

蒙古语言文字数字资源建设与共享工程
信息处理用蒙古文相关标准
MGC/01-02

信息技术 蒙古文拉丁转写规范

Information technology — Proposal of traditional Mongolian
Transliteration

内蒙古自治区民族事务委员会

内蒙古大学

2016 年 11 月

目 次

目次	I
前言	II
1 范围.....	1
2 术语和定义.....	1
3 基础研究	2
4 转写原则.....	2
5 蒙古文拉丁转写国际标准	3

前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

本标准由内蒙古自治区民族事务委员会提出。

本标准起草单位：内蒙古大学、内蒙古民族事务委员会。

本标准主要起草人：白音门德、正月、齐百灵

信息技术 蒙古文拉丁转写规范

1 范围

本标准规定了信息处理用蒙古文拉丁转写规则。

本标准适用于蒙古文的各种语料库、词汇集-词典且必须适应蒙古语语音合成和语音识别的需要。

2 术语和定义

下列术语和定义适用于本文件。

2.1

字符 character

共组织、控制或表示数据用的元素集合中的一个元素。

2.2

编码字符 coded character

字符及其编码表示。

2.3

编码字符集 coded character set

一组无歧义的规则,用于建立一个字符集和该字符集的字符及其编码表示之间的一一对应关系。

2.4

控制功能 control function

影响数据记录、处理、传输或解释的动作。

2.5

控制字符 control character

一种具有控制功能的字符。将它插在别的字符之间,是为了启动、修改或停止某种控制功能的执行。控制字符由机器解释,它不是图形符号。

2.6

名义字符 nominal form

蒙古文字母的主要形式。它适用于蒙古语的书面形式以及附加符号的表示、传输、交换、处理、存储、输入及显示。

2.7

文本 text

由编码字符构成的数据。

2.8

传统蒙古文单词 traditional Mongolian word

仅由传统蒙古文字母、控制符以及词中连接符(U+180A)构成的编码字符序列。

3 基础研究

本标准制定基于以下基础研究：

3.1

总结前人蒙古文拉丁转写的经验。

由于出版业的需要，尤其随着语言信息处理的发展国际上出现了非拉丁字母文字拉丁转写的问题。国际标准化组织（ISO）有专门的非拉丁字母拉丁转写标准。蒙古文属于非拉丁字母，因此也面临拉丁转写的问题。蒙古文拉丁转写有较长的历史，鲍培、李盖提、小泽重男、罗伊郭等国际知名的学者都创建过自己一套蒙古文转写系统。他们的研究以及所使用的蒙古文拉丁转写方案为我们的研究提供了非常好的参考。

3.2

解决一些蒙古文正字法中存在的理论问题。

如蒙古文deviser问题，蒙古文有没有复合元音问题。

3.3

从东西方文字规范的大环境中总结蒙古文正字法的历史经验。

蒙古文的规范历程中想尽一切办法区别字形的努力有相当长的历史。创制蒙古文的当初所借用的14个字符都没有点，在有些字符上加一点就是一个区别字形的努力。例如：

ᠠᠮᠤᠨ (人们)	ᠰᠡᠮᠤ (太陽)
ᠠᠮᠤᠨ (有)	ᠰᠡᠮᠤᠨ (山洞)
ᠠᠮᠤᠨ (平安)	ᠰᠡᠮᠤᠨ (秋)
ᠠᠮᠤᠨ (干)	ᠰᠡᠮᠤᠨ (硬的)
ᠠᠮᠤᠨ (咬)	ᠰᠡᠮᠤᠨ (旁边)

这些例子的后者如果没有点的话会出现意义混淆的可能性。这种教训出现过无数次。种种证据证明《蒙古秘史》是用没有点的蒙古文写的这一事实。§ 1 2 2 的ᠠᠮᠤᠨ这一人名在明朝用汉字转写的《蒙古秘史》中转写为「温真」，但在元朝文献《圣武亲征录》里记录为【嫩真】，拉史德的《史集》也是n。这些例子说明明朝时代用【温】字转写是错误的。还有在§ 2 5 8里把霍拉刺莫的名城ᠨᠢᠰᠠᠪᠤᠷ (Nišabur) 转写为「亦薛不儿」。

增加区别字(ᠠᠮᠤᠨ ᠠᠮᠤᠨ)是蒙古文规范历程中的一件大事。我们这里所说的区别字指的是「ᠠ」和「ᠡ」两个字符。这两个字符在蒙古语中不表达任何语音，只起到区别字形的作用。其中「ᠠ」只出现于词中，「ᠡ」只出现于词末。

4 转写原则

制定本方案时坚持以下原则：

- 1) 蒙古文同一个字符的不同变体只用一个拉丁字母来转写。
- 2) 此方案适用于电子、机械以及手写转写。
- 3) ISO/IEC10646 有关的编码能够自动转写。
- 4) 能够从传统蒙古文转写到拉丁文，从拉丁文转写的蒙古文的双向转写功能。
- 5) 懂得蒙古文的人能够准确读拉丁转写的文件。
- 6) 能够使用 SC2, ISO/IEC 10646 之 BASOC LATIN 和 LATIN-1 SUPPLEMENNT, LATIN EXTENDED A, B, COMBINING DIARITICAL MARKS, IPA EXTENSIVOS, LATIN EXTENDED ADDITIONAL 等的附加字符。
- 7) 充分研究蒙古文正字法的基础上提出需要转写的基本字符集。
- 8) 根据 ISO 的非拉丁文的拉丁转写基本原则和要求 and 我国国家标准的相关规定，根据蒙古文特点提出蒙古文拉丁转写方案，解决在计算机上蒙古文文本和拉丁转写文本的双向转写是这次要解决的难点和创新点。

- 9) 根据广大群众的蒙古语文学习使用需要,制定蒙古文拉丁转写规则,为蒙古语文教学服务的应用规则。
- 10) 适应蒙古语语音合成与语音识别的蒙古文蒙古文拉丁转写国家标准。

5 蒙古文拉丁转写国际标准

蒙古文拉丁转写国际标准见附表。

附表:

No	Character	UCS	Transliteration I	UCS
1	᠎	1800	&	0026
2	᠎	1801	...	2026
3	᠎	1802	,	002C
4	᠎	1803	.	002E
5	᠎	1804	:	003A
6	᠎	1805	#	0023
7	᠎	180A	-	2010
8	᠎	180B	'	0027
9	᠎	180C	"	0022
10	᠎	180D	`	0060
11	᠎	180E	_	005F
12	᠎	1810	' 0	0027, 0030
13	᠎	1811	' 1	0027, 0031
14	᠎	1812	' 2	0027, 0032
15	᠎	1813	' 3	0027, 0033
16	᠎	1814	' 4	0027, 0034
17	᠎	1815	' 5	0027, 0035
18	᠎	1816	' 6	0027, 0036
19	᠎	1817	' 7	0027, 0037
20	᠎	1818	' 8	0027, 0038
21	᠎	1819	' 9	0027, 0039
22	᠎	1820	a~a	0251 ~ 0061
23	᠎	1821	e	0065
24	᠎	1822	i	0069

25	o	1823	o	006F
26	u	1824	u	0075
27	ö	1825	ö	00F6
28	ü	1826	ü	00FC
29	ë	1827	ë	00EB
30	n	1828	n	006E
31	ng	1829	ng	006E, 0261
32	b	182A	b	0062
33	p	182B	p	0070
34	x	182C	x	0078
35	g~g	182D	g~g	0261 ~ 0067
36	m	182E	m	006D
37	l	182F	l	006C
38	s	1830	s	0073
39	š	1831	š	0161
40	t	1832	t	0074
41	d	1833	d	0064
42	č	1834	č	010D
F	ř	1835	ř	01F0
44	y	1836	y	0079
45	r	1837	r	0072
46	w	1838	w	0077
47	f	1839	f	0066
48	k	183A	k	006B
49	ķ	183B	ķ	1E33
50	c	183C	c	0063
51	z	193D	z	007A
52	h	183E	h	0068
53	ž	183F	ž	017E
54	lh	1840	lh	006C, 0068

55	𑌵	1841	ẓ	1E91
56	𑌶	1842	ĉ	0109
57	𑌷	1853	j	006A
58	𑌸	1858	g	0123
59	𑌹	185B	n	0144
60	𑌺	185C	z	017A
61	𑌻	1880	ṃ	1E43
62	𑌼	1881	ḥ	1E25
63	𑌽	1882	â	00E2
64	𑌾	1883	ö	014F
65	𑌿	1884	ô	00F4
66	𑍀	1885	˙	02D1
67	𑍁	1886	ː ˙	02D0, 02D1
68	𑍂	1887	a	0101
69	𑍃	1888	i	012B
70	𑍄	1889	k	1E35
71	𑍅	188A	n	1E49
72	𑍆	188B	c'	0188
73	𑍇	188C	t'	0165
74	𑍈	188D	t	0163
75	𑍉	188E	d	1E0F
76	𑍊	188F	n	0146
77	𑍋	1890	t	1E6F
78	𑍌	1891	d	1E13
79	𑍍	1892	p	0070+0304
80	𑍎	1893	ph	1E55+0068
81	𑍏	1894	s	015F
82	𑍐	1895	h	1E96
83	𑍑	1896	z	1E95

84	o	1897	3	0292+0304
85	'o	18A9	—	0332
		200C	^	005E
		200D	*	002A
		202F	-	002D