**自然语义的元语言阐释**

**张文彦**

**序言**

一直以来，人们对语言的研究都遵循着皮尔斯的三分法模式：语法、语义、语用。客观的讲，除了对语音和语法的研究取得一定的成果之外，语义和语用部分进展很缓慢，甚至有的语言学家认为语义研究不能看作一门科学，这是大多数语言学家都无法接受的。

然而事实情况是，不论在哲学领域、人工智能领域、还是词典学界，人们对于语义分析的需求越来越迫切。于是出现了各种各样的语义分析理论和观点。

今天我们主要介绍的是自然语义元语言的观点。简单说就是用元语言的办法来解释和处理语义。

（一）自然语义元语言的思想核心

1. 语言是用来表达意义的，因此对语言的研究应该从意义来着手。
2. 不论是词项、语音、还是句法结构，都有着各自的语义，可以称之为词汇语义，语法语义，语用语义等。甚至会有文化语义。
3. 和逻辑学家不同，自然语义元语言的支持者们认为语义研究的并不是符号和现实世界的关系，而是符号和人类认知的世界之间的关系。
4. 通过寻找基本义素（或者说是各个语言的共同语义最小片段）的办法，我们可以确定一些最初始的概念或词汇，以此来解释具有复杂语义的概念、词汇、或者句法结构。

（二）研究相关性

这样一种寻找基本义素生成语义的思想和形式语义学中对词项和修辞关系确定中所要求的初始谓词有着理念上的异曲同工之处。SDRT对这些初始谓词报以搁置的态度，究竟有多少初始谓词，应该以什么样的方式组合来生成新的概念，都还没有涉及。只是在需要用到初始谓词来确定语义时才凸显一下，非常不成系统。

如果自然语义元语言能够找到足够的基本义素，那么自然语义的形式化也就有了坚实的基础。

首先我们来看什么是元语言。

1. 元语言的发展历程

20世纪20年代，德国数学家Hilbert创造出metamathematics一词，借助于前缀meta-而构成的新概念也不断涌现出来。（如元科学、元理论、原问题）metalanguage也随之提出。

Meta是标志着研究层面转变的一种哲学语言，指人们的讨论对象由客体（object）转变为讨论（talk）本身。（安华林，2005:104-105）

在语言学界，其实很早就有了使用元语言来描述语言系统的典范：

Eoght Books（波你尼梵文语法）

这是一部研究印欧语最早的专著，最大的贡献在于用希腊语的词汇资源创造出了一套“技术性的元语言”。

后来，索绪尔等人创建并使用了技术性的元语言作为描写语言和语言系统的手段，（当然仅仅涉及到语法，并没有扩展到语义）使得现代语言学得到极大的发展。

在很多语言学家看来，元语言对语言学的作用非常之大。伦敦语言学派的Firth就说“每个科学领域都应该开发出适用于其性质的一套特殊语言，这个过程代表着科学工作的最基本组成部分。”甚至有学者认为，如果没有元语言，也就没有逻辑学，更不会有符号学和语言学。

上述内容告诉我们，使用元语言的办法分析自然语义，不仅是有依据的，也是十分必要的。

不同领域内对元语言一词的使用也是有所出入的。下面来看几个和我们有关的元语言定义。

1. 元语言的类型
2. 语义哲学界

20世纪波兰逻辑学家Tarski认为，当人们判断一句话是真还是假时，往往会把这句话的客观真实性与这句话的存在真实性混淆在一起。因此，在区别语言与语言所指称的事物之间的关系时，就有必要把真实语言和形式语言区分开。（这里所说的真实语言即对象语言，形式语言即元语言）

《西方哲学英汉对照辞典》里定义如下：第一语言的表达式的名称以及这些表达式之间关系的名称，都属于第二语言，后者叫做元语言。

1. 词典学界

在词典学界，也有一批学者使用元语言的办法来解释和处理词典中所收录的词语。他们称之为释义元语言。代表人物是波兰裔澳大利亚学者韦日比茨卡。主要观点如下：

1. 任何语言的词典中都存在不可定义的词，它们数量较少且自成系统，它们的作用是用来定义其他词语的。
2. 不可定义的词是可列举的，语言中的其它词项可以用它们来定义。
3. 不可定义的词在不同的语言中虽然各有不同，但却能够互相照应，语义上是等价的，可以将其视为“普遍词汇”。
4. 人工智能界

在自然语言处理界，自然语言的语音、语法部分的形式化工作基本完成，而语义部分的计算机处理一直是个难题。一个解决的基本思路就是创造出一些原始语义，通过一个翻译算法将自然语言转化成用原始语义进行的语义表示，其中的原始语义本身不能再消减成其他同类实体。

1. 异同点

共同点：1、都认为基本义素应该具有最小性和不可分解性。

2、都认为义素具有生成性，由基本义素组建成其他语义。

不同点：哲学界和人工智能界认为元语言应该是抽象，形式化的，高于事实语言的。词典学界则认为元语言就是事实语言的一部分，或者说是事实语言的一个真子集。

至于哪个观点是正确的，我想在实际应用中我们能够通过比较二者的优劣而得到答案。目前来看这一分歧并不影响我们分析语义。

下面我们分别来看一下各个领域相关的工作状况

1. 简述

尽管逻辑学界、人工智能界和词典学界对于语义元语言的设想和期待都很靠近，但各自展开研究的手段和方式却不尽相同。目前来看，取得较多成果的是词典学界，这里我们先从词典学界开始介绍。

1. 词典学界

代表人物：韦日比茨卡

1、理论基础：任何思想的产生和形成都是建立在对以往理论的批判、继承和创新基础上的。韦日比茨卡的元语言思想也同样受到前辈学者的影响。在不同的语言中析取其中共同的核心语义要素的思想早已存在。

洪堡特就曾经说过，“不管是在词汇中，还是在语法里，总是存在着所有语言围之旋转的核心。”

莱布尼兹（德国）创建的“人类思维字母表”也体现了这一思想，并且提出了要从自然语言中来寻求这些最小的基本概念。这个设想层一度被认为是“乌托邦”而遭到否定，但莫斯科语义学派以及一些波兰学者则深信这一设想一定会实现。

（简单介绍一下莫斯科语义学派：又称俄罗斯语义学派。

代表人物：阿普列相，莫斯科语义学派的奠基人之一，论证了词汇是具有严密组织的分类体系和运作体系。

主要观点：认为语义元语言是自然语言最简便和标准的描写语言的判断，提出了语言的形式模式和借助句法描写语义的新概念，研究了同义现象与同义词典的词位分析性注释的方法论。语言单位意义在话语中相互作用的规则等问题，阐释了词典学类型和词典学肖像等体系性词典学研究的纲领、理论和基本方法。）

近20年来对儿童语言研究的成果表明：儿童学习语言的首要任务是获得知识。而在获得知识的过程中，他们不是被动的接受者，而是配载有某些先天的基本概念的积极行为者。儿童语言学家鲍尔曼指出：她看到了儿童最早的概念和普遍语言以及语义最小成分之间的关系。

还有一个不容争议的事实是，尽管在不同的语言中存在大量具有民族文化特征的概念，但这些并没有阻碍人们成功地进行跨文化交流，这就使得自然语义的元语言思想有了很强的现实性。

以上就是自然语义元语言思想的理论和事实依据。下面我们看一下其他语言学派对语义的研究。

2、理论背景：在普遍语义元思想产生之前，语言学研究主要受布龙菲尔德的行为主义和乔姆斯基的转换生成语法的影响，但两者都忽视了意义的研究。

**（行为主义：**把语言看成为一系列[刺激](http://baike.baidu.com/view/65500.htm)和反应，并据此提出了著名的传递[公式](http://baike.baidu.com/view/645857.htm)：Sàr…sàR：S表示实际的刺激；R表示实际的反应；r表示语言的替代性反应；s表示语言的替代性刺激。

**转换生成语法：**主要包括基础和转换两个部分，基础部分生成深层结构，深层结构通过转换得到[表层结构](http://baike.baidu.com/view/1956536.htm)，语义部分属于深层结构，它为深层结构做出语义解释。语音部分属于[表层结构](http://baike.baidu.com/view/1956536.htm)并为表层结构做出语音解释。）

布龙菲尔德坚信语言学是一门严谨的科学，但对语义部分的研究又很难做到严谨，因此，为了保持语言学的科学性，他把意义排除在了语言研究的范围之外。

乔姆斯基也曾经在做完句法研究后在别人的建议下处理过语义，但并没能得出令人满意的结果。总体来说，意义的研究有些停滞了。

3、韦日比茨卡的观点和传统大不相同。

和传统的三分法不一样，她认为句法、语用这两个层面在本质上都是语义的，这三者之间并没有明显的界限，原则上可以用统一的语义语言对他们进行描写。

（在我看来，语音也是为表达意义而生，本质上也没有区别，手段不同而已）

和传统的逻辑学语义观点也不同，韦日比茨卡认为语义并不是单纯的符号和外部世界的关系，而是和人文为中心的、带有认知色彩的世界之间的关系。

4、基本方法

在传统的词典和语义学中使用的大多是循环定义的办法，例如：

狗：犬，学名“家犬”……

犬：本义狗，古时特指大狗，后犬、狗通名……

（出自《现代汉语辞典》）

循环定义的缺点显而易见，很难实现对语义的准确把握和形式化表述。

自然语义元语言试图使用最基本、最简单的词来表征和解释复杂的语义。类似于化学分析中把复杂的分子式还原为简单的基本化学元素，代数中把复杂的数学表达式还原为最简单的数学表达式，这种方法称之为“还原释义法”。

韦日比茨卡认为“每种语言都有一个不可再次还原的核心，通过这个核心，说话人可以理解所有复杂的思想和言语，不同语言的核心是相等的，它们反映了人类不可还原的基本思想。”这些核心包括：

A 一些无需定义的语义最小成素（semantic primitive）

B 制约组合的一些原则（可以看成是句法或文化规约）

C 体现语言特征的结构（类似认知语义学中的“脚本”script）

（脚本：认知心理学名词，源自图示schema,由Schank和Abelson在1977年提出。他们认为在人们理解的过程中存在着比概念表达式更大的知识单元，这就是脚本。例如人们在日常生活中经常遇到的一些事件，如就餐，旅游等，都是由一系列相对固定的行为序列和特定的环境条件所组成的。脚本将这些行为序列和环境条件组织在一起，形成关于这些事件的知识单元。脚本不仅是一个知识存储结构，同时也是潜在的语境。）

接下来就是问题的关键，即如何寻找和确定这些语义最小成素呢？

韦日比茨卡的答案是：试验，即大量的尝试和调查测试。

目标是寻找出数量最小，解释力最强，适用性最广的语义最小成素集。

目前韦日比茨卡得到的最新版本如下：

1. 元语言词库：15种类型，58个初始义素

Substantives: I,YOU,PEOPLE,BODY,SOMEONE,SOMETHING

Determiners: THIS,THE SAME,OTHER

Quantifiers: ONE,TWO,SOME,MANY/MUCH,ALL

（在many 和much这个初始义素上我们可以看出汉语的优势，因为英语里有摆脱不掉的“可数”和“不可数”义素附着在单词里，汉语则没有）

Attributes: GOOD,BAD,BIG,SMALL

…

1. 元语言句法：韦日比茨卡认为，人们在形成和表达思想的过程中，不仅仅需要一定数量的初始义素，而且还需要适用于这些初始义素的普遍组配模式，也就是句法。（这部分韦日比茨卡做的不太好，还没有成熟的结果。）
2. 最后在（1）和（2）的基础上就能造出如下的基准句：

You did something bad.

I know it happened.

I want to see this.

Because of something, something happened.

…

这些基准句就能够用来解释词汇或语句的语义。

1. 最终分析结果举例如下：

Make:

首先make主要出现在使役结构语句中。根据使役主体与使役对象的性质，又可以分为两大类：人际间使役关系和非人际间使役关系。

A 以make为中心词的人际间使役关系又可细分为6种：

a “make something happen to someone”

eg: He made her loose her temper.

这种情景所表示的语义用自然语义元语言表示为：

Someone X did something.

Because of this, something bad Z happened on somebody Y.

If X did not do this thing, Z would not happen on Y.

b “making someone think something”

eg: He made her forget her troubles.

这种情景所表示的语义用自然语义元语言表示为：

Someone X did something.

Because of this, someone Y think of Z.

If X did not do this thing, Y would not think of Z.

…

B 后面还有以make为中心词的非人际间使役关系的5种类型。略。

5、优缺点

主要有如下几个缺点：

一是通用句法模式的探求发展非常缓慢，远远落后于通用词汇的探求。

二是初始成素的数量逐渐扩大。

莱布尼茨认为，各个初始成素应该是互相独立的，假如两个成素之间有共同成分，则这两个成素在语义上都不是简单的，换句话说就是这两个成素都是由更基本的初始成素组合而成的。

但是为了减轻语义分析的负担，能够让人更直接的理解和接受，韦日比茨卡抛弃了莱布尼茨的这一设想。这一点是我非常不赞同的。虽然增加初始义素能够帮助人们直观理解，但是这不利于语义分析的严格形式化。

三是韦日比茨卡的最初义素都是从以英语为主的欧洲语言中提取的，有局限性，就像我们刚才提到的many和much，显然没有汉语中的“很多”基本和单纯。此外还有一些初始义素值得推敲。（如wish和want）

四是同一个初始义素在不同的语言中地位并不同，有些是独立的词汇，有些则只是些后缀。

如：“She left.”和“她走了。”表过去时的义素在英语和汉语中的地位不尽相同。也就是说不是所有的元语言都能用自然语言的子集来表示。在某些时候，我们还得引进人工语言。

1. 认知语义学界

关于语义的分析，认知学界也有着自己的理解。正如上文所提到的，认知语义学寻找初始义素（当然他们称之为认知义素，在我看来没有什么不同）的出发点是从“脚本”（这里称之为“认知事件”）开始的，而不是像韦日茨比卡那样漫无目的的大量尝试。并且从“认知事件”开始寻找义素，也从某种巧合上寻找到了义素的组合方式（即句法），也就是韦日比茨卡所欠缺的关于句法的内容得到了补充。

另外一点是认知语义学所寻找的认知义素并不局限于某一个语言中，这也就解决了刚才我们谈到的有些语言中的初始义素不是单独的词汇的问题。我的观点是就认知义素而言，汉语最起码要比英语更加符合认知语义的特点。后面我们会看到实例。

1、认知语义学所分析的初始义素（这里称认知义素）

目前来看,认知义素究竟需要覆盖哪些方面才能保证语句生成的过程中不出现语义违反,还是未知,众多的认知语义学家们还在不停寻找答案.单从动词的分析来看,我们现在就已经有了如下几种方面的认知义素:运动,方式,原因,结果,工具,路径等.

从动词角度来看，不同的语言动词包含的认知义素方面不同,如:汉语中动作的路径或结果等义素是独立于动词单独表达的(通常为状语或补语),如”杀死”,”穿过花丛”;而英语中很多表示方式或结果等的义素就嵌入到了动词内部，如“hammer”，“kill”。

从这个角度看汉语的字更接近于认知语义要素单位。如：

I killed John.

我杀死了约翰。

“kill”强调的是“死”这个结果，动作是抽象的，这里用“弄”来代替抽象动词。也就是说，[抽象施动动作]和[死]才是单词kill对应的认知义素。（我们用加中括号的方式表示义素，用没有括号的方式表示语言使用单位）

(这里我使用了汉语认知义素作为公共认知义素，这是我的一个主张，并不是为了简便。当然还有待验证。)

英语和汉语各自的特点决定了一个英语单词在很多时候对应的是几个分离的汉语语义单位（汉语的语义单位是什么同样没有定论，这里我们采用的观点是：汉字是最小的语义单位）。

2、认知义素的确定方式

当然动词也有很多种，具体哪些动词包含哪些初始义素，这些初始义素之间又是怎样的组合关系，还是要看不同的事件结构对动词的要求。下面我们来看一个认知事件。

事件结构为“施事者+施动行为+受事者+结果+受事者+方式+工具”的动词句

例如：I killed John by starving him.

我们用如下方式表示动词的义素成分：

I :={[我],[施动者]}

Killed:= {[抽象施动动作],[死],[了]}

By := {[用……的方式]}

Starving := {[饿]}

Him := {[他]，[受事者]}

可以看出，上述这些认知义素完全是由该事件结构所涉及的成分所确定的。那么这些类型的认知义素能否达到通用义素的效果呢？（也就是说在其他语言里是否适用）我们可以按照汉语中该事件结构的组合要求将这些义素组合成语句检验一下。

在未调整语法顺序的情况下，仅从认知义素本身我们就可以实现第一步翻译为：

我弄死了约翰用饿他的方式。

基本上达到了交流的效果。

显然，句是不符合汉语表达顺序的。（在表述“行为/方式/结果”事件的汉语语句中，表达顺序应为：主语+方式+行为+结果+宾语，这是相同事件结构句法表达中英语和汉语不同的部分）调整语序后的汉语表达为：

我用饿死他的方式弄死了他。

至此，翻译工作应该可以看作结束了。如果还要考虑简洁度的话，就需要使用SDRT中代词指代确定原则和汉语表达“行为+方式”特有的结构来消去多余部分（过程较繁锁，略去），最终简化为：

我饿死了他。

从汉语检验的角度来看，认知义素提供的义素种类还是足够的。

3、最终目标

一是对语言做出合适的义素切分,目标---切分后的义素属于一个共同的义素集.切分依据: 认知语义学的相关结果

二是找出语言使用中重点使用的相对完整的语义搭配种类.(即各种事件结构),或者称之为初始句法。

试想一下，如果有一天我们真正实现了标记出各个语言单位所蕴含的义素,并且通过共同的事件结构确定了语言各自特有的表达方式（也就是句法）.那么使用形式化的办法表达语义的目标就更进一步了。（当然，这是个美好的设想）

4、优缺点

优点：认知语义学所提供的认知义素和事件结构能够帮助机器理解语义，（即可以借鉴来做形式语义学的初始谓词和组合句法），也可以很好地帮助我们实现英译汉等机器翻译的实现。

缺点：事件结构的寻找工作还在探索阶段，而且也仅涉及到了对动词认知语义的帮助，其他性质的词语还有待研究。

（三）逻辑学界

周老师提出的concept system,郑值上次讲到的DRT等，都是从逻辑学的角度处理自然语言语义的积极尝试。共同点是逻辑学界没有出现概念互相定义的问题，都是认可初始义素的存在的。至于究竟从何找起，还没有特别系统的方法和相对完整的结果。

周老师的concept system把对象分为了个体、性质、类、动作、关系、时间、空间、状态等。是不是也等同于在某种程度上对初始义素所做的分类？至于这些分类内部到底包含哪些初始义素？我还没有完全看到。但concept system对各个概念节点的上下位关系以及概念之间的关系给出了一个详细的描述。在我看来这和初始义素的组合方式也有着很密切的关联。（如哪些义素可以组合，哪些不能组合，以什么样的方式组合，当然不完全等同。例如周老师提到“鸟”如果有“会飞”的内涵项，“鸟”的内涵包里就要有“飞”这个节点的地址。如果还有“有羽毛”的内涵项，那么怎么体现这两种内涵的不同呢？或者说内涵的地位都是一样的？）

还有周老师提到的外延树上的各种节点以及他们之间内涵的传递关系（即上位节点的内涵项也是下位节点的内涵项。例：“鸟”有“会飞”的内涵项，那么“燕子”也有“会飞”的内涵项，当然允许例外）是对很大一部分名词类词语非常系统的分析，非常值得借鉴。

DRT也对一些词项做了初始语义表述，但基本上也是从自身使用的角度来确定哪些是初始义素，哪些不是。没有一个系统的确定办法。以至于SDRT后来要处理句子间修辞关系问题的时候涉及到关键词项的特定语义要求时只能留一个空白在那里。

结语：

综合来看，对于初始义素的存在，是得到大多数学者的认可的。（当然也有反对的学者，如：）究竟该如何认定这些初始义素？是用句法还是别的形式来表示义素的组合方式？百科知识这部分内容又该出现在哪里？（如日本人在听到别人说“你家真漂亮”时会一直谦虚的说“哪里哪里”，而欧洲人则会回答说“thank you.”不同的文化背景决定了不同的回答。）这是各个理论主要的分歧所在。

我觉得可以借鉴认知语义学和俄罗斯语义学派等词典语义学的分析成果，结合各家所长，最终实现我们科学分析语义的目标，争取早日摆脱“语义学不是一门科学”的帽子。

参考文献：

[1] 自然语义元语言：理论与实践，张哲，张发祥，赵国栋，北京，科学出版社，2007

[2] 认知语义学，束定芳，上海，上海外语教育出版社，2008

[3] 语言整体性描写与体系性词典学，阿普列相著，杜桂枝译，北京，北京大学出版社，2011

[4] 俄罗斯语义学，张家骅，北京，中国社会科学出版社，2011

[5] 形式语义学引论，蒋严，潘海华，北京，中国社会科学出版社，1998

[6] 概念系统Cons.X及其结构方案，周北海，

[7] 安华林，现代汉语释义基元词研究，北京，中国社会科学出版社，2005

[8] 西方哲学英汉对照辞典，步宁，余纪元，北京，人民出版社，2001

[9] 认知心理学，布丽基特 贾艾斯编，黄国强，林晓兰，徐愿译，黑龙江科学技术出版社，2001