

锡伯文满文通用键盘字母数字区布局解决方案

Solution for Sibe and Manchu Universal Keyboard's Alphanumeric Zone Layout

■ 新疆大学¹ 乌鲁木齐市索贝特数码科技有限公司² 察布查尔锡伯自治县满文翻译中心³
付 勇^{1,2} 佟 靖³ 贺智程³

摘 要 分析了 GB/T 28038-2011 中锡伯文、满文通用键盘字母数字区布局中 K、G、H 键上的几个辅音字母以及 U、V 和 I 键上元音字母 u、v、大“i” (AM I) 和小“i” (AJI I) 的冲突问题和根源，并给出了解决这些问题的可行方案。

关键词 GB/T 28038 锡伯文 满文 通用键盘 布局

Abstract: Analyzed a few letters of the conflicts and their causes, including the consonants on the K, G, H keys and the vowels u, v, AM I and AJI I on the U, V and I keys which are described in GB/T 28038-2011 about Sibe and Manchu universal keyboard layout of the alphanumeric zone, and then presents a feasible scheme to solve these problems.

Keywords: GB/T 28038; Sibe language; Manchu language; universal keyboard; layout

1 引言

现行锡伯文、满文的通用键盘字母数字区布局依据是国家推荐性标准 GB/T 28038-2011 《信息技术 通用多八位编码字符集 蒙古文 通用键盘字母数字区布局》。然而，这一标准中锡伯文、满文键盘布局的设计是基于 GB/T 26226-2010 《信息技术 蒙古文变形显现字符集 和控制字符使用规则》中锡伯文、满文名义字符和变形显现字符处理方式，因此二者在几个字符的输入上存在不可调和的冲突问题。

GB/T 28038-2011 规定的满文通用键盘字母数字区布局如图 1 和图 2 所示。

2 几个字母的冲突问题

锡伯文、满文通用键盘字母数字区布局

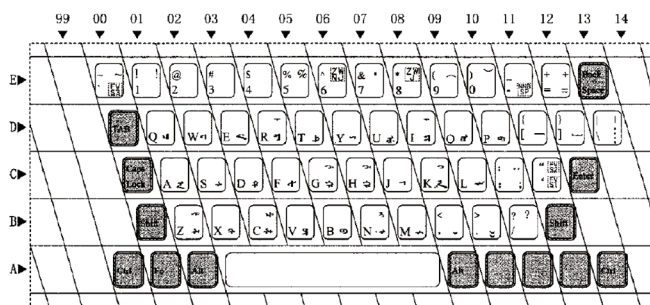


图 1 GB/T 28038-2011 中满文通用键盘字母数字区布局

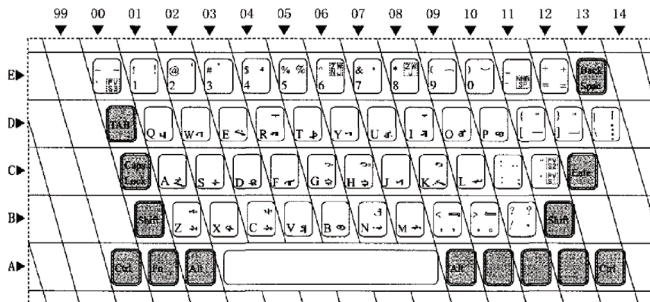


图 2 GB/T 28038-2011 中锡伯文通用键盘字母数字区布局

项目来源：新疆维吾尔自治区科技型中小企业技术创新基金无偿资助项目“锡伯文、满文输入法的研发”，项目编号：2014 531069。

中字母的冲突问题，本质上指通过该键盘布局输入字母时其对应输入的 ASCII 码值转换为锡伯文、满文字符编码时存在二义性、同一个字母的不同形式存在不同编码定义、以及字母冲突很难或根本不能通过软件或字库脚本进行处理的问题。

2.1 K、G、H 键上字母定义的冲突

请关注图 1 中 K、G、H 三个键键盘上指定字符的表示。这三个键上档键位分别标注的是满文名义字符“ᡤ”、“ᡤᡠ”、“ᡤᡠᡳ”，这三个字符在 GB/T 26226 中对应的字符编码为 183A、186C 和 186D；而下档键位上分别标注的是满文名义字符“ᡤᡠ”、“ᡤᡠ”、“ᡤᡠ”（注：“ᡤᡠ”与“ᡤᡠ”是同一个字母，其中“ᡤᡠ”是词尾形式，“ᡤᡠ”是词首形式），对应的字符编码是 1874、1864 和 1865。从 GB/T 26226 中我们还可以看出下档键的这三个字符与元音 e、i、u 拼写时还有对应的变形显现形式“ᡤᡠ”、“ᡤᡠ”和“ᡤᡠ”。

事实上，满文中“ᡤᡠ”、“ᡤᡠ”、“ᡤᡠ”存在与满文第六元音 v 的拼写形式“ᡤᡠ”、“ᡤᡠ”、“ᡤᡠ”，并且，满文（不是锡伯文）中“ᡤᡠ”和“ᡤᡠ”也存在与元音 v 的拼写形式“ᡤᡠ”和“ᡤᡠ”（注：GB/T 26226 忽略了这个问题）。从键盘输入来看，“ᡤᡠ”和“ᡤᡠ”输入的 ASCII 码的序列都是“kv”，显然存在二义性，因此，连续输入 k 和 v 时，是转换为“ᡤᡠ”还是“ᡤᡠ”，就是一个明显的冲突问题；同样，连续输入 h 和 v 时，也存在是转换为“ᡤᡠ”还是“ᡤᡠ”的问题，因为它们 ASCII 码输入序列都是“hv”。从理论上讲，这种冲突除了采用人工控制外，无论是通过软件控制还是通过字库的脚本转换，都是不可能解决的。

2.2 锡伯文元音 u 的冲突

1991 年颁布的《现代锡伯文学语言正字法》中明确规定，锡伯文只有 a、e、i、o、u 五个元音，满文中第六元音 v 的词中形式“ᡤ”和词尾形式“ᡤ”在锡伯文中是属于元音 u 的，且仅与“ᡤ”、“ᡤ”、“ᡤ”相拼写。但是，从图 2 所示的锡伯文键盘布局中可以看出，除了字母键 U 的下档键位上标注了元音 u 的名义字符“ᡤ”（1860），字母键 V 的下档键上还标注

了“ᡤ”（1861），同时，GB/T 26226 中甚至明确说明“ᡤ”和“ᡤ”都是锡伯文字母，这事实上认定锡伯文还有第六个元音，显然违背了锡伯文只有 5 个元音的规定，或者说是同一个锡伯文元音 u 却采用了两个字符编码，不能不说这是 GB/T 26226 和 GB/T 28038 造成的二义性，在实际应用中难免会造成误解和冲突。

2.3 元音大 i 和小 i 的冲突

无论是锡伯文还是满文，元音 i 的独立形式存在两种形式，一个是“ᡤ”，俗称“大 i”，另一种是“ᡤ”，俗称“小 i”。从图 1 和图 2 可以看出，元音 i 是标识在 I 键的下档键位上。然而在实际应用中，这两个独立体字符在输入时，其中一个只能通过人为控制加以转化，在 GB/T 26226 中规定采用变形选择符 FVS1 加以变换。在某些输入法中默认为“ᡤ”，加上控制符再转化成“ᡤ”。事实上，这种方式在实际应用中造成了一定的困难甚至是麻烦。常常有人会使用输入“e”和“i”两键的方式来构造“ᡤ”。另外，“主义”在锡伯文中拼写为“ᡤᡠ”，“同意”在锡伯文中拼写为“ᡤᡠ”，发音同汉字发音相似，其中发 /i/ 音的锡伯文字符部分是独立发音的，GB/T 26226 和 GB/T 28038 试图用音节隔离符“ᡤ”（1807）和“ᡤ”的词尾形式构成。音节间隔符“ᡤ”在图 2 的 4 键（E04）的上档位键上，但在实际应用中往往会有人使用“a”的词中形式“ᡤ”和“i”的“ᡤ”的词尾形式构成“大 i”的词尾形式“ᡤ”，这就使该单词的编码出现错误。表面上看，这些都是使用者的问题，但本质上是设计上的友好性问题，是“大 i”和“小 i”在同一个键位上的冲突问题。

3 问题分析与解决方案

3.1 K、G、H 键上字母的定义和部署

出现 2.1 节中的冲突问题，其根本原因是没有分清字母的特征和属性，将字母形态错误地划分有直接的联系。在 GB/T 26226 以及 GB/T 28038 中将字符“ᡤ”、

“ㄱ”、“ㄴ”分别与“ㅇ”、“ㅏ”和“ㅑ”作为同一个字母的两种不同形变形式，给予相同的字符编码并在键盘布局中分别安置在 K、G、H 键的下档键位上。对应的上档键位上则部署的是“ㅓ”、“ㅕ”、“ㅗ”。

可以从 1993 年出版的满汉大辞典中满文辅音字母表来看 K、G、H 键上字母属性，该字母表如图 3 所示。从图 3 中我们可以看出，满汉大辞典将“ㄱ”、“ㄴ”、“ㄷ”作为独立的三个辅音字母，而将“ㅇ”、“ㅏ”、“ㅑ”分别与“ㅇ”、“ㅏ”、“ㅑ”看做是一个字母的两种不同形式。我们认为这种划分是正确的。从语音学属性来看，“ㄱ”、“ㄴ”、“ㄷ”是在小舌部位发音的，其国际音标分别为 /q/、/q/、/x/，而“ㅓ”、“ㅕ”、“ㅗ”与“ㅇ”、“ㅏ”和“ㅑ”则同是在软腭部分发音的，其国际音标分别为 /k/、/k/、/x/；从字母拼写特征来看，满文的“ㄱ”、“ㄴ”、“ㄷ”是和元音 a、o、v（注：锡伯文是 u 中形同满文 v 的词中和词尾形式）有拼写关系，“ㅓ”、“ㅕ”、“ㅗ”是和元音 a、o 有拼写关系，而“ㅇ”、“ㅏ”和“ㅑ”是和 e、i、u、v（注：“ㅏ”不与 v 拼写）有拼写关系。因此，无论是从历史上字母的划分来讲，还是从语音学的分析来看，抑或是从拼写关系而言，“ㅓ”、“ㅕ”、“ㅗ”与“ㅇ”、“ㅏ”和“ㅑ”没有任何冲突，分别是同一个字母依据拼写关系的两种不同书写形式，反倒是“ㄱ”、“ㄴ”、“ㄷ”与“ㅇ”、“ㅏ”、“ㅑ”在元音 v 上存在冲突。显然，将“ㅓ”、“ㅕ”、“ㅗ”看做是“ㄱ”、“ㄴ”、“ㄷ”的形变，将其部署在同一键位上是错误的。

故而，解决方案很简单，只需将“ㄱ”、“ㄴ”、“ㄷ”部署在 K、G、H 键的上档键位上，而将“ㅓ”、“ㅕ”、“ㅗ”与“ㅇ”、“ㅏ”和“ㅑ”部署在 K、G、H 键的下档键位上。当然，同时必须修改 GB/T 26226 中这些字符的对应关系。

3.2 锡伯文元音字母 u 的定义和部署

锡伯文中只有五个元音，形同满文第六元音 v 的词中和词尾的字符形态是锡伯文元音 u 仅仅与辅

音标	字母	字头	字中	字尾
n	ᠨ	ᠨ	ᠨᠠ	ᠨᠢ
q'	ᠠ	ᠠ	ᠠᠠ	ᠠᠢ
q	ᠡ	ᠡ	ᠡᠠ	
x	ᠢ	ᠢ	ᠢᠠ	
p	ᠣ	ᠣ	ᠣᠠ	ᠣᠢ
p'	ᠤ	ᠤ	ᠤᠠ	
s(u)	ᠥ	ᠥ	ᠥᠠ	ᠥᠢ
ṣ	ᠦ	ᠦ	ᠦᠠ	ᠦᠢ
t'	ᠦᠢ	ᠦᠢ	ᠦᠢᠠ	ᠦᠢᠢ
t	ᠦᠡ	ᠦᠡ	ᠦᠡᠠ	
l	ᠦᠣ	ᠦᠣ	ᠦᠣᠠ	ᠦᠣᠢ
m	ᠦᠠ	ᠦᠠ	ᠦᠠᠠ	ᠦᠠᠢ
ᠬ'(to')	ᠤ	ᠤ	ᠤᠠ	
ᠬ(to)	ᠤ	ᠤ	ᠤᠠ	
j	ᠤᠠ	ᠤᠠ	ᠤᠠᠠ	
k'	ᠤᠢ	ᠤᠢ	ᠤᠢᠠ	ᠤᠢᠢ
k	ᠤᠡ	ᠤᠡ	ᠤᠡᠠ	
x	ᠤᠢ	ᠤᠢ	ᠤᠢᠠ	
r	ᠤᠠ	ᠤᠠ	ᠤᠠᠠ	ᠤᠠᠢ
f	ᠤᠡ	ᠤᠡ	ᠤᠡᠠ	
v	ᠤᠢ	ᠤᠢ	ᠤᠢᠠ	
ᠪ'	ᠤᠢᠢ	ᠤᠢᠢ	ᠤᠢᠢᠠ	
ᠪ	ᠤᠢᠢ	ᠤᠢᠢ	ᠤᠢᠢᠠ	
ᠵ	ᠤᠢ	ᠤᠢ	ᠤᠢᠠ	
s	ᠤᠢ	ᠤᠢ	ᠤᠢᠠ	
ᠰ'	ᠤᠢ	ᠤᠢ	ᠤᠢᠠ	
ᠰ	ᠤᠢ	ᠤᠢ	ᠤᠢᠠ	
ᠲ			ᠤᠢ	ᠤᠢᠢ













图 3 《满汉大辞典》辅音字母表














音“ㄱ”、“ㄴ”、“ㄷ”相拼写时才有的字符形态，因此，在锡伯文中，u 不应该有两种字符编码。由于在 3.1 节中，我们将“ㄱ”、“ㄴ”、“ㄷ”部署在 K、G、H 键的上档键位上，从键盘上输入的 ASCII 码的序列为“Ku”、“Gu”或“Hu”时，分别输出“ᠬᠤ”、“ᠭᠤ”或“ᠬᠤ”、“ᠭᠤ”或“ᠬᠤ”即可。

故而，解决方法很简单，在 3.1 的解决方案的基础上，在锡伯文键盘布局中，V 键不部署任何字符。并且改变 GB/T 26226 中关于 u 与“ㄱ”、“ㄴ”、“ㄷ”拼写时的形变关系。

3.3 元音大 i 和小 i 的定义和部署




















由于“大 i”“ᠢ”和“小 i”“ᠣ”都有独立存在的情况，用字符的位置关系无从做出取舍，用变形控制符进行转换也绝非是好的方式。另外，将锡伯文“主义”的“义”的锡伯文字符“ᠢ”看做是

音节隔离符“”(1807)和“”的词尾形式构成,也绝非是好的方式。例如:“ langgu”和“ lang-gu”是两个形、音、意均不同的单词。其中第一个单词的两个音节中第二个音节由于发音的连带关系,产生音变,而第二个单词中第二个音节是独立发音的,其中两个音节之间用类似字脊的字符形式作为音节隔离符,这在锡伯文和满文中常常使用,与 GB/T 26226 的定义完全不同。因此,用“”作为音节隔离符是不妥的,也容易与锡伯文、满文中元音 a 的词中形式、元音 e 的第二词中形式以及辅音 n 用作收音的词中形式混淆,字符形式“”则更适合用作音节隔离符。由于“”是独立发音的,且与“大 i”“”有很大的相似之处,因此,我们更愿意将“”作为一个整体,看做是“大 i”“”的词尾形式。另外,在 GB/T 26226 这种方式下,“”和“”在独立出现时,则根本不可能进行拉丁转写的双向可逆的变换。


故而,“”和“”应该分别给予编码,在键盘上,“大 i”的独立形式“”及其词尾形式“”应该分离出来,部署在 I 键的上档键位上。这样,用户在输入“大 i”的“”和“”直接用上档键位即可,而输入“小 i”“”及“小 i”的其他字符形态只需使用下档键位。另外,上档键位上原本部署了字符“”(185F),在于上档键位部署“大 i”“”,也无冲突问题。这是因为,“”是独立形式,而“”只有词中和词尾形式,“”虽说是词尾形式,但其仅出现在元音和辅音“ng”之后,而“”

及其词尾形式“”只与辅音 c、s 有拼写关系。

3.4 其他几个字符键位部署的调整

这里说到的几个字符,只需在键位上做一点调整即可。音节隔离符的字形更改为“”,并将其调整到“|”,D13 上档键位上。辅音字符“”与辅音“”发音相同,应部署在与辅音“”对应的上档键位上,同理,辅音字符“”应部署在与辅音“”对应的上档键位上。这样,才更利于记忆,方便使用。另外,在实际应用中,键盘布局图中不应该仅标注所谓的名义字符,例如,T 键的下档键位上应该标注辅音 t 的两种字符形式“”和“”;D 键的下档键位上应该标注辅音 d 的两种字符形式“”和“”;K、G、H 键的下档键位上应该分别标注“”和“”,“”和“”,“”和“”;I 键的上档键位应该标注“”、“”和“”。这样,才能更有利于用户学习和使用。

4 结语

本文谈到的 GB/T 28038-2011 中看似是键盘布局上的简单问题,然而,其中许多问题还是来源于 GB/T 26226-2010 设计的缺陷。若对锡伯文、满文未作深入的研究,就想做出锡伯文、满文信息化建设用的纲领性文件,一定会产生问题。希望相关部门和人士深思。

(收稿日期:2014-12-29)

(上接第 49 页)

北京:中国标准出版社,2011.

- [6] 确精扎布,何正安,唐英敏,等.GB 26122.1-2010 信息技术 通用多八位编码字符集 蒙古文名义字符与变形显现字符 32 点阵字型 第 1 部分:白体[S].北京:中国标准出版社,2011.
- [7] 吕肖庆,确精扎布,何正安,等.GB 26122.2-2010 信息技术 通用多八位编码字符集 蒙古文名义字符与变形显现字符 32 点阵字型 第 2 部分:新闻体[S].

北京:中国标准出版社,2011.

- [8] 确精扎布,陈壮,那顺乌日图,等.GB/T 28038-2011 信息技术 通用多八位编码字符集 蒙古文通用键盘字母数字区布局[S].北京:中国标准出版社,2012.
- [9] 嘎日迪,张主,斯日古楞,等.GB 8046-1987 信息处理 交换用蒙古文字符集键盘字母区的布局[S].北京:中国标准出版社,1988.

(收稿日期:2015-01-05)