

语言文字规范标准研究

几部语言文字规范中的问题探讨

邵霁吉 盐城师范学院文学院

[摘要] 近几十年以来,国家发布了多部语言文字规范,其中有一些规范适时进行修订,推出了更加完善的新版。但有些规范公布之后一直没有修订,比如《信息处理用 GB13000.1 字符集汉字部件规范》《现代常用独体字规范》《现代常用字部件及部件名称规范》等。随着时代的发展,科学研究的进步,适时修订这些语言文字规范,是很有必要的。

[关键词] 语言文字规范;修订;部件;独体字

一、引言

近几十年以来,国家发布了多部语言文字规范。随着时代的发展,科学研究的进步,有些已经发布的语言文字规范又经过修订,有了新的版本,新版本比旧版本更加完善。比如《标点符号用法》就有 1951 年 9 月中央人民政府出版总署公布的 1951 年版,有 1990 年 3 月国家语言文字工作委员会、中华人民共和国新闻出版署修订发布的 1990 年版,有 1995 年 12 月国家技术监督局批准并发布的 1995 年版,有 2011 年 12 月中华人民共和国国家治理监督检查检疫总局、中国国家标准化管理委员会发布的 2011 年版。再如《出版物上数字用法》有 1987 年版、1995 年版、2011 年版,《汉语拼音正词法基本规则》有 1988 年版、1996 年版、2012 年版,等等。但也有些语言文字规范标准自公布之后就一直没有修订,比如《信息处理用 GB13000.1 字符集汉字部件规范》《现代常用独体字

规范》《现代常用字部件及部件名称规范》等。笔者认为,精益求精、适时修订这些语言文字规范标准,是很有必要的。本文拟对几部语言文字规范标准中的一些问题作一点探讨,期望能为这些规范标准的修订做一点参考。

二、关于《信息处理用 GB13000.1 字符集汉字部件规范》

《信息处理用 GB13000.1 字符集汉字部件规范》是国家语委 1997 年 12 月 1 日发布的,它规定了 GB13000.1 字符集汉字的“基础部件”560 个,归纳为 393 组^[1]。20 多年来,这部规范在中文信息处理领域的设计、管理、科研、教学和出版等方面发挥了很好的作用。但它也存在一些小的瑕疵。如果修订《信息处理用 GB13000.1 字符集汉字部件规范》,笔者有以下几个建议。

(一) 建议增加基础部件的名称

《信息处理用 GB13000.1 字符集汉字部件规范》列出了独立使用的基础部件 560 个,但是没有给这些基础部件命名,没法称说,给它的使用和教学造成一定的困难。

胡乔木《关于当前文字改革工作的讲话》曾经提出部件“要能够通用,并便于称说”的问题,他说:应该“减少汉字的部件,并尽可能使汉字的部件独立成字;不能独立成字的部件,也要能够通用,并便于称说”^[2]。而《信息处理用 GB13000.1 字符集汉字部件规范》没有给出 560 个基础部件的名称,肯定是不“便于称说”的。

据专家回忆,“《信息处理用 GB13000.1 字符集汉字部件规范》研制时,已经提出对部件称说的有关规范。在规范报批审议时,国家语委标准化委员会对于部件的称说部分提出缓议,需要进一步调研,并建议在调研的基础上首先要研制供中小学语文教学使用的部件称说规范,在此基础上进一步规范信息处理用的 GB13000.1 字符集汉字部件称说。”^[3]但是,到现在,20 多年过去了,尚未见到《信息处理用 GB13000.1 字符集汉字部件规范》增加“部件名称”,不“便于称说”的状况依然如故。

所以,我们希望在修订《信息处理用 GB13000.1 字符集汉字部件规范》的时候,增加汉字基础部件的名称,以“便于称说”。

(二) 建议修订“《汉字基础部件表》使用规则”

《信息处理用 GB 13000.1 字符集汉字部件规范》中有一个“《汉字基础部件表》使用规则”,但其中有些内容没有紧扣“使用”讲规则。其中第一条(5.1)、第二条(5.2)讲的都是“如何从整字中拆分出部件”的:

5.1 用本部件表进行部件拆分时,相离、

相接可拆;交重不拆(可拆成笔画)。极少数不影响结构和笔数的笔画搭挂,按相接处理。

5.2 字形符合理据的,进行有理据拆分;无法分析理据或形与源矛盾的,依形进行无理据拆分。对多部件的汉字进行拆分时,应先依汉字组合层次做有理据拆分,直至不能进行有理据拆分而仍需拆分时,再做无理据拆分。

这两条“使用规则”侧重于讲整字的“拆分”(如何从整字中拆分出部件)问题,跟“使用”(如何使用表中的部件)有一些距离。可能是把“从整字中拆分出部件的规则”和“《汉字基础部件表》的使用规则”两方面的内容杂糅到一起了。在这部规范(标准)发布 12 年后,教育部和国家语委 2009 年 3 月发布的《现代常用字部件及部件名称规范》,是把与这差不多的内容分开写的,一章是“部件拆分原则和规则”,一章是“《现代常用字部件表》使用规则”。

我们建议,将来修订时,可以像《现代常用字部件及部件名称规范》那样,把这一章一分为二,一章讲“从整字中拆分部件的规则”;一章讲“《汉字基础部件表》的使用规则”。

(三) 建议对 560 个基础部件作适当增删

《信息处理用 GB13000.1 字符集汉字部件规范》的“汉字基础部件表”共列出 560 个基础部件,我们认为,对它可以作必要的增删调整。

1. 建议将第 246 号部件“冂”一分为二,增加一个部件“𠂇”。该规范把“节、爷”的下边、“卫”字的上边部件“冂”,跟“印、卯、印”等字的右边部件“冂”合在一起作一个基础部件看待,我们认为不甚妥当。因为“𠂇”和“冂”在形体上有较大区别,其区别在于接点不同,“冂”两笔是“首首相接”,“𠂇”两笔是身段被接,分为两个基础部件最为妥帖。《信息处理用 GB13000.1 字符集汉字部件规范》也是重视这个区别的,比如 195 号部件“匚”和 196 号部件“匚”,就是因为接点不同而区分为两个部件的,

212 号部件“𠂇”和 213 号部件“ㄣ”也是因为接点不同而区分为两个基础部件的。

2. 建议将第 250 号部件“冊”一分为二, 增加一个部件“𠂇”(“𠂇”的右边)。因为“𠂇”的右半边不是“冊”, 而是一个比“冊”多一笔的部件, 而这个部件又是一个笔笔相交的部件, 不能把它跟所多的那一画拆分为两个部件。肖金卯、武云翠《关于〈信息处理用 GB13000.1 字符集汉字部件规范〉中若干问题的讨论》曾经指出: “在部件规范第 250 号部件‘冊’的例字中列有汉字‘𠂇’, 而‘𠂇’右边是两横, 而部件‘冊’只有一横, 不是一回事”^[4]。

3. 建议减少“食、缶、示、习、肉、玉、犬、貝、見”等基础部件。在《信息处理用 GB13000.1 字符集汉字部件规范》的 560 个基础部件中, 有多个基础部件还可以进一步拆分为两个或三个基础部件。我们认为, 从“减少汉字的部件”着眼, 凡能够拆分为几个“小部件”的“大部件”, 都应该继续进行拆分, 这样可以减少不少基础部件。比如“食”可拆分为[人]+[丶]+[艮](本规范中“良”已拆分为[丶]+[艮]), “缶”可拆分为[匚]+[十]+[凵](本规范中“午”已拆分为[匚]+[十]), “示”可拆分为[二]+[小], “习”可拆分为[冫]+[彳], “肉”可拆分为[内]+[人], “玉”可拆分为[王]+[丶](本规范中“王”必须拆分为[王]+[丶], 因为基础部件中没有“王”), “犬”可拆分为[大]+[丶](本规范中“太”已拆分为[大]+[丶]), “貝”可拆分为[目]+[八], “見”可拆分为[目]+[儿], 等等。

三、关于《现代常用独体字规范》

《现代常用独体字规范》是教育部、国家语委 2009 年 3 月 24 日发布的, “规范”给出了现代常用独体字的范围、定义、现代常用独体字规范的制定原则、确定规则、“现代常用独体字表”说明、“现代常用独体字表”^[5], 篇幅不长, 简明实用。但小瑕疵还

是有的。如果要修订《现代常用独体字规范》, 笔者有以下几个建议。

(一) 建议删除溢出“现代常用字”范围的两个独体字

《现代常用独体字规范》是“现代”“常用”的“独体字”的规范, 照我们的理解, “现代”就是“现代汉语”, 这是一个大的范围, “常用”就是“常用字”, 这是一个小的范围, 两者合起来, 它就应该“是现代汉语常用字”这个范围中的“独体字”的规范。

而“现代汉语常用字”的范围所指, 国家语委、国家教委是发布了一个“语言文字规范”的。这就是国家教委 1988 年 1 月发布的《现代汉语常用字表》, 该表规定了现代汉语中 3500 个常用字, 其中常用字 2500 个, 次常用字 1000 个。

不过, 《现代常用独体字规范》给出的 256 个独体字中, 只有 254 个在《现代汉语常用字表》中, 另有两个不在《现代汉语常用字表》中, 这两个独体字是“囟”和“𠂇”。这说明, 《现代常用独体字规范》中的这两个独体字, 溢出了“现代”“常用字”的范围。

国务院 2013 年发布的《通用规范汉字表》有“一级字表”, 表首“说明”指出: “一级字表为常用字集, 收字 3500 个”^[6], 这个“常用字集”中也没有“囟”和“𠂇”。

既然没法证明“囟”和“𠂇”是“现代汉语常用字”, 那么它们被收入《现代常用独体字规范》的“现代常用独体字表”, 就不是十分恰当的。应当予以删除。

(二) 建议修改“独体字”的定义

《现代常用独体字规范》给出了一个“独体字”的定义, 这个定义是这样说的: 独体字是“由笔画组成、不能或不宜再行拆分、可以构成合体字的汉字”。这个定义从“由笔画组成”“不能或不宜再行拆分”

“可以构成合体字”三个方面规定了独体字跟合体字的区别。

但这个定义是有问题的。苏培成(2014)《现代汉字学纲要》(第3版)从三方面分析了这个定义的“缺陷”。关于“由笔画组成”,苏培成指出:“所有的汉字,不论是独体字还是合体字,拆分到最后都是笔画(‘由笔画组成’),这不是构成独体字的必要条件。”关于“不能或不宜再行拆分”,苏培成指出:“什么样的笔画组合‘不能’切分,什么样的笔画组合‘不宜’切分,本身没有可操作性,还要有明确的界定。”关于“可以构成合体字”,苏培成指出:“我们检查了《规范》给出的256个独体字,发现其中有19个字在通用字范围内不能构成合体字。”^[7]我们也觉得其中“个、卫、书、凸、甩、再、年、秉”等独体字,确实不可以用于“构成合体字”,不符合《现代常用独体字规范》给出的“独体字”必须“可以构成合体字”的要求。如果在定义中坚持独体字必须“可以构成合体字”,那么有些该规范给出的独体字就不是独体字了。

我们觉得,“由笔画组成”“可以构成合体字”这两个限定语是画蛇添足,而“不能或不宜再行拆分”这一限定语基本上还是可以的。说得明白一点,“不能或不宜再行拆分”就是“不能或不宜再行拆分(出部件)”的意思,独体字就是“不能拆分出部件或不宜再行拆分出部件”的汉字。

(三) 建议删除可以分析为合体字的独体字

《现代常用独体字规范》给出的“现代常用独体字表”列出了“现代常用独体字”256个,按理来说,它们都是“不能或不宜再行拆分的汉字”,但是我们发现,其中有65个是完全可以拆分的,如果依据《信息处理用GB13000.1字符集汉字部件规范》,它们就是“合体字”,而非“独体字”,应该从“现代常用独体字表”中把它们删除。这65字是:

么、凡、义、卫、刃、叉、与、马、开、

天、无、云、太、尤、互、少、午、壬、升、
禾、勿、乌、六、斗、办、击、正、术、丙、
平、卡、生、矢、失、用、甩、匆、鸟、主、
半、头、出、再、百、页、夹、朱、血、囟、
亦、产、亥、农、严、来、卤、囟、羌、弟、
肃、隶、面、首、兼、鼠

依据《信息处理用GB13000.1字符集汉字部件规范》,这65个字每一个都可以拆分为两个或三个基础部件,例如“么”可以拆分为[丿]、[厶]两个部件,“凡”可以拆分为[几]、[丶]两个部件,“义”可以拆分为[丶]、[乂]两个部件,“刃”可以拆分为[刀]、[丶]两个部件,“叉”可以拆分为[又]、[丶]两个部件,“囟”可以拆分为[丿]、[口]、[乂]三个部件,“弟”可以拆分为[ㄣ]、[巾]、[丿]三个部件,等等。

《信息处理用GB13000.1字符集汉字部件规范》是1997年发布的语言文字规范,早于《现代常用独体字规范》12年,是大家都应该尊重的国家标准。我们觉得,国家标准就是国家标准,具有它的权威性,大家应该尊重它。

四、关于《现代常用字部件及部件名称规范》

《现代常用字部件及部件名称规范》是中华人民共和国教育部、国家语委2009年3月24日发布的,该规范给出了现代常用字的范围、汉字部件的定义、部件拆分的原则和规则、部件名称命名规则、“现代常用字部件表”说明、“现代常用字部件表”使用规则、现代常用字部件表、常用成字部件表,还有两个附录,附录A是“现代常用字部件构字数表”、附录B是“现代常用字部件笔画序检索表”^[8],内容比《信息处理用GB13000.1字符集汉字部件规范》丰富,特别是它给出了“部件名称命名规则”,又在“现代常用字部件表”中具体给出了各个部件的名称,便于

称说,增强了它的实用性。它明确了该规范的选字范围是“现代汉语 3500 常用汉字”,没有溢出“现代汉语常用字”的范围,名副其实。

如果修订《现代常用字部件及部件名称规范》,笔者有以下几个建议。

(一) 建议修改“现代常用字部件表”这个名称

“现代常用字部件表”是《现代常用字部件及部件名称规范》的核心内容,表中列出从现代汉语 3500 个常用字中拆分出来的部件 441 组 514 个,表头栏目有“部件、序号、组号、部件名称、例字”,比《信息处理用 GB13000.1 字符集汉字部件规范》的《汉字基础部件表》的栏目设置多出了“部件名称”一列,因而更合理、实用。

只是,我们觉得这个表的名称“现代常用字部件表”值得商榷,其中的“部件”应该改为“基础部件”。因为“部件”和“基础部件”是内涵和外延都不同的两个概念。“部件”是大概念,包括“基础部件”和“合成部件”。我们看《现代常用字部件及部件名称规范》的“现代常用字部件表”中的“部件”,应该都是些“基础部件”,不宜笼统地称为“部件”,应该称之为“基础部件”。《信息处理用 GB 13000.1 字符集汉字部件规范》不给出“汉字部件表”而给出“汉字基础部件表”,做法是严谨的。作为国家发布的语言文字规范,应该严格区分“部件”和“基础部件”两个术语,不能够混淆使用。

大家也都知道,汉字的“合成部件”数量很多,估计有上千个乃至数千个,如果《现代常用字部件及部件名称规范》的“现代常用字部件表”中的“部件”包括“合成部件”的话,那数目肯定会多得多,而不是“514”这个数。而且,该规范前面还讲过“合成部件”,所举的两个合成部件“相、音”都不在“现代常用字部件表”的“部件”中,也证明这个“现代常用字部件表”中的“部件”实际上就是个“基础部件”。所以,建议把“现代常用字部件表”这

个名称也改为“现代常用字基础部件表”。

(二) 建议改掉前后自相矛盾之处

《现代常用字部件及部件名称规范》的“现代常用字部件表”中还有一些前后自相矛盾的问题,需要改正。比如:

部件	序号	组号	部件名称	例字
卡	197	169	卡 (kǎ)	卡

部件	序号	组号	部件名称	例字
上	309	267	上 (shàng)	让卡叔威

既然认为“卡”是一个部件(序号为 197,组号 169),那它就是一个独体字,是一个“基础部件”,不能再拆分成两个“基础部件”;如果认为“卡”可以作部件“上”(序号 309,组号 267)的例字,那就是认为“卡”可以拆分为[上]和[卜],则“卡”是个合体字。表格内容自相矛盾。再如:

部件	序号	组号	部件名称	例字
朱	500	430	朱 (zhū)	珠株

部件	序号	组号	部件名称	例字
未	381	331	未 (wèi)	妹朱株

既然认为“朱”是一个部件(序号为 500,组号 430),那它就是个独体字,是一个“基础部件”,它就不能被拆分成两个“基础部件”;如果认为“朱”可以作部件“未”(序号 381,组号 331)的例字,那就是认为“朱”可以拆分为[丩]和[未],则“朱”就是个合体字,自相矛盾。再如:

部件	序号	组号	部件名称	例字
头	365	315	头 (tóu)	买实读牒

部件	序号	组号	部件名称	例字
大	63	55	大 (dà)	驮夸头因

既然认为“头”是一个部件（序号为 365，组号 315），那它就是一个独体字，一个“基础部件”，不能再拆分成两个“基础部件”；如果认为“头”可以作部件“大”（序号 63，组号 55）的例字，那就是认为“头”可以拆分为〔两点〕和〔大〕两个“基础部件”，则“头”是个合体字。表格内容自相矛盾。

诸如此类，前后矛盾的地方都应该改正。

（三）删除可以拆分的部件

《现代常用字部件及部件名称规范》的“现代常用字部件表”列出了部件 514 个，其中还有不少是可以拆分的，有的可以拆分为两个基础部件，有的可以拆分为三个基础部件。例如：

乇、么、勺、凡、义、卫、刃、马、开、天、无、云、太、尤、互、少、午、壬、升、夭、仑、今、勿、欠、风、乌、印、爻、六、斗、办、击、正、去、术、丙、平、卡、矢、失、用、甩、匆、卯、鸟、主、半、头、穴、司、出、发、丝、老、再、戌、百、页、夹、至、朱、舌、乒、兵、血、向、囟、后、伞、争、色、赤、亥、农、丞、赤、严、豆、丽、辰、来、卤、卤、谷、龟、奂、免、角、卯、辛、弟、其、直、妻、具、隹、卑、鱼、京、单、肃、隶、南、面、韭、首、段、鬲、高、兼、隹、董、黄、兜、庸、鹿、鼎、鼠

这些字，依据《信息处理用 GB 13000.1 字符集汉字部件规范》，都应该拆分为两个基础部件或多个基础部件，例如“么、勺、凡、义、卫、刃、风、爻”等可以拆分为两个部件，“亥、奂、肃、黄、囟、卤、谷”等可以拆分为三个部件，它们是合体字（合成部件）。

我们建议把这些可以拆分为两个基础部件或三个基础部件的部件，从该规范的“部件表”中删除。

五、余论

《信息处理用 GB13000.1 字符集汉字部件规范》《现代常用独体字规范》《现代常用字部件及部件名称规范》都是国家发布的语言文字规范标准，但这三者之间有一些相互矛盾的地方，笔者觉得是不应该的。如果同一个汉字，依据这一部语言文字规范是这么分析，依据另一部语言文字规范又是另一种分析，是很不严肃、很不科学的。如果上述语言文字规范能够修订，应该考虑相互照应，和谐一致。

笔者在《〈通用规范汉字表〉独体字统计与思考》中说，盼望国家早日制定和公布《通用规范汉字独体字规范》和《通用规范汉字部件规范》^[9]。如果有了专门的基于《通用规范汉字表》的部件规范和独体字规范，那我们在对通用规范汉字作结构分析时，就有了新的统一的标准了。

参考文献

- [1] 国家语委：《信息处理用 GB13000.1 字符集汉字部件规范》，语文出版社 1998 年版。
- [2] 胡乔木：《关于当前文字改革工作的讲话》，《胡乔木谈语言文字》，人民出版社 1999 年版。
- [3] 张普：《张普应用语言学论文集》，北京语言大学出版社 2012 年版。
- [4] 肖金卯、武云翠：《关于〈信息处理用 GB 13000.1 字符集汉字部件规范〉中若干问题的讨论》，《中文信息处理技术研讨会论文集》2004 年。
- [5] 教育部、国家语委：《现代常用独体字规范》，《语言文字规范手册》，商务印书馆国际有限公司 2014 年版。
- [6] 教育部、国家语委：《通用规范汉字表》，语文出版社 2013 年版。
- [7] 苏培成：《现代汉字学纲要（第 3 版）》，商务印书馆 2014 年版。
- [8] 教育部、国家语委：《现代常用字部件及部件名称规范》，《语言文字规范手册》，商务印书馆国际有限公司 2014 年版。
- [9] 邵霭吉：《〈通用规范汉字表〉独体字统计与思考》，《盐城师范学院学报》2021 年第 3 期。

Discussion on Issues in Several Language and Writing Standards

SHAO Aiji

(College of Arts, Yancheng Normal University)

Abstract: In recent decades, many language and writing standards have been issued by the government. Some of these standards have been timely revised and updated with improved versions. However, there are some standards that have not been revised since their publication, such as “GB13000.1 Character Set Chinese Character Components Specification for Information Processing,” “Modern Com-

monly Used Single-Component Character Specification,” and “Specification for Components and Component Names of Modern Commonly Used Characters.” With the development of the times and progress in scientific research, it is necessary to timely revise this language and writing standards.
Key words: language and writing Standards; revision; components; single-component characters

作者简介

邵霭吉, 盐城师范学院文学院教授。主要研究方向: 汉语言文字学和《现代汉语教材》编写。

(学术编辑 石绍浪)