

汉语介词短语（np）的蒙古语转换规则研究

巴达玛敖德斯尔

（内蒙古大学 蒙古学学院语言所，内蒙古 呼和浩特 010021）

摘要：本文在借鉴汉语语法关于名词（n）、名词短语（np）的研究成果和传统蒙古语文翻译关于汉语n、np的译法的基础上，根据“政府工作报告”语料，以“达尔罕”汉蒙机器辅助翻译系统为依托，面向汉蒙机器翻译的实际需要，尝试以形式化的方式对汉语np的蒙古语转换规则进行描述的同时，探讨了有关np转换方面的语法和翻译问题。

关键词：汉语；蒙古语；n；np；转换规则；汉蒙机器翻译

中图分类号：H059 **文献标识码：**A

跨语言的研究证明，名词是人类语言中最基本的词类范畴和构成句子的最基本要素之一，在语法系统中具有非常重要的作用。名词短语是汉、蒙两种语言最基本的短语。长期以来，汉语语法学界对名词和名词短语的研究取得了丰硕的成果。蒙古语名词研究方面也发表了一些论文。特别值得一提的是，面向信息处理的蒙古语名词短语结构规则研究也有比较好的开头。但是，汉语和蒙古语之间名词短语的对比研究以及短语转换规则的研究，一直是一个空白。汉蒙名词短语转换规则研究，对于汉蒙翻译，对基于规则的和基于实例的汉蒙机器翻译系统的研制开发，对于汉蒙双语对比研究及教学、汉蒙双语词典的编纂等都具有重要的理论和应用价值。

从汉语 np 的内部组成情况看，大致可以分为附加结构、“所”字结构、联合结构、定中结构等四个类型。本文对汉语这几种 np 的蒙古语转换规则进行形式化描述的同时，探讨有关 np 转换方面的语法和翻译问题。

1. 附加结构 np

这里讨论的是 n 加“们、者、性”等语素（g）而构成的附加式 np 或 npbar（名词性准短语）的规则。从政府文献语料看，主要涉及以下几种结构类型。

(1) np->!n g<们> :: \$.内部结构=单词,%n.语义类=人,\$!=%n @复数,\$.复数=Yes

=> NP(!N/n F/g)

这里的汉语规则限制了加“们”的那个 n 的语义类为“人”，而且不具有“复数”这个属性。从“=>”开始是蒙古语的转换规则。意思是说，汉语一个名词加“们”字而构成的 np，在蒙古语中也可以转换为名词短语（NP），但蒙古语名词（N）后面要接一个表示复数的附加成分（标记为 F）。当然，蒙古语里可以接在表示人的词后面表示复数的附加成分不止一个，有 NAR，-D，-S，-CVD|-CUD 等好几个，但在这里暂时可以限定为就有一个形式 NAR。还有一个问题是，NAR 在蒙古语中既可以连写，又可以分开写，我们这里硬性

规定分开写。所以，我们在词典中把汉语的“们”翻译成了 NAR。请看与这条规则发生关联的汉语“们”的翻译在汉蒙机译词典中的写法：

\$\$ 们

**{g}g \$=[]

=> F<NAR>

(2) np->!n g<者>

汉语“者”对应于蒙古语的“-TANI-TEN”、“-GCI|-gCI”等表示“拥有该事物的人”、“人物”的名词的构词附加成分。蒙古语构词法的情况比较复杂，我们还没有对怎么自动生成构词附加成分的问题进行很好的研究，所以这种转换规则具体怎么写，是还在进一步研究当中的问题。

(3) np->!v g<性> ::\$.内部结构=单词

=> NP(V/v !N/g)%V.TTT=TEMDEG,%V.CAG=CAGODO

这是汉语动词(v)加“性”构成np的分析规则及其蒙古语的转换规则。蒙古语规则的意思是，汉语动词加“性”构成的np在蒙古语里也可以转换为NP，而且这个NP是由一个V和一个N构成的。从“%V.”开始是对V进行的限制，即这里的V采取的变化形式应为动词结尾形式(TTT)之一的形动词(TEMDEG)的现在时形式(CAG=CAGODO)。该规则跟机译词典中的条目“性”发生关联，其蒙古语翻译为“CINAR”。

在汉语中不仅在动词上加语素“性”构成短语，也可以在名词、形容词(a)后面附加“性”而构成短语。所以相应的分析和转换规则还可以有两条：

(4) np->!n g<性> ::\$.内部结构=单词

=> NP(N/n !N/g) %N.CASE=HAR /*政治性*/

(5) np->!a g<性>::\$.内部结构=单词

=> NP(A/a !N/g) %A.CASE=NER /*一致性*/

蒙古语转换规则中的“%N.CASE=HAR”、“%A.CASE=NER”分别表示规则右部第一个N和A的格变化形式为领格和主格。这样，“政治性”、“一致性”就翻译成VLVS TORO-YIN CINAR 和 NIGEDULTEI CINAR。

2. “所”字结构np

“所”字结构np在政府文献中出现了诸如“所+v+的”、“np+所+v”、“np+所+v+的”等结构，比如“所起作用”、“所占比重”、“可持续发展所面临的问题”、“为…所进行的调整”等。

“所”字结构 np 的情况比较复杂，因为在汉语中“所”字前面的名词是所字后面动词的主体，但翻译成蒙古语的时候没法判断这个名词是处于领格还是向位格，如果取的是领格形式，那么它是主语，如果取的是向位格形式，那么它必是宾语无疑。这个跟“所”字后面的动词有关，比如，如果是汉语动词“说”，翻译成领格，如果是汉语动词“面临”的话，翻译成向位格。汉语完全没有这个区别，凡在“所”字前面的名词都是主体，或者是主语。

还有一个问题是，像汉语中的“所”这样有其实体，但在蒙古语中翻译不出一个词或不需要翻译出来的单位怎么处理是值得研究的问题。我们现在的处理方法是，把这些词的译词

或翻译形式写成F<_>, 即=>F<_>, 暂时称之为“空位”, 相当于数的空位“O”。

3. 联合结构 np

汉语联合结构 np 主要有: 有连词的、无连词的和含标点符号的三种类型。联合或并列的成分可以是 np 和 np, 也可以是 vp 和 np。与此相应的规则有以下几种。

(1) np->!np c np :: \$.内部结构=联合,\$.连接成分=%c,%c.连体=是,%c.yx=和|跟,%%np.内部结构=~联合,

IF %np.内部结构=定中,%%np.内部结构=定中 THEN %np.定语.cpcat=%%np.定语.cpcat ENDIF

=> NP(NP/np C/c !NP//np) /*国际形势和外交工作*/

这是用连词连接而构成的(或者说是有关标记的)汉语联合式np及其转换规则。“%c.连体=是”限制其中的连词是连接体词性成分的那种连词,比如“和、跟”等,而不能是那些连接谓词性成分的连词,如“并且、而”等。“%%np.内部结构=~联合”描述的是,第二个np不能是联合结构。它的作用是限制了“小学中学和大学”这样结构的层次组合顺序是[[小学中学]和大学],而不是[小学[中学和大学]]。从“IF”开始是一种约束条件,它限制了并列的前后两个np可以是定中式np,比如“国际形势和外交工作”。

蒙古语转换规则与汉语规则稍有区别,即并列的两个NP的后一个NP是中心成分,因为整个短语往外再构成更大的短语时需要发生相应的形态变化的是它。

(2) np->!np np :: \$.内部结构=联合,%np.内部结构=单词,%%np.内部结构=单词,%np==%%np @语义类 @内部结构 @名词子类 @代词子类 @ccat @处所单位

=>NP(NP/np !NP//np) %NP.CASE=NER/*政治思想 思想道德 专家学者 武器装备*/

这是有关无标记的联合式np的一条规则。从这个规则可以看出,对无连词或无标记的联合式np的限制条件是比较严格的:要求规则右部的前后两项np的一些属性取值相同(%np==%%np),这些属性是:“语义类”、“内部结构”、“名词子类”、“代词子类”、“处所单位”等等。

值得提出的是,要正面描述两个成分构成无标记联合结构的条件,比较困难。对此,我们采取了一个策略。其做法是,类似于区分全局规则和局部规则的方法,将全局规则中的联合式结构规则跟其他规则区分开,把联合式短语规则的优先级降低,即硬性规定:在其他规则分析不出结果的情况下,系统才调用联合式短语的分析规则。这样,系统首先会把“专家学者”分析为定中np,然后才是联合式np。这显然是不得已而为之的一种权宜之计。但目前只能是这样。讨论联合式ap和vp的时候也会遇到同样的问题。请看对短语“专家学者”的分析和转换结果:

专家学者

MERGEJILTEN=U ERDEMTEN NP:9259344[蒙古语短语分类:NP]{}(NP:9241760[蒙古语短语分类:NP, CASE: HAR]{}(!N: 9463424< MERGEJILTEN> [蒙古语词类:N, YX: MERGEJILTEN, NSUBC:NCONT, NPLUR: D, GENDER: FEMALE, CASE: HAR]{}))! NP: 9244272 [蒙古语短语分类: NP]{}(!N: 9452544< ERDEMTEN> [蒙古语词类: N, YX: ERDEMTEN, NSUBC: NCONT, NPLUR: D, GENDER: FEMALE]{})) np: 9259344 [汉语短语分类:np,内部结构:粘中]{中心语:np<学者>,定语:np<专家>}(np:9241760[汉语短语分类:np,内部结构:单

词]{定语:[],中心语:[]}(!n:9463424<专家>[汉语词类:n,名词子类:na,数量名:YES,个体量词:位|名|个,前名:YES,前动:YES,后名:YES,名状语:NO,临时量词:NO,语义类:人,yx:专家,bx:专家,音节:2,汉字:专|家]{连接成分:[]})!np:9244272[汉语短语分类:np,内部结构:单词]{定语:[],中心语:[]}(!n:9452544<学者>[汉语词类:n,名词子类:na,个体量词:个|位,前名:NO,前代:YES,语义类:身份,yx:学者,bx:学者,音节:2,汉字:学|者]{连接成分:[]}))

MERGEJILTEN-U ERDEMTEN NP:9260624[蒙古语短语分类:NP] {} (NP: 9241760 [蒙古语短语分类:NP,CASE:HAR] {} (!N:9463424<MERGEJILTEN>[蒙古语词类:N,YX:MERGEJILTEN,NSUBC:NCONT,NPLUR:D,GENDER:FEMALE,CASE:HAR] {})! NP: 9244272[蒙古语短语分类:NP] {} (! N: 9452544<ERDEMTEN> [蒙古语词类:N,YX:ERDEMTEN,NSUBC:NCONT,NPLUR:D,GENDER:FEMALE] {})) np:9260624[汉语短语分类:np,内部结构:定中]{定语:np<专家>,中心语:np<学者>}(np:9241760[汉语短语分类:np,内部结构:单词]{定语:[],中心语:[]}(!n:9463424<专家>[汉语词类:n,名词子类:na,数量名:YES,个体量词:位|名|个,前名:YES,前动:YES,后名:YES,名状语:NO,临时量词:NO,语义类:人,yx:专家,bx:专家,音节:2,汉字:专|家]{连接成分:[]})!np:9244272[汉语短语分类:np,内部结构:单词]{定语:[],中心语:[]}(!n:9452544<学者>[汉语词类:n,名词子类:na,个体量词:个|位,前名:NO,前代:YES,语义类:身份,yx:学者,bx:学者,音节:2,汉字:学|者]{连接成分:[]}))

MERGEJILTEN ERDEMTEN NP:9253776[蒙古语短语分类:NP] {} (! NP: 9241760 [蒙古语短语分类:NP,CASE:NER] {} (!N:9463424<MERGEJILTEN>[蒙古语词类:N,YX:MERGEJILTEN,NSUBC:NCONT,NPLUR:D,GENDER:FEMALE,CASE:NER] {})NP : 9244272[蒙古语短语分类:NP] {} (! N: 9452544<ERDEMTEN> [蒙古语词类:N,YX:ERDEMTEN,NSUBC:NCONT,NPLUR:D,GENDER:FEMALE] {}))np:9253776[汉语短语分类:np,内部结构:联合,复数:YES]{定语:[]}(!np:9241760[汉语短语分类:np,内部结构:单词,语义类:身份,汉语词类:n,yx:专家,名词子类:na,处所单位:NONE]{定语:[],中心语:[]}(!n:9463424<专家>[汉语词类:n,名词子类:na,数量名:YES,个体量词:位|名|个,前名:YES,前动:YES,后名:YES,名状语:NO,临时量词:NO,语义类:身份,yx:专家,bx:专家,音节:2,汉字:专|家,处所单位:NONE]{连接成分:[]})np:9244272[汉语短语分类:np,内部结构:单词,语义类:身份,汉语词类:n,名词子类:na,处所单位:NONE]{定语:[],中心语:[]}(!n:9452544<学者>[汉语词类:n,名词子类:na,个体量词:个|位,前名:NO,前代:YES,语义类:身份,yx:学者,bx:学者,音节:2,汉字:学|者,处所单位:NONE]{连接成分:[]}))

采取该策略的结果是,先把“专家学者”分析成定中式np,也就相应地翻译成定中式NP“MERGEJILTEN-U ERDEMTEN”,然后才分析成联合式np,也就相应地翻译成联合式NP“MERGEJILTEN ERDEMTEN”。

(3) np->!np w<"> np :: \$.内部结构=联合,%np.内部结构=定中,%np.内部结构=定中,\$.标点符号=%w

=>NP(NP/np W/w !NP//np)

这是用标点符号(w)连接起来的联合式np的规则。从“\$.内部结构=联合”开始是对这条规则做的整体说明。“%np.内部结构=定中,%np.内部结构=定中”说明规则右部标点符号前后的两项np可以是定中式np,比如“社会生产力、综合国力”。

(4) np->!np w<"> np :: \$.内部结构=联合,%np.内部结构=单词,%np.内部结构=单词,\$.标点符号=%w

=>NP(NP/np W/w !NP//np)

这也是有标点符号的联合式np规则。与上一个规则的不同点是，顿号前后的两个np都是单词，如“水果、肉”。

(5) np->!np w<"、"> g :: \$.内部结构=联合,%np==%g @语义类,\$!=%np @复数,\$.复数=Yes,%np.内部结构=单词|联合

=> NP(NP/np W/w !N/g)

这条规则说明的是一个np和一个语素(g)用标点符号连接而构成更大的np的情况。要求规则右部标点前面的np和后面的g在语义类属性方面相同，其中前项np可以是单词(“水果”)，也可以是联合结构(如“水果、肉”)。

(6) np->!vp np :: \$.内部结构=联合,%np.内部结构=单词,%vp.内部结构=单词

=>NP(NP/vp !NP/np)

这是关于一个vp和一个np无标记联合而构成联合式np的规则。蒙古语规则中把汉语vp转换成了NP。用这条规则可以把“运输邮电”翻译成“TEGEGEBURI SIVDAN HQLBQG_A”，其树形结构为：

NP/np(NP/vp(/v<运输>)!NP/np(/n<邮电>))

4. 定中结构 np

在汉语中，定中式 np 大致上有以下几种组合模式。下面分别描述每一种组合模式的规则。

(1) np->vp !np

这种组合有几条规则。

(1a) np->vp !np :: \$.内部结构=组合定中,\$.定语=%vp,\$.中心语=%np,%np.内部结构=定中,%vp.yx=~持续

=> NP(VP/vp !NP/np) %VP.TTT=TEMDEG,%VP.CAG=CAGODO

(1b) np->vp !np:: \$.内部结构=粘合定中,\$.定语=%vp,\$.中心语=%np,%np.内部结构=单词,%np.yx=运动,%vp.内部结构=单词,%vp.yx=健身

=>NP(VP/vp !NP/np) %VP.TTT=TEMDEG,%VP.CAG=CAGODO /*健身运动*/

(1c) np->vp !np :: \$.内部结构=粘合定中,%vp.内部结构=单词,%vp.后名=是,\$.定语=%vp,\$.中心语=%np,%np.内部结构=单词

=>NP(AP/vp !NP/np) /*综合国力*/

=>NP(NP/vp !NP/np) %NP.CASE=HAR /*发展时期*/

=>NP(VP/vp !NP/np) %VP.TTT=TEMDEG,%VP.CAG=CAGODO /*自治权利*/

汉语中有的动词能够作定语，而大部分是不行的①。比如像“企图、料想、看重”等动词不能直接修饰名词，而像“设计、研究、综合”等动词可以直接修饰名词。规则(1c)中的“后名”属性就是用来判断一个动词能否直接修饰名词的。“后名=是”的意思是可以直接修饰名词。像“发展时期、综合国力”等短语的翻译就可以用到这条规则②。规则(1b)是用来处理“健身运动”这个短语的一条局部规则。把它放在这里，是为了说明“%vp.内部结构=单词”这条约束条件的作用的。运用规则(1b)翻译“健身运动”的结果是“BEY_E

MAHABVD-I CIHIRAGJIGVLHV HODELGEGEN”（树形结构为：NP/np(VP/vp(v<健身>) !NP/np(/n<运动>)))，如果把规则中的“%vp.内部结构=单词”改成“%vp.内部结构=宾述”的话，翻译就变成“BEY_E MAHABVD-I CIHIRAGJIGVLVN_A HODELGEGEN”，显然没有走这条规则路线。从转换的角度看，汉语vp加np而构成的np可以转换成NP，但内部结构不一样，汉语的vp分别对应于蒙古语的AP，NP，VP。

(2) np->vp !n

np->vp !n :: \$.内部结构=粘合定中,%vp.内部结构=单词|联合,%vp.后名=是,%n.前动=是,\$.定语=%vp,\$.中心语=%n,

IF %vp.内部结构=单词 THEN %vp.兼类=~n ENDIF,

IF %vp.内部结构=单词,%n.yx=~会 THEN %n.音节=2 ENDIF

=> NP(VP/vp !N/n) %VP.TTT=TEMDEG,%VP.CAG=CAGODO /*学习文件*/

汉语这条规则的np与前面几条规则中的np不同，其中心语是n，而不是np。相应地蒙古语规则也有一个N。

(3) np->r !np

np->r !np :: IF %np.内部结构=单词 THEN \$.内部结构=粘合定中 ELSE \$.内部结构=组合定中 ENDIF,\$.定语=%r,\$.中心语=%np

=> NP(R/r !NP/np) /*某些领域 */

蒙古语代词（R）修饰名词的时候，或者是在名词前作定语的时候需要变什么格，取决于它的子类。若是人称代词，一般变领格，如果是指示代词，则是主格。如，“我家”翻译成“MINU GER”，“这个问题”翻译成“ENE ASAGVDAL”。有关蒙古语代词语法属性的确定，代词子类划分与语法的关系等问题的研究是薄弱环节，有待加强。

(4) np->ap !np

这种组合有几种规则。

(4a) np->ap !np :: \$.内部结构=粘合定中,%ap.内部结构=单词|联合|状中,%ap.连接成分=No,%np.中心语.内部结构=~的字,%np.内部结构=单词,%np.yx=~地,\$.定语=%ap,%ap.主体=%np,\$.中心语=%np

=> NP(R/ap !NP/np) /*所有投资*/

=> NP(NP/ap !NP/np) %NP.CASE=HAR /*大型企业*/

=> NP(AP/ap !NP/np) /*国有企业*/

(4b) np->ap !np :: \$.内部结构=组合定中,\$.定语=%ap,\$.中心语=%np,%ap.内部结构=的字,%ap.中心语.内部结构=单词

=>NP(TP/ap !NP/np)%TP.CASE=HAR

=>NP(AP/ap !NP/np)/*新的步伐 新的水平*/

=>NP(VP/ap !NP/np)%VP.TTT=TEMDEG,%VP.CAG=CAGIRE

(4c) np->ap !np :: \$.内部结构=组合定中,\$.定语=%ap,\$.中心语=%np,%ap.内部结构=

的字,%ap.中心语.内部结构=~单词

=>NP(AP/ap !NP/np) /*比较好的发展*/

(4d) np->ap !np :: \$.内部结构=组合定中,\$.定语=%ap,\$.中心语=%np,%ap.内部结构=的字

=> NP(NP/ap !NP/np) %NP.CASE=HAR /*人民群众的物质生活*/

(4e) np->ap !np :: \$.内部结构=组合定中,\$.定语=%ap,\$.中心语=%np,%ap.中心语.内部结构=联合,%ap.内部结构=的字

=>NP(NP/ap !NP/np) /*民族团结和社会稳定的局面*/

(4f) np->ap !np :: \$.内部结构=定中,\$.定语=%ap,\$.中心语=%np,%ap.内部结构=单词,%ap.ccat=v

=>NP(AP/ap !NP/np) /*多渠道*/

(5) np->mp !np

这样的组合还有几种规则。

(5a) np->mp !np :: \$.内部结构=组合定中,%np.数量名=是,%np.ccat=~r,\$.定语=%mp,\$.中心语=%np,%mp.主体=%np

=>NP(MP/mp !NP/np) /*一个新的台阶*/

=>NP(NP/mp !NP/np) /*一批矿产资源*/

(5b) np->mp !np :: \$.内部结构=定中,\$.定语=%mp,%mp.内部结构=数量,\$.中心语=%np,%np.内部结构=组合定中,%np.定语.内部结构=的字

=>NP(TP/mp !NP/np)%TP.CASE=HAR /*一九九八年政府工作的建议*/

(5c) np->mp !np :: \$.内部结构=组合定中,\$.定语=%mp,%mp.内部结构=数量,%mp.yx=一批,\$.中心语=%np,%np.内部结构=组合定中

=>NP(MP/mp !NP/np)%NP.CASE=HAR /*第十五次全国代表大会精神鼓舞*/

(6) np->mcp !np

np->mcp !np ::\$.内部结构=组合定中,\$.定语=%mcp,%mcp.ccat=m,\$.中心语=%np,%np.内部结构=单词,%np.ccat=~q

=>NP(MCP/mcp !NP/np)/*二十一世纪*/

(7) np->np !np

(7a) np->np !np :: \$.内部结构=组合定中,\$.中心语=%np,\$.定语=%np, %%np.连接成分=No,%np.内部结构=单词|粘合定中,%np.ccat=~r,%np.中心语.ccat=n,%%np.内部结构=组合定中,%%np.ccat=~q

=> NP(NP/np !NP//np)%NP.CASE=HAR /*全国代表大会精神鼓舞*/

(7b) np->np !np :: \$.内部结构=组合定中,\$.中心语=%np,\$.定语=%np,%np.内部结构=粘合定中,%np.ccat=~r,%%np.内部结构=粘合定中

=> NP(NP/np !NP//np)%NP.CASE=HAR /*计算机在企业管理中的应用 计算机对英语的分析> */

(7c) np->np !np :: \$.内部结构=定中,\$.中心语=%np,\$.定语=%np,%np.内部结构=联合,%np.ccat=~r,%np.内部结构=单词

=> NP(NP/np !NP//np)%NP.CASE=HAR /*京九、南昆铁路*/

(7d) np->np !np ::\$.内部结构=粘合定中,\$.中心语=%np,\$.定语=%np,%np.内部结构=单词

=> NP(NP/np !NP//np) %NP.CASE=HAR /*钢产量*/

(8) np->sp !np

(8a) np->sp !np :: \$.内部结构=组合定中,\$.中心语=%np,\$.定语=%sp,%np.内部结构=定中

=> NP(NP/sp !NP/np)%NP.CASE=HAR

(8b) np->sp !np :: \$.内部结构=定中/*,\$.中心语=%np,\$.定语=%sp,%sp.内部结构=联合*/

=> NP(OP/sp !NP/np) /*中西部地区*/

(8c) np->sp !np :: \$.内部结构=组合定中,%np.内部结构=~联合,%np.ccat=~r,\$.中心语=%np,\$.定语=%sp, IF %sp.内部结构=单词,%sp.ccat=s THEN %sp.后名=是 ENDIF

=> NP(NP/sp !NP/np) %NP.CASE=HAR /*本届政府 全国人民*/

=> NP(OP/sp !NP/np) %OP.CASE=HAR /*海外同胞*/

|| IF %sp.内部结构=单词,%np.内部结构=单词,%np.yx=部 FALSE

=> NP(NP/sp !NP/np) %NP.CASE=HAR /*国外机器翻译的现状 户外天线*/

|| %sp.内部结构=单词,%np.内部结构=单词,%np.yx=部

=> NP(NP/sp) %NP.CASE=HAR /*西北部 */

(9) np->tp !np

(9a) np->tp !np :: \$.内部结构=定中,%tp.内部结构=定中,%tp.后名=是,%np.语义类=~时间,%np.ccat=~r,\$.中心语=%np,\$.定语=%tp

=> NP(NP/tp !NP/np) %NP.CASE=HAR /*本届政府*/

(9b) np->tp !np :: \$.内部结构=定中,%tp.内部结构=附加,\$.中心语=%np,%np.内部结构=组合定中,\$.定语=%tp

=> NP(OP/tp !NP/np)

(10) np->m a !np

np->m a !np :: \$.内部结构=组合定中,%m.数词子类=系数,%a.音节=1,\$.中心语=%np,\$.定语=%a, IF %np.内部结构=单词 THEN %a.形容词子类=aa ENDIF

=> NP(M/m A/a !NP/np) /*两大城市 */

(11) np->c !np

np->c !np ::\$.内部结构=扩展,\$.连接成分=%c,\$.作连接=Yes,\$==%c @作照应,\$!=%np @作照应 @作连接 @连接成分,%c.yx=~和~兼~与~跟~由于~因为, %c.连体=Yes, IF %c.yx=无论|不管|不论,%np.语气=疑问 THEN %np.语气=疑问, \$!=%np @语气 ENDIF

=> NP(!NP/np S/c) /*无论什么企业 无论什么人*/

(12) np->dj !np

np->dj !np ::\$.内部结构=定中,\$.定语=%dj,\$.中心语=%np,%np.内部结构=单词

=>NP(AP/dj !NP/np) /*有中国特色的社会主义*/

(13) np->dp !np

np->dp !np ::\$.内部结构=定中,\$.定语=%dp,\$.中心语=%np,%np.内部结构=单词

=>NP(AP/dp !NP/np)/*总里程*/

除此之外, 还有以下几种短语规则。

(14) np->np !ap

np->np !ap :: \$.内部结构=定中,\$.定语=%np,\$.中心语=%ap,%ap.yx=稳定|团结|健康|安全, %np.语义类=空间|集体, %ap.内部结构=单词

=>NP(NP/np !NP/ap)%NP.CASE=HAR /*民族团结 */

=>NP(NP/np !AP/ap)%NP.CASE=HAR /*社会稳定 国家安全 身体健康 大众健康*/

此规则中的“\$.中心语=%ap”一条说明的意思是, 该名词短语的中心词不是名词, 而是形容词。“%ap.yx=稳定|团结|健康|安全”对形容词本身也进行了界定, 即作为中心词的形容词可以是那几个原形中的一个, “|”表示的是“或关系”。

(15) np->ap !mp

np->ap !mp :: \$.内部结构=组合定中,%mp.量词子类=~不定,\$.定语=%ap,\$.中心语=%mp

=>NP(OP/ap !NP/mp) /* 第八届全国人民代表大会第一次会议以来的五年 */

=>TP(TP/ap !TP/mp) /*过去的五年*/

(16) np->np !v

np->np !v :: \$.内部结构=粘合定中,%np.后动=Yes,%v.前名=是,%v.兼类=~n,\$.定语=%np,\$.中心语=%v,%np.内部结构=单词|联合

=>NP(NP/np !N/v) %NP.CASE=HAR /* 思想道德建设*/

(17) np->np !vp

(17a) np->np !vp :: \$.内部结构=定中,%vp.前名=是,\$.定语=%np,\$.中心语=%vp,%np.内部结构=单词

=>NP(NP/np !NP/vp)%NP.CASE=HAR /*市场竞争*/

(17b) np->np !vp :: \$.内部结构=定中,%vp.yx=调控,\$.定语=%np,%np.yx=宏观
=>NP(NP/np !NP/vp) /*宏观调控*/

(17c) np->np !vp :: \$.内部结构=定中,%vp.yx=投入,\$.定语=%np,%np.yx=农业
=>NP(NP/np !NP/vp)%NP.CASE=HAR/*农业投入*/

(18) np->ap !vp

np->ap !vp :: \$.内部结构=组合定中,\$.定语=%ap,\$.中心语=%vp, %ap.内部结构=的字,
%vp.内部结构=~述宾~连谓,%vp.yx=~是,%vp.兼类=~n~p~d,%vp.音节=~1, IF %vp.复数主
语=是,%ap.中心语.cpcat=~pp THEN %ap.中心语.复数=是 ENDIF /*理论与实际的结合
*/

=>VP(NP/ap !VP/vp) %NP.CASE=OGH /* <<对国有大中企业财务的> <监督> */

=>NP(NP/ap !NP/vp) %NP.CASE=HAR /* 企业的改组 */

(19) np->sp !vp

np->sp !vp :: \$.内部结构=粘合定中,%sp.后名=是,%vp.前名=是,\$.定语=%sp,\$.中心语
=%vp, IF %sp.内部结构=单词,%sp.ccat=r FALSE /*~这里工作*/

=>NP(NP/sp !NP/vp)%NP.CASE=HAR /*全国代表*/

(20) np->d !vp

np->d !vp :: \$.内部结构=粘合定中,%d.后名=是,\$.定语=%d,\$.中心语=%vp,%d.内部结
构=单词

=>NP(O/d !NP/vp) /*对外开放*/

值得说明的是,上述(14) — (20)几类短语的中心词不是名词,而是其他词类。有些汉语语法学者认为,一个短语的功能类应该跟其中心成分功能类保持一致,比如名词短语应该由名词充当中心成分,动词短语应该由动词充任中心成分③。这种说法在原则上是正确的,但实际情况并不都是那么单一。汉语中就存在短语功能类跟组成成分的功能类性质不一致的情况,这在操作上应该是允许的。

参考文献

- [1] 达胡白乙拉. 面向信息处理的蒙古语短语结构规则研究[M]. 内蒙古大学硕士学位论文, 2002.
- [2] 巴达玛教德斯尔. 汉蒙机器翻译词典中的蒙古语词语语法属性描述[J]. 民族语文, 2002, (4): 61-63.
- [3] 那顺乌日图, 刘群, 巴达玛教德斯尔. 关于“汉蒙机器辅助翻译系统”[J]. ALTAI HAKPO (JOURNAL OF THE ALTAIC SOCIETY OF KOREA), 2001, (No. 11), Seoul: Seoul National University, Korea.
- [4] 那顺乌日图, 刘群, 巴达玛教德斯尔. 面向机器翻译的蒙古语生成[J]. 张普, 曹佑琦主编. 自然语言理解与机器翻译(第六届全国计算语言学联合学术会议论文集)[C]. 北京: 清华大学出版社, 2001.
- [5] 清格尔泰. 蒙古语语法[M]. 呼和浩特: 内蒙古人民出版社, 1991.
- [6] 吕叔湘. 现代汉语八百词[M]. 北京: 商务印书馆, 1980.

[7] 刘顺. 现代汉语名词的多视角研究 [M]. 学林出版社, 2003.

[8] 詹卫东. 面向中文信息处理的现代汉语短语结构规则研究 [M]. 北京: 清华大学出版社, 2000.

注释

①据李晋霞研究, 在“V_双+N_双”(“双”指双音节)定中结构的构成上, 非心理动词(如“保护、报复”等)、行为动词(如“出租、调查”等)、非特征动词(如“活动、害怕”等)比心理动词(如“佩服、羡慕”等)、非行为动词(如“作为、等于”等)、特征动词(如“穿梭、后怕”等)更易于构成定中“V_双+N_双”。参见李晋霞(2003), 论格式义对“V_双+N_双”定中结构的制约, 载《中国语文》, 2003年第2期。

②汉语语法学界对“发展、综合”等词到底是动词, 还是动名兼类词意见不一致, 我们在这里是把它们当作动词来看待的。

③参见宋国明(1997), 句法理论概要, 北京: 中国社会科学出版社, p. 44~51。原话为: “介词组以介词为首, 名词组以名词为首, 动词组以动词为首(本书中的词首就是中心语或中心词—引用者)”。

Mongolian transformation rules for Chinese noun phrases (np)

Badma-odsar

(Mongolian Language Institute, Academy of Mongolian Studies, Inner Mongolia University, Huhhot
010021, China)

Abstract: Based on the research on Chinese nouns(n), noun phrases(np) and Mongolian translation of Chinese n,np,the author tries to formalize the Mongolian transformation rules for Chinese np,which are practically needed in Chinese-Mongolian MT studies,the grammar and translation problems,concerning the transformation of np,are also talked in this paper.The author carries out his experiment on the government-work-report corpus and “Darhan”Chinese-Mongolian translation system.

Key words: Chinese,Mongolian,noun,noun phrases, transformation rules, Chinese- Mongolian MT

收稿日期: 2003-12-03

项目基金: 国家“863”计划资助(863-306-ZT04-05-3)

作者简介: 巴达玛教德斯尔(1965-), 男, 蒙古族, 内蒙古通辽市奈曼旗人, 内蒙古大学语言博士, 副教授, 硕士生导师。从事语言学理论、蒙古语语法、蒙古文信息处理、机器翻译、对外蒙古语教学等领域的研究工作。