

从词阶理论看满通古斯语族的语支分类

王国庆¹, 赵杰²

(1. 中央财经大学 国际文化交流学院, 北京 100081; 2. 北京大学 外国语学院, 北京 100871)

摘要:满通古斯语族语言主要分布在中国和俄罗斯境内, 中外学者对满通古斯语族诸语言的语支分类一直存在分歧。我们通过考察词阶理论和讨论中俄内满通古斯语族诸语之间的亲疏关系, 验证前人对语支的分类。词阶的比较发现, 中俄境内满通古斯语族分为满语支和通古斯语支, 其中, 通古斯语支由南北两个分语支构成, 前者包括乌利奇语、奥罗克语、那乃语、奥罗奇语和乌德盖语, 后者包括埃文基语、埃文语和涅吉达尔语。

关键词:词阶理论; 满通古斯; 语言系属

中图分类号: H54 **文献标志码:** A **文章编号:** 1674-6627(2016)03-0123-04

满通古斯语族语言主要分布在中国和俄罗斯境内, 包括满语、乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语、埃文基语、埃文语和涅吉达尔语共9种语言。中外学者对满通古斯语族内部语支的分类意见不一, 大致可以概括为二分法、三分法和四分法。二分法认为, 满通古斯语族分为满语支和通古斯语支两大语支, 典型观点见于史禄国(1924年)、秦其乌斯(1949年)、鲍培(1965年)、朝克(2000年)等学者的论述。三分法认为, 满通古斯语族由北语支、东语支和西语支3个语支构成, 阿弗洛林(1955年)和迭茹夫叶尔(1978年)的观点具有一定的代表性。四分法认为, 满通古斯语族分为4个语支, 兰司铁(1926年)、池上二良(1974年)均持此观点^[1]。

陈保亚(1996年)提出了词阶理论, 目的在于鉴别不同语言间有音义对应关系的关系词是同源还是接触造成的。该理论的原理在于, 基本词汇是各语言中最为稳定的, 很少被相互借用, 而非基本词汇则容易被借用或渗透。他根据斯瓦迪士(M. Swadesh)先后制成的100核心词表和200核心词表确定基本词汇中的一级词阶和二级词阶。这种词阶分类经研究表明, 在各种语言的比较中具有相当普遍的适用性。他把100词表作为高阶词(一阶词), 是因为这些词汇是人类语言中基本核心词汇中最稳定的核心词, 200词虽然核心性略差, 但也是非常稳定的核心词。如果不同语言之间的关系词比较后的结果是高阶词(一阶词)的比例高于低阶词(二阶词)的比例, 则这些语言的关系为同源关系, 反之, 则为接触关系。这种研究方法不但适用于确定亲属关系, 同样也适用于确定亲疏关系, 也就是说, 一阶词与二阶词的比率呈下降或持平趋势, 这两种语言的关系就比较密切, 如果相反, 关系就比较疏远^[2]。

为了直观、清楚地反映出九种满通古斯语族语言之间词汇关系的亲疏远近, 本文拟尝试运用同源词词阶定量分析的方法, 通过各语言之间核心词的分布, 考察和讨论它们的亲疏关系。迄今为止, 利用词阶理论来证实中外满通古斯语族语言的亲疏关系还是第一次。我们先对上述九种满通古斯语言进行100核心词比较, 进行统计和分析, 然后再用200核心词^①进行统计和分析, 并根据统计结果来验证九种满通古斯语族语言的同源关系。

收稿日期: 2016-01-16

作者简介: 王国庆(1971-), 男, 黑龙江庆安人, 中央财经大学国际文化交流学院副教授, 博士, 主要从事比较语言学研究; 赵杰(1953-), 男, 吉林伊通人, 北京大学外国语学院教授, 北方民族大学北方语言研究院教授, 博士生导师, 主要从事语言学及民族学研究。

① 本文的100词和200词由近400个同源词中选出。同源词的资料取自秦其乌斯 V. Y. 主编的《满通古斯语言词源词典》(第1卷、第2卷), 列宁格勒: 科学出版社, 1975、1977年。本文中“涅吉”指“涅吉达尔语”。

表 1:100 词表九种满通古斯语族语言相关核心词的百分比

语言	埃文基	埃文	涅吉	满	乌利奇	奥罗克	那乃	奥罗奇	乌德盖
埃文基	—	0.92	0.98	0.77	0.82	0.85	0.91	0.87	0.89
埃文	0.92	—	0.92	0.74	0.83	0.82	0.90	0.83	0.90
涅吉达尔	0.98	0.92	—	0.78	0.87	0.87	0.93	0.90	0.91
满	0.77	0.74	0.78	—	0.82	0.78	0.89	0.77	0.80
乌利奇	0.82	0.83	0.87	0.82	—	0.90	0.96	0.89	0.85
奥罗克	0.85	0.82	0.87	0.78	0.90	—	0.90	0.89	0.88
那乃	0.91	0.90	0.93	0.89	0.96	0.90	—	0.93	0.91
奥罗奇	0.87	0.83	0.90	0.77	0.89	0.89	0.93	—	0.91
乌德盖	0.89	0.90	0.91	0.80	0.85	0.88	0.91	0.91	—

表 2:200 词表九种满通古斯语族语言相关核心词的百分比

语言	埃文基	埃文	涅吉	满	乌利奇	奥罗克	那乃	奥罗奇	乌德盖
埃文基	—	0.92	0.92	0.68	0.73	0.75	0.75	0.75	0.71
埃文	0.92	—	0.89	0.74	0.76	0.74	0.80	0.75	0.75
涅吉达尔	0.92	0.89	—	0.68	0.78	0.76	0.78	0.78	0.73
满	0.68	0.74	0.68	—	0.63	0.60	0.69	0.65	0.63
乌利奇	0.73	0.76	0.78	0.63	—	0.69	0.80	0.70	0.74
奥罗克	0.75	0.74	0.76	0.60	0.69	—	0.75	0.70	0.70
那乃	0.75	0.80	0.78	0.69	0.80	0.75	—	0.75	0.74
奥罗奇	0.75	0.75	0.78	0.65	0.70	0.70	0.75	—	0.73
乌德盖	0.71	0.75	0.73	0.63	0.74	0.70	0.74	0.73	—

在用词阶理论考察语言之间的同源或接触关系时,可用表 3 说明。

表 3:同源词比率与词阶分布规律

	关系远	关系近
一阶和二阶同源词比率	上升显著	下降(少数平直)
词阶分布	无阶性分布差别大	无阶性分布差别小

同源关系的阶曲线一般是下降的,少数平直,反映了从高阶词到低阶词的分布是“极显著下降”“显著下降”“下降”和“平直”的关系。高阶词和低阶词在 10~20 之间的差别反映了“显著下降”的关系,以此类推,10 以下为“下降”的关系,没有差别则为“平直”的关系。我们使用这种统计方法来说明满通古斯语言的亲疏关系,即高阶词(一阶词)的比例高于低阶词(二阶词)的比例,则这些语言的关系比较密切,反之则比较疏远。数据结果表明,满通古斯语族诸语言每一对语言核心词的比例出现两种情况:有阶分布和无阶分布。所谓“有阶分布”是指 100 核心词和 200 核心词的比例是有差距的,而“无阶分布”则是指两种核心词的比例没有差距。在比较中,“有阶分布”可以确定亲疏关系,“无阶分布”则表明关系密切。表 1 和表 2 的结果表明,1. 高阶词和低阶词之间的关系阶曲线为“显著下降”的有:①埃文基语和奥罗克语、那乃语、奥罗奇语,②涅吉达尔语和满语、奥罗克语、那乃语、奥罗奇语、乌德盖语,③满语和乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语,④乌利奇语和奥罗克语、奥罗奇语、乌德盖语,⑤奥罗克语和那乃语、奥罗奇语、乌德盖语,⑥那乃语和奥罗奇语、乌德盖语、奥罗奇语和乌德盖语;2. 高阶词和低阶词之间的关系为“下降”的有:①埃文基语和埃文语、涅吉达尔语、乌利奇语,②埃文语和涅吉达尔语、满语、乌利奇语、奥罗克语、那乃语、奥罗奇语和乌德盖语;3. 高阶词和低阶词之间的关系为“平直”的有:埃文基语和埃文语、埃文基语和涅吉达尔语、埃文语和涅吉达尔语。通过验证说明这九种语言无疑属于同源语言。下面我们将通过“有阶分布”来验证不同语言之间的亲疏关系,并以此作为语支分类的根据。

(一)埃文基语、埃文语、涅吉达尔语的关系

孟和达来、黄行(1997 年)发现了突厥语族语言中两个重要的现象:1. 当两个语言的相关核心词很多时,两阶核心词的差别从概率统计上看不显著,但同源程度越高的语言,相关核心词词阶分布差异(无阶性)越小;2. 当两个语言的相关核心词较少时,两阶核心词的差别从概率统计上看不显著,但相关核心词词阶分布的差异大小仍然能识别同源程度的高低^[3]。我们对满通古斯语族诸语言相关核心词分析的结果和孟和达来、黄行(1997 年)对突厥语族语言的研究发现有共通之处,即无论相关核心词概率统计如何,相关核心词词阶分布的差异,即无阶性的大小,同样可以说明这些语言的亲疏程度。埃文基语、埃文语、涅吉达尔语关系词每两种语言之间高阶词的平均比例为 94%,低阶词的平均比例为 91%,高阶词和低阶词的比例差别为 3%,这印证了当两个语言的相关核心词很多

时,两阶关系词的差别从概率统计上看虽不显著,但从相关核心词阶分布差异(无阶性)来看,埃文基—涅吉达尔语最为接近,埃文基—埃文语次之,其次是埃文—涅吉达尔语。

陈保亚指出,“在第一百词的关系词比例分布中,同一语支的语言之间高于非同一语支语言之间的关系词比例”^{[2] (232)},因此可以这样认为,高阶词的比例越高,语言的亲属关系越密切。以此原理来考察亲疏关系,埃文基语、埃文语、涅吉达尔语之间高阶词的比例最高,说明这三种语言比较接近。埃文基语、埃文语和涅吉达尔语之间关系很近,还可以从另一个方面反映出来,即高阶词和低阶词的比例关系。“当200核心词中的关系词大于180时,阶分布是一般的下降,个别平直,不再出现极显著下降和显著下降的情况。从整个阶分布的平均情况看,也不显著”;“汉语北方方言各次方言之间分化年代不长,同源词比例比较高,第一阶词和第二阶词同源词的差距不太大,尤其是同源词比例在90%以上时,差距较小”^[4]。从上表可见,这三种语言的阶分布差异不大,基本处于平直状态,说明埃文基语、埃文语和涅吉达尔语分化年代不长,彼此接近。

埃文基语、埃文语、涅吉达尔语之间高阶词(1阶词)的比例高,埃文基和埃文语共有关系词数量为92,埃文基和涅吉达尔语为98,埃文语和涅吉达尔语为92,关系词之间的比例在92%~98%之间;200核心词的关系词数量,埃文基和埃文语为184,埃文基和涅吉达尔语为190,埃文和涅吉达尔语为181,均超过180,据此我们可以认定,埃文基语、埃文语、涅吉达尔语关系最近。由此我们得出两个结论:1. 高阶词中语言之间的关系词比例高;2. 200核心词的关系词的数量大于180时,高阶词和低阶词的差异呈现出的阶分布是一般的下降或者平直,分化时间较晚。这从埃文基语、埃文语和涅吉达尔语的比较就可以看出,孟和达来、黄行对突厥语族中维吾尔语同哈萨克语、柯尔克孜语、乌孜别克语、塔塔尔语的关系验证的两条标准,在验证满通古斯语族语言时也有效。

(二) 乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语的关系

乌利奇语、奥罗克语、那乃语、奥罗奇语和乌德盖语中相关核心词体现出两个特点:1. 高阶词中相关核心词比例高,约为91%;2. 低阶词平均比例约为73%,高阶词和低阶词之间的比例不同于埃文基语、埃文语和涅吉达尔语,差距较为明显,平均比例差别为18%。高阶词的相关核心词比例分布中,同一语支的语言之间高于非同一语支语言之间的相关核心词比例,乌利奇语、奥罗克语、那乃语、奥罗奇语和乌德盖语关系词中高阶词的比例很高,超过了90%,说明这五种语言的亲属关系密切。高阶词和低阶词之间的比例大于埃文基语、埃文语和涅吉达尔语高阶词和低阶词的比例,说明这几种语言的分化年代比埃文基语、埃文语和涅吉达尔语的分化年代早。

(三) 乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语同埃文基语、埃文语、涅吉达尔语的关系

乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语同埃文基语、埃文语、涅吉达尔语两组语言之间高阶词的平均比例约为87.3%,低阶词的比例约为75%,高阶词和低阶词的比例平均差别为12%,说明这两组语言的差别比较大。

(四) 满语和埃文基语、埃文语、涅吉达尔语的关系

满语与埃文基语、埃文语、涅吉达尔语中关系词的情况说明,无论是高阶词还是低阶词,比例均小于埃文基语、埃文语和涅吉达尔语高阶词和低阶词的比例。满语同埃文基语、埃文语和涅吉达尔语高阶词平均比例为77%,低于后三种语言高阶词比例的94%;满语同埃文基语、埃文语和涅吉达尔语低阶词平均比例为71%,低于后三种语言高阶词比例的91%。前者四种语言中高阶词和低阶词的比例差别为6%,后者三种语言高阶词和低阶词的比例差别为3%。虽然两个数字的差别不大,但是代表的意义却不同。根据孟和达来、黄行发现,满语和埃文基语、埃文语、涅吉达尔语两阶语言关系词差别概率不显著,为6%。这是因为满语和埃文基语、埃文语、涅吉达尔语三种语言每两种之间的相关核心词相对于埃文基语、埃文语、涅吉达尔语每两种语言之间的核心关系词来说,数量较少,两阶关系词的差别从概率统计上也不显著,埃文基语、埃文语和涅吉达尔语之间两阶相关核心词差别概率小,说明它们关系接近,而它们与满语相关核心词差别概率略高,说明关系较远。

(五) 满语和乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语的关系

满语与乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语关系词中,高阶词的比例较高,平均为81%,而低阶词的比例较低,平均为64%,高阶词和低阶词的比例差别平均为17%。就高阶词平均比例来看,满语与乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语中关系词的比例比同埃文基语、埃文语、涅吉达尔语中关系词的比例高4%。

高阶词的比例越高,说明语言之间的关系越密切。乌利奇语、奥罗克语、那乃语、奥罗奇语和乌德盖语同埃文基语、埃文语、涅吉达尔语的关系词中每两种语言高阶词的比例平均约为87%,满语同埃文基语、埃文语和涅吉达尔语每两种语言高阶词平均比例为77%,满语与乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语每两种语言高阶词平均比例为81%,由此可以得出初步的结论:乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语同埃文基语、埃文语、涅吉达尔语的关系比满语同埃文基语、埃文语和涅吉达尔语的关系更近。下面我们再进一步来看乌利奇语、奥

罗克语、那乃语、奥罗奇语、乌德盖语同埃文基语、埃文语、涅吉达尔语；满语与乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语；满语同埃文基语、埃文语和涅吉达尔语相关核心词中共有高阶词的分别情况。

表 4: 满通古斯语族诸语共有高阶词分布比例

语言	满—埃文基— 埃文—涅吉	满—乌利奇— 奥罗克—那乃— 奥罗奇—乌德盖	乌利奇—奥罗克— 那乃—奥罗奇— 乌德盖—埃文基— 埃文—涅吉	乌利奇—奥罗克— 那乃—奥罗奇— 乌德盖	乌利奇—奥罗克— 那乃—奥罗奇— 乌德盖
高阶词比例	69%	71%	77%	84%	91%

三组语言的情况分析结果如下:第一组满语同埃文基语、埃文语、涅吉达尔语高阶词的共有词的比例为 69%,第二组满语同乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语高阶词的共有词的比例为 71%,第三组乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语同埃文基语、埃文语、涅吉达尔语高阶词的共有词的比例为 77%。第三组语言的高阶共有词比例比第一组语言高 8%,这进一步说明,乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语同埃文基语、埃文语、涅吉达尔语之间关系更近,加上乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语同埃文基语、埃文语、涅吉达尔语每两种语言中高阶词的平均比例也很高,大约达到了 87%,因此可以把乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语同埃文基语、埃文语、涅吉达尔语列入一个大的语支。考虑到乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语每两种语言的高阶词和低阶词比例的平均差别较大,差别为 18%,明显高于埃文基语、埃文语和涅吉达尔语之间的 3%,因此可以把两族语言划分为同一语支下的不同分支。满语同上述两组语言之间高阶词的相关核心词比例相当,分别为 69% 和 71%,可以列为独立语支的语言。

由此,我们对中俄境内满通古斯语族语言语支分类如下:满通古斯语族语言分为满语支和通古斯语支,满语支包括满语;通古斯语支有两个语群构成:埃文基语、埃文语、涅吉达尔语为一个语群,乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语为另一个语群,也有学者把前者称为通古斯北语支,后者称为通古斯南语支。

综上所述,学界对满通古斯语族诸语的内部系属划分结果一直存在着不同的结论。我们运用词阶理论,选择斯瓦迪士 100 和 200 核心词表,对所选的九种满通古斯语族语言关系进行了考察。按照词阶理论,斯瓦迪士 200 核心词分为两个层次:高阶词(一阶词)和低阶词(二阶词)。根据高阶词比例的高低,以及高阶词和低阶词比例的差异,考察了九种语言的亲疏关系,为满通古斯语族语言进行语支分类。我们将所考察的语言分为满语支和通古斯语支,其中通古斯语支由南北两个语群构成,前者由乌利奇语、奥罗克语、那乃语、奥罗奇语、乌德盖语构成,后者包括埃文基语、埃文语和涅吉达尔语。

参考文献:

- [1] 朝克. 关于满通古斯诸语的分类[J]. 世界民族, 2000(3).
- [2] 陈保亚. 语言接触与语言联盟——汉越(侗台)语源关系的解释[M]. 北京: 语文出版社, 1996.
- [3] 孟和达来, 黄行. 蒙古语族和突厥语族关系词的词阶分布分析[J]. 民族语文, 1997(1).
- [4] 陈保亚. 再论核心关系词的有阶分布[J]. 民族语文, 1998(3).
- [5] 李永海. 满一通古斯语族通论(上)[J]. 满语研究, 2003(1).

【责任编辑 李小凤】

Language Branches Classification of Manchu Tungusic from Word Order Theory

WANG Guo-qing¹, ZHAO Jie²

(1. School of International Cultural Exchanges, Central University of Finance and Economics, Beijing 100081, China;

2. School of Foreign Languages, Peking University, Beijing 100871, China)

Abstract: Manchu Tungusic language is mainly distributed in the territories of China and Russia. Scholars both in China and abroad have respective ideas on the branch classification. In this article, the authors will make an analysis of the languages relationship to re-classify the branch language.

Key words: Word Order Theory; Manchu Tungusic; Language Family