doi: 10. 16104/j. issn. 1673-1883. 2022. 03. 002

彝文古籍文献术语标准化与数据库建设研究

王愈欣

(西南民族大学中国语言文学学院,四川 成都 610041)

摘 要:彝文文献是中华民族的宝贵财富,在文献信息化迅猛发展的今天,彝文古籍文献术语标准化已成为彝文文献研究与发展中亟待解决的问题。通过对大量彝文古籍文献内容进行筛选与整理,找出其中词频最高的彝文词语,并对其进行汉语规范标注,制定彝文古籍文献术语翻译标准原则,在技术指导下建立彝文古籍术语标准化数据库,为彝文古籍数字化管理提供更加高效便捷的手段,助推彝文古籍传承与保护工作更进一步发展。

关键词:彝文;文献术语;标准化;数据库

中图分类号:G255.1 文献标志码:A 文章编号:1673-1883(2022)03-0006-06

Research on Standardization of Terminology and Database Construction of Yi Ancient Books

WANG Yuxin

(College of Chinese Language and Literature, Southwest Minzu University, Chengdu, Sichuan 610041, China)

Abstract: Yi ancient books are the precious wealth of the Chinese nation. With the rapid development of document informatization today, the standardization of terms in Yi ancient books has become an urgent problem to be solved in the research and development of Yi ancient books. Firstly, this paper through to a large number of Yi ancient books content filtering and sorting, finding out one of the highest frequencies in Yi language words, and carrying on the Chinese specification tagging. Secondly, making Yi ancient books standard terminology translation principles. Finally, setting up ancient books database of Yi language terminology standardization in the technical guidance, to provide more efficient and convenient for Yi language ancient books digitization management means, then we will further promote the inheritance and protection of Yi ancient books.

Keywords: Yi language; document term; standardization; database

一、引言

彝文古籍文献是中华文化宝库中的璀璨明珠,其种类繁多且数量庞大,涉及彝族社会生活的方方面面,记载着许多彝族文明成果和文化现象,是中华民族古籍的重要组成部分,蕴含着巨大的文化价值和实用价值。近年来,彝族古籍文献的价值不断得到学界的重视,在文献搜集与整理、文献翻译与出版方面研究成果颇丰,特别是在文献信息化日益普及的今天,彝文古籍文献的研究领域也日渐扩大。应用现代数字技术对彝文古籍文献进行保护、开发与

利用也成了当今时代彝文文献研究的首要课题。

由于彝文自身表音性强等语言特点,目前 学界对于彝文古籍文献术语翻译还没有规范的 原则和统一的标准,这就使得在彝文古籍文献 研究中出现了如在文献翻译时同字不同译等情况,造成信息冗余等问题。这些问题同时也阻 碍着彝文古籍文献的信息化处理研究,更影响 着彝文古籍文献的传承与保护。因此,彝文古籍文献术语规范化、标准化是目前彝文古籍文献的核对与保护、整理与研究以及文献数字化建设中的突出问题,必须引起重视。

收稿日期:2022-04-11

基金项目:国家社科基金重大项目:大小凉山彝文经籍文献语音资源库建设(19ZDA284);西南民族大学中央高校基本科研业务费专项资金优秀学生培养工程项目:彝文古籍文献术语标准化数据库建设(2021SYYXSS04)。

作者简介:王愈欣(1996—),女,山西太原人,硕士研究生,主要研究方向:中国少数民族文献、彝族文献。

二、术语与术语标准化工作

(一)术语定义

术语是在特定专业领域中一般概念的词语指 称[1]。术语是约定性的语言符号,它通过语音或文 字来表达或限定科学概念,是思想和认识交流的工 具,具有专业性、科学性、单义性等特征。术语是一 种特定的学术语言,是官方正式的用语,属于名词 范畴。术语分为一般性术语和专业用术语。一般 性术语指由人们生活和工作习惯形成的,在正常生 活中使用频次非常高的词语,其含意往往比较含 混,一个一般性术语往往在不同的语境中表达不同 的概念。一般性术语与所表达的概念并不是—— 对应的,如"火车的车头""自来水管的放水活门"和 "杰出人物的领袖"都可以用"龙头"这个一般性术 语来表示。因此,一般不要求一般性术语的概念表 达做到严格精准。专业用术语指在特定专业范围 领域中表达固定概念的词语,如LAN (Local Area Network)局域网,指在某一区域内(一般是方圆几千 米以内)由多台计算机互联成的计算机组。专业用 术语多用于专业技术文献、标准和规定中,其含义 表达要求严格精确,不允许模棱两可,也不允许因 使用人的不同而有所差异,专业用术语的概念表达 必须做到准确无误。本文讨论的"彝文古籍文献术 语"就属于专业用术语的范畴。

(二)国家术语标准化工作

我国一直以来都十分重视术语标准化工作。 关于术语标准的编订,国家标准化管理委员会于 2019年8月发布了中华人民共和国国家标准《术语 工作 原则与方法(GB/T 10112-2019)》,其中规定了 制定和编纂各专业领域术语集的基本原则和方法, 确立了术语构成和定义表述的一般原则。同时,在 大部分专业学科与生活领域中,均有相关术语标 准,如《颜色术语(GB/T 5698-2001)》《电工电子产品 环境条件术语(GB/T 11804-1989)》《畜禽环境术语 (GB/T 19525.1-2004)》等。在古籍文献领域,我国 有《古籍著录规则(GB/T 3792.7-2009)》对古籍著录 项目、著录单元以及著录标识符号等进行标准规 范,但其中并没有关于古籍文献的术语标准。虽然 《术语工作 原则与方法(GB/T 10112-2019)》适用于 各专业领域的术语标准化工作,但对于民族古籍文 献来说,各民族古籍文献均有自己的民族特性,该 标准只可起到参考作用,应该在此基础上建立适用 于民族古籍文献术语标准化的原则与方法,以方便 民族古籍文献的研究与开发利用。

三、彝文古籍文献术语现状

目前在不同的彝文古籍文献中,常出现彝文词语同义异译、同音异译的现象,甚至在同一本彝文古籍文献的不同版本中,相同彝文词语的汉译形式也不尽相同。彝文古籍文献术语不统一的主要原因是没有建立规范与标准。造成目前彝文古籍文献术语翻译用词混乱的原因主要有两点:

一是汉语存在大量同音字。汉语同音字太多,用什么汉字转写对应的彝文词语并没有可遵循的标准原则,翻译用字的主观性极强。这使得在对部分彝文文献中的专业名词术语进行汉文转写时,人们往往根据彝文的读音翻译书写,常出现同一彝文古籍文献术语概念有多个不同形式汉文表达,如彝族传世的经典文献名称的翻译转写,常见的有"玛牧特依"与"马牧特衣","支呷阿鲁"与"支格阿龙""支嘎阿鲁"等等。此外,在涉及彝文古籍文献中的地名、人名、器具等概念时,往往也会由于音译差别,夹杂着研究者的主观因素,产生许多同物不同名的现象。

二是彝语方言差异大。仅现代彝语就有北部、东部、南部等6个方言,其中包括5个次方言与25个土语。不同方言区的彝语语音差异大,甚至不能交流。这使得不同方言区的彝文文献术语在翻译时会出现同物异音异译的形式,如彝文古籍文献中常见的表达"彝族传统宗教中的祭司"这一概念的词语,根据其读音或意义进行翻译转写,其汉文形式有"毕摩""奚波""西波""鬼师""拜玛""白莫""比姆""比母""笔母""阿毕"等十几种写法,这是方言差异和同音异译的结果。再如,表示"彝族民间脍炙人口、广为传诵的诗体口传文学"这一概念的词语,在翻译转写时常有"克哲""克智"等汉文形式。

彝文古籍文献术语翻译用词混乱,缺乏统一的规范标准,这对彝文古籍的著录工作与信息化处理带来极大的阻碍与挑战,不仅严重影响了彝文古籍文献的研究效率,甚至可能还会影响到不同民族间的文化交流。由此看来,制定彝文古籍文献术语标准化翻译原则,建立彝文古籍文献术语标准化数据库迫在眉睫,这是彝文古籍文献研究发展的需要,也是信息技术兴起的需要。

四、彝文古籍文献术语标准化原则

彝文古籍文献术语标准化,即运用标准化的原理和方法,根据彝文古籍文献中使用最高频的彝文字词特点,通过制定彝文古籍文献术语标准,使之

达到一定范围内的术语统一,从而获得最佳秩序和 社会效益。彝文古籍文献术语标准化研究主要针 对彝文音译性名词术语,同时兼顾意译性名词术语 展开。

(一)彝文古籍文献术语标准化总则

彝文古籍文献术语标准化原则的制定需要符合《中华人民共和国国家通用语言文字法》,维护国家主权和民族尊严,加强国家统一和民族团结。彝文古籍文献术语汉译需使用规范汉字,且符合普通话的标准语音。同时,彝文古籍文献术语也应符合国务院批准的《彝文规范方案》,参考正字法原则以及新词术语规范原则,统一使用规范彝文进行书写。此外,彝文古籍文献术语标准化还应遵循以下四点总则:

1.单义性

彝文古籍文献术语应做到表意唯一,即一条彝文术语对应唯一概念,且对应唯一汉译形式。对于表意相同但汉译字形不同的彝文古籍文献术语,需要在已有几个同义词之间选择最具代表性的一个词语形式,整合辨析同义术语、变体术语,选择能较好满足术语表达要求的字形与意义概念。例如"阝H&Q",其对应"彝族道德经""彝族训世经""彝族劝善经""彝族教育经典"等多个同义概念,应将其统一整合为"教育人们做人做事的经书"这一概念。同时,其汉译形式有"玛木特依""马木特依""玛牧特衣""玛穆特依""玛牡特依"等多种,应根据单义性总则以及下文所述简明性与约定俗成性总则,参考下文所述各细则,选取"玛牧特依"作为唯一的规范汉译形式。

2. 简明性

术语的信息交流要求术语尽可能简明,以此提高效率。因此,在彝文古籍文献术语汉译形式的选择上,应尽量选取简单易记的汉字作为其书写形式,且字与字的组合是无歧义的,术语要尽量精炼且表述准确,要便于使用与交流,如"引出Q",选择"阿惹妞"为标准术语形式,而不选"阿热牛""阿惹妞"等汉译术语形式。

3.约定俗成性

对于彝文古籍文献中使用频率较高、出现范围较广,且已经在彝族社会得到普遍认可,符合使用族群语言习惯的词语,其符合术语的约定俗成性,即使在研究或使用中有个别不理想之处,若没有重要原因或重大变故,则不宜轻易变更。如"划88 出",滇、川、黔、桂四省区彝族共同尊奉的祖先,选择

"阿普笃慕"为术语标准汉译形式。虽其用字不符合简明性的要求,但其已在彝族社会得到普遍认可与广泛使用,应按照约定俗成的原则收录为标准术语汉译形式。

4.非歧视性

国家通用语言文字的使用应当有利于维护国家主权和民族尊严,有利于国家统一和民族团结,有利于社会主义物质文明建设和精神文明建设。彝文古籍文献术语的制定要符合非歧视性原则,不使用有伤民族感情或损害社会公共利益的词语,不使用带有歧视性与侮辱性的文字,不带有褒贬等感情色彩的意蕴,特别是在音译术语用字的选择上。

(二)彝文古籍文献术语标准化细则

彝文古籍文献术语标准化除要符合上述总则外,还要根据彝语与汉语对译的实际情况进行编订,细则如下:

1. 意译性名词术语需遵从原义, 灵活表达

2. 音译性名词术语要基于国家标准普通话读音尽量原音对译

因彝语和汉语的音位系统和音位组合关系不同,所以在进行彝文古籍文献术语翻译时,首先要考虑两种语言音位系统的对译问题。

从声母来看,现代汉语声母由21个辅音声母与1个零声母组成。现代彝语声母有43个,其数量远远大于现代汉语的声母数量,再加上现代彝语声母中有舌根音、喉音以及清化鼻音等特殊发音部位与发音方法,这就造成了现代彝语声母与现代汉语声母无法实现——对应。因此,在声母的对译上,参照《彝文规范方案》中,"汉语借词应尽量用与汉语普通话语音相当的字书写"这一条规定,将现代彝语声母与有相同或相近发音部位以及发音方法的现代汉语声母相对应,依据实际情况,参考上述总则,选择最恰当的汉字来实现对译。具体对应情况如表1所示。

从韵母来看,现代汉语韵母共39个,按结构、韵 尾以及开头元音发音口形可分为多种类型。现代 彝语韵母共10个,分别是i、ie、a、uo、o、e、u、ur、y、yr,

表1 汉彝音译声母对应表

汉语声母	彝语声母	汉语声母
b	q	q
p	X	X
m	zh\rr\nr	zh
f	ch	ch
d	sh	shi
t	r\ss	r
n	z\zz\nz	z
1	c	c
g	s	s
k	y	y
h	v\w	W
j		
	b p m f d t n l g k	b q p x m zh\rr\nr f ch d sh t r\ss n z\zz\nz l c g s k y h v\w

其发音时的舌位高低前后以及唇形圆展都与现代 汉语韵母的发音状态基本一致,但现代彝语韵母在 发音时根据声带的松紧又可分为松元音(i、o、e、u、y)和紧元音(ie、a、uo、ur、yr)的对立,因此,在进行彝 汉术语对译时,要注意一个彝语韵母可以灵活对应 多个汉语韵母,如:i→i\ü\ei\in\ing,具体对应情况如 表2所示。

表2 汉彝音译韵母对应表

彝语韵母	汉语韵母	彝语韵母	汉语韵母	
i	i\ü\ei\in\ing	e	e\ai\en\ün\eng	
ie	ie	u	u\ui\üe\un	
a	a\an\ang	ur	u	
uo	uo\ao	y \yr	\mathbf{r}	
0	o\ou\iu\eng			

从声调来看,现代汉语普通话有四个调,即阴平(55)、阳平调(35)、上声(214)、去声(51)。现代彝语也有四个调,即高平调(55)、次高调(34)、中平调(33)、低降调(21)。从调值来看,只有彝语的高平调(55)与汉语的阴平调(55)相等,其余在汉语中都没有相当的调值和调类。因此,按照现代彝语声调与现代汉语声调对应的曲直规律来看,可以将彝语的高平调(55)与汉语的阴平(55)相借,将彝语的低降调(21)与汉语的去声(51)相借,彝语次高调(34)根据实际需要与汉语上声(214)或阳平(35)相借,彝语中平调(33)根据实际需要与汉语阳平(35)

或阴平(55)相借。具体对应情况如表3所示。

表3 汉彝音译声调对应表

	汉语声调
55(高平调t)	55(阴平)
34(次高调x)	214(上声)\35(阳平)
33(中平调)	35(阳平)\55(阴平)
21(低降调p)	51(去声)

彝汉两种语言词汇术语的互相转写,除了少部分有对等词(字)以外,大部分都需要近音转写,这往往会附加转写者的很多主观因素。为使用规范和统一标准,特制定"汉彝音译声母对应表""汉彝音译韵母对应表"以及"汉彝音译声调对应表"供彝文古籍文献术语对译参考。

3. 人名、神名、地名等尽量采用和原读音相一致的 汉字

在对译古籍中的人名、神名、地名时,应遵从一个彝语读音要与一个汉字相对应的原则,尽量选取和原读音相一致的汉字,避免在彝族社会产生误解与使用偏差。若无与原读音一致的汉字,则应根据谐音对应汉字的方法,选择读音相近的汉字对译,同时还要使用不同的汉字对彝语读音相近的音节做出细微的辨别,如:"云身了8*(ka³zu-³ho³phu³bo³)"译为"甘约伙普山","糾氧①*(ba³dzo³zì³o³bo³)"译为"巴久以俄山"等等。

4.保留彝文古籍术语中特有的性别用词(字)

如"惹""嫫""妮"等,"≵淫钅θ(tsl³⁵ tsl²¹ ni³³ndz,a⁵⁵)"为彝族女性题材文献,故选用"姿子妮乍"作为其汉译形式,而不选用"姿子尼乍"。

5.尽量少用或不用音译外国人名中的常用汉字

在对彝文古籍术语进行汉译转写时,要遵从彝族本民族文化特征,继承自己民族的优秀传统,尽量避免选择让人产生异域文化特征臆想的词(字),如"石一"与"矢野"等。

除上述总则与细则外,还有一点另需说明:在前期对大量词条数据进行整理、分析与翻译的过程中,暂未发现各项总则与各项细则相冲突的情况,若后续彝文古籍文献术语研究工作中出现总则与细则冲突的情况,建议采取"细则优先,总则辅助"的处理方式,即在符合细则已做出规定的情况下,优先适用细则,并结合总则辅助调整翻译后的术语;在细则无规定的情况下,术语翻译要以总则为主。

五、彝文古籍文献术语标准化数据库建设

古籍文献标准化中的术语问题,就是专业领域内的概念标准化问题。这是古籍信息化处理的首要工作之一。随着彝文文献信息化处理的发展,彝文古籍文献中彝文与汉字的双语术语转换需求也越来越大,双语术语转换技术在语言自动处理、机器翻译、跨语言信息检索等领域都有重要应用,而双语术语转换首先需要建立庞大的术语数据库。彝文古籍文献术语标准化工作除制定标准化术语编订原则外,也需要建立彝文古籍文献术语标准化数据库,以便于日后对彝文古籍文献做进一步地研究与应用。

(一)数据来源与数据整理

彝文古籍文献术语的数据获取主要有两种途径:一是由专业语言工作机构提供;二是人工采录。在前期的彝文古籍文献术语收集整理工作中,共采集到彝文古籍文献术语词条1110项,其中凉山州语委提供了"部分彝文典籍词汇汉译资料"共530项词条,同时,从经典彝文古籍如《西南彝志》等文献中人工采录得到彝文古籍文献术语词条580项。将采集到的词条数据进行整理、分析与归纳,根据词频保留较高频名词性术语词条,剔除重复词条项,整合归并同义词条(如"划章00""划章05"归并为"划章00,阿嘎伙俄"),同时对词条进行彝文注音与国际音标注音,并根据上述彝文古籍文献术语标准化各项原则确定其唯一的汉译形式。通过对词条数据进行处理,最后共获得655项有效词条,用于彝文古籍文献术语标准化数据库建设。

(二)数据库建设流程概述

彝文古籍文献术语标准化数据库采用 MySQL 进行建设,其具有体积小、速度快、成本低、开放源码等特点。通过 Navicat Premium 12 工具对 MySQL 数据库进行连接与管理,创建彝文古籍文献术语数据库 IT 术语对照表,表格中主要包括"彝文id""彝文""彝文注音""国际音标""汉译"以及"其他"六项,如表 4 所示。将搜集到的 655 条有效术语词条按类别导入 MySQL 数据表中(图1),完成数据的导入与储存。随后根据目标需求进行双语术语的添加、修改、删除、查找和浏览等操作。

因网页文件本身是一种文本文件,所以在数据库建设过程中通过HTML超文本标记语言中的一系列标签来制作可以说明文字、图形、声音、表格、链接等的HTML文档,再通过Web浏览器读取HTML文档,从而创建前端页面,以此实现网页显示的

表4 彝文古籍文献术语数据库IT术语对照表

名称	类型	主键	外键	备注
id	int(11)	yes	no	彝文id
yiwen	Char	no	no	彝文
phoneticNotation	text	no	no	注音
intPhonAlphabet	text	no	no	国际音标
translation	text	no	no	汉译
comment	text	no	no	其他

id	yiwen	phoneticNotation	intPhonAlphabet	translation	comment
1	11 €	a dur	a33tu33	阿都	(Null)
2	41441米莱	a dur yit mo bbo	a33tu33 zi55mo33b	阿都依某山	(Null)
3	11≥010	a gga ho o	a33ga33xo33o33	阿嘎伙俄	(Null)
4	11≥0 <u>₿</u>	a gga ho pux	a33ga33xo33phu34	阿嘎伙普	(Null)
5	김홍반일	a gga lit tuo	a33ga33li55tho33	阿甘乃拖	(Null)
6	礼を用る	a tur sha lo	a33thu33§a33lo33	阿图沙洛	(Null)
7	1 1!	ap bbu	a21bu33	阿布	(Null)
8	440€	ap ggep ma lo	a21gu21ma33lo33	阿格马罗	(Null)
9	北京北京	ap hly ap ddip	a21t133a21di21	阿勒阿迪	(Null)
10	44110	ap hly ggop o	a21t133go21o33	阿尔果峨	(Null)
11	4.L	ap ly	a21l ₁ 33	阿尔	(Null)
12	450m	ap mop qo lu	a21mo21t6ho33lu33	阿莫曲日	(Null)
13	1/4 M =	ap nrep qit tuo	a21ndzw21tchi55th	阿哲且拖	(Null)
14	4. 注	ap nry	a21ndzj33	阿哲	(Null)
15	4114	ap nyip hxo lo	a21nj21ho33lo33	阿尼伙洛	(Null)

图 1 MySQL的图形界面

功能。

其次,通过构建SSM框架作为后端进行开发, 实现对前端页面操作的响应功能。SSM框架是通过 使用Spring架构实现业务对象管理、Spring MVC架 构负责请求的转发和视图管理以及 Mvbatis 架构作 为数据对象的持久化引擎来搭建的。构建的SSM 框架包括四层,分别为Dao层、Service层、Controller 层和 View 层,其中 Dao 层负责与数据库进行联络任 务的封装,完成Service层提交的数据库操作,如增、 删、改、查等:Service 层负责业务模块的逻辑应用设 计,将复杂的业务拆分为多个细粒度的Dao层操作, 调用Dao 层获取数据然后封装为结果;Controller 层 负责具体业务模块流程的控制,通过接收表单请 求,将请求划分为细粒度业务,再调用相应Service 完成业务; View 层和控制层紧密结合,负责前台 JSP 页面的显示,如显示返回的数据、页面提示与返 回等。

最后,点击搜索按钮后,搜索框里的内容会在 SSM中与数据库进行连接查询,并返回查询值,由此 实现彝文术语与汉译形式的查询功能(图2)。本数 据库支持彝汉双语查询,同时显示对应的彝文注音 与国际音标注音,受众范围广,可以满足用户的不 同检索需求,极大地便利了彝文古籍文献的学习与 研究。



图 2 彝文古籍文献术语标准化数据库查询界面

六、结语

构建彝文古籍文献术语标准化数据库,可为彝文古籍文献术语标准和规范的制定提供翔实的数据参考,有利于协调、规范和统一彝文古籍的数字化管理。同时,基于彝文古籍术语标准化数据库的

建设,可以进行彝文古籍文献标准术语翻译软件的 开发,这对于彝文古籍的保护、开发与利用和彝族 文化的研究与传播有着积极的推动作用。彝文古籍文献术语标准化数据库的建立也将大幅提高学者的研究效率,对彝族历史文化保护与传承有着重大的影响。

综上所述,在大量实践与探索的基础上,基于 彝文文献数字信息化处理的研究背景,结合文献 学、文学、语言学、民族学以及计算机科学等多学科 领域的研究方法,得出了彝文古籍文献术语标准化 的具体原则。基于此建立的彝文古籍文献术语标 准化数据库是在理论指导下的实践创新成果,在后 续的研究工作中对彝学研究、文献著录信息化以及 彝文古籍文献的数字化保护与传承有着重要的 意义。

参考文献:

- [1] GB/T 15237.1-2000, 术语工作词汇第1部分: 理论与应用[S].
- [2] 王秀英.现代彝语概论[M].成都:四川民族出版社,2017:33-63.
- [3] 马鑫国,乌尼乌且.汉译翻译教程[M].成都:四川民族出版社,2004.
- [4] 沈乃文,曹淑文.古籍著录标准化的名词术语问题[J].大学图书馆通讯,1986(4):9-15.
- [5] 吉差小明,朱崇先,彝文古籍数字化研究现状与展望[J],哈尔滨师范大学社会科学学报,2016,7(3):111-113.
- [6] 兰旭辉,熊家军,邓刚.基于MySQL的应用程序设计[J].计算机工程与设计,2004,25(3):442-443+468.
- [7] 木合亚提·尼亚孜别克,古力沙吾利·塔里甫.汉哈双语IT领域术语资源库检索系统的设计与实现[J].信息技术与标准化,2020(11):36-38.
- [8] 张淑娜,陈丽云,严世芸.术语学视域下的术语标准化[J].中国科技术语,2022,24(1):80-86.