

文章编号: 2095-2163(2019)02-0089-03

中图分类号: TP391

文献标志码: A

现代藏文音节结构分析研究

陈小莹

(西藏民族大学 信息工程学院, 陕西 咸阳 712082)

摘要: 本文主要对现代藏文音节结构进行了分析和研究,通过分析藏文的文字特点和拼写规律同时结合藏文正字法知识,对藏文正确的拼写结构进行了归纳和总结,同时依据小字符集编码方案对藏文音节不同的部件进行划分和确定,进而最终达到对藏文音节结构进行判断和识别藏文音节结构中不同的部件。

关键词: 藏语; 小字符集; 音节结构; 占位辅音

Determination of Tibetan syllable structure based on small character set encoding

CHEN Xiaoying

(School of Information Engineering, Xizang Minzu University, Xianyang Shanxi 712082, China)

[Abstract] This paper mainly analyzes and studies the modern Tibetan syllable structure. By analyzing the Tibetan characters and spelling rules and combining Tibetan orthographic knowledge, the correct spelling structure of Tibetan is summarized, and based on the small character set, the coding scheme divides and determines the different parts of the Tibetan syllables, finally achieves the judgment of the Tibetan syllable structure and the identification of different parts in the Tibetan syllable structure.

[Key words] Tibetan; small character set; syllable structure; occupied consonant

1 藏文文字特征

藏文是属于一种逻辑格语法体系的拼音文字,按照从左到右顺序进行书写。书面藏语的音节字可以只包含一个字母辅音,也可以包含多个字母符号。藏文的音节字中包含一个主要的辅音字符称为基字,其余部位的字母符号按照其相对于基字的位置分别命名。其中,上加字位于基字上方,下加字位于基字下方,前加字位于基字的前面,基字后面的字母符号称为后加字和再后加字。藏文音节中除了基字不可或缺,其它位置上都可为空。藏文文字以音节为基本单位,各个音节之间用音节符“·”隔开,不同的藏文句子之间一般都有分隔符单垂符“|”的存在,不同的段落用双垂符“||”分开^[1]。

与汉语拼音和英文字符不同,藏文音节中的字符并不只是从左到右横排的,还存在竖排结构。藏文字符同时具有横向和纵向拼写的特征也决定了其处理的复杂性。前加字、基字、后加字与再后加字主要进行横向拼写,纵向拼写的主要是基字结构,基字部位上下可能存在上加字、下加字和元音的纵向显示。一般而言,把所有构成藏文文字的字符符号统称为藏文音节字的部件。藏文音节结构形式比较多样化,总体

的构造规则相对固定,基字是一个藏文音节结构中必不可少的一个部件,其它的加字都可以空缺^[2]。

2 藏文拼写特点

在现代藏文文本中,各部件按照藏文正字法组成藏文音节,正字法规定了一套严格而完整的部件构造规则。但是也存在极少部分不符合藏文正字法的音节,这些音节可称为梵音藏文。由于梵音藏文不具备规律性,且在藏文文本中出现的频率较低,所以本文的研究主要针对现代藏文。根据藏文部件构造规则,藏文音节字横向上最长为4个字符位置,纵向上可最多存在4层叠加,横向和纵向以基字为中心,藏文的一个音节最多可以包含7个部件^[3]。藏文音节结构如图1所示。



图1 现代藏文音节结构图

Fig. 1 Modern Tibetan syllable structure diagram

依据藏文正字法知识,同时结合藏文部件构造规则,现代藏文音节字必须满足以下2个条件:

基金项目: 西藏科技厅项目(XZ2017ZRG-56)的阶段性研究成果。

作者简介: 陈小莹(1983-),女,硕士,实验师,主要研究方向:藏文信息处理。

收稿日期: 2018-12-25

(1)藏文音节结构中所有的部件组合形式必须满足现代藏文音节的结构。

(2)藏文音节结构中所有部件的组合搭配要符合藏文音节部件构造规则。

一般情况下,在对现代藏文音节分析时,首先要正确判断出藏文音节的结构形式,确定出音节中的基字;其次,再确定出藏文各部件的编码序列;最后结合藏文正字法知识以及小字符集编码特点,判别是否符合现代藏文音节的组合方式。

3 现代藏文音节正字法知识

藏文中30个辅音字母均可以作为基字,其中一些在基字字性分类中原都属于弱音势的阴性字(浊音字母)“གདབ་མཐོང་མཐོང་”可以做前加字,当将其改变位置作为前加字与特定的基字组合时,辅音字母的字性会改变,对应的也会发生音变^[4-5]。前加字与基字的组合规则为^[6]:阳性前加字“མ”除同组字母“མ་མཐོང་”外还能与阳性基字“གཅན་ཚེ་”和阴性基字“གཅན་ཚེ་ནུང་མཐོང་མཐོང་མཐོང་”结合,但是不能与中性基字“མཐོང་མཐོང་”和其它字母结合;阴性前加字“མ”能与阴性基字“གཅན་ཚེ་”和中性基字“མཐོང་མཐོང་”结合,不能与其它基字结合;中性前加字“ག”能与阳性基字“ཅན་ཚེ་”和阴性基字“ཅན་ཚེ་ནུང་མཐོང་མཐོང་”结合,不能与其它字母结合;中性前加字“མ”能与阳性基字“མཐོང་མཐོང་”和阴性基字“གཅན་ཚེ་”结合,不能与其它字母结合;极阴前加字“མ”能与中性基字“མཐོང་མཐོང་”和阴性基字“གཅན་ཚེ་ནུང་མཐོང་མཐོང་”结合,不能与其它字母结合。10个辅音字母“གཅན་ཚེ་ནུང་མཐོང་མཐོང་མཐོང་མཐོང་མཐོང་མཐོང་མཐོང་མཐོང་མཐོང་”可以作为后加字,在10个可以作后加字的辅音字母中,有4个后加字“གཅན་ཚེ་”后面还可以再加一个辅音字母,作为再后加字。后加字是置于基字所含元音韵母后面做韵尾的,可以与所有的基字横向组合。此外,还有4个辅音字母“མ་མཐོང་མཐོང་མཐོང་མཐོང་”可以在藏文音节中作为下加字存在。藏文辅音字母中有3个辅音字母“མ་མཐོང་མཐོང་”可以加在一部分基字上作为上加字。

藏文正字法知识规定了藏文音节结构中不同位置上的部件组合方式,在具体分析藏文音节结构时,分别从各位置部件组合时应满足的条件进行分析说明^[1],对此可做阐释解析如下。

(1)前加字、上加字和下加字应满足如下条件:

①如果音节中包含有上加字“མ”,则纵向结构的非占位辅音只能存在于集合{ག, མ, ས, ས, ས, ས, ས, ས, ས, ས, ས, ས}中。

②如果包含有前加字“མ”,则后面的基字只能存在于集合,{ག, མ, ས, ས, ས, ས, ས, ས}中。

③如果同时包含有上加字“མ”和前加字“མ”,则基字为非占位辅音,且只能存在于集合{ག, མ, ས, ས, ས, ས, ས}中。

(2)元音、后加字和再后加字应满足如下条件:

①藏文音节中有可能不显示元音,但是一般会在发音时加上/a音,也可以是{མ, མ, མ, མ}中之一;

②如果存在后加字,则后加字必在集合{ག, མ, ས, ས, ས, ས, ས, ས, ས}中。

③如果存在再后加字“མ”和“མ”,则一定有后加字。当再后加字为“མ”时,后加字在集合{ག, མ, ས, ས}中,当再后加字为“མ”时,此时后加字在集合{ག, མ, ས}中。

4 现代藏文音节结构特征

4.1 藏文编码方案

藏文文字在计算机中是以编码的形式显示的,但是因目前对藏文编码的研究各不相同,造就了不同的藏文编码方式,从而藏文编码形式各异。从目前对藏文的研究来看,藏文的编码方案可以分为大字符集编码方案和小字符集编码方案两种。其中,大字符集编码方案将藏文中纵向结构的字丁组合按照一个单元进行横向编码,小字符集编码方案通过对藏文基本字符建库,藏文文字通过藏文基本字符动态组合形式显示,不同显示方式的藏文基本字符对应不同的编码^[7]。

4.2 小字符集编码藏文音节结构特征

小字符集编码方案中对藏文中前加字、基字、上加字、下加字、后加字与再后加字等部件进行编码。带有纵向结构的字丁组合在计算机显示时只占据一个字符,字符的宽度由纵向字丁组合中的第一层辅音决定,此时第一层辅音为占位辅音,对应的编码为占位辅音编码,该占位辅音下方的字符不单独占宽度,称为非占位辅音,对应的辅音编码称为非占位辅音编码^[7]。在藏文文字中,元音都在辅音的上方或下方以叠加方式存在,依附于纵向结构的其它字丁,不单独占宽度,因而其也为非占位字符。(下转第95页)