

# 藏缅语族语言的数理分类及其分析

邓晓华 王士元

[提要] 本文运用词源统计分析法, 对中国境内的藏缅语族语言作出数理分类以及亲缘关系程度的描述, 并通过树枝长短来表示距离关系。显示藏缅语族语言的类簇和分级层次, 分析其形成过程。

藏缅语族是汉藏语系中语种最多、分布最广、内部差异最大的一个语族。在中国, 使用藏缅语族的民族有: 藏、门巴、珞巴、景颇、独龙、怒、羌、普米、彝、哈尼、阿昌、拉祜、基诺、傈僳、纳西、白、土家等 17 个民族, 据 1987 年《中国语言地图集》(以下简称《地图》) 公布中国境内的藏缅语族语言共有 23 种。而据最近的研究(1997), 中国境内的藏缅语族语言则已增加到 36 种, 原因是发现了一些新的语种, 例如, 木雅、道孚、扎巴等。就国内外藏缅语族语言的数量来说, 分歧较大, 国内学者认为 100 余种, 而马提索夫则认为至少有 250 种。主要原因是划分语言和方言的标准不统一。藏缅语族语言人口约 1500 万。

## 一 藏缅语族语言的传统分类

19 世纪 50 年代开始认识藏语跟缅语有同源关系。最早英国学者 Konow Sten 《印度语言调查》(1903—1928) 藏缅语分 9 个分支。1937 李方桂《中国的语言和方言》采用谱系树分类, 分藏缅语族为藏、景颇、缅、彝 4 个语群; 分语群和语言 2 级。罗常培、傅懋勳《国内少数民族语言文字概况》藏缅语族分类近似李, 分 4 个语支。而《地图》则分藏缅语族 23 种语言为 7 类: 藏、嘉戎、错那门巴、仓洛门巴; 羌、普米; 博戈尔珞巴、义都珞巴、格曼僜语、达曼僜语; 独龙、景颇; 载瓦、阿昌; 彝、纳西、白、傈僳、怒、拉祜、哈尼、基诺; 土家语。

戴庆厦(1989) 分藏缅语族为南北两大语群。孙宏开分藏缅语族为 5 个语支, 增设羌语支, 即藏、羌、彝、缅, 景颇。孙 1994 年又增加了中国境外藏缅语族语言, 作全面的藏缅语族分类。此外, 本尼迪克特和马提索夫也对藏缅语族进行了分类。

不难看出, 上述分类的标准主要有: 语音面貌, 语法范畴, 词汇系统即同源词的多寡和对应关系的远近。以上分类除了国外藏缅语, 涉及国内藏缅语问题的主要分歧有: 1. 羌语是否单列一个语支, 瞿蔼堂归藏语支, 戴庆厦归嘉戎—独龙语支, 孙宏开单列。2. 白语的归属, 两种意见: 一是归汉语族。格林伯格 1953 《历史语言学和文字语言》认为白、汉有发生学关系, 主张白语归汉语族, 本尼迪克特同意此观点, 划归汉语族。郑张尚芳同此观点。二是归藏缅语族。或单列独立语支, 或归入彝语支。戴庆厦单立、孙宏开、吴安其归入彝语支。他们都认为白、汉是接触关系, 而非发生学关系。

## 二 藏缅语族语言的计量分类

### 1. 分类的标准和方法

100 词及数量方法。本文主要运用词源统计分析和方法来作藏缅语族语言的亲缘关系分类。词源统计分析法的基本观念是两种具有亲缘关系的语言分离的时间深度，可以通过它们继承的词共享程度来判断。词源统计分析法最初由 Swadesh 提出，近年来，由于理论的发展和计算机的巨大进步，词源统计分析法又受到历史语言学家的重视。王士元（1995）曾运用该方法作出壮侗语族的分类。

我们主要依据 Swadesh 的 100 词表作同源词的数理分析，同时参考选择雅洪托夫的 35 词表，具体分析了藏缅语族的 12 个语言，<sup>①</sup>最后画出藏缅语族语言树形图。树形图包含两个重要的信息：（1）语言集团的呈阶级式的聚合分类；（2）树枝的长度可以反映语言从祖语分离的时间距离以及各语言间亲缘程度。

### 2. 数理分类的主要步骤

#### （1）相似矩阵（Similarity Matrix）

首先优选出同源词，编制同源词表，然后计算出每对语言的同源百分比。

	拉	阿	羌	普	嘉	景	緬	彝	哈	纳	白	土	OC
拉		92	54	49	53	44	52	51	55	54	33	35	31
阿			52	50	52	43	52	52	55	51	31	33	31
羌				79	73	56	60	58	60	65	40	40	34
普					74	46	65	63	66	64	34	32	34
嘉						49	63	60	60	59	34	37	33
景							54	43	48	50	34	33	28
緬								67	69	68	34	35	32
彝									72	61	31	37	30
哈										72	34	38	32
纳											32	33	31
白												23	33
土													20

#### （2）距离矩阵（Distance Matrix）

由于数理树形图是通过分枝的长度来反映语言间的距离的，所以，我们必须把上面的相似矩阵转换为距离矩阵(Distance Matrix)。

	拉	阿	羌	普	嘉	景	緬	彝	哈	纳	白	土	OC
拉		6.15	29.3	33.5	30.1	38.2	30.9	31.8	28.5	29.3	50.7	48.1	53.4
阿			30.9	32.6	30.9	39.2	30.9	30.9	28.5	31.8	53.4	50.7	53.4

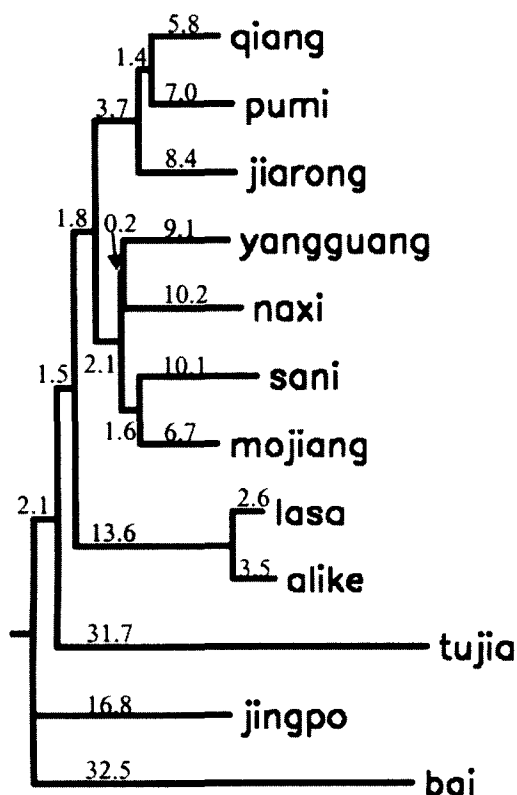
<sup>①</sup> 该文表中汉语缩写“拉”代表 Lasa、“阿”代表 Alike、“羌”代表 Qiang、“普”代表 Pumi、“嘉”代表 Jiarong、“景”代表 Jingpo、“緬”代表 Yangguang、“彝”代表 Sani、“哈”代表 Mojiang、“纳”代表 Naxi、“白”代表 Bai、“土”代表 Tujia，“OC”代表古汉语。文中不再作说明。

羌	12.8	16.2	27.7	24.7	26.2	24.7	21.2	42.3	42.3	49.4
普		15.6	36.3	21.2	22.6	20.6	21.9	49.4	52	49.4
嘉			33.5	22.6	24.7	24.7	25.4	49.4	45.7	50.7
景				29.3	39.2	34.4	32.6	49.4	50.7	57.8
缅					19.9	18.6	19.3	49.4	48.1	52
彝						16.8	24	53.4	45.7	54.8
哈							16.8	49.4	44.6	52
纳								52	50.7	53.4
白									66.4	50.7
土										72.4

(3) 从无根树到有根树

从距离矩阵转换成无根树有许多种方法，所有这些方法都需要大量的计算。生物学家为生物种系发生分类设计出很好的计算程序。我们采用的计算程序，则是最有影响的，1967年由 Fitch 和 Margoliash 发明的程序，以及 1987年由 Saitou 和 Nei 发明的程序。

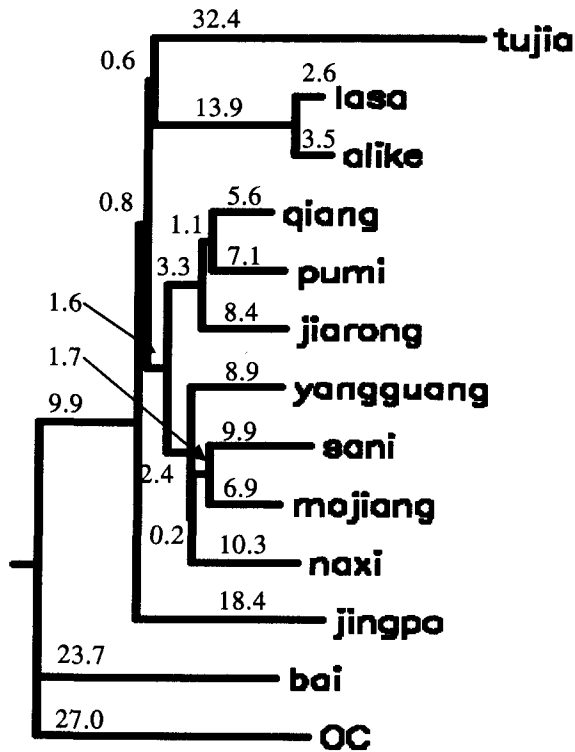
树图一 (12 languages:Neighbor joining:bai as an out-group) :



藏缅语族语言数理树形图 (树图一) 显示的藏缅语族 12 支语言, 可分为 6 个较大的聚类, 即白、景颇、土家、藏 (拉萨、阿力克)、彝缅 (哈尼、彝、纳西、缅)、羌 (嘉戎、普米、羌)。其分级和层次可假设为: 因白、景颇树根分叉点的数字为零, 可以认为树图有四个层次: 白、景颇、土家等; 土家与藏 (拉萨、阿力克) 等; 藏与彝缅 (哈

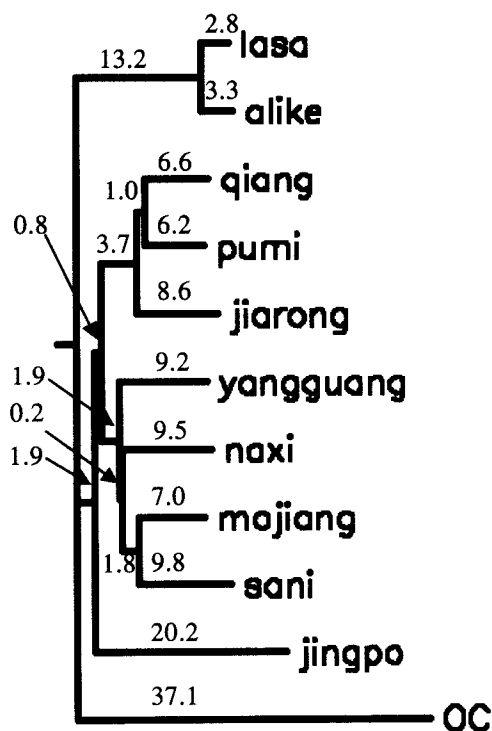
尼、彝、纳西、缅)、羌(嘉戎、普米、羌);彝缅与羌,显示平行关系。彝缅的远近关系依次为哈尼、彝、纳西、缅;羌的远近关系依次为普米、羌、嘉戎。树图的数字代表树枝的长短距离,反映语言之间亲缘关系的远近,树枝长的表示两种语言亲缘距离远,树枝短的则表示亲缘距离近。而同一个小簇类里的语言关系则比外簇类的语言关系近。树图一显示在藏缅语族 12 支语言中,白与土家的线条特别长,这可理解为:如果我们假设语言分化后,各自演变的速度本来应是均等相同的,但是由于白与土家与其他语言接触多,借词成份特别多,所以其变化速度特别快,这就显示出白与土家与其他语言的不同演变结果。

增加古汉语的数据,变成 13 支语言,发现图一计算出来的 12 支藏缅语族语言分类格局仍然未变,见下面的树图二 (neighbor joining):



树图二显示 13 支语言,可分为 7 个较大的聚类,即古汉、白、景颇、土家、藏(拉萨、阿力克)、彝缅(哈尼、彝、纳西、缅)、羌(嘉戎、普米、羌)。由于景颇、土家的树根分叉点的数字近为零,所以不能认为土家跟藏(拉萨、阿力克)为一个小簇。可以假设认为树图的第一层为三分,即古汉、白与景颇等;第二层为四分,即景颇、土家、藏(拉萨、阿力克)、彝缅(哈尼、彝、纳西、缅)和羌(嘉戎、普米、羌)。第三层为二分,即彝缅(哈尼、彝、纳西、缅)和羌(嘉戎、普米、羌),其内部距离远近如图一所示。

Using Fitch OC as out-group



树形图说明：

我们使用 Neighbor Joining 和 Fitch 两种计算方法，结果都一样。各分级层次的类簇相同。

增加汉语数据，由 12 支语言变成 13 支语言后，分类的情况仍然相同。

我们曾经考虑过白与土家的线条特别长，是否因为白与土家这两支语言的借词成分特别多，我们采用了排除的方法，即先排除白与土家语，看看其他语言的分类结果是否跟未排除前一样，结果显示跟原先的分类基本一致，这就说明分类结果是可信的。

从三个不同的树图比较来看，我们有个重要的发现：随着语言数的增加和减少的数据的变化，树图的反映很敏感，都会体现在树图的不同变化上；而最重要的变化是白、景颇、土家这三支语言在图中出现的位置，即摆放的位置不同，这就反映了这三支语言在 12 支藏缅语言中的特殊地位不同，而传统的分类正是认为白、土家系属未定，而景颇地位特别，学术界为此常常引起争论。数理分析的结果跟传统的定性分析相合。

### 3. 几点结论

(1) 据树图藏缅语族语言可分为 6 个较大的聚类，即白、景颇、土家、藏（拉萨、阿力克）、彝缅（哈尼、彝、纳西、缅）、羌（嘉戎、普米、羌）；这种分类结果与戴根据不同材料、不同的方法作出的分类完全一致。

(2) 树图表示白、景颇、土家各自为独立的一个聚类，可视为单一语支；特别是白语，单一性特别明显，这个结论与《地图》将白语归入彝语支的观点有很大差异。但是白语虽然跟古汉语在假设的同一个分级层次上，两者关系特殊，但并没有跟古汉语形成同一个簇类。从亲缘关系的距离看，白语离树根最远，其次是景颇语和土家语。同其他

藏缅语族语言比较，白语与彝语聚类最近，特别是跟哈尼语最近；这跟传统分类往往把白语划入彝语支相似。

(3) 本尼迪克特认为白语归汉语，藏缅语族语言以景颇语为中心，向四周其他语言辐射。树图表明第一层次上的分类即景颇语、白语及其他藏缅语的三大聚类。这在一定程度上，跟本氏设想相合。

(4) 藏缅语族语言有相当高的同质性，确实存在一个藏缅语族语言共同体。

(5) 羌语支语言为独立的一个聚类，可视为独立于藏的另一语支。黄布凡认为羌语支保留古藏文已没有的小舌音和 9 个辅音韵尾及其他复辅音韵尾，保留古老面貌，独具特点。树图的结果支持把羌独立为一个语支的观点。

#### 4. 计量分类跟传统分类的比较

最重要的差别：(1) 传统分类只重视语言差异的定性分类，而无法作语言间亲缘关系的距离的程度的量的分析。而计量分类可作亲缘关系程度的描述，并通过树枝长短来表示距离关系。(2) 虽然使用的材料和方法不同，但分类的结果却大致一样。(3) 可以显示语言的类簇和分级层次。(4) 传统分类都没有细分出藏缅语族各语言的较小的聚类和关系程度，我们的树图显示，羌语支内部羌语跟普米的距离最近，其次是嘉戎；而彝语支内部哈尼与彝关系最近，其次是纳西，最远的是缅。

#### 5. 关于白语的系属问题

白族是整个西南民族汉化最早、最深刻的一个民族。白语是否属于古汉语的一支？虽然白语中的汉语借词高达百分之 60 以上，包含了古今汉语各个时期的新、老借词，但是我们仍然认为白语属于藏缅语族语言。主要证据有：(1) 语言学内部证据，在 106 个核心同源词中，与羌语同源 40 个，与彝语同源 34 个，有些是与其他藏缅语族语言交叉同源，有些单读同汉字，但与藏缅语族同源部分，保留在合成词的语素中；如此统计则数量更多。有些同源词虽然可以追溯到与上古汉语同源，但同源词音韵变化形式与其他藏缅语族语言相同或相似，走的是同一条音变路子。例如：“太阳、地、树、根、叶子、狗、鱼、虱、鸟、尾巴、羽毛、犄角、脖子、头、眼睛、鼻子、耳朵、嘴、脚、牙齿、乳房、胸、腹、舌头、肉、血、心脏、皮、名字、人、男人、妇女、说、吃、知道、死、杀、燃烧、给、飞、来、站、大、小、冷、满、新、好、干、白、黄、绿、你、我、我们、那（近）、那（远）、胖、肥、什么、一、二、没有”。统计结果表明白语在 106 词中有超过 45 个词与藏缅语族语言同源。据此，白语与藏缅语族语言是发生学关系。如果将白语与汉语一对一的单独进行比较，很可能会认为白语中许多词是古汉语借词，但经过多语言的全面比较后，则会有新的看法。例如“新”貌似借词，实际上来源于古汉藏语。(2) 人文科学证据，即语言学外部证据。据多学科研究，云南的汉文化直到明代才占统治地位，成为“主流文化”，明代之前汉文化为“非主流文化”，即明之前以“土著文化”的“地方化”为主，明之后以“主流文化”的“汉化”为主，所以云南的少数民族包括白族操明清时形成的西南官话，而不同于闽、粤、客方言保留唐、宋时代的中原汉语音韵。白族自称“民家”，其认同形成于明代，明政府将云南居民划分为官户、民家、夷，其中官户、民家系汉族，民家为一般汉人，夷则为云南少数民族，虽然汉族仍称白族为“夷”，但白族精英坚持自称为“民家”，以示认同汉文化，以别于云南其他少数民族。白族汉化很早，白族政权南诏国与唐王朝关系极密切，但据日本发现的宋元写本《南诏图传》看，白族文化与汉文化仍有较大差异。元代白族模仿汉字创造白文。

宋元时，白族地主经济以及科举文化水平已与汉族地主无异。白族是自新石器时代生活在甘、青地区的氐羌的后裔，春秋、秦汉称“滇焚”，三国时称“叟人”，南北朝时称“樊人”，唐代“白蛮”，宋元“白人”，明清“民家”。自称“爨子”，“白子”，“白尼”，“白伙”等。唐宋以前仍实行“石棺葬”和“火葬”，文化类型同彝缅民族，而不同于汉族。

### 三 藏汉分离及藏缅语族内部各语言分离时代

计算出语言分裂的时间深度是历史语言学最具有挑战意义的工作之一。Swadesh 受到化学上发明的碳-14 年代测定法的启示，认为可用基本词汇作测量单位，他假设基本词在漫长的时间跨度上的取代比率，是相对稳定的，每 1000 年的保留比率为 80%，即  $r=0.8$ 。如果两种语言 La 和 Lb 分离的时间 Td 是在 2000 年前，那么，La 和 Lb 所保留的比率为  $0.8 \times 0.8$ ，即最初词汇的 0.64。如果我们用 P (La, Lb) 来代表两种语言中保留的基本词汇的比率，词源统计分析法的公式可写为：

$$P(La, Lb) = r^{2t}$$

$$t = \lg(P(La, Lb)) / (2 \lg r)$$

根据以上公式，如果  $t=2$ ，那么 P (La, Lb) 的值接近 41%。如果  $t=3$ ，那么 P (La, Lb) 的值接近 26%。如果已知 P (La, Lb) 的值，那么，就可计算出任何一对相关语言的 t 值，得出这两种语言分离的时间 Td。比如说，如果两种语言保留有 26% 的共有基本词汇，它们分离的时间则在 3000 年前。

下面是我们根据 12 支藏缅语族语言的共享的基本词汇比率，运用词源统计分析法，计算出藏缅语族内部各语言分离以及藏汉分离的时代。表中数字代表距今的相对年代，例如古汉语与藏语分离距今约 2755 年左右，而羌与藏分离距今约 1511 年左右。这仅仅是个大致的推算，并非绝对年代。

Tibeto-Burman Time distance Estimation (retention rate=80%)

	拉	阿	羌	普	嘉	景	缅	彝	哈	纳	白	土	OC
拉	0	317	1511	1729	1553	1970	1596	1639	1470	1511	2615	2483	2755
阿	317	0	1596	1684	1596	2022	1596	1596	1470	1639	2755	2615	2755
羌	1511	1596	0	659	836	1430	1275	1351	1275	1096	2184	2184	2548
普	1729	1684	659	0	805	1871	1096	1166	1062	1131	2548	2684	2548
嘉	1553	1596	836	805	0	1729	1166	1275	1275	1313	2548	2358	2615
景	1970	2022	1430	1871	1729	0	1511	2022	1775	1684	548	2615	2983
缅	1596	1596	1275	1096	1166	1511	0	1028	962	995	2548	2483	2684
彝	1639	1596	1351	1166	1275	2022	1028	0	867	1238	2755	2358	2828
哈	1470	1470	1275	1062	1275	1775	962	867	0	867	2548	2299	2684
纳	1511	1639	1096	1131	1313	1684	995	1238	867	0	2684	2615	2755
白	2615	2755	2184	2548	2548	2548	2548	2755	2548	2684	0	3424	2615
土	2483	2615	2184	2684	2358	2615	2483	2358	2299	2615	3424	0	3737
OC	2755	2755	2548	2548	2615	2983	2684	2828	2684	2755	2615	3737	0

假定两种语言失去联系后，在词的保留方面互不影响，那么，共享的基本词汇的比率必定是两种独立保留的基本词汇的产物。两种语言的核心词共享程度的比率不同，其分裂年代的时间深度是不一样的。学术界采用每 1000 年的保留比率为 80% 或 85% 这两种计算方法。我们认为采用每 1000 年的保留比率为 80% 的计算结果较符合人文学科的研究结论。当然，我们还要考虑到树图的每个分级层次的结合点，往往会因新的层次的接触产生新的词的共享度而跟上一层次的语言词的共享度有所差别，所以，注意排除各个层次的借词是很重要的。

人文学科表明，夏、商、周三代时期是前藏缅文化与前汉文化反复冲突接触的时期，以黄帝为代表的华夏集团以中原为中心建立夏王朝，这应是汉语族的直接先民；以炎帝为代表的东夷、“三苗”等南方民族取代夏，建立商王朝，但中心仍在中原。实现南方文化与北方文化的大融合。以西北地区甘、青文化圈为主的羌戎民族南下中原，建立周王朝，这是前藏缅文化与中原地区的华夏文化的大融合。所以，直至诗经时代，藏汉关系仍很密切。

藏汉分离年代与人文证据大致相同。而藏、羌分离时代跟唐代藏族“吐蕃”王国的建立的时代相合。

#### 四 藏缅语族与其他语族的关系——区域文化共同体接触融合的结果

##### 藏缅语族与汉语的关系

学术界一般公认藏缅语族与汉语之间存在着发生学关系。检查 106 核心词，藏缅语族与汉语的同源词远高于苗瑶语族与汉语同源词，以公分母算有 38 个，其中以与羌语同源最多，有 34 个。“我”藏缅语族与汉语古今音义几乎完全一样。

藏缅语族和苗瑶语族各自的区域特征明显，各自有一批不同的特征词。例如“我、牙齿、火、头、手、鸟、血、烟、死、杀、盐”等（因限于篇幅，相关词表省略）。

藏缅语族、苗瑶语族和汉语同源的，例如“月、干（干枯）、脖子、舌头、圆、满、来、角、眼睛、鼻子、耳朵、名字、路、给”等。

特别值得注意的是，有些词，南岛语跟藏缅语族、苗瑶语族、壮侗语族和汉语同源，例如“箭（射、弓、弩）、脚（骹）、风（飞廉）、舌头、月亮、米（稻）、路、豕（猪）、飞、树、皮肤、油脂、马”等。

而南岛语跟苗瑶语族、壮侗语族同源的核心词则更多，例如“死、杀、穿山甲、地（田）、盐（卤）、我、火（苗瑶外）”等。

#### 五 关于东亚语言与民族演化进程的几点思考

##### 1. 文化和文明，多样性与同一性

文明的交流和传播。古藏缅语很可能与斯达罗斯廷 (Sergai Starostin) 主张的“汉语—高加索”语系有关。

M. Rulén 将南亚语、苗瑶、台、南岛划归南方语系 Austric，而将汉藏语归入德纳—高加索语系 (Dene-Caucasian)。最近的考古研究证明小麦、青铜器、铁、车马、土坯建筑技术、先进的农耕灌溉技术为西方欧亚草原传入，经新疆—甘青—华北（龙山文化华夏区），铁是古藏缅人——周人（羌戎民族）用马车载到华夏族的中原地区的。家马、骑

兵、驯马技术由乌克兰草原传入。青铜器、车等突然出现在晚商时期，并非中原地区新石器时期以来的延续性发展，而是欧亚北方草原文明传播的结果。例如，“马”在藏缅语语言中形式变异很大，但显示同质性，所以中国学者往往构拟了一个非常复杂的上古汉语形式，具有词头、词尾和复杂的复辅音形式。实际上，“马”应是西亚借词。考古学证明，中国北方地区龙山——夏商时期没有发现家马的证据，马骨很少，与人没有关系，但中、西亚马骨、马车大量出现，5000年前已有驯马文化。家马和成熟的马车突然出现在晚商殷墟墓地。“马”的大量出现和传播不超过4000年。所以，其形式应该并不复杂。

## 2. 区域文化的交互作用圈理论

一方面，我们应认识到前汉藏语时期多种语言文化系统同时并存的格局，另一方面应认识到各系统间文化交流的重要性。例如，青铜器的制作中心在黄河中游地区，而矿产资源却来自长江中游地区，显示黄河文明与长江文明之间的交互作用。但最终黄河文明影响长江文明，藏缅语汉语影响苗瑶语壮侗语。这就是我们对不同文化区域的共同性和特殊性的认识。语言的分化与融合是语言发展中的两种不同形态，龙山——夏商时期主要为语言的融合时期。语言形成因素的多源性和复杂性。一些语言主要受一个来源的语言因素的影响，形成单一构成语言形态长期稳定发展的局面。另一方面，还有一些语言则同时或先后接受了几个不同来源的语言因素的影响，形成完全变异的复合语言形态。所以，我们主张语言的演化为多线演进模式，而非单一的演进模式。著名的考古学“重瓣花朵”理论即强调区域文化的交互作用，也强调文化中心区域向周边地区的辐射作用。这个理论很值得语言学借鉴。

3. 地域文化传统是语言分布格局的基础。环境对语言发展的限制、阻隔封闭作用和语言对环境的适应这是方言形成和变化的最主要因素。前藏缅民族——羌周原先是定居的农耕社会，但4000~5000年前气候突然变化，沙漠化，前藏缅民族——羌周由定居的农耕社会转变为畜牧业为主的狩猎游牧社会，因争夺生存资源，其与中原关系则更为密切，藏缅语和汉语发生大融合。

本文首次运用词源统计法作藏缅诸语言关系及其时间深度的研究，与传统分类相比较，虽然材料和方法不同，但分类的结果与传统分类大致相同，并能发现传统定性分类的缺陷，用数理方法测量出语言亲缘关系的相关“程度”。同时树图研究证明白语、羌语、土家应当分别独立为一个语支，这对解决长期以来的学术界因为传统的分类而产生的争论，提供一个数理分类的样品。以100核心词为对象，计算及分析汉藏诸语言的同质性和多样性、变异性，语言的同源和接触等各种复杂关系。计算出藏缅诸语言分裂的年代，其结果与人文科学研究结论大致相同。

说明：本课题得到香港政府“中国南方濒危语言研究”（项目批准号：9040237）和“汉语在亚洲语言中的地位”（CERG#9040781）及台湾中研院“中国各民族的起源及形成过程”的专项研究基金的资助。课题由王士元教授主持并指导。本文是本课题的阶段成果之一。课题组成员南开大学石锋教授、北京大学陈保亚教授、香港城市大学向倩先后为本课题的资料收集整理做了大量的工作。课题组成员香港城市大学柯津云小姐为本文的数据计算付出许多辛勤的劳动。藏缅语专家孙宏开教授、孙天心教授为本文的同源词识别提出很好的意见。在此一并致以深深的感谢。

## 参考文献

- 藏缅语语音和词汇编写组：《藏缅语语音和词汇》，中国社会科学出版社，1991年。
- 戴庆厦、傅爱兰、刘菊黄：《关于我国藏缅语族的系属分类问题》，《云南民族学院学报》1989年第3期。
- 邓晓华：《南方汉语中的古南岛语成分》，《民族语文》1994年第3期。
- 丁邦新、孙宏开主编：《汉藏语同源词研究》（一），广西民族出版社，2000年。
- 丁邦新：《丁邦新语言学论文集》，商务印书馆，1998年。
- 黄行：《苗瑶语方言亲疏关系的计量分析》，《民族语文》1999年第3期。
- 黄布凡主编：《藏缅语族语言词汇》，中央民族学院出版社，1992年。
- 瞿霭堂、劲松：《汉藏语言研究的理论和方法》，中国藏学出版社2000年。
- 梁敏、张均如：《壮侗语族概论》，中国社会科学出版社，1996年。
- 列·谢·瓦西里耶夫：《中国文明的起源问题》，文物出版社，1989年。
- 马学良主编：《汉藏语概论》，北京大学出版社，1991年。
- 倪大白：《汉藏语系语言的系属问题》，《中国语言学报》1995年第6期。
- 孙宏开、江获：《汉藏语言系属分类之争及其源流》，《当代语言学》1999年第2期。
- 孙宏开：《原始藏缅语构拟中的一些问题——以“马”为例》，《民族语文》1989年第2期。
- ：《论藏缅语族中的羌语支语言》，中研院语言学研究所（台北），《语言暨语言学》2001年第2卷第1期。
- 王辅世、毛宗武：《苗瑶语古音构拟》，中国社会科学出版社，1995年。
- 王士元、沈锺伟：《方言关系的计量表述》，《中国语文》1992年第2期。
- 王士元：《语言的探索——王士元语言学论文选译》，北京语言文化大学出版社，2000年。
- ：《王士元语言学论文集》，商务印书馆，2002年。
- 王文光：《中国南方民族史》，民族出版社，1999年。
- 吴安其：《汉藏语同源研究》，中央民族大学出版社，2001年。
- 邢公畹：《汉台语比较手册》，商务印书馆，1999年。
- 张光直：《考古学专题六讲》，文物出版社，1986年。
- 张增祺：《中国西南民族考古》，云南人民出版社，1990年。
- 郑锦全：《汉语方言亲疏关系的计量统计》，《中国语文》1988年第2期。
- 中国社会科学院、澳大利亚国立大学：《中国语言地图集》，朗文出版社，1988年。
- Benedict, P.K., *Austro-Thai: Language and Culture, with a Glossary of Roots*. New Haven, HRAF Press. 1975.
- Cavalli-Sforza, L.L. and W.S.-Wang. *Spatial distance and lexical replacement*. *Language*. 62. pp38-55. 1986.
- Dyen, Kruskal. and Black. *An Indo-European Classification: A Lexicostatistical Experiment*. *Transactions of the American philosophical Society*, Vol. 82, Part 5. 1992.
- Ellegard, A., *Statistical measurement of linguistic relationship*. *Language*. 35. 1959.
- George van Driem, *Tibeto-Burman replaces Indo-Chinese in the 1990s: Review of a decade of scholarship*, *Lingua* 111 79-102. 2002.
- Joseph H. Greenberg, *The Methods and Purposes of Linguistic Genetic Classification*, *Language And Linguistics* Volume 2 Number 2. 2001.
- Li, F.K., *Languages and dialects of China*. *Journal of Chinese Linguistics* 1.1. 1973.
- Matisoff, J.A., *Issues in the subgrouping of Tibeto-Burman in the post-Benedict Era. Paper on 30<sup>th</sup> ICSTL*.

- Merrit Ruhlen, *A Guide to the World's Languages: Vol. 1: Classification*, Stanford University Press. 1991.
- Qiao, Sanzheng and Wang, William S-Y., *Evaluating Phylogenetic Trees By Matrix Decomposition*. *Anthropological Sciences*. 106 (1), pp1-22. 1998.
- Society 96. 1952.
- Swadesh, M., *Lexico-statistic dating of prehistoric ethnic contacts, with special reference*, *Philosophical*
- Swadesh, M., *Time depths of American linguistic groupings*, *American Anthropologist* 56. 1955.
- Teeter, K.V., *Lexicostatistics and genetic relationship*. *Language* 39. 1963.
- Wang, William S-Y., *Glottochronology, Lexicostatistics, and other Numerical Methods*. *Encyclopedia Of Language and Linguistics*. Pergamon Press. 1994.
- : *A quantitative study of Zhuang-Dong Languages*, pp81-96, 1995. 余霭芹、远藤光晓编《桥本万太郎纪念——中国语言学论集》，内山书店。
- : *Linguistics diversity and language relationships*. In *New Horizons in Chinese Linguistics*. James Huang and Audrey Li, editors. Kluwer Academic Publishers. 1995.
- : *The Ancestry of the Chinese Language*. *Journal of Chinese Linguistics Monograph No.8*. 1995.
- 1997.
- William H. Baxter, *A Handbook of Old Chinese Phonology*, Moutou de Gruyter, Berlin, New York. 1992.

#### Abstract

Traditional studies of the grouping of Tibetan-Burman languages have focused only on analysis of the differences among languages. Few works have investigated the genetic distances between languages quantitatively. The research reported in this paper adopts the Lexicostatistics approach to give a numerical analysis of the genetic relationships among the Tibeto-Burman languages within China. The genetic distances between languages are indicated by the branch lengths of the constructed trees, which show the clustering of the languages and their separation. Also, the time depths of the languages in the family are estimated and a hypothesized process of the formation of the language family given.

(通信地址: 361005 厦门 厦门大学人类学博物馆  
香港城市大学电子工程学系)