

论邵语在中国台湾南岛语中的系属地位*

范志泉

[提要] 我国台湾高山族邵语的系属地位问题,其实质在于如何确定语言间纵向分化的亲缘关系以及横向传递的接触关系。邵语的南岛语同源词依旧保持在50%以上。邵语与布农语、西部平埔语和北部泰雅语群均有密切关系,或许使用这些语言的人群在历史上有过频繁的接触与交流。

[关键词] 中国台湾南岛语 邵语 系属地位 演化

高山族为居住在台湾省的少数民族,是我国统一多民族大家庭中不可分割的一部分。台湾高山族使用南岛语言(Austronesian,旧称“马来—波利尼西亚语”)。高山族邵人约有800多人^①,主要分布在我国台湾省南投县鱼池乡日月村,以及水里乡大平林聚落等地。邵语是邵人的母语,目前能流利使用邵语的邵人不多,且都是70、80岁以上的老人,邵语已是一种濒危语言(简史朗2018:4)。

关于邵语的系属地位,学界主要有几种不同的观点。戴恩·伊赛多(1984)认为邵语属于“排湾语群”,陈康(1992:3)和孙宏开等(2007:2111)沿用这一分类。费罗礼(Ferrell 1969:25)进一步将戴恩的排湾语群下分为2个分支,邵语被其归入“排湾语群第一分支”。李壬癸(Li 1990)将邵语与布农语归为一支,土田滋(1998)进一步将邵语与布农语归入“南台湾语群”。白乐思(Blust 1996, 1999)和何大安(1998)认为邵语与道卡斯、巴布扎(也称“猫雾椏”)、巴布拉、洪雅等语言组成“西部平埔语”,为原始南岛语的一个独立分支。李壬癸(1999:59, 2006)将道卡斯、巴布扎、巴布拉、洪雅等语言统称为“西部平埔语”,又将邵语与西部平埔语划为“西北支”,西北支语言与赛夏语、巴宰语等组成“北部台湾语群”的大西北语支^②。吴安其(2009:267)认为邵语与布农语、卑南语和阿美语构成“卑南—布农语支”,这一语支与邹语、卡那卡那富语、拉阿鲁阿语、鲁凯语和排湾语构成的“邹—鲁凯语支”一起组成“邹

* 本文系国家社科基金青年项目“台湾南岛语的语音对应规律数据库建设与研究(22CYY047)”的阶段成果。论文曾在“《民族语文》第十五届学术研讨会”(昆明2022.9.30-10.1)上宣读,承蒙与会专家悉心指导。匿名审稿专家也对本文提出了宝贵的修改意见。谨此一并致谢。文中如有错误,概由笔者负责。本文语料来源:泰雅语、赛德克语、太鲁阁语、邹语、卡那卡那富语、沙阿鲁阿语、鲁凯语、布农语、排湾语、卑南语、邵语、赛夏语、巴宰语、噶玛兰语、撒奇莱雅语、阿美语、雅美语引自台湾省民族事务主管部门发布的《族语学习词表》(<http://ilrdc.tw/research/athousand/explain.php>);巴宰语引自李壬癸、土田滋(2001);西部平埔语及原始西部平埔语引自李壬癸(Li [2003]2004b);原始泰雅语引自郭育贤(2022);原始布农语引自柴田海(2020);原始南岛语构拟引自白乐思(Blust Robert)和特鲁塞尔(Stephen Trussel)主编的The Austronesian Comparative Dictionary(web edition),网址为:<https://www.trussel2.com/ACD>。

^① 数据来自:<http://www.apc.gov.tw/portal/chart/chartindex.htm>。

^② 本文沿用李壬癸(1999, 2006)的分类。

—卑南语群”。沙加尔 (Sagart 2014) 认为邵语与泰雅语、Enemish 语支属于 Limaish 分支。

专题研究邵语系属地位的主要论著有白乐思 (Blust 1996) 和李壬癸 (1999)。在分类标准上, 费罗礼 (Ferrell 1969)、白乐思 (Blust 1999)、何大安 (1998)、土田滋 (1998)、李壬癸 (1999)、吴安其 (2009) 等均采用历史语言学的标准, 主要依据语音特征和演变规则, 即利用辅音、元音等的共同创新 (Shared Innovation) 特征, 是一种定性研究。沙加尔 (Sagart 2014) 则是基于词汇特征。也有学者采用词汇统计的定量方法, 如戴恩·伊赛多 (1984) 和李壬癸 (Li 1990)。

总的来说, 前人及时贤在邵语系属问题的探讨方面取得了诸多研究成果。但是囿于不同的分类理论、方法及材料, 邵语系属地位研究依旧存在诸多分歧与争议。核心问题在于如何断定邵语与布农语、西部平埔语等语言之间的关系。本文先从语音演变、核心词比较两个方面将邵语与其他台湾南岛语进行比较, 在此基础上, 结合系统发生学模型理论与方法, 对邵语的系属地位进行探讨。

一 邵语的音位及其语音演变

(一) 邵语的音系与书写系统

李方桂先生是最早系统研究邵语的语言学者。1956年, 他与陈奇禄、唐美君共同发表《邵语记略》一文。后续学者如李壬癸、白乐思等大都沿袭《邵语记略》的研究框架。根据现有研究, 邵语的音位系统 (Li 1976; Blust 1998, 2003:18; 黄美金 2000:43; 简史朗 2018:11) 共有 3 个元音, 即 a、i、u^①; 21 个辅音, 即 p、^ʔb、t、^ʔd、k、q、ʔ、 ϕ 、 θ 、 δ 、s、 ζ ^②、h、m、n、 η 、 ɬ 、l、r、w^③、j。

2005年12月, 台湾省民族事务主管部门联合教育部门共同颁布台湾南岛语言书写系统。邵语的书写系统中, 字母 b 表示前喉塞唇音 [^ʔb], d 表示前喉塞舌尖音 [^ʔd], ' 表示喉塞音 [ʔ], f 表示双唇清擦音 [ϕ], th 表示齿间清擦音 [θ], z 表示齿间浊擦音 [δ], sh 表示硬腭清擦音 [ζ], ng 表示舌根鼻音 [η], lh 表示舌尖清边音 [ɬ], y 表示半元音 [j]。

(二) 邵语的语音演变

邵语在语音演变中, 除仍保留原始南岛语 (Proto-Austronesian, PAN)^④ 的部分语音特征外, 也产生了一些独有或者与其他台湾南岛语相同的演变特点。

^① 《邵语记略》(李方桂等 1956) 中没有元音 [ə]。李壬癸 ([1992]2004b:1463) 表示元音 [ə] 是 /i/ 的音位变体, 即前高元音 [i] 与辅音 [ʔ、h、r] 相邻时舌位降低, 读为前半高元音 [e], 而与小舌音 [q] 相邻时中间有过渡的央元音 [ə]。

^② 在有关邵语的研究文献和教材中, [j] 常用来表示硬腭清擦音 [ζ]。为方便跨语言比较及行文统一, 本文统一使用 [ζ] 表示硬腭清擦音。另外, 还有用字母 c 表示齿间清擦音 [θ] (黄美金 2000:43; Blust 2003:18), g 表示舌根鼻音 [η] (Blust 2003:18), S 表示硬腭清擦音 [ζ] (李壬癸 1992[2004b]; 黄美金 2000:43), L 表示舌尖清边音 [ɬ] (黄美金 2000:43)。为避免混淆及行文统一, 有关语料本文均作统一处理。

^③ 邵语的半元音 [w] 位于元音之前时, 则读为双唇浊擦音 [β] (简史朗 2018:23)。

^④ 据白乐思 (Blust 2022:1178), 原始南岛语共有 4 个元音, 25 个辅音。具体为: *i、*u、*e、*a; *p、*b、*t、*d、*D、*k、*g、*j、*q、*S、*s、*h、*C、*c、*z、*m、*n、*ñ、*ŋ、*l、*N、*R、*r、*w、*y。其中, 字母 D 表示卷舌浊塞音 [d], 只出现在词尾, 仅有 3 个例词。j 表示腭化舌根浊塞音 [g], 只出现在词中和词尾。c 表示硬腭清塞擦音 [tc], 仅有 4 个例词。z 表示硬腭浊塞擦音 [dz]。ñ 表示硬腭鼻音 [ɲ], 仅有 6 个例词 (根)。c、z、ñ 只出现在词首和词中。r 表示闪音 [r]。x 仅有 6 个例词, 但在总的音位表里没有该音位。

表1 原始南岛语在邵语中的语音演变

PAN	邵语	例词 ^①
*e[ə]	i, u, Ø	*aken “我” > y-akin “我(宾格)”; *nipen “牙齿” > nipin; *telu “三” > turu; *tenun “编织” > tunún; *baqeRuh “新的” > φαqtu; *bineSiq “(播)种” > φα-φινçiq
*s[ç]	t	*esa/isa “一” > tata; *susu “乳” > tutu; *asu “狗” > atu; *Caŋis “哭” > θ-m-anit
*b[b]	φ	*batu “石头” > φatu; *tebuS “甘蔗” > tuφuç; *babuy “野猪” > φαφuy; *qaleb “膝盖” > qaruf
*C[ts]	θ	*Cau “人” > θau; *Cawa “笑” > ma-θaθawa; *kuCu “虱子” > kuθu; *maCa “眼睛” > maθa; *sakeC “鹿, 山羌” > takiθ; *RameC “根” > †amiθ
*j[g]	ð	*pajiS “敌人” > paðiç; *pajay “稻米” > paðay; *kuja “如何” > mi-a-kuða; *bajaq “知道; 会; 懂” > ma-φαðaç; *pijax “多少” > lapiða
*N[l̥]	ð	*ma-puNi “白色” > ma-puði-puði; *aN-aNak “小孩” > að-aðak; *daNum “水” > saðum; *tiaN “肚子” > tiað; *bulaN “月亮” > φurað
*d[d]	s	*tuduq “(水)滴” > tusuq-in; *qudem “黑; 暗” > ma-qusu-qusum; *kudkud “脚” > kuskus; *likud “背” > rikus; *daqis “脸” > çaqiç
*z[ɖ]	s	*quzaN “雨” > qusað-in; *zalan “路” > saran; *t<um>uzuq “指向” > tmusuq; *zaRum “针” > †a†um
*S[s]	ç	*duSa “二” > tuça; *bineSiq “(播)种” > φα-φινçiq; *bukeS “(毛)发” > φukiç
*ŋ[ŋ]	n	*daŋdaŋ “烤; 煮; 温暖” > pa-sansan “烤(地瓜)”; *laŋaw “苍蝇” > ranaw; *mujiŋ “鼻子” > muðin; *luCuŋ “猴子” > ruθun
*R[r̥]	†	*baqeRuh “新的” > φαqtu; *kiRim “寻找” > k-m-iŋim; *qasiRa “盐” > qti†a; *siNaR “太阳” > ti†að; *gaRaŋ “螃蟹” > ka†an
*l[l̥]	r	*ala “拿” > m-ara; *bali “风” > φari; *bulaN “月亮” > φurað; *dilaq “舌头; 舔” > s-m-iraq; *bual “温泉” > (matata wa) φuar

根据白乐思 (Blust 1996)、李壬癸 (1999:91-97) 等研究, *b>φ、*C>θ、*j 和 *N>ð、*d>s、*S>ç、*R>† 等均为邵语独有的音变特征。而 *e>i 或 u、*s>t、*z>s、*ŋ>n、*l>r 等则是与其他台湾南岛语共有的音变特征 (见表 1)。具体如下:

1. 原始南岛语的 *e 演变为邵语的 i、u, 即 PAN *e>i 或 u。这也是泰雅语、布农语的音变特点,但在各语言中的分布不规则。如: PAN *qaRem “穿山甲”, 邵 qatun, 原始泰雅 *qagum, 原始布农 *qalum; PAN *nipen “牙齿”, 邵 nipin, 原始泰雅 *gipun, 原始布农 *nipun; PAN *tenun “编织”, 邵 tunún, 泰雅 t-m-inun, 布农 ma-teindun。

2. 原始南岛语的 *s 演变为邵语的 t, 即 PAN *s>t。这也是排湾语、巴宰语、西部平埔语的音变特点。如: PAN *esa/isa “一”, 邵 tata, 排湾 ita, 道卡斯 ta-anu, 巴布拉 ta-nu, 巴布

^① 文中例词为笔者结合前人研究, 从使用的语料中收集所得。

扎 na-ta, 洪雅 itta/m-eta/a-ta; PAN *susu “乳”, 邵 tutu, 排湾 tutu, 巴布扎 chi-tu, 洪雅 tu-tu; PAN *asu “狗”, 邵 atu, 排湾 vatu, 巴布拉 hatu, 巴布扎 atu, 洪雅 atu。

不过, 只有当 *s 位于词末时, 在巴宰语中才会演变为 t。如: PAN *Caŋis “哭”, 巴宰 m-angit; PAN *bales “回答”, 巴宰 pa-baret; PAN *bukas “解除、释放”, 巴宰 bukat “解除陷阱、卸下陷阱”。但当 *s 位于词首或者词中时, 在巴宰语中则演变为 z。如: PAN *sepsep “吸”, 巴宰 mu-zezeb; PAN *busuR “(打猎用的)箭”, 巴宰 buzux; PAN *basuq “洗”, 巴宰 mu-bazu; PAN *asu “狗”, 巴宰 wazu。而在西部平埔诸语中, 当 *s 位于词末时, 则倾向于丢失, 即 *-s>Ø(Li [1985]2004a:823)。如: PAN *Caŋis “哭”, 道卡斯 s-am-ani, 巴布扎 s-um-ani, 巴布拉 s-am-ani, 洪雅 s-um-ni; PAN *ma-hemis “甜的”, 道卡斯 mme, 巴布拉 mame, 洪雅 mami。

3. 原始南岛语的 *z 演变为邵语的 s, 即 PAN *z>s。台湾省中部的拉阿鲁阿语也有这种音变。如: PAN *quzaN “雨”, 邵 qusað-in, 拉阿鲁阿 usaŋi; PAN *zalan “路”, 邵 saran, 拉阿鲁阿 saraŋa。

4. 原始南岛语的 *ŋ 演变为邵语的 n, 即 PAN *ŋ>n。西部平埔语也有这种演变。如: PAN *Caŋis “哭”, 邵 θ-m-anit, 道卡斯 s-am-ani, 巴布扎 s-um-ani, 巴布拉 s-am-ani, 洪雅 s-um-ni; PAN *qaNuaŋ “牛”, 邵 qnuan, 巴布扎 luan, 巴布拉 loan, 洪雅 loan; PAN *Caliŋa “耳朵”, 邵 ɬarina, 道卡斯 sarina, 巴布扎 harina, 巴布拉 sarina, 洪雅 sarina。其他台湾南岛语没有这个音变特征, 因此该特征常被作为邵语与西部平埔语具有亲密关系的一个证据(李壬癸 1999:96)。

5. 原始南岛语的 *l 演变为邵语的 r, 即 PAN *l>r。巴宰语、西拉雅语、原始泰雅语等也有此类音变。如: PAN *Cemel “草”, 邵 ɬmir, 巴宰 semer; PAN *bali “风”, 邵 φari, 巴宰 bari, 西拉雅 vare; PAN *kalih “挖”, 邵 k-m-ari, 巴宰 sa-a-kari “(挖土的)棍子”, 西拉雅 kari; PAN *bulaN “月亮”, 邵 φurað, 原始泰雅 *bural; PAN *Cali “芋头”, 邵 ɬari, 原始泰雅 *cari?。西拉雅语还存在 PAN *l>l。如: PAN *dilaq “舌头; 舔”, 邵 s-m-iraq, 西拉雅 da-dila。

此外, 有学者认为邵语前喉塞唇音¹b、前喉塞舌尖音²d、喉塞音³?、喉擦音 h 及浊边音 l 等音位并非对原始南岛语的反映, 而是借自布农语(参 Li 2013)。

如前文所述, 历史语言学的系属分类方法大多采用共同创新的定性标准, 凭借个人学术经验择取若干条语言特征, 并以此类特征为依据对语言进行聚类分析。然而不同学者采用不同的共同创新特征, 其分类结果往往也有差异。尤其是当分类特征数目较多且不同特征所支持的分类结果有差异时, 人工的判断与择对往往力不从心。

二 基于词源统计的邵语演化模型的构建与分析

除定性分类研究外, 也有学者采用词汇统计的定量方法。定性的共同创新法能提供的特征信息十分有限, 缺乏具体的量化指标, 难以准确测量语言间亲缘关系的远近程度。因此, 基于定量分析的数理统计法便应运而生。相对来说, 定量分类特别是基于分类特征的数理统计, 要求分类特征具有确定性和可统计性以及可操作性, 在一定程度上能尽量排除人为主观因素的干扰。目前主要有基于特征统计和基于距离统计两种定量分析。在基于特征统计的定量分析方法中, 语言的语音、词汇等不同属性的特征都可以被赋予不同的“特征值”, 可以说

这是历史比较方法在量化问题上的延伸与扩展（范志泉 2021:115）。与特征统计法不同，距离统计法采用的是以分类对象整体之间的亲缘关系远近程度进行分析。距离统计法主要有两种计算方法：一是通过计算同源词百分比，生成同源词百分比相似距离，后再根据相关算法将相似距离转换成遗传距离；二是通过计算语音字符串之间的编辑距离（范志泉 2021:117）。

此外，随着语言接触研究的深入，语言学者发现生物学中尝试用来反映物种之间的移位、水平基因转移、杂交等横向事件的系统发生学网络模型，同样也能适用于语言间的演化研究（范志泉 2021:118-119）。相对于树形图，网状结构不仅包含了分类对象的纵向分化信息，同时也能显示其横向传递信息（张梦翰、韦远诚 2019）。本文将在词源统计分析的基础上，采用生物学的系统发生学理论与方法，生成邵语分类图。并通过邵语的语音演变、核心词分布，结合模型图尤其是网络模型图中盒状结构的冲突信号，检验邵语在台湾南岛语言中的横向交流度量，尝试借助合适的分类模型来阐释其系属地位及演化历史。

（一）构建邵语演化模型的材料

由于凯达格兰语、西拉雅语、龟仑语、猴猴语、道卡斯语、巴布扎语、巴布拉语、洪雅语等 8 种语言已消失，因此未纳入本文讨论范围。本文以台湾省民族事务主管部门公布的《族语学习词表》（<http://ilrdc.tw/research/athousand/explain.php>）和李壬癸、土田滋（2001）收录的 17 种台湾南岛语为基础材料，即泰雅语、赛德克语、太鲁阁语、邹语、卡那卡那富语、沙阿鲁阿语、鲁凯语、布农语、排湾语、卑南语、邵语、赛夏语、巴宰语、噶玛兰语、撒奇莱雅语、阿美语、雅美语（又称“达悟语”）。在此基础上，我们整理出 17 种台湾南岛语的斯瓦迪士 100 核心词，通过识别同源词，编制同源词表；从而计算每对语言之间的同源百分比，最终生成同源百分比的相似矩阵。相似矩阵反映的是同源词的数量多少，而数理统计生成的演化模型是通过分支线段的长度来反映语言之间的距离远近。所以，需要将相似矩阵通过负对数公式 $d = -\log s$ （ d 代表距离； s 代表同源词相似数字）转换成能反映语言间亲缘关系的距离矩阵。距离矩阵将作为生成语言演化网络模型的基本输入数据（范志泉 2021:142）。

（二）邵语的演化网络模型生成与分析

基于以上材料与方法，本文分别采用 Neighbor-Net 法（Bryant & Moulton 2004）和 Neighbor-Joining 法（Saitou & Nei 1987），以雅美语为外群，生成邵语的演化网络图（图 1-A）和演化树形图（图 1-B）。

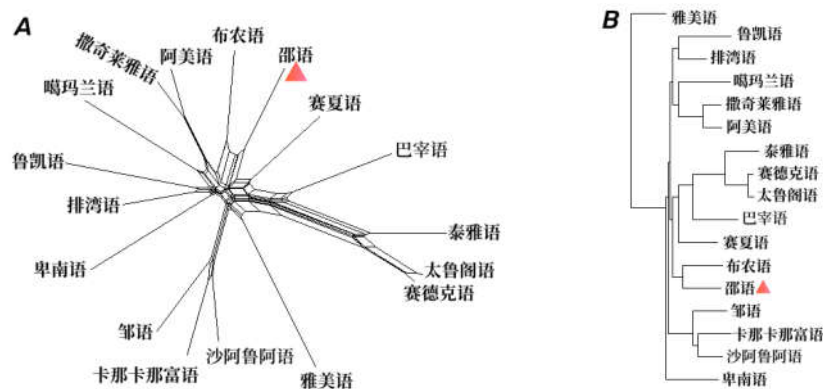


图 1 邵语的演化网络图和树形图

位于兰屿岛上的雅美语是台湾省唯一一种属于“马来—波利尼西亚语群”的南岛语，这一点争议不大。图示说明：网络模型图中有许多网状结构，网状结构反映出一定的盒状信号，而盒状信号可能是语言间横向接触交流的表现特征^①。此外，为了解语言接触交流程度，在网络模型图中，还可以用 Delta 值来衡量两两语言之间的交流度量 (Holland et al. 2002)。Delta 值的范围一般在 0 到 1 之间，数值越接近于 1 则说明该网状结构中的横向交流事件多 (Gray et al. 2010)。目前已有不少学者尝试采用计算 Delta 值对印欧语、波利尼西亚语、藏缅语、壮侗语台语支等语言的网络模型图进行分析 (见范志泉 2021:118)。本文根据 Delta 值计算公式得出了图 1-A 的 Delta 值为 0.394，且在图 1-A 网状结构中各个语言点的 Delta 值为：阿美 (=0.449) > 邵 (=0.446) > 噶玛兰 (=0.438) > 雅美 (=0.431) > 鲁凯 (=0.417) > 布农 (=0.411) > 卑南 (=0.406) > 排湾 (=0.404) > 巴宰 (=0.392) > 赛夏 (=0.387) > 沙阿鲁阿 (=0.377) > 赛德克 (=0.360) > 邹 (=0.355) > 卡那卡那富 (=0.327) > 泰雅 (=0.315)。总的来说，Delta 值由南向北逐步递减，显示台湾南部南岛语言较北部南岛语言而言，彼此之间应该有较多的接触交流。而就单个语言点的 Delta 值来看，邵语 (=0.446) 在整个网络模型图中的 Delta 值也较高。

通过比较后可知，由 Neighbor-Net 法生成的演化网络图 (图 1-A) 所显示的语言聚类与通过 Neighbor-Joining 法生成的演化树形图 (图 1-B) 所得到的聚类呈现相对一致的分类结果。即：图 1 中 A、B 两幅图都较为一致地显示，邵语不仅与布农语聚类在一起，同时还与泰雅、赛德克、巴宰、赛夏等台湾北部南岛语也倾向聚类于一簇。与演化网络图比较，演化树形图确实更为简单明了。它不仅能直观表现邵语的纵向演化，也能清晰地表现邵语与其他台湾南岛语之间的亲疏远近关系。不过树形图终究只是一种理想化的描述语言演化历史进程的理论模型，它容易忽略语言之间除了纵向分化外的横向接触与交流历史 (徐通锵 1991:44)。演化网络图或可弥补这一缺陷，相较于传统的树形结构，演化网络的网状结构不仅能描述物种之间的纵向传递，还能兼容横向传递。

此外，本文图 1 所显示出的分类结果与李壬癸 (2006) 的分类结果基本一致，不同之处在于邵语与布农语的关系。为进一步了解邵语与布农语的关系，本文对邵语斯瓦迪士 100 核心词进行了考察。

(三) 邵语 100 核心词的分布

通过对邵语斯瓦迪士 100 核心词的考察可以看出，在 100 核心词中，邵语保留的原始南岛语核心词有 51 个；邵语独有且与其他台湾南岛语不同源的核心词有 34 个；与西部平埔语相似或相近，但与其他台湾南岛语不同的核心词有 3 个 (Li [2003]2004b:945)，即“冷的”“名字”“米”。“冷的”，邵 *ma-tiθi*；原始西部平埔语 **ma-siCi*；巴宰 *lamik*，布农 *ma-kaðav*。“名字”，邵 *łanað*；原始西部平埔语 **RanaN*；阿美 *ŋaŋan*，布农 *ŋaan*。“米”，邵 *ðaçuq*；原始西部平埔语 **NaSuq*；泰雅 *beras*，布农 *tilas*。

再者，邵语与布农语及其方言相似或相近，与其他台湾南岛语不同的关系词有 12 个：“全部、大的、女性、树、皮、(动物)角、尾巴、脖子、坐、烟、山、黄色的”。除“大的、山”为本文补充外，其余 10 个关系词，李壬癸 (Li 2013) 认为是源于布农语的借词^②。试比较：

^① 盒状信号大并不意味着语言有过密切的接触历史，也有可能是由语言的平行演化造成的。

^② 李壬癸 (Li 2013) 共收录邵语 193 个布农语借词。

表2 邵语与布农语相似关系词比较

词义	邵语	布农语	其他台湾南岛语
全部	aðʔað	aðʔað 邵群方言	巴宰 dadua; 赛夏 saβsaβæh; 排湾 savid; 卑南 peniya
大的	ma-raʔin	ma-daiŋ	巴宰 mataru; 赛夏 soβæ:æh; 噶玛兰 Raya
女性	ʔbinanauʔað	binanauʔað 卓群·卡群·丹群· 密群方言	巴宰 mamais; 赛夏 minkorinŋ; 阿美 fafahiyan
树	riʔbuç	libus 卓群·卡群·丹群·密群· 邵群方言	巴宰 kahuy; 赛夏 kæhoy; 阿美 kilan; 噶玛兰 paRin
皮	çapa	sapa “毛皮” 邵群	巴宰 rapay; 赛夏 βaŋəS; 噶玛兰 lubuŋ
(动物)角	waqa	vaqa 卓群·卡群·丹群·密群方言	巴宰 uxuŋ; 赛夏 kahʔoŋ; 阿美 puhuŋ
尾巴	wa-wiçwiç	visvis 卓群·卡群方言	巴宰 dulut; 赛夏 kiko:; 排湾 iku
脖子	ʔbuqtur “颈颌; 脖子”	butqul “喉结; 喉咙” 密群方言	巴宰 hahur; 赛夏 warəŋ; 卑南 takulaw; 排湾 liqu
坐	mi-luŋqu	ma-luŋqu 丹群·密群	巴宰 mituku; 赛夏 minaSa.əŋ; 原始泰雅 *matamaʔ
烟	qumʔbu	qumbu 卓群	巴宰 busubus; 赛夏 kasβol; 排湾 tsevuʔ
山	huʔdun	ludun	巴宰 binayu; 赛夏 koko:ol; 排湾 gadu; 阿美 lutuk
黄色的	ma-ʔbulaw “成熟; 黄的”	ma-bulav 卓群·卡群方言	巴宰 tabarak; 赛夏 haβiyalan; 阿美 kulawlaway

数百年来, 由于邵人与布农人接触密切, 通婚频繁, 邵语深受布农语的影响 (李壬癸 1999:97)。因此, 李壬癸 ([2001]2004a:897-898) 认为上述关系词是源自布农语的借词。例如:

“全部”, 布农语 amin “全部; 都; 所有”, 排湾语 pu-amin “结束”。原始南岛语构拟为 PAN *amin “全部; 完成”。很明显, 布农语来自原始南岛语。不过, 布农语表示“全部; 都”义的还有 aðaʔaða。参考: 邵语 aðʔað “全部; 都”。根据李壬癸 (Li 1988) 和白乐思 (Blust 2003:73-79) 的研究, 布农语的齿间浊擦音 ð < PAN *y[j], 如: PAN *pajay “米” > 布农 paað; PAN *qaCay “肝” > 布农 qatað; PAN *ma-aCay “死” > 布农 matað。而邵语的齿间浊擦音 ð 则来自 PAN *j[gʲ]。如: PAN *pajay “米” > 邵 paðay; PAN *bunaj “沙” > 邵 ʔbunað; PAN *kuja “如何” > 邵 mi-a-kuða。很明显, 邵语的 aðʔað “全部; 都”借自布农语。

“树”, PAN *kaSiw “树; 木”, 排湾语 kasiv, 卑南语 kawī, 原始泰雅语 *kahuy, 邵语 kawī “木柴; 木头; 树干”。显然, 邵语来自原始南岛语。邵语还有 riʔbuç “树; 大树林, 山林”。参考: 布农语 libus “森林; 山林”。据李壬癸 (Li 1988) 和白乐思 (Blust 2003:73-79), 邵语的颤音来自 PAN *l, 而 PAN *l 在布农语中丢失。如: PAN *Caliŋa “耳朵” > 布农 tanja; PAN *lima “手; 五” > 布农 ima; PAN *bulaN “月亮” > 布农 buan。显然, 邵语既保留来自原始南岛语 *kaSiw 的 kawī “木柴; 木头; 树干”, 也有借自布农语 libus 的 riʔbuç “树; 大树

林，山林”。

“黄色的”，布农语_{卓群·卡群} ma-bulav。原始南岛语构拟为 PAN *bulaw “金色”。原始南岛语双唇浊塞音 *b 演变为布农语的 b，邵语的 φ。如：PAN *batu “石头” > 邵 φatu，布农_{邵群} batu；PAN *tebuS “甘蔗” > 邵 tuφuç，布农 sibus；PAN *babuy “野猪” > 邵 φαφuy，布农 babu。显然，邵语 ma-²bulaw “成熟；黄的” 应是借自布农语。

此外，经比较，邵语与布农语的同源词比率最高，达到 44.2%，高出其他台湾南岛语 6.75%-21.37% 左右。这显示与其他台湾南岛语相比，邵语与布农语更倾向于聚在一起。这一结果也与戴恩·伊赛多（1984）和李壬癸（Li 1990）的研究结论相似。

不过需要注意的是，用作比较的基本词词目数量越多，就越难排除语言之间相互借用的成分。斯瓦迪士最初提出 200 词，是将其作为人类语言最稳定的核心词，然而在实际研究中，他发现所谓 200 词也可以被借用。后续学者同样也验证了这个说法。由此，斯瓦迪士又从 200 词中筛选出了他认为最稳定的 100 核心词。陈保亚（1996:185）指出：“如果我们从无界有阶性的角度看，尽管 100 词也可以借用，但比起 200 词的借用率要小得多。”由此，他将斯瓦迪士从 200 词中挑选出来的 100 词，称为“第一阶词或高阶词”，而余下的词再经过适当合并与调整，称为“第二阶词或低阶词”。经过实践，他认为前 100 高阶词比余下的 100 低阶词更稳定，更难以借用。汪锋、王士元（Wang & Wang 2004）也得出类似的结论。邓晓华、王士元（2009:25）认为，斯瓦迪士 100 基本词“是经过许多语言检验过的一组能反映语言纵向分化的词表，这组词不容易受到借用的影响，词义有较高的稳定性。所以做统计时可以保证因借用而造成统计的干扰很小。”

三 结语与思考

邵语的系属地位是学界一个颇具争议的论题。邵语系属地位分歧的实质在于如何确定语言间纵向分化的亲缘关系与横向传递的接触关系。因此，本文从邵语的语音演变、词源统计等方面入手，基于生物种系发生学理论与方法，构建邵语的演化网络模型。

相较于演化树形图（图 1-B）只关心纵向分化的聚类问题，演化网络图（图 1-A）还显示邵语不仅与布农语聚类在一起，同时也与泰雅语群等明显倾向聚类于一簇。此外，邵语与泰雅语群、布农语两者之间的盒状结构都比较大，网状结构的边长也较长。这说明邵语与布农、泰雅等语言的特征所蕴含的冲突信号较多，反映了邵语不仅与布农语，而且还与泰雅、赛德克、巴宰等语言有着一定的接触与交流。历史上，邵语与西部平埔诸语或许有过较为密切的关系，两者不仅有相同的音变规律，而且也有少量同源词。但是由于西部平埔语消失殆尽，目前已有的证据不足以说明它们是否同出一源。另一方面，在历史的发展进程中，由于邵人与布农人之间的密切接触与频繁通婚，导致邵语在一定程度上深受布农语的影响。

总的来说，前人及本文的分类结果暗示着邵语在我国台湾北部、西部地区的纵向分化和接触交流并不仅仅是单一事件。相对于树形图而言，本文的研究结果再次显示网络模型确实能够同时描述分化和接触两种演化，可为阐释有争议的语言系属问题提供一定的研究视角和理论依据，不过其适用性还需进一步检验与实践。

参考文献

- 白乐思. 2022. 《南岛语言》，李壬癸等译，新北：联经出版事业股份有限公司。
- 柴田海. 2020. 《原始布农语音韵系统与词汇的重建》，台湾清华大学硕士学位论文。
- 陈保亚. 1996. 《论语言接触与语言联盟——汉越（侗台）语源关系的解释》，北京：语文出版社。
- 陈康. 1992. 《台湾高山族语言》，北京：中央民族学院出版社。
- 戴恩·伊赛多. 1984. 《台湾高山族语言》，陈康译，载中国社会科学院民族研究所语言室编《民族语文研究情报资料集》（第3集）第77-106页。
- 邓晓华、王士元. 2009. 《中国的语言及方言的分类》，北京：中华书局。
- 范志泉. 2021. 《台湾高山族语言的谱系分类研究》，厦门大学博士学位论文。
- 郭育贤. 2022. 《泰雅语音韵系统、重建与分群》，台湾清华大学博士学位论文。
- 何大安. 1998. 《台湾南岛语的语言关系》，《汉学研究》第2期。
- 黄美金. 2000. 《邵语参考语法》，台北：远流出版事业股份有限公司。
- 简史朗. 2018. 《邵语语法概论》（第二版），新北：“原住民族”委员会。
- 李方桂、陈奇禄、唐美君. 1956. 《邵语记略》，《台湾大学考古人类学刊》第7期。
- 李壬癸. 1999. 《台湾原住民史：语言篇》，南投：台湾省文献委员会。
- 李壬癸. [2001]2004a. 《邵族的地位——兼评白乐思（Blust 1996）的邵族地位说》，载李壬癸著《台湾南岛语言论文选集》（第二册）第891-906页，台北：中研院语言学研究所。
- 李壬癸. [1992]2004b. 《中国语文台湾南岛语言的语音符号系统》，载李壬癸著《台湾南岛语言论文选集》（第二册）第1413-1471页，台北：中研院语言学研究所。
- 李壬癸. 2006. 《从本土到国际：台湾南岛语言研究的契机》，载刘翠溶主编《四分溪论学集：庆祝李远哲先生七十寿辰》第727-748页，台北：允晨文化实业股份有限公司。
- 李壬癸、土田滋. 2001. 《巴宰语词典》，台北：中研院语言学研究所筹备处。
- 孙宏开、胡增益、黄行主编. 2007. 《中国的语言》，北京：商务印书馆。
- 吴安其. 2009. 《南岛语分类研究》，北京：商务印书馆。
- 徐通锵. 1991. 《历史语言学》，北京：商务印书馆。
- 张梦翰、韦远诚. 2019. 《台语支系统发生模式及横向交流度量》，载徐丹、傅京起主编《语言接触与语言变异》第376-391页，北京：商务印书馆。
- 土田滋. 1998. 「言語研究への誘い」. 日本順益台湾原住民研究会編『台湾原住民研究への招待』第137-139ページ. 東京：風響社。
- Blust, Robert. 1996. Some remarks on the linguistic position of Thao. *Oceanic Linguistics* 35(2): 272-294.
- Blust, Robert. 1998. Some problems in Thao phonology. In Shuanfan Huang (ed.), *Selected Papers from the Second International Symposium of Languages in Taiwan*, pp. 1-20. Taipei: Crane.
- Blust, Robert. 1999. Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In Elizabeth Zeitoun & Paul Jen-kuei Li (eds.), *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, pp. 31-94. Taipei: Institute of Linguistics, Academia Sinica.
- Blust, Robert. 2003. *Thao Dictionary*. Taipei: Institute of Linguistics (Preparatory Office), Academic Sinica.
- Bryant, David & Vincent Moulton. 2004. Neighbor-net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution* 21(2): 255-265.
- Ferrell, Raleigh. 1969. *Taiwan Aboriginal Groups: Problems in Cultural and Linguistic Classification*. Taipei:

- Institute of Ethnology, Academia Sinica.
- Gray, Russell D., David Bryant & Simon J. Greenhill. 2010. On the shape and fabric of human history. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences* 365(1559): 3923-3933.
- Holland, Barbara R., Katharina T. Huber, Andreas Dress, et al. 2002. δ Plots: A tool for analyzing phylogenetic distance data. *Molecular Biology and Evolution* 19(12): 2051-2059.
- Li, Paul Jen-kuei. 1976. Thao Phonology. *Bulletin of the Institute of History and Philology* 47(2): 219-244.
- Li, Paul Jen-kuei. 1988. A comparative study of Bunun dialects. *Bulletin of the Institute of History and Philology* 59(2): 479-508.
- Li, Paul Jen-kuei. 1990. Classification of Formosan languages: Lexical evidence. *Bulletin of the Institute of History and Philology* 61(4): 809-834.
- Li, Paul Jen-kuei. [1985]2004a. The position of Atayal in Austronesian family. *Selected Papers on Formosan Languages* (volume 2), pp. 805-834. Taipei: Institute of Linguistics, Academia Sinica.
- Li, Paul Jen-kuei. [2003]2004b. The internal relationships of six western plains languages. *Selected Papers on Formosan Languages* (volume 2), pp. 941-952. Taipei: Institute of Linguistics, Academia Sinica.
- Li, Paul Jen-kuei. 2013. Thao loans from Bunun. *Bulletin of Chinese Linguistics* 7(2): 225-241.
- Sagart, Laurent. 2014. In defense of the numeral-based model of Austronesian phylogeny, and of Tsouic. *Language and Linguistics* 15(6): 859-882.
- Saitou, Naruya & Masatoshi Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4(4): 406-425.
- Wang, Feng & William S.-Y. Wang. 2004. Basic words and language evolution. *Language and Linguistics* 5(3): 643-662.

The Genetic Position of Thao in the Austronesian Languages of Taiwan, China

FAN Zhiquan

[Abstract] The genetic position of the Thao language within the Austronesian languages of Taiwan, China lies in how to determine its genetic relationship of vertical differentiation and contact relationship of horizontal transmission. Thao retains over 50% of Austronesian cognates. It exhibits close linguistic ties to Bunun, Western Plains and northern Atayalic languages, suggesting frequent historical contact and interaction among their speaker communities.

[Keywords] Austronesian languages in Taiwan, China Thao genetic position evolution

(通信地址: 350007 福州 福建师范大学闽台区域研究中心
350118 福州 福建理工大学南岛语族研究院)

【本文责编 吴雅萍】