

《现代韩国语动词语义组合关系研究》

毕玉德著

民族出版社，2005年11月出版，

ISBN 7-105-07337-3/H·559

序言

冯志伟

2004年我曾在韩国科学技术院（KAIST）计算机科学系担任访问教授，毕玉德在年底时也到那里做博士后研究，他向我介绍了过去他在做博士期间研究现代韩国语动词语义组合关系的情况，我觉得这是一项很实在的研究工作。现在他的研究成果就要出版了，我感到由衷的高兴。

在自然语言的计算机处理中，动词的分析是一个关键性的问题。动词是句子的中心，而句子中所表述的事件就好比一场戏，戏里面要描述参与演戏的角色、场景、工具、时间、地点等等，它们都要以动词为中心来安排。动词的语义组合关系好比演这场戏的脚本，句子中的其他成分都要按照这个脚本来实现它们各自的功能。没有脚本就不能演戏，没有动词的语义组合关系，句子也就不能成其为句子了。所以，动词语义组合关系的研究是非常重要的。

本书首先研究了现代韩国语动词的论元结构，作者把动词的论元（argument）叫做角色（role），他把语义角色系统分为三个层级，第一级分为主要角色和次要角色，第二级进一步分为主体角色、客体角色、关联角色、环境角色、根由角色和凭借角色。第三级又进一步分为施事、经验、当事、领事、致使、受事、结果、内容、感事、受益、来源、伴随、涉事、基准、原始状态、结果状态、系事、属事、分事、

工具、材料、方式、资格、范围、起点、终点、场所、数量、路径、方向、时间、条件、程度、根据、原因、目的等36个角色。

20世纪70年代末和80年代初，我在法国格勒诺布尔理科医科大学研制汉-法/英/日/俄/德多语言机器翻译系统FAJRA时，曾经根据Tesnière的依存语法（*grammaire de dépendence*），对汉语的动词、形容词和部分名词的论元结构进行了初步的探索，当时我提出的论元共30个。我使用计算机求解了这些论元信息，把汉语自动地翻译成5种外语，顺利地完成了多语言机器翻译实验。这是世界上第一次汉语到多种外语的机器翻译系统，这次机器翻译的实践说明了我提出的30个论元是正确的。本书中根据韩语的事实所提出的论元结构系统与我在研制机器翻译时根据汉语的事实所用的论元结构系统很接近，如果用本书中提出的这个论元结构系统做与韩语相关的机器翻译，估计也会取得比较满意的结果。这说明，论元结构不依赖于特殊的语种，具有语言上的普遍性，不论是韩语还是汉语，它们的论元结构应该相同或者相近。我认为，作者的这个论元结构系统具有普遍性，是科学的、合理的。

本书还研究了韩语动词的语义分类，建立了语义特征集，共提出了23种语义特征，其中，有些特征是动词本身直接反映出来的，而有些特征是通过语义搭配的主体和客体间接反映出来的。对于动词的语义分类，很多研究者往往只注意表示行为的动词的分类，而忽视表示状态和关系的动词的分类。本书把动词语义场分为状态语义场、关系语义场和行为语义场三大类，比较全面地反映了动词的语义特征。

我认为，这样的动词语义分类也是科学的、合理的。

本书还提出了结构体的概念，在结构体中，分别给出语义分类码、汉语释义、词类范畴、句法表达式、语义特征、语义表达式和实例等信息，作者采用这样的方法，建立了韩语动词的知识库。这样的知识库，对于机器翻译和其他自然语言处理系统的研制，是极为宝贵的语言资源。

本书在语义组合关系框架内，根据作者已经提出的动词语义特征集，推导出了动词系统的语义分类树，但是，这样的推导是在作者已经定义好的语义信息范围内进行的，推导过程中并没有产生出这个范围之外的新信息，因此，我们认为，本书中建立的韩语动词的语义关系基本上还是隐性的、潜在的、静态的，要把这种隐性的、潜在的、静态的语义关系转化为显性的、具体的、动态的语义关系，还必须把它们放到句子中去，使它们与句子中的名词、代词、形容词等其他成分相互作用，彼此碰撞，才能显现出句子中所描述的事件，从而演出生动活泼、有声有色的戏剧来。语言知识库中储存的动词语义信息只是静态的脚本，这个脚本必须通过句子中的名词、代词、形容词等其他成分来担任角色，赋予它们以特定的句法功能和语义功能，才能发挥真正的作用。这些在句法语义分析过程中产生的句法功能和语义功能信息，是在名词、代词和形容词的原先的知识库中没有的，它们是新产生的信息，因而都是动态的信息。这些动态信息的自动推导，是自然语言处理系统智能性的重要标志。

我希望作者在本书研究成果的基础上，进一步研究韩语名词和形

容词的语义系统，把动词的语义与名词和形容词的语义结合起来，以动词的语义信息为基础，结合名词和形容词的语义信息，建立各种运算规则来动态地推演出句子中的名词和形容词等成分所扮演的角色。知识库中存储的信息是静态的，而通过规则运算推演出的信息则是动态的，静态信息的制定要根据词类、句法和语义系统的规范，这个规范应该以“知识本体”（ontology）为基础，动态信息的求解要根据产生式规则，产生式规则的基本形式是“条件-动作”偶对。如果我们把基于知识本体的静态信息与基于产生式规则的动态信息求解规则结合起来，就可以构成一个完整的自然语言处理系统，例如，机器翻译系统、自然语言理解系统等等。

1991年，美国计算机专家尼彻斯（R. Niche）等在完成美国国防部高级研究计划局（Defense Advanced Research Projects Agency，简称DARPA）的一个关于知识共享的科研项目中，提出了一种构建智能系统方法的新思想，他们认为，构建的智能系统由两个部分组成，一个部分是“知识本体”（Ontology），一个部分是“问题求解方法”（Problem Solving Methods，简称PSMs）。知识本体涉及特定知识领域共有的知识和知识结构，它是静态的知识，而PSMs涉及在相应知识领域进行推理的知识，它是动态的知识，PSMs使用知识本体中的静态知识进行动态的推理，就可以构建一个智能系统。我认为，尼彻斯等学者提出的这种构建智能系统的思想与我们提出的关于基于知识本体的静态信息与基于产生式规则的动态信息求解规则相结合构建自然语言处理系统的思想是一致的，这应该是构建智能系统的基本

方法。如果本书作者能够按照这样的方法继续进行研究，一定会为韩语的信息处理做出更大的贡献。

冯志伟

2005年7月19日于北京