

CSSCI 来源集刊

語言學論叢

ESSAYS ON LINGUISTICS

(第四十一辑)

北京大学汉语语言学研究中心

《语言学论丛》编委会编

商務印書館

2010年·北京

菲律宾塔伽洛语(Tagalog)的词类范畴

罗仁地

提要 本文透过 Himmelmann (2007) 和 LaPolla (2008) 所提出的有关词的分类法以及两位语言学家对菲律宾塔伽洛语的一些看法,探讨塔伽洛语的词类问题。主要目的在于滤清对塔伽洛语词类的一些争议,并进一步展现塔伽洛语词类的独特性。也由此对汉语词类问题进行反思。

关键词 菲律宾 塔伽洛语 词类 汉语

1 简介

塔伽洛语(Tagalog,塔加路语、他加洛语、他加禄语)^①是菲律宾的官方语言之一(另一官方语为英语)。属于南岛语系西部马来一波利尼西亚语族中部菲律宾语支。使用人口:88,574,614(2007年人口普查),在菲律宾,目前以塔伽洛语作为母语的使用者大约占全国人口的百分之三十,主要分布在吕宋岛和周围的部分岛屿。

对于研究句法结构的语言学家来说,塔伽洛语是很有挑战性的语言。语言学者对于塔伽洛语语法的看法主要在两大方面有比较大的争议:一方面是有关语态系统,另一方面是有关实词的词类^②。看法不一主要是因为塔伽洛语的语态系统和词类在概念上有别于大多数语言(包括英语在内)。

为了让语言学界对塔伽洛语的词类有比较清楚的了解,本文根据 Himmelmann (2007) 和 LaPolla (2008) 所提出的有关词的分类法的理论概念,并以此理论概念对塔伽洛语的词类进行分析讨论。本文虽然只讨论词类的问题,但是,语态系统和词类这两个方面并非

是完全没有关联的。

2 词的分类法

讨论词类的时候,我们必须要先分清所论及的词类范畴指的是哪一种范畴。是词义范畴,还是语法范畴?是语法范畴的话,是以形态来定的,还是以句法分布来定的?

a. 词义范畴。我们可以根据某一个词的意思,也就是该词是用来指“人”或是“物”或是“行为”,抑或是“状态”等来划定该词的范畴。从这个层面出发所划定的词的范畴是该词的词义范畴。

b. 语法范畴。包括两个层面。

1b 形态层面:当我们从某个词的形态变化划定该词的范畴的时候,我们所划定的是该词的形态范畴;

2b 句法层面:当我们从某个词的句法分布,也就是根据该词在句子中出现的位置来定范畴,那么从这个层面出发所划定的是该词的句法范畴。

一般来说,当我们在讨论词类范畴的时候,我们指的是词的语法范畴。但是我们必须要注意的是,如上所述,语法范畴包括两个层面,即形态层面和句法层面。印欧语系语言里的词,从形态和句法这两个层面所划定的词的范畴是一致的。但一直以来,当我们在讨论词的范畴的时候,总是理所当然地认为所有语言的词的语法范畴都和印欧语系一样,形态和句法两个层面是一致的。可是事实上,有些语言从这两个层面所定的范畴是不一致的。也就是说有些语言的词,从形态层面定的范畴和从句法层面定的范畴是不一样的。塔伽洛语就属于这一类。

Himmelman 把语言的词类在形态与句法层面上所呈现的不同情况列表如下(Himmelman 2007:264):

表一 语言的词类在形态与句法范畴上可能呈现的不同情况

语言	形态层面		句法层面
1a	范畴明确	≠	范畴明确
1b	范畴明确	=	范畴明确
2	范畴不明确		范畴明确
3	范畴明确		范畴不明确
4	范畴不明确		范畴不明确

呈现情况 1a 的语言,在形态层面和句法层面都有明确的词类范畴,但两个层面的范畴不一致;

呈现情况 1b 的语言,在形态层面和句法层面都有明确的词类范畴,而且两个层面的词类范畴是一致的;

呈现情况 2 的语言,在形态层面没有明确的范畴,可是在句法层面有明确的范畴;

呈现情况 3 的语言,在形态的层面有明确的范畴,可是在句法层面没有明确的范畴;

呈现情况 4 的语言,在两个层面都没有明确的范畴。

印欧语系大部分的语言呈现情况 1b。Himmelman 认为塔伽洛语和北美洲的 Salishan 和 Cayuga 语言都呈现情况 3。

3 塔伽洛语的词类

有些学者认为塔伽洛语呈现情况 3 的原因是因为塔伽洛语的词根没有预设的语法范畴(即词根是 pre-categorical),也就是说塔伽洛语的词根本身是没有范畴的,要等到那个词附加形态标志出现在句子中才根据那个词在句子中出现的位置而定词类(如 Foley 1998)。另外,有些学者认为塔伽洛语所有的实词都是名词(如 Kaufman 2009)。Himmelman (2007) 并不认同塔伽洛语的词根没有预设的语法范畴的说法,因为他认为塔伽洛语的词根是可以单独使用的,也就是说没有附加词缀的词根可以用来指东西、有生命

体、财物、状态、动作的结果、动作的名称等等,而且词根在形态的层面呈现不同的词类范畴,只不过这些词类不是名词或动词这样的词类,而是根据能和哪些词缀搭配而分的词类,这是塔伽洛语独特的词类。Himmelman 认为从词义层面来讲,塔伽洛语的词可以用来指“人”、“物”、“行为”或“状态”,这是词的词义层面范畴,可是这一层面的范畴跟语法层面的范畴是没有关系的。就是说,从塔伽洛语词的词义并不能看出该词属于哪个词类。Himmelman 提出了一个对词类的探讨很关键的观点,就是如果词义范畴和所设定的语法范畴完全一致,那么就没有真正的语法范畴,只有词义范畴。

Himmelman (2007) 进一步说明,例举塔伽洛语词有两种类型:A类和B类,并以使用广泛的前缀 *ma-* 为例列表如下(Himmelman 2007:270):

A类词例:

ma+词根	词义	词根	
mabahay	有很多房子	<bahay	房子
mabaha	水涨得厉害	<baha	洪水
maganda	很(有)美	<ganda	美
madali	很(有)快	<dali	快
marami	很(有)多	<dami	(大的)量

B类词例:

ma+词根	词义	词根	
malutas	能够解决	<lutas	解决
mahinog	能够熟	<hinog	熟
mamura	能够便宜	<mura	便宜
madurog	变成碎	<durog	碎
maluto	能够煮熟	<luto	煮
maputol	变成断了	<putol	断
mabutas	变成有洞	<butas	洞
matapos	能够完成	<tapos	完成
magalit	变成生气	<galit	怒
magutom	变成饿	<gutom	饿

A类词,前缀 *ma-* 十词根,前缀 *ma-* 表示“有(很多)词根所指

的事、状态或物”；

B类词:前缀 *ma-* + 词根,前缀 *ma-* 表示“形成词根所指的事、状态或物”。

B类的词可以带语态和体的标志,比如:*magalit*“变成生气”的实然整体体形式^③是 *nagalit*“生气了”,实然非整体体形式是 *nagagalit*“开始(正在)生气”,而非实然非整体体形式是 *magagalit*“将要生气”。可是A类的词一般不可以带语态和体的标志。A类的词可以重叠词根的第一个音节来表示复数:比如 *magaganda*“漂亮的(复数)”。然而B类词的重叠形式却表示非整体体。部分A类的词,如果加中缀 *-um-*,就表达类似B类“变成词根所指的事、状态或物”的意思(即类似B类的词带前缀 *ma-* 的意思),如 *ganda*“美”,*gumanda*“变美”。同样地,部分B类的词可改变重音位置来表达类似A类带前缀 *ma-* 的词的意思,即“有(很多)词根所指的事、状态或物”的意思;比如 *bútas*“洞”,把重音改在第二个音节上,*butás* 就是“破了个洞”的意思。就是说,这两类词可以用不同的形态方式来表达同样的语义概念。由此,我们可以推断这些范畴之间的差异是语法差异,不是语义差异。

虽然塔伽洛语在形态层面可以分出以上这两种词类,但是这两种词类可以充当谓语,也可以充当论元,也可以充当修饰成分。请看以下以A类词 *baha*“淹水”和B类词 *patay*“死亡”为例的句子:

分析 A1 *baha na ang kalsada*

淹水 CSM SPEC 马路
马路淹水了

分析 A2 *kalsada ang baha*

马路 SPEC 淹水
淹水的是马路

分析 A3 *baha na kalsada*

淹水 LNK 马路
淹了水的马路

B1 patay na ang baboy

死 SPEC 猪

猪死了

B2 baboy ang patay

猪 SPEC 死

死的是猪

B3 patay na baboy

死 LNK 猪

死猪

A类词 *baha* 在 A1 中充当谓语,在 A2 中充当论元,而在 A3 中充当修饰成分。B类词 *patay* 也一样,在 B1 中充当谓语,在 B2 中充当论元,而在 B3 中充当修饰成分。

塔伽洛语的词根虽然可以单独用,可是词根不能用来指正在进行的动作或某个特定的动作。只有带语态标志的词才能用来表示具体动作,或在进行中的动作。而且带语态标志的词同时标志动作的某一个参与者。比如:

Iniabot ni Tong ang mga perlas, at itinuloy ang tanong.

[[in-i-abot]_{PRD} [ni Tong] [ang mga perlas]_{TOP}]

RPUT-CON-到达 REL PN SPEC PL 珍珠

at [[i-in+tuloy]_{PRD} [ang tanong]_{TOP}]

CONJ CON-RPUT+继续 SPEC 问题

Tong 把珍珠递过去了,然后继续问他的问题。

在这个句子里,*iniabot* 指的是“递过去”这个具体动作,同时标志句子的主题是所递过去的东西,而 *itinuloy* 指的是“继续”这个具体动作,同时标志句子的主题是“问问题”这个动作。

因为塔伽洛语带语态 (voice) 标志的词和不带语态标志的实词的形态和语义不同,所以 Himmelmann (2007) 认为塔伽洛语有语态标志的词是塔伽洛语独有的形态—词类范畴 (morpho-lexical category),就是在形态层面是一个独特的词类,可是在句法层面跟

其他实词一样,可以当论元也可以当谓语。在这一点上与英语有很大的差异。英语的形态词和句法词是相互交叉的,譬如,英语的 *die* “死亡”,由于这个词在形态层面是不及物动词,因此它同时在句法层面也是动词而不可能是名词。塔伽洛语却不一样,塔伽洛语有语态标志的词都是衍生词,而英语的却不一定。塔伽洛语有语态标志的词的衍生包括形式和语义两方面。在形式方面,词根转为有语态标志的形态一词类范畴。在语义上,指东西、有生命体、财物、状态、动作的结果或动作的名称等等的词根衍生为指动作的具体进行(不是动作的结果)的词,同时该衍生词也指动作的某一个参与者。此外,英语非衍生的基本动词都是主动态(active voice),而塔伽洛语带语态标志的词都是衍生的,没有一个基本形式。

4 塔伽洛语的词组

通常在写参考语法的时候,我们以短语(词组)来分章节,就是有一章讨论名词短语,一章讨论动词短语,在概念上,通常名词短语是以名词为中心词的短语,动词短语是以动词为中心词的短语。但是,如果一个语言的词类不包括名词和动词这两类词的话,那么构成句子的成分是不是形成不同类型的短语?如果是的话,所形成的短语是不是名词短语和动词短语呢? LaPolla (2008) 就以这个问题来分析塔伽洛语的长篇语料,发现除了单词可以作为句子成分以外,塔伽洛语的句子还呈现两种不同结构的短语,而这两种短语可以充当谓语,也可以充当论元,而且在这两种短语中,任何一个实词都可以充当短语的中心词。因为这两种短语跟名词短语、动词短语没有关系,LaPolla (2008) 把它们叫做“X 短语”和“Y 短语”,或者“=ng/na 短语”和“ng ([naŋ]) 短语”。

X 短语有两个或两个以上的成分,成分之间出现一种连接标志。X 短语的连接标志有两种形式: =ng 或者 na。如果连接的两个词,前面的词有辅音韵尾的话就用助词 na 形式,na 出现在两个

词之间。但是如果连接的两个词,前面的词有元音韵尾的话就用附属词 =ng 形式,由 =ng 黏着在前面的词后面。这种结构不仅表达成分之间存在着一种修饰和被修饰的关系,还可以表示很多不同的修饰关系:状态、数词、容量词、关系子句、指示词、副词与实词之间的关系,谓语或论元内部非拥有性质的关系,间接引语,非主要子句成分之间的关系等。X 短语所表达的修饰关系有时候很难确定,好像没有具体的语意,只是表示短语内的成分形成短语而已。X 短语里面的成分常常可以转换位置,而不影响短语的意思,比如 *magandang babae* 和 *babaeng maganda* 同样都是“漂亮的女人”的意思。因此我们没有办法从 X 短语里面的词的位置认定它是否为短语的中心词,因此 X 短语没有以语法定中心词,只有根据语义定的中心词。

Y 短语也有两个或两个以上的成分,成分之间也出现一种连接标志。如果标志后面的词表示人名,标志的形式是 *ni*,要不然其形式是 *ng* ([*nanj*])。这个结构所表达的意思是成分之间存在着一种修饰和被修饰的关系,包括整体一部分之间的关系、拥有关系、抽象的拥有关系、谓语和状语之间的关系以及谓语和非主题论元之间的关系。这个结构里面的第一个成分是整个短语的中心词。

以下请看这两种短语结构在一篇长篇故事语料的几个片段中的用法 (LaPolla (2008)):

(1) *Alamat ng Gubat*

[*alamat* _Y *ng* _X *gubat*]_{X/TITLE}^①

传说 REL 森林

森林的传说

(2) *Noong unang panahon,*

[*noon = ng* _Y *una = ng* _X *panahon*]_{X/TOP}

那时 = LNK 第一 = LNK 时间/时代/天气

很久以前(在第一个时代的时候),

(3) *sa isang liblib na kaharian,*

[sa [isa=ng liblib na ka-hari-an]_{X/LC}]_{LP}

LOC —=LNK 偏远 LNK ASSOC-王-LFS

在一个偏远的王国,

(4) *sa ilalim ng dagat,*

[sa [ilalim ng dagat]_{Y/LC}]_{LP}

LOC 底 REL 海洋

在海洋的深处,

(5) *ay may nakatirang maganda at mabait na sirena,*

ay [may [naka-tira=ng

FT EXIST OS-住=LNK

[ma-ganda at ma-bait]_{CONJP} na sirena]_{X/EX}]_{PRED}

STAT-美 CONJ STAT-善 LNK 鱼美人

有个美丽善良的鱼美人住在那儿。

(6) *Però wala siyang kinalaman*

pero [wala <siya>_{TOP}=ng[in-ka-alam-an]_{EX}]_{X/PRED}

但是 N. EXIST 3单 T=LNK RPUT-ASSOC-知道-LFS

sa kwentong ito.

[sa [kwento=ng ito]_{X/LC}]_{LP}

LOC 故事=LNK 这

但是她无关,跟这个故事。(但是她跟这个故事没有关系。)

(7) *Inabot s'ya ng pitong araw sa paglalakbay ...*

[[in-abot]_{PRED} <siya>_{TOP} ng [pito=ng araw]_X]_{Y/CL}

RPUT-到达 3单 T REL 七=LNK 日

[sa pag-la-lakbay]_{LP}

LOC ACTNOM-REDUP-旅游

他的旅程用了七天……

(8) *Idinura ni Buwaya si Maya na naglilinis ng ngipin n'ya.*

[[i-in+dura]_{PREP} ni Buwaya]_Y

CON-RPUT+吐 REL 鳄鱼

[si [Maya na [nag-li-linis

SPEC PN LNK RIAT-REDUP-干净

ng [ngipin niya]_Y]_Y]_X]_{TOP}

REL 牙齿 3 单 NT

鳄鱼把正在给它刷牙的 Maya 吐出来了。

(1) 是故事的题目,形成 Y 短语。(2) 到 (5) 是一个长句。(2) 是个 X 短语。X 短语可以充当句子的任何成分。

塔伽洛语的句子主要有四种成分:谓语、主题论元(有 *ang* 标志)、非主题论元(有 *ng* ([*naŋ*]) 标志)^⑤以及处所/时间论元(有 *sa* 标志)。处所短语的构成是(处所标志 *sa* + 处所),单词、X 短语、Y 短语都可以充当表示处所的成分。在(3)里,充当表示处所的成分是 X 短语;但在(4)里,充当表示处所的成分是 Y 短语。

(5) 句 *ay* 以后的短语是这个句子的谓语。塔伽洛语句子的谓语一般是句子的第一个成分,但是有时候主题或处所论元可以出现在谓语前面。在这样的情况下,谓语前面要加 *ay*。这里,充当谓语成分的是存在短语。存在短语的结构是[存在量词(existential quantifier) *may/mayroon* + 表示所存在的东西的单词、X 短语或 Y 短语]。这一行里有 *may* + X 短语,而这个 X 短语的前三个成分都是带语态词缀的词。

如果所出现的存在量词是 *mayroon*,存在量词和表示所存在的东西的单词、X 短语或 Y 短语形成 X 短语,如 *mayroong karamdaman* “有病”。如果谓语是否定存在短语,那么谓语的否定存在量词和表示不存在的东西的单词、X 短语或 Y 短语形成 X 短语,如: *walang karamdaman* “没有病”。

(6) 的第三人称单数主题代词指的是句子的主题(代词分主

题、领属、向格 (dative) 三套形式), 从词序来说, 出现在谓语里面, 可是从句法来说, 不是谓语的成分, 是一种“第二位置附属词” (second-position clitic), 出现在谓语第一个词后面。这里的谓语, 如果没有第三人称单数主题代词, 就是 *walang kinalaman*。从这个例子我们可以知道, 形成 X 短语的成分不一定出现在一起, 而且连接标志 =ng 所附属的词不一定是 X 短语的成分。另外, 在 (6) 里, 表示所存在的东西的词 *kinalaman* 是个带语态词缀的词。

从 (7) 和 (8) 句中我们比较容易看得出塔伽洛语的句子本身是一种 Y 短语。(7) 里的谓语和论元“七天”形成 Y 短语 (第三人称单数代词又是第二位置附属词, 因此出现在谓语后面), 而且表示“七天”的短语是 X 短语, 就是说, X 短语可以出现在 Y 短语里面。同样, Y 短语可以出现在 X 短语里面。

在 (8) 句里, 表示施事者的词和谓语形成 Y 短语, 而表示受事者/主题的短语是 X 短语, 在这个 X 短语里面包含一个 Y 短语, 这个 Y 短语里面又包含一个 Y 短语 (领属代词修饰另外一个词的时候便与这个词形成 Y 短语)。

5 总结

从以上简短的讨论中, 我们可以看出塔伽洛语词类的独特性。在此可以滤清的是塔伽洛语并不是没有“词类”, 关键是当我们在谈词类的时候, 我们对词类的定义是什么。如果是根据英语或印欧语对词类的概念出发, 那么塔伽洛语是没有那种词类。正如上文所指出的, 在讨论词类范畴的时候, 我们必须注意的是我们是从哪一个层面来讨论词类范畴。在印欧语系语言里, 无论是在形态层面或是句法层面, 词的类范畴都是一致的, 因此, 从印欧语系的观点出发来讨论词类范畴, 理所当然地, 没有必要去考虑形态与语法这两个层面。可是并不是所有语言的结构都与印欧语言一样, 有些语言, 在讨论其词类范畴的时候, 必须先考虑到词类范畴的层面, 是从形态层面

还是从语法层面。因为有的语言的词类,从形态层面定的范畴和从语法层面定的范畴有时候是不一致的。塔伽洛语就属于这一类。塔伽洛语的实词在形态层面是可以分词类的,可是这些词类跟语义范畴没有直接关系,即不能说是名词或动词,而是根据能和哪些词缀搭配而呈现的词类。塔伽洛语从语义层面定的词类范畴跟从形态层面定的词类范畴是没有关系的。也就是说,从塔伽洛语的词义并不能看出该词是属于哪个词类。塔伽洛语词类的独特性,也在以上 LaPolla 根据塔伽洛语语段分析的结果很明显地展现。

6 汉语的分类情况

在讨论塔伽洛语的同时,我们也想到有关汉语词类的讨论。Himmelmann (2007:264)在讨论词类的可能呈现的不同情况(表一)时,提到极端独立性的语言可能呈现情况 2。就是说,如果某一个语言完全没有形态标志,那就没办法以形态来定词类;如果分词类的话,就要靠句法分布。当然,极端独立性的语言也可能呈现表一里的情况 4,不但在形态层面范畴不明确,而且在句法层面范畴也不明确。这就让我们想起汉语:赵元任(1968:501)认为汉语虽然有少数的词可以用形态来定词类(比如“桌子、桃儿、骨头”等),但大部分的词只能以句法分布来定词类。这就是说,汉语大部分的词在形态的层面范畴不明确,可是在句法层面的范畴明确,那也就是 Himmelmann 所说的情况 2。

缩写符号

3 单 NT	第三人称非主题代词	LNK	连接附属词/助词= <i>ng/na</i>
3 单 T	第三人称主题代词	LP	处所短语
ACTNOM	行为名物化前缀	N.EXIST	否定存在连词
ASSOC	共同行为的构词前缀	OS	状态持续的构词前缀
CL	子句	PL	复数助词

CON	运送构词前缀	PN	人名
CONJ	连词	PRED	谓语
CONJP	连词短语	REL	连接助词 <i>ng</i> [nɑŋ], <i>ni</i>
CSM	情况变化助词	RIAT	标实然非整体主题的前缀
EX	所存在的人或物	RPUT	标实然整体受事者主题的中缀
EXIST	存在量词	REDUP	重叠形式
FT	主题前置助词	SPEC	标志有具体所指的助词
LC	处所	STAT	状态构词前缀
LFS	处所构词后缀	TITLE	题目
LOC	处所助词	TOP	主题

附注

①笔者认为“塔伽洛”比较接近原音。

②所有研究塔伽洛语的学者都认为塔伽洛语有虚词和实词之分,可是有关实词的性质和分类的意见不一致。

③我们这里把 *realis perfective aspect* 译为“实然整体体”。常常有人把 *perfective* 译为“完成体”,但这是不恰当的,因为 *perfective* 的意思并不是指动作已经完成了,而是把动作看成是个整体,着重点不在动作的进行。

④短语后面加上了“结构类型/功能”的标签。

⑤其实 *ng* ([nɑŋ])不是论元的标志,*ng* 是短路的连接标志。由于 *ng* 出现在短路的成分之间,因此,*ng* 后面的论元不能前置,即不像 *ang* 和 *sa* 后面的论元那样可以出现在谓语前面。

参考文献

- Chao, Yuen Ren (赵元任) (1968) *A Grammar of Spoken Chinese* (中国话的文法). Berkeley and Los Angeles: University of California Press.
- Foley, William A. (1998) Symmetrical voice systems and precategoriality in Philippine languages. Paper presented at the 3rd LFG conference, Brisbane.
- Himmelmann, Nikolaus P. (2007) Lexical categories and voice in Tagalog. In Peter Austin and Simon Musgrave (eds.), *Voice and Grammatical Relations in Austronesian Languages*, 246-293. Stanford, CA: CSLI Publications.
- Kaufman, Daniel (2009) Austronesian Nominalism and its consequences: A

Tagalog case study. *Theoretical Linguistics* 35-1, 1-49.

LaPolla, Randy J. (2008) Constituent structure of a Tagalog text. Plenary address to the 10th Philippine Linguistics Congress, University of the Philippines, 10-12 December, 2008.

(澳大利亚拉筹伯大学)