成字偏旁。^[11]我们认为这个解释比较合适。

杭苇在《部首查字法的几个问题》中专设了“关于字的切分问题”,他强调“最理想的是切分

为二,取其一”。^③范可育进一步提出:两分“是最科学、最方便的分解方法”^{[8](P106)}。他们的论述极有代表性。汉字中左右、上下结构字占绝大多数,两分法符合汉字的结构特点,容易掌握。但是有些字不可以两分,如曼、夔、囊、羸、羸等应该三分或多分。

上述认识适用于简单结构字的切分,但是遇到由复合偏旁构成的字就不那么简单了。例如:甬、蠡、蝨、硎、够都有包围偏旁,甬、硎还有独体偏旁。如何切分?独体偏旁与独体结构字应当保持一致,都不能切分;包围偏旁的内外位置隐讳难以取部而不宜作切分。因此,甬、硎、蠡、蝨首先两分,切分出的偏旁不是部首;再对偏旁两分,甬、硎的偏旁由独体偏旁和包围偏旁组成,不能切分,只能从起笔位置取单笔部首,蠡、蝨的上头是包围结构偏旁不作切分,下底可以两分取“虫”作部首。比较复杂的是𧈧字,需要多分,仍取不到部首;然后在两端切分,上不能切分,就对由“分”构成的下底两分;不取中间的“八”,取“刀”为部首。蠡,也不取中间的“日”为部首。上述切分结果具有较深的应用基础,请看现代较有影响辞书的取部结果:

蠡、蝨、𧈧,《汉语大字典》、《汉语大词典》、《现代汉语规范字典》、《现代汉语规范词典》、《中华字海》取虫为部首。

𧈧,《辞海》^④、《汉语大字典》、《汉语大词典》、《现代汉语规范字典》、《中华字海》取刀为部首。

甬,《汉语大字典》、《汉语大词典》、《现代汉语规范字典》、《现代汉语规范词典》、《中华字海》取一为部首。

在合体结构字中出现的复合偏旁,体现出汉字结构的复杂性。我们曾经在四角设立取部位置,企图从四角切分取部。但是后来发现不适合就不用了,改用偏旁分析法。对整字或偏旁采用两分法或者多分法之后取不到部首,要对复合偏旁进行切分,遇到独体偏旁或包围(含

① 与传统的独体字、合体字略有不同,因此本文不用旧称。

② 见《基础教学用现代汉语常用字部件规范》(征求意见稿)。

③ 载《辞书论集》,知识出版社,1987年。文章说杭苇参加了1964年“汉字查字法整理工作组”下设的四个小组之一的部首查字小组。当时指定由辞海编辑所制定方案,由杭苇具体负责。

④ 《辞海》的取部位置和部首与其他5本辞书不同,其归部结果常常无法作比较的对象。

全包、半包)偏旁则不作切分。

四、结束语

确立合理的切分依据和原则,促使切分标准获得统一,其重要性不言而喻。如果汉字都是简单的左右、上下、包围结构字,如:环、岩、患、冑、笨、深、把、周、还、困类字,占据了汉字的绝大多数,一次切分就取到了部首,那么切分就变得简单而不值得研究了。切分难点的研究主

要集中在少数需要多次切分才能取到部首的复杂结构字,以及一些较难确定切分位置的独体结构字中,虽然字数较少,但是阻碍了切分问题的彻底解决,因而在部首法切分中占有较重要的地位。研究切分问题,要考虑简单和复杂的切分,确定相对合理并一以贯之的,能够覆盖全部汉字的切分标准。合理的切分标准应当符合汉字的特点,基本符合使用汉字人群的切分习惯,符合汉字部首法的发展趋势,就会使内部系统性与社会应用性获得高度统一。

参考文献:

[1] 汉字部首排检法工作组.统一部首表(征求意见稿)[J].文字改革,1983(11).
[2] 李宇明.规范汉字和规范汉字表[J].中国语文,2004(1).
[3] 杜定友.查字法问题的基本认识[N].光明日报.1961-05-31.
[4] 苏宝荣.汉字部首排检法规范化试探[J].辞书研究,1995(5).
[5] 李青梅.从康熙字典的归部看汉字的归部[J].语文建设,1997(2).
[6] 苏培成.现代汉字学纲要[M].北京:北京大学出版社,2001.
[7] 国家语言文字工作委员会.信息处理用 GB13000.1 字符集汉字部件规范[Z].北京:语文出版社,1998.
[8] 苏培成.现代汉字学参考资料[Z].北京:北京大学出版社,2001.
[9] 李行健.学生规范字典[Z].北京:中国大百科全书出版社,2001.
[10] 傅永和.汉字的部件[J].语文建设,1991(12).
[11] 陶晓东.也谈部件的定义[J].语文建设,1993(6).

The Research on Single Character Segmentation of
Modern Chinese Character Radical Method

CHEN Yan

(The Literature Collage, Tianjin Normal University, Tianjin 300387, China)

Abstract: The segmentation of Modern Chinese Character Radical Method is mainly used in the compound characters, which plays an important role in the selection of radical. The related research has been carrying out only a little more than 20 years, which shows the trend of Modern Chinese Character Radical Method to be deeper and more deliberated. On the present achievements, the paper integrates the hard points encountered in the practice, brings up the solutions, makes definite the idea of to separate character on the basis of outer structure, puts emphasis on the unification of systematicness and social application.

Key words: the segmentation of Modern Chinese Character Radical Method; the basis of segmentation; the principle of segmentation; the hard points of segmentation