

# DB65

## 新疆维吾尔自治区地方标准

DB 65/T 3691—2015

---

### 现行维吾尔文与西里尔维吾尔文 编码字符转换规则

Coded Character Conversion Rules between Uyghur and Cyrillic Uyghur

2015 – 01 – 13 发布

2015 – 03 – 13 实施

新疆维吾尔自治区质量技术监督局 发布



## 前 言

本标准依照GB/T 1.1-2009《标准化工作导则 第1部分:标准的结构和编写》、GB 13000-2010《信息技术 通用多八位编码字符集(UCS)》、GB 21669-2008《信息技术 维吾尔文、哈萨克文、柯尔克孜文编码字符集》的规定编写。

本标准由新疆维吾尔自治区民族语言文字工作委员会提出。

本标准由新疆维吾尔自治区经济和信息化委员会归口。

本标准起草单位:新疆维吾尔自治区标准化研究院、新疆西北星信息技术有限责任公司、中科院新疆理化技术研究所、新疆维吾尔自治区民族语言文字工作委员会。

本标准主要起草人:董军、吾守尔·斯拉木、石常海、刘金龙、亚森·伊明、茹克百提·艾尔肯、周喜、阿里木·哈萨尼、申文英、胡彬华、吐尔根·伊布拉音、王磊、塔依尔·阿布都外力。

## 引 言

现行维吾尔文与西里尔维吾尔文都是用于表达维吾尔语言的拼音文字，但两种文字表达维吾尔语言字母的方式有许多显著区别，主要包括以下几个方面：

- 现行维吾尔文是从右向左书写的文字，西里尔维吾尔文是从左向右书写的文字。
- 现行维吾尔文字母不区分大小写，西里尔维吾尔文字母区分大小写。
- 维吾尔文有 8 个元音字母。现行维吾尔文用字符 **ۋ**（该字符的编码为 0x0626）与另外 8 个字符结合构成 8 个双字符字母来表示这 8 个维吾尔文元音字母，且根据上下文环境的不同，字符 **ۋ** 有时需要出现，有时不需要出现；西里尔维吾尔文用 8 个字符分别表示这 8 个维吾尔文元音字母。
- 缩略语在西里尔维吾尔文中用每个单词首字母的大写形式表示，而在现行维吾尔文中则用每个单词的首字母加空格表示。
- 现行维吾尔文有 32 个字母，西里尔维吾尔文有 40 个字母。

为了保证现行维吾尔文与西里尔维吾尔文编码字符转换过程中信息的完整性，应对这两种文字的书写方式进行必要的规范。因此，本标准提出了与编码字符转换有关的两种文字的书写规则，并制定两种文字编码字符的转换规则。

# 现行维吾尔文与西里尔维吾尔文编码字符转换规则

## 1 范围

本标准规定了现行维吾尔文与西里尔维吾尔文编码字符转换的术语和定义、书写规则及转换规则等。

本标准适用于现行维吾尔文与西里尔维吾尔文编码字符的转换工具和各类信息系统。

## 2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅所注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB 13000 信息技术 通用多八位编码字符集(UCS)

## 3 术语和定义

下列术语和定义适用于本文件。

### 3.1

现行维吾尔文 `uyghur`

指阿拉伯字母为基础的维吾尔文。

### 3.2

西里尔维吾尔文 `cyrillic uyghur`

指西里尔字母为基础的维吾尔文。

### 3.3

字母 `letter`

注音及拼读的符号。

### 3.4

字符 `character`

供组织、控制或表示数据用的元素集合中的一个元素。

### 3.5

编码字符 `coded character`

字符及其编码表示。

### 3.6

单字符字母 `mono letter`

由一个字符表示的字母。

### 3.7

**双字符字母** dual letter

由两个相邻字符按固定的先后顺序构成字符序列表示的字母。

### 3.8

**阿拉伯字符** arabic character

用于表示阿拉伯字母和符号的编码字符，在GB 13000中，有5个阿拉伯字符的编码区域，分别是基本区（编码从0x0600到0x06FF）、补充区（编码从0x0750到0x077F）、扩展区（编码从0x08A0到0x08FF）、变形显现区A（编码从0xFB50到0xFDEE）、变形显现区B（编码从0xFE70到0xFEFF）。

### 3.9

**西里尔字符** cyrillic character

用于表示西里尔字母和符号的编码字符，在GB 13000中，中有4个西里尔字符的编码区域，分别是基本区（编码从0x0400到0x04FF）、补充区（编码从0x0500到0x052F）、扩展区A（编码从0x2DE0到0x2DEF）、扩展区B（编码从0xA640到0xA69F）。

### 3.10

**分隔符** delimiter

字符 '，主要用于音节分隔或区分双字符字母和单字符字母，该字符的编码为0x0027。

### 3.11

**镜像字符** mirrored character

在GB 13000中，具有Bidi Mirrored属性的字符，例如各种括号。

### 3.12

**布局和格式控制字符** layout and format control character

用于控制字符的连接行为、双向顺序行为和选择显示格式行为的字符，例如：编码为0x200C的零宽不连接字符，编码为0x200F的从右向左标记字符等。

### 3.13

**零宽空格** zero width space

编码为0x200B的字符，宽度为零，主要用于指示单词的边界。

### 3.14

**零宽不中断空格** zero width no-break space

编码为0xFEFF的字符，宽度为零，主要用于指示字符的前后即不分词也不断行。

### 3.15

**一对一的编码字符转换** conversion of one coded character to one coded character

一个编码字符转换为一个编码字符。

## 3.16

一对多的编码字符转换 conversion of one coded character to several coded character

一个编码字符转换为多个编码字符。

## 3.17

多对多的编码字符转换 conversion of several coded character to several coded character

多个编码字符转换为多个编码字符。

## 3.18

多对一的编码字符转换 conversion of several coded character to one coded character

多个编码字符转换为一个编码字符。

## 4 与编码字符转换相关的现行维吾尔文书写规则

## 4.1 缩略语的每个字母之间应有一个空格（见示例1）。

示例1: ش ئۇ ئار (شەنجاڭ ئۇيغۇر ئاپتونوم رايونى، 新疆维吾尔自治区)。

## 5 与编码字符转换相关的西里尔维吾尔文书写规则

## 5.1 句首字母应大写（见示例2）。

## 5.2 双引号内的首字母应大写（见示例3）。

## 5.3 专有名词(人名、地名、机构名称)的首字母应大写（见示例4）。

## 5.4 缩略语中的字母应大写（见示例5）。

示例2: Мән мәктәптә оқуймән. (我在学校读书)

示例3: Алим мундақ деди: "Мән бурун бундақ егиз адәмни көрүп бақмиған."。(阿里木说：“我以前没有见过这么高的人。”)

示例4: Алим (阿里木), Турпан (吐鲁番)

示例5: ШУАР (新疆维吾尔自治区)

## 6 现行维吾尔文编码字符转换为西里尔维吾尔文编码字符的规则

## 6.1 总体要求

## 6.1.1 扩展区编码字符的转换

阿拉伯字符变形显现区A和变形显现区B的编码字符表达的现行维吾尔文，应首先转换为阿拉伯字符基本区编码字符表达的现行维吾尔文，然后再转换为对应的西里尔维吾尔文编码字符。

## 6.1.2 现行维吾尔文应首先转换为西里尔维吾尔文的小写形式

如无特别说明，现行维吾尔文编码字符应首先转换为对应的西里尔维吾尔文编码字符的小写形式，然后再根据西里尔维吾尔文的书写规则进行必要的大小写转换。

## 6.1.3 镜像字符的处理

如无特别说明，镜像字符保留不变。信息系统不能依据上下文环境选择镜像字符的正确字形用于显示的，需将一个镜像字符的编码字符转换为与之镜像的另一个编码字符。表1列出了部分镜像字符的镜像关系，完整的镜像字符列表可以参照GB 13000。

表1 部分镜像字符

文字	字符									
现行维吾尔文	( 0028	) 0029	[ 005B	] 005D	{ 007B	} 007D	« 00AB	» 00BB	‹ 2039	› 203A
西里尔维吾尔文	) 0029	( 0028	] 005D	[ 005B	} 007D	{ 007B	» 00BB	« 00AB	› 203A	‹ 2039

6.2 一对一的编码字符转换

6.2.1 现行维吾尔文 21 个辅音字母，其编码字符应直接转换为对应的西里尔维吾尔文文字的编码字符，具体见表 2（示例见 6 至 9）。

表2 一对一的编码字符转换

文字	字符										
现行维吾尔文	ب	پ	ت	ج	خ	د	ر	ز	ژ	غ	ف
西里尔维吾尔文	Б	П	Т	Ж	Х	Д	Р	З	Ж	Ғ	Ф
现行维吾尔文	ق	ك	گ	ڭ	ل	م	ن	ھ	ۋ	ي	
西里尔维吾尔文	Қ	К	Г	Ң	Л	М	Н	Һ	В	Й	

- 示例6: قول (手) 应转换为 кол  
示例7: باش (头) 应转换为 баш  
示例8: پۈت (足) 应转换为 пۈт  
示例9: كۆز (眼) 应转换为 көз

6.2.2 现行维吾尔文文字 س 与两个西里尔维吾尔文文字 ц 和 с 对应，当出现在词首时，其编码字符需根据专用词汇表确定转换为西里尔维吾尔文文字 ц 或 с 的编码字符。若专用的词汇表中没有包含需要转换的词条，则转换时现行维吾尔文文字 с 默认转换为西里尔维吾尔文文字 с 的编码字符；当出现在非词首时，且前面没有现行维吾尔文文字 ت ، 则直接转换为西里尔维吾尔文文字 с 的编码字符，具体见表 3（示例见 10 至 12）。

表3 现行维吾尔文文字 с 的转换

文字	字符	
现行维吾尔文	с	с
西里尔维吾尔文	ц	с

- 示例10: سۆيگۈ (爱情) 应转换为 сөйгү  
示例11: سىرك (马戏) 应转换为 цирк  
示例12: سىياسەت (政治) 应转换为 сиясәт

6.2.3 如表 4 所示，现行维吾尔文文字 ش 与两个西里尔维吾尔文文字 ш 和 щ 对应。因此，现行维吾尔文文字 ش 的编码字符需根据专用词汇表确定转换为西里尔维吾尔文文字 ш 或 щ 的编码字符。如果专用的词汇表中没有包含需要转换的词条，则转换时现行维吾尔文文字 ش 默认转换为西里尔维吾尔文文字 ш 的编码字符，具体见表 4（示例见 13 至 14）。



表4 现行维吾尔文字母 ش 的转换

文字	字符	
现行维吾尔文	ش	ш
西里尔维吾尔文	ш	ш

示例13: ئىشپىون (特务) 应转换为 ишпион

示例14: جىياڭشى (江西) 应转换为 Жианцзи

6.2.4 现行维吾尔文字母 چ 与两个西里尔维吾尔文字母 ч 和 ш 对应。因此, 现行维吾尔文字母 چ 的编码字符需根据专用词汇表确定转换为西里尔维吾尔文字母 ч 或 ш 的编码字符。如果专用的词汇表中没有包含需要转换的词条, 则转换时现行维吾尔字母 چ 默认转换为西里尔维吾尔文字母 ч 的编码字符, 具体见表 5 (示例见 15 至 16)。

表5 现行维吾尔文字母 چ 的转换

文字	字符	
现行维吾尔文	چ	ч
西里尔维吾尔文	ч	ш

示例15: چاق (轮胎) 应转换为 чак

示例16: چوتكا (刷子) 应转换为 щетка

6.2.5 现行维吾尔文元音字母是双字符字母, 当出现在非词首位置时, 根据上下文环境字符 ۋ 有时需要出现, 有时不需要出现。当字符 ۋ 不出现时, 则这 5 个元音字母直接转换为对应的西里尔维吾尔文元音字母的编码字符, 具体见表 6 (示例见 17 至 20)。

表6 现行维吾尔文元音字母在非词首位置的转换

文字	字符					
现行维吾尔文	ا	ۋ	و	ۇ	ۆ	ۈ
西里尔维吾尔文	a	ə	o	y	ɵ	ʏ

示例17: كىتاب (书) 应转换为 китаб

示例18: ۋەتەن (祖国) 应转换为 вәтән

示例19: تومۇر (经脉) 应转换为 томур

示例20: كۆمۈر (煤) 应转换为 көмүр

6.2.6 现行维吾尔文元音字母 ئې 是双字符字母, 且与两个西里尔维吾尔文字母 е 和 э 对应。当现行维吾尔文字母 ئې 出现在非词首位置时, 根据上下文环境字符 ۋ 有时需要出现, 有时不需要出现。当字符 ۋ 不出现时, 现行维吾尔文字母 ئې 需根据专用词汇表确定转换为西里尔维吾尔文字母 е 或 э 的编码字符。如果专用的词汇表中没有包含需要转换的词条, 则转换时现行维吾尔字母 ئې 默认转换为西里尔维吾尔文字母 е 的编码字符, 具体见表 7 (见示例 21)。

表7 现行维吾尔文元音字母 ئې 在非词首位置的转换

文字	字符	
现行维吾尔文	ئې	ئې
西里尔维吾尔文	е	э

示例21: لېكتور (讲师) 应转换为 лектор

6.2.7 现行维吾尔文元音字母 ئى 是双字符字母, 且与两个西里尔维吾尔文字母 и 和 ы 对应。当现行维吾尔文字母 ئى 出现在非词首位置时, 根据上下文环境字符 ۋ 有时需要出现, 有时不需要出现。当字符 ۋ 不出现时, 现行维吾尔文字母 ئى 需根据专用词汇表确定转换为西里尔维吾尔文字母 и 或 ы

的编码字符。如果专用的词汇表中没有包含需要转换的词条，则转换时现行维吾尔字母 ئى 默认转换为西里尔维吾尔文字母 и 的编码字符，具体见表 8（见示例 22）。

表8 现行维吾尔文元音字母 ئى在非词首位置的转换

文字	字符	
现行维吾尔文	ئى	ئى
西里尔维吾尔文	и	ы

示例22: كىتاب (书) 应转换为 китаб

### 6.3 多对一的编码字符转换

6.3.1 现行维吾尔文元音字母是双字符字母，当出现在词首位置时，应忽略词首的字符 ۋ，双字符字母中的另一个字符直接转换为对应的西里尔维吾尔文元音字母的编码字符，具体见表 9(示例见 23 至 28)。

表9 现行维吾尔文元音字母在词首位置的转换

文字	字符					
现行维吾尔文	ئا	ئە	ئو	ئۆ	ئۆ	ئۆ
西里尔维吾尔文	а	ә	о	у	Ө	У

示例23: ئانار (石榴) 应转换为 анар

示例24: ئەنجۈر (无花果) 应转换为 әнжүр

示例25: ئوردا (皇宫) 应转换为 орда

示例26: ئۇرۇش (战争) 应转换为 уруш

示例27: ئۆردەك (鸭子) 应转换为 өрдәк

示例28: ئۈزۈم (葡萄) 应转换为 үзүм

6.3.2 现行维吾尔文元音字母 ئې 是双字符字母，且与两个西里尔维吾尔文字母 е 和 ә 对应。当现行维吾尔文字母 ئې 出现在词首位置时，应忽略词首的字符 ۋ，双字符字母中的另一个字符需根据专用词汇表确定转换为西里尔维吾尔文字母 е 或 ә 的编码字符。如果专用的词汇表中没有包含需要转换的词条，则转换时现行维吾尔字母 ئې 默认转换为西里尔维吾尔文字母 е 的编码字符，具体见表 10(示例见 29 至 30)。

表10 现行维吾尔文元音字母 ئې在词首位置的转换

文字	字符	
现行维吾尔文	ئې	ئې
西里尔维吾尔文	е	ә

示例29: ئېلان (广告) 应转换为 елан

示例30: ئېكۋاتور (赤道) 应转换为 экватор

6.3.3 现行维吾尔文元音字母 ئى 是双字符字母，且与两个西里尔维吾尔文字母 и 和 ы 对应。当现行维吾尔文字母 ئى 出现在词首位置时，应忽略词首的字符 ۋ，双字符字母中的另一个字符需根据专用词汇表确定转换为西里尔维吾尔文字母 и 或 ы 的编码字符。如果专用的词汇表中没有包含需要转换的词条，则转换时现行维吾尔字母 ئى 默认转换为西里尔维吾尔文字母 и 的编码字符，具体见表 11（见示例 31）。

表11 现行维吾尔文元音字母 ئى在词首位置的转换

文字	字符	
现行维吾尔文	ئى	ئى
西里尔维吾尔文	и	ы

示例31: ئىت (狗) 应转换为 ит

6.3.4 单词中有表 12 所示的字母序列,其编码字符分别转换为对应的西里尔维吾尔文 字母的编码字符 (示例见 32 至 34)。

表12 现行维吾尔文 字母序列的转换

文字	字符		
现行维吾尔文	يا	يۇ	تس
西里尔维吾尔文	я	ю	ц

示例32: سىياسەت (政治) 应转换为 сиясәт

示例33: ئىيۇن (6 月) 应转换为 Июнь

示例34: ئاۋىياتسىيە (航空) 应转换为 авиация

6.4 多对多的编码字符转换

6.4.1 现行维吾尔文 5 个元音字母是双字符字母,当出现在非词首位置时,根据上下文环境字符 ۋ 有时需要出现,有时不需要出现。当字符 ۋ 出现时,则字符 ۋ 转换为分隔符 ' 的字符编码,双字符字母的另一个字符转换为对应的西里尔维吾尔文元音字母的编码字符 (示例见 35 至 40)。

表13 现行维吾尔文元音字母在非词首位置的转换

文字	字符					
现行维吾尔文	ئا	ئە	ئو	ئۇ	ئۆ	ئۈ
西里尔维吾尔文	'a	'ə	'o	'u	'ө	'ү

示例35: ئىنىكئانا (奶妈) 应转换为 иник'ана

示例36: ئەسئەت (艾斯艾提) 应转换为 Ә'сәт

示例37: ئارخېئولوگىيە (考古) 应转换为 архе'ологийә

示例38: مەسئۇل (负责) 应转换为 мәс'ул

示例39: قارىئۆرۈك (乌梅) 应转换为 қари'өрүк

示例40: ئائۇمىد (失望) 应转换为 на'үмид

6.4.2 现行维吾尔文元音字母 ئى 是双字符字母,且与两个西里尔维吾尔文 字母 е 和 э 对应。当现行维吾尔文 字母 ئى 出现在非词首位置时,根据上下文环境字符 ۋ 有时需要出现,有时不需要出现。当字符 ۋ 出现时,字符 ۋ 的编码字符转换为分隔符 ' 的编码字符,双字符字母中的另一个字符需根据专用词汇表确定转换为西里尔维吾尔文 字母 е 或 э 的编码字符。如果专用的词汇表中没有包含需要转换的词条,则转换时现行维吾尔文 字母 ئى 默认转换为西里尔维吾尔文 字母 е 的编码字符 (见示例 41)。

表14 现行维吾尔文元音字母 ئى在非词首位置的转换

文字	字符	
现行维吾尔文	ئى	ئى
西里尔维吾尔文	'e	'ə

示例41: ئىتئېيىق (狗熊) 应转换为 ит'ейик

6.4.3 现行维吾尔文元音字母 ئى 是双字符字母,且与两个西里尔维吾尔文 字母 и 和 ы 对应。当现行维吾尔文 字母 ئى 出现在非词首位置时,根据上下文环境字符 ۋ 有时需要出现,有时不需要出现。当字符 ۋ 出现时,字符 ۋ 的编码字符转换为分隔符 ' 的编码字符,双字符字母中的另一个字符需根据专用词汇表确转换为西里尔维吾尔文 字母 и 或 ы 的编码字符。如果专用的词汇表中没有包含需要转换的词条,则转换时现行维吾尔文 字母 ئى 默认转换为西里尔维吾尔文 字母 и 的编码字符 (见示例 42)。

表15 现行维吾尔文元音字母 ئى在非词首位置的转换

文字	字符	
现行维吾尔文	ئى	ئى
西里尔维吾尔文	'и	'ы

示例42: جەمئىي (共计) 应转换为 жәм'ий

6.4.4 表16所示的现行维吾尔文元音字母序列, 当出现在非词首时, 其编码字符直接转换为西里尔维吾尔文元音字母 ц 的编码字符 (见示例43)。

表16 现行维吾尔文元音字母序列 ئى在非词首位置的转换

文字	字符
现行维吾尔文	ئى
西里尔维吾尔文	ц

示例43: سارتىن (撒尔金) 应转换为 Сартин

## 6.5 Ё 和 ё 的处理

现行维吾尔文单词在转换为西里尔维吾尔文单词的过程中, 需要根据专用词汇表判断单词中是否需要增加西里尔维吾尔文元音字母 Ё 和 ё 的编码字符, 以及判断在单词的什么位置增加西里尔维吾尔文元音字母 Ё 和 ё 的编码字符 (示例见44至45)。

示例44: مەنئىي (禁止) 应转换为 мән'ий

示例45: رېلىس (铁轨) 应转换为 рельс

## 6.6 缩略语的编码字符转换

相邻的两个或两个以上现行维吾尔文元音字母左右及之间都有且仅有1个空格, 则忽略字母之间的空格, 并将现行维吾尔文元音字母转换为对应的西里尔维吾尔文元音字母的大写形式 (示例见46至48)。

示例46: ج ك پ (中国共产党) 应转换为 ЖКП

示例47: ج خ (公安) 应转换为 ЖХ

示例48: ش ئۇ ئار (新疆维吾尔自治区) 应转换为 ШУАР

## 6.7 标点符号的编码字符转换

现行维吾尔文标点符号的编码字符应转换为对应的西里尔维吾尔文标点符号的编码字符, 具体见表17 (示例见49至51)。

表17 现行维吾尔文标点符号的转换

文字	字符		
现行维吾尔文	، 060C	؟ 061F	؛ 061B
西里尔维吾尔文	، 002C	؟ 003F	؛ 003B

示例49: بۇ نېمە؟ (这是什么?) 应转换为 бу не́мэ?

示例50: مەن، سەن (我，你) 应转换为 мән,сән

示例51: ئادەم؛ ھايۋان (人；动物) 应转换为 адәм; Һайван

## 6.8 非现行维吾尔文编码字符的转换

### 6.8.1 现行维吾尔文中西里尔字符的处理

现行维吾尔文中的西里尔字符应保留不变，同时在西里尔字符的前后分别增加零宽空格和零宽不中断空格。

#### 6.8.2 零宽空格和零宽不中断空格的处理

忽略现行维吾尔文中的零宽空格和零宽不中断空格，同时零宽空格和零宽不中断空格之间的字符保留不变。

#### 6.8.3 布局和格式控制字符的处理

现行维吾尔文中的布局和格式控制字符保留不变。

#### 6.8.4 连接符的处理

当连接符出现在现行维吾尔文单词中，转换为西里尔维吾尔文时，忽略连接符。

#### 6.8.5 单个不带字符 ۋ 的现行维吾尔文元音字母的处理

现行维吾尔文的8个元音字母由包括字符 ۋ 在内的两个字符构成的双字符字母，根据上下文环境，字符 ۋ 有时需要出现，有时不需要出现。当不带字符 ۋ 的现行维吾尔文元音字母单独出现的时候，在转换为西里尔维吾尔文时其编码字符保留不变。

#### 6.8.6 其它非现行维吾尔文编码字符的处理

现行维吾尔文中的其它非现行维吾尔文编码字符保留不变。

#### 6.9 现行维吾尔文与西里尔维吾尔文编码字符对照见附录 A。

### 7 西里尔维吾尔文编码字符转换为现行维吾尔文编码字符的规则

#### 7.1 总体要求

##### 7.1.1 扩展区编码字符的转换

如无特别说明，西里尔维吾尔文编码字符转换为对应现行维吾尔文字母名义形式的编码字符就完成了转换工作。信息系统不能依据现行维吾尔文字母名义形式的编码字符，选择正确的变形显现的，需将现行维吾尔文字母名义形式的编码字符依据现行维吾尔文的书写习惯进一步转换为正确的变形显现形式编码字符。

##### 7.1.2 西里尔维吾尔文字母大小写形式的处理

如无特别说明，西里尔维吾尔文编码字符转换为现行维吾尔文编码字符的过程中，对同一个西里尔字母的大写形式和小写形式采用同样的方式处理。

##### 7.1.3 镜像字符的处理

如无特别说明，镜像字符保留不变。信息系统不能依据上下文环境选择镜像字符的正确字形用于显示的，需将一个镜像字符的编码字符转换为与之镜像的另一个编码字符。表18列出了部分镜像字符的镜像关系，完整的镜像字符列表可以参照GB 13000。

表18 部分镜像字符的转换

文字	字符									
西里尔维吾尔文	( 0028	) 0029	[ 005B	] 005D	{ 007B	} 007D	« 00AB	» 00BB	‹ 2039	› 203A
现行维吾尔文	) 0029	( 0028	] 005D	[ 005B	} 007D	{ 007B	» 00BB	« 00AB	› 203A	‹ 2039

## 7.2 一对一的编码字符转换

7.2.1 西里尔维吾尔文 24 个辅音字母，其编码字符应直接转换为对应的现行维吾尔文文字的编码字符，具体见表 19（示例见 52 至 55）。

表19 西里尔维吾尔文辅音字母的转换

文字	字符											
西里尔维吾尔文	б	п	т	ж	ч	х	д	р	з	ж	с	ш
现行维吾尔文	ب	پ	ت	ج	چ	خ	د	ر	ز	ژ	س	ش
西里尔维吾尔文	ғ	ф	к	к	г	н	л	м	н	һ	в	й
现行维吾尔文	غ	ف	ق	ك	گ	ڭ	ل	م	ن	ھ	ۋ	ي

示例52: кол (手) 应转换为 قول

示例53: баш (头) 应转换为 باش

示例54: пут (足) 应转换为 پۈت

示例55: көз (眼) 应转换为 كۆز

7.2.2 西里尔维吾尔文 ш 的编码字符与现行维吾尔文 ш 和 ش 对应。需根据专用词汇表确定西里尔维吾尔文 ш 转换为现行维吾尔文 ш 和 ش 的编码字符。若专用的词汇表中没有包含需要转换的词条，转换时西里尔维吾尔文 ш 默认转换为西里尔维吾尔文 ш 的编码字符，具体见表 20（示例见 56 至 57）。

表20 西里尔维吾尔文 ш 的转换

文字	字符	
西里尔维吾尔文	ш	ш
现行维吾尔文	ш	ش

示例56: щетка (刷子) 应转换为 چوتكا

示例57: Жяңци (江西) 应转换为 جياڭشى

7.2.3 表 21 中的 6 个西里尔维吾尔文元音字母出现在非词首时，转换为对应现行维吾尔文元音字母不带字符 ۋ 的形式的编码字符（示例见 58 至 61）。

表21 西里尔维吾尔文元音字母在非词首位置的转换

文字	字符					
西里尔维吾尔文	а	ә	о	у	ө	У
现行维吾尔文	ا	ە	و	ۇ	ۆ	ۋ

示例58: китаб (书) 应转换为 كىتاب

示例59: вәтән (祖国) 应转换为 ۋەتەن

示例60: томур (经脉) 应转换为 تومۇر

示例61: көмүр (煤) 应转换为 كۆمۈر

7.2.4 表 22 中 4 个西里尔维吾尔文元音字母出现在非词首时，转换为对应现行维吾尔文元音字母不带字符 ۋ 的形式的编码字符（示例见 62 至 63）。

表22 西里尔维吾尔文元音字母在非词首位置的转换

文字	字符			
西里尔维吾尔文	е	э	и	ы
现行维吾尔文	ي	ي	ى	ى

示例62: лектор (讲师) 应转换为 لېكتور

示例63: книга (书) 应转换为 كىتاب

### 7.3 一对多的编码字符转换

7.3.1 表 23 中的 6 个西里尔维吾尔文元音字母出现在词首位置时, 转换为对应现行维吾尔文双字符字母的编码字符 (示例见 64 至 69)。

表23 西里尔维吾尔文元音字母在词首位置的转换

文字	字符					
西里尔维吾尔文	а	э	о	у	ө	ү
现行维吾尔文	ئا	ئە	ئو	ئۇ	ئۆ	ئۈ

示例64: анар (石榴) 应转换为 ئانار

示例65: энжүр (无花果) 应转换为 ئەنجۈر

示例66: орда (皇宫) 应转换为 ئوردا

示例67: уруш (战争) 应转换为 ئۇرۇش

示例68: өрдәк (鸭子) 应转换为 ئۆردەك

示例69: үзүм (葡萄) 应转换为 ئۈزۈم

7.3.2 表 24 中的 4 个西里尔维吾尔文元音字母出现在词首位置时, 转换为对应现行维吾尔文双字符字母的编码字符 (示例见 70 至 72)。

表24 西里尔维吾尔文元音字母在词首位置的转换

文字	字符			
西里尔维吾尔文	е	э	и	ы
现行维吾尔文	ئي	ئي	ئي	ئي

示例70: елан (广告) 应转换为 ئېلان

示例71: экватор (赤道) 应转换为 ئېكۋاتور

示例72: ит (狗) 应转换为 ئىت

7.3.3 表 25 中的 2 个西里尔维吾尔文元音字母 Я 和 Ю, 转换为现行维吾尔文时, 编码字符直接转换为对应的现行维吾尔文元音字母序列的编码字符 (示例见 73 至 74)。

表25 西里尔维吾尔文元音字母 Я 和 Ю 的转换

文字	字符	
西里尔维吾尔文	я	ю
现行维吾尔文	يا	يو

示例73: сиясәт (政治) 应转换为 سياسهت

示例74: Июнь (6 月) 应转换为 ئىيۈن

7.3.4 西里尔维吾尔文元音字母 Й, 当在词首时, 与现行维吾尔文元音字母 س 和字母序列 تس 对应。因此, 需根据专用词汇表确定西里尔维吾尔文元音字母 Й 的编码字符转换为现行维吾尔文元音字母 س 或字母序列 تس 的编码字符。如果专用的词汇表中没有包含需要转换的词条, 则转换时西里尔维吾尔文元音字母 Й 默认转换为现行维吾尔文元音字母 س 的编码字符; 当在非词首时, 直接转换为现行维吾尔文元音字母序列 تس 的编码字符, 具体见表 26 (示例见 75 至 76)。

表26 西里尔维吾尔文字母 Ц 在词首位置的转换

文字	字符	
西里尔维吾尔文	Ц	Ц
现行维吾尔文	تى	س

示例75: Сарцин (撒尔金) 应转换为 سارتسين

示例76: цирик (马戏) 应转换为 سيريك

#### 7.4 Ъ 和 ь 的处理

西里尔维吾尔文字母 Ъ 和 Ь 的编码字符在转换为现行维吾尔文时忽略 (示例见77至78)。

示例77: мәнЪий (禁止) 应转换为 مەنئىي

示例78: рельс (铁轨) 应转换为 رېلىس

#### 7.5 分隔符编码字符的转换

当分隔符 ' 在西里尔维吾尔文元音字母之前, 应转换为字符 ۋ 的编码字符; 分隔符 ' 在其它位置, 则保留不变 (示例见79至80)。

示例79: инник'ана (奶妈) 应转换为 ئىنىكئانا

示例80: Әс'әт (艾斯艾提) 应转换为 ئەسەت

#### 7.6 缩略语编码字符的转换

相邻的多个大写西里尔维吾尔字母编码字符转换为现行维吾尔文编码字符时, 直接转换为对应的现行维吾尔文编码字符, 然后在现行维吾尔字母的编码字符之间加1个空格编码字符 (示例见81至83)。

示例81: ЖКП (中国共产党) 应转换为 ج ك پ

示例82: ЖХ (公安) 应转换为 ج خ

示例83: ШУАР (新疆维吾尔自治区) 应转换为 ش ئۇ ئار

#### 7.7 标点符号的转换

7.7.1 西里尔维吾尔文标点符号的编码字符应转换为对应的现行维吾尔文标点符号的编码字符, 具体见表 27 (示例见 84 至 86)。

表27 西里尔维吾尔文标点符号的转换

文字	字符		
西里尔维吾尔文	002C	003F	003B
现行维吾尔文	060C	061F	061B

示例84: бу немә? (这是什么?) 应转换为 بۇ نېمە؟

示例85: мән, сән (我, 你) 应转换为 مەن, سەن

示例86: адәм; һайван (人; 动物) 应转换为 ئادەم; ھايۋان

#### 7.8 非西里尔维吾尔文编码字符的转换

##### 7.8.1 西里尔维吾尔文中阿拉伯字符的处理

西里尔维吾尔文中的阿拉伯字符应保留不变, 并在阿拉伯字符的前后分别增加零宽空格和零宽不中断空格。



### 7.8.2 零宽空格和零宽不中断空格的处理

忽略西里尔维吾尔文中的零宽空格和零宽不中断空格,同时零宽空格和零宽不中断空格之间的字符保留不变。

### 7.8.3 布局和格式控制字符的处理

西里尔维吾尔文中的布局和格式控制字符保留不变。

### 7.8.4 其它非西里尔维吾尔文编码字符的处理

西里尔维吾尔文中的其它非西里尔维吾尔文编码字符保留不变。

### 7.8.5 有特殊要求需要保留不变的字符或字符串

有特殊要求需要保留的字符或字符串保留不变,例如罗马数字(I、II、III、IV、X)、字母名词(WTO、VCD、DVD)等。

### 7.9 西里尔维吾尔文与现行维吾尔文编码字符对照见附录 B。

附 录 A  
(规范性附录)  
现行维吾尔文与西里尔维吾尔文编码字符对照

A.1 现行维吾尔文与西里尔维吾尔文编码字符对照

现行维吾尔文与西里尔维吾尔文编码字符对照见表A.1。

表 A.1 现行维吾尔文与西里尔维吾尔文编码字符对照表

序号	现行维吾尔文					西里尔维吾尔文	
	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式	大写形式	小写形式
1	ئا 0626+0627	ئا ا FBEA FE8D			ئال FBEB FE8E	A 0410	a 0430
						'A 0027+0410	'a 0027+0430
2	ئە 0626+06D5	ئە ە FBEC FEE9			ئە ە FBED FEEA	Ә 018F	ә 04D9
						'Ә 0027+018F	'ә 0027+04D9
3	ب 0628	ب FE8F	ب FE91	ب FE92	ب FE90	Б 0411	б 0431
4	پ 067E	پ FB56	پ FB58	پ FB59	پ FB57	П 041F	п 043F
5	ت 062A	ت FE95	ت FE97	ت FE98	ت FE96	Т 0422	т 0442
6	چ 062C	چ FE9D	چ FE9F	چ FEA0	چ FE9E	Ж 0496	ж 0497
7	چ 0686	چ FB7A	چ FB7C	چ FB7B	چ FB7D	Ч 0427	ч 0447
						Щ 0429	щ 0449
8	خ 062E	خ FEA5	خ FEA7	خ FEA8	خ FEA6	Х 0425	х 0445
9	د 062F	د FEA9			د FEAA	Д 0414	д 0434
10	ر 0631	ر FEAD			ر FEAE	Р 0420	р 0440
11	ز 0632	ز FEAF			ز FEB0	З 0417	з 0437
12	ژ 0698	ژ FB8A			ژ FB8B	Ж 0416	ж 0436

表 A. 1 (续)

序号	现行维吾尔文					西里尔维吾尔文	
	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式	大写形式	小写形式
13	س 0633	س FEb1	س FEb3	س FEB4	س FEB2	С 0421	с 0441
						И 0426	и 0446
14	ش 0634	ش FEB5	ش FEB7	ش FEB8	ش FEB6	Ш 0428	ш 0448
						Ил 0429	ил 0449
15	غ 063A	غ FECD	غ FECF	غ FED0	غ FECE	Ғ 0492	ғ 0493
16	ف 0641	ف FED1	ف FED3	ف FED4	ف FED2	Ф 0424	ф 0444
17	ق 0642	ق FED5	ق FED7	ق FED8	ق FED6	Қ 049A	қ 049B
18	ك 0643	ك FED9	ك FEDB	ك FEDC	ك FEDA	К 041A	к 043A
19	گ 06AF	گ FB92	گ FB94	گ FB95	گ FB93	Г 0413	г 0433
20	ك 06AD	ك FBD3	ك FBD5	ك FBD6	ك FBD4	Ҳ 04A2	ҳ 04A3
21	ل 0644	ل FEDD	ل FEDF	ل FEE0	ل FEDE	Л 041B	л 043B
22	م 0645	م FEE1	م FEE3	م FEE4	م FEE2	М 041C	м 043C
23	ن 0646	ن FEE5	ن FEE7	ن FEE8	ن FEE6	Н 041D	н 043D
24	ه 06BE	ه FBAA	ه FBAC	ه FBAD	ه FBAB	Һ 04BA	һ 04BB
25	ئو 0626+0648	و ئو FBEE FEED			و ئو FBEF FEEE	О 041E	о 043E
						'O 0027+041E	'o 0027+043E
26	ئو 0626+06C7	ۇ ئو FBF0 FBD7			ۇ ئو FBF1 FBD8	У 0423	у 0443
						'y 0027+0423	'y 0027+0443
27	ئۆ 0626+06C6	ۆ ئۆ FBF2 FBD9			ۆ ئۆ FBF3 FBDA	Ө 04E8	ө 04E9
						'Ө 0027+04E8	'ө 0027+04E9

表 A.1 (续)

序号	现行维吾尔文					西里尔维吾尔文	
	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式	大写形式	小写形式
28	ئۇ 0626+06C8	ۋ ئۇ FBF4 FBDB			ۋ ئۇ FBF5 FBDC	Y 04AE	Y 04AF
						Y 0027+04AE	'Y 0027+04AF
29	ۋ 06CB	ۋ FBDE			ۋ FBDF	B 0412	в 0432
30	ئې 0626+06D0	ې ئې FBF6 FBE4	ې ئې FBF8 FBE6	ې FBE7	ي ئې FBF7 FBE5	E 0415	е 0435
						'E 0027+0415	'e 0027+0435
						Э 042D	э 044D
						'Э 0027+042D	'э 0027+044D
31	ئى 0626+0649	ى ئى FBF9 FEEF	د ئى FBFB FBE8	د FBE9	ى ئى FBFA FEF0	И 0418	и 0438
						'И 0027+0418	'и 0027+0438
						Ы 042B	ы 044B
						'Ы 0027+042B	'ы 0027+044B
32	ي 064A	ي FEF1	ي FEF3	ي FEF4	ي FEF2	Й 0419	й 0439
33	ئى 0626	ئى FE89	ئى FE8B	ئى FE8C	ئى FE8A	، 0027	، 0027
34						Ъ 042A	ъ 044A
35						Ь 042C	ь 044C
36	تى 062A+0633	تى FE97+FE82	تى FE97+FE84	تى FE98+FE84	تى FE98+FE82	Ц 0426	ц 0446
36	يا 064A+0627	يا FEF3+FE8E	يا FEF3+FE8E	يا FEF4+FE8E	يا FEF4+FE8E	Я 042F	я 044F

表 A. 1（续）

序号	现行维吾尔文					西里尔维吾尔文	
	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式	大写形式	小写形式
37	يو 064A+06C7	يو FEF3+FBD8	يو FEF3+FBD8	يو FEF4+FBD8	يو FEF4+ FBD8	Ю 042E	ю 044E
38	‘ 060C	‘ 060C	‘ 060C	‘ 060C	‘ 060C	‘ 002C	‘ 002C
39	؟ 061F	؟ 061F	؟ 061F	؟ 061F	؟ 061F	？ 003F	？ 003F
40	؛ 061B	؛ 061B	؛ 061B	؛ 061B	؛ 061B	؛ 003B	؛ 003B

附 录 B  
(规范性附录)

西里尔维吾尔文与现行维吾尔文编码字符对照表

B.1 西里尔维吾尔文与现行维吾尔文编码字符对照

西里尔维吾尔文与现行维吾尔文编码字符对照见表B.1。

表 B.1 西里尔维吾尔文与现行维吾尔文编码字符对照表

序号	西里尔维吾尔文		现行维吾尔文				
	大写形式	小写形式	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式
1	А 0410	а 0430	ئا 0626+0627	ا ئا FBEA FE8D			ئا ا FBEB FE8E
2	Б 0411	б 0431	ب 0628	ب FE8F	ب FE91	ب FE92	ب FE90
3	В 0412	в 0432	ۋ 06CB	ۋ FBDE			ۋ FBDF
4	Г 0413	г 0433	غ 06AF	غ FB92	غ FB94	غ FB95	غ FB93
5	Д 0414	д 0434	د 062F	د FEA9			د FEAA
6	Е 0415	е 0435	ئې 0626+06D0	ي ئې FBF6 FBE4	ب ئې FBF8 FBE6	ب FBE7	ي ئې FBF7 FBE5
7	Ж 0416	ж 0436	ژ 0698	ژ FB8A			ژ FB8B
8	З 0417	з 0437	ز 0632	ز FEAF			ز FEB0
9	И 0418	и 0438	ئى 0626+0649	ى ئى FBF9 FEEF	د ئى FBFB FBE8	د FBE9	ى ئى FBFA FEF0
10	Й 0419	й 0439	ي 064A	ي FEF1	ي FEF3	ي FEF4	ي FEF2
11	К 041A	к 043A	ك 0643	ك FED9	ك FEDB	ك FEDC	ك FEDA
12	Л 041B	л 043B	ل 0644	ل FEDD	ل FEDF	ل FEE0	ل FEDE
13	М 041C	м 043C	م 0645	م FEE1	م FEE3	م FEE4	م FEE2
14	Н 041D	н 043D	ن 0646	ن FEE5	ن FEE7	ن FEE8	ن FEE6
15	О 041E	о 043E	ئو 0626+0648	و ئو FBEE FEED			و ئو FBEE FEED

表 B.1 (续)

序号	西里尔维吾尔文		现行维吾尔文				
	大写形式	小写形式	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式
16	П 041F	п 043F	پ 067E	پ FB56	پ FB58	پ FB59	پ FB57
17	Р 0420	р 0440	ر 0631	ر FEAD			ر FEAE
18	С 0421	с 0441	س 0633	س FEb1	س FEb3	س FEB4	س FEB2
19	Т 0422	т 0442	ت 062A	ت FE95	ت FE97	ت FE98	ت FE96
20	У 0423	у 0443	ئۇ 0626+06C7	ۇ ئۇ FBF0 FBD7			ۇ ئۇ FBF1 FBD8
21	Ф 0424	ф 0444	ف 0641	ف FED1	ف FED3	ف FED4	ف FED2
22	Х 0425	х 0445	خ 062E	خ FEA5	خ FEA7	خ FEA8	خ FEA6
23	Ц 0426	ц 0446	س 0633	س FEb1	س FEb3	س FEB4	س FEB2
			تس 062A+0633	تس FE97+FEB2	تس FE97+FEB4	تس FE98+FEB4	تس FE98+FEB2
24	Ч 0427	ч 0447	چ 0686	چ FB7A	چ FB7C	چ FB7B	چ FB7D
25	Ш 0428	ш 0448	ش 0634	ش FEB5	ش FEB7	ش FEB8	ش FEB6
26	Щ 0429	щ 0449	ش 0634	ش FEB5	ش FEB7	ش FEB8	ش FEB6
			چ 0686	چ FB7A	چ FB7C	چ FB7B	چ FB7D
27	Ъ 042A	ъ 044A					
28	Ы 042B	ы 044B	ئى 0626+0649	ى ئى FBF9 FEEF	د ئى FBFB FBE8	ى FBE9	ى ئى FBFA FEF0
29	Ь 042C	ь 044C					
30	Э 042D	э 044D	ئى 0626+06D0	ى ئى FBF6 FBE4	پ ئى FBF8 FBE6	پ FBE7	ى ئى FBF7 FBE5
31	Ю 042E	ю 044E	يا 064A+0627	يا FEF3+FE8E	يا FEF3+FE8E	يا FEF4+FE8E	يا FEF4+FE8E
32	Я 042F	я 044F	يۇ 064A+06C7	يۇ FEF3+FBD8	يۇ FEF3+FBD8	يۇ FEF4+FBD8	يۇ FEF4+ FBD8
33	Ғ 0492	ғ 0493	غ 063A	غ FECD	غ FECF	غ FED0	غ FECE

表 B. 1（续）

序号	西里尔维吾尔文		现行维吾尔文				
	大写形式	小写形式	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式
34	Ж 0496	ж 0497	چ 062C	چ FE9D	چ FE9F	چ FEA0	چ FE9E
35	Қ 049A	қ 049B	ق 0642	ق FED5	ق FED7	ق FED8	ق FED6
36	Ҳ 04A2	ҳ 04A3	ك 06AD	ك FBD3	ك FBD5	ك FBD6	ك FBD4
37	Ү 04AE	ү 04AF	ئو 0626+06C8	و ئو FBF4 FBDB			و ئو FBF5 FBDC
38	Һ 04BA	һ 04BB	ه 06BE	ه FBAA	ه FBAC	ه FBAD	ه FBAB
39	Ә 04D8	ә 04D9	ئه 0626+06D5	ه ئه FBEC FEE9			ه ئه FBED FEEA
40	Ө 04E8	ө 04E9	ئۆ 0626+06C6	ۆ ئۆ FBF2 FBD9			ۆ ئۆ FBF3 FBDA
41	‘ 0027	‘ 0027	‘ 0027	‘ 0027	‘ 0027	‘ 0027	‘ 0027
			غ 0626	غ FE89	غ FE8B	غ FE8C	غ FE8A
42	‘ 002C	‘ 002C	‘ 060C	‘ 060C	‘ 060C	‘ 060C	‘ 060C
43	؟ 003F	؟ 003F	؟ 061F	؟ 061F	؟ 061F	؟ 061F	؟ 061F
44	؛ 003B	؛ 003B	؛ 061B	؛ 061B	؛ 061B	؛ 061B	؛ 061B