

# 面向信息处理的蒙古语时位词短语结构规则研究

斯琴<sup>1</sup>，张建梅<sup>2</sup>

(内蒙古大学 蒙古学学院，内蒙古 呼和浩特 010021)

**摘要：**本文面向蒙文信息处理的实际需要，尝试以形式化的方式对现代蒙古语时位词短语结构规则进行描写。概括介绍了现代蒙古语时位词短语的范围、时位词短语组成成分之间的句法关系和组合方式。用不同标准对时位词短语进行了不同层次的分类，归纳出简单时位词短语的6种结构规则和复合时位词短语的33种结构规则。总结出现代蒙古书面语时位词短语的四个特点。

**关键词：**信息处理；现代蒙古语；时位词短语；结构规则

**中图分类号：**H085.6

**文献标识码：**A

## 一、时位词短语的范围

蒙古语中，以时位词为中心词的短语称时位词短语。在时位词短语中，时位词在意义及功能方面起主导作用，这体现为时位词短语在句子中总是作为修饰成分来表示相关成分（词或短语）的时间或位置意义，并具有名词的形态变化。

在时位词短语中，时位词一般位于其它成分的后面。需要说明的是，我们认为《DOTOR\_A NI、HOJIM NI》等由时位词和语气词组成的结构是简单时位词短语、“SEDHIL DOTOR\_A NI、BIDEN-U DEGER\_E BOL”等结构是复合时位词短语。把“NASI CASI、HOYAR-VN HOGORONDO GVRBA-YIN DVMDA、HOBİ-BAN HOGEGE/JU HVLA-BAN VNV/JV YABV、NIS/HU ONGGOCA、NEYIGEM JIRVM”等视为固定短语。对固定短语的结构，本文不进行分析。由固定短语和时位词组成的短语（如：NIS/HU ONGGOCA DEGER\_E、NEYIGEM JIRVM OGEDE），按简单短语处理。

## 二、时位词短语的句法关系和组合方式

蒙古语时位词短语的句法关系以辅助关系为主，其中时位词一般处于辅助地位，而且一般位于其他成分的后面。仅在并列结构的时位词短语（JEGUN-TEI BARAGVN-TAI）、支配关系的时位词短语（JEGUN TAL\_A-YIN）、由时位词和语气词构成的时位词短语结构中，时位词才位于其它成分的前面。时位词短语中的时位词一般处于辅助地位，但它却成了时位词短语的中心词，这是传统语言学中的蒙古语词组与面向信息处理的蒙古语短语的基本区别之一。所谓短语的“中心词”是指具有本身的结构特征和语法功能的词；而不是处于被补充地位的具有独立意义词。在时位词短语中，作为中心词的时位词使其它成分以名词的主格、属格、从格、不定“N”形式或领属词尾形式变化，并本身也具有名词的主格、与格、从格及领属词尾等形式变化。

需要说明的是，除了领属词尾形式的标记引用了吉仁花所用的[YHAM1]和[YHAM2]之外，还用到了[YHAM3: := -YIN-IYEN\IYAN，/N-V-BAN，/N-U-BEN，-VN-IYAN，-UN-IYEN，-YIN-HI，/N-V-HI，/N-U-HI，-VN-HI，-UN-HI]和[YHAM4: := -ACA-BAN -ECE-BEN]等标记。《现代蒙古语文数据库》中的-TAI和-TEI形式，有时是共同格形式（BAG-VN DARVG\_A-NAR-TAI HAMTV），有时是构词附加成分（HEGEL-TEI），有时还起并列连接词的作用（BARAGVN-TAI，JEGUN-TEI）。我们在归纳时位词短语结构规则时，把以上三种情况下的-TAI和-TEI分别表示为[HAM]、[-TAI]或[-TEI]、[/TAI]或[/TEI]。另外，语料库中的动词“UGEI”的词类标注是“UIL”，但无法说明它有何动词语法形式，因此直接用“[UGEI]”表示；语气词和连词也没有形态变化，所以直接用词语本身[NI]、[BOL]、[BOGED]等来表示。

## 三、时位词短语的种类及其结构规则

以不同标准，对时位词短语进行不同层次的分类。首先，按组成时位词短语的词数，把时位词短语分为简单时位词短语和复合时位词短语两大类。其次，按被修饰词的词类对简单时位词短语分6类。第三，按被修饰短语的类别对复合时位词短语分3类。第四，按组成成分的词数和词类，对以上三类的复合时位词短语进行再分类。简单时位词短语是指由两个词组成的短语；其中，前面的词可以是名词、形容词、代词、时位词或动词，后面的词一般是时位词，有时是语气词。例如：JAM-VN HAJAGV-DV, HALAGVN DEGER\_E, ENE HOGORONDO, JEGUN TAL\_A, CILOGELEGDE/GSEN-ECE NASI , DOTOR\_A NI 等。复合时位词短语是指由短语和时位词（或语气词）组成的时位词短语。其中的短语包括：名词短语、时位词短语、动词短语等。例如：OROI-YIN HOGOLA/N-V-BAN DARAG\_A , HOMON VNTA/GSAN-V HOYIN\_A , MANGNAI DEGER\_E NI 等。

### （一）简单时位词短语的种类及其结构规则

根据简单时位词短语的前一成分的词类，蒙古语简单时位词短语可分为以下6类。

#### 1. CP1.1 {JIN CON} R01.1 (由名词和时位词组成的简单时位词短语)

这类简单时位词短语的结构规则是：

R01.1 IF {JIN[HAR] | [BUDN] | [NER] | [YHAM3]&CON[OGH] | [NER] | [GAR] | [YHAM4] | [YHAM1] | [UID]} THEN {JIN CON} CP1.1。如：OROI-YIN HIRI-DU（傍晚时分），ARAD-VN DOTOR\_A（在民间），SIREGE/N DEGER\_E（桌子上），CASV/N DOOR\_A-ACA（从雪底下），EGUDE VRVGV（向门），MORI/N DEGER\_E-BEN（在马上），HONI/N DOTOR\_A-BAN（羊群中），ECIGE-YIN-IYEN OGEDE-ECE（向其父），SIREGE/N-U-BEN EMUN\_E（在桌子前），HOTA-YIN-IYAN ARV-BAR（在浩特的后面），[ ]ARSLAN-V ARV-ACA（从阿尔斯冷的后面），HOMOS-UN ARV-BAR（在人们后面）。

#### 2. CPA1.2 {TEN CON} R01.2 (由形容词和时位词组成的简单时位词短语)

这类简单时位词短语的结构规则是：

R01.2 IF {TEN[NER]&CON[NER] | [YHAM1]} THEN {TEN CON} CP1.2。如：HALAGVN DEGER\_E（趁热），AS DEGEGSI-BEN（直上）。

#### 3. CPA1.3 {TON CON} R01.3 (由代词和时位词组成的简单时位词短语)

这类简单时位词短语的结构规则是：

R01.3 IF {TON[NER] | [GAR] | [BUDN] | [HAR]&CON[NER]} THEN {TON CON} CP1.3。如：ENE HOGORONDO（其间），EDE/N DEGER\_E（他们这里），EDE/N DEGER\_E（他们这里），BIDE/N-U DEGER\_E（在我们上面）。

#### 4. CPA1.4 {CON CON} R01.4 (由时位词和时位词组成的简单时位词短语)

这类简单时位词短语的结构规则是：

R01.4 IF {CON[NER] | [-TAI]&CON[HAR] | [NER] | [-TAI] | [YHAM1]} THEN {CON CON} CP1.4。如：JEGUN TAL\_A-YIN（左侧的），BARAGVN-TAI，JEGUN-TEI（右侧，左侧）。

#### 5. CPA1.5 {UIL CON} R01.5 (由动词和时位词组成的简单时位词短语)

这类简单时位词短语的结构规则是：

R01.5 IF {UIL[GAR]&[NER]} THEN {UIL CON} CP1.5。如：《CILOGELEGDE/GSEN-ECE NASI（解放以来），CILOGELEGDE/GSEN-ECE HOYISI（解放后）》等。

在以上五种简单时位词短语中，时位词总是位于组成成分的后面。

#### 6. CPA1.6 {CON SVL} R01.6 (由时位词和语气词组成的简单时位词短语)

这类简单时位词短语的结构规则是：

RO1.6 IF {CON[NER]&SVL[NI]} THEN {CON SVL} CP1.6。只有在这类简单时位词短语中，时位词才会位于其它成分的前面。例如：“DOTOR\_A NI（里面）、HOJIM NI（后来）”等。

需要说明的是，蒙古语中虽有“NIGE TON:02H19\* 7 DOOR\_A CON:02H19\* 7”（一块儿）等由数词和时位词所组成的结构，但因在固定短语库里有这种结构，所以，就不认为短语种类中有这种由数词和时位词所组成的短语。另外，“OROI-YIN HIRI-DU，UDE-YIN HOYIN\_A，NAMVR-VN SARA/N DOOR\_A，JIRGVGA/N JIL-UN EMUN\_E”等短语中的“OROI，UDE，NAMVR，JIL”等词在面向信息处理的词语分类来看，应是时位词。但在我们所利用的语料库中，其词类标记并不是时位词，而是名词。所以，目前就把这种短语归类于由名词和时位词组成的短语种类中。

## （二）复合时位词短语的种类及其结构规则

现代蒙古语中，复合时位词短语有由名词短语和时位词组成的复合时位词短语、时位词短语和时位词组成的复合时位词短语、动词短语和时位词组成的复合时位词短语、时位词短语和语气词组成的复合时位词短语等四种类型。在归纳这四种复合时位词短语的结构规则时，组成短语的词数及词类起决定性作用。所以，可依此特征，进一步分为三个类型加以说明。

### 1. CP2.1 {NP CON} R02.1（由名词短语和时位词所组成的复合时位词短语）

按组成短语的词数及词类的特征，可将此类复合时位词短语分为25子类。因为篇幅有限，下面仅举出6个子类的结构规则。

(1) 由名词-名词-时位词组成的复合时位词短语。其结构规则是：

RO2.1.1 IF {NP {JIN[NER] | [HAR] | [YHAM3] | [BUDN]&JIN[NER] | [HAR] | [GAR] | [BUDN] | [YHAM3] | [HAM]} &CON[NER] | [GAR] | [YHAM1]} THEN {JIN JIN CON} CP2.1.1。例：[]BOROHUU-YIN DVGVIYILANG VRVGV（向宝日夫的小组方向）

(2) 由形容词-名词-时位词组成的复合时位词短语。其结构规则是：

RO2.1.2 IF {NP {TEN[NER]&JIN[NER] | [HAR] | [BUDN]} &CON[GAR] | [NER]} THEN {TEN JIN CON} CP2.1.2。例：BODATAI AJIL DEGER\_E-ECE（在实际工作上）

(3) 由代词-代词-名词-时位词组成的复合时位词短语。其结构规则是：

RO2.1.6 IF {NP {TON[NER]&NP {TON[BUDN]&JIN[HAR]}} &CON[NER]} THEN {TON TON JIN CON} CP2.1.6。例：ENE HEDU/N JIL-UN DOTOR\_A（在这几年里）

(4) 由动词-动词-名词-时位词组成的复合时位词短语。其结构规则是：

RO2.1.10 IF {NP {VP {UIL[JER]&UIL[TEMONG]} &JIN[HAR]} &CON[OGH]} THEN {UIL UIL JIN CON} CP2.1.10。例：BOGO/JV TALBI/GSAN EBESU/N-U-BEN JAH\_A-DV（在捆好的草旁）

(5) 由动词-名词-动词-名词-名词-时位词组成的复合时位词短语。其结构规则是：

RO2.1.19 IF {NP {VP {NP {TEN[NER]&JIN[NER]} &UIL[TEMONG]} &NP {JIN[HAR]&JIN[BUDN]}} &CON[NER]} THEN {UIL JIN UIL JIN JIN CON} CP2.1.19。例：TVNVMAL GEREL CAYIRA/GSAN NAMVR-VN SARA/N DOOR\_A（在散发着清澈光芒的秋月下）

(6) 由名词-名词-名词-动词-形容词-名词-名词-时位词组成的复合时位词短语。其结构规则是：

RO2.1.24 IF {NP {NP {NP {NP {JIN[NER]&JIN[HAR]} &NP {VP {JIN[NER]&UIL[TEMONG]} &NP {TEN[NER]&NP {JIN[NER]&JIN[YHAM]}}}} &SVL[NI]} &NP {JIN[NER]&JIN[BUDN]}} &CON[GAR]} THEN {JIN JIN JIN UIL TEN JIN JIN JIN CON} CP2.1.24。例：[]LINGHW\_A HOGSIN-U TOSO BOL/O/GSAN HURENG XdALINGBVV DALINGBV% JVBCAGAN=V=HI&NI ENNGER-UN SILBI/N DEGER\_E-ECE（在莲花老人油腻的棕色褙褙布吊面皮袍前襟纽襟上）

### 2. CPB2.2 {VP CON} R02（动词短语和时位词组成的复合时位词短语）

按组成短语的词数及词类的特征，可将此类复合时位词短语分为 6 子类。

(1) 由名词-动词-时位词组成的复合时位词短语。其结构规则是：

RO2. 2. 1 IF {VP {JIN[NER] | [JIG] | [YHAM1]&UIL[HAR] | [GA]}&CON[NER]} THEN {JIN UIL CON}

CP2. 2. 1。例：[]HOMON VNTA/GSAN-V HOYIN\_A (人睡了之后)

(2) 由形容词-动词-时位词组成的复合时位词短语。其结构规则是：

RO2. 2. 2 IF {VP {TEN[JIG]&UIL[HAR]}&CON[NER]} THEN {TEN UIL CON} CP2. 2. 2。例：OLAN-I

SAGV/GSAN-V DARAG\_A (在大家坐下之后)

(3) 由名词-动词-动词-时位词组成的复合时位词短语。其结构规则是：

RO2. 2. 3 IF {VP {VP {JIN[YHAM1]&UIL[JER]}&UIL[TEMQDQ]}&CON[YHAM1]} THEN {JIN UIL UIL CON}

CP2. 2. 3。例：MAL-IYAN NEGULGE/JU IRE/HU JAGVR\_A-BAN (在将牲畜迁移过来的时候)

(4) 由代词-名词-动词-动词-时位词组成的复合时位词短语。其结构规则是：

RO2. 2. 4 IF {VP {VP {TON[NER]&VP {JIN[YHAM1]&UIL[JER]}}&UIL[/HU-YIN]}&CON[UID]} THEN {TON

JIN UIL UIL CON} CP2. 2. 4。例：BIDE MAL-IYAN HAMAGALA/JV TEMECE/HU-YIN JERGECEGE-BER  
(我们在为保护牲畜而斗争的同时)

(5) 由名词-名词-名词-动词-动词-时位词组成的复合时位词短语。其结构规则是：

RO2. 2. 5 IF {VP {NP {JIN[NER]&JIN[NER]}&VP {VP {JIN[NER]&UIL[NOHJER]}&UIL[/GSAN-

ACA]}&CON[NER]} THEN {JIN JIN JIN UIL UIL CON} CP2. 2. 5。例：NIDVNVN NAMVR HOTA  
NEYILE/JU BAGV/GSAN-ACA NASI (从去年秋比邻而居以来)

(6) 由代词-数词-名词-名词-动词-时位词组成的复合时位词短语。其结构规则是：

RO2. 2. 6 IF {VP {NP {TON[HAR]&NP {NP {TON[BUDN]&JIN[HAR]}&JIN[NER]}}&UIL[/GSEN-U]}&CON[NER]}

THEN {TON TON JIN JIN UIL CON} CP2. 2. 6。例：MAN-V TABV/N JIL-UN TOLOBLEGE  
BEYELEGDE/GSEN-U DARAG\_A (在我们的五年计划实现后)

### 3. CP2. 3 {CP SVL} RO2. 3 (时位词短语和语气词组成的复合时位词短语)

按组成短语的词数及词类的特征，将此类复合时位词短语分为以下 2 种类型。

(1) 由名词-时位词-语气词组成的复合时位词短语。其结构规则是：

RO3. 1 IF {CP {JIN[NER]&CON[NER]}&SVL[NI]} THEN {JIN CON SVL} CP2. 3. 1。例：MANGNAI

DEGER\_E NI (在其额头上)

(2) 由代词-时位词-语气词组成的复合时位词短语。其结构规则是：

RO2. 3. 2 IF {CP {TON[BUDN]&CON[NER]}&SVL[BOL]} THEN {TON CON SVL} CP2. 3. 2。例：BIDE/N-U

DEGER\_E BOL (在我们的上面是)。

## 四、结语

从以上诸结构规则中可知：第一、在由语气词和时位词（或时位词短语）组成的时位词短语中，语气词的位置在后面。其它类时位词短语中，时位词位置在后面。第二、时位词短语具有名词的形态变化形式，并且在句子里主要表示相关成分的时间和位置。第三、在复合时位词短语中，只出现了“短语+时位词”的结构，而没有出现“短语（或词）+时位词短语”的结构。第四、不论复合时位词短语由多少词组成，时位词总是与其前面的整体结构相关联。也就是说，复合时位词短语中没有“时位词+时位词”这种连续接加时位词的情况。因“BARAGVN HOYITV DEGER\_E、GER-UN BARAGVN HOYITV HAN\_A-YIN TOLOGAI DEGER\_E、DVGVIYILANG-VN BARAGVN EMUN\_E”等短语中的“BARAGVN=HOYITV、BARAGVN=EMUN\_E”等结构被看作是固定短语，所以，把这种结构就视为一个时位词来处理了。

## 参考文献

- [1] 内蒙古大学中国语言文学系蒙语教研室. 现代蒙古语[M]. 下册, 呼和浩特: 内蒙古人民出版社, 1964.
- [2] 清格尔泰. 蒙古语语法[M]. 呼和浩特: 内蒙古人民出版社, 1991.
- [3] 确精扎布. 蒙语语法研究[M]. 第一册, 呼和浩特: 内蒙古大学出版社, 1989.
- [4] 那顺乌日图. 蒙古文信息处理[M]. 赤峰: 内蒙古科学技术出版社, 1998.
- [5] 华沙宝. 蒙古语词类标注系统—AYIMAG[A]. 计算语言学文集[C]. 北京: 清华大学出版社, 1999.
- [6] 德·青格乐图. 面向信息处理的蒙古语固定词组研究[M]. 呼和浩特: 内蒙古教育出版社, 2001.
- [7] 德·青格乐图. 蒙古语固定短语语法属性库框架设计[M]. 国家社会科学基金项目(02CY007).
- [8] 吉仁花. 面向信息处理的蒙古语形容词短语结构规则研究[M]. 硕士学位论文, 内蒙古大学蒙古学学院, 2004年6月.

## A Study of Mongolian Time and Place Phrase Constructing for Information Processing

SI QIN<sup>1</sup>, ZHANG Jian-mei<sup>2</sup>

(Mongolian Language Institute, College of Mongolian studies, Inner Mongolia University, Huhhot, 010021 China)

**Abstract:** It is the target of this essay that a set of formal constructing rules of Mongolian time and place phrase is concluded for satisfying MLIP(Mongolian language information processing) practical needs. In this article, six simple Mongolian time and place phrase and thirty three complex constructing rules on the base of introducing scopes, syntactic relations, combinatorial ways and types of Mongolian time and place phrase is made up.

**Key word:** Information Processing; Contemporary Mongolian language; time and place phrase; constructing rules.

**收稿日期:** 2004-8-23;

**基金项目:** 本文得到国家社会科学基金项目《对蒙古语语料库的短语标注》(02BBY036)和国家自然科学基金项目《蒙古语短语的自动切分与标住》(60263001)资助。

**作者简介:** 1. 斯琴(1964-), 女, 蒙古族, 内蒙古奈曼旗人。内蒙古大学蒙古学学院新闻系副教授。主要从事现代蒙古语方面的研究。2. 张建梅(1975-), 女, 蒙古族, 内蒙古奈曼旗人。内蒙古大学蒙古学学院蒙古语文研究所硕士研究生。主要从事蒙文信息处理研究。