

俄语动词 бороться 的句法语义能力构建

孙爽

(东北林业大学, 哈尔滨 150040)

提 要: 句法语义详解组合词典是机器翻译系统必不可少的组成部分。本文根据 С. Ниренбург 和 В. Раскин 的本体论语义学, И. Мельчук 《意思↑文本》语言学模型理论和 Ю. Апресян 的语言整合描写与系统词典学理论, 试图建立可应于俄汉机器翻译系统的句法语义详解组合词典。本文在对谓词进行语义分类的基础上, 对谓词 бороться 的义项做详细的句法语义描述。

关键词: 俄汉机器翻译; 句法语义; 详解组合词典; бороться

中图分类号: H35

文献标识码: A

1 引言

在自然语言处理系统中, 句法语义详解组合词典是必不可少的重要组成部分, 详解组合词典通过属性描述包含了基本的句法语义信息, 直接影响到该系统的性能与译文质量。因此, 详解组合词典的构建历来受到计算机专家和语言学专家的高度重视。

在国外的研究中, 西方国家以 WordNet、MindNet、FrameNet 为代表的知识库构建中, 都是根据词义来描述词汇信息而建立语义词典。(Miller, 1990; Richardson, 1998; Bake, 1998) 俄罗斯研发的 ЭТАП-1, ЭТАП-2 和目前尚在研发的 ЭТАП-3 是基于 И. Мельчук 《意思↑文本》语言学模型理论和 Ю. Апресян 的语言整合描写与系统词典学理论开发而成, 通过建立一种统一的、可全面地揭示语言单位意义和词义之间本质差异的描写性语言, 为词典编撰服务。(Апресян 1989, 1995; Мельчук 1984)

近年来, 董振东研发的知网 (How Net) (董振东 1998), 鲁川、林杏光编撰的《人机通用现代汉语动词大词典》, 黄曾阳创立的概念层次网络 (HNC) 理论等, 都为汉语语义分类和语义角色描写建立了相应的解释模型。

在国内外现有研究的基础上, 就目前俄汉机器翻译系统的译文质量来看, 计算机处理文本意义的的能力仍有待提高。因此, 解决语义问题不仅是人类社会交际的迫切需要, 也是计算机对自然语言理解的迫切需要。本文顺应现代语言学和自然语言处理的发展需要, 深入探讨句法语义研究的紧密联系。在深入研究谓语动词和名词性成分之间语义关系的基础上, 以便为俄汉机器翻译提供强有力的语言学支持。

2 动词语义分类及属性标注

动词语义分类是建立详解组合词典的主要内容之一, 也是一直困扰学者们的一个主要问题。于鑫指出, 建立全面、客观的词汇语义分类系统是词汇语义学, 及至句子语义学研究的

基础。动词和名词分类系统的建立对于语言研究和自然语言处理都具有重要的意义。(于鑫 2008)词的语义分类是标明一个词的语义属性的常用手段之一。(詹卫东 1997) 动词是一个句子的中心,主要是用来表示事件和情景。基于共性特征的动词语义分类,学者们根据不同的标准,从不同的侧面已作过一些探索。传统的动词语义范畴有存在、状态、关系、行为等。H. C. Авилова 将动词分为“行为、过程、状态、性质、处所、潜能、联系、结果”八个语义类别。(Авилова 1976: 151) Z.Vendler 将英语动词分为四种基本类型:状态动词 (states)、无界活动动词 (activities)、有界行为动词 (accomplishments)、瞬间结果动词 (achievements)。(Vendler 1967: 103) Ю.С. Маслов 根据动词的词汇语义与体的句法语义之间的深层联系,将俄语动词分为单体未完成体、单体完成体和对偶动词三大类型。M. B. Всеволодова 把动词首先分为独立意义动词和构造动词,而独立意义动词又分为行动动词、存在动词、状态动词、关联动词和性质动词,构造动词又分为系动词、调节动词、引导动词和补偿动词。

但上述语义分类只能得到信息颗粒度较大的信息,远远满足不了信息处理的需要。词语的个性比较突出,同属于一个语义类的词语,其内部属性及其搭配要求却有很大的差别。语义分类能够满足解释上的充分性,但不能满足描述上的充分性。(俞士汶 2000) 对于每一个词语或概念而言,分类的方法给出的是一个“单一标记”;而属性标注的方法给出的是一个“复杂特征标记”。分类的方法的着眼点是上下位的树状关系;而属性标注的方法的着眼点是多元、多层次的网状关系。(董振东 2001) 因此,我们在为俄汉机器翻译系统研制详解组合词典时,考虑借鉴更具探索价值的自然语言处理原理。目前,S.Nirenburg 和 V.Raskin 的本体论语义学在构建语义机制方面提出了极具参考价值的研究成果。他们所提出的本体论语义学是一种关于自然语言意义的理论,一种关于自然语言加工的方案,它把经构造而成的世界模型或本体论作为提取和表述自然语言文本意义的基本框架,作为从文本中推出知识的前提。在语义模型中建立本体论图示,即一种概念框架。本体论提供的是描述一种语言的词汇单元的意义所需的元语言,以及说明编码在自然语言表征中的意义所需的元语言。

(Nirenburg 2004) 在我们构建俄汉详解组合词典的过程中,将根据本体论语义学中所定义的元语言进行语义属性描述。以 бороться 为例,бороться (с кем-чем) 表示某人同某人或某物反抗、抵抗、反对、阻挠的意思,бороться 属于行为活动动词 (SOCIAL-ACTIVITY) 中的反抗类动词 (COUNTERACTION), 同类的常用动词还有 биться, воевать, драться, конкурировать, соперничать, соревноваться 等等。бороться 还有另外一个支配形式,即 бороться за что 表示为达到某种目的或结果而做的努力 (ACHIEVEMENT OF AIMS OF PURPOSES), 同类的动词如 добиться, завоевать, заработать, преодолеть, стремиться 等等。

3 бороться 的词条描述

范晓曾指出,理想的动词配价词典应是某种语言生成句子的词典。(范晓 1996: 7) 因此,我们在构建该词典时,对动词的形式特征和句法模式尽量进行更为详细的描述,并对语义范围进行更为全面的标注。构建该词典根据莫斯科语义学派的整合性语言学描写和系统性词典学研究成果 (Мельчук 1974; Апресян 2009), 借鉴本体论框架图型,在语义分类的基础上,依靠属性描述来刻画每一个谓词,完善详解组合词典的构建,以期更好地为自然语言处理系统提供高性能和准确度高的保障。我们根据前期的研究,结合俄语的语言特点,对俄语谓词采用句法语义一体化描写,将句法语义信息整合在一起。在词典结构设计上,汲取对自然语言进行整合描写的理念,将词典和语法融为一体。一个义项构成一个词条,把谓词的每一个义项看成一个结构体,每个词条都有清晰的标准结构,包括词的形态信息、支配模式及属性限制等信息,支配模式同时体现被释的句法题元和语义配价,所提供的具体实例可直观反映该词组合搭配的关系,为正确分析判断该类句子起到重要作用。鉴于篇幅所限,本文

仅以动词 *бороться* 的词条描述为例，展示详解组合词典的设计。我们前期的研究发现，在从俄语国家语料库中检索的 500 个带 C+N5 结构的例句中，在 196 个结构受动词、动名词、或形容词中心词的支配的例句中，包含 *бороться с кем-чем* 结构的句子共有 15 个，占有一定比重。（孙爽，2003）可见，动词 *бороться* 是较为重要的一个动词，研究该词具有重要的意义。以下即为动词 *бороться* 的词条描述。

БОРОТЬСЯ
 БОРОТЬСЯ-1
 POR: V
 TRANS: 角力比赛（摔跤）
 SYNT: Ø /intr
 SEM: COUNTERACTION
 MY_1

| | | |
|--------|------------------------------------|---------------------------------------|
| 句法题元 | 1 = X (NP _{nom, sg/pl}) | 2 = Y [Ø / (c NP _{ins})] |
| 语义角色 | 1 – Agent_1 | 2 – Ø / Agent_2 |
| 语义属性描述 | 1 – HUMAN, ANIMAL | 2 – HUMAN, ANIMAL |

例句：

Мальчики долго боролись, но все-таки Андрюша одолел Игоря, неожиданной и резкой подножкой повалив на землю.（两个小孩许久缠抱在一起，但还是安德烈战胜了伊戈尔，用突然，猛的一记脚绊儿摔倒在地上。）¹

上表为 *бороться* 的支配模式 1，表格横栏分别表示句法题元、语义题元和语义受限。纵栏表示在不同的句法结构中句法题元、语义题元和语义受限的具体对应。*Бороться₁* 用莫斯科语义学派元语言解释为：X *бороться₁* = X 比赛（力气），从中可以看出，*бороться₁* 可以搭配一个名词性成分，即“主体”配价，配价的要求是“人或动物”，且为复数，如上例；*бороться₁* 也以搭配两个名词性成分，即一个“主体”和一个“第二主体”²，与前置词 *с* 连用，表示“和谁”。

БОРОТЬСЯ-2
 POR: V
 TRANS: 战斗/抗争
 SYNT: intr
 SEM: COUNTERACTION
 MY_2

| | | |
|--------|---------------------------------|---|
| 句法题元 | 1 = X (NP _{nom}) | 2 = Y (c NP _{ins}) |
| 语义角色 | 1 – Agent | 2 – Antagonist |
| 语义属性描述 | 1 – HUMAN, ANIMAL, ORGANIZANION | 2 – HUMAN, ORGANIZANION, DISEASE, NATURE, ARTIFACT, EMOTION, STATION, ACTION, ILLNESS, IDEOLOG, POLITY, CHARACTER, ABSTRACTIO |

例句：*Они борются с терроризмом.*（他们同恐怖主义斗争。）

Бороться₂ 用莫斯科语义学派元语言解释为：X *бороться₂* с Y = X 同 Y 斗争，即 X 和 Y 在某方面对立，相互冲突，一方力求在智力、精神或专业等方面战胜另一方。从中可以看出，*бороться₂* 有两个语义配价。主体 X 可为表人、表动物、表某组织的名词，客体可为表人、动物、某组织、某中感情、状态、行为、意识形态、疾病、政治制度或条例（*бороться с штампами*）、性格或其他抽象概念（*бороться с крайностями*）等意义的名词。

БОРОТЬСЯ-3

POR: V

TRANS: 克制 / 消除

SYNT: intr/pos

SEM: COUNTERACTION

МУ_3

| | | |
|--------|---------------------------------------|--|
| 句法题元 | 1 – NP nom | 2 – (с NP ins) / (против NP pos) |
| 语义角色 | 1 – Agent | 2 – Contractor |
| 语义属性描述 | 1 – HUMAN, ORGANIZANION, MATTER | 2 – FEELING, ILLNESS, SOCIAL_CONDITION, PHENMENON |

例句: *Мысли вслух помогли ей бороться со сном.* (一边想, 一边喃喃自语帮助她消除了睡意。)

бороться₃用莫斯科语义学派元语言解释为: X бороться₃ с Y / X бороться₃ против Y = X克制Y, 即X抵抗某种感情或状态Y, 努力战胜并消除它。在该结构中, бороться₃有两个语义配价, 分别为“主体”和“客体”配价。主体为“人”或“某物质”例如: *Витамины С и Е, являясь антиоксидантами, активно борются с процессом старения.* (维他命C和E是抗氧化剂, 能有效地防止老化。)中的витамины С и Е。客体为表示某种感情、疾病、社会状态、现象等意义的名词。

БОРОТЬСЯ-4

POR: V

TRANS: 为反对...而斗争

SYNT: intr

SEM: COUNTERACTION

МУ_4

| | | |
|--------|----------------------------|---|
| 句法题元 | 1 – NP nom | 2 – (с NP ins) / (против NP pos) |
| 语义角色 | 1 – Agent | 2 – Patient-Source |
| 语义属性描述 | 1 – HUMAN, ORGANIZANION | 2 – ORGANIZANION , ACCIDENT, ACTION, IDEOLOG, POLITY |

例句: *Почти два года строители Армении борются с последствиями землетрясения.* (差不多两年时间亚美尼亚的建设者们同地震的后果作斗争。)

Борься₄用莫斯科语义学派元语言解释为: X бороться₄ с Y / X бороться₄ против Y = X为反对Y而斗争, 即X与负面的、不良的政治或经济过程、行为Y相互冲突, X力图战胜Y。从中可以看出, бороться₄有两个语义配价, 一个主体配价和一个客体配价。主体为某人、某人的团体或组织, 客体可为某组织、某行为、某些政治立场等。

БОРОТЬСЯ-5

POR: V

TRANS: 为争取...而斗争 (为...而奋斗)

SYNT: object

SEM: ACHIEVEMENT OF AIMS OF PURPOSES

МУ_5

| | | |
|------|------------|-----------------|
| 句法题元 | 1 – NP nom | 2 – (за NP obj) |
|------|------------|-----------------|

| | | |
|--------|-------------------------|--|
| 语义角色 | 1 – Agent | 2 – Aim-Purpose |
| 语义属性描述 | 1 – HUMAN, ORGANIZATION | 2 – HUMAN, OBJECT, EVENT, ACTION, POLITY, ABSTRACTIO |

例句: *Создавая феминистические организации, женщины борются за равноправие с мужчинами.* (由于成立了女性组织, 妇女为争取同男性的权力平等而斗争。)

Борься₅用莫斯科语义学派元语言解释为: X бороться₅ за Y = X 为争取 Y 而斗争, 即 X 克服重重困难, 为达到既定目标 Y 坚持并不断追求。可见, 该支配模式中 бороться₅ 有两个语义配价。主体为某人或组织, 客体可为某人、某客体、某事件或行为、某政治制度、抽象概念 (борься за справедливость, бороться за чистоту нации) 等等。

以上是关于 бороться 的词条描述, 为检验该词典的实用性, 我们从国家语料库中选取了前100个带 бороться 的例句, 研究结果, 含 бороться с чем 的结构共有45个; 含 бороться за что 的结构共有41个; 含 бороться против чего 的结构共有2个; 无补语的结构共12个。研究发现, 在无补语的结构中, 未出现“角斗”意义的句子, 12个句子都具有“战斗”意义, 而且常与表示工具意义的第五格连用, 表示“用……来战斗”, 如 *Нужно бороться идеологией партии, а не кадрами партии.* (需要用党的思想, 而不是党的干部战斗。) 或是常与表程度的副词连用, 如: *Однако к чести "Единства" стоит отметить, что, даже безнадежно уступая, они продолжали бороться до самого конца.* (至于说到“统一”队, 应当指出的是, 即便注定失利, 他们一直战斗到最后。) 在很多句子中, 出现省略主语的情况。

4 结语

本研究尝试将多种分析方法相结合, 将多理论综合加以利用, 尽可能充分地发掘构建详解组合词典的描述机制, 希望该课题研究能够推进俄汉信息化处理研究, 有助于构建语义驱动的结构语言模型和发展俄语的理解技术, 从而促进俄汉机器翻译系统研究的发展。

附注

1 例句取自俄罗斯国家语料库[EB/OL].<http://www.ruscorpora.ru>, 2014-03-20.

2 Ю.Д. Апресян 在《新编俄语同义词解析词典》(Апресян, 1997, 2000) 中使用了 16 个语义配价, 称对称谓词(结识、交谈、争论、战斗等)的第二语义配价为第二主体。(张家骅 2003)

参考文献

- [1]Bake, C. F., C. J. Fillmore, and John B. Lowe. The Berkeley FrameNet Project[A]. In Proceedings of COLING, 1998.
- [2]Miller, G., et al. Introduction to WordNet: an on-line lexical database[J]. International Journal of Lexicography 1990(3).
- [3]Nirenburg S., Raskin V. Ontological Semantics[M]. Cambridge, MA: MIT Press, 2004.
- [4]Richardson, S. D. William B. Dolan, and Lucy Vanderwende. MindNet: acquiring and structuring semantic information from text[A]. In Proceedings of COLING, 1998.
- [5]Vendler Z. Linguistics in Philosophy Ithaca[M]. New York: Cornell University Press, 1967.
- [6]Авилова Н.С. Вид глагола и семантика глагольного слова[M]. Москва: Наука, 1976.
- [7]Апресян Ю.Д. и др. Лингвистическое обеспечение системы ЭТАП-2[M]. Москва: Наука, 1989.
- [8]Апресян Ю.Д. Исследования по семантике и лексикографии. Том1. Языки славянских культур[M]. Москва, 2009.

- [9]Апресян Ю.Д. Лексическая семантика. Синонимические средства языка[M]. Москва: Наука, 1995.
- [10]Апресян Ю.Д. Новый объяснительный словарь синонимов русского языка, выпуск I[M]. Москва, 1997.
- [11]Апресян Ю.Д. Новый объяснительный словарь синонимов русского языка, выпуск II[M]. Москва, 1997.
- [12]Мельчук И.А. Опыт теории лингвистических моделей «Смысл↕Текст»[M]. Москва, Наука, 1974.
- [13]Мельчук И.А., Жолковский А.К. Толково-комбинаторный словарь русского языка: Опыты семантико-синтаксического описания русской лексики[M]. Вена: Wiener Slavistischer Almanach, 1984.
- [14]Ожегов С.И. и Шведова Н.Ю. Толковый словарь русского языка[Z]. Москва, 2003.
- [15]董振东, 董强. 面向信息处理的词汇语义研究中的若干问题[J]. 语言文字应用, 2001(3).
- [16]董振东. 语义关系的表达和知识系统的建造[J]. 语言文字应用, 1998(3).
- [17]范晓. 动词的配价与句子的生成[J]. 汉语学习, 1996(1).
- [18]黑龙江大学俄语语言文学研究中心辞书研究所. 大俄汉词典[Z]. 北京: 商务印书馆, 2008.
- [19]孙爽. 面向俄汉机器翻译的 C+N₅结构的语义识别研究[J]. 中国俄语教学, 2013(3).
- [20]于鑫. 关于俄罗斯当代语义学[J]. 俄语语言文学研究, 2008(2).
- [21]俞士汶. 计算语言学文集(4)[C]. 北京: 北京大学计算语言学研究所, 2000.
- [22]詹卫东, 刘群. 词的语义分类在汉英机器翻译中所起的作用以及难以处理的问题[A]. 语言工程, 全国第四届计算语言学联合学术会议论文集[C], 1997.
- [23]张家骅. 莫斯科语义学派的配价观[J]. 外语学刊, 2003(4).

Syntactic and Semantic Ability Construction of Russian Verb *бороться*

Sun Shuang

(Northeast Forestry University, Harbin 150040, China)

Abstract: Syntactic semantic explanation combination dictionary is a necessary part of machine translation system. Based on ontological semantics of Nirenburg S. and Raskin V., “Meaning↔Text” linguistics model theory of Melchuk and language integrated description and systematic dictionary theory of Апресян, this paper tries to establish the syntactic semantic explanation combination dictionary which can be used in Russian-Chinese machine translation system. Based on the semantic classification of the predicate, the paper gives a detailed syntactic semantic description of meanings of predicate *бороться*.

Key words: Russian-Chinese machine translation; syntactic and semantic; explanation combination dictionary; *бороться*

基金项目: 本文系国家社科基金项目(11CYY063), 中央高校基本科研业务费专项资金资助项目(2572014CC10)之系列文章。

作者简介: 孙爽, 东北林业大学外国语学院副教授, 博士, 研究方向为计算语言学、语义学。

收稿日期: 2016-03-31

[责任编辑: 叶其松]