

## 《意思⇌文本》模式的深层句法结构与句法转换规则

薛恩奎

(解放军外国语学院, 河南 洛阳 471003)

**摘要:**《意思⇌文本》模式是莫斯科语义学派的核心理论,是一种对语言的编码和解码的工具。其主要的工作原理是从“意思”到“文本”,或从“文本”到“意思”对语言各个层次进行同义转换。深层句法结构是《意思⇌文本》模式中跨语际信息转换的桥梁。本文介绍部分句法转换规则与相应的词汇转换规则的操作方法与转换过程,并提出构建汉语句法转换规则和俄汉机器翻译用俄汉双语详解语义词典的设想。

**关键词:** 深层句法结构; 同义转换; 转换规则; 详解组合词典

**中图分类号:** H354.3

**文献标识码:** A

### 1 《意思⇌文本》模式的工作原理

《意思⇌文本》(«Смысл⇌Текст»)模式是莫斯科语义学派的核心理论。《意思⇌文本》模式是由 А.К.Жалковский, И.А.Мельчук 等人于 20 世纪 60 年代提出的语言总合一体描写方法。它既是一种语义学理论体系,又是一种语言描写强有力的工具。经过 И.А.Мельчук 等人近四十年的耕耘,这套理论体系在语言学的宏观领域和微观领域都取得了丰硕的成果,并广泛应用于跨语言的信息处理和机器翻译等领域,是俄罗斯当代语义学中的一枝奇葩。(张家骅, 2001)

莫斯科语义学派的理论体系区别于其他语义学理论的显著特点是:模仿人类掌握自然语言的能力,从综合与分析两个相对的方向,将人类获得语言综合能力和分析能力的过程作为模式化对象,将语言从义素到句型、从词法到句法、从语义到语用等领域研究的内容采用一致的描写原则,纳入统一的描写对象。

Ю.Д.Апресян 认为,人类掌握语言显示出 3 个方面的能力: 1) 用各种不同的同义方式表达既定意思的能力,即综合能力; 2) 从一确定的句子中获取意义,以及识别各种同音现象的能力,即理解分析能力; 3) 区别正确和错误的语言句子的能力。(Ю.Д.Апресян, 1995) 从这一意义上说,《意思⇌文本》模式现阶段还只是关于句子的语言学理论。《意思⇌文本》模式中的“文本”(ТЕКСТ)是相对于“意思”(СМЫСЛ)而言,“文本”不仅代表书面语句,还代表有声的语句,而非一般意义上的“文本”(语篇)。《意思⇌文本》模式正是根据人类掌握语言这三个方面的能力,对语言进行综合、分析和过滤。从“意思”到“文本”的综合过程,以“意思”为输入对象,对语义、句法、词法、语音从深层到表层,由下至上逐级逐层描写,是编码过程;从“文本”到“意思”的分析过程,以“文本”为输入对象,对语音、词法、句法、语义从表层到深层由上向下逐级逐层描写,是解码过程。针对人类识别语言句子错误的的能力,《意思⇌文本》模式设计了八种类型的过滤器,对各级次、各层次的综合与分析的结果进行过滤:

1) 一般的共性过滤器,剔除语义合成结果中所有包含虚拟词汇的深层句法结构

( глубинно-синтаксическая структура, ГСС ), 它们在同义词汇函数转换 (ЛФ-синонимическое перифразирование) 过程中是必要的中间过渡环节, 完成转换之后, 没有必要再进入表层句法;

- 2) 剔除同义 ГСС 集中所有包含“空位”关键词函数的 ГСС;
- 3) 保障积极语义价和句法价“饱和”过滤器;
- 4) 词和词组的组合性能限制过滤器;
- 5) 词序规则过滤器;
- 6) 表层句法成分限制过滤器;
- 7) 形态、构词限制过滤器;
- 8) 优化文本过滤器。

莫斯科语义学派的描写语言采用一套“深层句法语言”( глубинно-синтаксический язык, ГСЯ)。ГСЯ 的词汇系统由四个方面组成: 1) 自然语言的词汇语义单位; 2) 熟语; 3) 词汇函数; 4) 人造虚构词。由于深层句法语言的句法具有普遍性, 所以深层句法关系不具有语义性 (асемантическое), 只表示抽象的句法关系。具有语义内容的表层句法关系在深层句法中用一组阿拉伯数字表示概括语词之间的句法关系。深层句法语言词汇之间的句法关系较为抽象, 在向表层句法转换时, 可体现更多的不同表层句法关系。在《意思⇌文本》模式中共有六种深层句法关系, 分别用阿拉伯数字 1—6 表示, 1—4 是动词的语义配价关系(《意思⇌文本》模式中最多只考虑四价动词), 5 是广义限定关系, 6 是并列顺序关系。(张家骅, 2002)

《意思⇌文本》模式的主要工作原理就是“同义转换”( синонимичное преобразование)。所谓“同义转换”, 在该系统中是以“意思”(смысл)为“常体”(инвариант) (意思⇌文本), 由一种意义结构转换为另一种同义结构的过程。系统通过专门的语义描写语言对“意思”进行描述, 体现为语义图表 (семантический граф, СГ), 即“语义表示”(семантическое представление), 并通过词汇转换规则 (лексическое правило) 与相应的词汇函数 (лексическая функция) 进行词汇转换, 再通过句法转换规则 (синтаксическое правило) 调整相应的句法结构, 生成大量的以“意思”为原型的同义深层句法结构。

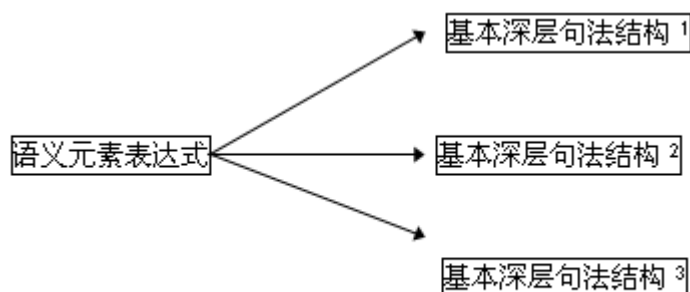
同义现象 (синонимия) 是任何语言不可避免的一种属性, 传统的语言学著作研究同义现象时, 一般把着眼点放在各种同义表达方法之间的细微意义区别上, 即研究同义现象的不同方面; 与传统方法不同的是: 《意思⇌文本》模式从同义现象的相同方面着手, 即强调各种不同表达手段的同一性方面。只要是表达“所指”(означаемое) 相同的手段, 《意思⇌文本》模式都认为同义现象。同一种现象, 不同的描写角度, 也会产生不同的效果。《意思⇌文本》模式中所指的同义现象, 或同一“所指”, 与传统意义上的同义词 (синоним) 不是一个概念。《意思⇌文本》模式中所指的同义现象是对传统意义上同义词的一种概括、抽象, 是词语之间的一种换算关系, 是对词语的一种语义解释。例如: 昨天=[一天前]; 明天=[一天后]。关键是取决于系统的定义和描写的角度。在该系统中, 动词 предшествовать 与前置词 до (перед) 是同义关系, 动词 мыть, стирать 可以认为是同义现象, 因为它们都是[借助液体使……干净]。系统认为, 同义词是一种个别现象。如果认为 X, Y 是同义词, 那么 1) 'X' = 'Y', 2) X, Y 具有相同的句法类别和句法价。(И.А Мельчук, 1995) 正是利用这种同义现象之间的转换原理, 《意思⇌文本》模式在各个级次生成大量的同义结构, 再通过各种过滤装置剔除合乎自然语言规则的结构, 筛选出合格的文本 (张家骅, 2002)。

俄罗斯科学院信息传输所计算语言学实验室基于《意思⇌文本》模式理论研制成功的多语种机器翻译系统 ЭТАП-3 是典型的转换型机器翻译系统。该系统的核心部件就是同义转换模块。(易绵竹, 薛恩奎, 2002)

《意思↔文本》模式转换系统的信息源是《俄语详解组合词典》(И.А.Мельчук, А.К.Жолковский, 1984)。在《俄语详解组合词典》中专门设有“词汇函数”字段, 标明关键词全部可能的函数取值范围和基本深层句法结构。深层语法分为词法和句法, 并配有词汇转换规则和句法转换规则。词汇转换规则根据《俄语详解组合词典》中对相应的词汇语义单位所做的“语义表示”, 将同一意思转换为各种不同词汇表达手段。句法转换规则服务于词汇转换规则。句法转换规则根据基本深层句法结构, 利用词汇转换规则提供的同一意义的各种不同词汇表达手段, 生成大量的同义深层句法结构。

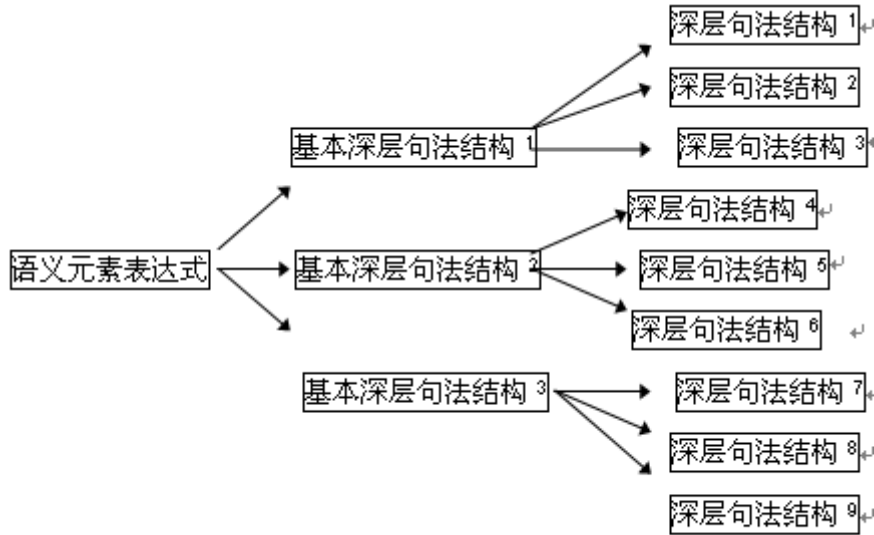
## 2 深层句法的构成

深层句法结构是“语义”与“表层”之间的一种过渡层级的表达手段。深层句法结构按其性质可分为“基本深层句法结构”(базовая ГСС)和“同义函数深层句法结构”(ЛФ-синонимичные ГСС)。基本深层句法结构是同义函数深层句法结构集中最具代表性的共性结构, 是“生产结构”(производящая структура), 而“同义函数深层句法结构”是“派生结构”(производная структура)。深层句法结构的句法表现形式(синтаксическое представление, СП)是依存关系树形图和一套形式语言。从“意思”向“文本”转换时, 深层句法结构的输入端是语义元素(семантический компонент), 即以语义图表(семантический граф)为表现形式的意义原子(смысловые атомы)连续统, 输出端是句法元素, 即表层句法结构, 其流程图: 语义元素↔深层句法表示↔句法元素。转换过程由抽象逐渐到具体。以语义图表为体现形式的语义元素是超语言的抽象结构, 还没有完成句子基本结构单位的切分, 是一个混沌的结构体。在转换为深层句法结构的同时, 也完成了句子基本结构单位的切分。句子的基本结构单位体现为深层句法树的节点, 用深层句法语言的词汇标注, 句法关系用阿拉伯数字1—6标注。同一语义元素表达式因取值不同, 句法关系也不同, 因而, 可以生成同一意义的不同深层句法结构; 每一级次的基本转换原则是: “一进↔多出”(один вход↔много выходов)(图1)。



(图 1)

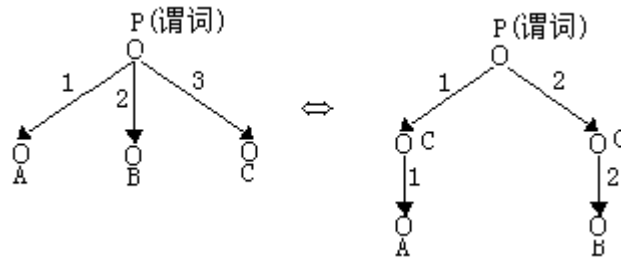
严格地讲, 从语义元素表达式到深层句法结构的第一步还不是完全意义上的转换, 而是一种语义切分, 是将以语义图表为表现形式的语义元素切分为抽象的句法结构单位。第一步得到的深层句法结构是基本深层句法结构。在第二步中基本深层句法结构变成了输入端。词汇转换规则通过各种词汇函数对基本深层句法结构树的节点词汇取值, 生成同一意思的不同词汇表达手段。再通过相应配套的句法转换规则, 调整句法结构方案, 生成更多的同一意思深层句法结构(图2)。



(图 2)

### 3 句法转换规则

深层句法转换规则 (синтаксическое правило, СП), 即转换系统中的句法规则, 它是保障一个深层句法结构转换为另一个同义深层句法结构的规则, 如同代数式中的换算律:  $(a+b)c = ac+bc$ , 相应地: 如果 A 在 C 的方面超过 B, 则 A 的 C 超过 B 的 C:



Он превосходит её по силе. (他在力气方面超过她。) ⇔ Его сила превосходит её силу. (他的力气超过她的力气。)

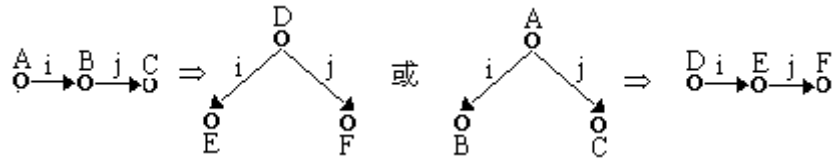
如前所述, 句法规则服务于词汇规则, 根据词汇规则的转换机制, 提供同一意思的不同词汇表达方式作相应的句法结构调整。但也有少数只作句法结构调整, 没有词汇变化的独立句法规则 (如上例)。

深层句法关系有三种类型: 谓词题元关系、广义限定关系和并列关系。转换系统根据三种不同的句法关系对句法树节点的生长作出了一些限制: 谓词的题元关系 (1—4) 一个节点只能生长出一个树枝; 并列关系 (6) 只以最左边的一个为代表与谓词发生联系, 也只有最左边具有代表性的并列成分能生长一个树枝; 标号 (5) 为广义的非同等级限定关系, 根据系统对自然语言现实文本所做的试验表明, 一个节点 (5) 的生长树枝不超过十个。因此, 关系 (5) 的一个节点的生长树枝极限为十以内。由于对节点生长树枝数量的限制, 句法转换规则的数量也得到了限制。在早期的试验系统中共有 80 条句法转换规则 (Гладкий – Мельчук, 1971), 其中:

12 条分解关系 (расщепление):  $A \Rightarrow B \overset{i}{\rightarrow} C \mid$  或  $f(A) = B$ , 或  $f(A) = C$  [ $i = 1-6$ ]

例: Он помог мне. (他帮助了我。) ⇔ Он оказал мне помощь. (他给予了我帮助。)

62 条转换移植关系 (перевешивание):



[i, j = 1—6 并且 i ≠ j 或 i = j = 5]

例: Иван выиграл партию у Джонса.(伊万赢了琼斯一盘棋。) ⇔

Джонс проиграл партию Ивану.(琼斯输给伊万一盘棋。)

6 条合成关系 (склеивание), 合成关系与分解关系之间是转换关系:  $A \xrightarrow{i} B \Rightarrow C [i = 1—6]$

例: Мы постоянно оказываем ему поддержку. (我们经常给予他支持。) ⇔ Мы постоянно поддерживаем его. (我们经常支持他。)

这 80 条规则构成了一部深层结构的“普遍句法”(универсальный синтаксис), 各种自然语言的深层句法结构都可以通过这 80 条规则进行转换。为了提高系统的直观性, 研究者对现行系统中需要多步骤操作的规则作了筛选, 剔除了一些过渡环节, 并将现行规则分为 4 大类:

1) 限定关系操作——将树顶改写为限定关系:

通过词汇转换规则将关键词 ( $C_0$ ) 转换为相应的限定关系词或词组, 句法转换规则对句法结构作相应的调整, 将原句的相关题元 (A、B 或 C) 提升为谓词:

(1) 题元 A 改写为谓词, 由原来的谓词派生出同义熟语:

Рост цен сопровождается ( $C_0$ ) кризис. (危机伴随着物价上涨。)

$C_0 \Leftrightarrow Adv_{0A}(C_0)$  [Adv 表示由  $C_0$  派生的与  $C_0$  同义的熟语; 下标 A 表示将  $C_0$  的第一题元提升为谓词, 0 表示零价的无题元: *сопровождать* → *вместе с*]

$$A \xleftarrow{1} X \xrightarrow{2} (B) \Leftrightarrow A \xleftarrow{5} Y \xrightarrow{2} (B)$$

Цены растут вместе с кризисом. (危机加剧的同时物价也上涨。)

(2) 题元 B 改写为谓词, 由原来的谓词派生出同根副词:

Он поспешил ( $C_0$ ) выйти. (他着急出去了。)

$C_0 \Leftrightarrow Adv_{1B}(C_0)$  [Adv( $C_0$ )——由  $C_0$  派生的同义副词; 下标 B 表示将  $C_0$  的第二题元提升为谓词, 1 表示说明第一题元: *поспешить* → *поспешно*]

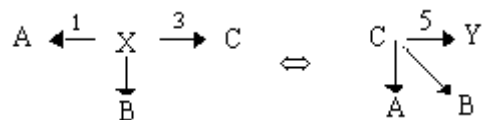
$$A \xleftarrow{1} X \xrightarrow{2} \Leftrightarrow A \xleftarrow{1} B \xrightarrow{5} Y$$

Он поспешно вышел.(他匆忙出去了。)

(3) 题元 C 改写为谓词, 由原来的谓词及第二题元派生出熟语:

Он закончил ( $C_0$ ) поход посещением Самарканда.(他以参观撒马尔汗结束了旅行。)

$C_0 \Leftrightarrow Adv_{1c}(C_0)$  [ $Adv(C_0)$ ——由  $C_0$  派生的同义熟语；下标 C 表示将  $C_0$  的第三题元提升为谓词，1 表示说明第一题元：закончить поход  $\rightarrow$  в конце похода ]

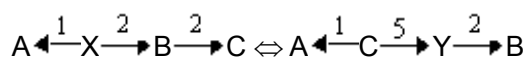


В конце похода он посетил Самарканд. (旅行结束时他参观了撒马尔汗。)

(4) 题元 C 改写为谓词，由原来的主体性熟语派生出限定性熟语：

Она сдержала своё обещание прийти( $C_0$ ). (她遵守了自己的许诺，来了。)

$Real_1(C_0) \Leftrightarrow Adv_{2B}Real_1(C_0)$  [ $Real_1(C_0)$ 表示  $C_0$  的“执行者”； $Adv_{2B}Real_1(C_0)$ 为复合函数，将第一次派生的结果再次派生出熟语；下标 B 表示同义熟语是由  $C_0$  的“执行者”的动词派生而来，2 表示熟语是句子的第二题元：сдержать обещание  $\rightarrow$  в соответствии с обещанием ]



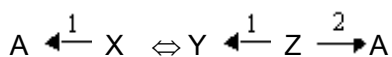
Она пришла в соответствии со своим обещанием. (她按自己的诺言来了。)

2) 合成、分解操作——将一个节点分解为两个节点，即一个关键词汇语义单位分解成同义的两个词汇语义单位；或将两个节点合成为一个节点，即两个词汇语义单位合成为同义的一个词汇语义单位：

(1) 将关键词( $C_0$ )分解成同义的名词和相关的动词：

Он тоскует ( $C_0$ ). (他很忧伤。)

$C_0 \Leftrightarrow S_0(C_0)Func_1(S_0(C_0))$  [复合函数， $S_0(C_0)$ 表示由  $C_0$  先派生出同义名词，再由  $Func_1(S_0(C_0))$  派生出相应表示“属于，有，来自”的动词：тосковать  $\rightarrow$  одолевать, тоска]

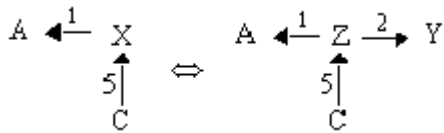


Его одолевает тоска. (他很忧伤。)

(2) 关键词的义素由两部分组成：Copul( $C_0$ )+  $C_0$  [某种行为的人+相应系词=实施某种行为、职业的动词= Pred( $C_0$ )]:

Он учительствует( $C_0$ ) Рязани. (他在梁赞当教师。)

$Pred(C_0) \Leftrightarrow Copul(C_0) \xrightarrow{2} C_0$  [Pred 是个特殊的动词函数，表示  $C_0$  相当于“系词+名词”，相当于系词函数 Copul (Pred = Copul+ $C_0$ ): учительствовать  $\rightarrow$  работать, учитель ]

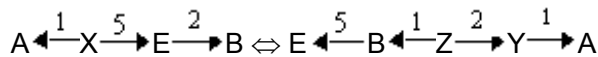


Он работает учителем в Рязани. (他在梁赞当教师。)

(3) 由关键词(C<sub>0</sub>)派生出表示 [实施相应行为地点] 的名词(S<sub>loc</sub>), 并由(S<sub>loc</sub>)派生出相应系词与具体的称名单位构成种属关系。

Гражданин Петров постоянно проживает (C<sub>0</sub>) в Москве. (公民彼德洛夫常住莫斯科。)

$C_0 \xrightarrow{5} \text{Loc}_{in} \Leftrightarrow S_{loc}(C_0) \xleftarrow{2} \text{Copul}(S_{loc}(C_0))$  [Loc<sub>in</sub> 为处所前置词函数, S<sub>loc</sub>(C<sub>0</sub>) 表示由 C<sub>0</sub> 派生出同义处所名词, 再由 Copul(S<sub>loc</sub>(C<sub>0</sub>)) 派生出相关的系词: проживать → местожителство, являться]

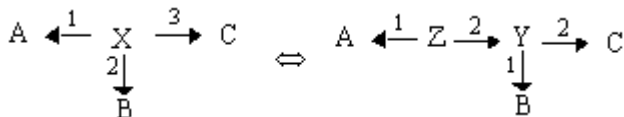


Постоянным местожителством гражданина Петрова является Москва. (公民彼德洛夫的常住地是莫斯科。)

(4) 由关键词(C<sub>0</sub>)所及间接对象分离出特征与特征的承载者 (Oper<sub>1</sub>), 并改变相应题元的关系。

Мы считаем (C<sub>0</sub>) её честной девушкой. (我们认为她是个诚实的姑娘。)

$C_0 \Leftrightarrow S_2/A_2(C_0) \xleftarrow{2} \text{Oper}_1(S_2/A_2(C_0))$  [S<sub>2</sub>/A<sub>2</sub>(C<sub>0</sub>) 表示派生出与 C<sub>0</sub> 第二题元相关的形容词和名词, 再由 Oper<sub>1</sub>(S<sub>2</sub>/A<sub>2</sub>(C<sub>0</sub>)) 派生出形容词特征的“承载者”: считать → кто, честный]

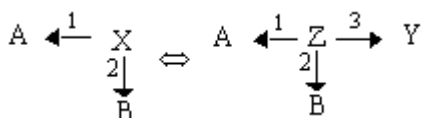


Мы считали, что она честная девушка. (我们认为, 她是个诚实的姑娘。)

(5) 将关键词(C<sub>0</sub>)的直接客体分解为种属概念和类属概念词, 并改变它们的题元关系。

Я лечу (C<sub>0</sub>) Петра. (我给彼德看病。)

$C_0 \Leftrightarrow S_2(C_0) \xrightarrow{3} \text{Labor}_{21}(S_2(C_0))$  [S<sub>2</sub>(C<sub>0</sub>) 表示派生出与 C<sub>0</sub> 第二题元相关的名词, Labor<sub>21</sub>(S<sub>2</sub>(C<sub>0</sub>)) 是由该名词派生的动词, 并将该名词与 C<sub>0</sub> 的直接客体相连接, 下标 21 表示将 C<sub>0</sub> 的直接客体改写为结构的主体语义题元: лечить → иметь кого кем ]



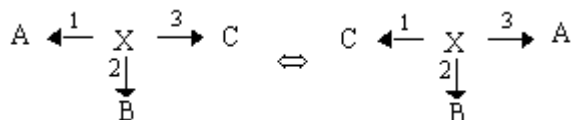
Я имею Петра своим пациентом. (彼德是我的病人。)

3) 变换层次操作——改变相应题元关系:

(1) 将关键词(C<sub>0</sub>)的主体题元与客体题元调换位置, 改变句法关系。

Партизаны оказали карателям упорное сопротивление(C<sub>0</sub>). (游击队员们顽强抵抗宪兵。)

Oper<sub>1</sub>(C<sub>0</sub>) ⇔ Oper<sub>2</sub>(C<sub>0</sub>) [Oper<sub>1</sub>(C<sub>0</sub>)表示主体语义题元与 Oper<sub>2</sub>(C<sub>0</sub>)客体语义题元调换语义关系: оказать... → натолкнуться на... ]

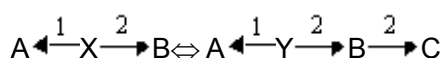


Каратели натолкнулись на упорное сопротивление партизан. (宪兵遭到游击队员们的顽强抵抗。)

(2) 由关键词(C<sub>0</sub>)派生出关键词的行为所产生的结果名词, 并与主体语义题元连接, 将原来的直接客体改变为间接客体。

Он копирует (C<sub>0</sub>) документ. (他拷贝文件。)

C<sub>0</sub> ⇔ S<sub>res</sub>(C<sub>0</sub>)  $\xleftarrow{2}$  Oper<sub>1</sub>(S<sub>res</sub>(C<sub>0</sub>)) [S<sub>res</sub>(C<sub>0</sub>)表示派生出由 C<sub>0</sub> 行为所产生的结果名词, 并由 Oper<sub>1</sub>(S<sub>res</sub>(C<sub>0</sub>))派生出与该名词相关的动词: копировать → снимать, копия ]

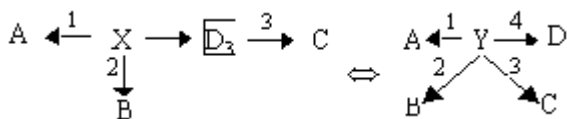


Он снимает копию с документа. (他拷贝文件。)

(3) 从关键词(C<sub>0</sub>)中分解出隐含的题元, 并将原题元 (3) 改为题元 (4)。

Иван берёт (C<sub>0</sub>) землю у Петра. (伊万向彼德租了一块地。)

C<sub>0</sub> ⇔ S<sub>0</sub>(C<sub>0</sub>)  $\xleftarrow{3}$  Labor<sub>12</sub>(S<sub>0</sub>(C<sub>0</sub>)) [S<sub>0</sub>(C<sub>0</sub>)表示由 C<sub>0</sub> 派生出与 C<sub>0</sub> 相关的性质名词, 并作为第三题元通过 Labor<sub>12</sub>(S<sub>0</sub>(C<sub>0</sub>))与 C<sub>0</sub> 相连接: брать → брать, аренда ]



Иван берёт землю в аренду у Петра. (伊万租了彼德一块地。)

(4) 在对关键词(C<sub>0</sub>)进行语义分解的同时, 相应的限定成分也变换层次, 改变说明对象。

Противник сопротивлялся (C<sub>0</sub>) двое суток. (敌人抵抗了两天两夜。)

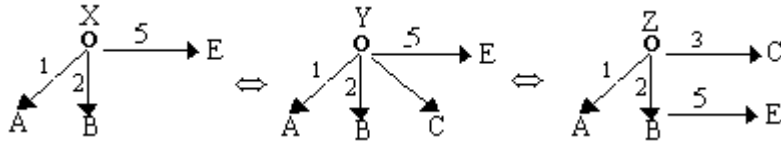
C<sub>0</sub> ⇔ S<sub>0</sub>(C<sub>0</sub>)  $\xleftarrow{2}$  Oper<sub>1</sub>(S<sub>0</sub>(C<sub>0</sub>)) [S<sub>0</sub>(C<sub>0</sub>)派生出与 C<sub>0</sub> 同义名词, 再由 Oper<sub>1</sub>(S<sub>0</sub>(C<sub>0</sub>))派生出相关动词作为该名词的主体语义题元。因该主体语义题元不处于句法主体位, 因此相关的限定成分也改变说明对象: сопротивляться (Adv) → оказывать, сопротивление(Adv)]



$$A \xleftarrow{1} X \xrightarrow{5} E \Leftrightarrow A \xleftarrow{1} Y \xrightarrow{2} B \xrightarrow{5} E$$

Противник оказывал сопротивление двое суток. (敌人进行了两昼夜抵抗。)

但是有些谓词，在语义分解之后，仍可以和相关的限定成分搭配，即两种同一意思的不同句法结构，如 Мы постоянно поддерживаем его. (我们经常支持他。) 可分解成：

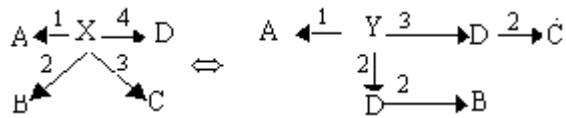


Мы постоянно оказываем ему поддержку. (我们经常给予他支持。)

Мы оказываем ему постоянную поддержку. (我们给予他经常性的支持。)

(5) 绝大部分句法规则都与相应的词汇规则配套使用，但有少数句法规则因没有词汇的变化，只改变句法关系结构层次，所以可单独使用。

Он сравнил мясо с рыбой по цене. (他比较肉和鱼的价格。)



Он сравнил цену мяса с ценой рыбы. (他对比肉的价格和鱼的价格。)

4) 转换操作——并列节点之间相互调换位置，有单层转换和双层转换。

(1) 第一题元与第二题元调换，改变句法关系。

Учёный проводит (C<sub>0</sub>) опыты. (科学家在做实验。)

$C_0 \Leftrightarrow \text{Conv}_{21}(C_0)$  [Conv 转换函数，表示变换语义题元关系，下标 21 表示 C<sub>0</sub> 的第一题元与第二题元变换位置：проводить → проводиться]

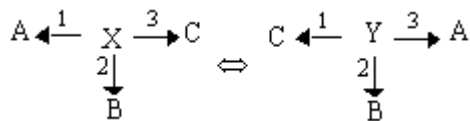
$$A \xleftarrow{1} X \xrightarrow{2} B \Leftrightarrow B \xleftarrow{1} Y \xrightarrow{2} A$$

Опыты проводятся учёным. (实验由科学家做。)

(2) 第一题元与第三题元调换，改变句法关系。

Пётр купил (C<sub>0</sub>) книгу у Ивана. (彼德买了伊万一本书。)

$C_0 \Leftrightarrow \text{Conv}_{321}(C_0)$  [Conv 同上，下标 321 表示第一题元与第三题元变换位置：купить → продать]

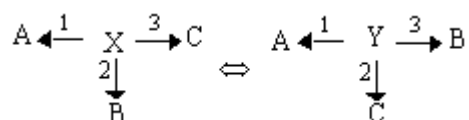


Иван продал книгу Петру. (伊万卖给彼德一本书。)

(3) 第二题元与第三题元调换，改变句法关系。

Петя обучает (C<sub>0</sub>) их языку. (别佳教他们语言课。)

C<sub>0</sub> ⇔ Conv<sub>132</sub>(C<sub>0</sub>) [Conv 同上, 下标 132 表示第二题元与第三题元变换位置: обучать → преподавать]

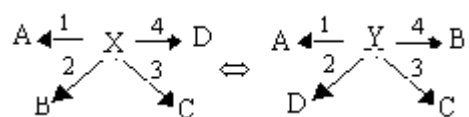


Петя преподаёт им язык. (别佳给他们上语言课。)

(4) 第二题元与第四题元调换, 改变句法关系。

Он купил (C<sub>0</sub>) у Петра книгу за 15 рублей. (他以 15 卢布向彼德买了一本书。)

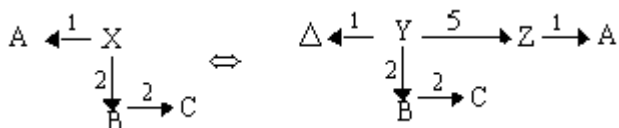
C<sub>0</sub> ⇔ Conv<sub>1432</sub>(C<sub>0</sub>) [Conv 同上, 下标 1432 表示第二题元与第四题元变换位置: купить → заплатить]



Он заплатил Петру за книгу 15 рублей. (他付给彼德 15 卢布买了一本书。)

(5) 第一题元变为广义限定语 (工具), 相应的谓词变为无人称形式, 用 (△) 表示。独立句法规则。

Снаряд оторвал (C<sub>0</sub>) ему ухо. (弹片削掉了他的耳朵。)

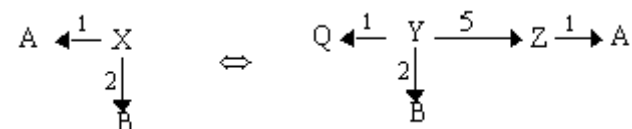


Снарядом ему оторвало ухо. (他的耳朵被弹片削掉了。)

(6) 第一题元变为广义限定语 (地点), 相应的谓词变为不定人称形式, 用 (Q) 表示, 关键词一般是国家、组织等的名称。

Франция (C<sub>0</sub>) запустила спутник. (法国发射了一颗人造卫星。)

C<sub>0</sub> ⇔ Loc<sub>in</sub>(C<sub>0</sub>)  $\xrightarrow{2}$  C<sub>0</sub> [Loc<sub>in</sub>(C<sub>0</sub>) — 由 C<sub>0</sub> 派生的处所前置词, 并将 C<sub>0</sub> 作为直接支配对象, C<sub>0</sub> 由主体语义题元变为处所: Франция → в, Франция]



Во Франции запустили спутник. (在法国发射了一颗人造卫星。)

(7) 关键词 (C<sub>0</sub>) 的限定主体语义题元转换为收受客体语义题元, 并改变相应的句法关系。

Его роман описывает (C<sub>0</sub>) жизнь индейцев. (他的小说是描写印第安人的生活。)

Func<sub>2</sub>(C<sub>0</sub>) ⇔ Func<sub>1</sub>(C<sub>0</sub>) [Func<sub>2</sub>(C<sub>0</sub>) 将 C<sub>0</sub> 的限定主体语义题元转换为 Func<sub>1</sub>(C<sub>0</sub>) 所属

客体语义题元: описывать → принадлежать ]

$$C \xleftarrow{1} A \xleftarrow{1} X \xrightarrow{2} B \Leftrightarrow B \xrightarrow{2} A \xleftarrow{1} Y \xrightarrow{2} C$$

Ему принадлежит роман о жизни индейцев. (描写印第安人生活的一本小说是他写的。)

(8) 主体语义题元与客体语义题元转换, 并改变相应的句法关系结构。

В поддержку закона выступают (C<sub>0</sub>) центристы. (中立派支持这项法律。)

Oper<sub>1</sub> (C<sub>0</sub>) ⇔ Oper<sub>2</sub> (C<sub>0</sub>) [Oper<sub>1</sub> (C<sub>0</sub>)关键词主体语义题元, 通过 Oper<sub>2</sub> (C<sub>0</sub>)转换为客体语义题元: выступать→пользоваться]

$$A \xleftarrow{1} X \xrightarrow{2} B \xrightarrow{2} C \Leftrightarrow C \xleftarrow{1} Y \xrightarrow{2} B \xrightarrow{1} A$$

Закон пользуется поддержкой центристов. (这项法律得到中立派的支持。)

在莫斯科语义学派的模式理论中, 转换系统, 无论是词汇转换规则, 还是句法转换规则, 尽管具有高度的概括性和抽象性, 但绝不是纯形式的相互转换, 而是对现实交际情景的具体刻画, 为具体条件下的现实情景提供更多的情景选择参数。尽管人类语言各不相同, 但人类生存的世界是共同的, 这是人类相互交流的共同基础。莫斯科语义学派的同义结构转换理论, 是现实交际情景的概括和抽象, 为跨语际的自动化信息交流创造了条件。不过, 《意思⇔文本》模式理论与实用技术是在形态变化丰富的俄语基础上提炼出来的, 对印欧语系的大部分语言具有广泛的适用性。但汉语属于分析型语言, 是一种孤立语。汉语句法缺乏形态控制, 语序的变化往往引起句法关系和语义的变化。汉语中, 词序的变化不仅会引起整个句子意义的变化, 甚至一个词中的义素次序变化, 不仅义素值发生变化, 同时也会引起整个词意义的变化。例如: “案犯—犯案”中“犯”的意义有别, 前者指“罪犯”, 后者指“发生”; “人工—工人”中“工”, 前者指“一个人一个工作日的劳动”, 后者指“工作”(曹炜, 2001)。因此, 《意思⇔文本》模式的句法转换规则除少数情况外, 大部分句法规则对汉语句法不完全适用。如:

(各族人民支持这项法律。) Oper<sub>1</sub> (C<sub>0</sub>) ⇔ Oper<sub>2</sub> (C<sub>0</sub>)  
 $A \xleftarrow{1} X \xrightarrow{2} B \xrightarrow{2} C \Leftrightarrow C \xleftarrow{1} Y \xrightarrow{2} B \xrightarrow{1} A$  (这项法律得到各族人民的支持。)

汉语的各句法单位之间、各词类之间没有明确的形式界限, 但这并不意味着汉语句法结构没有一定的结构规律。语序是汉语的重要语法手段, 语序的不同会导致语法关系、语义关系、甚至语用含义的差异。但在形态发达的俄语中, 语序的变化对句子的整体含义同样具有重要的作用。两种语言之间的差别不在语序本身, 而在句法联系, 并因此而产生的句法单位再组合。所谓形态控制语法, 主要是句法单位因为有形态标记, 变换位置以后仍然与原来的句法单位保持着句法联系的轨迹, 不会导致与新环境中的句法单位发生新的联系。而汉语句法单位位置的变化, 因没有形态标记, 会失去与原句法单位的联系, 很容易与新环境的句法单位建立新的句法联系, 构成新的组合, 从而导致整个句子的语法关系、语义关系的变化, 甚至会导致无序排列——“非句”。认识并把握这一点非常重要, 汉语缺乏形态变化有可能促使我们摆脱形态学语法观念的干扰, 探索出一套适合汉语特点的分析方法和描写工具, 建立我们自己的方法论体系。

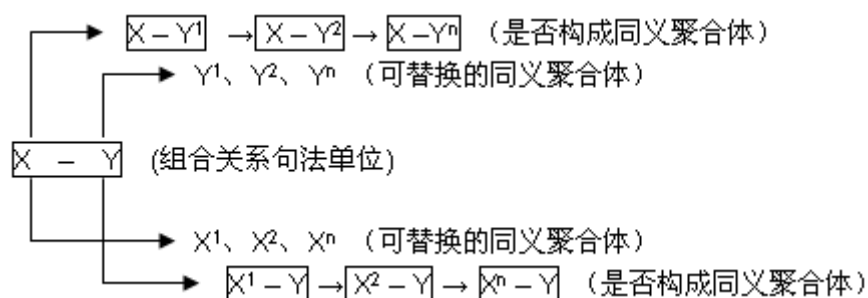
#### 4 面向俄汉机器翻译的汉语研究设想

莫斯科语义学派的语言一体化描写体系主要由“词典”和“语法”两部分组成。《俄语详解组合词典》在一体化描写的框架内, 用形式化的方法从语言的综合过程和分析过程两个方向, 对词语的义类关系、构词关系、搭配关系、句法关系展开系统的描写。其中既有语言的微观信息, 又有宏观信息, 不仅是语言的知识库, 也是使用语言的知识库。这种把词汇知识、句法知识、语义知识和语用知识融为一体纳入词典的编纂范畴是一种全新的视角, 对汉语的语

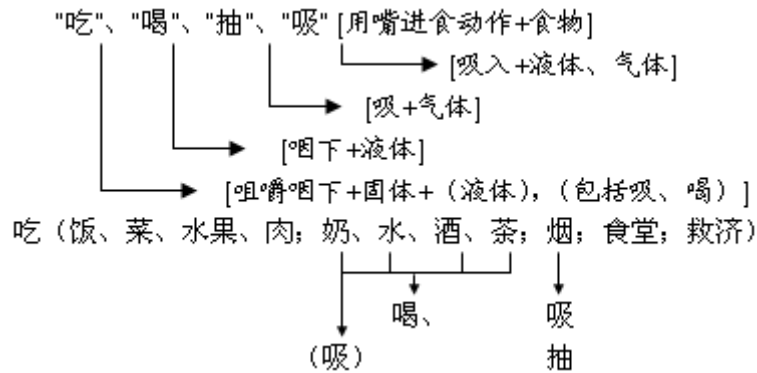
法、语义、语用三个平面各抒己见的研究方法颇有借鉴作用。语言作为人们的交际工具，其本身并没有什么语法、语义、语用之分，是一个浑然的整体。语言学家为了研究语言，将其从交际活动中分离出来，并加以简化、抽象，形成了语言学理论。针对汉语缺乏形态特征的实际，我们认为，完全可以运用莫斯科语义学派的一体化描写理论和实用工具，对汉语进行对比分析研究。从微观领域入手，以解决宏观问题。把词的义素关系、搭配关系与句子的组合关系归入一个平面加以研究，以词为主线。一个词往往可以指称一个生活片段，构成一个“事件”。把词的特征纳入句子的框架中来描写，以探索句法规则的各种必需参数。句法规则只要能保障信息传递具有一定的确定性，达到一定的可理解度就可视为合格的规则。自然语言有其复杂性和模糊性，句法规则不可能具有绝对性。《俄语详解组合词典》是一个综合的语言知识规则库，同时也是一个受限的语言规则知识库。实际上我们人类的知识在自然界面前，在自然语言面前都是受限的。针对俄汉机器翻译的语言学研究，不仅需要一部俄语详解词典，更需要一部汉语详解词典。我们应该从务实的角度对汉语的微观领域展开系统的研究。中国人民大学《多国语言机器翻译系统》课题小组花费四五年的时间以汉语动词为专门研究对象，编辑出版了一部《现代汉语动词大词典》。这部词典对二千多现代汉语常用动词的义项以及与周围名词性成分的语义组合关系逐一作出了详尽的描写。这一研究成果在国内外产生的影响充分证明了它的研究路线与方法的正确性，得到了国内外专家的广泛好评。“《现代汉语动词大词典》的编辑出版对我国的语言信息处理事业作出了开拓性的贡献”。（黄昌宁，《现代汉语动词大词典》序，1994）

语际间的信息转换（即翻译）至少涉及两种或两种以上的自然语言，转换的基础是依据存储在机器中的“词典”。这种“词典”与人们通常使用的“词典”有本质的不同。对人工翻译来讲，除了通常的“词典”，还有一部存储在脑子里的“使用语言知识词典”。供计算机自动翻译用的“电子词典”不仅需要相关语言的“语言知识词典”还需要一部相关语言的“使用语言知识词典”。两部“词典”应是同一框架内有机联系的网络化规则库。转换的基本条件，一是要在两种自然语言之间建立一座“桥梁”，这座“桥梁”就是两种语言转换的“等价物”。它一般是由在某种自然语言的基础上提炼出来的、表达一确定思想的“意思结构”（*смысловая структура*）构成的体系，一般称之为“中介语”（*язык-посредник, interlingua*）。这种“中介语”不仅具有本语言的共性特征，还应具有一般自然语言的共性特征，具有普遍性。我们称这样的“中介语”为“一级中介语”；二是要对每一参与转换的自然语言展开系统的描写，提炼出由具有本语言共性特征的“意思结构”构成的“中介语”，与转换系统的

因此，我们认为，面向计算机语言信息处理的汉语研究，应在词和句子两个层面展开系统的规约化研究。不仅需要《现代汉语动词大词典》那样对具体词的组合关系逐一作详尽的描写，还应对构成组合体的各要素的同义聚合体，以及同义聚合体之间的替换对句法关系的影响等进行规约化描写。假设 X、Y 可以构成一个组合关系的句法单位，那么 X 对 Y 的句法语义特征有选择性，Y 对 X 也同样具有选择性。X 是否可以被  $X^1$ 、 $X^2$ …… $X^n$  替换，Y 是否可以被  $Y^1$ 、 $Y^2$ …… $Y^n$  替换。这种关系可表示为：



如汉语动词“吃”、“喝”、“抽”、“吸”都具有 [用嘴进食] 相关动作的共性特征，同时每个词又具有自己的个性特征和各自的搭配范围，以及可替换的可能性：



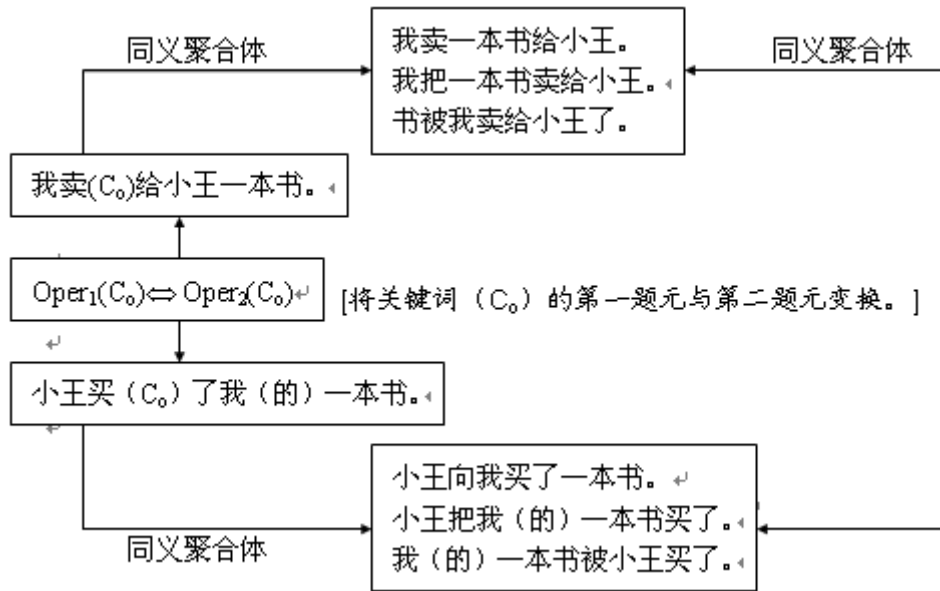
句子层面的研究应把握两个方面的变化规则。首先是在没有词汇语义单位变化的情况下，句法单位序列变化对整个句子的句法、意义的影响；其次是同义词汇语义单位替换对整个句子的句法、意义的影响。例如：

1) 语句内部要素的次序变化：

①我送他一本书。→②我把书送给他了。→③我把书送他了。→④我送一本书给他。→⑤书我已经送他了。→⑥书已被我送给他了。①

三价动词“送”的最大变换式为六种，实际上可看成是四种，因为②、③和⑤、⑥中的介词是可选项：“常体”式、“把”式、“被”式、“介词”式。变换式的意义没有变化，但是否具有普遍性，须进一步研究；

2) 同义聚合体转换：



语句转换之后意义没有发生变化，只是语义价有所不同。但是这种转换规则对汉语是否具有普遍性，我们不敢断言。语句结构是由具有一定功能的词汇语义单位纵向的配列（聚合关系的选择）和横向的搭配（组合关系的选择）构成的线性序列单元，是显性的线性序列体现的隐性立体网络。针对俄汉机器翻译的汉语研究，须在句法领域探索出适应汉语特点的规则化系统与俄语句法规则系统接轨。

《意思↔文本》模式的“深层句法”是在《俄语详解组合词典》对俄语语言知识翔实分析、

描写的基础上提炼、概括、抽象出来的一套词汇转换规则和句法转换规则，以及配套的词汇函数。这套“语法体系”经多语种机器翻译系统 ЭТАП-3 的实践证明是有效的，成功的，并具有普遍性。其理论体系和大部分词汇转换规则以及配套的词汇函数可直接用于汉语词汇的义素分析和组合规律分析，是建构俄汉机器翻译用双语词典可借鉴的实用工具。

#### 附注

①例句选自《现代汉语动词大词典》(1994)，北京语言学院出版社，第34页。

#### 参考文献

- [1] Апресян Ю. Д. 1995 Интегральное описание языка и системная лексикография [M]. М.
- [2] Апресян Ю. Д., Цинман Л. Л. 1998 Перефразирование на компьютере // Семиотика и информатика вып. 36 [C]. М.
- [3] Апресян Ю. Д. 1974 Лексическая семантика [M]. М.
- [4] Жолковский А. К. 1964 О правилах семантического анализа // Машинный перевод и прикладная лингвистика №8 [C].
- [5] Жолковский А. К. 1964 Предисловие к сборнику «Машинный перевод и прикладная лингвистика» // Машинный перевод и прикладная лингвистика №8 [C].
- [6] Жолковский А. К., Мельчук И. А. 1969 К построению действующей модели языка «Смысл↔Текст» // Машинный перевод и прикладная лингвистика №11 [C].
- [7] Жолковский А. К., Мельчук И. А. 1967 О семантическом синтезе // Проблемы кибернетики вып. 19 [C].
- [8] Мельчук И. А., Жолковский А. К. 1984 Толково-комбинаторный словарь современного русского языка [M]. Вена.
- [9] Мельчук И. А. 1999 Опыт теории лингвистических моделей «Смысл↔Текст» [M]. М.
- [10] 曹 炜 2001 现代汉语词义学 [M], 学林出版社。
- [11] 林杏光等主编 1994 现代汉语动词大词典, 北京语言学院出版社。
- [12] 易绵竹, 薛恩奎 2002 一种与 UNL 接口的机器翻译系统 ЭТАП-3 概要 // 机器翻译研究进展 [C], 电子工业出版社。
- [13] 章宜华 2002 意义—篇章模型及其释义功能 [J], 现代外语, 第 1 期。
- [14] 张家骅 2001 莫斯科语义学派 [J], 外语研究, 第 4 期。
- [15] 张家骅 2002 莫斯科语义学派的“预设”观 [J], 外语学刊, 第 2 期。
- [16] 张家骅 2002 “词汇函数”的理论和应用 [J], 外语学刊, 第 4 期。

## The Deep Syntactic Structure and Its Transformational Rules of the 《Meaning↔Text》 Model

XUE En-kui

( PLA University of Foreign Languages, Luoyang 471003, China )

**Abstract:** The 《Meaning⇔Text》 Model is the nucleus of the Moscow semantic school. It is a tool of encoding and decoding language. The main principle of the model is to produce for any given meaning all the texts that can serve as its expressions and for any given text – the meaning(s) it conveys. The model consists of several successive levels. One of the levels is dwelt upon at the set of rules for transforming synonymous deep syntactic structures into one another. The article describes the working process of the deep syntactic rules with the aid of lexical rules and assumes some linguistic applications for creating Russian-Chinese machine translation system.

**Key words:** deep syntactic structure; synonymous transform; syntactic rules; explanatory combinatorial dictionary

收稿日期: 2003-07-06

基金项目: 教育部人文社会科学研究重大项目 (01JAZJD740006), 黑龙江大学俄语语言文学研究中心承担。

作者简介: 薛思奎, 男, (1954 - ), 安徽庐江人, 解放军外国语学院教授, 黑龙江大学博士研究生, 主要研究方向: 计算语言学、俄语语义学、计算机辅助外语教学多媒体应用技术。

**[责任编辑:靳铭吉]**