

俄语动词义素聚合场与组合场的建立

薛恩奎 *易绵竹 南振兴

(解放军外国语学院计算语言学研究室 河南 洛阳 471003)

(*黑龙江大学俄语语言文学研究中心, 哈尔滨, 150080)

E-mail: xek@sinohome.com; yimianzhu@china.com

摘要: 动词是句子的 DNA, 义素是动词的 DNA。对这两类 DNA 的分析, 应以动词本身的 DNA 分析为先导。它不仅涉及动词作为一种词类的内部结构, 还涉及句法层面的方方面面, 与超句统一体也有不可忽视的联系。本文依据俄语动词本身的义素特征, 对动词的概念聚合类型和句法组合能力及其相互关系作了概括分析描述, 提出“先聚合、后组合”的技术路线。作者认为, 动词的语义聚合与语义组合分析是自然语言信息处理中最为关键的基础性工作, 也是目前自然语言信息处理领域的瓶颈。

关键词: 动词义素; 聚合场; 组合场; 自然语言处理

Building the Paradigmatic and Syntagmatic Fields of Russian Verbs' Sememes

Xue Enkui *Yi Mianzhu Nan Zhenxing

(Section of Computational Linguistics PLA Foreign Languages University, Luoyang 471003)

(*Research Center for Russian Language & Literature, Heilongjiang University, Harbin, 150080)

E-mail: xek@sinohome.com; yimianzhu@china.com

ABSTRACT: The analysis of the verb, equivalent to the DNA of a sentence, should precede the analysis of the sememe, comparable to the DNA of the verb. Factors involved in this approach include the internal structure of the verb itself, various aspects on the syntactic level, and the super-phraseological entity which is a closely-related issue. Based on the sememic features of the Russian verbs, this paper describes in general the conceptual paradigmatic types of verbs, syntactic constituent capability, and the mutual relations between them, proposing a method of paradigmatic priority over syntactic constituency. The central theme is that the analysis of semantic paradigm and syntagm of verbs, the bottleneck in NLP, form the groundwork in natural language information processing.

KEY WORDS: the sememe of verbs; semantic paradigm and syntagm; NLP

1、语义聚合场与组合场概述

语义场理论(The Theory of Semantic Fields)是德国学者特雷尔(J.Trier)最早提出来的,其主要论点是,词在语义上是相互联系的,并共同构成一个完整的系统。只有把语言中的词

汇放入一个完整的系统，通过分析词与词之间的语义关系，才能确定一个词的真正涵义。这一理论受到语义学界的普遍接受，对语义学研究具有较高的理论意义和实用价值。

但是，特雷尔的语义场理论是针对词汇语义学而提出的，是关于词汇语义分类和确定词汇语义类别层级关系的理论。我们这里所说的语义聚合场和组合场是基于动词固有的语义特征，揭示句法语义层面上所显示的聚合关系和组合关系。动词最大的特点是具有两个层面的意义特征：①体现情景要素的概念特征：行为、状态、关系等；②体现情景参与者及其特征、功能角色等。动词的物质意义特征反映了动词的所指特征，即受所指特征制约的客观内容，它具有一定的等级层次性；而动词的过程性意义特征反映在事态、事件与现实的联系之中，它是一种粘合要素，不具有等级层次性。语义聚合场是基于动词所指意义特征类型而形成的**框架结构场**，语义组合场则是基于动词过程性意义特征，能够反映某类意义特征句法表现的形式化**参数结构场**。当我们要表达某种“思想”时，实际上在我们的头脑中已经形成了关于某种“思想”的抽象概念和表达意向。概念之间总是存在上下位的包含与隶属关系。意向的选择是自上而下的纵向行为，是一个由抽象到具体的过程。形式化表达结构是横向的线性序列，序列之间是独立的平等关系，不存在上下位包含关系。抽象的概念、意向在选择表达形式的过程中不断具体化，使表达意向与表达形式“魂体相附”，最终生成可被理解的语言表达式，从而达到交际的目的。

我们采取纵向概念意义选择与横向句法表达结构相匹配、且概念意义优先的路线，它与现实言语交际过程是一致的。质言之，交际的目的就是表达或传递某种信息，形式结构只是表达、传递信息的手段而已。俄裔加拿大籍著名语言学家梅里丘克(И.А.Мельчук)长期致力于建立《意思 ↔ 文本》(Смысл ↔ Текст)的语言学模型，他认为言语交际过程有三个重要方面：①被传递或感知的信息 — 意义内容；②负载信息的物理符号 — 形式文本；③信息与信号之间的协调关系 — 转换编码。语言学家只能从大量的言语活动中观察语言功能产生的结果，而不能观察语言功能产生结果过程的转换机理。我们不可能打开“黑匣子”(черный ящик)去分析言语的过程机理，而只能通过研究已产生的言语结果，尽量去理解、描述言语背后产生言语结果的语言系统装置的运作机制。因此，基于纵向概念类别选择和横向组合模式建构的方法描写动词的语义结构，对于揭示语言系统机制和面向计算机的词汇语义学的研究都具有现实意义。

2、 语义聚合场类型

自然语言中大多数句子都是以动词为中心展开的，而动词最本质的特征是表示运动类概念。动词与静词的本质区别就是动词既包含聚合意义要素，又包含组合意义要素。根据聚合意义要素，动词构成聚合概念场；根据组合意义要素，动词构成句子的形式语义模式 — 组合场。静词只包含聚合意义要素，静词聚合意义要素是构成词汇语义场的主要依据。所以，静词不能独立构建句子的形式模式。静词的聚合意义要素只具有填充动词语义框架中情景参与者语义角色槽的功能，是动词语义框架的填充物，使动词潜在的组合意义要素具体化。例如：俄语动词 поступать, держаться (表现, 举止)是“表现”类动词的基准词，这类动词都具有抽象的意义要素“表现”，根据这一抽象的意义要素，可构成独立的“表现”类动词子场。根据它们的组合意义要素，形成组合序列，由此建立抽象的形式语义模式：

概念：“表现”（поступать）——> [表现行为 + 表现行为主体 + 表现行为评价]

形式化语义模式中动词行为的主体位和评价位必须由各种类型的静词填充，行使其各自的职能，使动词的组合意义要素具体化，使形式化的语义模式与客观现实产生联系并生成表示具体含义的句子。动词是表示“过程”（процесс）的词类，“存在”（бытийность）与“变化”（изменчивость）是动词最本质的概念义素特征，也是极为抽象的概念义素，其具体内容只有通过众多词位（lexeme）的义素分析才能得以确定。但是，概念具有不同的层级，我们可以根据动词包含概念义素的抽象程度，将动词分为不同的层级类别：

一级抽象动词：静态动词（статические глаголы）和动态动词（динамические глаголы）。

“动”和“静”是现实事、物事件在人们头脑中形成的高度抽象的模糊概念意义，二者构成世界素朴的语言图景（наивная языковая картина мира）；

二级抽象动词：“行为动词”（глаголы действия）、状态动词（глаголы состояния）和关系动词（глаголы отношения），是人们对客观世界事物存在、变化及运动方式的初步类化；

三级抽象动词：“运动”（движение），“活动”（деятельность），“思维”（мышление），“举止表现”（поведение），“平等关系”（равенство），“超越关系”，（превосходство）“符合关系”（соответствие），“领属关系”（принадлежность），“制约关系”（обусловленность），“情感”（чувство），“空间状态”（положение в пространстве），“外部变化”（изменение внешнего вида），“物理状态”（физическое состояние）等，是人们对事物运动规律与语言表达系统之间相互联系的细化与操作，是形成动词词汇意义的初始元素。

可见，三级抽象动词抽象程度最低，它们已获得相当具体的语义特征，许多语言学著作中把此类动词称为“词汇语义群”（Лексико-семантические группы，简称 ЛСГ）。

归属于某一类 ЛСГ 的动词自身都具有一套区别于其他类 ЛСГ 的“非实体义素”（несубстанционные семы），由此构成区别于其他类的聚合体。同一种 ЛСГ 之间不构成纵向的聚合层级关系，也不构成相互依存的横向链状序列，而是平等地从属于本类的“核心义素”（ядерная сема）。例如，运动动词的 ЛСГ 包含以下一些非实体义素：

“Среда”（环境）——> лететь（飞），идти（走），плыть（游）；
“Способ”（方式）——> идти（步行），ползти（爬行），ехать（乘车）；
“Интенсивность”（强度）——> бежать（跑），нестись（漫步）等等。

上面所区分的“环境”，“方式”，“强度”等非实体义素均从属于此类 ЛСГ 的核心义素“运动”。因此，非实体义素是动词词汇意义的最基本组成部分。需要强调的是，不同类型的 ЛСГ 之间会存在一些交叉现象，为使分析系统具有良好的操作性，对交叉现象可作两类处理，也可并作一类处理，但必须保持系统内部体系分类遵守唯一性原则。对个别受上下文表层结构制约的动词可作个别处理。如 растворить（打开）只能与“门”，“窗”类词汇搭配组合，对这类动词的处理需要考虑表层句法结构的制约性。

此外，实体义素（субстанционные семы）是动词词汇句法意义的组成部分，是组合性义素。该类义素的词汇意义就在于体现情景参与者的“品质”、“功能”、“关系”、“数量”等物质性意义特征，使参与者的物质性意义特征在一定的语境中具体化、实际化。

3、语义组合场模式

客观世界中万事万物都存在某种相互关系，关系与物质是互为存在的统一体。没有关系以外的物质，也没有物质以外的关系。关系是在一定的物质、环境、范围之内形成的。关系与物质、现象均为现实之物。《知网》(How-net) 创立者董振东先生明确提出，意义是关系的系统，关系是词汇语义的灵魂，只有通过“关系”才能教会计算机懂得或在某种程度上懂得“意义”，也只有通过“关系”才可能教会计算机对意义进行运算。

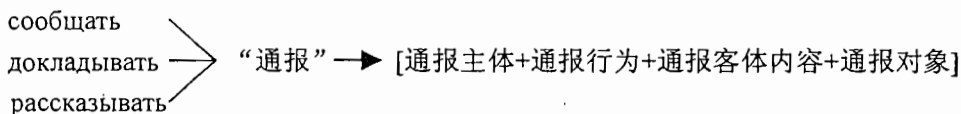
语言反映客观现实脱离不开一定的环境 (среда)，正所谓“无水不存鱼”。言语交际情景是实现言语交际必不可少的环境场所，也是为言语交际设定的封闭范围 (有些动词当其在超出一定情景范围时，词义会有所改变，组合模式也有所不同)。所谓“形式化语义模式”

(формально-семантическая модель)，实际上是根据动词组合义素特征人工刻画的交际情景框架，情景参与者、抽象位 (позиция) 及其相互间的抽象关系是其主要组成要素。情景参与者本身具有实体的物质性，而建立抽象位与实体参与者之间的联系正是动词所固有的实体义素。因此，动词的实体义素也是一种组合义素。动词的“过程”性意义正是表现在它的“物化” (материализация) 过程中，通过“动”与“静”的组合，生成具有现实意义的句子。静词只有在进入动词框架，填充到相应的“位”之后才能行使相应的功能，发挥其相应的作用。动词的组合义素决定情景参与者的数量，即语义价，亦称语义位或语义角色 (关于角色的数量和功能分类，作者将另文论述)。组合义素分为可选义素和必选义素。**必选义素**是语言结构单位，是隐性的义素。如 Он поступал *правильно* (他做得对)，动词 поступать 包含的隐性义素表示的评价意义特征 *правильно* 不能从动词的词汇意义中得到反映。**可选义素**是言语结构单位，是显性的、自由的义素，是具体话语结构意义的扩展。

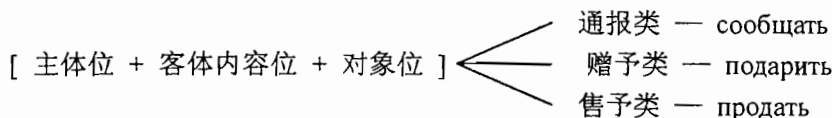
根据必选组合义素，俄语动词可分为：一价，二价，三价。四价以上的动词为数不多，尽管按动词的组合义素分类都属于必选价，但有些价位仍属于可选价。因为在现实交际中有三个价位出现就能完全满足交际需要，在特定的句子中其余价位应视为冗余价位，不可能所有的必选价在一个句子中同时出现。但是，哪些价位应该出现，哪些价位不应该出现，则取决于上下文对某些价位的制约，这就是所谓的“语用价”要讨论的问题。至于语用价问题可参照俄语句法实义切分理论，在此不作赘述。

值得注意的是，对于聚合义素和组合义素都相同的动词，或只是聚合义素相同，或只是组合义素相同的动词在系统分类时，可按聚合分类，也可按组合分类。例如：

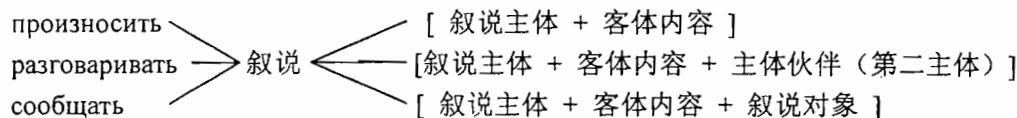
◆聚合义素、组合义素都相同的动词：



◆只是组合义素相同的动词：



◆只是聚合义素相同的动词：



正因为有这种交叉的现象，在实际言语交际中广泛存在着同义结构相互替换的可能性。

我们建立俄语动词义素聚合场和组合场的技术路线是：正确认识“语言 \leftrightarrow 思维 \leftrightarrow 现实”三者之间的辩证关系，将语言看作是现实的编码系统，充分揭示世界的概念图景到语言表达的映射关系。我们以概念为聚焦点，在确定合适的动词概念类之后，再寻绎相应的语言表达式。

4、 俄语动词概念意义类别及其形式组合模型

为了减少动词语义分类系统的运算量，我们首先根据组合义素必选项对动词进行配价分类，再根据概念意义建立聚合场，每一聚合类都有相应的语义框架模型：

(1) 一价动词的 ЛСГ:

- ①ЛСГГ: “存在” быть, существовать \rightarrow [存在 + 存在主体];
- ②ЛСГГ: “物理或心理状态” устать, работаться \rightarrow [状态 + 状态载体]
- ③ЛСГГ: “状态变化” таять, расти \rightarrow [变化 + 变化主体]
- ④ЛСГГ: “能力” летать, плавать \rightarrow [能力 + 能力载体]
- ⑤ЛСГГ: “工作活动” учиться, работать \rightarrow [活动 + 活动主体]
- ⑥ЛСГГ: “声响” шуметь, журчить \rightarrow [声响 + 声响源]
- ⑦ЛСГГ: “光亮” светить, мерцать \rightarrow [光亮 + 光亮源]

(2) 二价动词的 ЛСГ:

- ①ЛСГГ: “创造” производить, создавать \rightarrow [创造行为 + 创造主体 + 创造结果];
- ②ЛСГГ: “感知” ощущать, видеть \rightarrow [感知行为 + 感知主体 + 感知对象];
- ③ЛСГГ: “发现” открыть, найти \rightarrow [发现行为 + 发现主体 + 发现对象];
- ④ЛСГГ: “位移” нести, тащить \rightarrow [位移行为 + 位移主体 + 位移对象];
- ⑤ЛСГГ: “运动” идти, подниматься \rightarrow [运动行为 + 运动主体 + 运动方向];
- ⑥ЛСГГ: “存在” находиться, статья \rightarrow [存在行为 + 存在主体 + 存在地点];
- ⑦ЛСГГ: “举止表现” поступать, держаться \rightarrow [表现行为+表现主体+表现评价];
- ⑧ЛСГГ: “领属关系” иметь, принадлежать \rightarrow [领属关系+领属主体+领属客体]。

(3) 三价动词的 ЛСГ:

- ①ЛСГГ: “通报” рассказывать \rightarrow [通报行为+通报主体+通报信息内容+通报对象];
- ②ЛСГГ: “给予” подарить \rightarrow [给予行为+给予主体+给予物+给予对象];
- ③ЛСГГ: “放置” положить \rightarrow [放置行为+放置主体+放置客体+放置方位];
- ④ЛСГГ: “逻辑评价” считать \rightarrow [评价行为+评价主体+评价对象+评价结果];
- ⑤ЛСГГ: “情感关系” относиться \rightarrow [关系行为+关系主体+关系对象+关系态度]。

此外，俄语中零价动词为数不多，且大部分表示自然现象，可视作整体结构单独处理：

Светаёт. (天亮。) Вечереет. (天黑。) Морозит. (天气上冻。) Моросит. (下毛毛细雨。)

Гремит гром. (打雷) Сверкает молния. (闪电。) Идет дождь. (下雨。)

5、 余论

总之, 动词是一种极其复杂的词类, 其语义特征涉及语言系统的各个层面, 除本文分析的聚合特征与组合特征外, 动词的时向特征、体特征也是语义研究的重要课题。我们知道, 俄语的时间表示有一套完整的词汇系统, 弗谢沃洛多娃 (М.В.Всеволодова) 曾以两本专著专门论述俄语时间系统。但是, 任何事物有其内在联系, 必有外在联系。时间系统与动词时向、体、态等范畴有着密切的关系 (有关于这方面的内容, 作者将在专著《俄语动词》中作详细论述)。对于动词的概念义素与组合义素, 必须纵横统筹考虑, 综合分析, 才能揭示其本质特征。由于篇幅所限, 我们只对动词的上述义素进行了概括性的描述。为使我们的分析系统进一步完善, 在具体动词的分类描述、特殊动词的处理方面还有许多需要解决的问题。我们近期工作的目标是建立相对完整的俄语动词语义分类及编码体系, 为研制面向俄文信息处理的通用语义码模式奠定基础。

参 考 文 献

- (1) 伍谦光编著. 语义学导论. 湖南教育出版社, 1992.
- (2) 董振东、董强. 关系—词汇语义的灵魂. 《第二届汉语词汇语义学研讨会论文摘要》. 北京大学计算语言学研究所, 2001. 05. 14-18.
- (3) 易绵竹. 语义信息代码化的操作原理. 《外语学刊》, 2001年第2期.
- (4) Русская грамматика. Том I. Изд. Наука. -М., 1980 .
- (5) Гайсина Р.М. Лексико-семантическое поле глаголов отношения в современном языке. Изд. Саратовского университета, 1981 .
- (6) Мельчук И.А. Опыт теории лингвистических моделей 《Смысл ⇔Текст》. - М., 1999 .
- (7) Васильев Л.М. Исследования по семантике. Вып.2. Уфа, 1976 .
- (8) Всеволодова М.В. Способы выражения временных отношений в современном русском языке. Изд. Московского университета, 1975.