

汉语交叉依存类非投射性现象

**EDGE-CROSSING NON-PROJECTIVE
PHENOMENA IN THE CHINESE
LANGUAGE**

王 跃 龙

WANG YUE LONG

新加坡国立大学中文系
NATIONAL UNIVERSITY OF SINGAPORE
2012

汉语交叉依存类非投射性现象

**EDGE-CROSSING NON-PROJECTIVE
PHENOMENA IN THE CHINESE
LANGUAGE**

王 跃 龙

WANG YUE LONG

新 加 坡 国 立 大 学 中 文 系
博 士 学 位 论 文

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF CHINESE STUDIES
NATIONAL UNIVERSITY OF SINGAPORE
2012

Acknowledgement

I would like to express my sincere appreciation to my supervisor, Dr. Wang Hui at the Department of Chinese Studies in National University of Singapore, for her continuous encouragement and guidance in all the time of research for and writing of this thesis. I would also like to express my gratitude to the members of my dissertation committee, Dr. Peng Rui and Dr. Lua Kim Teng, for their helpful suggestions and comments that are crucial to the completion of this thesis.

My colleagues from the Department of Chinese Studies supported me in my research work. I want to thank them for all their help, support and valuable hints. Especially I am obliged to LIN Jinzhan, BAI Xiaopeng, XU Tingting and QIN Shaokang.

Finally, I would like to give my deepest thanks to my parents. Their patient love enabled me to complete this work.

目 录

ACKNOWLEDGEMENT	I
ABSTRACT	IX
1 绪 论	1
1.1 课题的提出	1
1.2 研究对象	3
1.3 研究方法与目的	7
1.4 本文的结构安排	8
2 文献综述	10
2.1 依存语法综述	10
2.1.1 依存语法研究概况	10
2.1.2 非投射依存现象的研究	22
2.2 计算语言学领域内的依存分析	25
2.3 交叉依存现象相关的语言学研究	32
2.3.1 主谓谓语句研究	32
2.3.2 相关动补结构研究	40
2.4 语言网论研究	47
2.5 小 结	53
2.5.1 研究现状的总结	53
2.5.2 本文的理论框架	55
3 交叉依存现象分类及句法语义描述	57
3.1 汉语中交叉依存现象的分类	57
3.1.1 数据说明与分类原则	57
3.1.2 交叉依存的类型	62
3.2 单个谓词下的交叉依存现象	73

3.2.1 名词与量词分离类	73
3.2.2 形容词与名词分离类	77
3.2.3 名词与名词分离类	80
3.2.4 副词与数词分离类	82
3.2.5 框式介词类	85
3.3 多个谓词下的交叉依存现象	91
3.3.1 名词与动词分离类	91
3.3.2 副词与动词分离类	94
3.3.3 动词与数词分离类	97
3.3.4 框式副词类	99
3.4 小 结	101
4 句法依存的属性和谐原则	104
4.1 词语的属性	104
4.1.1 汉语词类间依存关系的特点	104
4.1.2 依存描写的细化	107
4.1.3 属性的选取	112
4.2 动词意义的属性	113
4.2.1 动词意义的过程属性	117
4.2.2 动词意义的位置属性	123
4.2.3 动词意义的情感属性	126
4.3 名词意义的属性	128
4.3.1 名词意义的量属性	128
4.3.2 名词意义的情感属性	130
4.4 副词意义的属性	132
4.5 介词的属性	137
4.6 小 结	139
5 环形语序域原则	142

5.1 网络视角下的线性一维化	142
5.1.1 语序成因的现有解释	142
5.1.2 语言立体网络到一维线性序列的投射	146
5.1.3 从网络投射看汉语的语序	152
5.2 汉语语序的事件域	160
5.2.1 第一域	161
5.2.2 第二域	169
5.2.3 第三域	172
5.2.4 第四域	179
5.3 域的交错	180
5.4 域的性质	191
5.4.1 相对性	191
5.4.2 递归性	194
5.4.3 拓扑等价性	197
5.5 小 结	200
6 依存图库的构建	202
6.1 相关的理论探讨	202
6.1.1 从依存树到依存图	204
6.1.2 事件作为中间单位	207
6.2 事件标注	211
6.2.1 单个事件的标注	211
6.2.2 复杂事件的标注	214
6.3 依存图标注规则	225
6.3.1 词汇信息标注	226
6.3.2 句法关系标注	228
6.3.3 语序信息标注	234
6.4 小 结	235

7 结 论	237
7.1 本文研究工作的总结	237
7.2 下一步的研究计划	239
附 录	241
附录一 名词与量词分离类	241
附录二 名词与形容词分离类	241
附录三 名词与名词分离类	242
附录四 副词与数词分离类	243
附录五 框式介词类	256
附录六 名词与动词分离类	262
附录七 动词与副词类	268
附录八 动词与数词分离类	274
附录九 框式副词类	278
参 考 文 献	281

图表目录

插图一览

图 1 交叉依存现象分布图	71
图 2 交叉依存现象的语序分布	102
图 3 语言立体网络图示	147
图 4 三维图到二维平面的投射	148
图 5 事件结构二维环型图示	149
图 6 句子的一维线性图示	149
图 7 语序三个层面关系示意图	188
图 8 域的嵌套	195
图 9 域的叠合	195
图 10 域的链接	197
图 11 词类依存关系层级	208

表格一览

表 1 词类范畴依存关系表	105
表 2 动词位置属性示例	125
表 3 动词意义的感情属性上组合关系表	126
表 4 名词意义情感属性与动、形组合关系表	132
表 5 介词属性描写示例	139
表 6 二维图概括能力在不同语言类型中的例证	150
表 7 “生”各义项属性维度标注示例	227

Abstract

Dependency-based methods for syntactic parsing have become increasingly popular in recent decades for the usefulness of bilexical relations and parsing efficiency. However, most of them are only restricted in projective dependency. In order to improve parsing precision, it is inevitable to involve non-projective dependency in dependency parsing. Therefore, it is also an important topic for dependency research.

For Mandarin Chinese, it is unclear how many kinds of non-projective phenomenon involved. For that reason, it is necessary to do basic data analysis. We choose edge-crossing dependency, one of the non-projective phenomena, as our research topic. According to the syntactic criteria proposed by Mel'čuk (1988), we analysed language data involved in U-penn Chinese Treebank to find edge-crossing phenomena. In the 15162 sentences, we found 2489 sentences including edge-crossing phenomena. Except for the edge-crossing phenomena happened among clauses or caused by coordinate structure, others can be classified into nine different kinds.

By observing their syntactic and semantic conditions, we found two points in common. First one is that syntactic dependencies in Chinese are based on systematic semantic relations, the other one is that words are gathered around verbs in word order.

Therefore, we proposed two principles to explain the reason for these edge-crossing phenomena formed. First, it is the semantic attribute harmony principle, which determines the syntactic relation among words. According to the realities of Mandarin Chinese, we propose to use systematic semantic attributes as the judging criterion for syntactic dependencies. Second is the ring-shaped field principle of word order, which determines the linear order distribution of words. Different word categories distribute in different field. Meanwhile, free word order can happen inside of each field. These two principles jointly restrict the formation of edge-crossing dependency.

Using attributes as judging evidence for syntactic dependency is a feasible method to fine dependency description for Mandarin Chinese. It is beneficial to conduct the correct parsing. Meanwhile, ring-shaped word-order principle can expand linear grammar research to planar perspective. Many controversial problems can be explained from the planar word-order perspective.

These two principles can also be used to conduct corpus construction. We proposed new standards for dependency-graph bank annotation. Abandoning the tree-structure assumption used before and the single-root node, single-parent node and pure node golden standards demanded therefrom, we use graph-form to present sentences. The dependency-graph bank can be used as data learning and evaluation platform for statistic-based parsers.

Keywords:

Non-projective dependency; dependency grammar; edge-crossing dependency ;

Ring-shaped field of word order; attribute; dependency-graph

汉语交叉依存类非投射性现象

1 绪 论

1.1 课题的提出

近年来，依存语法（dependency grammar）的研究与应用越来越广泛。依存引起人们广泛的研究的兴趣不仅在于其理论本身的长处，更直接的原因在于其在信息处理中的应用优势。可具体体现在关系抽取（relation extraction）、机器翻译（machine translation）、消歧（disambiguation）和句法分析（parsing）等诸多方面。典型的应用系统如 IBM 公司的国际机器翻译系统（International Machine Translation Project）和欧洲的机器翻译系统 EUOTRA (Johnson, King and Tombe, 1985)，都是基于依存语法分析器的。并且都取得了很好的应用效果。

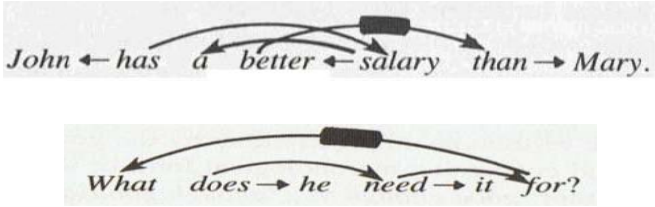
但是，现有的依存句法分析器多是建立在投射性（projective）的原则之下的。Mel'čuk (1988) 对一个句子具有投射性的描述是：在一棵依存树中 ①两个边没有交叉；②没有边跨过最高的结点。这里“结点”代表的是单个词语，“边”是指两个词之间依存关系的连线。Robinson (1970) 对投射性分布的假设则具体描述为：“如果成分 A 直接从属于成分 B，而成分 C 在句子中位于 A 和 B 之间，那么，成分 C 或者从属于 A，或者从属于 B，或者从属于 A 和 B 之间的某一成分。”可以看出，Mel'čuk (1988) 对投射性依存的定义范围要小于 Robinson (1970) 的定义。另可参看 Lecerf (1960)、Hays (1964)、Gaifman (1965) 等关于形式化表示的论述。这里我们采用 Mel'čuk (1988) 对投射性依存较小范围的标准，以求得可

分析的非投射性依存现象范围的最大化。

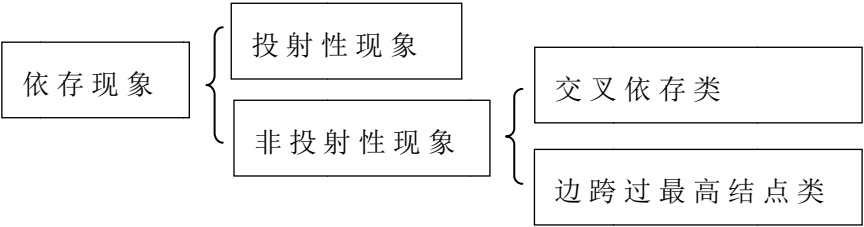
因此，如果出现了两个边的交叉或者有边跨过最高结点，则称其为非投射性现象。相应地，非投射性现象可以分为两种，一种为交叉依存类，一种为有边跨过最高结点类的现象。分别如下两图中加黑方块的线条部分所示。



其在语言中的具体例子如下¹：



这样，整个依存的现象可以做如下的分类：



从句法分析器评测的结果来看，依存分析器的表现是现有分析方式中最好的。但是，并不能说依存的分析已经尽善尽美。现有依存的表示与分析主要是基于投射性的假设，但每种语言中都有部分现象并不符合这个假设，当然其数量相对于投射性的现象来说要少很多，一般为 15%~25% (Nivre & Nilsson, 2005)。要想进一步提高依存分析的精度，必然要开始关注到这些非投射性的现象的正确分析与处理。而且，刘海涛（2009）的研究发现，汉语词汇间的依存

¹ 引自 Mel'čuk (1988)《Dependency syntax: theory and practice》第 37 页。

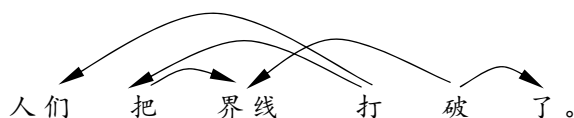
距离为 2.81, 这个数据远超过英语和日语的依存距离(英语 1.386, 日语 1.43)。因此, 很可能非投射性现象在汉语中所占的比例会更高。

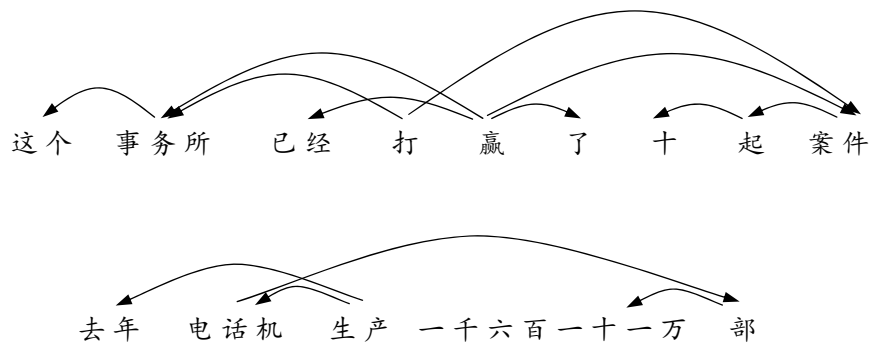
近年来, 非投射性现象的分析与处理已经开始被学者们关注, 但主要是在计算语言学界。而且, 针对汉语的分析研究很少。当然, 投射性和非投射性的概念本就是计算语言学的概念, 并不能因此推论其在以往语法研究中没有得到重视。汉语中的非投射性现象也是广泛存在的, 涉及到以往语法研究中的主谓谓语句、动补结构、副词、介词等多个部分。但是, 以往的研究是分散的, 根据投射性和非投射性的形式划分得到的具体实例, 散布于研究中的或是某一类, 或是某一类中的某些现象。大家以往并不认为这些现象之间有什么必然的联系。但对于依存的机器处理来说, 投射与非投射的假设是基本标准的问题。随着人们对正确处理非投射性现象越来越关注, 亟需一个系统性的理论来把以往这些分散的研究整合在一起。以期在非投射性现象的分析上能有统一的操作标准。

1.2 研究对象

鉴于非投射性依存现象对大规模真实文本分析准确性提高的重要性与目前研究的状况, 本文选取非投射性依存现象中的一个子类----交叉依存现象作为考察的切入点。

下面几个例子是我们汉语中经常会遇到的非投射性结构中交叉依存类的情况。分别涉及以往研究中的“把”字句、动补结构、移位理论等。这里的几个句子只是举例性质的, 后文中会详细谈到汉语中更多的交叉依存现象及其分类。





选择交叉依存现象作为我们研究的对象，是基于以下两个方面的考虑：

首先，对研究材料的实际操作性和应用价值上的考虑。（1）交叉依存现象涉及到汉语中一些研究的热点，如动补结构和主谓谓语句等。这几种结构的分布在汉语中是非常广泛的，且一直是语法研究中的重头戏。对这几种结构的综合考察，如果能统一在同一个框架之下，给出合理的解释，将具有较大的理论价值；（2）对这几类现象各自的现有研究已经取得了丰硕的成果，可以作为本文研究的基础与参照。（3）从应用的方面来说，这类现象的正确考察与处理，将会为句法分析器效能的提高具有直接的推动作用，使得分析精度得到显著的提高。同时，统一的理论框架对于语料库建设也具有指导意义。

第二，基于对组合关系和语序的重新思考。

Mel'čuk（1988）指出，语言中只有两种手段来表达信息，一种是词汇手段，另一种是非词汇的手段。非词汇手段又可以分为三种：一是线性的词序；二是韵律（音高、停顿、重音）；三是屈折变化。他们传达的信息又可以分为两种，一种是语义信息，一种是句法信息。句法信息又可以间接地反映语义信息。我们把词序、韵律、屈折统称为句法手段。三种句法手段中，屈折是应用最少的，有的语言中根本没有屈折变化。而且，在所有的语言中也存在着句法上与别的词语有联系，但无屈折变化的词语。韵律则是处于中

等的地位。词序应用最为广泛。它存在于所有语言的所有句子当中。这是由语言在物理上的线性特征所决定的。

词序、韵律、屈折，都是构成句法结构的手段，因而，不能同时又作为表示句法结构的手段。也即，句法结构的表示手段不能是句法结构本身。短语结构的表示方法既是用句法结构本身来表示本身的做法。对于英语这样的组块式的语言来说，这样的所谓表示只能看做是对其组块结构的承认，而不能说是对其给与了表示。而且，对于词序灵活的语言来说这样的表示方式是无法接受的。句法结构传递的信息是比较复杂隐晦的，有的时候不够明确。比如，同一语序可以形成不同的结构，同一个结构可以用不同的语序呈现。因而，这种模糊性在句法结构的表示中是应该被避免的。所以，我们付诸于依存语法来对句子承载的各种信息给与表示。

使用依存的研究方法可以引起我们对以往组合关系研究的重新思考。我们知道，组合和聚合是语法研究中经常提到的基本关系之一，但是人们往往默认了组合是只能发生在相邻的两个单位之间的²。而这个基本的前提假设是否成立，却很少有人争论。正如 Mel'čuk (1988) 指出的，组块分析方法的提出者必定是以英语这样的组块语言为基础的。那么建立在这种语言模型基础上的分析方法能否适用于其他的语言是需要验证的。汉语的研究在借鉴这种分析方法的时候，首先也需要反思汉语是否是组块型的语言。或者换一种说法，是否能用组块的分析来涵盖所有的语言现象是值得争论的。以现有的研究成果来看，基本上都是默认了组合是发生在相邻的两个成分之间的。这里，我们提出组合是否必须相邻这个问题，是希望引起在方法论上的重新思考，如果组合并不都是发生在相邻成分之间的，那么词序作为语法研究中重要的关注点，很可能是一个独立的层面。同时，这样的研究也在另一个角度说明了依存描写

² 并不必然是相邻的两个词，可以是 X (YZ) 的形式。因此我们用相邻的两个单位来指称。

的强大能力。对交叉依存现象的关注，即是我们对这个问题思考的一个切入点。

正如 Mel'čuk (1988) 所说的，研究非投射性现象的允许条件是个非常重要的任务³。对汉语中交叉依存现象的分析与研究具有较重要的理论与实践意义。

投射性假设下的依存表示与跟短语结构的表示具有很多类似性。在投射性假设的基础上，Gaifman (1965) 曾论述了依存语法与短语结构语法的等价性。由于在投射性范围内的等价，人们认为既然短语结构语法已经被广泛接受，依存语法又与之等价，就没有必要另外采用一种表示方式。这就直接导致了依存语法及依存表示长期以来式微的局面。

投射性的假设对于自然语言的描述是不全面的。其割裂了分属不同组块的词语之间的句法联系。其最初的假设是基于这样的考虑，如果包括非投射性的现象，则会使得句法分析的复杂性增高，降低系统处理的效率和健壮性。因此，面向真实大规模真实文本的依存句法分析器基本都限于投射的规则。同时，由于非投射的现象在整个语言中所占的比例不高，一般为 15%~25%，在总体句法分析准确率不高的情况下，这些情况可以暂时不予考虑。但是，从理论的角度来说，这些现象的正确处理是实现最终目标研究的渐进过程中一个必须要面对的问题。我们知道，句法驱动的语义分析是按照组成性原则来进行的。在这样的语义分析中，语义组成部分应该与句法的组成部分匹配。如果不能够分析出相应的句法树，则语义的正确处理就没有可能。因此，对于非投射性现象的考察研究显得越来越重要，而且，这是依存分析的真正优势之所在，对其现象规则的形式化，可以直接促进信息处理的效率提高。

³ “The study of the conditions under which non-projectivity is allowed is a fascinating task.” (Mel'cuk 1988), P. 38.

1.3 研究方法与目的

本文的研究工作是在依存语法的理论框架之下进行的。因此使用的是依存语法的判断和描写标准。具体包括依存关系成立的判断原则、依存关系种类的区分原则和依存关系描述方式等。我们从依存语法的角度来考察汉语中交叉依存现象有多少种，以及其成立的句法、语义条件，进而探讨这些现象形成的制约原则及深层机制，及其在信息处理领域的一些应用。

具体来讲，首先通过观察语料库中实际的语料，根据句法依存的判定标准，挑选出符合交叉依存现象的句子。由于对汉语中交叉依存现象的分类目前没有前人的研究可以借鉴，我们必须从语料中自己寻找例句，然后考察其具体的分类。我们选择宾州大学汉语树库（U-penn Chinese treebank）的语料作为研究的基础，是由于宾州大学树库的语料是比较权威的，经常用作各种句法分析器学习和评测的基础。

第二步，通过依存关系的判断标准得到若干交叉依存现象的句子，根据其最明显的区别特征作出大致的分类和细分的小类。对每一个小类的现象，都考察其成立的句法、语义条件，以及其语序上的特点，寻找不同的交叉依存现象之间共有的特征。

第三步，在观察这些共同点的基础上，由交叉依存现象句法语义方面的共性来分析其形成的制约原则。考察是什么原则制约着交叉依存现象的发生，以及这个原则背后的深层机制；

我们提出两条基本的原则来对这些交叉依存现象形成做出统一的理据性解释。我们认为，交叉依存现象的发生是由两个原则共同制约形成的：

第一个原则是词语组合时的属性和谐原则，句法依存关系的存在是以语义上的属性和谐为基础的；

第二个原则是环形语序域原则，不同语言对词序的安排上具有

不同的辖域。

第四步，探讨交叉依存研究结果的应用。针对应用的需要，把研究得到的系统性理论应用于指导语料库的构建工作。建立依存图构成模式以及依存图库构建的具体原则。

本文的主体内容即是通过交叉依存现象的研究对这两个原则进行具体的解释与证明。也就是说，我们的研究目的是考察汉语中交叉依存现象的分布情况，在此基础上提出一个能够系统分析和处理此类现象的理论。并且反思现有依存树库构建的问题，提出依存图库构建的具体规则。

1.4 本文的结构安排

本文的内容分为七个章节：

第一章介绍了本文的选题缘由、具体的研究对象、研究方法和研究目的等；

第二章对交叉依存现象研究涉及到的几个方面的以往研究成果进行了简要的回顾，指出以往研究的成绩与不足之处，从而梳理出本文的研究方向和研究框架；

第三章详细描述了交叉依存现象在汉语中的分布情况及其不同的类别，考察了各类现象的句法和语义条件，以及不同类别现象之间的相通之处；

第四章提出了制约交叉依存现象形成的两个基本原则之一的属性和谐原则，这是句法依存关系成立的制约原则。这一章对这条原则进行了深入地探讨；

第五章提出了环形语序域原则，这是词序分布上的制约原则，也是交叉依存现象语序的制约原则之一。在这一章对这条原则进行了深入地探讨；

第六章是应用部分，提出了依存图库构建原则，本文研究的交

叉依存现象也包括在其中；

第七章对全文进行了总结，指出了本文的创新点以及不足之处，并且提出了下一步的研究方向。

2 文献综述

2.1 依存语法综述

2.1.1 依存语法研究概况

依存思想的语法体系曾经在历史上的若干个世纪中占有主导地位，特别是在古代欧洲和斯拉夫地区。有学者（Kruijff, 2002）把依存的源头追溯到 Pānini 时代的梵语语法。但一般来说，人们还是把法国语言学家 Tesnière 作为现代意义上依存语法的创始人。从 Tesnière 开始，依存的句法体系开始复兴，有很多类似的语法体系开始出现。

依存的概念来源于 Tesnière（1959）的《句法结构》一书。依存是指构成句子的词语之间的两分的不对称关系。在 Tesnière（1959）《句法结构》的第一章中表述为：“句子是一个有机的整体，其组成部分即是一个个的单词。在相邻的词汇之间，我们可以感知他们之间的联系，这种联系构成了句子的结构。这种结构的联系建立了词汇之间的依存关系。有直接联系的词语之间具有不对称的关系。上级词语可以称之为‘支配词’，下级词句称之为‘被支配词’⁴。”也有很多学者分别称之为中心词（head）与依存词（dependent）。鉴于“中心词”这个术语容易与 Bloomfield 所定义的专指向心结构

⁴ La phrase est un *ensemble organisé* dont les éléments constitutifs sont les *mots*. [1.2] Tout mot qui fait partie d'une phrase cesse par lui-même d'être isolé comme dans le dictionnaire. Entre lui et ses voisins, l'esprit aperçoit des *connexions*, dont l'ensemble forme la charpente de la phrase. [1.3] Les connexions structurales établissent entre les mots des rapports de *dépendance*. Chaque connexion unit en principe un terme *supérieur* à un terme *inférieur*. [2.1] Le terme supérieur reçoit le nom de *régissant*. Le terme inférieur reçoit le nom de *subordonné*. Ainsi dans la phrase *Alfred parle* [. . .], *parle* est le régissant et *Alfred* le subordonné. [2.2] (Tesnière, 1959, 11–13, emphasis in the original)

中的中心词概念相混淆，我们建议使用“支配词”与“被支配词”的称谓。

Tesinière 的研究一般认为是属于句法层面的，但也有一些争议。随着研究的深入，依存的概念已经细分为适用到语言的各个层面——形态的、句法的、语义的、语用的层面等。依存语法可以把语言若干不同层面的研究统一在同一理论体系之下。

对于句法依存关系，Mel'čuk (1988) 有过较为详细的表述，其经常用作依存关系研究的参照。Mel'čuk 指出，讨论句法依存关系有三个基本的假设：

首先，讨论的只是表层的句法。其次，仅仅讨论语法完整的句子，省略现象与口语不在讨论的范围之内。第三，只讨论词汇间的不对称关系。

在这个前提之下，我们讨论句法依存的具体规则。句法依存的规则应该包括三个方面的标准：首先，判断是否存在句法依存的标准；第二，判断句法依存方向的标准；第三，区分依存类型的标准。

判断是否存在句法依存，有两个基本的条件。

A 线性相关；在一个句子中，如果 W_1 与 W_2 两个词中某一个的线性位置必须参照另一个才能够确定，那么我们就说这两者之间存在直接的句法依存。但是，这并不意味着一个词对另一个词位置的决定方式是唯一的。其实质是说，当我们提到 W_2 的位置时，必须提到 W_1 。比如冠词，当我们提到冠词的时候，必然会同时指出有名词位于其后。这样，冠词与名词就具有直接的句法依存。

B 韵律相关；即 W_1 与 W_2 可以形成一个韵律单位。⁵

这两个条件需要同时具有，才是确立存在句法依存的依据。

判断句法依存的方向，也有两个基本的标准。

⁵ 韵律可以表示组成部分的边界，因此一个韵律单位之间往往具有句法依存关系，以至于 Mel'čuk 做为判断句法依存的一个依据。但韵律单位与句法单位之间并不是绝对的对应关系。有时候为了满足韵律的需要，会牺牲一些句法单位的完整性。例如，“一衣带水”的句法单位划分应该为“一/衣带/水”，为满足双音节韵律的需要，在使用中这个成语被划分为“一衣/带水”的韵律单位。

首先，是看其对句法结构所能充当的句法成分的影响力的大小，影响力大的词语作为这个句法结构的支配词。例如，数量短语可充当的句法配价的能力是由其名词所决定的，所以其句法依存的支配词为其中的名词。

第二，是选择与结构外词语有形态依存关系的词语。

两个标准不是等同的，在第一规则不能奏效的情况之下才考虑第二规则。同时需要注意的是，有两个标准不能够作为判断依存方向的标准。一个是语义的可省略性；二是预测能力的强弱。一般来说，被支配词对整个句法结构的预测能力要大于支配词，比如形容词对名词的预测能力要大于名词对形容词的预测能力。但是也有不同的情况。比如介词，对后面的名词预测性要大大高于名词对介词的预测性，但是由于介词在这个句法结构中对于其充当句法配价的功能具有更大的决定性，我们依然选择介词为支配词。

区分依存类型的标准有三个：

第一，如果一个词对内部词序的不同可以形成不同的语义关系，则依存关系不同；

第二，如果子树可以循环替代，则依存关系不同；

第三，相同词类对同一节点的可替代与否不同，则依存关系不同。

句法依存具有以下几个特点：

第一，与形态依存不同，句法依存是广泛存在于所有语言当中的。没有一种语言可以不存在句法依存，没有一种语言中的词语可以游离于句法依存之外。在所有的线性语言中，句法依存连接句子中所有的词语，他们总是形成一个连接的结构(**connected structure**)。

第二，句法依存不能够是双向的，即两个词语不能同时即 A 依存于 B，又 B 依存于 A。这是依存树表示语言的一个形式特征。

第三，在一个句子中，一个词语只能有一个父节点，即支配词的唯一性。(Mel'čuk1988)

以上的论述主要参照的是 Mel'čuk (1988) 的标准。我们将会在这些句法依存的原则之下进行交叉依存的考察。

如果把基于依存的基本假设的语法理论都划归到依存语法这个大类之下，可以形成一个庞大的依存家族。包括 Word Grammar (Hudson, 1984, 1990), Functional Generative Description (Sgall et al., 1986), Dependency Unification Grammar (Hellwig, 1986, 2003), Meaning-Text Theory (Mel'čuk, 1988), Lexicase (Starosta, 1988) 等语法体系。另外，基于特征限定的 (constraint-based) 依存理论 also 具有很强的传统，有 Constraint Dependency Grammar (Maruyama, 1990; Harper & Helzerman, 1995; Menzel and Schröder, 1998), Weighted Constraint Dependency Grammar (Schröder, 2002)。在 Constraint Grammar (Karlsson, 1990; Karlsson et al., 1995) 基础上发展出来的 Functional Dependency Grammar (Tapanainen and Järvinen, 1997; Järvinen and Tapanainen, 1998), Topological Dependency Grammar (Duchier and Debusmann, 2001), Extensible Dependency Grammar (Debusmann et al., 2004) 等。另外，还有把依存语法与范畴语法综合起来的 Dependency Grammar Logic (Kruijff, 2001) 等。

这些理论都基于这样的一个基本的假设：句子的结构是由具有两分的不对称关系（依存关系）的词汇组成的。在这个基本假设的基础上，不同的理论针对具体的情况又各有自己不同的看法，形成了在细节上各有特点的语法理论体系。因此，依存结构的最一般特征即是没有像组块理论那样的非终极结点。

依存语法的兴盛是由其自身的优势所决定的：

首先，基于依存的表示方式与语义关系具有较天然的对应性，容易从句法层面对应到语义层面。而且，依存语法体系可以贯穿形态、句法、语义、语用各个层面。我们知道，目前对语义的表示基本上都使用的是支持谓词论元结构的形式，如一阶谓词验算表示法、

语义网络表示法、概念依存图表示法和基于框架的表示法等（冯志伟 2004）。如果在句法层面采用短语结构的表示，语义层面又采用其他的表示形式，两个层面就相当于运用了不同的语法体系，由于理论的差异，在句法与语义的对应上必然会出现困难。如今，语义的分析与处理已成为语言学及计算语言学研究的重头戏，更被一些应用处理系统所青睐。所以依存语法与应用的研究越来越受到人们的重视。依存语法体系可以整合语言的若干层面，有利于语言的立体研究。

第二，自二十世纪七十年代开始，依存理论开始得到了认知语言学的赞同以及心理学试验的有力佐证。

从结构主义的布龙菲尔德开始，一直到主张转换生成语法的乔姆斯基都沿用了组块分析的范式，组块分析的方法其实是源于心理学上模块认知的基础。模块认知的理论影响到语言学研究，形成了在近半个多世纪以来组块分析占主导的现象。从二十世纪 70 年代开始，直接成分分析语法在很多语言的具体问题上受到了质疑。同时，认知心理学上网络认知的观点也开始影响学者们重新思考基于依存的语法在对语言现象解释上的适用性。例如，扩散性激活（spreading activation）等认知模型可以解释很多以前模型认知理论所不能解释的现象。而且也得到了实验的证据（Swinney 1979）。

Hays（1964）曾说过，容易得到的调查数据往往由于学校教育的影响而不能真实反映语言习得的事实。为了避免这种偏颇，有些心理语言学家通过对学前儿童语言习得的研究得出了依存语法更符合儿童语言习得规律的结论。Ninio（1994; 1998）指出，短语结构基础上的语言习得理论不能够解释儿童最早期时的词语合并现象，使得儿童语法系统分为两个非连续的阶段，前语法阶段（pregrammatical stage）和语法阶段（grammatical stage）。但是，无证据证明儿童习得语言过程中有截然不同的两个阶段，因而不是一个完善的理论体系。她认为，依存语法可能更加符合人类习得语

言的本质，人类习得语言的过程就是习得一个个词语对之间的依存关系的过程。Ninio 通过实验分别研究了儿童习得英语和希伯来语三词句结构的情况，发现儿童对于支配词（head）直接相邻的结构习得要比被支配词（dependent）与支配词间隔结构的习得更加容易。而且，依存的方向及被支配词数量也影响其习得的难度。

彭艳（2007）通过短语结构语法与依存语法的对比，也从认知的几个方面论证了依存语法更符合人类语言习得的本质的观点：

- 依存距离作为衡量人脑记忆处理负担的标准要更自然一些，长距离依存关系在头脑中的处理要比短距离依存关系更难。
- 语言中有要求依存方向一致性的趋势，虽然认知简单，但处理更难，功能的要求决定了依存方向上的多样性。
- 语法关系是依存关系的次分类，人们对依存关系的认知与非依存关系的认知是一样的。
- 依存关系范畴与其他范畴一样符合原型理论，有原型的依存关系形式，同时也有少量的例外现象。
- 人脑使用词汇及其之间关系的形式来储存词汇知识，要比使用短语结构的储存形式更容易。
- 人脑对依存结构可以通过词语的相邻出现而类推习得，并通过语言实践中逐渐变得复杂化。

第三，Chomsky 形式语法的广泛应用一个重要原因在于其在计算应用上的效率。在依存语法早期的形式化模型中，由于 Gaifman (1965)把依存限定在了投射性方面，且论证了在此范围内依存语法与短语结构语法的等价性，在当时美国语言学派占主导的潮流下，依存必然处于了一个比较边缘的位置。二十世纪七十年代以后，语言研究开始由注重个别语言的语法规则向更具普遍性的语法规则转向。随着考察语言种类的增多，现实语言的多样性使得人们寻找能够描写更多语言种类的语法体系。依存的表示与研究是基于更广泛语言研究的需要。特别是针对词序比较灵活的语言而言。主流的

形式语法学派，如短语结构语法（PSG）、中心词驱动的短语结构语法（HPSG）对于大部分的语言可以说是适用的，特别是类似于英语的组块式语言。但是在对于非连续结构的表示与分析上依然会捉襟见肘。而且，HPSG本就是注意到了这些词序灵活的现象而修改而来的。对于词序灵活的语言来说，如拉丁语、德语、捷克语、土耳其语等，这种劣势更加明显。随着考察语言种类的逐渐增多，人们开始寻找能够在更广范围内形式化描写不同语言的语法体系。依存语法就成为了一个比较合适的选择。这样的语法体系，可以允许在同一理论框架下进行更多不同语言间的比较研究，因而具有更大的适用性。而且，依存的表示在词序限制上比较松散，比较适合描写真实语料，这也是计算语言学界更多地付诸于依存语法的原因之一。同时，词汇主义的兴起也影响到了主流语言学的发展。短语结构语法基础上演化出来的 HPSG 和树邻接语法（Tree Adjoining Grammar）等都可以看作是对词汇在语法系统中重要作用的确认和强调。

在认同依存关系是词汇之间的不对等关系的基础上，不同的语法体系针对不同的语言实际又各有所发展，各自具有了不同的特色。这里不再对每一个语法体系作详细的介绍。总体来说，依存的语法体系包括两个基本的方面：一方面是语言分析的理论，另一方面是语言表示的理论。依存的复兴，是与其分析的便捷性和表示的明晰性紧密相关的。对于语言理论研究来说，对第一方面的特点可能更侧重一些，在计算语言学领域，则是两个方面都要兼顾的。

强调语言分析还是强调语言表示导致了对依存理论的不同侧重。以往的理论中对具体依存关系的判断标准、关系种类等细节的问题探讨很多，至于表示方面的特点则观点比较一致，像允许比较灵活的词序等。但是随着语言信息处理的需要，依存的语言分析中的词序问题逐渐凸显出来。从而再次引出了依存理论中对于句法关系与词序关系的争论。在以往各种依存语法理论中，即有关于句法

关系与线性词序关系是否一致的争论。词序与句法之间的关系，可以有两种看法，一种认为依存的句法关系与线性的语序无关，如，Tesinière（1959）就认为依存关系是属于结构次序，不同于线性的词序。因此，句法依存的表示中是不考虑词序的。大部分依存语法也随着 Tesinière 的观点，认为词序是独立于句法层面的。由此可以得出依存语法具有对词序要求不那么严格的优势。同时，另一种观点则认为句法关系与语序是紧密相关的，语序是句法关系表达的重要途径。

对这个问题，我们需要辩证地看待。语序本身是表达句法关系的一个重要工具，但是，语序不是句法关系本身。所以，认为依存的句法关系与语序无关有其合理性。但是，所有的语言都有用语序表达句法关系的情况，认为句法关系和语序紧密相关也是正确的。在论述依存的优势时，经常会提到其允许较自由的语序。对于这个特点，我们需要从以下两个方面来理解。第一，依存的表示形式中通过边的表示把原来由语序承担的句法关系明确化了，因而在表现形式上可以不再考虑语序的制约，使用比较自由的表现形式。第二，在语序灵活的语言中，一定程度上的语序灵活其句法关系不变的说法是从宽泛意义上来说的。严格来说，语序的不同所反映的句法关系还是有细微的不同的。在语言分析的过程中，必须要参照语序所承担的信息。把句法关系与语序分离的观点在客观上导致了依存研究中对表示方式和句法的细节研究一头独大的局面，对语序问题的关注上不够。这是由于早期的语法研究主要是人工进行的，可以先理解再分析，语序问题就不那么突出。而对于机器的分析来说，必须是先分析了才能理解，是不同的过程。

对语序这个概念，学者们也有不同的看法。一种认为词序是从句子的整个布局来看的，可以称之为宏观的语序；另一种是从微观的角度来看的，即具体词汇之间的相对位置。这又可以分为两种，一种是沿用了结构主义的划分，即先划分一定的结构，然后再观察

结构内词序的特点，另一种则是纯粹的依存划分，即相对词序的划分，只关注具有依存关系的词语之间的相对词序，而不管具体的功能结构。

句法结构的实现在与实际的词序结合在一起表示时，则会发现一个有趣的现象，即称之为投射性与非投射性的现象。相应地，在对具有词序关系的句子来进行机器的句法分析时，则必然要考虑词序与句法依存的关系，则怎样处理非投射性现象的问题。

立足于组块理论的短语结构语法虽然预设了词序与句法结构之间的对应，或者说是把词序也看成了句法结构的一个重要部分，但也认识到非投射性现象的存在是其理论的一个巨大挑战。因此很多学者用短语结构的语法来尝试解释论元分离、移位等现象。从逻辑的角度来看，这是与其组块的假设相违背的。因此不得不用深层到表层的转换来作为解释。在深层坚持了组块的假设，而在表层出现语序的不同。在表层的语序上，也是描写的多，解释的少。

一般来说，大多数依存处理系统是假设依存结构的投射性特征的，而大多依存语法语言理论的研究则没有假设投射性的特征。更准确地说，大多数依存语法的形式化理论以投射性的要求为标准，同时也认识到对非投射性依存结构（例如，长距离依存）表示的必要性。对非投射性依存现象关注较少主要是因为这种现象在语言中所占的比例相对较小，经常采用简单的处理方法绕过了这个问题。

依存语法理论单纯从句法关系来看依存，不考虑语序的做法对于面向人的语言分析可能问题不大。因为大部分的依存关系都是发生在相邻的成分之间，个别的长距离的依存关系因其突出性容易被人察觉到。因此可以只指出特殊部分依存关系的存在而不深入探讨制约语序的原因。或者如范畴语法（category grammar）中把语序的描写附加到词语的合并规则中去（Mary 1993）。但是面向机器的依存分析必须非常重视语序，否则就无法判断依存关系的存在。因此语序在这里显得特别重要。同时，对语序的忽视也使依存理论对

各种语言语序的成因上解释的力度较差。这是一个完整的语言理论需要补足的地方。

一些面向应用的理论系统在这个问题上做出了改进，开始把词序作为一种相对独立的层面来处理。从而形成了依存多层性的理论，这些理论在某些特定层面允许非投射性现象，把词序作为一个独立的层面来处理。例如，Functional Generative Description 的表示中 tectogrammatical 层面是投射性的，而在 analytical 层面则允许非投射性依存的表示 (Sgall et al., 1986)。鉴于德语语序比较自由，非投射性依存普遍存在的现象，一些学者把依存在德语中区分了词序依存 (Topological dependency) 层面与句法依存 (syntactic dependency) 层面，把词序层面独立描述。根据 Reape (1994) 论述的德语词序域原则，Gerdes 和 Kahane 把德语的词序划分为五个主辖域 (main domain), 【Vorfeld, left bracket, Mittelfeld, right bracket, Nachfeld】，并且形式化为独立的拓扑依存语法 (Topological dependency grammar)。假设在词序层面 (Linear precedence) 的表示必须是投射性的，而在句法层面则是无序的，句法层面的表示可以通过扁平化 (flating) 对应于词序层面，因而允许非投射性的表示 (Duchier and Debusmann, 2001)。句法层面与词序层面两者之间是正交且相互限制的。这样的做法便于形式化，对于我们的研究具有启发意义。

不同的理论对句子依存表示的分层是不同的，这与语言的特点有关。仅就句法关系来说，Mel'čuk 的 MTM 模型把句法依存关系分为浅层和深层两种。在浅层句法层面，所有的词语都在句法树中得以呈现；在深层句法的表示中，则省略了助词和功能词，并把俗语用一个节点表示，以便于向语义网络的对应。在汉语的分析中，首先需要的是分词的操作，因此可以直接把惯用语和成语用一个节点来表示。同时，汉语中不存在很多类似印欧语中的助词和功能词，不必特意区分深层句法和浅层句法，只做一个句法层面即可。

汉语的依存研究始于冯志伟（1983）发表在《国外语言学》上的一篇介绍性文章。此后，汉语研究上出现了研究配价理论和依存语法的热潮。严格来说，汉语学界对配价的研究与依存的研究是受不同的直接来源影响的。配价的研究主要是受德国配价语法的影响，主要研究动词各个层面所要求带的必有成分，即动词的价，后来扩展到对名词和形容词配价的研究。汉语的配价研究往往与传统语法和结构主义方法的研究相结合，形成了颇具中国特色的配价理论。而依存的研究则大致是直接从法国 Tesnière 那里学习而来的。德国配价的源头也是特氏的学说。因此两种研究基本上又是同质的。只是配价的研究侧重具体词语的研究，主要研究价携带者的补足语问题，而依存语法则要同时考虑补足语和说明语两个部分。补足语与说明语的分别是等级的划分，而非层次的分析，是以语义为基础的。配价研究属于对部分的研究，依存研究则更重视整体。配价的研究是依存语法重要的组成部分。

配价的研究在汉语语言学界的影响很大，渗透到语言研究的各个层面----句法、语义、逻辑等。很多以往的争议问题在配价的思想下得到很好的解决。最早把配价应用于汉语研究的是朱德熙。他的《“的”字结构和判断句》即是用配价的思想很好地解决了“的”字结构的问题，引起了学者们对配价语法的广泛兴趣。刘涌泉、乔毅（1991）更是介绍了配价语法理论及其成果对机器翻译等信息处理方面的作用。配价研究代表性的学者很多，如沈阳，郑定欧（1995）、陈昌来（2002）、袁毓林（2010）等，都有关于配价的专著出版。袁毓林把“价”的概念分化为有“联、项、位、元”四个平面构成的配价层级，形成了全新的配价思想。至于研究性的论文则更是不胜枚举。如文炼（1982）《词语之间的搭配关系》、文炼和袁杰（1990）《谈谈动词的“向”》、范晓（1991）《动词的“价”分类》等。刘丹青（1987）《形名同现及形容词的向》和袁毓林（1992）《现代汉语名词的配价研究》更是把配价的研究拓展到形容词和名词的范围。

值得指出的是袁毓林的研究，他提出的四个平面的配价层级，已经开始把句子宏观的词序配置与词语的配价结合在一起，即配位方式的研究。配位方式指的是“施事、受事等语义成分跟主语、宾语等句法成分之间的连接和对齐关系。”⁶只是其配位的研究还是延续了结构主义的研究思路，即在组块的范围内探讨语序的配置，对分属不同组块之间的词语没有做统一的探讨。

依存的研究则在计算语言学的学者中更为广泛一些。以冯志伟（1983；1999）、黄昌宁、周明（1994）、刘海涛（2009）等为代表的学者们把汉语依存语法的研究推向了深入。冯志伟不仅最早在汉语界介绍了 Tesnière 的依存语法，而且最早使用依存语法进行了自然语言处理的尝试，取得了不错的处理效果。刘海涛、冯志伟（2007）还提出了“概率配价模式”，实现了对依存关系进行定量的研究。刘海涛（2009）研究发现，汉语的依存距离为 2.81，大于英语 1.386 和日语 1.43 的依存距离。这将有助于把语言作为一个复杂网络来研究。

英语中，词语的词性和功能之间具有较为直接的对应关系。因此，按照词性进行的片段分割是与功能的片段分割相对应的。汉语中，词性与功能之间则不存在单纯的对应关系。所以根据词性序列划分出来的片段必然会有多种功能的表达。因此，屈折语中的句法结构层次性是功能的层次性相对应的。汉语中则不能按照词性的序列来推导功能序列。这值得我们思考按照短语结构的片段划分来研究汉语是否适当。组合的词性越多，则功能的种类越多。我们认为，用最小单位的组合即两个词语之间的组合关系的语法体系（依存语法）来研究汉语可能会更加适当。

但是，现有关于汉语的依存研究也存在很多的问题。其理论多是直接沿用的以往，并没有深入讨论其在汉语中适用的情况。依存的一般做法是探讨词类范畴之间的依存关系，但是对于词类与功能

⁶ 袁毓林《汉语配价语法研究》2010，页 112。

之间无简单对应关系的汉语来说，其复杂性是远远大于其他语言的。因此对于依存的概念需要作出细化的修订。目前对理论假设的前提讨论不多，更多的情况是与现有的不同流派研究结合在一起，着眼于对具体问题的解决上。这主要体现在配价的部分，对依存理论本身的系统性上考察不足。正如冯志伟（2009）所说的：“在具体的细节研究方面取得了很好的成果，但是在系统性和整体性方面还显得很不足”⁷。

而且，汉语的依存研究还存在一个重要问题，即是从分析的角度谈的多，从生成的角度谈的较少。从分析的角度或从生成的角度来研究语言其侧重点是不同的，其涉及的因素也不尽相同。从分析的角度需要研究词语之间配价、依存的关系，以及怎样形式化的表示，这也正是目前依存研究的主要内容。但是，同时从生成的角度来看是非常必要的。这是一个问题的两个方面，不应该顾此失彼地只强调某一个方面。这里所指的生成不是生成语法讨论的生成，而主要是指对线性序列形成的制约原则。在应用中例如机器翻译的从源语言到目的语转换中将是极其重要的。怎样形成合法的线性序列是以往依存研究中较为忽视的一个方面。

2.1.2 非投射依存现象的研究

非连续性现象在所有语言中都是存在的，其中大部分都属于非投射性依存现象。从 20 世纪 80 年代开始，部分依存语法的学者开始关注到非投射性依存现象的研究。

Mel'čuk（1988）列举了拉丁语、法语、德语、英语和俄语中的几类非投射性依存现象。并且指出，所有非投射性依存现象都在一定程度上是有标的（marked）。其原因是由于强调（emphatically）、语体风格（stylistically）、互动（communicatively）的因素，特别

⁷ 引自冯志伟为《依存语法的理论与实践》（刘海涛 2009）所写的序言。

是在口语语体中，经常会出现这类现象，又或是由于句法上保有特殊的附着形式。Mel'čuk 同时强调，研究在哪些条件下允许非投射性依存现象是极有意义的工作，特别是在语言类型学研究方面。而且，这类现象只能在依存的框架下得以明确地表示。Dikovsky(2001)指出，罗马语和德语中代词和 wh 词的 extraction、疑问句、关系从句、话题化、连词或介词的连用、非连续的否定、法语中的代名词附着等都属于非投射性依存现象。

心理语言学的研究数据也表明，形式上投射与非投射的不同并不必然伴随着处理难度的不同。Bach(1986)甚至发现，人脑对荷兰语(Dutch)中跨系列依存(cross serial dependency)的非投射性现象的处理要比对德语中 Nested dependency 的现象容易。因此，并不能把处理困难作为逃避处理非投射性依存现象的理由。

对非投射性依存现象，在依存语法的理论范畴内的探讨数量上并不是很多，只是列举性地指出了一些非投射依存的种类，并没有系统地探讨其形成的原因和制约原则。Mel'čuk(1988)和 Dikovsky(2001)的论述也只是列举性地指出了几种语言中的几类现象，并没有就具体一种语言做深入的分类和考察，对其形成的原因和制约因素以及几类现象之间的联系都较少涉及，因而只是一个初步的思路，需要更多进一步的研究。在计算语言学方面的研究则多是具体分析的算法和应用，也没有做理论的分析 and 指导。

在转换生成语法中，也有对部分非投射性依存现象的讨论，一般称之为长距离依存(long-distance dependency)现象。包括 wh- 疑问(wh-question)、关系从句(relative clauses)、话题化(topicalization)、easy 的用法(easy-adjectives)等。举例如下：

Wh-questions:

(1) What did you find? (2) Tell me who you are looking for?

Relative clauses:

(3) the item that I found (4) the guy who(m) I talked to.

Topicalization:

(5) The book, I can't find (6) Mary, you should talk to.

Easy-adjectives:

(7) My place is easy to find. (8) That man is hard to talk to.

在转换生成语法中，用移位来解释这个结构的形成。其中形成隔断（gap）的词语之间格一致关系、介词的选择都在移位前应用。这些结构一直以来都被认为是由深层到表层转换的证据。

由于依存语法理论中对语序层面的问题探讨不是很多，在形式化表示与分析、语料库标注中，都是沿用了以往投射性的假设。因而对与语序紧密相关的非投射性现象几乎较少关注。这是现有研究中一个很大的空白地带。随着句法分析准确率提高的要求，对这类现象系统性考察的必要性也凸显出来。

目前对非投射性依存现象的探讨，有两个明显的特点，第一，涉及的多为有边跨过最高结点的类，对交叉依存类的现象涉及较少。第二，探讨到的语种中极少涉及到汉语。而在汉语中，没有类似印欧语的疑问句的前置，定语从句，非连续的否定，附着形式等现象，只有前面谈到的话题化、连词或介词的连用现象存在于汉语中。而汉语中一些明显的交叉依存类现象是依存研究中没有被注意到的，而且很可能是较复杂的类别。

从依存语法角度对汉语中非投射性依存现象的专门研究，目前还没有见到。这或许是因为，第一，研究依存语法的学者数量本身就少，且主要集中在计算语言学方面；第二，对依存语法目前还是基本停留在借用的阶段（尽管对作为依存语法组成部分的配价理论有很多方面的发展），还没有关注到这些较少类的现象。

汉语中有多少类非投射性的现象，各有什么特点，其形成的原因是什么，这些理论问题都需要更多的研究。投射性假设理论指导

下的分析过程把非投射性结构按照投射性结构分析，句法结构的分析结果必然会出现错误，在对应到语义关系时也不可能正确。因此，需要修正这些错误。对非投射性现象的关注，无论是在理论的角度还是在应用的角度来看都是非常必要的。

2.2 计算语言学领域内的依存分析

依存语法在计算语言学领域的广泛应用，是与其自身的特点密切相关的。从语言工程的角度来看，相较于短语结构的句法分析，依存分析也具有很多的优势。

首先，在基于依存结构的分析器中，每个词语都是单独的结点。依存分析器的任务只是连接现存的结点，并不假设新结点，因此分析的任务更加直接。而且，即使一个结点分析错误也不影响其他结点的分析结果。而在短语结构的分析中，由于假设了很多非终极节点，一个结点的分析错误会直接影响后续部分分析的结果。同时，基于依存的分析一次只处理一个词语，比起短语结构分析要等整个短语分析成功才能完成一个步骤来说，速度要快很多。所以依存分析器的处理效率更快。这对于面向大规模真实文本的机器翻译系统来说是个很大的优势(Convington, 2001)。另外，边交叉的情况也是上下文无关语法不能够而只能在依存树下形成的。

依存句法分析大致来说有两种基本的策略：一种是规则驱动（Grammar-Driven）的策略；另一种是数据驱动（Data-Driven）的策略。两种策略及其内部的具体方法之间并不是绝对互斥的，规则驱动和数据驱动只是个大致的分类。也有很多尝试把两者结合在一起的处理思路。

规则驱动的方法大致来说又可分为三类：

第一类方法是依存语法规则的直接形式化。

这类方法最早可以追溯到 Hays (1964) 和 Gaifman (1965) 对依

存语法的形式化表述。Gaifman (1965)概括为三条形式规则。

① L1: $X(Y_1, \dots, Y_i, *, Y_{i+1}, \dots, Y_n)$ Y 分别表示 X 左右的从属词, $*$ 表示 X 在句中的位置。

② L2: $X(*)$ 表示 X 在句中没有任何从属性, 是终极型规则;

③ L3: $*(X)$ 表示 X 不是任何词的从属词, 是初始型规则。

形式化的依存语法与短语结构语法之间具有紧密的联系。Gaifman (1965) 证明了这种依存体系与上下文无关语法在一定范围内的等价性, 即依存系统可以转换为相等的上下文无关语法。这里所指的一定范围, 即是投射性的限定。鉴于这种等同性, 可应用于上下文无关语法的算法也可以应用于依存语法。Hays (1964) 使用的自底而上的动态规划算法就非常类似于上下文无关语法应用的 CKY 算法与 Earley's 算法。而且, 动态规划算法也成为近年来语法驱动的句法分析方法中一个显著的趋势。Eisner (2000) 提出的 Bilexical grammar 可以基本概括这个处理的传统。在 Eisner 的形式化中, Bilexical 语法由两个部分组成:

1. 由终极符号组成的词汇集合 V , 包括根节点符号 Root;
2. 两个决定性有限状态机 l_w 和 r_w 。每个自动状态机接受符合规则的子集 V^* 。

由 bilexical 依存语法 G 定义的语言 $L(G)$ 分析过程为 l_w 从右到左地接受词序序列 (可能为空), r_w 自左到右地接受词汇序列。Eisner 提出的分析算法也是动态规划的算法, 计算的复杂性由 $O(n^5)$ 减少到 $O(n^3)$ 。

第二类方法为基于受限满足的分析方法。

基于这种思路的第一个分析系统为 CG (Constraint Grammar) (Karlsson, 1990; Karlsson et al., 1995)。其核心思想为通过一定的形式框架把依存句法分析看做是加标与排除不当分析的过程。后来 Maruyama (1990) 扩展为 CDG, 其分析另外包括了词汇间依存边标记与中心词。受限满足的方法大致来说是 NP 满足的问题。其特

点在于把句法分析问题转变为标记(taging)和分类(classification)的问题。但是由于其只利用局部信息(local information), 处理的效果不甚理想。

在后来发展的 TDG (Topological Dependency Grammar) 中, Duchier (1999, 2003) 使用了多层面的表示, 把限制特征分化为分属不同层面的限制特征, 例如, 直接支配(Immediate Dominance)层面与线性序列(Linear Precedence)层面。

受限满足的方法没有把分析的句子限定在投射性依存的范围之内, 因而可以处理部分非投射性依存的现象。因此在分析的精确度上会比投射性依存分析的结果要高。但是, 对于真实文本来说, 受限满足的处理方法也会出现两个问题, 一种是可能出现没有分析结果满足所有的限制条件的情况, 这将影响到系统的健壮性(robustness)。另一种则是会出现超过两种分析结果的情况, 这样会引起消歧的问题。

第三类方法的思路比较简单, 为决定性分析策略(Deterministic parsing strategy)。Covington (2001) 把其形式化为如下的过程:

把句子中的词语从左向右依次接受, 依次把新接受的词语连接于前面的词语或作为新的支配词。

这种分析的策略能与很多不同的形式语法兼容。Covington 的算法的时间复杂性则是 $O(n^2)$ 。

第二种分析的策略为数据驱动的分析策略。

最初的数据驱动的策略是跟前一种规则驱动的策略结合紧密的, 大多数分析策略为依存语法与生成概率的结合运用, 基本上通过语料库学习得到概率模型然后用于消歧。例如 Carroll & Charniak (1992) 提出的概率上下文无关语法。其上下文无关语法是与 Hays 和 Gaifman 定义的依存语法是等同的。他们通过在任意创造的语料库上的无导学习得到概率模型, 然后用于分析句子, 但是结果不是很好。

一种更加成功和有影响的方法是 Eisner (1996a,b) 提出的。他定义了多个概率模型以进行依存句法分析。数据用有导的方法从宾州树库中的华尔街时报部分习得。在后续的研究中, Eisner (2000) 把几种概率模型都归并到了 **bilexical grammar** 的概念之下, 介绍了一种比较成功的方法为 **weighted bilexical grammar (WBG)**。在前面提到的规则分析的基础上, 赋予每个自动机 l_w 和 r_w 不同的权重, 取得了不错的分析效果。权重的习得是通过条件概率。

Collions 等应用生成概率分析模型进行依存句法分析, 从布拉格依存树库中习得数据, 先把依存结构转换成为扁平的短语结构以习得数据, 分析得到的短语结构树经过转换成为依存结构的表示方式。

Kudo 和 Matsumoto (2000) 以及 Yamada 和 Matsumoto 提出了 **deterministic discriminative** 的方法, 使用空间向量机训练分类器预计下个动作来构建依存结构。分析算法采用 **shift-reduce** 的分析动作。

数据驱动的句法分析策略有两个值得注意的特点, 其分析结果可以与单纯规则驱动的句法分析形成比较, 而且两者的分析可以结合起来, 共同作用以取得更好的分析效率。

但是, 对于本文所关注的非投射性依存现象, 在以往的处理系统中是关注不多的, 只是在最近几年才开始得到学者们的重视。这涉及到多方面的原因: 一方面, 在依存语法形式化的初期即有限定于投射性依存的习惯。很多系统也沿用了这个习惯。同时, 也是因为针对上下文无关语法的各种算法比较成熟, 比较容易把适用于上下文无关语法的算法平移或者稍微改良后使用到依存的语法体系中来。另一方面, 尽管大家认识到上下文无关语法不能够生成边交叉的情况, 但由于现有系统的整体处理效果一般, 而对占整个语言现象中较少部分的非投射性依存现象的处理对提高系统处理精确

度意义不大，因此受到的关注不多。而且，非投射性的假设也直接影响到句法分析的复杂性。在一定意义上，非投射性是与句法分析的效率相排斥的。寻找一种在表达性（expressivity）和有效性（effectiveness）两方面都可接受的形式化理论是很难的。但是，也有一些系统开始注意和涉及到了部分非投射性依存现象的处理问题。这些处理思路也相应地可以分为两种，一种是基于规则的，一种是基于统计的。

（1）基于规则的处理非投射性依存现象的思路主要有以下几个：

第一种思路是在投射性处理的后期引入非投射性结构的处理。如 Link Grammar (Sleator and Temperley, 1991, 1993)，前期把依存分析看做是词汇化了的上下文无关语法的系统，限定为投射性的结构。但是允许在句法分析结束后的后期处理中引入非投射性结构。

第二种思路则是同步的处理，较多使用的是受限满足的思路。Tapanainen & Jarvinen (1997) 提出了一个基于规则的可允许非投射性结构的分析器。其思路是源于 Fred Karlsson (1990) 提出的 Constraint Grammar。Constraint Grammar 的基本思路属于受限满足的方法。这里的不同之处是其表示形式是基于依存的。而且增加了挂接 (linking) 与修剪 (pruning) 的手段，在处理歧义方面比传统的做法取得了更好的效果。

Kahane 等 (1998) 也提出了一个可处理名词前置和疑问代词 wh- 前置类结构的非投射性依存语法。其处理的核心步骤为上移 (lifting)，把原依存树中非投射性部分的支配词的边 (arc) 上移，把边连接的两个词语中的支配词从其句法支配词上移到词序支配词。在形式语法中指出需要上移的具体条件，使一个非投射性的依存图转变为一个投射性的依存图。

第三种处理思路为把线性词序和句法词序明确地区分开的做法 Broker (1998)。在德语中发展出了拓扑依存语法 (Topological dependency grammar) (Duchier & Debusmann 2001)，把依存分为句

法依存与线性依存两个独立的层面。其线性依存层面可以看做是由直接依存上移而得到的。Nivre (2008) 提出, 非投射性依存结构的句法分析可以分为两个独立的过程, 首先把句子中的词语调整为标准的词序, 然后再按照投射性依存的分析方法来进行句法分析。

Bartha (2006) 对匈牙利语 (Hungarian) 的分析中采用了一种折衷的方案, 没有把词序和句法关系完全分离开。而是把词序的描写结合在句法描述中, 使用特征结构的思路在词库中增加句法依存对词序的限定, 使得句法的分析和词序限定结合在一起。匈牙利语中的非投射性依存的现象能够在这样的思路下得以分析。

第四种处理思路的代表为 Lecomte (1992) 提出的方法。其主张的依存树是通过推导而来的。依存树由一些推理的规则----可换性(commutativity)、关联性(associativity)、多模态(multi-modality)----等假言推理规则来决定, 可以形成非投射性的依存现象。

(2) 基于统计的处理思路有以下几个。

Samuelsson (2000) 提出了一个包括了依存类型且允许非投射依存的分析模型。这个模型包括两个随机的过程, 一个过程自顶而下的分析形成树 x , 一个过程自底而上形成线性序列 y 。然后计算 y 条件下 x 出现的概率。 $P(x, y) = P(y) P(x / y)$

McDonald 等 (2005) 提出了用寻找最大跨度树 (Maximum Spanning Trees) 的算法来分析非投射性依存现象。即把依存句法分析形式化为在一个有向图中寻找最大跨度树的过程。

Nivre (2006) 定义了非投射性结构的度 (degree) 的概念。度是指在边 $e(i, j)$ 跨度之内的结点中不被 e 中的 i 或 j 支配的结点的数量。一个非投射性句子的度取其最大值。Nivre 考察了两个依存树库 Prague Dependency Treebank (PDT) 和 Danish Dependency Treebank (DDT), 通过统计发现, 投射性依存的限制过于严格。如果在句法分析中使用投射性的特征, 在 PDT 中将排除大约 25% 的非投射性的结构, DDT 中将排除大约 15% 的非投射性结构。而如

果把非投射结构的度限定在 1，则可分析 98% 的句子，限定在 2，则可分析 99.5 以上的句子。这意味着，如果在算法中整合如非投射度的限制，将会提高系统的准确率和效率。

可以看出，投射性依存的处理在计算语言学方面已经有较长的一段时间，在技术上已经比较成熟。但是，对非投射性依存现象的关注则是近几年才开始的，而且只是针对部分典型的非投射性现象。不仅在覆盖面上不很全面，而且在技术上也不很成熟。不论是在预处理时把非投射性现象处理为投射性的或者是在分析完后再修正为非投射性结构，都是把非投射性现象的处理放在一个相对边缘的位置。

而且，前述几种处理的非投射性依存现象，基本上都是基于印欧语言的，这些语言具有明确的屈折变化，而且非投射性依存的举例主要集中在的疑问词提前的格式，对属于非投射依存中一个子类的交叉依存类现象的探讨很少。对于汉语来说，可能有更多的问题需要考虑。比如，现有对非投射性依存现象的分类在汉语中是否全部存在？汉语中名词前置的非投射性依存现象有什么不同之处？汉语中交叉依存类非投射性现象都有哪些？现有的处理思路是否可以用来借鉴处理汉语？这些都是值得思考的问题。

目前基于规则的研究对交叉依存的涉及较少，缺乏系统性理论来指导句法分析的过程。在词序的研究上只限于描写了部分特殊语序，在解释上做的研究不够。统计研究的情况也多为研究怎样通过统计的方法从语料库中习得非投射类依存现象的数据以及在实际的语料中的处理情况。对于汉语来说，连可以作为统计学习基础的语料库建设都还是不足的。因此，要实现更高目标的依存处理，现有的研究成果还是远远不够的。

2.3 交叉依存现象相关的语言学研究

交叉依存现象涉及到以往汉语研究中很多的领域，比较典型的如主谓谓语句、论元分裂、动补结构等。由于研究视角的不同，这些现象都分别散布于不同的主题之下，以往并不认为几类现象之间有任何的关联性。

2.3.1 主谓谓语句研究

主谓谓语句是指主谓短语充当谓语的句子。其最初的解释源于陈承泽（1921）“得以句为说明语”⁸的论述。主谓谓语句作为汉语一种特有的句式，基本上已经被学者普遍认同。朱德熙（1985）论述为：“主谓结构做谓语的格式是汉语中最常见最重要的句式之一。应该看成是正好跟‘主-动-宾’相匹配的基本句式”。但是也存在部分的争论。这些争论大致可以分为两种，第一种争论是方法论上面的，是关于使用线性分析还是层次分析方法的争论；第二种争论则是由于对主谓谓语句的判定标准不同引起的，是使用形式标准、语义标准还是语用的标准。

第一种争论，表现在使用的分析方法上。以范继淹为代表的一些学者主张句子的线性分析，认为线性分析的方法在教学上会更为方便。而以朱德熙、丁声树等为代表的结构主义学者则坚持使用层次分析的方法。对于主谓谓语句，范继淹（1984）称之为“多项NP句”，结构主义学者则命名为主谓谓语句。

第二种争论，表现在对主谓谓语句范围的划分以及形成的探讨上。具体来说，主谓谓语句划分范围的分歧，主要集中在对第一位的名词、第二位名词与谓语动词的关系的界定上。根据其关系的不同，有人认为是主谓谓语句，有人则认为不是。其争论的本质，在

⁸ 参见赵元任《汉语口语语法》第57页脚注，商务印书馆1979年版。

我们看来是关于主语概念应该用形式标准、语义标准还是语用标准的争论。

划界的争论主要表现在两个名词是否是主语以及那个是主语的界定上，因此要梳理清楚这些问题，我们有必要重新审视一下对于主语的定義。

那么到底什么是主语呢？主语的概念最早可以追溯到亚里斯多德。他认为主语与谓语是句子的两个基本构成成分。但同时又指出，在英语中，主语与动词的数、人称上要保持一致，在其他屈折语言中，还可能一致于动词的性、格等。可见，这样的定义其实是有点模糊的，这也引起了大家对这个概念的不同使用。在屈折语中，可以同时满足这两个条件。但是在汉语中没有形态的变化，导致学者们对主语这个概念有不同的理解和使用情况。

胡裕树（1995）曾指出，语法学家对主语的概念主要有三种理解，第一种认为主语是针对谓语动词而言的。第二种认为主语是谓语陈述的对象；第三种认为主语是话题。朱德熙（1982）即主张主语与话题是合一的，句子的主语也是语用上表达的话题。汤廷池（1978）则主张明确区分主语和话题两个概念。不能用语用层面的话题来替代主语的概念。Li & Thompson（1990）认为，主语和话题的区别主要在于，主语和谓语之间必须具有选择关系，主语必须是谓语核心的一个论元成分，而话题成分则不必跟谓语核心有选择关系，即不必然是一个论元。而且，话题必须是有定的或者是已知的成分。曹逢甫（1979）则认为主语是动词左边首个有生名词的短语，他的看法更像是结构与语义的综合产物。不同语法学家选用了不同的主语定义，导致了对主谓谓语句划界的不同。

值得指出的是，主谓谓语句是汉语独有的提法，在其他语言的语法著作中没有这种说法。例如，“那个人我认识”的句子，用英语的语法系统来解释，“那个人”是应该分析为宾语的。如“Him I've saw”中的 Him 是句法上的宾语。这样来看，英语中的主语是使用

的第一种定义。即其语法形式其实是语义的抽象体现。汉语由于没有类似于英语这样的屈折形式，所以多采用第二种或者第三种的主谓谓语句定义。关于主谓谓语句的争议，很多是术语指称范围的不同所引起的。由于本文所探讨的非投射性依存类现象只会涉及到部分的主谓谓语句，我们的综述也将只着重于与本文相关部分的争论。这部分主谓谓语句主要有三类：

(9) 那个人我认识。

(10) 他饭没有吃。

(11) 我肚子笑疼了。

(9)与(10)两类是争论较多的部分，(11)类则是比较原型的主谓谓语句。最早的主谓谓语句定义的范围只是(11)类，即大小主语之间有领属关系的类。如王力（1985）《中国现代语法》中关于“叙述句”或“描写句”，都只限于形容词性谓词作谓语，且大小主语之间具有领属的关系。邢欣（1995）也主张把主谓谓语句只限于这一类。后来受结构主义的影响，主谓谓语句的范围扩大到了包括(9)、(10)两类。丁声树（1961）、赵元任（1968）以及朱德熙（1982）为代表的结构主义论者根据 NP_1 和 NP_2 的语序，将他们分别称为大小主语，既不考虑它们的语义属性，也不考虑是否能够挪回到动词的后面，认为宾语提前的说法没有根据，也不考虑是否有代词复指。通通称之为主谓谓语句。这样的处理，从结构主义理论的角度来看是比较整齐划一的，为纯粹的形式标准。同时，朱德熙指出，汉语中有用主语表示已知的确定的事物的倾向。因此，主谓谓语句中的主语多为定指的。但是，朱德熙（1982）在《重印题记》里否认了把句首的处所词和时间词一律看成是主语的观点，认为有的是状语，例如：屋里坐，明天见。

胡裕树（1995）和吕叔湘（1979；1986）虽也主张主谓谓语句，

但把前文（10）类的句子看做是宾语前置而来的。同时，吕叔湘认为，称其为“宾语提前”只是一种方便的说法。胡裕树在《现代汉语（重订本）》里也保留了（10）类结构“宾语提前”的说法。因为（10）类句子中动词的受事此时与动词处于同一个平面，而大主语则跟动词处于不同的平面。大致来说，胡与吕开始还是承认主谓谓语句的。但是，随着三个平面理论的提出，胡裕树（1982）中又论述到，这种句式应该是语用层面分析的结果，因而在句法上没有其应有的地位。胡的这种观点跟 Givón (1984) 的功能主义观点是大致相同的。这两种观点对主语概念的使用都属于主语的第三种指称情况。黄伯荣、廖序东（1991）的《现代汉语（增订版）》中主谓谓语句的范围与胡本的内容基本类似，但不包括句首时间词作状语的句子。

邢欣（1995）根据胡裕树三个平面理论的观点，主张可以把主谓谓语句的范围限定在具有领属关系的类。把类似于（9）类的句子中的大主语在句法上都处理为修饰语、提示语（胡裕树、范晓，1985），（10）类的句子为焦点前移句，看做是宾语提前。而具有领属关系的类，如果激进一些，也可以看做是提示语句。当大主语与小主语之间有副词时候才是主谓谓语句。Charles Li 和 Sandra Thompson 在《Mandarin Chinese》一书中则主张 A 类句子的句首成分看做话题，同时承认它是宾语。⁹ 类似的，沈家煊（1999）、Givón (1984) 等也称之为 OSV 句式。不少学者（如：陈平，2004）也从话题—陈述结构的角度进行过专门的研究。徐枢（1988）把（9）类称为“名_受+名_施+动”句式，指出这一格式的使用相当复杂，涉及到语法语义和语用多方面的因素，有些句子对这一格式的使用具有强制性，有些则可用可不用。名_受的有定性是这一句式成立的必要条件。且动词和名_受之间的语义搭配是不可忽视的。

⁹ 转引自徐枢（1988）从语法、语义和语用角度谈“名_受+名_施+动”句式，《语法研究和探索（四）》。北京大学出版社，第 50 页。

孟维智（1984）认为，（9）类的“那个人我认识”不应该是主谓谓语句，而是宾语前置句，为“宾、主、谓”；“他饭没有吃”为“主、宾、谓”。孟维智的分析，其实是使用的传统语法的分析方法，完全从意义出发的。其使用的“主、谓、宾”术语应该指的是具体的主词、谓词与宾语词，而非抽象意义上的主语谓语。以意义作为判断标准的思想实际源于黎锦熙《新著国语文法》的传统。

关于（9）（10）两类句式，方梅（1995）曾论述到，相对于“施事+受事+VP”结构，“受事+施事+VP”为常态句式。另外，还有一些只针对特定现象的分类研究也与两类句式相关，如李敏（1996）对 S₁ 隐含介词情况的分析，李芳杰（1998）对大主语或小主语是“的”字结构的情况的分析等。

（9）（10）两种分类，大约分别代表两种非投射性依存的现象。只是在前面提到的这两种结构的后面再增加部分成分，增加的成分为谓词前的第一主语或第二主语的修饰语。（11）类则是主谓谓语句与动补结合在一起形成的非投射性依存现象。（9）类代表有边跨过最高结点的情况，（10）（11）两类则是交叉依存的现象。在以往的主谓谓语句研究中，重在探讨第一第二主语与谓语动词间的关系，以及两个主语之间的关系，而对其他的部分关注不多，大多是把谓词及其后部分通通称之为谓语部分。例如：“我球赛倒看了不少（范继淹 1984）”。对谓词后部分于前两个主语的关系的探讨上是很少的。

如果考虑到动词后部分的词语，则主谓谓语句的研究就和论元分裂的研究结合在一起了。论元分裂是汉语中以转换生成语言学为理论框架的另外一个研究的问题。沈阳（1995）论述了数量词在名词短语移位中的作用和特点，即数量 NP 跨越动词的前移只能是名词前移，成为受限的非数量形式；如果 NP 的后移，则只能是数量词后移。在这些分析的基础上，认为在包含了数量词的 NP 中，数量词才是语义和句法的核心。这样，主谓谓语句中的大主语或小主

语就解释为是论元分裂而形成的。沈阳(1996)在论述移位理论时,把原有的看法推广到部分非投射性依存现象中的(9)类。沈阳把(9)类句式解释为论元 NP 的分裂移位。且认为动词后论元 NP 分裂移位时,只能是语义所指范围比较大的成分分裂前移,而语义所指范围小的成分留在原位。所谓语义范围大小,并不决定于某个成分是原名词词组的中心词还是修饰语。同时,沈认为论元 NP 向前分裂移位都不是真性论元 NP 移位。留在动词后原位置的成分才是受动词直接支配的成分,或者说动词后才是成分支配位置,分裂前移的成分则不再受动词支配。蔡莉(2008)指出分裂移位的成立的条件为名词短语内部的语义关系为领属性、限定性与描述性三种。领属性与限定性语义关系的名词短语分裂其中心语须为实体名词。描述性关系的名词短语的移位则需要满足修饰语对中心语有区别性特征的要求。蔡莉的论述是针对(9)类现象来说的,并没有论述到(10)类的交叉依存现象。但其结论对交叉依存的句法语义探讨有一定参照意义。

陆俭明(1988)曾讨论了数量结构的成句功能,沈阳的研究则把其拓展到论元分裂的问题中来,但是这样的解释也有掩盖了其他层面制约因素的嫌疑。例如下面两个句子可以通过增加数量短语来使得句子成立:

? 张三掉了门牙。--张三掉了两颗门牙。---张三门牙掉了两颗。
? 王奶奶瞎了眼。--王奶奶瞎了一只眼。--王奶奶眼睛瞎了一只。

我们也可以通过替换动词为双音节动词而使得句子成立,因此其制约成句的作用是多方面的。

(12) ? 张三掉了门牙。---- 张三碰掉了门牙。

(13) ? 王奶奶瞎了眼。---- 王奶奶哭瞎了眼。

对这类现象，还有两种解释的思路。一种认为是动词的论元类型可以在不同的结构中发生变化，即“动词变元说”，另一种认为是省略了名词，动词后应出现的论元 NP 中的名词因与前第一或第二名词相同而省略。可以称之为“名词省略说”（转引自沈阳 2001）。这两种解释都具有片面性，动词变元的解释会导致动词论元的不确定性，名词省略说不能解释很多不能够补充出修饰语的现象。

另外一种解释的思路即认为首个名词是话题，把话题上升到与主语相并行的句法概念（Li&Thompson 1990）。话题化的解释修改了以往的句法定义，因而引起了很多的争议。如朱德熙（1985）即认为是把结构和语义两个平面混淆了，语义平面上的施受事与结构平面上的主语杂糅。另外，就本文的研究来说，话题化的只能解释部分非投射性依存现象，对其中的交叉依存类的现象则不能够解释。

综上所述，不同学者对主谓谓语句范围的划分，是分别基于不同的理论出发点。因为所采用的语法理论和分析方法的不同，并不具有横向比较优劣的可能。所以，只要其坚持了理论的系统性，都有其一定的合理性。但是，也有些问题是比较明显的，如单纯根据词序划分主语的做法使得主谓谓语句的范围无限扩大，会出现严重的句型交错，而且主谓结构如果可以重复嵌套则失去了其分类的意义。而参照语义的分类，也往往由于语义类别边界的模糊性引起较多的争议。争论的关键还是集中在关于是不是主语上，主语的本源是指名词性成分与谓词之间的一致关系，并不必然限定于施事或受事，因而并不是绝对的语义标准。同时也并不是绝对的语用标准。而汉语中并不存在这样的一致形态变化，无论是参照语义还是语用，都不应惯之于主语的称呼，不能整齐地对应于印欧语的主语概念。有人主张用停顿来判断主语，以此来证明主谓谓语句，同样也会出现逻辑上的悖论，因为可以停顿的成分不仅限于名词性成分。而且，作为韵律表现形式的停顿，很可能属于不同于句法的韵律层面（冯胜利 1996），韵律对句法有影响，并不能作为句法层面判断的标准。

把话题与主语并列作为句法成分的做法同样也存在着定义不清的问题。可能我们在最初引入主语宾语这些概念的时候并没有很好地参照汉语的实际作出明确统一的定义，才导致了立足点不同的争论，从实际效果来看是不尽如人意的。而且还会引起许多后续的问题。

在我们看来，主谓谓语句的研究还存在着如下的几个问题：

首先，边界划分问题会引起逻辑上的悖论。主谓谓语句中两个名词主语之间的关系一般认为只是语义上的所属或相关，不是句法上的偏正结构。这样，主谓谓语句与偏正关系的名词短语做主语的句子在划界上会存在问题。人们在研究偏正关系时，一般认为不加“的”的修饰关系在组合上要紧于带“的”的偏正关系¹⁰。而在主谓谓语句中又认为两个名词之间没有句法关系，这样的处理思路不能使人信服，是值得商榷的。寿永明、朱绍秦（2002）虽一再强调 S_1 与 S_2 之间语义上的隶属关系不同于结构上的偏正关系，但其观察的角度是语用的角度，并没有从句法上论证两者的划界悖论。我们知道，名词也是有配价的。一些主谓谓语句中的第二个主语为一价名词。一些部件类的一价名词划归为主谓谓语句，如“奶奶眼睛哭瞎了”，而同为一价名词的亲属关系类则更可能是一般主谓句。如：小明奶奶去世了。从依存的角度来看，无论是部件名词或是亲属名词，都属于一价名词，要求有名词与之同现，因此两个名词之间是有句法上的依存关系的。这种同现激活的关系比有形式标记的“的”的关系可能要更近些。仅仅从结构的角度来判断主谓谓语句是不能够成功的，必须参照语义关系。因而在汉语中对语义关系的参照更为重要。

第二，讨论较多集中于两个名词之间的关系以及各自与动词的关系，对谓词后部分的讨论较少。对谓语部分的笼统处理使得主谓

¹⁰ “一般来说，不用‘的’，整个词组结合得非常紧密；用上了‘的’，前边词语的修饰性就更加明显了。”胡裕树《现代汉语（重订本）》上海教育出版社，1995年版，P306.

谓语句中修饰语与名词分离的现象没有受到应有的关注。

依存语法给我们提供了一个不同的角度来看待这些问题。作为句法关系最初代表的屈折变化，在本质上来看是语义范畴，是系统性的语义关系。因此，我们可以把汉语中系统性的语义范畴作为句法标准来使用，作为区分句法关系的基础。对主谓谓语句范围的争论，我们可以从依存的角度来考察词语之间关系的异同，避免术语的争论。如果两个名词与动词之间的关系是不同的，那么就不能用同一个关系标记。我们也可以从依存理论的角度来观照到在语序上分离的成分之间的关系。与名量词共现的只能是名词，与动量词共现的只能是动词。因此，动词后的名量式短语只跟名词具有依存关系，而不是动词直接支配的成分。相应地，类似于“他饭没有吃一口”的句子中，与“口”有依存关系的只能是“饭”，而不应该是“吃”。从语义的层面来看，受“吃”支配的也是两个名词性成分，一个是施事，一个是受事。沈阳用语义范围的大小来谈可能有点绝对化了。依存理论在分析过程中对于词序的限定上不那么严格，所以可以对类似的句子做比较一致的解释。虽然其受事在动词的前面，但依然可能是受动词直接支配的成分。

第三，分裂移位的说法因没有形式上的证明而难于自圆其说。一些认为是移位的成分可以移回到原位，但也有很多例句中认为是移位的成分并不能够移回到原位去。这样就使得其理论的系统性打了折扣。

2.3.2 相关动补结构研究

动补结构又叫述补结构，是由一个动词和另一个谓词性成分组合而成的。动补结构是汉语中独具特色且极为重要的一种句法结构，是汉语区别于印欧语的重要特点之一。对其特点的认识关系到对汉语语法特殊性和整体特征的认识（薛红 1987）。其与“把”字结构、

被动结构、受事主语句、动词拷贝结构、话题结构等汉语其他特殊结构有着非常紧密的关系，可谓牵一发而动全身。所以总是不同语法理论研究的重点内容与检验理论是否适当的重要指标。对其历史来源、句法、语义、语用等方面的研究历久不衰。

关于动补结构的特点，李临定（1980）曾指出，“从逻辑意义上来分析，它包含着两个表述（两个主谓），但从结构形式来看，它却只是一个单句”。朱德熙（1982）把动补结构分为粘合式与组合式两种。组合式述补结构可分为表示可能性和表示状态的两类，两个谓词中间有助词“得”出现。粘合式述补结构又分为结果补语组成的述补结构和由趋向动词充任补语的两种。同时朱德熙指出，带结果补语的述补结构在语法功能上相当于一个动词。吕叔湘（1980）也将动补结构称之为短语式动词，其动趋式与动结式两种命名基本相同于朱德熙所说的粘合式动补结构。以后许多学者都沿用了“动结式”的说法，但是并没有强调作为一个词来看待，更多的从短语的角度来探讨的。

在早期的研究中，有学者把动补结构作为一个词来看待。赵元任（1979）的《汉语口语语法》中即称之为“动补复合词”。汤廷池（1977）的《国语变形语法研究》中把黏合式动补结构称之为“结果式动词”，是“复合动词”的一种。有一些海外学者也把动补结构当成一个复合词来看待。日本学者太田辰夫（1958、1987）将补语是不及物动词的动结式称为使成复合动词，补语是形容词的动结式称为结果复合动词。志村良治（1984、1995）则全部称为使成复合动词。¹¹

看成是词或者是短语，是基于不同的理论出发点。由于动补结构的能产性极高，很多动词和形容词都可以结合到动补结构之中，如朱德熙所说的“述语和补语的组合是极其自由的（朱德熙 1982）”，

¹¹ 转引自施春宏（2008）《汉语动结式的句法语义研究》，北京：北京语言大学出版社，第6页。

因此看做是一个结构可能会更加合理一些。毕竟作为一个词应该是相对固定的，而动补结构中的组成部分却往往是可以灵活替换的。

动趋式结构中的趋向动词，有的表示动作的趋向，有的用于引申的意义（吕叔湘，1980），其语义指向多与述语动词同指或指向述语动词，因此相对比较单纯¹²。从依存的角度来看，两者之间的关系是第二谓词对前动词表示的动作本身的修饰，因而两者之间的依存关系是比较明显的。而动结式述补结构的情况则比较复杂，成为学者们关注的热点。交叉依存类现象涉及到的动补结构也主要为动结式的动补结构。所以，我们的综述将主要集中于动结式的动补结构。“动结式”的名称是一个较为宽泛的指称，两个谓词之间并不都是致使的关系，可以是比较松散的集合。

了解动结式的研究，首先需要把动结式述补结构与述补式复合词区别开。动结式述补结构与一般述补式复合词在句法功能上具有几个明显的不同之处：

第一，述补结构可以用“得、不”扩展，述补式复合词不能扩展。（朱德熙，1982）第二，述补式复合词支配名词性成分的数量是确定的，而动补式则随组合类型的差异而有不同；第三，述补式复合词的宾语与 V_1 和 V_2 都有语义选择关系，而动结式中述语动词的宾语与补语动词未必一定有语义上的联系。第四，有的动结式可使用动词拷贝式，而述补式复合词都不能使用动词拷贝式。（施春宏，2008）

其次，要注意动结式的类别划分的不同。

动结式的分类可以有不同的标准。马真、陆俭明（1997）将形容词充当结果补语的述补结构分为四种。A. 预期结果的实现；B. 非理想结果的实现；C. 自然结果的出现；D. 预期结果的偏离。刘

¹² 趋向动词除了有“动词说”外，还曾有不同的称谓。陆志韦《北京话单音词词汇》中叫做“后附的副词”，在《汉语的构词法》中也曾称之为“构词成分”。俞敏、林汉达等称之为“词尾”，房玉清、徐静茜称之为“动态助词”。这些称谓的不同，反映出其不同于一般动词的特点。

子瑜（2004）把动趋式之外的动结式分为三种，结成述补结构、状态述补结构和结度述补结构。施春宏（2008）则把动结式分为致役式、评述类和自变类三种。

第三，要注意动结式述补结构内容的研究，大致来说有以下几种研究的思路。

第一种研究是从历时语言学角度来探讨其形成的历史过程以及其与其他句式之间的源流关系。

对于动补结构的产生时间，学者们争议很多。周迟明（1958）认为，动补复合动词产生于殷代。王力（1958）认为动补结构最早出现于汉代。梅祖麟（1991）以“V死O”格式作为动结式产生的标准，认为动结式产生于六朝。持类似观点的还有蒋绍愚（1999）、吴福祥（1999）等。

更多的学者是从不同的角度出发来构拟动补结构的形成与发展的历史过程。潘允中（1980）提出的VV并列结构，志村良治（1984、1995）的双音节研究，太田辰夫（1958）和梅祖麟（1991）提出的自动变他动，蒋绍愚（1999）曹广顺（1999）提出的隔开式与合用式的同时产生，黄正德（1995）的核心左移，曹广顺、吴福祥（1999）的语义分析等观点，都对我们了解动补结构的发展轨迹具有启发意义。在这些研究的基础上，一些学者（冯胜利，2002；刘子瑜，2004）从更全面的角度来构拟其形成过程。石毓智（2003）在动补结构研究的基础上构拟了整个汉语语法系统建立的过程。

从“V+N+V”到“VVN”的历史演变过程，是汉语系统由简单到复杂发展过程的典型体现。汉语由此变得复杂与精细，“最能体现汉语结构简练而寓义丰富的特点”¹³。从事件的角度来看，则是组合方式从简单的链接方式向整合方式的质变。

第二种是从共时的角度来研究的。

共时角度的研究经历了从单纯的分类与描写到解释与生成的

¹³ 李临定 1986《现代汉语句型》，第198页。

发展趋势。正如袁毓林（2008）指出的，“以往对动结式的研究大多侧重于局部问题，难于构建动结式的句法语义生成系统。”¹⁴一些学者以动补结构为入手点，考察其对整个句子的生成制约，即动结式论元结构的整合过程。尝试解释动结式结构中动词论元与补语论元的出现规律，以及操作的具体过程。

郭锐（1995）、王红旗（1995）提出了底层论元提升的概念及规则，认为动结式中述语动词的主体直接提升，客体不提升，补语动词的论元都能提升，分别给出了动结式配价的计算方法。袁毓林（2001）指出了其自下而上分析的不足，建立了一套自顶向下的述结式的论元准入规则，称之为配位方式。具体指同一句中动词的配项之间的同现选择限制以及主宾语等句法成分对施事受事等予以成分的选择关系。认为这种同现选择限制是由说话人的透视域决定的。郭锐（2002）对郭锐（1995）有所调整，认为述语动词的客体论元也可以有条件地提升为主体论元，并明确了底层论元提升的方向和优先序列。¹⁵施春宏（2008）提出了界限原则来解释动结式语义特点对句法结构整合过程的影响。界限原则是时间顺序原则和突显原则共同作用的结果。

刘街生（2006）通过与动词并列式、连动式和兼语式的比较，认为动结式组构成分之间既有一定的并立性，又有一定的说明性的特点。司马翎、沈阳（2008）采用“小句理论（SCT）”来解释动结式的形成过程与操作程序，尝试用同一种理论把不同的动结式的形成原因作出解释，具有创新性。而且，这样的处理与汉语“把”字句的分析可以很容易地联系起来。

第三种，可以把针对特殊类别动结式的研究划归为这一类。这些研究虽然着眼于小处，却往往有很强的理论意义。如，任鹰（2001）对主宾可换位的动结式的研究。考察了换位机制形成的理据，认为

¹⁴ 参见袁毓林为施春宏《汉语动结式的句法语义研究》一书所写的序言。

¹⁵ 转引自施春宏（2008）第18页。

“述语成分的两个配项都有充任宾语的条件，也都有充当谈话起或语句焦点的可能，是主宾可换位结构形成的内在理据所在。”吴为善、吴怀成（2008）研究了双音述宾结果补语类动结式的现象，认为其补语成分所表示的结果状态多属于非预期的。宋文辉（2006）通过主副事件来解释宾语前置的动结式结构中某些宾语必须前置的原因。认为其机制在于补语所代表的结果可预测程度低，凸显程度高，成为句子的焦点。随着认知上独立性的增强，原来依附性强的不占据独立句法位置的补语要求占据独立的句法位置，因此将动结式后原来属于宾语的位置占据，宾语必须前置。

我们从依存的角度来看，印欧语中两个谓词之间的连接总有明显的形式标记，或者是一个使用不定式，或者用动名词的形式区别于主动词，两个谓词之间的主从关系一目了然。而汉语中无此类形式标记，因此动补结构中两个谓词的关系仅在形式上不能确定哪个词才是中心，只能从语义出发。动趋式的组合中两个谓词之间的依存关系比较明晰，因而不存在争议。只有动结式的争议较多。这是因为两个谓词之间在语义上并无必然的联系，无强制共现的关系，不同于一个谓词充当另一个谓词的配价成分的情况。因此，对其处理会有两种方式，一种是像动趋式一样看做两者之间有依存关系，一种是看做两个无句法关系的谓词。从事件的角度来说，看做两个事件易于对各自支配成分的把握。而且，目前对于配位的研究也是就各个谓词单独来说的，没有强调两个谓词之间的关系。所以，不如看做是两个无句法关系的词，对于机器分析来说会更加便捷。

由于两个谓词都可以带有自己的支配成分，连接不同词语之间的依存关系边就很容易产生交叉的情况。交叉依存现象涉及的动补结构中，主要涉及到的是致役类的动结式。因为其两个谓词都分别有其自身的论元结构。如果动补结构的受事可以移动到主语之后。如，大娘哭瞎了双眼，可以表述为“大娘双眼哭瞎了”，这样的结构即可形成交叉依存的结构。施春宏（2008）指出可移动的受事与

主语两者之间存在着领属关系，且补语动词的论元都是一价名词。袁毓林（1994）曾把一价名词分为亲属名词、属性名词和部件名词三种。能构成交叉依存的一价名词其实不能为亲属名词，因为亲属名词放在名词之后可能会发生歧义。王跃龙（2008）指出，这种N+N结构中的亲属名词为称呼语。所以，能构成交叉依存关系的只能为属性名词、部件名词及其配价名词的共现。

综合以上的研究，可以初步得出以下的结论，即大家都认可动补结构内部比较复杂，不宜作为一个词来看待，看作是两个谓词的组合的处理可能更为妥当。但是对具体问题的看待上分歧还是很多的，不同的学者分别从不同的角度对动补结构的不同现象进行了解释。特别是对其论元成分隐现及位置的配位研究上的争议更多。配位概念的提出是一个全新的概念，开始把配价的思想从只关注词语的价拓展到了解决位置的问题，把认知作为解释的工具，是个不错的思路，但是也有把语义、语音和语序等不同层面的制约因素混合在一起的嫌疑。至于从依存语法整体角度来作的分析则基本没有看到。

从不同的理论角度来看，现有的研究都有其一定的理论价值。但是，如果从投射非投射的形式区分的角度来看，这些研究仍然显得是比较零散的。特别是从信息处理的应用实际来看，还需要一个系统的理论来把现有的研究整合在一起，研究这些现象之间的共有的特性和制约原则。

首先，现有的研究分类也是比较绝对的，对于一些渐进的现象划界存在问题。例如，“①他喝酒很厉害。②他喝酒喝醉了。③他喝醉了酒。”三句中都有“喝酒”一词，在①主谓谓语句中处理为主语，在②重动句和③动补结构中则处理为谓词性成分。我们有疑问的是，他们之间有无根本性差别？如没有，为什么分析为不同的成分？如有，差别又是什么？这些都是我们尝试在本文中作出统一解释的。

其次，对形成动补结构的两个谓词的配位研究中，只是指出了可以提升的条件和位置，但不同的论元之间前后位置的不同的制约原因是什么，并没有解释。但对与投射非投射性的区分来说，却是至关重要的一个现象，位置的不同可以导致投射非投射的截然两分，对其原因的解释是非常必要的。

2.4 语言网论研究

语言网络的思路是受认知心理学研究的启发。认知心理学认为人脑对知识的储存是以有连接的结点构成的网络形式存在的。在以网络形式存在的知识结点之间，一个结点的激活会扩散到其直接相连接的结点的激活。这即是扩散性激活（spreading activation）模型。这是与以往的认知理论所不同的。较早的认知理论认为人是通过模块来认知的，称之为“模块理论”。模块理论认为知识可以分为不同的模块，然后通过有限的连接把不同的模块连接起来。短语结构语法的认知基础即是建立在模块理论之上的。随着网络认知理论的提出，一些学者也通过扩展激活理论的模型来探讨网络结构的语言观。认知语言学的基本假设即是认为语言的概念使用的是像其他概念一样的组织方式。Goldberg（1995）称之为“语言的知识也是知识”¹⁶。一般性知识通常被认为是以网状的形式存在的，那么也可以推论语言的知识也是以网状形式存在的。

大致上来说，语言网络的观点兴盛分别有几种不同的驱动来源。一种是来源于语言研究内部的；一种是计算语言学领域的，图形计算技术的发展为其提供了思路；另一种则是受其他相关的科学体系的影响。

由语言研究内部发起的语言网研究又可以分为两个方向。一个是研究各个词语之间聚合关系的词汇语义方向，另一个则是研究词

¹⁶ “Knowledge of language is knowledge”

语之间组合关系的句法方向。刘海涛（2010）称之为静态网络与动态网络的差别。

把词语之间的关系看作是一个网络结构在词汇语义学范围内具有很久的传统。其中最具代表性的成果是 WordNet（Miller et al., 1990; Miller, 1995）的建设。WordNet 是一种网状的语言结构，表示词语（概念）之间的聚合关系。他们根据心理学理论来定义单词的意义，形式化为“单词—意义”、“意义—意义”之间的联系。在 WordNet 中有两种节点：词与概念。词与词之间通过不同的关系（如同义、反义）来连结。同义的词语形成一个 synset，每个概念都有一个 synset。概念之间通过关系（上下位关系、部分整体关系）来联结。这些关系的表示共同形成了一个复杂的网络结构。Wordnet 作为一种知识的表示体系，现在已经成为研究词语间关系、计算词语之间语义距离的重要基础资源。类似的观点下构建的资源有 HowNet、中文的 CCD、知网等。这些网络多是基于词典等语言资源构建的，同义、反义、上下位、部分整体等关系都不涉及词语在句中的使用，属于静态的网络建设（刘海涛 2010）。

从词语组合的层面把语法结构看成是由词语之间依存关系构成的网，其代表是 Hudson 的 Word grammar。Word Grammar（WG）是 Hudson（1984, 1990, 2007）在二十世纪八十年代建立的语言结构理论。他自言是受到 Systemic Functional Grammar（Halliday 1985）、Stratificational Grammar（Iamb 1966, Bennett 1994）与 Dependency Grammar（Tesnière 1959）的影响。经历了初期 Daughter-dependency Grammar 后为 Word Grammar。WG 认为，语言是一个由词语的概念、意义及其词性综合构成的知识网络。这个词语的网络是单层的，而不是以往认为的由深层转换而来的。语言网络是由词语之间的依存关系构成的，而不是短语结构。可以通过 label 表示词语之间的语法及功能关系。WG 认为，人们在使用概念的时候使用的是原型概念而不是分类概念，所有的概念特征在地

位上都是相等的，有些是因为语用的原因被废弃。在这个网中，不同领域的知识之间的分界并不是清晰的，特别是对词的内部与外部意义的分界，所以语法与社会语言学联系比较紧密。语言网还具有 Small-Word 和 Scale-free 的特征。组合关系的网络是词语在动态使用中形成的网络，属于动态的网络建设。

聚合与组合的划分，是从其源头来说的。随着研究的深入，两者之间的界限已不再那么明显。聚合的研究往往要参照组合关系，组合研究也需要兼顾聚合关系，已经成为研究中的常态。以往特别关注聚合研究的词汇语义学也开始关注组合研究的问题即是一个典型的代表。

Hudson (2001) 详细论述了认知网络的特征以及在语言中的体现。语言网络具有扩散性激活 (spreading activation)、概念距离影响 (effects of conceptual distance)、默认继承 (default inheritance)、词法与语法的统一性 (the unity of grammar and lexicon)、非模块性 (non-modularity)、可处理性差异 (degrees of accessibility)、两分不对称关系 (binary relations) 等特征。这对网络性的语言观具有重要的指导意义。扩散性激活是指一个词语结点的激活可以扩散到与其直接相连词语的激活。概念距离是指概念之间相连通所跨过的连接数的多少。默认继承指上位概念的特征在下位概念中可以得到默认的继承。词法与语法的统一是指词语的构成原则与语法之间并没有明显的界限，应该是类似的构成方式。非模块性是指语言网络并没有明显的模块界限，类型的扩散性激活、默认继承等特征是贯穿于整个网络的构成之中，而不是局限于某一个具体的模块。可处理性差异是指词语之间联系的权重会影响到认知的速度。两分不对称关系是指任何概念的定义都是与其他概念有关系的，语言网络就是由词语以及词语之间的关系所构成的。

由于现有的语言理论多是在模块认知理论的影响下形成的，人们长期以来一直认为词法和句法是两个互不相关的领域，即构词的

方法和造句的方法是两个层面的概念，不具有统一性。在汉语的研究中这种观念也有很深的影响。尽管有学者指出汉语的构词与造句上有相同之处，但是并没有给与深入的探讨。以致有如下的观点：“汉语的构词方式与造句方式有相同之处，这对于说明汉语的发展情况有参考价值，但是对于说明用词造句的规律一点没有用处。（胡裕树 张斌 1981）”。网络语言观的提出，指出了词法与语法很可能是统一的，使得我们有机会来重新审视汉语中的相关现象。

随着计算技术的发展，相较于以前的树形计算技术，图（网）型计算的技术也逐渐走向了成熟。而且，图型的数据结构比起树型的数据结构来说在信息共享上有更多的优势。这个时候，很多学者开始考虑用图论的思路来进行语言的分析，以及相关的语言资源建设。自 Word Grammar 开始，关于语言网络的表述在计算语言学界开始蔚然成风。对于汉语研究来说，Ji Donghong（2000, 2002, 2004a, 2004b）最早提出用图（graph）的方式来标注汉语的短语结构。他认为，传统的依存表示在汉语中并不能区分细致的语义关系，而且不能够表示出多重依存关系以及无支配的关系。所以需要用图才能给予妥善的表示，并且给出了依存图表示的形式化表示方式。这是目前对于汉语图论形式化表述的最早表述。但是其参照的语法标准还没有完全脱离开短语结构的形式。其最近的图论观点扩展到句子的层面（Wang & Ji, 2007）。但其研究主要是从语义的角度来看的，包括其最近的特征结构理论都是从语义的角度来探讨语义关系的表示方式，在在句法层面上着墨不多，是需要加强的地方。

客观来说，从组合角度来看待语言的基于图论的语言学理论体系不是很多，而且还不是很成熟。但是由于其强大的应用优势与前景，已经有不少的计算语言学家开始致力于图论的句法分析算法的应用，且主要是基于统计的。这也跟语法理论研究的薄弱有关系。

图论观点指导之下，有一批基于依存语法的语料库开始构建。

词汇共现网(Word co-occurrence networks)(Dorogovtsev & Mendes, 2001, 2003a; Milo et al., 2004)是一个标注了词汇共现关系的语料库。句法依存网(syntactic dependency networks)是在词汇共现网的基础上得到的一个句法依存语料库,每个句子都用依存的形式给予表示。针对汉语的图型标注有武汉大学的特征结构语料库。把句子分解为<实体、特征、特征值>的三维数组形式,其最鲜明的特点是不使用泛语言的标记来表示词语之间的关系,而是用实际语言中的词语来表示词语,这些词语称之为“特征”。

另一个对于语言“网”理论的研究是发源于其他科学领域的,并逐渐影响到语言的网研究。复杂网络是对复杂系统的抽象和描述,现实世界的复杂系统都可以看做是复杂网络的结构。它们广泛存在于社会、经济、生物等众多领域。随着互联网、生物网络、社会网络、生态网络等研究的发展,复杂网络的研究作为一种基本的研究理论已逐渐形成,并且影响到对语言的看法。人们开始把语言也看成是由不同的单词构成的复杂网络。与其他网络观研究语言不同的是,这些研究不是从具体的构成来研究网络的,而是从宏观的角度出发,考察其整体的特征及其应用。

很长时期以来,人们对语言网络的假设只有两种,一种是完全规则的网络,一种是完全随机的网络。自 Watts 和 Strogatz (1998)在《Nature》上发表关于动态系统具有小世界的特征之后,一些学者开始用这种理论来分析和研究语言的网。Amaral (2000)和 Strogatz (2001)发现,语言的网不但共有一些生物网和技术网的特征,而且具有不同于数图理论下的随机图特征。在连接数量上不同于古典的随机图的 poisson 分配。这意味着词汇的网应该倾向于归于互联网的家族之中。Cancho & Solé (2001)指出,语言的网也跟其他复杂网一样具有小世界(Small-world)和无标度(Scale-free)两个显著的统计特征。所谓小世界特征可以表述为,大部分节点只与其邻近的节点相连,同时某些节点也可以与非

邻近节点直接相连。因此，其图形具有高聚合度，但是节点之间的平均路径距离却很小。这样我们就可以解释以下的现象，虽然人脑储存有巨大量的词汇，我们都可以把任一个词语通过平均小于三次的连接跟其他任何词语发生联系。无标度的特征是指网络的度分布服从于幂律分布，幂指数值在 2~3 之间。存在着少量度相对较高的节点，但绝大多数节点的度相对很低，也叫做胖尾。这可以解释语言网络中的新词会优先与出现频率较高的词语相连接的现象。出现频率高的词语为核心词汇，核心词汇是多用途的，可以用来表示几乎所有的意思。其连接数量越多，越容易成为核心词汇。

在这些研究的基础之上，刘知远（2007）论述了汉语词语的共现网络、句法网络也具有这两个特征。刘海涛（2009）指出，汉语语义网络也具有这两个特征。

这类关于网络的整体特征的研究，大多是从数学统计的角度来看的，但是其对语言微观研究的指导意义目前还不是很明确。虽然刘海涛（2010）指出，复杂网络的研究可以反映出语体特点，相关特征可以作为区分语言不同层面的手段，还可以研究某些词类对整个语言系统的影响。但在可操作性上依然有待验证。因此，语言的网路观大多只是一种宏观的看法，并没有渗透到语言研究的内部。而且，从数学的角度来看待一个网中的节点，每个节点之间的地位是等同的，只是在连通的数目上有差距。但是从依存的角度来看，构成一个命题的词语之间的重要程度是存在差别的，平等的对待可能与事实不相符。

从网论的角度来看待语言，是一个全新的视角。使得我们能够从新的角度来审视以往的研究，而且也为我们提供了一个新的研究方向。现有针对汉语网络的研究还不够充分，首先表现在很多研究的成果虽然命名为网络的结构，但实际上有些并不是真正的网，而只是平面图或者树结构的形式。虽然在一定程度上平面图或者树型图也可以看做是网的形式，但在数据之间的信息共享上是存在着不

同的，所以，其得到的数据并不能说一定会符合网络结构的特征。其次，一个更为重要的一个问题是，对网络结构与语言线性结构之间的关系探讨的很少。尽管有学者指出语言线性表面的背后是语言的的网络结构，但具体网络结构是怎么对应于线性结构的，线性结构怎样才能转换为网络结构，却很少有学者提到。但这恰恰是一个问题的两个重要方面，必须能够相互观照才是一个合理的理论体系。否则，网络结构与线性语言之间的关系论述就是脱节的，只能看做是一种宏观的看法。网络结构到线性结构的关系，是从宏观角度观察与分析语言的重要方面。

语言网络研究的深入是与计算机的处理能力相适应的。基于树的处理思路与树形计算技术的发展为基础。随着图形计算技术的发展，图论的观点才会在语言领域有了发展壮大的土壤。在这些观点之下，我们也应该开始反思现有树库构建模式存在的不足之处。现有的句法树库的标注有两个问题，第一是树形结构而不是真正的网络结构；第二，现有的投射性原则之下的标注实际上是词序的依存，而不是真正意义上句法的依存，因而需要进行修改与完善，使之能够更好地反映语言句法关系的现实，使其成为学者们研究与统计学习的有利工具。

2.5 小 结

2.5.1 研究现状的总结

可以看出，交叉依存现象的研究涉及到语言研究的多个方面，包括语法、词汇语义、计算语言学和语料库等。其涉及的具体专题研究更多。因此，不可能把所有与之相关的研究做一个全面的综述。本章只是对与交叉依存现象紧密相关的几个主要部分做了简要的回顾和总结，因此必定会有很多的疏漏之处。这些先贤的研究成果

都在不同的方面促进了研究的深入，也是我们研究交叉依存现象的基础。

从中可以看出，目前依存语法理论和应用的研究都已经引起很大的关注，也取得了很多的成果。但对于交叉依存类非投射性现象的研究是不够的。虽然在计算语言学方面有一些关注，但尚未见到有依存理论下系统性的理论研究成果作为应用的指导。相较于德语、土耳其语、捷克语的研究来说，汉语中的非投射性现象的存在可能更加广泛，因而需要引起我们的重视。汉语的依存研究有几个需要注意的地方。一是理论研究的人比较少。依存的概念只是在计算语言学领域有学者应用较多，在理论语言学范围内则主要关注“配价”的观念。我们知道，配价只是依存理论体系内的一个组成部分，没有扩展到整个句子的层面；二是基础研究材料上的缺乏，现有几个依存树库的建设是按照非投射性依存的要求构建的，没有对非投射性结构的标注。所以，我们需要从更为基础的地方着手研究，比如通过实际语料来调查汉语中有哪几种非投射性结构，各有什么特点，以及怎样构建一个能作为统计学习与测试基础的包括非投射性现象的树库等。第三，根据刘海涛（2010）的研究，汉语中词语之间的距离要大于英语和日语，所以汉语中可能存在更多的非投射性依存现象，这是我们需要注意的。无论对于语言理论研究还是对于中文信息处理来说，非投射性依存现象的正确分析和处理将具有更加重要的意义。

由于观察视角的不同，交叉依存类现象散布在以往不同的研究主题之中。因而其结论都是针对个别类的现象，且服务于一定的理论框架的。本文的研究将需要从依存理论的角度出发对交叉依存类的现象进行系统性的考察，研究表面不同的现象之间所具有的共同制约原则。同时，也希望能够把研究成果运用在指导语言信息处理，具体的如语料库建设和机器分析上，期望能对句子分析效率和精度的提高能有一定的帮助。交叉依存的概念本就是从计算语言学角度

发现的一个研究角度，理论的思考如果能反哺到实际应用的层面，将会是一个相得益彰的好事情，实现理论和应用的双重进步。

2.5.2 本文的理论框架

前文简单回顾了相关问题的历史研究及现状。综合学习和分析已有的研究成果，可以梳理出本文的大致研究思路。

首先，可以看出依存语法的理论和应用优势已经得到很大范围的认同。但是，如果对依存语法应用的句法分析依然限定在投射性依存的标准之下，则会影响到分析精度的继续提高。因此，对非投射性依存现象的研究是出于理论和应用两个方面的需要。对于汉语来说，将需要从基础的研究开始，包括具体考察其种类和描写其句法语义条件，以探讨其形成的原因等；

第二，由于我们在引入依存概念时没有针对汉语的实际情况深入思考，使得依存的概念目前还停留在词类范畴之间的关系层面，没有针对汉语的实际进行细化的操作。屈折语中的形态标记在本质上来看是系统性的语义关系的外在表现（胡明扬 1994）。句法范畴层面上的依存关系也可以转换为是系统性的语义的表现，那么我们就可以把系统性的语义关系看做是句法关系来使用，这样我们就可以找到一条细化依存关系的途径。对交叉依存中涉及到的在语序上分离但存在依存关系的词语之间的描述，可以用这种系统性的语义制约来解释；

第三，对于交叉依存现象涉及到的几种现象，以往的研究采用的是组块的分析和解释方式。其对语义关系的描写上是非常深入的，但是对语序分离的解释上不是很信服。这是由于一直在用线性序列的角度来做解释，没有把线性序列与语言网络的观点结合起来。参照语言网络的研究成果，我们需要寻找一种在语言网络观点下从网到线性序列的语序解释方式，从而能够解释交叉依存中有依存关系

的词语之间的语序分离的现象；

第四，鉴于非投射性依存现象的研究直接服务对象是应用系统，而现有标注了句法关系的汉语依存语料库建设都是基于投射性依存假设的，我们需要为应用做一些基础性的资源建设，即语料库的建设。我们应该给出包括了非投射性依存现象的语料库标注标准，从而为非投射性依存的研究提供经过了标注的语料，同时也可以作为统计句法分析器学习和测试的平台。

3 交叉依存现象分类及句法语义描述

3.1 汉语中交叉依存现象的分类

3.1.1 数据说明与分类原则

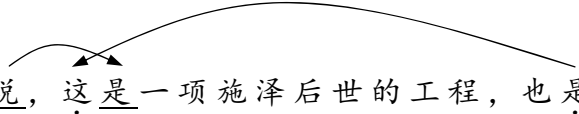
本文研究的对象是汉语中的交叉依存类非投射性现象。选择此类现象作为研究对象的原因已在第一章有了较为详细的论述。对汉语交叉依存类现象的研究，首先需要解决的问题是搞清楚这类现象到底有多少种。迄今为止，好像还没有学者给出一个明确的答案。“特点因比较而显，没有比较就没有特点（朱德熙 1985:2）”，基本问题的答案缺失，使得我们不可能通过与其他语言中交叉依存现象的比较来发现汉语交叉依存现象有什么特点。因此，我们需要从最基本的资料分析入手，来做一些基础性的工作。即通过观察语料库来考察汉语中交叉依存类现象有多少类，然后才能谈得上进一步的分析和研究工作。

鉴于 Tesnière 对依存关系的定义比较笼统，而且存在着是语义还是句法的争论，我们对句法依存关系的判断，参照的基本是 Mel'čuk (1988) 的标准。Mel'čuk 指出了判断两个词语之间具有句法上的依存关系的两个基本原则，分别为线性相关原则和韵律相关原则。只有两个原则同时成立时，依存关系才能够成立。这是针对印欧语来说的，在汉语中，韵律原则并不是必须的条件，经常有句法关系与韵律不相符的情况。例如“一衣 | 带水”的韵律即与句法上“一 | 衣带 | 水”的划分不对应，这是汉语双音节化的要求而形成的。而且，冯胜利 (1996) 曾指出，韵律结构可以是一个独立的平面，对句法现象有影响和限制的功能。当然，线性相关与韵律相

关不一致的现象是比较少的。因此，判断句法依存的唯一必须的条件是线性相关的原则，韵律相关依然是极其重要的参照条件。

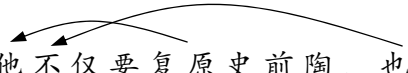
我们认为，交叉依存现象可从不同的层面来观察。一种是在句子的范围内来看；由于句子内部的关系经常涉及到分句之间名词、代词、动词、副词等词类的共享，以及框式连词、分句做事件宾语等，这些都可以引起句法表示形式上的交叉，因而会是一种比较广泛的现象。如下例：

(14) 他说，这是一项施泽后世的工程，也是政府最急切要办的事，我有责任尽一份力。



句中的代词“这”分别与两分句中的动词“是”都有依存关系。同时，两个分句都与动词“说”有被支配关系，表示为两个动词“是”都与动词“说”有一个依存边。这样，“说”与第一个“是”的边就会与“这”与第二个“是”的边发生交叉的现象。另例如：

(15) 他不仅要复原史前陶，也要还原史前陶的精神。



框式连词“不仅...也...”两个词语之间也有共现的要求，即线性相关，因而也具有依存关系。其依存边与代词“他”和动词“复原”之间的依存边也会形成交叉依存的现象。

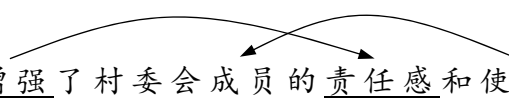
这些现象在句中是广泛存在的。我们调查了美国宾州大学汉语树库(U-penn Chinese Treebank)中 15162 句的语料，共发现 2489 句具有交叉依存的现象，占到总体语料比例的 16.416%。其中单纯由于分句之间成分的共享与框式介词等引起的交叉依存现象为 777 句，占总语料比例的 5.124%。其余 1712 个句子中则都有发生在分句内部的交叉依存现象，占总语料比例的 11.291%。

另一种观察的角度，则是把交叉依存现象限定在分句的范围内，

即在逗号分隔开的范围内。这里之所以不适用单句的标准，是因为单句形式中可以嵌套入复句的句子，形成复杂单句的形式。从而把分句之间成分共享形成的交叉现象与其他现象混淆在一起。分句范围内的交叉依存现象才是更为基本的类型。语料库中 15162 个句子可分为 42415 个分句的形式。共有 1826 句具有交叉依存的现象，占总比例的 4.305%。

其中值得注意的现象是词语并列的现象，如名词、动词、形容词等的并列，往往会形成交叉的现象。如下列：

(16) 民主选举增强了村委会成员的责任感和使命感。

A diagram illustrating cross-dependence in a sentence. The sentence is "(16) 民主选举增强₁了村委会成员₂的₃责任感₄和使命感₅。". Arrows show the dependency of the verb "增强" (enhance) on the noun "成员" (members) and the parallel nouns "责任感" (sense of responsibility) and "使命感" (sense of mission). Specifically, an arrow goes from "增强" to "成员", and another from "增强" to "责任感". A third arrow goes from "成员" to "使命感", crossing the arrow from "增强" to "责任感". This crossing represents the cross-dependence phenomenon where the verb "增强" is shared by both "责任感" and "使命感", but the noun "成员" is shared by both "责任感" and "使命感", creating a cross in the dependency graph.

其中名词“责任感”与名词“使命感”之间是并列的关系，“成员”是两个名词共同支配的成分，而且“责任感”与“使命感”都受动词“增强”的支配。因而会发生交叉的现象。具体来说，是动词“增强”与名词“责任感”的边和名词“成员”与名词“使命感”的边之间形成交叉。

词语并列关系在所有语言中都存在，对并列关系的处理在投射性依存体系中也曾有过很多的争论，有若干种处理的方法。我们认为，如果不局限于投射性依存的强制要求，那么词语分别支配的成分之间的依存关系应该在句法依存图中得以表现，因而也必然会形成交叉依存的关系。而且，并列关系的处理与复句的处理在实质上应该是统一的。分句间主语名词的共享，可以看做是分句内的动词并列共享主语情况的泛化，分句间框式连词的依存关系可以看做是分句内框式结构的拓展。

除并列关系的处理外，还有 835 个分句内的交叉依存现象，为汉语中比较特殊的类型，将是我们重点研究的部分。

在依存语法中，谓词（动词形容词）具有优先的地位，处于句子的中心，支配其他的成分。谓词对其支配的成分具有选择性。而

调查得到的交叉依存现象也都与谓词紧密相关，最明显的差别即是谓词数量的多少。我们可以根据谓词数量的多少把现有的动词和形容词作谓语的交叉依存现象分为两大类。一类是单个谓词下的交叉依存现象，另一类是多个谓词下的交叉依存现象。单个谓词的句子中的交叉依存现象的表现不同于两个或多个谓词连用的时候的交叉依存，我们需要分别来说明。

得到的交叉依存类型，需要用形式化的手段表示出来。如果一组词语 (a, c) 两者之间存在着句法依存关系的同现要求。而同时有另一组词语 (b, d) 之间也具有句法依存关系，两组依存关系共现，同时具有线性排列顺序 $a < b < c < d$ 。那么就出现了交叉依存的现象，可以形式化表示为：

$$\text{linked}(a, c) \wedge \text{linked}(b, d) \wedge a < b < c < d$$

“ \wedge ”表示合取，“ $<$ ”表示前于。这里我们使用了“前于”的概念。冯志伟（2004）对于无句法关系的结点之间的关系是这样论述的：“从属树中的两个结点，只有当它们之间没有支配关系的时候，才能够从左到右的方向上排序，这时，这两个结点之间就存在着前于关系。”冯对于前于关系的论述，其实是把词序也加入到了依存树的表示当中来。对于本文谈到的交叉依存来说，我们可以借用其中一个概念，即“前于”这个术语，但把“前于”扩大到适用于所有的词语上，即不考虑是否存在依存关系的先后序列。我们不仅可以说结点 b 与 c 之间具有“前于关系”，a 与 c 之间，b 与 d 之间也都存在着前于的关系。

德语的拓扑依存表示中专门有一个层面 (Topological dependency) 来表示词序之间的依存关系，可以看做是对这种词序关系高度重视的一个表现。或者，我们可以采取对每个词语加标的做法，在句法树中标示出其原来线性的句法先后关系，这样的表示

才能避免单纯句法树传达的信息的遗漏。一直以来，我们认为短语结构语法的一个重要缺陷就在于对分属不同语块的语言单位之间的句法联系的遗漏，依存的表示则可以弥补这种缺憾。同样的，我们也需要在依存的表示体系中对无句法依存但有线性词序位置关系的信息，这样，我们才能说依存的表示体系比短语结构的体系表示的语言信息更精细而全面。那么这种无句法依存关系的强制线性位置是怎么造成的呢？需要我们给出合理的解释。

在依存的形式化表示中加入词序的强制要求，使之与句法依存关系放在一个相等的地位上，是我们根据汉语的现实对依存语法做的进一步扩展。以往我们认为依存的表示可以允许较自由的语序，是沿袭了印欧语的看法，且是针对特定的语法现象的。在汉语中，语序承担的句法作用要远远大于屈折语。正如任鹰（2001）指出的“语序有时显得灵活多变，并不是因为语序本身具有随意性，而是因为汉语语序负载着较多的语法信息。”而且，无依存关系的词语之间的语序制约，以往并没有很多的论述。对交叉依存的现象描述来说，恰恰是非常关键的制约因素。

我们认为，词序可以分为两个层次，一个是宏观的词序，一个是微观的词序。无依存关系的词语之间的词序是由宏观的词序域决定的。两个词语之间的强制共现的依存关系是微观的词序。宏观和微观的词序结合语用等因素，共同决定了线性的词序。这将是本文重点阐述的一个观点。在后面的文章中将会做详细的解释。

我们假设的由四个结点（词）a、b、c、d形成的交叉依存现象属于最简的一种形式。在实际的语料中，可以有两个以上的不同的词语介于这四个结点之间，因而是比较复杂的形式。如果四个节点之间又加入了新的结点，那么这个句子的形式化表示也会变得复杂。因此，前面的形式化表示需要修正，扩充为如下的形式：

$$\text{linked}(a, c) \wedge \text{linked}(b, d) \wedge a < \dots < b < \dots < c < \dots < d$$

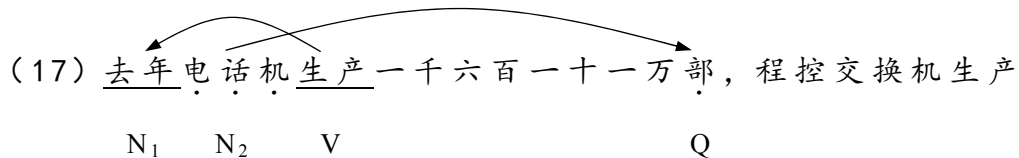
3.1.2 交叉依存类型

除去词语并列支配成分共享构成的交叉依存现象是所有语言中普遍存在的之外，另外还有一些现象是汉语中比较独特的，是我们研究的重点。这些现象可以分为九个小类的交叉依存类现象。根据句子中心谓词的多少，我们把这些现象分为单个谓词下的交叉依存和多个谓词下的交叉依存现象两大类：

第一， 单个谓词下的交叉依存现象；

单个谓词的情况是指句子的中心是由一个动词或者形容词充当，整个句子表达的是一个事件。这几类形成的交叉依存现象中，总有一组词语的依存对是跟中心谓词相关的。而另外一对词语则是不同的。根据另一对词语词性的不同，我们分为五个小类，分别如下：

1. 名词与量词分离类



三百三十四万线，微机生产十二万台，产量分别成倍增长。

在这个句子中，名词“去年”与动词“生产”之间具有依存关系，名词“电话机”与量词“部”之间也有依存关系。可以分别用连线来表示词语间的这种依存关系，称之为依存边。这样的两个边会形成交叉的现象。

我们如果用词类范畴符号来代表具体的词语，用 N_1 来代替名词“去年”， N_2 代替名词“电话机”， V 代替动词“生产”， Q 代替量词“部”，可以形式化描述为如下的条件：

- A. 存在 (N_1, V) 、 (N_2, Q) 两个依存词对；
- B. 在语序上有 $N_1 < N_2 < V < Q$ 的要求。

更多的例句如：

(18) 目前国内提供高阶主管进修课程的公私立大学共有台大、交大、中山、太叶、元智等十馀所，为什么企管所特别受社会人士青睐？

(19) 常州市沿运河两岸零点五平方公里范围内，大型厂矿企业有一百八十六家。

(20) 无锡市沿运河企业有五百多家，超亿元产值的大型企业有五十多家，成为无锡市经济发展的主骨架。

(21) 镇江市近年依托大运河建厂的外商独资企业就有八家，合资企业更多。

2. 名词与形容词分离类

(22) 为突出城市园林绿化的亚热带风光及民族特色，

南宁市 大量 种植 棕榈科植物 和开花乔木、花灌木、彩叶植物以及地

N₁ Adj V N₂

被植物、草坪等，市区的一百多条主要街道基本实现了绿化、彩化、果化、香化。

在这个句子中，名词“南宁市”与动词“种植”之间具有依存关系，形容词“大量”和名词“植物”之间也存在依存关系。可以分别用连线来表示这种依存关系。这样的两个依存边会形成交叉的现象。

我们用词类范畴符号来代表具体的词，用 N₁ 来代替名词“南宁市”，Adj 代替形容词“大量”，V 代替动词“种植”，N₂ 代替名词“植物”，可以形式化描述为如下的条件：

我们可以用词类范畴符号来代表具体的词，用 N_1 来代替名词“张学良”， X 代替“在”， N_2 代替名词“意识”， V 代替动词“强”，形式化描述为如下的条件：

A. 存在 (N_1, N_2) 、 (X, V) 两个依存对；

B. 在语序上有 $N_1 < X < N_2 < V$ 的要求。

更多的例句如：

(28) 三峡工程虽然投资巨大，但在全国所占比重很小。

(29) “博格坎普在世界杯结束前状态还不错，但是今年下半年，他的状态就差多了，”胡子长过头发的维阿里说。

(30) 戴胜通反问：「台湾如果环境不好、不好赚钱，大企业为什么都把他们在国外的第二代叫回来接班，在大企业国际化潮流如此普及的今天，他们尽可以让年轻一代留在外面创业！」

(31) 跳水选手已全部抵达罗马，并开始赛前训练。

4. 副词与数词分离类

(32) 据美英宣布，这次空袭行动共击中了100多个伊拉克

N Adv V M

军事目标，达到了预定目标。

这类结构中的数量副词一般是当做状语成分来看待的，状语修饰后面的动词。但在意义上看则是跟动词后的成分有联系。从依存的角度来看，则是副词与数词之间具有依存关系。

这句话中，副词“共”与数词“100”之间有依存关系，名词“行动”与动词“击中”之间也有依存关系。可以分别用连线来表示这种依存关系。这样的两个边会形成交叉的现象。

我们用词类范畴符号来代表具体的词，用 N 来代替名词“行动”，

A. 存在 $(N, V)(Adv, M)$ 两个依存对；
B. 具有 $N < Adv < V < M$ 的语序要求。

(33) 其中大专院校、大型科研院所和大中企业开办的企业约占百分之七十。

(35) 他们於毕业后须在香港至少工作三年。

(36) 本届残障奥运，亚洲国家中，南韩队成绩最优异，总共夺得五面金牌。

(37) 去年实现进出口总值达一千零九十八点二亿美元, 占全

由于介词连用而引起的边交叉的情况是在语言中广泛存在的。在这句话中，两个介词“由”与“到”之间存在着依存关系，名词“比重”与动词“提高”之间也存在依存关系。如果用连线来表示这种词语间的依存关系，这样两个依存边之间就会形成了交叉。

66

形式化描述为如下的两个条件：

A. 存在 $(N, V)(P_1, P_2)$ 两个依存对；

B. 具有 $N < P_1 < V < P_2$ 的强制语序。

更多的例句如：

(38) 同期平均每万人口在校大学生人数，由八点九人增加到二十五点七人。

(39) 同样的三千名员工，月产量则从二十万片暴增到二亿片。

(40) 欧洲统计局还说，欧洲联盟 1 1 月份的通胀率从一个月前的百分之一一点一降至百分之一点零。

第二，多个谓词下的交叉依存现象；

多个谓词的情况是指句子的中心是由两个动词或者一个动词加一个形容词充当。多个谓词做中心的句子，表达的是多个事件的组合。形成交叉依存的两个边基本上与两个谓词都有关。我们根据与正常语序相异的部分不同而分为四个小类，分别如下：

1. 名词与动词分离类；

(41) 人类社会刻意筑起的界线与杆格，是人为加上去的，

最後也终需 人们 将它打破。

$N_1 \quad N_2 \quad V_1 V_2$

这类现象与以往研究中的“把”字句和动补结构都有关系。这句话中，名词“人们”与动词“打”之间具有依存关系，名词（代词）“它”与动词“破”之间也有依存关系，用连线来表示这种依存关系，就会形成两个依存边交叉的现象。

我们用词类范畴符号来代表具体的词，用 N_1 来代替名词“人

们”， N_2 代替代词“它”， V_1 代替动词“打”， V_2 代替动词“破”，可以形式化描述为如下的两个条件：

- A. 存在 $(N_1 V_1) (N_2 V_2)$ 的依存对；
- B. 具有 $N_1 < \text{把} < N_2 < V_1 < V_2$ 的强制语序。

更多的例句如：

(42) 蚊子进食时就会把植物细胞毒素吸入体内。

(43) 威迪奥诺表示将把此意见转达给联合国秘书长。

(44) 经过一段时期的“投石问路”，越来越多的外商象美国百时美施贵宝公司一样，看好上海的投资环境和市场潜力，纷纷追加投资扩大再生产，并逐步把上海建成其在中国甚至亚太地区的经营、销售中心。

2. 动词与副词分离类；

(45) 近年来，这个律师事务所已在国际上打赢了十余起涉外

$N \quad Adv \quad V_1 V_2$

经济案件，为国内企业和公民追回赔款近千万美元。

这句话中，副词“已”与动词“赢”之间具有依存关系，名词“事务所”与动词“打”之间也有依存关系。用连线来表示这种依存关系，就会形成两个依存边交叉的现象。

我们可以用词类范畴符号来代表具体的词，用 N 来代替名词“事务所”， Adv 代替副词“已经”， V_1 代替动词“打”， V_2 代替动词“赢”，形式化描述为如下的两个条件：

- A. 存在 $(N V_1) (Adv V_2)$ 两个依存对；
- B. 具有 $N < Adv < V_1 < V_2$ 的强制语序。

更多的例句如：

(46)家住湖北省新洲县的叶章荣在今年罕见的大洪水中失去了家园，不过在元旦到来之前，他们一家已经住进一幢崭新的二层小楼里。

(47) 不知不觉间浓荫几乎完全封死了马路的上空，原来马路已经变窄了，并且有了坡度。

(48) 调查人员将从野马的尸体中取出子弹，以帮助缉拿凶手。

(49) 事后，舟山受害方在有关各方大力帮助下，终于调查到了肇事船船东。

3. 动词与数词分离类:

(50) 据统计，至去年年底，宁波保税区累计完成进出口贸易额八点一

N V₁ V₂ M

二亿美元，仅去年一年通过保税区海关的进出口贸易额就达三点六
五亿美元。

这句话中，名词“保税区”与动词“完成”之间具有依存关系，动词“累计”与数词“八点一二亿”之间也有依存关系。用连线来表示这种依存关系，就会形成两个依存边交叉的现象。

我们可以用词类范畴符号来代表具体的词，用 N 来代替名词“保税区”，V₁ 代替动词“完成”，V₂ 代替动词“累计”，M 代替数词“八点一二亿”，形式化描述为如下的两个条件：

A. 存在 $(N, V_2) (V_1, M)$ 两个依存对;

B. 词语之间具有 $N < V_1 < V_2 < M$ 的强制语序。

更多的例句如：

(51) 十省市合计出口一千六百二十一点九亿美元，占全国的

百分之八十八点八。

(52) 上海统计局公布的数字显示，一九九二年至一九九六年上海国内生产总值年均增长百分之十四点二，今年的增长速度也将达到百分之十三以上。

(53) 文化部的统计显示，一至十二月中旬，两岸文化交流项目总计达二百七十三项、参加人数达两千六百七十三人次。

3. 框式副词类；

(54) 使这些经济活动·一出现就被纳入法制轨道。

N Adv₁ V Adv₂

这句话中，副词“一（刚、才）”与“就”经常连用，形成一种固定的搭配形式，来表示两个动作之间的先后关系。因此两者之间有共现的关系，即句法上的依存关系。我们仿照 Greenberg (1995) 框式介词的称呼，把这类现象称之为“框式副词”。名词“活动”与动词“出现”之间也有依存关系。用连线来表示词语之间的依存关系，两个依存边之间形成了交叉。

我们可以用词类范畴符号来代表具体的词，用 N 来代替名词“活动”，Adv₁ 代替副词“一”，V₁ 代替动词“出现”，Adv₂ 代替副词“就”，形式化描述为如下的两个条件：

A. 存在 (Adv₁, Adv₂) (N, V₁) 两个依存对；

B. 词语之间具有 $N < Adv_1 < V_1 < Adv_2 < V_2$ 的强制语序。

更多的例句如：

(55) 贾坚一回忆，民国五十年代的台湾，经济困窘，民众根本无财可理；又为了怕打仗，一攒到钱就去买黄金、换黑市美钞。

(56) 连续剧卖座，电视台就没完没了地拍下去，边拍边播就

成了台湾八点档的特色，品质当然也无法太讲究。

(57)「她爸爸（吴建乐）二十二岁就结婚了，女孩子怎麼能拖？」阿嬷又爱又怜地说。

(58) 对于有足球王国之称的巴西来说，1998年是一个失望多于兴奋，胜利与失败并存，成绩既令人高兴但又让人感到缺憾的一年。

3.1.3 交叉依存现象的各类型数据分布

不同类型的交叉依存现象在汉语中的分布情况是不同的。在宾州树库的语料中，我们限定在分句的范围内，收集了名词与量词分离类 14 例，名词与形容词分离类 14 例，名词与名词分离类 16 例，副词与数词分离类 307 例，框式介词类 100 例，名词与动词分离类 126 例，副词与动词分离类 142 例，数词与动词分离类 88 例，框式副词类 28 例，共计 835 例。我们依次编号为 1-9 九个类，可以表示为如下的分布图形式：

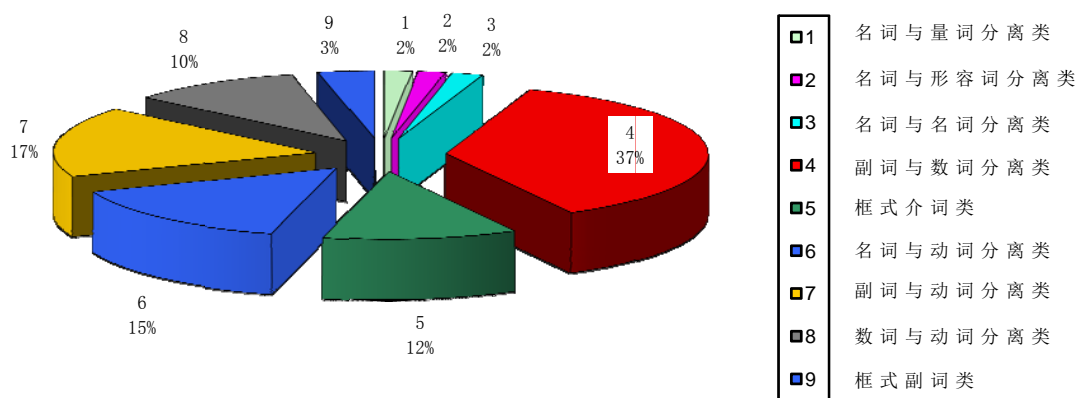
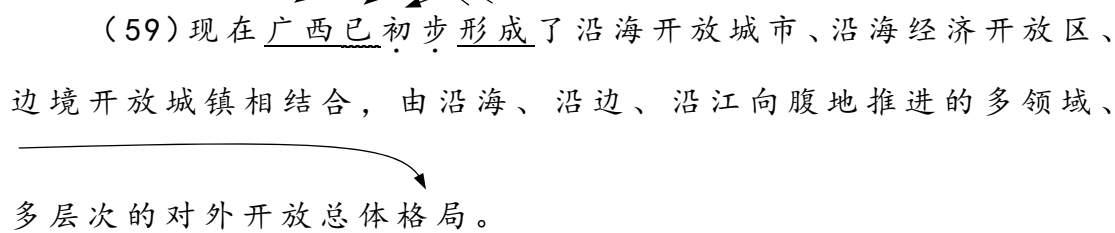


图 1 交叉依存现象分布图

可以看到，其中第四种副词与数词分离类的现象是最多的。这与语料的来源与语体有关。宾州汉语树库的语料来源是新闻语料，属于受限的语体。语料中较多的论述是各方面发展的成果与数据的

表达，因此这类现象会显得特别多。另外，第七类的副词与动词分离类的现象为第二多的，说明动补结构是非常多的。这些结构的正确分析，会对句法分析精确度的提高具有重要的意义。

另外需要指出的是，不同种类的交叉依存现象之间不是孤立的。相互之间经常叠加，可形成多重交叉的现象。如：



(59) 现在 广西 已初步 形成 了沿海开放城市、沿海经济开放区、
边境开放城镇相结合，由沿海、沿边、沿江向腹地推进的多领域、
多层次的对外开放总体格局。

这里实现了至少两个交叉的现象，一个是形容词“初步”与名词“格局”之间的边与动词“形成”与名词“广西”之间的边的交叉。属于形容词与名词分离的现象。同时，由于形容词与名词的隔断，使得副词“已”与动词“形成”之间也形成了隔断，它们之间的边也与“初步”与“格局”的边构成了交叉。因而是多重交叉的现象。但是，形容词与名词的分离是最基本的交叉现象。因此，本文的分类是针对最基本类型来说的。

3.2 单个谓词下的交叉依存现象

在这一节，我们将对单个谓词下的各种交叉依存现象进行分析。考察各种现象形成交叉的句法及语义条件。对比以往分析器分析的结果，指出其分析的不足。同时，还会构拟新的依存语法分析过程。

这里需要说明的是，每一节中所构拟的依存语法分析模型都将是针对讨论中的具体一类来说的。为了说明的简洁及集中，句法规则集只是定义到足够分析此类现象为止。另外，本章对于交叉依存现象的句法图示中，只是指出了哪些词语之间存在着依存关系以及各自的依存方向，并不给出具体依存关系种类的描述。这一方面是为了突出重点，即指出是哪些边之间形成了交叉，另一方面也是为了避免在对依存关系类别定义标准上的不同争论。在本文第五章关于语料库构建的部分，将具体谈到依存关系的定义标准问题。为使讨论的集中，便不在这里做过多的涉及。

3.2.1 名词与量词分离类

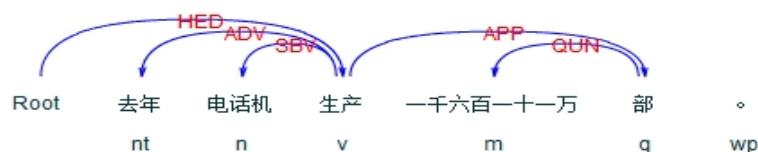
在 3.1.2 一节中，我们谈到了这类现象中形成交叉的两个词对分别是名词与动词边与名词与量词的边。依旧以语料中的用例为例，我们分析其句法语义条件：

(60) 去年，电话机生产一千六百一十一万部，程控交换机生产三百三十四万线，微机生产十二万台，产量分别成倍增长。

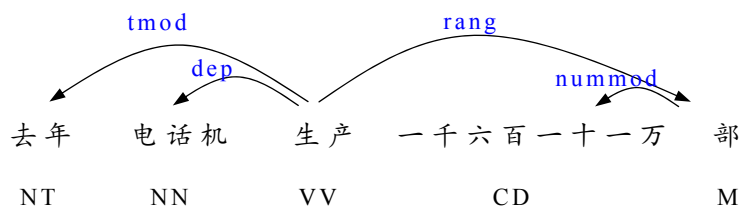
为了讨论的简洁与更有原型性，我们把枝节的成分去掉，取“去年电话机生产一千六百一十一万部”部分作为讨论的原型。

从句法分析来看，这个句子的依存结构应该还是比较简单的。但

是由于假设了投射性依存，现有的句法分析器不能分析得到正确的句法结构，下图是哈工大依存分析器的结果可以作为对比的参照。



可以看出，哈工大的分析结果把量词“部”直接依存于了动词“生产”。同样的，另一个依存分析器的代表-----斯坦福依存分析器，其分析结果也把名量词“部”依存于了动词“生产”。



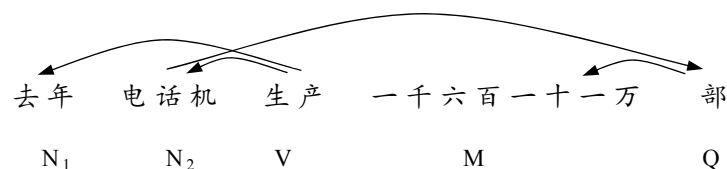
两者的句法分析结果对名量词“部”的处理都是错误的。这必然会影响到其后语义的正确处理，所以需要修正。

对这类现象，形式学派认为是发生了论元分裂移位。根据沈阳（1996）的观点，认为这个句子是动词后论元 NP 发生了分裂移位，语义所指范围比较大的成分分裂前移，语义所指范围小的成分则留在了原位。所谓语义范围大小，并不决定于某个成分是原名词词组的中心词还是修饰语。但同时，沈认为论元 NP 向前分裂移位都不是真性论元 NP 移位。留在动词后原位置的成分才是受动词直接支配的成分，或者说动词后才是成分支配位置，分裂前移的成分则不再受动词支配。

从依存的角度来看，我们会得出不同的结论。依据句法依存的判定标准，动词与名词之间有强制共现的要求，即线性相关，所以

有依存关系。就本句来说，“生产”为二价动词，要求有两个名词与之共现。而名量词与动词之间是没有直接的依存关系的，只跟名词有共现的要求。名量词要求必须有一个名词与之同现，即具有线性相关性，因而具有句法依存关系。量词与数词之间也有同现的要求，因而也具有句法依存关系。“电话机”无论是位于“生产”的之前或之后，其与“生产”之间的依存关系基本是一样的。名量词“部”与“电话机”之间也存在依存的关系。依存关系的共现是一种相对的词序，而不是绝对固定的词序。所以，上例中的名量词不能说是其受动词的支配。本句中两个名词“去年”与“电话机”之间无线性相关，且不能形成一个韵律单位，因此无句法依存关系。

本句完整的句法依存关系可描述为：二价动词“生产”与名词“电话机”有直接的配价关系。名词“去年”充当可有论元。量词“部”与名词“电话机”有依存关系。数词“一千六百一十一万”与量词“部”有依存关系。其依存关系可以表示为如下的形式：



我们定义一个简单的依存语法系统来模拟这个例句的机器分析过程：

$$G = \langle C, S, \Sigma, T \rangle$$

其中，C 为句法范畴符号的集合，S 为起始符号， Σ 为依存规则的集合；T 为终极符号集合。

$$C = (N, V, M, Q)$$

$$S = V$$

$$\Sigma = \{$$

$$V \rightarrow N;$$

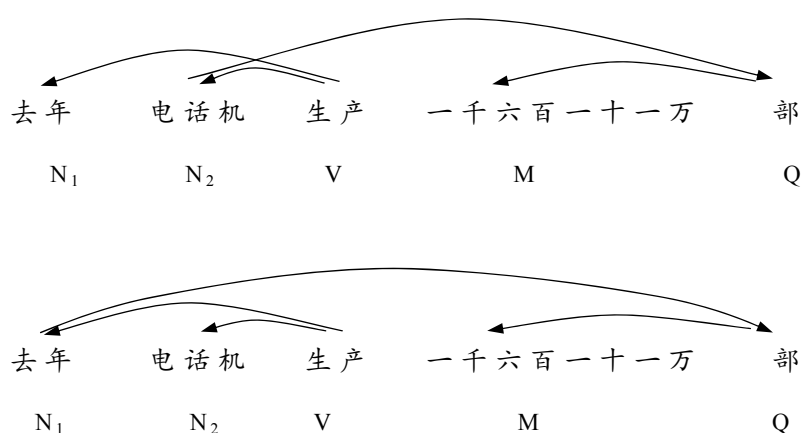
$Q \rightarrow M;$

$N \rightarrow Q;$

}

$T = \{\text{去年} = N、\text{电话机} = N、\text{生产} = V、\text{一千六百一十一万} = M、\text{部} = Q\}$

舍弃了投射性依存的限定条件，那么可能的句法分析结果形式会有两种。分别如下两图所示：



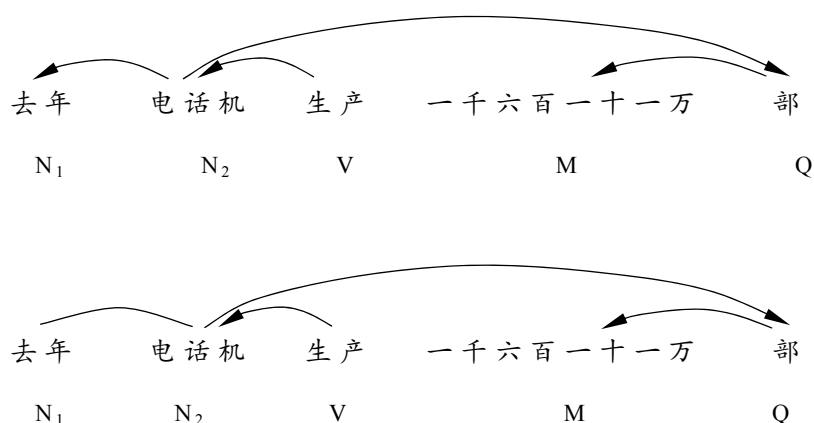
句法分析得到的两棵依存树，在加入了具体词汇的语义限定以后，将会排除掉图第二种的分析结果。因为量词“部”与名词“去年”在语义上不能够组合成功。这样只能有一种正确的分析结果，即第一个图的形式，这就达到了完全的句法分析的目的。当然，这里构建的形式语法系统是最简的，省略掉了对于两个名词之间依存关系的判断，如果加上两个名词的判断，则需要再增加几条规则：

① $N \rightarrow N;$

② $N \dashv\dashv N;$

规则 ①描述的是名词修饰名词的情况，规则 ②则是两个名词同位的现象。

如果加上这两条规则以后，则会增加另外两种分析的可能：



但是在加入词汇意义的限制之后，就会逐个地排除，达到最后正确的分析结果。

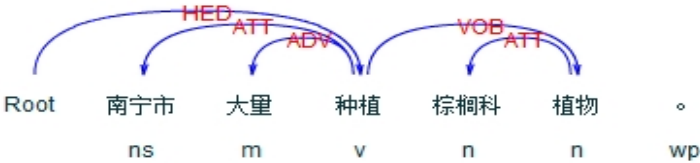
3.2.2 形容词与名词分离类

在 3.1.2 一节，我们谈到了这类现象中形成交叉的两个词对分别是名词与动词的边和形容词与名词的边。依旧以语料中的用例为例，我们分析其句法语义条件：

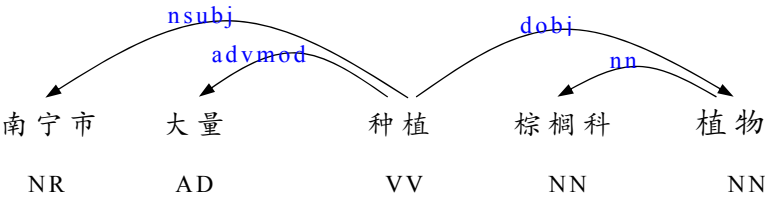
(61) 为突出城市园林绿化的亚热带风光及民族特色，南宁市大量种植棕榈科植物和开花乔木、花灌木、彩叶植物以及地被植物、草坪等，市区的一百多条主要街道基本实现了绿化、彩化、果化、香化。

为讨论的简洁，我们可以剪除枝节的成分，构拟一个更加原型的句子作为讨论的基础，为“南宁市大量种植棕榈科植物”。把复句简化为一个简单单句来分析。以往的语法分析中，一般把“大量”看做是状语成分，是直接修饰后面动词的。我们认为，是形容词“大量”与名词“植物”之间有句法依存关系，但是在语序上处于分离的位置。

下图是哈工大依存分析器的分析结果，可以看成是投射性依存假设下的分析代表：

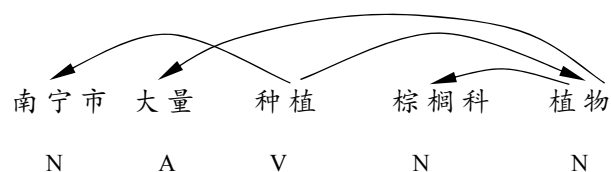


可以看出，哈工大的分析器中把形容词“大量”直接依存于了动词“种植”。同样的，斯坦福依存分析器的分析结果也把形容词“大量”直接依存于了动词“种植”。



两个句法分析器都坚持了投射性依存的条件。不同的是在依存关系上，哈工大把名词“南宁市”与动词“种植”的关系处理为定中关系，斯坦福分析器则处理为主语关系。

从依存语法的角度来分析，汉语中的形容词跟名词或动词之间都有可能发生依存关系，因而是比较复杂的。需要从具体的内容上去判断其修饰的部分。“大量”作为一个表数量的属性词，跟名词之间有共现的关系。在本句中，由于有三个名词，理论上都有可能是形容词修饰的对象。一个是“南宁市”，一个是“棕榈科”，一个是“植物”。但只有“植物”与“大量”两者之间在属性上是相匹配的。两者之间有句法上的依存关系。同时，其语序虽然分离，但依然遵循着域的原则，即形容词与名词同属于第三域内。这是在下一章环形词序域部分将着重论述的内容。整个句子的依存关系可以图示如下：



我们可以定义一个简单的依存语法系统来模拟这个例句的机器分析过程：

$$G = \langle C, S, \Sigma, T \rangle$$

其中， C 为句法范畴符号的集合， S 为起始符号， Σ 为依存规则的集合； T 为终极符号集合。

$$C = \{ N, V, A \}$$

$$S = V$$

$$\Sigma = \{$$

$$V \rightarrow N;$$

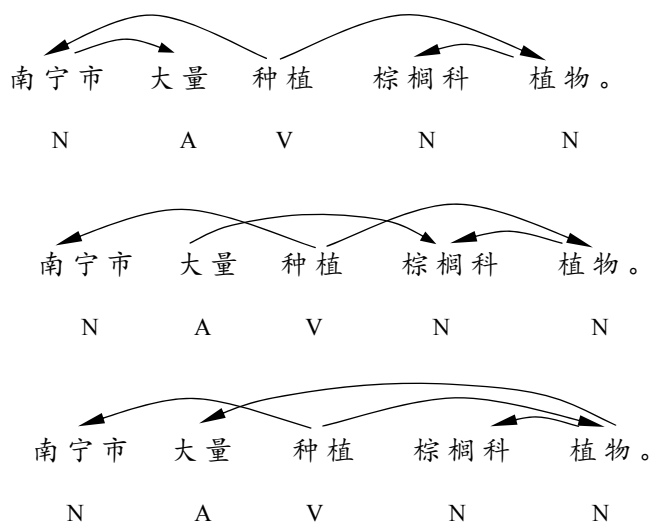
$$N \rightarrow N;$$

$$N \rightarrow A;$$

$$\}$$

$$T = \{ \text{南宁市} = N, \text{大量} = A, \text{种植} = V, \text{棕榈科} = N, \text{植物} = N \}$$

这个依存分析器分析的结果将有三种可能，分别为如下的形式：



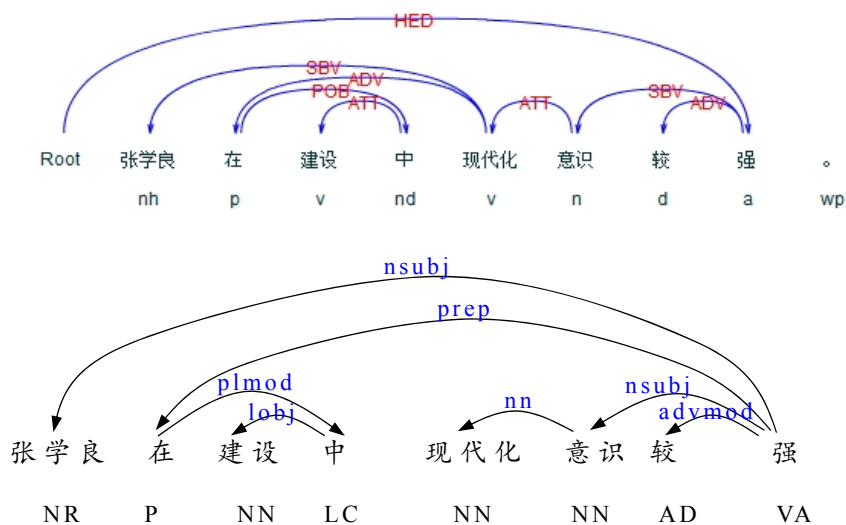
可见，单凭词类范畴本身是不能够很好地消歧得到唯一的分析结果的。必须借助于词语本身之间的意义。由于多个名词的出现，使得作为修饰语形容词与名词有了若干种组合的可能。这里省略了名词“棕榈科”与“植物”之间修饰关系的分析过程。

3.2.3 名词与名词分离类

这类现象中，形成交叉的两个词对分别是介于两个名词之间的词语与形容词的词对与两个名词构成的词对。依旧以语料中的用例为例，我们分析其句法语义条件：

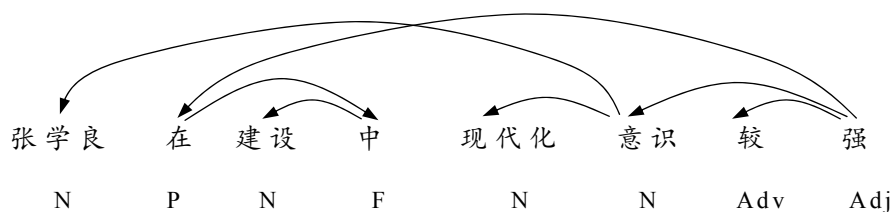
(62) 张学良在东北建设中现代化意识较强，大胆开创。

名词“张学良”与名词“意识”之间的领有关系是明显的，因而具有句法上的依存关系。我们把这个句子简化为“张学良在建设现代化意识较强”作为讨论的基础。首先，我们看哈工大依存分析器和斯坦福分析器的分析结果，依次如下两图所示：



可以看出，两者的分析还是差别很大的。哈工大的分析器中关注到了“张学良”与“意识”之间的关系，但却错误地依存于“现代化”。而且，把“在”连接的成分依存于了“现代化”，也是错误的分析结果。斯坦福的分析器则没有关注到两个名词“张学良”与“意识”之间的依存关系。

从依存语法的角度来看，介词与动词之间有强制共现的要求，因此一般是与动词有依存关系。就本句来说，介词“在”应该是依存于动词“强”。而“意识”与“张学良”之间也有强制共现的关系。这句话正确的句法依存应该表示为如下的形式：



我们可以定义一个简单的依存语法系统来模拟这个例句的机器分析过程：

$$G = \langle C, S, \Sigma, T \rangle$$

其中，C 为句法范畴符号的集合，S 为起始符号， Σ 为依存规则的集合；T 为终极符号集合。

$$C = (N, P, F, Adv, Adj)$$

$$S = Adj$$

$$\Sigma = \{$$

$$Adj \rightarrow N;$$

$$Adj \rightarrow P;$$

$$Adj \rightarrow Adv;$$

$$P \rightarrow F;$$

$$F \rightarrow N;$$

$N \rightarrow N;$

}

$T = \{\text{张学良} = N、\text{在} = P、\text{建设} = N、\text{中} = F、\text{现代化} = N、\text{意识} = N、\text{较} = Adv、\text{强} = Adj\}$

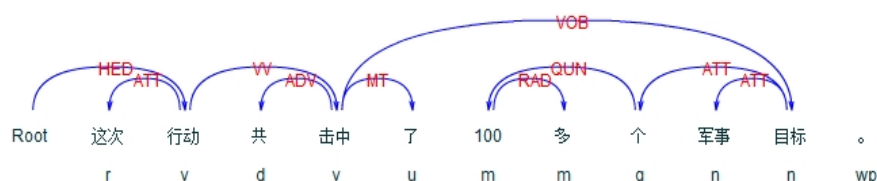
在增加方位词与名词左向相邻依存的条件下，可以把介词及其支配部分直接依存于形容词，但是几个名词之间的依存关系，却不得不借助词义才能得到唯一正确的分析结果。

3.2.4 副词与数词分离类

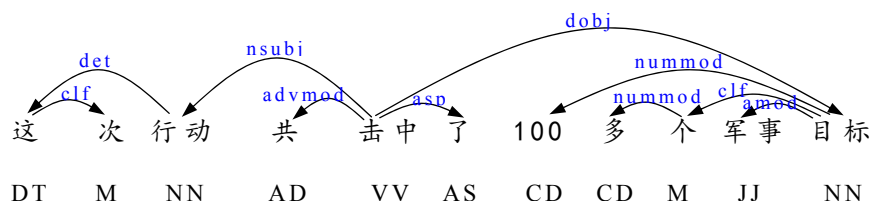
屈折语中副词一般与动词有同现关系。汉语中的副词则不仅可以跟动词同现，还可以跟数词、名词等更多的词类同现。有一些副词是只用来修饰数词的，与数词有句法上的共现要求。如果在语序上分离开，便可形成边交叉的情况。语料库中的用例如下：

(63) 据美英宣布，这次空袭行动共击中了 100 多个伊拉克军事目标，达到了预定目标。

我们把第一个句子简化为“这次行动共击中了 100 多个军事目标”作为讨论的原型来看现有句法分析的情况。副词“共”与数词“100”之间有句法依存关系，名词“行动”与动词“击中”之间也有句法依存关系，两个边形成交叉，因而是非投射性依存的现象。但是在投射性依存的假设下，却不能够得到正确的分析结果。



这是哈工大分析器的分析结果。可以看出，其为了遵守投射性的要求，把副词“共”直接依存于动词“击中”，而没有关注到其作为数量副词与数词之间的直接关联。同样的，斯坦福依存分析器也把副词“共”直接依存于动词“击中”。这是错误的。



一般认为，副词是直接依存于动词结点的，其充当的是动词的“状态元”。这是就印欧屈折语来说的。在汉语中，副词并不只是修饰动词的，还可以修饰数词、名词、形容词等。因此不能笼统来说副词都是动词的支配成分，而应该根据词义的不同做不同的判断。副词修饰数量词的用法有马真（1981）做过论述，其列举了 40 多个修饰数量词的副词，分别如下：

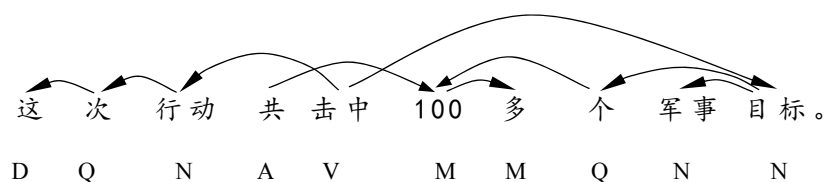
正好 刚好 恰好 恰巧 恰恰 刚 刚刚 已经 只 仅 仅仅
 就 才 不过 足足 足 都 也 大概 大约 约 约摸 大致
 大抵 也许 将近 最多 至多 顶多 最少 至少 的确
 真的 果真 果然 当真 共 总共 一共 一总

但是，能够形成交叉依存现象的在语料库中只发现了“只、足足、大概、约、大约、大致、最多、顶多、最少、至少、共、总共、一共、整整”等。另外，在语料库中还存在“整整”一词，是马真的论述中没有提到，我们把“整整”也划归到此类，大约可以划归到其七个分类中的“言多”类。词典中“整整”的释义为：

整整 圖 达到一个整数的：～忙活了一天 | 到北京已经～三年了。

对修饰数词的副词的位置，马真（1981）指出，当数量词处于宾语位置的时候，副词一般放在动词之前。动词后面带了“了”，这些副词才能直接放在数量词之前。（少量副词仍然不能直接放在数量词前）

从依存关系的角度来看，不考虑投射性假设的要求，副词“共”是跟数词有同现的要求的。“共”出现时，可以没有动词的出现，但是不能没有数词。而且从韵律的角度来看，可以形成“共 100 多个”的韵律单位，因此可以判定具有句法依存的关系。相反，“共”与“击中”既无语义上的关系，不能形成线性相关，也不能形成一个韵律单位。因此可判断两者并无句法上的依存关系。这句话正确的句法依存应该表示为如下的形式：



这样就出现了两个依存边交叉的情况。副词“共”与数词“100”之间的边分句与“行动”与“击中”的边、“击中”与“目标”的边都形成了交叉。这不同于投射性依存假设下的分析结果。我们可以定义一个简单的依存语法系统来模拟这个例句的机器分析过程：

$$G = \langle C, S, \Sigma, T \rangle$$

其中，C 为句法范畴符号的集合，S 为起始符号， Σ 为依存规则的集合；T 为终极符号集合。

$$C = (N, V, M, Q, \text{Art}, \text{Adv})$$

$S = V$

$\Sigma = \{$

$V \rightarrow N;$

$Q \rightarrow M;$

$Adv \rightarrow M;$

$N \rightarrow Q;$

$N \rightarrow M;$

$N \rightarrow N;$

$M \rightarrow Art;$

$\}$

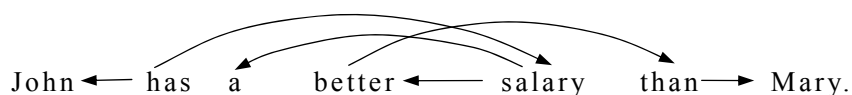
$T = \{ \text{这} = Art, \text{次} = M, \text{行动} = N, \text{共} = Adv, \text{击中} = V, 100 = M, \text{多} = M, \text{个} = Q, \text{军事} = N, \text{目标} = N \}$

这里我们可以增加一个原则，即量词在对数词和定冠词的支配上有左向相邻的要求。加上量词在与名词的依存上有右向优先的要求。两条原则一起作用，这样即可分析得到合法的句法分析结果。当然，这些具体的规则只是针对这个句子的，一个能够囊括全部语法现象的规则要复杂的多。例如副词还有依存于动词的规则，需要在语义层面才能得以消歧。

3.2.5 框式介词类

Greenberg (1995) 最初提出的框式介词的概念，是指前置介词与后置介词的连用而形成的框架结构。我们借用其概念，把介词连用的情况也称之为框式介词。把由于介词连用而形成的交叉依存现象称之为框式介词类的现象。

一些交叉依存现象是与介词有关的。Mel'čuk (1988) 在非投射性依存现象中的举例中即有一例涉及到介词，如下图所示：



英语中形容词“better”与介词“than”之间的依存关系是明显的，介词“than”的出现要求必然有一个比较级的形容词与之同现。

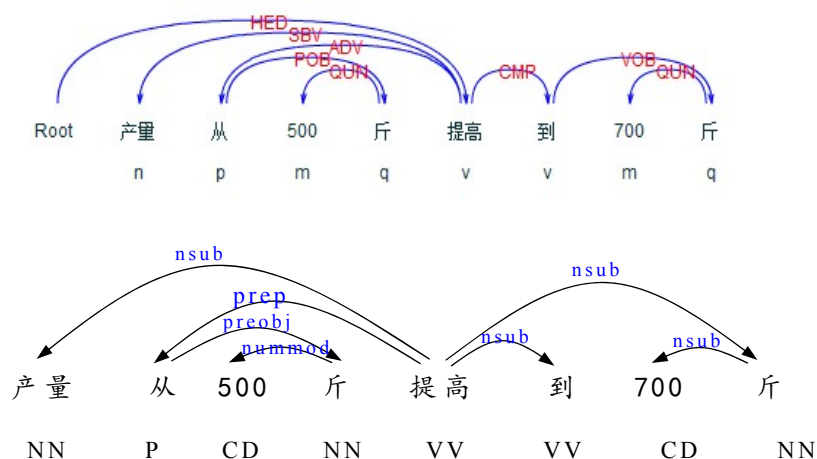
我们这里探讨的则是两个介词连用的情况。如语料库中提到的

(64) 去年实现进出口总值达一千零九十八点二亿美元，占全国进出口总值的比重由上年的百分之三十七提高到百分之三十九。

介词连用可以发生在复句的范围内，这样会使得交叉依存发生的范围变得很大，为讨论的简洁，我们限定在单句中出现的介词连用。同样也构拟一个原型的句子做为讨论的对象。

(65) 产量从 500 斤提高到 700 斤。

下两个图示依次分别是哈工大依存分析器和斯坦福依存分析器的分析结果：



可以看出，两个分析器都不认为两个介词之间有依存的关系，都只承认“从”是介词，同时把“到”分析为了动词。斯坦福的分析甚至不认为“到”与“斤”之间存在依存关系。

在汉语中，介词与谓词之间有依存关系的存在是大家都认同的。介词可以为动词不能直接赋元的成分赋予论元角色。汉语中的介词总体上是在动词前出现的多，少量的介词与其所带的名词放在动词之后，是古汉语中更为常见。据刘丹青（2003）的观察，介词放在动词后的情况主要有两类：一类是表示方向终点类，如“向、往、到”等；一类是与事的标记“给”。现有的处理多是把连用的介词分为两个部分，各个部分分别依存于动词。但“到”作为介词的用法是大家都基本同意的。

以往的介词研究，虽也谈到了一些介词与其他词类连用的情况，但较少关注到两个介词连用的时候有强制共现的要求。朱德熙《语法讲义》中提到了介词“连”与副词“都、也”连用的情况，刘丹青（2003）中把以往认为是名词中一个子类的方位词看做是后置的介词，前置的介词与后置介词共同组成一个框式结构。但对其他类的由两个介词构成的框式结构类则较少有人提到。我们这里谈到的介词框架，是两个介词的共现形成的。尽管“从。。。到。。。中的“到”是否是介词还是存在争议的，朱德熙先生（1982）即认为是动词，把表示终点的“到”看做是动词的补语。刘丹青（2003）看做是介词。对比日语中表示终点的介词“ni”，和英语中的“to”，我们认为处理为介词可能更加合理一些。这种现象 Li（1990）解释为动词与介词的“再分析”。冯胜利（1996）则进一步指出是汉语的韵律结构导致了介词与动词的紧缩。

“到”在有时体标记“了”的时候，往往用于介词之后。这可以认为是后置的介词还没有完全与动词分化开，这是演化过程中的一种现象。也可以看做是双音化的影响使得介词在音节上并入了动词。冯胜利（1996）也指出是韵律结构平面与句法平面相互作用形

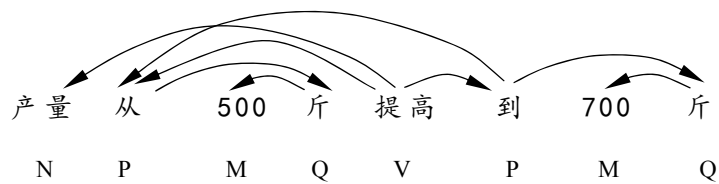
成的。我们认为，把前后两个介词看做是介词的组合出现可能更合理一些，这样可以把一部分以往认为是时间和地点宾语的成分分化出来，避免宾语范畴的外延过大，利于应用系统处理更加明确和方便。连用的介词对中，第二个介词也要求有第一个介词与之同现。正如例句中的“从...到...”“由...到...”。

在武汉大学特征结构语料库标注中，把“从...到...”“由...到...”类的介词连用现象作为一种“特征”来看待，表达的是两个实体之间的关系。这是纯粹从语义的角度来看的。但是其对句法处理的启发也是很大的，以往把两个介词短语分隔开的做法多是跟语法研究中忽视对后置介词的研究造成的。分割的做法丢失了很重要的语法信息，况且两者之间有强制共现的关系，做为“框式介词”是应该在句法层面得以反映的。如果两者之间的词语与两者之外的词语发生了依存关系，就必然会发生边的交叉，形成交叉依存类的非投射性现象。

“从、动词、到”三个词语之间的顺序是按照时间的先后来排列的。这是因为介词都是由动词虚化而来的，因而在属性特征上与动词相类似。介词的使用一般是与动词的属性相适应的。其线性序列的先后一般受时间顺序象似性原则的制约。以动词为参照，两个介词及其支配的词语分布于动词的两边。当然，象似性原则只表示一种倾向，并不是绝对的规则。

需要注意，介词的连用与关联词的连用是不一样的。关联词的连用表达的更多的是一种逻辑上的概念，介词的连用则有比较实在的意义。而且，关联词的使用在位置上也比较自由，而介词的位置则是相对稳定的。当然，连词都是从介词逐渐虚化而来的，两者之间的绝对界限也很难清楚地划界。

我们构拟的最简状态例句，其形成的交叉依存现象可以表示如下图。两个介词分别都与动词有依存关系，同时，两个介词之间也有句法上的依存关系。



我们可以定义一个简单的依存语法系统来模拟这个例句的机器分析过程：

$$G = \langle C, S, \Sigma, T \rangle$$

其中，C 为句法范畴符号的集合，S 为起始符号， Σ 为依存规则的集合；T 为终极符号集合。

$$C = (N, V, P, M, Q)$$

$$S = V$$

$$\Sigma = \{$$

$$V \rightarrow N;$$

$$V \rightarrow P;$$

$$Q \rightarrow M;$$

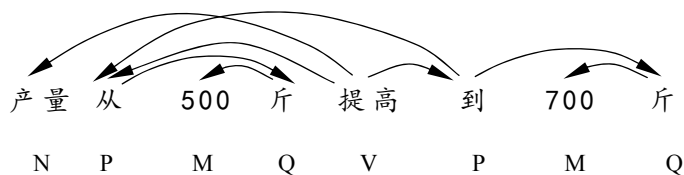
$$P \rightarrow Q;$$

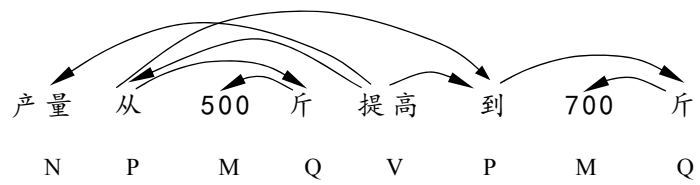
$$P \rightarrow P;$$

$$\}$$

$$T = \{ \text{产量} = N, \text{从} = P, 500 = M, \text{斤} = Q, \text{提高} = V, \text{到} = P, 700 = M \}$$

汉语中的介词在支配其他词类时是有明确的右向方向性的，所以这个分析器分析的结果将有两种可能，如下两图所示：





两个分析结果从句法的角度来看都是正确的，汉语中连用的介词，既有第一个介词要求有第二个介词共现的情况，也有第二个介词要求有第一个介词共现的情况。其消歧要借助具体的词汇语义。在加入语义的限制以后，会排除掉第二种分析的结果，是终点的“到”要求有起点的“从”共现。第一种分析为唯一正确的结果。

3.3 多个谓词下的交叉依存现象

多个谓词下的交叉依存现象不同于单个谓词下的交叉依存现象。在这一节，我们对多个谓词下的各种交叉依存现象进行分析。考察各种现象形成交叉的句法及语义条件。对比以往分析器分析的结果，指出其分析的不足。同时，还构拟新的分析过程。

与前一节相同的是，每一节中所构拟的依存语法分析模型都将只是针对讨论中的具体一类来说的。为了说明的简洁及集中，句法规则集只是定义到足够分析此类现象为止。另外，句法图示中，也只指出哪些词语之间存在着依存关系以及各自的依存方向，并不给出具体依存关系种类的描述。

3.3.1 名词与动词分离类

此类现象即以往研究中的一些“把”字句，且其动词部分由两个谓词构成，也即以往谈到的动补结构。因此与“把”字句和动补结构的研究都有关系。“把”也可用“将”来替换。“把”字句在汉语研究中具有很长的历史。研究的专家多，取得的成就也很大。在句法、语义层面的观察和描述上可以说基本上已经清晰，因此，我们在描写上可以有很多的借鉴。

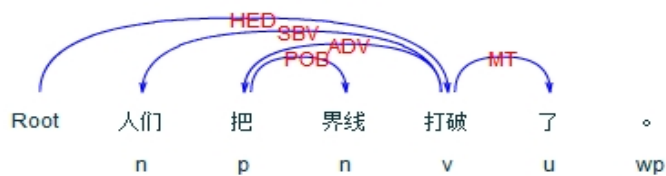
“把”字句相关的交叉依存现象主要是“名+把+名+动+动(形)”的形式。两个名词与两个谓词的依存边构成交叉。

按照学者们的观察，“把”之前的成分可以有九种语义角色。分别为施事、准施事、致事、工具、系事、因事、遇事。“把”后的宾语与动词的语义关系可以有受事、使事、处所、工具、成事、当事和施事等。“把”作为一个介词，有给动词不能直接赋元的成分赋元的功能（刘丹青 2003）。因此，“把”之后的名词不仅仅是

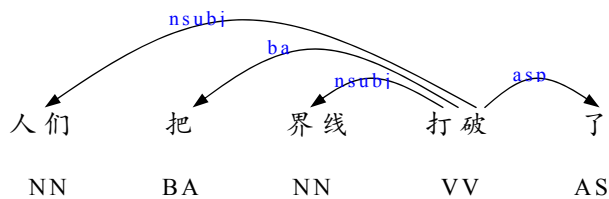
动词的受事。谓词成分的讨论也很广泛，对于本类涉及到的动补类，学者们指出补语的成分可以是结果、趋向、程度、情态类。而且，我们发现这些词语都是只跟“把”后的名词有直接的语义关系。如下例：

(66) 人类社会中刻意筑起的界线与杆格，是人为加上去的，最後也终需人们将它打破。

为讨论简洁，我们构建一个比较原型的句子“人们把界线打破了”来作为说明的基础。下图为哈工大依存分析器的处理结果：



可以看出，这里哈工大的处理把“打破”作为一个动词来处理，是一种简单化的处理方式。但从其他的处理来看，并没有保持其处理的一致性，有些动补也分作了两个词处理。另外一个依存分析器的代表——斯坦福依存分析器同样也把“打破”作为一个词来处理，但与哈工大分析器不同的是，并不认为“把”与“界线”之间有依存的关系。如下图所示：



首先，我们需要讨论的是把其中的“动+形”看做是一个词还是两个词的不同。如果看作是一个词，虽然在形式上看简化

了处理，但会引起很多后续的问题。如对配价的确定，所带成分的角色，都会有异于看做是两个词的情况。更重要的是，这种组合的可能性是无限多的，其可以形成的词语不具有可穷尽性而无法一一列举。因而，我们认同把动补看做是词语组合的方式。然后分别探讨各个谓词与其支配成分组合的关系。

我们把分析“名+把+名+动+动（形）”类交叉依存现象时定义的依存语法系统进行拓展，模拟机器分析的过程：

$$G = \langle C, S, \Sigma, T \rangle$$

其中，C 为句法范畴符号的集合，S 为起始符号， Σ 为依存规则的集合；T 为终极符号集合。

$$C = (N, Pre, V, Adj, Part)$$

$$S = V \& Adj$$

$$\Sigma = \{$$

$$V \rightarrow N;$$

$$Pre \rightarrow N;$$

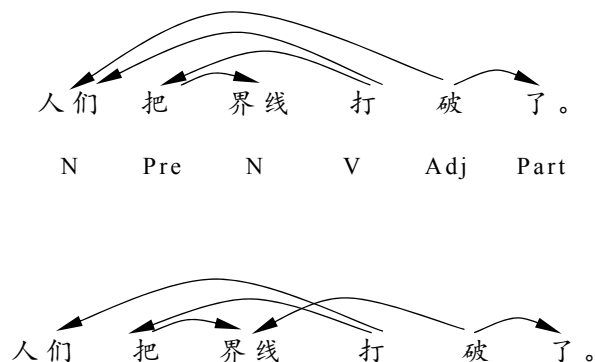
$$Adj \rightarrow Part;$$

$$Adj \rightarrow N;$$

$$\}$$

$$T = \{ \text{人们} = N, \text{把} = Pre, \text{界线} = N, \text{打} = V, \text{破} = Adj, \text{了。} = Part, \}$$

加入词序的限定规则，即前置介词的右向连接，这句话可以得到两个可能的句法分析结果，分别如下两图所示：



N Pre N V Adj Part

加入词语语义的限定后，可以排除第一个分析的结果，因此第二个为唯一正确的句法分析结果。

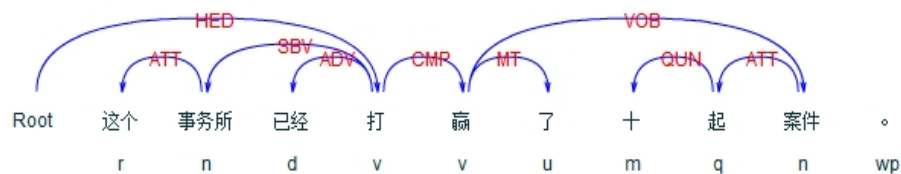
这样的分析中，名词“界线”有两个父节点，一个是“把”，一个是“破”。这就突破了以往依存分析中单一父节点的限定。因而不是树的形式，而是一个有向图的形式。单一父节点的假设，使得一个名词同时受两个或两个以上谓词支配的现象不能得以呈现，这样在语义的映射中就会丢失可能的正确语义形式。我们知道，句法分析的最终目的是为了语义的理解，所以，在句法层面可以做的事情没有必要推到语义的层面，这也正是以往依存分析存在的一个重要问题。因此还需要增加几种分析的可能。

3.3.2 副词与动词分离类

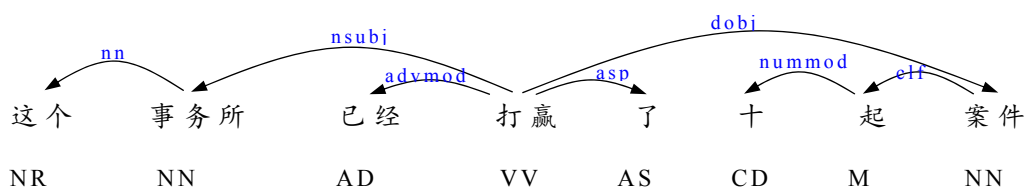
这类现象也多与以往研究中的动补结构紧密相关。一般采用的是“X+副词+动词+动词（形容词）”的形式。“名词+副词+动词+动词（形容词）”类现象在汉语中广泛存在，如下面这个实例：

(67) 近年来，这个律师事务所已在国际上打赢了十余起涉外经济案件，为国内企业和公民追回赔款近千万美元。

这是在宾州大学汉语树库中发现的实际语料。为便于讨论，我们可以简化为“这个事务所已经打赢了十起案件”。我们把这个句子作为讨论的原型。先以哈工大依存分析器的分析结果为例，来看投射性依存假设下的处理方式。



可以看出，哈工大的分析结果并没有像前一节中那样把两个谓词的组合形式看成是一个词，而是分为了“打”和“赢”两个动词。另外，把时间副词“已经”依存于动词“打”，而不与带有时体标记的“赢”有依存关系。而在斯坦福的分析结果中则直接作为一个词来看待。如下图：



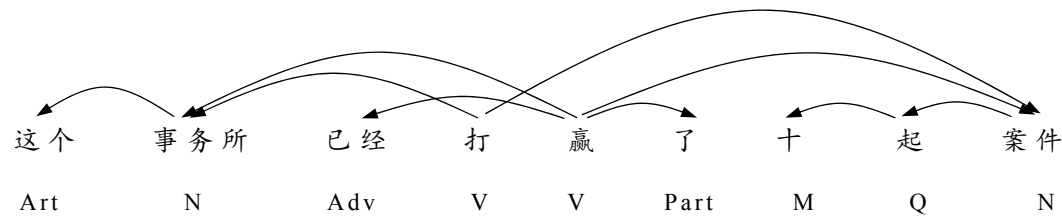
把两个谓词看做是一个词是简单化的处理方式，在从句法到语义的映射过程中容易出现问題。如果作为两个词来看，单纯依靠词性很难确定是第一个谓词还是第二个谓词决定了副词“已经”的位置。

Tenny (2000) 把动词分为了三大类，分别为 (verbs with core events and a measure or path)，有核心事件但无度量或路径的动词 (verbs with core events but no measure or path)，无核心事件的动词 (verbs without core events)。这里的有核心事件的动词则可以大致对应于汉语中的动补结构。相应地，他把副词也分为了三大类，度量或部分副词 (the measure or partitive adverbs)，重复副词 (the restitutive adverbs)，几乎类副词 (the adverbs exemplified by almost)。不同的副词类通过不同的方式与事件结构发生联系。Tenny 认为，副词是与核心事件的终结状态 (endstate) 相联系的。

根据 Tenny (2000) 的观点，副词应该是与第二谓词的直接关系更加明显。在下一章纬度和谐的部分，我们也将论述到，副词“已经”要求必须与带有时体标记的谓词与之共现，因此，我们可以选择比较接近于语义表达的方式，即让副词依存于第二个谓词。这样就形成了交叉依存的现象。

这种前于关系的发生，在表面上看则是两个谓词支配的成分之间的构成仿效。即第二个谓词支配的成分放到了第一谓词之前，且其句法位置遵照第一谓词的一般习惯。不能说“这个事务所打已经赢了十起案件”。在词序上，副词“已经”有放在动词之前的要求。其制约的原因我们将在下一章探讨。

这句话正确的句法关系表示应该如下面的形式：



我们构拟一个依存语法系统来机器分析的过程如下：

$$G = \langle C, S, \Sigma, T \rangle$$

其中，C 为句法范畴符号的集合，S 为起始符号， Σ 为依存规则的集合；T 为终极符号集合。

$$C = \{ \text{Art}, \text{N}, \text{Adv}, \text{V}, \text{Part}, \text{M}, \text{Q} \}$$

$$S = \text{V}$$

$$\Sigma = \{$$

$$\text{V} \rightarrow \text{N};$$

$$\text{V} \rightarrow \text{Part};$$

$$\text{V} \rightarrow \text{Adv};$$

$$\text{N} \rightarrow \text{Art};$$

$$\text{N} \rightarrow \text{Q};$$

$Q \rightarrow M;$

}

$T = \{\text{这个}=\text{Art}, \text{事务所}=\text{N}, \text{已经}=\text{Adv}, \text{打}=\text{V}, \text{赢}=\text{V}, \text{了}=\text{Part}, \text{十}=\text{M}, \text{起}=\text{Q}, \text{案件}=\text{N}\}$

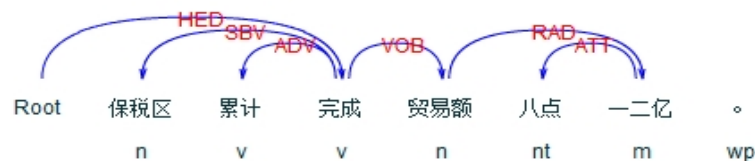
分析的结果将只有一种合法的结果，即上图的形式。当然，这里构拟的句法系统只是足够分析这一个句子而已，一个完整的句法系统规则要复杂的多，还需要增加词汇的、语序的限定规则等。

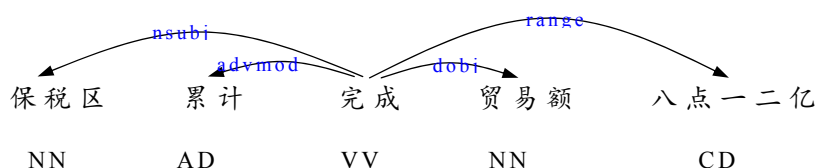
3.3.3 动词与数词分离类

在 3.2.4 一节中，我们谈到副词可以直接与数词有依存关系，其依存边与其他的边构成交叉的情况。这一节中的情况则是动词与数词有依存关系，动词与数词的依存边与其他的边形成了交叉的情况。例句如下：

(68) 据统计，至去年年底，宁波保税区累计完成进出口贸易额八点一二亿美元，仅去年一年通过保税区海关的进出口贸易额就达三点六五亿美元。

我们以第一句为基础，构拟一个原型的句子“保税区累计完成贸易额八点一二亿”作为讨论的基础。下两图依次分别为哈工大依存分析器与斯坦福依存分析器的分析结果，可以作为现有投射性依存假设之下依存分析的代表。





可以看出，无论是哈工大的分析结果还是斯坦福的分析都存在很多问题。在投射性的假设下，两个分析器都把“累计”依存于动词“完成”。斯坦福分析器甚至把动词“累计”的词性也错误分析为副词。

我们从依存语法的角度来分析，“累计”作为动词，在语义上要求有表示数量的词语与其共现，具有线性相关性，因此两者之间具有依存的关系。动词“完成”与名词“保税区”之间也存在依存边，其与“累计”与数词之间的依存边一起，必然会形成交叉的现象，形成交叉依存类的非投射性依存现象。类似的动词还有“总计、合计、平均”等。例句中两个动词“累计”与“完成”之间的先后关系并没有遵从严格意义上的时间原则，“累计”放在了前面。从词序的角度来看具有强制性，

其中，“贸易额”与“八点一二亿”之间的位置还是可以改变的，但是并不影响交叉依存现象的形成。只是所表达的语法意义有部分的差异，是下文将论述到的“表述”与“描述”的差异。“累计”与数量词之间的依存关系是先验性的，并不因为距离的隔断而中断。跟前面几种现象相类似，我们也可以看做是“累计”支配的成分后置于第二个动词，其词序参照 V_2 支配对象的语序。

我们构拟一个依存语法系统来模拟机器分析的过程：

$$G = \langle C, S, \Sigma, T \rangle$$

其中， C 为句法范畴符号的集合， S 为起始符号， Σ 为依存规则的集合； T 为终极符号集合。

$$C = (N, V, M, Q)$$

$$S = V$$

$$\Sigma = \{$$

$$V \rightarrow N;$$

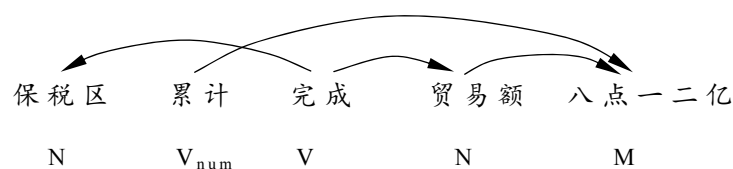
$$V_{\text{num}} \rightarrow M;$$

$$N \rightarrow M;$$

$$\}$$

$$T = \{ \text{保税区} = N, \text{累计} = V_{\text{num}}, \text{完成} = V, \text{贸易额} = N, \text{八点一二亿} = M, \}$$

分析的结果将只有一种合法的结果。



从图中可以看出，以往作为黄金标准的独根节点和单一父节点原则在本例中都被突破，不再是依存树的形式，而是依存图的形式。

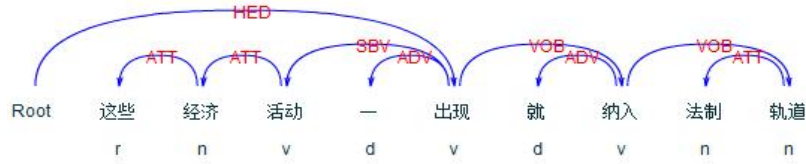
3.3.4 框式副词类

仿照前文框式介词的用法，我们把副词连用的现象称之为“框式副词”。框式副词在所有语言中都是广泛存在的。

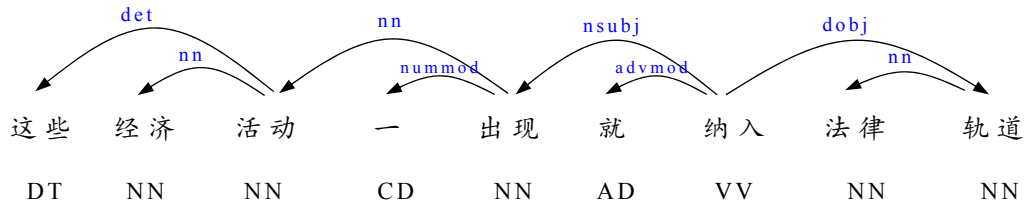
(69)对此，浦东不是简单的采取“干一段时间，等积累了经验以后再制定法规条例”的做法，而是借鉴发达国家和深圳等特区的经验教训，聘请国内外有关专家学者，积极、及时地制定和推出法规性文件，使这些经济活动一出现就被纳入法制轨道。

由于习用的缘故，词语之间的共现具有强制性，即线性相关，因而具有句法依存关系。副词组合在一定程度上与关联词的组合具有相似性，表示的是一种逻辑关系。如事理上的先后、因果、条件等。因而可做类似于连词的处理。但是副词单用的情况却是不同的，是对单个事件内性状的修饰，应该区别对待。由于副词连用经常出现在不同的分句之间，这样就会使交叉依存讨论的范围大大扩大，我们限定在单句内出现连用的情况。

为简化讨论，我们构拟“这些经济活动一出现就纳入法制轨道”这个句子作为讨论的原型。先看投射性依存假设下的哈工大依存分析器的分析结果：



可以看出，哈工大的处理结果不认为两个副词之间有依存关系。而且，把“活动”也分析为动词。同样的，作为依存分析另一个代表的斯坦福分析器也不认为“一”与“就”之间有句法上的联系。而且，斯坦福分析结果把动词“出现”错误分析为名词。如下图所示：



理论上讲，这里的“一”应该作为副词更为妥当。“一…就…”的格式属于一种固定的框架结构，表示动作的先后关系，因此可以

认为具有句法上的依存关系。作为整体处理可以使判断更加直接。同时，例句中的“活动”与“出现”之间也存在着依存关系，这样两个依存边就形成了交叉。

我们构拟一个依存语法系统来模拟机器分析的过程如下：

$$G = \langle C, S, \Sigma, T \rangle$$

其中，C 为句法范畴符号的集合，S 为起始符号， Σ 为依存规则的集合；T 为终极符号集合。

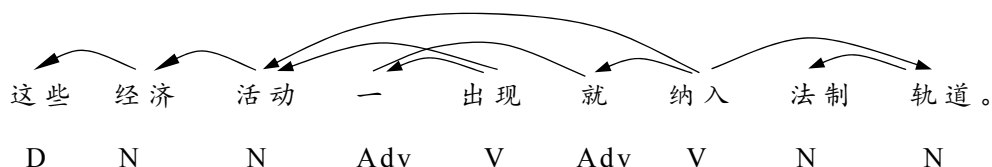
$$C = (D, N, V, M, Adv)$$

$$S = V$$

$$\Sigma = \{ \\ V \rightarrow N; \\ V \rightarrow Adv; \\ Adv \rightarrow Adv; \\ N \rightarrow D; \\ N \rightarrow N; \\ \}$$

$T = \{ \text{这些}=D, \text{经济}=N, \text{活动}=N, \text{一}=Adv, \text{出现}=V, \text{就}=Adv, \text{纳入}=V, \text{法制}=N, \text{轨道}=N \}$

加入词汇语义的限定和词序限定条件，分析的结果将只有一种合法的结果。其正确的分析结果应该如下的形式：



3.4 小 结

本章分别对汉语中不同类型的交叉依存现象进行了分析。这几

种交叉依存现象,是我们研究的切入点。在更大的语料范围内,可能还会发现新的类型。因此,在分析其共同之处时,需要注意概括的规则需要有较大的概括性。以上的几个小节详细描写了各类交叉依存现象的句法语义条件。比较上面几类交叉依存类的现象,我们可以发现如下的几点的相同之处:

第一,这几类依存边交叉现象的发生基本都是与谓词紧密相关的。形成交叉现象的两组具有句法依存关系的词对中,至少一组词对中有一个是与谓词(动词、形容词)有直接的依存关系的。

第二,语序上,除第九种框式副词情况较特殊外,其余现象中的谓词有紧靠在一起的趋向,居于中心的位置。介词与其引介的宾语则紧靠谓词。谓词所支配的名词以及名词支配的相关成分则处在整个谓词部分的两边,而副词则处在两个部分之间;如下图所示:

N ₁ N ₂		V	Q
N ₁ Adj		V	N ₂
N ₁ X N ₂		V	
N	Adv	V	M
N		P ₁ V P ₂	
N ₁ 把 N ₂		V ₁ V ₂	
N	Adv	V ₁ V ₂	
N		V ₁ V ₂	M
N		Adv ₁ V ₁ Adv V ₂	

图 2 交叉依存现象的语序分布

第三,两个谓词类的交叉现象中,一些第二谓词的被支配成分在线性位置上前于了第一谓词,且后于第一谓词支配的支配成分;或者是第一谓词的被支配成分的线性位置后于了第二谓词;前于或

后于的成分在位置上参照同构的第一、二谓词的词序。

第四，从语义方面来说，即是两个词语 W_1 与 W_3 在语义特征上具有组合性，词语 W_2 与 W_4 在语义特征上也具有组合性，但其组合关系并不因词语在线性距离上的隔断而中断。

这些句法和语义上的特点，为我们的研究提供了初步的思考方向。

首先，应该区分开语序和句法关系两个不同的层面。交叉依存现象的形成是跟两个方面都有关系的。句法依存的形成不仅跟词性有关，而且是有语义上的基础的，语义上的关联可以不因为语序上的隔断而中断。但是，现有依存关系的描述还是限于词类范畴之间，在句法分析中其限定能力是有限的，对于汉语来说需要更细化的规则。

其次，没有句法依存关系的词语之间的语序也是有一定的原则制约的。这个原则应该是交叉依存形成的重要影响因素。使得具有句法依存关系的词语在语序上处于分离的状态。

再者，我们应考虑到所有可能的句法依存关系，不应该受限于独根节点、单纯结点、单一父节点的要求。从信息处理的角度来看，现有的依存处理系统大多是假设依存是只发生在投射性的情况之下的，因而在实际的分析过程中会把非投射性的情况错误处理。这样做的直接结果就是导致了从句法到语义映射时候的偏差与错误。我们知道，句法驱动的语义分析基本是按照组成性原则来进行的，语义组成部分应该与句法的组成部分匹配。如果不能够分析出相应的句法树，则语义的正确处理就没有可能。系统性的语义可以作为句法层面使用的标准。

属于非投射性现象中一个子类的交叉依存现象，对其种类的考察与句法语义条件的研究，将对实际的应用系统提供直接的理论支持。能对处理结果的优化具有直接的促进作用。

4 句法依存的属性和谐原则

在前一章，我们论述了交叉依存的形成原因可以有两个观察的角度。第一，一组词语有共现的要求，因而有句法依存关系的存在；第二，有共现关系的两个词语在语序上处于分离的状态。相应地，我们提出两个原则来解释交叉依存现象形成的原因。第一，词语间句法依存关系在组合成立上的属性和谐原则；第二，词语在语序方面的环形语序域原则。在第一章，我们重点解释词语之间句法依存成立的属性和谐原则。

4.1 词语的属性

4.1.1 汉语词类间依存关系的特点

词语之间的句法依存关系，通常是从不同词类之间的同现关系来探讨的。典型的依存关系如介词与名词的共现，形容词与名词的共现等。一旦谈到句法的依存，一般都用这几类词语来作为说明的基础。这两类词语之间有强制共现的关系，因此说明两个词类范畴之间具有依存关系。在屈折语中，由于词语的形态变化比较丰富，词类与功能之间的对应关系往往比较固定，使得词类之间的依存关系也相对比较固定。较容易形成唯一的句法依存链，而且可以较容易地转化为语义关系。但是在汉语这样的孤立语中，依存关系的判定却是一件非常麻烦的事情，其词类与句法功能之间并没有像屈折语言那样整齐的对应性。一个词类可能对应很多的功能位置。最广泛的如名词、形容词、动词，可以对应主语、谓语、宾语、补语、定语、状语各种句法成分对应，可以跟很多的词类形成依存关系。

因此，汉语中的依存关系的可能情况将是个极大的数目。

我们以教材中比较通用的十三类的分类（胡裕树 1995），把汉语中的词分为名词、动词、形容词、数词、量词、代词、副词、介词、连词、助词、叹词、象声词、语气词共十三类。可以通过图表的形式来表示他们之间可能有的依存关系。

词类	名	动	形	数	量	代	副	介	连	助	叹	象声	语气
名	+	+	+	+	+	+	+	+	+	+	-	-	+
动	+	+	+	+	+	+	+	+	+	+	-	-	+
形	+	+	+	+	-	+	+	+	+	+	-	-	+
数	+	+	+	-	+	+	-	-	+	-	-	-	+
量	+	+	-	+	-	+	-	-	-	-	-	-	-
代	+	+	+	+	+	-	-	+	+	+	-	-	+
副	+	+	+	+	-	-	+	-	+	-	-	-	-
介	+	+	+	-	-	+	-	-	-	-	-	-	-
连	+	+	+	+	+	+	-	-	+	-	-	-	+
助	+	+	+	-	-	+	-	-	-	-	-	-	-
叹	-	-	-	-	-	-	-	-	-	-	-	-	-
象声	-	-	-	-	-	-	-	-	-	-	-	-	-
语气	+	+	+	+	-	+	-	-	+	-	-	-	-

表 1 词类范畴依存关系表

可以看出，相较于类似英语的屈折语，汉语中词汇范畴之间的依存关系显得要复杂得多。同一种词类的词语，可以与之形成依存关系的其他词类会比英语多很多。例如，在英语中，介词只能与名词、动词具有依存关系。在汉语中，介词则可以跟名词、动词、形容词、代词形成依存关系。在英语中，形容词主要跟名词发生依存关系，也可以与系动词发生句法依存。在汉语中，形容词则可以跟

介词、副词、代词等更多的词类形成依存关系。即朱德熙（1985）谈到的“形容词能够跟主语、谓语、宾语、补语、定语、状语各种句法成分对应，功能最为广泛”。名词也几乎可以跟所有其他的词类形成依存关系，这即是吕叔湘（1978）谈到的“与谓语动词有关的名词的身份难定”。这种复杂的句法依存可能性，使得在对汉语进行依存分析时具有了很大的困难，很难形成唯一的句法依存链。同时，汉语中一个词形在不同的义项上可以具有几种不同的词性。在面向生语料的信息处理中，词语词类的判断本身即是一个重要的难题。泛化的句法依存再加上无定的词类，更是使得在句法分析时难形成唯一的句法依存链。而且，在计算机应用处理中也将会带来巨大的计算量。因此，人们不得不求助于更加细化的语义特征或句法特征来增加规则的限定性，以区别表面上相同的句法关系。

与此相类似，一直以来作为语法主流的短语结构语法也被发现其规则的约束力不够，单一的范畴标记不能够充分消解短语结构类型和短语边界的歧义问题（黄昌宁 张小凤，2002）。因而也求助于复杂特征集的限制或从词汇主义角度重建语法体系。于是，短语结构语法和作为词汇主义代表的依存语法有了合流的趋向。这即是现在中文信息处理领域普遍适用“大规则、小词库”处理思路的根源。而且，复杂特征的描写方法可以实现语言知识递增式的开发和积累，能逐步提高分析的精度而不必对语法体系做大的改动。因而具有较强的适应性。在诸多依存句法分析中，基于约束的（constraint-based）依存分析方法已经成为了其中一个重要的分析流派，而且取得了不错的分析结果。其代表如 CDG（Constraint Dependency Grammar）（Maruyama 1990）和 WCDG（Weighted Constraint Dependency Grammar）（Schröder 2002）。

4.1.2 依存描写的细化

如果只把句法依存限定于宏观的词类之间的关系上，那么要想取得合法的唯一依存链，是非常困难的。必须依靠对具体词语的语法或语义特征所作的详细的描写。但是，目前无论是句法特征上抑或是语义特征上的描写标准从组合的角度来看都是缺乏系统性的。首先，语法特征的描写多是基于组块语法体系的。由于其没有与语序分离开，因而并不宜用来作为基于词汇语法的语法特征描写标准。第二，在语义特征的描写时对语义特征的选取缺乏系统性。至今没有一份明确的语义特征表作为描写语义的标准。Pusterjovsky (1991) 更是认为任何假设的定量的元素都不能完全表示自然语言。而且，对语义特征之间的边界和层级架构也存在诸多的争议。泛化的词类依存可能，加上无定的词汇特征描写，必然使得依存关系的判断变得非常困难。要解决这个问题，必须能够在两者之间构建一个通道。或者能够细化词类之间的依存关系，或者可以减少词汇特征描写的无定性。

由于汉语中不存在明显的词类标记，从形式上细分词类是不可行的。我们决定使用第二种解决问题的思路，即减少词汇特征描写的无定性，把系统性的语义特征作为句法层面的标准来使用。我们知道，在屈折语中有独立的形态层面表达数、性、格的概念，在汉语中则不完善。汉语词类形态的缺乏，使得在词类的判定上必须要依靠语义信息和分布特征。因此可以说，在汉语中作为句法的词类和语义之间并没有明显的界限。我们可以用一种较为朴素的理念来看待汉语中的句法依存现象。我们认为，句法依存关系是以语义上的联系为基础的（功能词联系的情况除外），句法关系是语义关系的间接反映，或者说是语义联系的泛化。也即认知功能语言学者所认为的。语义在很大程度上对语法起决定作用。语义的联系则又是以百科知识的联系为基础的。印欧语言形式上的屈折变化，其根本

依然是语义的范畴。只是随着使用时间和频率的增多而逐渐语法化。其作为语法判断标准的根本原因也在于其系统性，可以观照一定范围内所有的词语使用现象。因此，我们建议对汉语来说要 ①把词语所具有的系统性的语义特征归纳为区别性的句法特征，这些句法特征可以来作为区分依存关系的基础。并且，②深入到词语的义项内部来探讨语义特征。

这两点对于汉语来说不仅是可行的，而且是非常重要的。

第一，可以用来作为句法依存判断标准的区别特征必须是系统性的语义特征，而不是以往语义特征分析法中较为随意的语义特征分析，特别是仅采用几组词语对比的方法比较而来的语义特征。

语义特征分析又叫做义素分析法，通过比较得到语义特征的分析方法对于分化歧义句式应该还是比较有效的，但是也有几个重要的缺点是无法克服的。首先，语义特征是通过一定的句法格式得到的。换句话说即是只对分化特定的句式有用。这在汉语教学上可以说是功不可没，但是对信息处理来说不一定适用。因为至今为止，并没有一个明确的穷尽的语义特征集作为特征标注的标准。而且，由于能够进入某种格式的词语是有限的，因而其概括的语义特征不能够保证系统的封闭性，与由其他格式概括出的语义特征的关系在区分与划界问题上就会存在困难，因而并不是一个完善的理论体系。义素分析方法中虽然也有一条重要的原则即系统性原则，“义素分析必须在一定的语义系统中进行，……不同语义系统中的词语是无法相互比较的。”但是其所指的系统范围是很小的。“父亲”和“老师”因属于不同的系统而无法进行比较。因此，可以说“义素分析尚未形成客观的标准，有较大的随意性（邢福义 吴振国，2002:125）。”当然，现有的一些语义特征分析的研究成果也是我们必须借鉴和继承的。

具体的语义特征与系统性的语义特征两者之间的区别有时候不太明显，不易被人察觉。但是，仔细区分两者却是非常必要的。

系统性语义特征由于其规整性，可以作为宽泛意义上的句法依据。而具体的语义特征则需要一个词一个词地去分析和掌握。

我们可以以计算科学上对颜色的定义来类比，解释说明具体性语义特征与系统性语义特征的差别。我们知道颜色可以分为黑色、白色、黄色、绿色、紫色等若干种。这些具体的“黑、白、黄、绿”等都可以称之为具体的取值。但同时，任何颜色又都可以通过对比度、亮度、灰度等一组参数来定义，在计算科学上称之为属性。颜色的不同体现为各个属性参数的不同取值，相较于“黑、白、黄、绿”等颜色的离散取值，这些对比度、亮度、灰度的特征即为系统性的特征，可以用来描写所有的颜色。其与具体颜色名词的差异即类似于我们所说的具体性语义特征与系统性语义特征的区别。在语言现实中，我们也可以做类似的分析。例如以下几个动词：“抓、踢、顶”，在语义上都属于身体部位的动作，分别与手、足、头有着直接或间接的关系。在对组合词语的选择上分别要求具有“手、足、头”的语义特征。但是，这些语义特征彼此之间是无明显关联的，需要针对具体的词语一个一个去描写和掌握，因此是具体的语义特征。但是作为行为动词，其在过程属性上则是具有系统性的。其义项的分化也与具体过程取值有紧密的关系。我们可以把这些系统性的属性用作句法层面的标准，作为词语组合时限定选择时候的重要依据。系统性语义特征与个别性语义特征对于解释组合成立上作用的大小是有一定差别的。系统性的特征告诉我们组合的时候不能是什么，个别的语义特征则告诉我们必须是什么。个别性语义特征是建立在系统性语义特征基础之上的。

这样的处理思路与胡明扬（1999）谈到的“语义语法范畴”概念比较类似。胡明扬认为，“所有的语法范畴说到底都是语义语法范畴”。“汉语正是一种非形态语言，显性语法形式不发达，不少语法范畴是语义语法范畴，没有相应的显性语法形式，只有相应的隐性语法形式。”对于这个观点我们是基本赞同的。就像“数、性、

格”等语法范畴在实质上都是语义范畴且具有系统性一样，从语义的角度入手探讨组合的可能性和分布形式的语义语法范畴的方法是可行的。在此基础上，本文强调了区别具体语义特征和系统性语义特征的必要性，主张合理构建范语法畴内部的关系以避免范畴设立上的任意性。

第二，特征的分析必须深入到词语每个义项的内部。以往的句法依存关系的研究大致只是在宏观的词类之间谈句法的依存，对于汉语来说这是远远不够的。因为汉语中词类与功能之间并没有对应关系，而且，一个词的不同义项之间也会有很多属性上的差异。目前，在义项层面上的句法描写已得到较为广泛的响应。如配价描写深入到义项层面，对于各种应用来说是成效显著的。所以，有必要深入到具体的义项内部来探讨语义上的区别特征。

以往的语义特征研究，多是针对某个词类泛泛而谈的，并没有深入到义项的内部。例如马庆株（1988）对自主、非自主动词的研究。当然，已有的这些研究是需要辩证地去看待的，既是我们现有研究的学习基础，也蕴含着我们研究中需要改进的方向。我们知道，词语的义项之间的差异可以是很大的，无论是在语义上或者是在句法表现上。而词语的使用又是以义项为基本单位的，过于宏观的语义特征描写对于词语之间组合的限定能力是很弱的。因此，我们提议必须深入到义项的内部来发现系统的语义特征。同时，鉴于一些词在词类上存在争议，或者说存在兼类的情况，只强调其某一种词类上的特征是不全面的。我们可以在兼有两种或两种以上词类的词语上加标各个体系的语法和语义特征，以便于组合关系的判断。

另外，在以往的研究中，有两个术语“属性”和“特征”与本文的研究紧密相关。在这两个术语的使用上，以往并没有一个清晰的界限。一般来说，在计算语言学领域，受计算科学上的术语使用的影响，学者们对属性（attribute）和特征（feature）都有广泛的使用，相应地有属性值、特征值的用法，但是并没有做特意的区分。

在理论语言学领域，“特征”这个术语使用的会多一些，像语义特征、句法特征等。“属性”则多用在对属性名词的研究中。属性概念一般的使用是泛指的，事物所具有的任何特点都可称之为属性。袁毓林（1994）把属性名词和亲属名词、部件名词作为并列的三种一价名词。刘春卉（2008）定义属性为“大致来说，属性是人们在认识世界和改造世界的过程中感知或认识到的主客观对象的各种可供认知或比较的方面”。她从《现代汉语实词搭配词典》中选取的415个属性名词，也大多是具体的特征。这些具体的特征很少能覆盖所有的名词类。属性值的术语在用法上也比较模糊，无论是零散性的取值或系统性的取值一般都称之为属性值。

为避免在术语使用上引起的争议和便于与以往的研究做区分，我们在术语使用上把现有的研究与以往的区分开。我们给两者做明确的区分。系统性的语义特征我们称之为“属性”，零散的语义特征则称之为“特征”。而系统性的属性取值我们称之为“属性值”，即义项内的系统性区别特征。《现代汉语词典（第五版）》定义的属性概念为“是事物所具有的性质、特点。”这里我们参照词典对“属性”的定义本身来重新定义我们所指的“属性”和“属性值”两个术语及其两者之间的关系。

属性是事物所具有的系统性的性质和特点。

属性值则是事物属性上的系统性取值。

“属性”与“属性名词”是两个不同的术语。“属性名词”是指作为认知对象的“属性”在语言中的自然称代形式。但并不是所有的属性都在实际语言中有明确的自然形式，也可以是人新造的指称。而且，更为明显的是“属性名词”所表示的特点是比较零散的，不同于本文所指的系统性的性质和特点。

以行为动词为例，“过程”是这类动词的一个明显的属性，而“开

始、持续、结束、结束后”则分别为过程属性上的四个系统性取值，我们称之为“属性值”。具体的时间点如“五点、昨天”等则是零散性的取值。词语之间组合关系的形成，必须在相应的属性取值上具有和谐性，即不能够发生相悖的情况，组合关系才可能成立。这个原则我们称之为组合成功的“属性和谐原则”。

另外，属性的取值与义项的划分也有紧密的联系。义项的分裂往往与属性取值的分化有关。把系统性的属性看作句法层面的标准，必然会涉及到使用中的标注问题。因此也是本文一个需要明确的问题。我们的研究也可以看作是对明确义项划分的标准上做的一种尝试，属性取值的统一性如果能够与义项的划分对应起来，这种义项划分的方式就可以为机读辞典构建提供的一个解决思路。

我们知道，义项划分的标准历来是引起大家争议的一个难题。一般来讲，义项划分的方法有很多种，历史分析法、逻辑分析法、经验分析法、功能分析法和搭配分析法是几种基本的方法。但是现在的词典编纂或者是在对义项划分的标准上不够明确的，缺乏操作性和整体性。或者是把通过不同标准分化得到的义项排列在一起。《现代汉语词典》的划分标准即是综合使用了这许多种的方法。换一种说法则是在划分标准上缺乏系统性。划分标准的无定必然导致在义项描写上的无定。对于人的学习来说可能是无碍的，但是对于面向机器的词典构建来说，则是必须首先要回避的问题。因为必须先有统一明确的义项划分原则，才可能在不同的词语组合时有合理的判定依据。在属性组合思路下的义项分化，不涉及词义的历时演变的情况，只在共时层面来看义项分立的可能，重在其系统性与可操作性。

4.1.3 属性的选取

属性的划分从理论上来说可以是无穷多的，本章只选取了有限

数目的属性来描写。所列举的属性是最基本、最重要的，因此可以允许后续的细化描写，具有可扩展性。而且这些属性也是大家所共知的，是较为确定的属性。足够以解释本文所涉及的交叉依存现象。

因此，本文所阐述的属性与属性值具有一定的举例性质。主要是为了帮助说明本文所要阐述的属性和谐的原则，以解释交叉依存中词语之间依存关系不因词序的隔断而中断的现象。至于具体属性之间宏观框架的安排，则可以根据实际需要的不同而做出不同的选择。

4.2 动词意义的属性

属性和谐是指词语组合时词语之间的相互选择在某一个属性取值上的无悖性。在前文论述交叉依存形成的原因时候我们曾经指出：交叉依存的形成是由于在具有句法依存关系的两个词语之间加入了具有与两者之外具有句法依存关系的词语，但是依存关系并不因为线性位置上的隔断而中断。这里所指的句法依存关系上的不中断即为词语组合时的属性和谐性。在这一节，我们将主要通过动词与其他词类之间在组合时相互选择的属性和谐的现象来说明动词的属性。

根据 Mel'čuk (1988) 的定义可知，句法的依存是指词语之间的线性相关。通常是指一类词与某类词之间的强制共现性，可以看做是抽象化了的语义关联。英国学者 A. Shaikeritch 也认为，词语如果在上下文中经常共现，则意义之间必定有联系。（转引自：王惠，2004）本文谈到的几类交叉依存的形成，也都是因为语义组合关系选择的唯一性使得句法上的依存关系不因位置的隔断而中断。从认知的角度来说，组合关系的成立是建立在人们对不同的词类分量一致性的先验理解基础上的。组合关系可以不因序列上的隔断而中断则是由于认知上的语义启动（semantic priming），是由于习得

过程中无数次的重复而习得的条件反射反映。

词语组合研究不仅是语法研究的重要内容，也是词汇语义学近期以来研究的重点。鉴于动词在整个句子中的核心作用，语法范畴内词类细分的研究一般是从动词开始的。在词汇语义范畴内，以王惠（2004）等为代表的学者研究名词义位的区分时也参照了名词与动词的组合表现。

动名组合的研究往往涉及到动词的分类问题，即动词的子类研究。动词的细分可以解释很多组合成立与否的原因，因而是非常有用的。动词内部的分类可以有两个标准，一个是意义的标准，一个是形式的标准。早期的动词分类研究多采用静态的意义标准，如《马氏文通》、《新著国语文法》。一般认为，静态的分类由于脱离了组合规律，无法解释词语组合的本质。后来动词分类的研究越来越注重动词和有关成分，特别是名词成分的关系。

目前的动词分类结果大致有以下几种：及物不及物动词、自主非自主动词、持续非持续动词，也有动作动词与状态动词的划分，按照不同配价的划分等。这些分类分别体现了意义和形式两个层面的标准。形式层面的标准，如及物不及物的划分；意义层面的标准，如自主不自主、持续非持续的划分。也有兼有两个层面的，如对动词的配价划分等。

这些分类体系当然都有其存在的合理之处，它们确实是立足于语法事实的，虽然比较粗疏，却能够合理地解释部分相关的语法现象。有的标准大约是可以推广到所有动词的类别的，如形式上的及物不及物¹⁷。但是有些分类标准只是针对一定范围的动词的，如果强行推广到所有的动词分类，则不具有可行性。如持续非持续的划分只是过程属性上的两个取值的不同，自主非自主的划分对于表存

¹⁷ 形式上及物不及物的分类也是有争议的。吕叔湘（1979，P34）认为“动词分成及物（外动，他动）和不及物（内动，自动），是很有用的分类，可也是个界限不清的分类。……问题在于‘宾语’的范围：是不是动词后面的名词都是宾‘语’？要是这样，汉语里动词，就真的像有些语法学者所说，很少是不及物的了。”

现的动词如“有”则意义不大。这些小类的提出，并没有顾及动词分类系统的完善。

动词的再分类因为牵涉到了语义，在动词的搭配问题上曾存在过是否属于语法分析的争论。我们认为，语义也是有系统性与非系统性之分的。即使在形式层面划分出的及物不及物子类，也是以其语义上的差别为基础的。可以把系统性的语义作为句法来使用。

我们可以把对动词划分次类的系统性语义依据看做是动词某方面的“属性”。通过对动词组合词语的观察来对动词的不同属性做出不同的分类。原有的及物不及物、自主非自主、持续非持续的划分，都可作为是对动词属性考察的参考角度。属性是动词分类的依据，但不是分类的唯一标准。对动词属性的考察西方语言学者研究的比较早，也比较深入。如 Vendler (1957) 根据时间的属性把英语动词分为了 states, activities, achievements 和 accomplishments 四大类，成为长期以来英语动词分类的参照标准。

对于这些有效的区分特征，我们给予借鉴并应用到我们的属性描写之中。我们相信，动词最初的意义是比较单一的，但是在使用过程中由于在一个属性的某个属性值上使用频率很高，则慢慢分化出了新的义项。这样，在不同的义项上则又分别有了原来动词本义的那些属性，并且有可能继续在各个分量上进行分化。因此，我们提出义项属性描写的概念。

而且，我们不仅要在整个词类层面要细分，在义项内部也要进行属性的描写。这样在判断组合成立的时候就具有了较高的准确性。而且，这些属性之间关系的架构也是一个非常关键的问题。虽然对于动词义项分立的标准与结果一直以来都存在不少的争议，但是如果搁置这些争议，仅仅从应用的实际来看，分义项标注对于提高一些应用系统如问答系统、机器翻译、信息抽取的效率都是非常直接的。所以，从语言信息处理的角度出发来看，对动词分义项的属性描写是非常必要的。其相对应的其他词类如果也描写完善，对于交

又依存类现象的判断将更加准确。

至于词语义项内部不同属性的考察，近期以来引起汉语语言学界学者的普遍重视。但是在属性标准等方面还没有一个明确的标准。对词语不同属性的考察，可以从动词入手，也可从名词入手。许多学者做了若干不同的尝试。Vendler (1957) 根据时间特征把英语动词分为了 *states*, *activities*, *achievements* 和 *accomplishments* 四类。这里的时间特征可以看做是动词的属性。黄居仁 (2000) 把汉语动词分为 13 种事件类型。这些都是一些重要的尝试，但也有不足之处。比如，其分类是针对所有动词的，其分类标准不一定能涵盖义项内的划分。黄的分类中把很多的动宾短语做为词来分析，也是值得商榷的。一些学者也尝试研究名词属性 (胡慊 2007)，希望通过名词的属性来考察与其组合的动词属性，但其可行性须待验证。我们提议需要从语义的角度出发，结合其组合的词语，探讨动词义项内部属性的描写。参照现有的研究成果，建立一个面向机器处理的句法语义属性描写体系。而且，动词属性的讨论可以引起相应的名词副词等属性描写体系的建立。这些属性的研究反过来又会促进动词属性研究的深入。

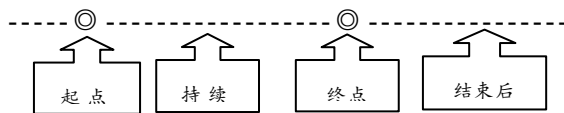
一般认为，世界是由物质与运动组成的。物质离不开运动，运动也总是物质的运动。而时间与空间则是物质运动的两个基本参照量，我们就首先通过这两个基本参照量来看义项的分化及其组合的要求。同时，人们在表述某个事件的时候总是带有一定的情感倾向，因而情感也可以看做是一个基本的属性。

结合本文研究的交叉依存现象，我们初步考察了动词的三个“属性”：**过程、位置、情感**。在“过程”属性上有四个取值：**起点、持续、终点、结束后**；位置属性上有四个取值：**起点位置、经过位置、目的位置和发生位置**。情感属性上具有三个取值：**褒义、中性和贬义**。

4.2.1 动词意义的过程属性

严格意义上来说，动词代表的所有运动都与时间有关，表现为具有过程结构。因为任何运动都不能够脱离时间和空间的范畴。Vendler(1957)对动词做分类的标准，其实是一个比较笼统的分类。其分类的结果是针对英语的，与汉语的实际不尽相符。尽管可以不完全同意其分类的结果，但其观察的角度却值得借鉴。郭锐(1993)则称之为过程结构。但其只分为了起点、终点和续段三个要素。我们认为，每个动词在过程上都是具有起点、持续、终点和结束后四个取值的。这些过程取值的不同都是伴随动词的不同义项的。只是可能某些词语的某些义项在具体取值上使用频率差异比较大，有些取值不常用而已。而且，词语义项的分化也与过程属性的不同取值有关。借鉴 Vendler (1957) 对动词分类的思路，并且深入到义项的层面来看待属性值的分化。我们认为：动词内部不同的义项应该是由基本义逐渐分化出来的，其分化的依据往往是因为其在某个属性值上使用的组合对象频率较高，我们就可以把它分立为单独的义项。当其分离为单独的义项之后，其在过程属性上又可以分离出不同的取值，形成不同的组合关系。

以最具代表性的行为动词为例，一个行为动作在过程轴上可以分为四个基本取值：起点、持续、终点、结束后。如下图所示：



以基本的行为动词“走”为例，可以划分出“走”的开始、持续、终点和终止后四个不同取值。如果动词的基本义在某个属性值上使用的频率很高，则在语义上会出现明显的分化趋向，并且逐渐固定下来。在词典中则会表示为分列的义项。目前，义项的分立在

不同的词典中使用的是不同的标准。总体上来说其划分的界限由于多重标准的使用而缺乏系统性。通过属性取值的概念，我们也在尝试提供一种较为系统性的义项分立标准。参照“走”的释义，可以考察其义项分化与属性取值的关系。

走

- ① 𢇛 人或鸟兽的脚交互向前移动：行～ | ～路 | 孩子会～了 | 马不～了。
- ② 𢇛 <书>跑：奔～。
- ③ 𢇛 (车、船等)运行；移动；挪动：钟不～了 | 这条船一个钟头能～三十里 | 你这步棋～坏了。
- ④ 𢇛 趋向；呈现某种趋势：～红 | ～热。
- ⑤ 𢇛 离开；去：车刚～ | 我明天要～了 | 请你～一趟吧 | 把箱子抬～。
- ⑥ 𢇛 指人死(婉辞)：她还这么年轻就～了。
- ⑦ 𢇛 (亲友之间)来往：～娘家 | ～亲戚 | 他们两家～得很近。
- ⑧ 𢇛 通过：咱们～这个门出去吧。
- ⑨ 𢇛 漏出；泄漏：～气 | ～风 | 说～了嘴。
- ⑩ 𢇛 改变或失去原样：～样 | ～调儿 | ～味儿 | 你把原意讲～了。
- ⑪ (Zǒu) 𢇛 姓。

除去作为名词的义项⑪和因历史原因产生的义项②，其他的各动词义项都可以看做是从基本义分化而来的。分化的不同是过程属性上不同的取值。“走”在结束后的取值上分化出来的义项有：

- ④ 𢇛 趋向；呈现某种趋势：～红 | ～热。
- ⑤ 𢇛 离开；去：车刚～ | 我明天要～了 | 请你～一趟吧 | 把箱子抬～。
- ⑥ 𢇛 指人死(婉辞)：她还这么年轻就～了。

⑧ 通过：咱们～这个门出去吧。

⑨ 漏出；泄漏：～气 | ～风 | 说～了嘴。

⑩ 改变或失去原样：～样 | ～调儿 | ～味儿 | 你把原意讲～了。

在开始取值上分化出来的义项有：

③ (车、船等) 运行；移动；挪动：钟不～了 | 这条船一个钟头能～三十里 | 你这步棋～坏了。

⑦ (亲友之间) 来往：～娘家 | ～亲戚 | 他们两家～得很近

属性取值的不同是纯粹的语义差别，是母语者能够感受到的不同。在分化为不同的义项后又可以在此义项上带有过程属性的四个不同要素。需要注意的是，系统性取值基础上的义项分化总是跟某些方面的具体语义特征有紧密关系的。例如，“走”在结束后取值上分化出来的“指人死”义项，总是跟“人”的语义特征相关。但是，具体语义特征是非系统性的，虽然也是义项分化的重要参照，但由于不具有系统性而不能作为分化的基础。其分化的系统性特征基础是时间属性上的不同纬度。

词语的某一个义项在与不同的词语组合时，可以使用不同的取值。但不同取值上的使用频率会有很大的差异。如“走”在表示“行走”的基本义项上，经常使用的为“开始”与“持续”的取值。而其在“结束后”取值上使用已经分化为若干个独立的义项，因而其本义“行走”在“结束后”的取值上就很少使用了。

(70) 金融监管开始走向规范化和法制化轨道。 开始取值

(71) 他坚持走了三个小时才休息。 持续取值

在一个句子中，如果其搭配的成分能够在 ①同一个义项的不

同取值上或 ②不同义项的不同取值上同时成立，那么往往会产生理解上的歧义。

(72) 他走了三个小时。

这句话中的动词“走”可以同时“行走”义项的持续取值和已经分化为新义项的“离开”的结束后取值上都成立，所以会产生理解上的歧义。因此，上例可以同时有两种解释的可能：第一，他行走了三个小时；第二，他离开了三个小时。

属性和谐的要求在词语组合中是具有强制性的，如果在组合的时候属性值的使用不是出现上面谈到的两可的情况，而是出现了直接的冲突，则句子不能够成立。例如：

(73) 他走了三个小时。(✓)

(74) 他开始走了三个小时。(×)

动词“走”要跟表示时间段的“三个小时”组合成功，必须在义项上使用“持续”或“结束后”的取值。(73)可以满足这个要求，因而是成立的。但是第二个句子中，“开始”要求“走”使用的是“开始”的取值，而“三个小时”要求使用“持续”或“结束后”的取值。同一个义项被同时要求使用两个不同的取值而引起了冲突，从而使得这样的组合不能够成立。

有一些动词在不同取值的使用频率上差别较大，总是倾向于使用其中一个取值，但偶尔也会使用其他的取值。因其没有分列为新的义项，当其搭配的词语能够同时在两个取值上成活的时候，受话者理解时则会产生歧义。但是，一般来讲，受话者理解为使用频率较高的取值的概率要高得多。例如汉语中经常讨论的句子：山上架着炮。之所以产生歧义，是因为动作表持续的取值跟结束后的取

值都可以用“着”来修饰。但是，由于其在结束后取值上的使用更为常见，人们还是倾向于理解为结束后表示结果状态的意义，即“山上架设了炮”。

类似的例子可以列举很多，以观察动词持续取值和结束后取值的差异：

(74) 冰箱里冻着冰。

S₁ 冰箱里冻着冰，冻好了你拿出来。（持续）

S₂ 冰箱里冻着冰，你拿点出来。（结束后）

(75) 这座房子盖了五年了。

S₁ 这座房子盖了五年了，还没有盖好。（持续）

S₂ 这座房子盖了五年了，原有的住户基本都搬走了。（结束后）

(76) 客厅里铺着地板呢。

S₁ 客厅里铺着地板呢，你要看紧这些工人。（持续）

S₂ 客厅里铺着地板呢，你要脱鞋才能进去。（结束后）

动词在起点和终点的分量上也可以引起理解的歧义，如果从其搭配的论元来看则是施事、受事与与事的不确定。

(77) 分了好几次房子，钟先生都不满意。

A₁ “钟先生”是“分”的施事。（起点）

A₂ “钟先生”是“分”的受益者。（终点）

英语中经常讨论到一些副词有做情态副词（manner adverb）和说话者朝向的副词（speaker-oriented adverb）两种用法的现象，其实是把动词在不同取值使用上引起的歧义强加于了副词。例如：

Lisa rudely departed. McConnell-Ginet (1982) 指出这个句子有两种意义：① It is rude of Lisa to depart. ② She departed in a rude manner. 在我们看来，这个句子之所以会产生两种意义，是因为动词可以同时在这两种不同的取值上成立。一个是开始的取值，一个是结束后的取值。在开始的属性值上使用，则副词 rudely 修饰的是“depart”行为的不礼貌。而在结束后的取值上，则是指“Lisa depart”这件事情是无礼貌的。

本文谈到的过程属性及其开始、持续、终点、结束后的划分，是从事件结构的角度来看的，有异于语法上所谈到的“时”的概念。语法上讲的“时”是指事件发生的时刻相对于说话者本身所处的时刻为参照系的相对关系，分为“过去时，现在时和将来时”三种。而事件内的时间属性则往往跟语法中谈到的“体”的概念糅合在一起表现出来。我们知道，体可以分为“进行体、完成体”两种，时体的结合更是增加了分量判别的难度。体是就整个事件来说的，属性则是就动词来说的，两者属于不同的层面。仅仅是一个完成体就可以容纳事件的过程属性上的四个取值。比如，副词“已经”组合的动词一般使用的是“现在时、完成体”，在事件内的属性上则四种分量都可能使用。

- | | | |
|------------------|-------|-------|
| (78) 他已经开始走了。 | <行走义> | (开始) |
| (79) 他已经走了三个小时了。 | <行走义> | (持续) |
| (80) 他们已经分了房子。 | <分配义> | (终点) |
| (81) 他们已经走了。 | <离开义> | (结束后) |

交叉依存现象中，与动词的过程属性关系最直接的为两个谓词下的交叉依存现象的第二类。即副词与动词分离的情况。以“已经”为例来解释。“已经”的词典释义如下：

已经 圖 表示动作、变化完成或达到某种程度：任务～完成 | 他～来了 | 天～黑了，他们还没有收工 | 我今年～七十了。

“已经+V₁+V₂+了”类的交叉结构中，“V₁+V₂”表示的是一个复合事件，是两个事件的组合。以往称之为动补结构，并且指出 V₂ 表示的是动作的结果，因此一般称之为补语。从“已经”的释义来看，其修饰的词语必须表示动作变化的完成，所以一般要求与带有时体标记的谓词相组合。而“V₁+V₂+了”中的谓词 V₁ 却没有任何时体的标记，V₂ 使用的正是时间属性上的结束后取值。因此，副词“已经”只能与 V₂ 有句法上的强制依存关系。V₂ 的位置上的谓词也可以是形容词，形容词本身一般表示的是静止的状态，因此“形容词+了”一般表示的是结束后的状态，因此也符合组合关系成立的条件。

过程属性的表现一般总是要求动词带有时体的标记“着、了、过”，因此我们把不带有时体标记的动词只看做是一种抽象的行为能力（现在时的情况除外），把带有了时体标记的动词作为实际的动词来看待。相应的，在同一个句子中，带有了时体标记的动词所要求的与之配价的词语在对于本句的重要性上要高于抽象行为能力（即不带时体标记）的动词所带有的配价词语。这可以突出体现在动补结构中的动词组合上，也即是本文研究中的名词与动词分离类的交叉依存现象。例如，“人们把界线打破了。”第二个谓词“破”所带有的配价成分“界线”在重要性上要比“打”所带的配价成分中的“人们”高。我们甚至可以省略“打”所带有的配价成分“人们”，而直接说“界线打破了”。但是不能反过来省略“破”所带有的配价成分“界线”说是“人们打破了”。

4.2.2 动词意义的位置属性

位置也即是处所。位置的概念在汉语中一般用“处所”这个术

语来指称。动作的发生总是在一定的空间中的。在前一节提到，空间主要是指方向与位置两个方面。方向的属性在前一节已经讨论过，这一节主要讨论位置的属性。原有的术语“处所”在表达一些相对位置时使人感到不太能接受。例如说“我左边”中“左边”作为方向的表达是以“我”为参照点的。但如果把“我”称为“处所”则与人的直觉相违。因而我们建议采用相对中性的物理概念术语“位置”来指称。

“位置属性”与“位置的表达”是两个不同的概念，需要加以区分。“位置属性”是对动词来说的。任何动作都是发生于一定的位置的。一般的动词则都有发生的位置，即使是一些心理动词、存现动词也都有明显的位置属性。如“他在心里想了很久”“他在北京有个弟弟”的用例。而“位置的表达”则是与动词的位置属性相对应的表述部分，可以有多种呈现的方式。动作总是发生在一定的位置，位置与动作的表达和表述的主体都紧密相关。

位置的表达有两种方式。一种是作为动词的配价成分，即必有成分。另一种则是充当可有的成分。充当配价成分时的位置多为目的位置，一般不用介词直接表达位置的概念。如经常谈到的处所宾语。而充当可有成分时则各种情况都有。这种情况一般通过介词来引介。介词具有给动词赋元的功能。

发生位置是一切动词都具有的。而“起点位置、经过位置和目的位置”三种则是针对位移动词来说的。或者说，发生位置是个笼统的概念，而起点经过与目的则是具体的概念。我们可以把位置属性做两分的处理，一种是发生位置，一种是位移位置。位移位置可以分为三种，分别是起点、经过和目的。以往认为的原点位置类似于本文所指的发生位置。位移动词可以在发生位置和位移位置上同时成立。如“我在家里从客厅到卧室走了好多趟。”位移动词在位置上的三个分量大致对应于在时间属性上的“开始、持续、结束”三个分量。

位置与方向的组合作为一个整体又可以表达一个完整的位置概念，如“椅子上”“屋子里”。所以，空间的表达可以是一个叠加的概念，如“在椅子上”“从屋子里”等，也有人称之为框式结构（刘丹青 2003），不过其把方位词称为后置介词。

动作与时间空间都不可分，因此动词的意义都具有位置的属性。动词与表达位置的词语之间在位置属性上相一致，组合能够成立。我们可以把这种系统性的语义作为句法层面来使用。位置属性可有分为两种，分别是发生位置和位移位置，其中位移纬度又可以细分为起点、经过、目的三种纬度。示例如下：

	动词的位置属性			
	位移位置			发生位置(原点)
	起点位置	经过位置	目的位置	
有				+
喜欢				+
走	+	+	+	+

表 2 动词位置属性示例

本文论述的交叉依存现象中，与动词的位置属性相关的部分为框式介词类。动词的位置属性与两个引介位置的介词都有关系。一个表示起点位置，另一个则表示终点位置。如：

(82) 去年实现进出口总值达一千零九十八点二亿美元，占全

国进出口总值的

比重

由

上年的百分之三十七

提高

到

百分之三十九。

N

P₁

V

P₂

“由”表示起点位置，“到”表示目的位置，都与动词的位置

属性相组合成功。另外，两个介词之间在位置属性上也有和谐性，一个表示起点，一个表示终点，构成一个框式的结构。这将在后面介词的属性部分详细论述。

4.2.3 动词意义的情感属性

人们在表述某个事件的时候不可避免的会带有自己的情感判断。因此，情感也是行为动词具有的另外一个基本属性。一个词语的情感属性对将与之组合的词语的情感色彩具有限定和影响的作用。在实际的句子使用中，互相组合的词语之间在情感上应该相容，如果不相容，则会产生错误或者其他的语用效果。

我们把系统性的情感属性作为句法层面的标准来使用。一般认为，情感可以有三个取值：即褒义、中性和贬义。词语在组合时关系的成立可以用下表来表示：

动词	组 合 词 语		
	褒 义	中 性	贬 义
褒 义	√	√	×
中 性	√	√	√
贬 义	×	√	√

表 3 动词意义的感情属性上组合关系表

动词的情感属性，对于所有与之有依存关系的词语都有限定作用。因此，与本文所谈的所有交叉依存现象都是有紧密关系的。因为形成交叉的两条边中总有一个是与动词相关。受动词支配的词语与动词之间在属性上应该相容。而且，特别需要指出的是，在能够形成交叉依存现象的动词并列中，同现的动词之间在情感属性上也应该是一致的。例如：

(83)中国是个男性主导的社会,女人往往被视为附属的地位,对于这样的不公平,做为一个端正的男人更应该爱护、尊重女性。

动词“爱护”与“尊重”都属于褒义词,因此可以形成并列。而且,动词对与之组合的名词也要求在情感的属性上不相违。例句中受动词支配的名词“女性”是中性词,可以与褒义动词组合成立,但是如果换成是贬义的名词如“贱人、娘儿们”,则组合关系不能够成立。

动词在使用过程中如果经常与表示不同感情色彩的词语共现,就会分化出新的义项。即是词语之间情感属性的影响。一般是两个词语组合中的贬义词逐渐影响到中性词分化出贬义的义项。如下两例“勾连”与“活动”的义项之间的情感差异,都是因为同现的名词多为贬义而影响到动词分化出贬义的义项:

勾连 ① 勾结:暗中~|他们~在一起,干了不少坏事。② 牵涉;牵连:一人做事一人担,绝不~别人。

活动 ② ③ 为某种目的而行动:抗战时这一带常有游击队~。

④ ⑤ 指钻营、说情、行贿:他为逃避纳税四处~。

这种影响是相互的,动词支配的词语的情感属性可以影响到动词义项的分化,动词的义项也能够影响到其配价成分的情感变化并分化出新的义项。因此,在义项中注明属性及其取值对于组合关系的判断是非常有效的。

4.3 名词意义的属性

名词的语义特征可以说是很芜杂的。由于物体的广泛性，其语义特征的数量也是无穷尽的。在词语的组合中我们可能用到的语义特征也是难以预测的。一些近义词之间的区别往往极其细微，只能通过少数的几个语义特征才能把他们区别开。与分析动词的语义特征相类似，我们也不想去涉及无穷尽的具体语义特征，因为其对于限定范围内的教学语法研究可能比较有用，但由于系统性不足，只能一个一个地标注。我们在这里只考察系统性的语义特征，作为句法判断的统一性标准。

名词的基本属性有多少个，具体是什么，这是个比较难于回答的问题。如果从语义特征的角度来寻找，那么可以有若干种不同的答案。因为不同的名词从不同的角度可以找出若干种不同特征。但是如果要求这种特征能够覆盖整个名词的词类，那么可能的属性类应该是有限的。

4.3.1 名词意义的量属性

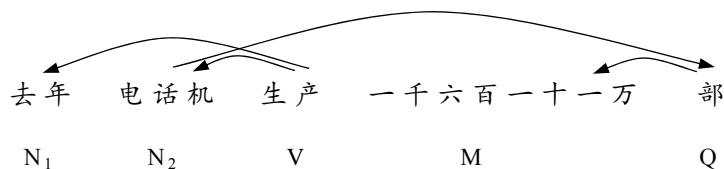
名词最容易理解和观察到的属性为量的属性。名词所指称的名物都是可以计量的。计量有两种方式，一种是自身数目的计量，一种是借助他物的计量。自身的计量一般称之为数，这样的名词一般称为可数名词。名词并不是都可以用数来计量的，所以才分为可数与不可数两种。因此，“数”的范畴并不能作为系统性的特征。而借助他物的计量则是可以覆盖所有的名词类的。自身计量与借助他物计量对于可数名词都是适用的，既可以说“一人”，也可以说“一屋子人”。不可数名词则只能通过借助他物的计量来表达量的属性。

以不可数名词“水”为例，可以借助外在的计量物来表达物质的量。如“一碗水”“两车水”，“碗”和“车”都是外在的计量工

具，这里临时用作了量词。这里的数词是不能够与名词直接组合的，必须借助量词的中介。也就是说，量词与名词之间是有强制共现关系的，因此两个词类之间具有句法上的依存关系。关系成立的基础是在量属性上的和谐，即量词与名词的量的属性上一致。量词的划分是很复杂的，可以有很多种次范畴的存在。从系统性特征的角度来看，我们把“量”作为一个属性是个比较妥当的层面，次范畴的细分则不能作为属性来看待了。虽然都属于量词，名量词与动量词是不同的。动量是针对动词的，名量词则是针对名物的。量词的术语是两者归并的类。

从词语共现的角度来看，是名量词决定了必然有名词的出现。但是在语义上，则是名词处在主导的地位，其对量词具体词语的选用有决定性的作用。由于名词充当的句法功能更加重要，一般把名词作为中心词。量词与名词的组合成功，取决于它们各自的语义特征。在宏观的语义上，凡是名词指称的事物，必然都有量的特征，因此可以看做是句法层面的标准。但是在微观的语义特征来看，不同的量词则是各有不同的具体特征。只有相应的名词也具有这样的语义特征，组合才能够成功。由于具体的语义特征差别很大，不具有统一性，因而需要在具体的词语中描写，不能上升到句法的层面。

本文讨论的交叉依存类现象中，名词与量词分离类的现象，即是对量属性上和谐要求的例证。“去年电话机生产一千六百一十一万部”其依存关系的表示如下图所示。



名量词“部”要求必须出现名词作为其依存的成分，其与名词“电话机”的量属性上组合能够成立，其依存关系并不因位置的隔

断而隔断。而“部”与名词“去年”的组合则不能成立，需要借助具体的语义特征消歧。句子中各个词语的语序位置是受环形语序域的制约，尽管量词“部”与名词“电话机”处于动词的两侧，但是是同一个域内。这将在下一章环形语序域部分将论述到的。

4.3.2 名词意义的情感属性

名词往往是充当动词的配价成分，反过来也就是说动词往往是要有名词性成分与其共现的。因而动词与名词之间有词类之间的依存关系。在论述动词属性时我们曾经谈到，动词起码有两个属性是与名词有紧密的联系的。一个是动词的情感属性，一个是动词的配价属性。因此可以推论，名词也有一个基本的情感属性与动词的情感属性相观照。情感属性是动词与名词都具有的一个基本属性。相应地，名词也具有“褒义、中性、贬义”三个分量的划分。在与动词组合时，组合关系的成立要求名词在情感的属性上与动词的属性不相斥。

名词的情感属性不仅与支配它的动词的情感属性相对应，而且还会限制其支配的形容词的情感属性选择。在由名词并列形成的交叉依存现象中，由于成分的共享，并列的名词之间在情感属性上也要求是一致的。例如：

（84）身为中华民族一分子，一个生活在香港的中国人，我谨代表所有香港同胞，向所有为此作出贡献的中华儿女，献上深深的敬意和感激。

动词“献上”是褒义词，要求其支配的名词在情感属性上可以为褒义或中性，但不能是贬义词。由于名词“敬意”和“感激”之间是并列的关系，这个要求分别适用于两个名词。它们可以都是褒

义或者中性，但不能都是贬义词。

另外，名词的情感属性跟形容词之间也有和谐的要求。这类现象在所有句子中都是广泛存在的。例如：

(85) 不过，在利物浦队队员中，参加下流活动的始终只有加拉切尔一人。

形容词“下流”和名词“活动”一个是贬义，一个是中性，可以组合成立，但是如果一个为褒义，一个为贬义的时候，组合关系就不能够成立。

情感属性上不同的取值可以影响到义项的划分。以往的研究中一般没有注意到情感作为义项划分的一个明确的标准，往往把情感的分别放在语用的层面来谈，我们认为，情感的不同取值对义项的分化上具有很大的参照意义，应该明确为义项划分的标准。

下面是在《现代汉语词典》中的两个例子：

【小人】 ①古代指地位低的人，后来地位低的人也用于自称。②指人格卑鄙的人：～得志|势利～

【神女】 ①女神。②旧时指妓女。

可以看出，两个词语的两个义项之间根本的区分之处是情感属性上的不同。义项①为中性，而义项②为贬义。类似的例子可以找到很多，例如“小鬼、小姐、手段”等。在词典义项的划分中应该明确这个划分的标准，同时在机器应用上也将有助于组合关系判断的正确性。

另外，名词的情感属性不仅跟动词组合时有在属性取值上的不相悖的要求，在跟形容词组合的时候，也有属性和谐的要求存在。形容词在情感属性上的不同取值，限定了其所修饰的名词的情感属

性取值范围。动词、名词、形容词在情感取值上都符合下表中描述的组 合限制。

	褒义	中性	贬义
褒义	+	+	-
中性	+	+	+
贬义	-	+	+

表 4 名词意义情感属性与动、形组合关系表

4.4 副词意义的属性

副词在不同的语言中所指称的范围是不尽相同的。在印欧语中，表时间和处所的成分都属于副词的范围，做句子的状语成分。而在汉语中，这两个部分一般是不划入副词范围之内的。所以，汉语中副词的总量相对来说是较少的，但数量少并不意味着不重要，副词的研究一直是学界非常关注的问题。因为其与句法和语义都有紧密的联系，副词的研究是探讨句法与语义接口问题的重要途径。

要想合理的解释副词在组合中的分布和特点，我们认为，必须参照事件结构来看。以往的很多研究是从静态的语义方面来看的，对组合关系成立的解释上不够精细。我们的研究是从组合的方面来看待副词与动词、形容词的关系。副词与这两类词之间组合关系的形成需要遵守属性和谐的原则，需要在一定的分量上能够组合成功。

副词的分类主语有三个不同的思路。一般认为，汉语的副词可以划分为时间副词、程度副词、范围副词、数量副词、否定副词、情态副词、语气副词、处所副词、疑问副词几大类。这样分类的标准是按照语义的。这种在句法研究的大框架之下根据语义进行细分的做法是典型的语义句法的做法，其对于解释和理解副词是有意义的。

从分布来看副词的分类，是副词分类的另外一个重要分支。Jackendoff (1972) 根据短语结构语法把副词分为了说话者朝向的副词(Speaker-oriented adverbs), 主语朝向的副词(subject-oriented adverbs), 方式、时间或程度副词(adverbs of manner time or degree), 焦点副词(focusing adverbs) 四类。Jackendoff 的分类大致是从语义的角度来分的，但是更重要的是参照了分布的特征。其中，说话者朝向的副词(Speaker-oriented adverbs) 和主语朝向的副词(subject-oriented adverbs) 为句子层面的副词，方式、时间或程度副词(manner、degree、time) 副词为短语层面的副词。句子层面的副词位置由句子结点控制，动词短语层面的副词的位置由VP 结点控制，焦点(focus) 副词只能处于助词的位置。这些分类是根据副词在句子中的分布情况所做的分类。应该看做是形式的角度。Jackdoff 的分类的重要之处在于实现了副词的语义类型分类与句法分布之间的对应。

通过事件结构来看副词的分类，是另外一个比较新的切入点。比较著名的如 Tenny (2000) 对副词的分类。从事件结构的角度来看，有两个元素必须在副词的分类中体现出来。一个是与核心事件和对事件的度量有直接关系的副词，另一个是其他类的副词。

核心事件与对事件的度量是复杂事件包括的两个重要成分。从而把副词分为度量副词(The measure adverbs)、反复副词(The restitutive adverbs)、几乎副词(the almost adverbs)。

在语序上可以把副词分为三个大类，一类是紧挨着动词出现的类，基本上为事件内的副词，一类是表示句子之间逻辑关系的词语，因而处于句子的连接域内。第三类则是比较自由的类，其位置随着修饰成分的变化而变化。

我们不采用组块分析法中把语义与分布捆绑在一起的做法，而是把副词的语义(副词并不像助词那样无实义) 与其词序的序列分离开，则会发现另外的一个可能视角。在属性上采用传统的语义分

类法，结合事件来考虑语序。句法上的依存关系大致是以语义上的联系作为基础的，而语序则是另外的一种限制规则（其形成的原因虽然有其语义的影响，但可以看做是一个独立的层面）。这样就可以理解语义指向理论中谈到的，汉语中很多副词可以有多个的“项”，在具体的语境中多只表现为一个，却使用的是同一种形式的现象。语义决定其可能联结的对象，词序决定其在句中所处的位置。两者属于不同的限制层面。一定语境中的副词使用多为单义的，同一种形式可以表达不同的语义联系即为词序制约的直接表现。

汉语的副词是实词（动词）在历史演变过程中逐渐语法化而形成的，其位置受其历史影响很深。在汉语中，有一类与动词有关的副词多紧挨着动词出现，且位于动词之前，这即是动词虚化为副词的过程中再分析决定的位置分布。副词位置的形成，将在下文的语序部分做详细的解释。这里我们只解释副词与动词在纬度上的和谐原则。

副词的许多属性是与动词的属性相对应的。在对动词属性的了解基础上，我们可以探讨副词与动词组合在纬度上的和谐现象。

副词与动词组合时候的属性和谐，其表现相对于名词与动词的组合上可能要简单一些。这是因为，副词的意义大多是由动词的某个义项逐渐虚化而来的，其分化为副词的时候是比较单义的。因此，在匹配时选择的过程较之名词与动词义项的相互限制要简单一些。

以时间副词“已经”为例，我们来说明副词与动词组合时候在过程属性上的和谐现象。“已经”的字典释义为：☐表示动词变化、完成或达到某种程度。其词义与动词的过程属性有紧密的联系。

（86）他已经开始忙了。 （开始）

（87）他已经忙了三个小时了。（持续）

（88）表已经停了。 （终止）

（89）他已经走了。 （结束后）


的动作概念，两个谓词所论述的主语甚至可以不同，必须明确给予区分。

两个动词“打”与“赢”，只有“赢”带有时体标记“了”，“已”只与“赢”有句法依存的关系。这是因为“赢”使用的是动词的结束后分量，而“打”则与“事务所”有句法依存关系，也可以说是具有系统性的语义依存关系。这样就形成了两个边交叉的现象。

我们在前面曾经论述到，交叉依存现象的解释需要两个方面，一个方面是依存关系为什么存在，另一个方面是为什么处在隔断的位置上。副词和动词组合关系的存在在分量上的和谐原则则是对第一个方面的回答。在下一章，我们将论述第二个方面，即副词与动词的依存关系并不因位置的间隔而隔断的问题。

副词的另一个小类为修饰数词的数量副词，在语序上基本属于较自由的类。其位置既可以与所修饰的数词相邻，也可以与动词相邻。这样就形成了副词与数词分离的现象。但是这些副词在语义上只与数词有关。所有的数词副词都具有这样的特点，因而可以看做是句法层面的依存关系。


(94) 据美英宣布，这次空袭行动共击中了 100 多个伊拉克军事目标，达到了预定的目标。



数量副词“共”与数词“100”有依存关系，动词“击中”与名词“行动”之间有依存关系，两个边之间形成了交叉。

另外，交叉依存现象中的框式副词类，是两个副词强制同现的情况。一般表示事件的先后关系。因此，副词除了各自与动词有属性上的组合关系外，两个副词之间一定程度上可以看做成成对的连词。

(95) 这些经济活动一出现就被纳入法制轨道。



4.5 介词的属性

介词的类在语言中是广泛存在的。从题元角色的角度来看，介词的主要作用是给动词不能直接赋元的名词赋予题元角色。在前文中论述到，动作不仅有过程的属性，还有位置的属性。位置的表达一般通过介词加名词的形式。因此，介词的一个重要属性即是位置。以往树形的表达中，只表达了动词分别于两个介词引介的名词之间的位置关系，对两个介词之间的关系没有很多的关注。

朱德熙《语法讲义》中指出，介词的作用在于引介与动作相关的对象（施事、受事、与事、工具）以及时间和处所等。因此，介词与动词的属性是紧密相关的。引介动作相关的对象是动词配价的延伸，引介处所与时间，则与动词的时间与位置属性紧密相关。我们知道，时间与空间是一种比喻的关系，在介词的使用上是相通的，是从具体到抽象的发展过程。因此，我们可以把介词大致划分为两类，一类是引介相关对象的，一类是引介时空概念的。

Greenberg（1995）提出了框式介词的概念。这种框式介词由一个前置词和一个后置词共同构成一个框。前置词多是由动词语法化而来的，后置词则多是由名词语法化而来的。在汉语中，这些后置词多为方位词中的一些词项。方位词是后置词的重要来源（刘丹青 2003）。如“在……后，在……上”的形式。这里，我们提出要注意由两个介词构成的另外一种框式，如“从……到……”的框式。

介词连用经常表达的是一个动作发生位移的起点与终点。由于大多数介词还是处在由实词向虚词的语法化过程之中，介词在很多时候与实词界限模糊，具有一定的连续性。所以很大程度上保留了某些实词的实义性。更多虚义的范畴也都是在实义范畴基础上的隐喻或引申，因此其句法关系与表达实义的位移所涉及的是一样的。

表达位移的动词跟两个介词引介的位置有关系，一个是位移的起点，一个是位移的终点。连用的介词之间的依存关系是先在的，

是人们对客观世界事物之间联系的直接体现。有了位移的终点，必然要求有起点的出现。因此可以确定两个介词之间有共现的句法依存关系。武汉大学使用的特征结构标注中，把类似“从……到……”的结构看做是关系特征来表示，即是对这种共现关系的重视与表达。在动词后的介词，能受动词影响念轻声。在节律上依附于动词，而与介词支配的名词有韵律上的界限。因而有时候时体标记也会附加在介词的后面。

与副词相类似，介词大部分也是由动词逐渐虚化而来的。其义项比较单纯，因而其具有的属性也相对比较单纯，引介时空概念的介词只有位置的属性¹⁸。连用的两个介词，一个表示起点，一个表示终点，在位置的属性上是相和谐的。位置虽然是语义的概念，但是由于其系统性，我们作为句法层面判断的标准。介词与动词的依存关系是词类之间的，介词与介词之间依存也可以是词类之间的。

介词的连用，可以形成交叉依存的现象，即两个介词短语之间出现动词。动词与其他词语形成的边会与两个介词形成的边形成交叉。介词分别处于动词的左右，是汉语在类型学上比较特殊的地方，一般认为，在一定的语言中介词短语或者放在动词前，或者在动词后，前后都出现的情况说明这种语言正处在演变的过程中。即一些学者（孙朝奋 1998）认为的汉语正处在前置词短语从动词后向动词前的发展过程中。

介词表达的位置属性，可以对应于动词的位置属性描写，分为四个取值。分别是发生位置、起点位置、经过位置和终点位置。在表达位移概念的动词中，可以有分别表示起点和终点两个取值的介词组合出现，作为组合成立的和谐表现。表达静止状态位置的介词可以看做是使用的终点分量。介词属性描写可以通过如下的形式：

¹⁸ 引介时间时使用的介词与引介位置时使用的介词是相同的，因此，我们可以把时间看做是位置属性的一种隐喻。

介 词	介 词 位 置 属 性			
	位 移 位 置			发 生 位 置
	起 点 位 置	经 过 位 置	终 点 位 置	
从	+	+		+
到			+	
由	+	+		+
至			+	
在		+	+	+
自			+	+
于			+	+

本文研究的交叉依存现象中的框式介词类,除了动词与介词之间在组合时有在位置属性上的一致要求外,两个介词在位置属性上的也具有组合性。其系统存在于引介时空的介词中,我们把这类系统性的语义特征看做是句法层面的特征,是依存关系成立的判断依据。

(96) 去年实现进出口总值达一千零九十八点二亿美元，占全

国进出口总值的 $\frac{N}{P_1}$ 比重由 上年的百分之三十七 $\frac{V}{P_2}$ 提高 到 百分之三十九。

4.6 小 结

本章论述了词语组合关系成立的属性和谐原则。属性和谐是指词语之间在语义上的系统性相关，词语只有在属性及其取值上不相斥才可能组合成功，这是句法依存关系成立的根本原则，也是本文

研究的交叉依存现象中分离的两个词语之间依存关系存在的原因。而且，句法依存关系的存在并不因词语在线性距离上的隔断而中断。属性和谐原则虽然是从交叉依存现象的研究中启发得到的，但同样适用于所有词语间组合关系成立的情况，因而是一条普遍性的原则。

句法范畴从实质上讲即是系统性的语义范畴。作为属性来使用的都是系统性的语义特征，我们把其作为句法层面的标准来使用。当然，还有更多的语义特征是缺乏系统性的，需要逐个词地学习和掌握。在对词语组合成立的限定上，往往是一些细节性的语义特征起到的限定作用更大。系统性的语义特征则是其基础。因此，句法与语义两个层面不仅需要区分对待，而且需要综合来看。对于汉语这样的孤立语来说，很可能仅靠句法范畴是不能够自足的，还需要结合语义。这正如刘丹青（1994：11）所说的：“汉语的语法系统不是句法规则和其他规则的结合，而是句法规则和其他规则的融合”。在句法层面的属性和谐基本可以解释交叉依存现象，但个别的现象在遵守这些大的原则之下还需要借助更加精细的个体语义特征。

而且，把系统性的语义作为句法层面的标准来使用，也可以加深对句法概念的理解。印欧语中的形态我们也可以不认为是词类划分的依据，而是功能类别的标记。例如名词的单复数与动词的单复数表达的就是不同的概念。名词的单复数是语义的反映，而动词的单复数则是功能的概念，只有和主语名词组合的时候才会使用。词类之间的转换，其形态的标记也可以看做是功能的标记，而不必定是词类本身的变化。

因此，英语中用形态表示的词类差别仅仅是功能差别的表现，参照的是功能标准。但不能因此推导划分词类只能从功能入手。汉语中，这些功能的表达则是由语序，特别是相对的语序来承载的，并没有像英语那样使用形态这样的衍生物。

由此观之，并不能通过对古人根据语义对词语静态划分做法的

否定来作为现在按照功能分词的论据，按词义划分词类不应该受到如此多的攻讦。

相应地，一系列由此推演出的看法也需要重新审视。如对于按照词义给动词划分次类的论断。根据词义的分类，“由于脱离了汉语动词特有的组合规律，这种分类往往无助于描写汉语句法中隐藏得较深的一些范畴，无法揭示词语搭配的本质。（申小龙 1987:76）”这个论断是从组合研究的单方面来说的，其结论可能过于绝对。表象类似的组合中既有可能包括共同的部分，也可能只是表象相同而已。而从语义出发的分类，则具有先验的一致性。在此基础上考察其表现的异同，同样具有重要的意义。从现象的相同寻找共同的范畴与从语义出发来寻找共同的表现是语法研究的两个方面，应该互相观照才能共同促进。

5 环形语序域原则

在第四章，我们论述了句法依存的属性和谐原则。这是制约交叉依存现象形成的两条基本原则之一，属性和谐是依存关系成立的基础。同时，交叉依存现象中词语间的语序分离现象则受环形语序域原则制约。这同样是制约交叉依存现象形成的原则之一。在这一章，我们将详细论述之。

5.1 网络视角下的线性一维化

5.1.1 语序成因的现有解释

语序的研究一直是语法研究中关注的一个热点。语序不仅是重要的语法形式或语法手段，而且是一种基本的语法分析的方法。语序分析在不同语言中都得到广泛的运用和不断的发展。特别是在语言类型学的研究中，语序更是其一个核心的领域，人们把语序的特点作为考察语言类型特征的一个重要标志。

“词序”和“语序”是相关研究中经常遇到的两个术语。这两个术语在汉语研究中的具体使用情况比较复杂，在理解上有很多分歧。有的学者不加以区分，有的则分为不同的指称。具体来说，“词序”和“语序”用来指称语法单位的排列顺序，或用来指称结构成分之间的排列次序的情况在学者中都有发现。也有学者区分两个术语分别为“单位序”和“成分序”。范晓（2001）即用“词语的序列”和“结构成分的序列”来区分词序与语序，把两者区分为不同层面的概念。词语的序列和结构成分的序列的区别是从短语结构语法框架之下做出的区分。短语结构作为介于词和句子之间的中间单

位，对语法研究具有非常重要的意义。但是，在依存语法系统中是不认为有短语结构这样的中间语法单位的。整个句子只是若干个词语通过句法关系组合而成的序列。因此，“词序”与“语序”两个术语在依存语法的视角下并没有区分的必要，而且在英文中的指称都是“Word order”，很难说其实质上有什么差别。本文在对这两个术语的使用上除非特殊注明，一般不做特意区分。

句子所呈现的线性序列形式不是无序的，而是有其内在规则的制约。不同的语言具有不同的线性表现。是否能够透过这些线性列表面的不同去发现其共同的规律是语言类型学学者们长期以来一直在努力探索的一个方向。作为当代语言类型学鼻祖的Greenberg(1966)曾指出，如果一种语言中单句为OV(object-verb)式，其名词短语多为M-N(modifier-nominal)式结构，且后缀形态占主导。如果单句为VO式，则其名词短语多为修饰语后置的N-M式结构，前置形态占主导。Givon(1984)的《syntax》一书中也指出，在单句条件下，语言可以划分SVO、SOV、VSO、VOS几种主要的语序语言类型。并列举了不同类型词序语言的代表。然后分别举例探讨了其分句、定语等词序特点上的类型性。其他很多语言类型学家也在致力于探讨不同语言的共同之处与不同之处。但是，这些研究都是从线性角度来解释语序共同之处的。

对于汉语来说，语序的研究一直都是关注的一个焦点。学者们尝试从不同的角度来解释语序的制约因素。从二十世纪八十年代的意义角度(李英哲 1983; 陈平 1994)认为语义单位的排列次序是反映的概念顺序，到后来的语用(方梅 1993)和认知上的时间顺序、信息处理难易度原则(戴浩一 1985; 张敏 1998; 袁毓林 1999)等角度，提出了不少对汉语语序形成机制的主张。但是，总体上来说，这些规律都是倾向于举例性的，只能够解释部分的语言事实，缺乏整体性。而且，语用与语义可以作为说明语序特点的一个途径，并不是语序形成的根本因素。语序是一种手段与现象，句法、语义

和语用都是可以由语序承载的，用语义、语用解释语序的形成，在逻辑上是本末倒置的。正如体态作为一种手段，可以表达一定的意思，达到一定的交际效果，但不能说是意义或者交际效果决定了体态的手段一样。在不同的文化语境内，同一个体态可能表达完全不同的意思。如，摇头在有些地方表示认同，有些地方表示不认同。所以，语序问题的根本原因，不能从语用或者语义的角度来解释。

同时，把语言看做是一个网络结构的观点也开始从词汇语义领域向句法领域拓展，使得两个领域的研究有了沟通的趋向。人们开始从网络的视角来看待语言组合的构成。许多学者表示了认同并且试图给予证明。心理学上一个普遍接受的观点认为，人们是通过网络的形式来储存社会的知识的。是否语言也是以网络的结构存在呢？一些心理学家也试图通过实验的方法来证明。例如，Miller 等人对 WordNet (Miller et al., 1990; Miller, 1995) 的建设尝试。Ninio (1994; 1998) 的实验中，也认为依存语法可能更符合人类习得语言的本质，人们习得语言的过程是一对一的词语依存关系的习得过程。但是 Ninio 的论述中并没有解决语言的线性呈现与网络结构之间的对应问题。在句法学上，也有 Hudson 用网络的观点来论述句子的结构。但其主要是针对具体语言的，并没有针对人类语言的共同性与差异性来看。而且其研究的角度是从线性的句子来看网络的构成，并不是从网络到线性的投射。这应该是两个有紧密联系但不尽相同的研究角度。从网络结构的角度来探讨不同语言的线性呈现之间的相同与不同的理论目前还不多见。而且，平面的网络和立体的网络也是不同的，立体的网络可以投射为不同的平面网络图像。本文的研究则是尝试从立体网络的角度出发，看网络的投射与线性的序列之间有何相通之处。

这里需要注意的是，语言网络观需要区分两个不同的层面。第一层认为，关于现实世界的一般性知识是用网络的形式储存的；另一层则是认为，语言的知识也是用网络的形式在人脑中储存的。我

们这里所指的网络是两种关于网络看法的总和。即认为语言的网络是一般知识概念与语言知识概念共同构成的网络结构。这样的网络又分为静态与动态的两种（刘海涛 2010）。静态网络反映的是语言作为知识网络的状况，而动态网络则是真实使用中的句子构成的网络。

正如刘海涛（2010）所指出的，虽然语言网络的观点甚嚣尘上，但多是作为一种隐喻来使用的。大家都没有解释多维空间的网络是怎样对应到一维的线性序列的，而多是在线性序列上来观察和描写语言的语序问题。因此，语序问题的几个基本的问题还是没有很好地解决，并且必须要解释清楚才能更好地支持语言网络的观点。即网络结构与线性的语言现实之间是怎样形成对应的？是什么原因造成了这样不同的对应类型？这些类型之间有什么共同之处？在语言网络观的视角下，这是个无法回避的问题，必须给出一个较合理的解释才能是语言网络观与线性的语言序列形成统一。

相应地，如果要构建一个相对完整的语序研究体系，应该考虑包括至少以下三部分的内容：

第一，语言中的语序需要遵守什么样的原则；这又包括两个部分：首先，要考虑人类所有的语言有没有需要共同遵守的原则；其次，则是某一种语言有哪些具体的原则。

第二，不同的语言为什么要遵守共同的原则，又为什么有自己不同的特点；即语言表象背后的制约机制；

第三，在一种具体的语言现实中，经常会存在某些语序灵活变化的现象，为什么语言中会具有这样的灵活性；

这三个基本问题的合理解答，是构建一个合理的语序观点所不能偏废的。现有关于语序的研究，在描写层面做的较多，解释层面做的较少。即大多是在描写不同的语言有什么共同的大致规则以及各自不同的独特规则。在解释为什么遵循这些原则的时候感觉不是很透彻。类似于 Greenberg（1966）和 Givón（1984）等学者的论

述，也大致算是描写层面的。解释都是从线性的角度、功能的视角来论述的。Givón (1984) 认为，词序具有被用来区分不同的格角色的功能。同时，他也谈到词序在句法系统中还具有许多其他的功能。如保持话题的连续性，表达对比与强调等。在此认知的基础上，可以把语言区分为词序固定的语言和词序灵活的两种类型。这样的解释具有一定的合理性，但是把任何特殊的结构都归咎于功能的说法却是值得商榷的。且对相同语序的解释上不能够让人信服。戴浩一的象似性原则，袁毓林的信息处理难易度原则，也都只能解决部分现象，缺乏整体性。

5.1.2 语言立体网络到一维线性序列的投射

我们认为，功能虽然是影响语序的一个重要的因素，但是功能并不是决定语序的因素。而且，功能这个比较笼统的概念也不能够囊括对所有的语言现象的解释。功能的需求要与语言具体特点（语言形式可能有的区别方式）相适应，只有语言本身提供了形式变化的可能，才能有功能分化的可能。不同的语言外部形式应该有个共同的内部形式，才是人们能形成同一反映的基础。这个共同的语言内核形式，是值得我们探讨的。

从依存语法的角度来探究句子线性化规则的研究很少。根据 Mel'čuk(2011)的观察，目前只有两种从依存角度来定义的线性化的模型。第一种是 Mel'čuk (1965, 1974, 2011) 为代表的，从交际结构 (syntactic-communication structure) 来定义句子的线性化方法。第二种是 Gerdes 和 Kahane (Gerdes 2002, Gerdes&Kahane) 为代表的定义全部句子模式 (full-sentence pattern) 的方法。两者都没有解决从网络结构到线性结构的对应问题。从依存的角度出发，我们尝试做这样一个新的基本理论假设，来解释语言中既有相同又有不同的语序现象。

我们认为，语言是一个立体的网络结构，是空间内多维的网络。动态语言网络在使用中表示的是现实世界内的事件。赖金铤(1992)曾指出，语言的网络是六维的。当然，语言维度的具体个数是多少个，与作者的具体研究立场与目的有关，或许我们可以再商榷。但语言是个多维网络的观点却是基本上可以接受的。我们可以把语言的网络看做是由词语所反映的概念组成的事件组成的立体网络。同时，我们说出来和看到的书面语言序列则是一维线性的。这同样是语言网络的另一种表现方式。从多维的立体网络到一维的线性序列的呈现是一个逐步减维的过程。语言的线性实现即是立体网络的一维化。下图即为语言立体网络的简单图示。

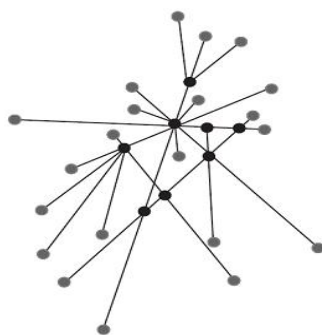


图 3 语言立体网络图示

语言符号的物理特性之一就是具有线条性 (linearity)。网络从多维空间到一维线性序列的转换过程就是学者们所说的事件结构到论元结构到句法结构的投射过程。从多维到一维的直接转化过程可能比较抽象，我们可以从比较熟悉的二维网络到一维的过程来说明。实际上，我们熟悉的树形图，无论是依存树或者是大家更为熟悉的短语结构树，都是语言的二维呈现形式。从二维平面图到一维线性序列的转化过程即为减维的过程。Tesnière 曾经有过类似的论述：“说一种语言，就是把结构次序转换为线性次序的过程。反之，

理解一种语言就是将线性次序转换为结构次序的过程。¹⁹”从空间的多维到平面的二维是一种位置的投射，再到线性的一维呈现也是一种位置投射，是由语言的生理、物理条件决定的。这即是索绪尔所说的能指的线条性。这其中的减维过程，也是形成本文讨论的主题即交叉依存现象的关键。而且，这个过程可以用数学的方式精确描述出来。即数学上的拓扑等价。从三维到二维的投射过程可以参照下面的草图。

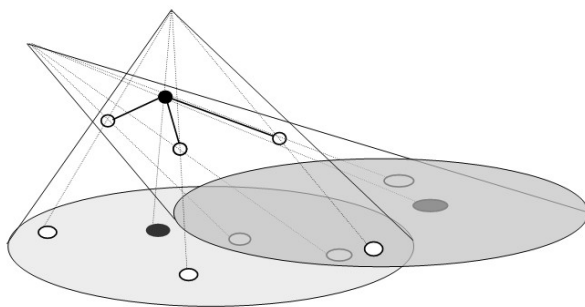


图 4 三维图到二维平面的投射

可以看出，同一个空间网络结构，由于观察角度的不同可以形成不同的投射角度，因而形成不同的二维平面投射结果。但两者反映的实质是等价的。

语言的网络是对客观世界的反映。客观世界是由物质与运动组成的，物质离不开运动，运动也离不开物质，反映到语言中来就是动词与名词两大基本的词类。一个事件结构最简单的组成即是动词和名词。以上图为例，如果让黑色的球代表动词，白色的球代表名词，那么基本可以形成一个以黑色球为中心，白色球环绕周围的平面图形。其他的词类都是聚集在这两大类词的周围。这些词语之间的联系可以用有向的连线来表示。在不同的层面这些连线表示不同的关系，有句法、语义和语用关系等。如果我们用平面图的形式来表示这种关系就是扁平化的二维平面图。动词与其名词的联结共同

¹⁹ 转引自刘海涛 2009《依存语法的理论与实践》第 120 页。

表达的是一个基本事件结构。在语序的层面，如果我们用“域”来称呼不同词类的范围，那么可以成为环形的语序域。

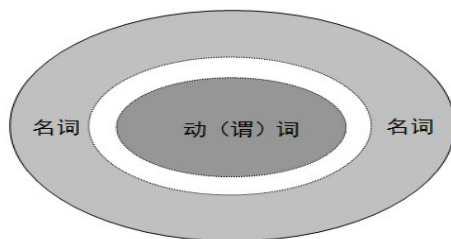


图 5 事件结构二维环型图示

为了便于说明本文讨论的问题，上图中只简单地列出了几个基本的词类。从单个事件表示的角度来说，一般是动词处于中心的位置，名词处于动词的外围。两大词类之间的界限基本上是比较清楚的。如果把这个二维的平面图转化为一维的线性序列以后，则为一个普通的句子序列。



图 6 句子的一维线性图示

理论上来说，投射的观察有几种不同的视角。以由一个动词与两个名词构成的简单句子来看，既可以动词为中心来观察投射，也可以其中一个名词为中心来观察投射结果。第一种即为“动词中心说”，与第二种类似的则有“主语中心说”（宾语中心的理论则不多见）。两者观察的角度虽然不同，但在实质上却是相同的。鉴于动词中心的观点为大家所一致认同，同时也是我们坚持的依存语法的基本立场，我们这里只采用以动词为中心的视角来观察投射的结果。

可以看出，从二维到一维转化的过程中，不同词类的位置可以具有一定的灵活性。即成为了 名词 -- 名词 -- 动词 -- 名词 -- 名词 的可能序列。在句子使用中，不同的名词可以分别担当主宾

语的角色，因而可以允许 SOV、OSV、SVO、OVS、VSO、VOS 六种不同结构的出现，这也正是语言类型的几种基本分类。因此，这些不同的语言类型可以具有共同的事件反映内核。也就是说事件结构是所有语言的共性，这才是不同类型的语言之间的本质相同之处。而且，我们由此可以推论不同的语言使用者之间可以对同一事件形成共同的头脑反映映像的原因。

我们可以通过下表中不同语言的例句来验证上文中的论述的概括能力。²⁰

语言类型	代表语言	例 句
SVO	英语	The man hit the ball.
SOV	夏 尔 巴 语 (Sherpa)	ti mi-ti-gi cenyi caaq-sung the nan-DEF-ERG cup break-PERF The man broke the cup.
VSO	哈 卡 尔 特 克 语 (Jacalteco)	xa' ix te' hum wet an gave CL-she CL-the book to-me pl She gave the book to me
VOS	马 尔 加 什 语 (Malagasy)	manasa lamba Rasoa wash clothes Rasoa Rasoa is washing clothes
OVS	Hixkaryana 语	kana yanimno biryekomo fish he-caught-it boy The boy caught a fish
OSV	Xavante 语	无

表 6 二维图概括能力在不同语言类型中的例证

²⁰ 表中用例引自 Givón, T. (1984) *Syntax : a functional-typological introduction*. 第 6 章。

具体的语种最初不同的语序选择会直接影响到语言的特点。在 SVO 的语言中，由于 S (subject) 与 O (object) 处于动词的不同侧，比较容易区分，因而较少需要或者不需要形态的标记来帮助。而在 SO 处于 V 的同一侧的语言中，主语和宾语往往需要借助形态的变化来区分不同的身份。或者两者都添加标记，或者给其中的某一个添加标记。在通过形态区分主宾语成分的前提下，又可以允许一定范围内比较灵活的语序改变。例如德语、俄罗斯语等。朱德熙 (1985: 3) 曾指出一个拉丁语中极端自由语序的用例，即在拉丁语中三个词语可以有六种合法的组合形式。

- (97) Paulus vidit Mariam. (98) Mariam vidit Paulus.
(99) Paulus Mariam vidit. (100) Mariam Paulus vidit.
(101) Vidit Paulus Mariam. (102) Vidit Mariam Paulus.

另如 Mel'čuk (2011) 举例的俄语中“我爱你”三个词语的六种合法的组合形式。

动词与名词构成的事件是一种最简的事件结构。如果增加其他的词类，则投射的过程将会复杂很多。在一个动词两个名词的基础上增加一种词类----副词，则会增加几种可能的形式，形成 副词 -- 名词 -- 副词 -- 名词 -- 副词 -- 动词 -- 副词 -- 名词 -- 副词 -- 名词 -- 副词 的可能序列。在不同的语言中，同样可以选择使用不同的投射方案。

由单个事件构成一个句子是理想状态下的情况，实际的语料中这样的句子相对还是较少的，更多的句子表达的是事件的组合，可以称之为复杂事件。复杂事件是单个事件通过不同的方式组合而成的。可以是单句的形式，也可以是复句的形式。我们一般用形式上的单句或者复句作为划分句子的不同形式标记，但是对于事件结构来说，单复句的区别没有实际的意义，单句中可以是复杂的事件，

复句也可以表示的是单个事件。

汉语中，事件组合的方式有并列、链接、拼合、接合、嵌套和交叉六种。在下一章语料库标注部分会做详细的解释。

5.1.3 从网络投射看汉语的语序

语序在汉语中具有特别重要的作用。因为汉语没有类似于印欧语言那样丰富的屈折体系，所以词序相比于屈折语言来说必然要承担更多的语法功能，在整个语法体系中占有至关重要的地位。

1. 汉语语序总体特点的再认识

对于汉语的语序特点，不同的人从不同的角度曾有不同的论断。有观点认为汉语的语序是比较固定的，如王力（1944）即认为：“词序固定是中国语的一大特征”，洪笃仁（1955）也指出：“汉语的词序是固定的”。同时，也有观点认为汉语的语序比较自由。吕叔湘（1986）在承认汉语句法固定的一面下，详细描写了汉语中移位、省略和动补结构多义性等灵活性的表现。针对这两种看似矛盾的观点，张斌、胡裕树（1984）指出，认为汉语语序固定的观点多是从短语结构的角度来看的，“把语序限制在语法结构的范围之内”，在先确定结构类型的前提下再观察语序的做法是很容易得到这样的结论的。如果跳出语法结构的限制，即容易得出语序在动态句子中比较灵活的观点。正如范晓（2001）所论述的，“如果语序没有相对的固定性，那语序也就没有规律可言；如果语序没有灵活性，那语序在运用中就变成僵化的教条。”固定性与灵活性是对立统一的。绝对地说汉语语序固定或者灵活都是不合适的。根据前文中我们对语序的看法，我们对语序问题持如下的观点：如果我们把动词与名词的位置不同看作不同的域，域层面上的语序称之为宏观的语序，域内词语之间的位置关系看作是微观的语序。我们认为，汉语的语

序在宏观语序上是比较固定的，但是在微观语序上则比较灵活。

2. 从投射看具体语序的一些争议

汉语语序的研究有很久的历史。自张世禄在二十世纪三十年代提出“语序论”的观点以来，语序研究作为一种基本的方法论，在语法研究上得到了广泛的运用。朱德熙（1982）论述到，汉语倾向于用主语表示已知的事物，用宾语表示不确定的事物的特点；文炼、胡附（1984）认为汉语的语序有语法、语义和语用三个不同的层面；戴浩一（1985）论述了汉语语序符合时间原则的特点；等等。除了这些理论性探讨之外，对细节的词序探讨也有很强的传统。参照申小龙（1986）的分类标准，可以把这些研究大致分为两大类，一类为论元成分的语序研究，另一类为非论元成分的语序研究。

对论元成分的语序研究，直接体现为在二十世纪六十年代关于汉语主宾语的大讨论，以及二十世纪末开始的国际汉学界对属于汉藏语系的汉语是 SVO 语言还是 SOV 语言的讨论上。

从《马氏文通》开始，大致是以 SVO 的框架来描写古代汉语的，以语义来区分主宾语，同时允许有条件的 SOV 语序。后来的黎锦熙、吕叔湘和王力在对汉语语序的论述上也是采用按照语义关系来划分主宾语的做法。但是，也有学者质疑从语义区分主宾语的做法。刘丹青（1995）针对钱乃荣（1990）主编的《现代汉语》列出的 16 类主宾语语义类型表，指出其“都是同时适用于主语和宾语，可见靠语义关系来区分主宾语是完全无望的”。从赵元任《国语入门》开始，结构主义在中国的影响逐渐增大，成为语法的主流。结构主义学者一般以语序作为确定主宾语的唯一标准。

由于汉语中不存在类似屈折语中的主语与谓语动词在形态上的一致关系，主宾的身份难定。因此，主宾的概念在汉语中就有不同的理解。有认为是形式上的，有人认为是语义的，还有认为是语用的。许多争议都是源于对主语概念界定的不同。结构主义代表的学

者主张纯粹的按照名词语序先后来确定主语的做法，可以看做是对语序重要性的强调。但是这样会引起主谓结构的无穷嵌套，对汉语研究来说意义不大。主宾讨论无疾而终，最后确定的“形式与意义相结合”的标准，依然不具有可操作性。

随着语言类型学研究在汉语界的引入，关于汉语语序类型的争论再次引起学者的关注。戴浩一（1973, 1976）认为汉语是 SOV 的语言，从古代汉语到现代汉语是在经历一个由 SVO 到 SOV 的过渡。Li & Thompson 也从历时的角度提出汉语是从 SVO 向 SOV 型的转变。汤廷池（1972）则更是认为现代汉语在深层是属于 VSO 或 VX 语序。而 Light（1979）和 Mei（1980）等则坚持汉语是 SVO 型语言的观点。

在二维图中可以看到，动词处于最中心的圈，动词直接支配的名词则都处于外圈内。在一维的线性表示中，才有了放到动词之前或之后的区别。但是，放到动词之前或之后并没有本质的区别，都是动词直接联系的论元成分。

因此，主宾语的争论我们可以这样来看待：投射视角的不同可以得出不同的前后顺序。空间网络的投射提供了处于左右两种可能，才有了功能分化的前提，也提供了语序变化的可能。在逻辑语义上基本是等同的。S 与 O 其实都是在 V 的外层的名词性成分。不同语言可能选择了 S、O 不同的位置来区别不同的意义。在三维空间是没有绝对意义上位置先后的区别的。本质上来说，主语概念在屈折语中是一种一致关系，主要体现为主语与谓语动词在“数”范畴上的一致。而不绝对是语义上的相关。因此，我们认为主语宾语这样的术语范畴对于汉语来说只有区分角色的参照意义，并不能对应于汉语中的具体类。

在事件表达的角度来看，这些表示在实质上是一样的，也即句子表达的逻辑语义是一样的。之所以有前后位置的区别，除了是根据表述起点的不同而选择的，同时也有避免歧义的考量。在没有歧

义的情况下，动词支配的成分的位置可以相对自由。但是如果放置位置不同的情况下会产生歧义，则要求的位置比较固定。如学界讨论过的主宾互换现象，不同的学者对其作出了详细的描写和不同的解释（李宇明 1987；李敏 1998；任鹰 2001）。我们认为，之所以两个部分可以互换，根本的原因是避免歧义的发生。“十个人吃一锅饭”之所以能说成是“一锅饭吃十个人”，是因为交换“吃”行为的主客体后不会发生歧义。如果可以发生歧义，则一般不能换用，如“老王批评小李”，换位后为“小李批评老王”，则跟原意相违，所以不能说。这个原则我们称之为“无歧义相对自由原则”。

另外，现代汉语与古代汉语相比有两个明显的特点，一个是借助前置词“把”的受事前置句，一个是代词宾语前置。这些貌似矛盾的古今演变现象在我们的理论下具有了统一解释的可能性。即同域内的词语前后位置的转换。

非论元成分的语序研究主要体现在对定语和状语位置的讨论上。关于状语的位置，学界的争论不是很大。基本上认可状语不仅可以放在句首、动词前，也可以放在动词后面。“倒装的状语一般起补充说明和突出描写的作用。”²¹但对于定语的位置上争论较多。定语能否后置，能否前行，都有不同的观点。以王力、陈望道等为代表的大部分学者都认为定语是可以后置的。张志公（1953）、李芳杰（1983）等则认为修饰语在前，被修饰词在后的语序不能够颠倒。史存直（1982）也认为对定语后置需要加以限制。

陈望道（1978）《文法简论》中认为“定语的前附带有限制描写的作用，后附则带有说明、补充的作用”。这样的说法其实是值得商榷的。所谓“限制描写”和“说明补充”，都是一个笼统的概念，只是对人们观察感觉到的定语位置不同的不同命名而已，并没有说明本质的不同。而且，“限制”与“说明”并不是两个相对的概念，

²¹ 申小龙（晓珑），1987，汉语语序研究述评，载《语文导报》1986年第9期。

不具有非此即彼的互斥关系，因而不能作为绝对的区分特征。

我们认为，定语后置跟状语后置一样是广泛存在的。最明显的例子即是交叉依存现象中的名词与量词分离类的情况。尽管处于动词的不同侧，修饰关系依然存在。因而是一个环形的语序域。

不同的语言在最初选择的线性序列上可能不同，而且其他的手段也逐渐完善了这种分化的功能。在形态发达的屈折语中，形态的变化担当了区分语义角色与表达时间概念的功能，因而在语序上会相对的灵活。如德语、俄罗斯语、土耳其语等。原本为屈折语的英语由于形态标记逐渐弱化，语序有变得越来越固定的趋向，这是分化功能载体由形态到语序演变的一种表现。当然，由于语言使用者的集体惰性，语言特点的演变总是很缓慢的。语言网络的环形投射结果，为语序的演化提供了可能。在类似于汉语的孤立语中，语序则承当了更多的区分角色与传递时间概念的功能。关于汉语的语序，一些认知语言学者认为其句法结构与概念结构之间有对应关系。戴浩一（1985）认为，两个句法单位之间的词序是由其表示的概念世界的时间顺序所决定的²²。这样的观点是有一定合理性存在的。但是我们认为，词序也只是反映客观现实的一种句法手段，不同的语言可以有不同的反映手段。在形态发达的屈折语中，较多地借助形态的改变来表示时间概念。在缺乏屈折变化的语言中，则更多的求诸于词序的句法手段。词序只是对现实世界的若干反映手段的一种，其反映还可以有其他形式。因此认为是概念世界决定了词序先后是比较绝对化的观点。

从空间网络到一维的线条，是从语序的角度来说的。语序的改变并没有改变词语之间的固有联系（语义的、句法的）。语序的关系与句法的关系是两个不同的层面，这跟 Tesnière 表述的“结构

²² The relative word order between two syntactic units is determined by the temporal order of the states which they represent in the conceptual world. (James H-Y Tai, 1985, P49-72)

顺序是二维的，线性顺序则是一维的”²³观点在本质上是相同的。句法关系经常借助语序关系来表达，但并不是完全相吻合的。句法理论的重要问题就是确定多维网络到一维线性顺序的规则。句法的表示应该能很容易地从多维的语义网络中产生，同时又能很容易地转换到一维线性的序列中去。最便捷的形式即是二维的平面图。因此，树形结构是一个不错的选择。这也是大部分理论采用树形句法表示的原因之一。但是，树形的句法关系到线性序列之间的转换并不是直接任意的，还需要符合环形语序域的描述。即在相对固定的语序域之下，允许一定灵活的微观语序。

3. 事件作为中间单位

句子形成的语言网络结构反映的是一个一个的事件。时间、空间、变化等客观世界的概念如何在句法和语义中表现出来，是词汇语义学与逻辑语义学领域的学者们一直努力的一个方向，是词汇语义与逻辑语义两个传统研究领域都在关注的一个主题。在把事件作为语法对象的主题下，两个研究领域现在有逐渐靠拢的一个趋向。词汇语义的研究需要从动词内部向外扩展到句子。逻辑语义的研究必须从句子整体向内关注到与事件相关的动词的语义特征。

在依存语法体系中，一般认为只有词和句子两级语法单位。在印欧语言中，一个句子只有一个主动词，主动词之外的动词需要使用其他的形式加以区别，如不定式或动名词（英语中）。因此句子单位和事件单位在形式上是基本一致的。而在汉语中，并没有明显的形式标记来区别主动词与其他动词，特别是在复杂的连谓结构中。这样在把句子作为单位的依存分析时，不同的谓词及其各自支配的论元的分析就容易出错。因此，我们认为应该构建依存分析的中间单位，即事件结构。一个谓词与其支配的所有论元（必有论元和可

²³ 转引自冯志伟《现代语言学流派》（陕西人民出版社 1999 年 4 月第 1 版）第 163 页。

有论元)的集合构成一个事件结构。再由事件结构组合成句子。事件结构中心不同于传统语法体系的句子中心和结构主义语法的短语结构中心。短语结构中的词和词语之间是组成关系,词组和句子之间则是实现的关系(朱德熙 1985)。事件结构中心从词语到事件结构再到句子都是部分和整体的组成关系。层次性这个语言的本质属性,在依存语法中也能得以体现,即事件又可以作为其他事件的一个组成部分。

Hudson (1984) 也指出,语言是由命题联结而成的网络结构(A language is a network of entities related by propositions.²⁴)。语言传递的信息是一个个的事件(命题)。事件可以分为动态和静态的两种。这与 Bach (1981) 对动词的分类有点类似。只是我们扩大到句子这个单位的层面。动态事件表述的是整个事件的过程,因而关注的是事件的整体。而静态的事件为一种描述,关注的是静态的场面,因而关注的应该是事件的某个局部特征。这类似于摄像中一般场景的拍摄和特写镜头的处理。一般的拍摄是对事件全景的交代,特写镜头关注的则是静态的局部,因而具有强调的意思。类似地,在静态的表述句中,在功能中也具有强调的意味。这是与人的认知习惯相同的。

我们把动态的事件称为“表述”,把静态事件称为“描述”以示区分。对汉语来说,静态的事件其中心可以是形容词或名词。动态的事件可以分为简单事件和复杂事件两种。一般来说,语言中表达动作的词语为动词,因此具有一个动词做谓语的句子可以称之为简单事件,两个或两个以上动词做谓语的句子叫做复杂事件。以简单事件为例,在网络空间中,一个动词的结点与不同的名词结点联结在一起,表达一个简单的事件。复杂事件是由若干个基本事件整合而成的,其整合的方式可以有不同的类型,如并列、链接、交叉等。这些事件的集合构成纷繁芜杂的语言立体网络。

²⁴ Hudson, R. A. Word Grammar. Oxford: Blackwell, 1984:1.

环形语序域原则的提出，使语序的研究从传统的一维线性视角拓展到了二维的平面视角。这将同时有利于促进二维视角的语法研究。以往的研究中较多谈到的是二维的树图表示方式，在对语言本身的研究上却一直是线性的语法研究。观察点从线性视角转变为平面视角，将引起语法研究的一系列变化。许多以往争议的老问题在平面视角下就可能找到新的解决方案。而且，一些新的研究问题也会被发现。因此，从线性的语法研究拓展到非线性的语法研究具有重要的意义。

一般来说，线性系统中的整体正好等于所有部分相加之和，而非线性系统中的整体则不等于所有部分的相加之和，而是很可能大于所有部分的相加。物理学、生物学或者社会学的研究都已经发现存在很多非线性系统。虽然微观上这些系统的基本组成个体和组成法则都是比较简单的，但是由于这些简单的组成单位之间能够自动地相互发生作用，于是出现了复杂性的组织结构。

一个系统的组成个体有无数可能的方式相互影响和作用。正是由于这些无数可能的相互作用，非线性系统展现出一系列以往线性认知完全不同的特点。其中最富有科学内涵和哲学魅力的结论是：一个非线性的混沌系统，一旦超越了它的多样化临界点，就会发生爆炸性的变化；而且原来的平衡一旦被打破，就不可能自行恢复。一个看上去十分微小的诱因，在内外因素共同参与下，有时会产生极其重大，极其复杂的后果。相应地，汉语中一个新的用法或一个新句式的产生，很可能就会产生极大的影响，从而影响到汉语整个系统的全局。前者如某些词语的语法化过程中后期的迅速类推的现象，后者如大家对动补结构的相关研究，动补结构的出现对整个汉语语法结构系统的影响是非常重大的，一系列相关的句式开始出现。因此，语言也是一个非线性的混沌系统。二维平面的视角可以为我们提供一个新的研究问题的角度。

5.2 汉语语序的事件域

我们在 5.1 中简单论述了事件结构在空间网络中的表现形式，以及其二维平面图和一维线性的表示方式。为了讨论的方便，我们只涉及了最简单的动态事件类型。显然，这是比较泛泛地针对所有的语言情况来说的，也就是说是一个可以大致概括所有语言类型的模型。而实际上的句子情况会复杂很多，不仅涉及到语言自身特点的不同，而且涉及到事件之间不同的整合方式。在这一节，我们将结合本文研究的交叉依存现象，针对汉语的情况详细地论述较为复杂的事件形式的表现涉及到的语序问题。分别探讨汉语不同词序域内的语言表达方式。

“域(field)”的概念借鉴于德语依存处理中的拓扑依存概念，他们把句法层面与词语的线性序列层面分离开，把一个句子中词语所处的位置分为不同的语序域。一定的词只能置于一定的域中，不能够跨界。德语中把句法层面与语序层面分开的做法，是针对其较自由的语序与相对固定的句法依存之间的现象的。汉语中有必要引入域的概念是基于以下的考虑。汉语中缺乏屈折变化，判断一个词在句子中充当的成分主要依据是语序。以往对词性的判断中否定了单纯的意义原则，提出所谓的分布。其实质即为位置与意义的综合。在词性问题上的诸多争论除了与词的复杂本身有关，还把一部分词序承担的功能强加到了词性上，造成了现在汉语句法分析处理的诸多争议。如“名动词（朱德熙 1982）”、动词的“名词化”（黎锦熙 1924）以及动词短语“名物化”（石定栩 2003）等观点的提出。

我们主张把词序承当的句法功能与词性本身区分开，把句法依存与词序关系两个层面分开。使用词序域的概念来说明词序本身承担的句法功能。屈折语言中的词性一般来说与其充当的句子成分是对应的，而汉语中则没有这种广泛的对应，因而词序在这里是一个极其重要的参照系。因此，汉语中的域是一个相对的概念，即词所

属的词序范围不是绝对的，而是一种相对的位置。

5.2.1 第一域

作为事件论述的核心成分，第一域具有先在性。一般来说，原型的事件结构中处于事件核心位置的词语为动词或形容词。动词为动态事件的核心，形容词则为静态事件的核心，这是大家比较容易理解的。如王希杰（1988）所说的动作句和描写句的区别。但是事件在添加了附加成分以后，能够处于第一域的词语词类就会比较复杂。现以词类为标准来分别说明。

第一种，动词和形容词

处于第一域内最多的词类为动词和形容词。多个谓词下的交叉依存现象，两个谓词都是处于第一域的。汉语中动词和动词性短语连用的情况，基本遵循的是“时间性原则”（戴浩一 1985），即按照事件发生的先后顺序排列。多个动词连用，表示的是多个小事件的组合。

但是，除去并列的情况外，两个谓词的重要性是有差异的。带有时体标记的谓词其重要性要大于不带时体标记的谓词。不能带时体标记的谓词倾向于只表示一种抽象的行为能力，只能处于带有时体标记的动词的外围。以交叉依存现象中名词与动词分离类和副词与动词分离类的两个例句来说明：

（103）人们把界线打破了。

（104）这个事务所已经打赢了十起案件。

“打”与“破”，“打”与“赢”都处于第一域，但是其重要性是不同的。“破”与“赢”可以带时体标记，因而处于核心的地位。

“打”不带时体的标记，倾向于表示一种抽象的能力，处于外围，而且经常可以省略。

“时间性原则”只是个大致原则，并不能涵盖汉语中全部的现象。而且在某些具体语言现象中也并不遵循这个原则。交叉依存现象中，动词与数词分离类的交叉依存现象即是这样的情况：

(105) 国家财政从一九九八年起五年内 **累计** 投入二十五亿元用于支持该项工作。

动词“累计”与“投入”之间，只有先“投入”然后才能够“累计”，因此，并没有遵循时间先后的原则。

动词及其支配的名词有时候可以构成拷贝结构，以往也称之为重动句。其两个动词性短语中的动词是一样的，表达的是同一个事件。这样的结构如果依旧用时间性原则来解释就显得比较牵强。我们认为，如果动词与动词性短语没有带体标记，则倾向于只表示一种抽象的行为能力，因而处于带体标记的动词和动词性短语的外围。重动句中的前一部分动词短语正是不带有体标记的。例如：

(106) 他 喝酒 喝醉 了。

(107) 他 说话 说错 了。

这里的“喝酒”“说话”都表示一种抽象的行为能力，更倾向于是一个词，一定程度上具有类似于体词的指称功能，因而不带有体标记“着、了、过”。这种表抽象行为能力的动词在划分上是有两可性的，如果有副词，则语序的归属比较明显。例如，“他已经喝酒喝醉了”中，“喝酒”是属于第一域的。但如果副词处于“喝酒”之后，“他喝酒已经喝醉了”中，“喝酒”作为抽象行为能力的指称则是可以归于名词域（即第三域）的。这是域的相对性的具体体现。

在上例重动句中第二部分的动补结构内部，其原则也是相同的（这里讨论的动补结构只包括补语为谓词的情况）。第一个动词不能带体标记，因而表示一种抽象的能力。但是副词只能处于这类抽象动词之前，所以这样的动词也属于第一域。在事件结构中，不带体标记的动词相对于带体标记的补语重要性相对较低，因而可以解释动补结构中以补语为中心的原因。²⁵“他喝醉了”，“喝”可以省略，为“他醉了”，但是不能省略为“他喝了”。在多个动词连用的结构中，不带体标记的动词也有类似的趋势，与其宾语一起表示一种抽象的行为能力。因而动词与其所带的宾语可以一起处于第一域内。动补结构的处理与本文涉及到的交叉依存类现象紧密相关。

动词中还有一种比较特殊的词，叫做祈使动词。祈使动词表示使令、容许的意思。在域的观点下，对其定界也是需要重新思考的。例如：

（108）良好的投资环境使外商在沪投资的成功率达到百分之八十。

（109）圣经上说：「你们若有信心像一粒芥菜子，就算叫这山移到那里，也没有不成的。」

（110）上海，让你想到什麼？

传统上，这几个句子都称为是兼语句。“使、叫、让”被称为祈使动词，其宾语既是祈使动词的宾语又是后面动词的主语。我们认为，以往认为没有关系的“环境”与“达到”、“你们”与“移”、“上海”与“想到”之间也是存在着关系的。只是这种关系是通过“使”、“叫”和“让”传递的，是一种致使关系。而“环境”与“成功率”、“你们”与“山”、“上海”与“你”之间的关系是间接的，

²⁵ 也有不少学者认为动补结构中的动词才是中心。本文认同的是李临定（1980）的说法，即补语是中心。

有点类似于主语与宾语之间的关系。即虽然都与动词相关，但是并不能说是直接的句法关系。从句法结构的角度来看，则应该是主语与动词的关系并没有仅仅限于施事、受事的语义关系。而以往认为的祈使动词的宾语在语义上才是动词的施事或历事。

在日语中，类似于汉语祈使动词的对应词语是划归为助动词的²⁶，因此可以认为其与动词的关系更加紧密。所以，我们把祈使动词看做是一种关系的标记可能更合理一些。既然是标记的一种，那么则不宜与动词等同看待。因此，可以与名词一起放到第三域内。²⁷在有副词明显定界的情况下，则划归为第一域。

而对于真正的兼语句的处理，则是比较明确的。“ V_1+N+V_2 ”的兼语格式中， V_1 可以带时体标记， V_2 则不可以，所以 V_2 和 N 一起划归到第三域。

- (111) 会议通过决定,任命姜兴长和沈德咏为最高人民法院副院长。(✓)
(112) 会议通过决定,任命了姜兴长和沈德咏为最高人民法院副院长。(✓)
(113) 会议通过决定,任命姜兴长和沈德咏为了最高人民法院副院长。(×)

第一域中另一类需要商榷是是部分趋向动词的处理。

陆俭明(2002)提出了动词后趋向补语和宾语的位置问题。在域的划分中，我们也需要合理解释与恰当处理。

“动+趋+宾”式的处理在域的划归中是比较明晰的。即“动+趋”属于第一域，宾属于第三域。如：

²⁶ お母さんは小明にドアを開けてもらう。てもらう(让)是助动词，需要跟动词同现。

²⁷ 英语中的助动词(Auxiliary Verb)，在疑问句的使用中与汉语的祈使动词有类似之处。助动词的功能本就在于辅助动词构成时态和语态，起区别的作用，自身没有词义，不能够单独使用。从域的观点来看，英语疑问句中的助动词虽然反映时体的概念，也宜看做是一种形式的标记而与名词一样放在了第三域中。

(114) 拿来一本书。

(115) 送上一本书。

可能引起争议的为“动+宾+趋”式。如：

(116) 拿一本书来

(117) 切一个西瓜来

“来”的域所属是值得研究的。如果看作是动词的一种，则会跟域的理论相违背。我们对比以下几个用例：

(118) 拿了一本书来 (✓) (119) 切了一个西瓜来 (✓)

(120) 拿来了一本书 (✓) (121) 切来了一个西瓜 (✓)

(122) 拿一本书来了 (×) (123) 切一个西瓜来了 (×)

“来”放在动词后与放在宾语后是不同的。放在动词后可以加时体标记“了”，放在宾语后则不可以加“了”。

同样，在一些以往认为是复合趋向动词的“出来、进去、上来、下去”等中的“来、去”也有类似的用法。

(124) 拿出了一本书来 (✓)

(125) 拿出来了一本书 (✓)

(126) 拿出一本书来了 (×)

因此我们可以推论，两个“来”是不同的词类。能够加标记“了”的“来”是典型的趋向动词，而不能加“了”的“来”则是虚化了的用法。我们知道，虚化的“来”有助词的用法，表示“已经发生”。如：“你到哪里去来？”语义上表示趋向而不能加“了”的“来”

则是正处于两者之间的一种过渡形式。并逐步向助词的方向发展。因此，我们可以把宾语后的“来、去”看做是类似助词处理，划归到名词后的第四域。

第二类，时体标记

“着、了、过”作为动作的体标记总是附着于动词的，不能够单独做谓语，所以其是与动词作为一个整体来表达事件的变化。我们也可以把它们看做是一种形态标记，跟动词一起属于第一域内部。

第三类，介词

另外，框式介词类的现象中，前后两个介词及其支配的成分与动词一起处于第一域。这是由于介词都是由动词虚化而来的。

这里所说的介词我们采用的是朱德熙先生的观点。“纯粹的介词只能用在连谓结构里，不能单独做谓语（朱德熙 1982）”。一般来讲，介词都是由动词演化而来的，因而很多时候介词其与动词的界限有时候并不是很明显。介词或与其引介的宾语一起，与不带体标记的动词一样，一般处于带体标记动词的外围。这里所说的介词只是介词其中的一个次类。典型的如“在、到、比”等。

(127) 中国进出口银行在日本获债券信用高等级。

(128) 上海计划到本世纪末实现人均国内生产总值五千美元。

(129) 实际外资额也比去年同期上升了百分之三十以上。

介词与其宾语不仅可以在动词之前出现，也可以出现在动词的后面。处于动词面的介词多为从古汉语中继承而来的。如“向、自、于”等。如：

(130) 中国高新技术开发区酝酿于八十年代初。

(131) 中国的影响主要来自家庭。

(132) 今日统一的德国正在从经济大国走向政治大国。

介词可以连用，可以同时居于动词的前部，也可以一个处于前部，一个处于后部。连用的两个介词之间有句法依存关系。不带体标记的动词与其带有的宾语，介词以及带有的宾语，是一个类似于词的整体，是可以存在于动词域之内的。动词与其带有的宾语共同表示一种抽象的能力，整体上具有指称的功能。这并不与名词处于动词外域的观点相冲突。

(133) 农村贫困人口由九十年代初期占农村总人口的百分之五十下降到目前的百分之十左右。

介词中有一些被认为是用来引介施事或受事成分的。如“把、被、叫、让、由”等。以“把”和“被”为例，一般认为“把”是引进受事的，“被”是引进施事的。比较典型的用法如：

(134) 区内企业把发展重点主要集中在光电子、计算机、新材料、生物工程等具有一定科研实力的领域。

(135) 防城港、钦州、北海和湛江，被人们称为西南出海大通道的“龙头”，南昆铁路建设给沿海四市带来了前所未有的发展机遇。

这几个词的介词地位是值得商榷的。从题元角色的角度来看，介词的主要作用是给动词不能赋元的名词赋予题元角色。时间、地点等作为事件的旁支成分在进入句子时需要靠介词引介，这是能够理解的。而施事与受事都是动作的主体，如果它们进入句子时也说是需要介词的引介，恐怕难以让人接受。因此，可能“被、叫、让、

由”与动词的联系要强于与施事名词的关系。王力（1958）从历时的角度论述过，“被”的最初作用不是引介名词，而是直接修饰动词的。因而把“被”字句中的“被”统一看做是助动词。这点上有点类似于英语的助动词，只是辅助动词构成时态和语态，并没有实义。李胜梅（2006）则更是指出在“被+V”的“被”词汇意义更加虚化，有类前缀的趋向。近几年中的网络热词“被就业、被增长、被自杀”的用例，恰恰也是“被”字词缀化的有力旁证。另外，陈淑梅（2006）关于鄂东方言被动结构的调查也可以作为“被”其实是词缀观点的一个旁证。

邵敬敏（2006）则明确指出，认为“把”和“被”是介词的看法是错误的。“把”和“被”的作用并不是引进跟动词相关的语义角色，而应该是“标记一种特殊句式，并且显示该句式特殊的语用意义²⁸”。

我们赞同不能把“把”和“被”等看做是介词的观点，但是并不能用语用的标记来解释其句法的作用。我们认为，“被、叫、让、由”不仅是以往所说的引介的工具，而且是一种句法上分割的标记，是为了避免歧义的发生而使用的角色区别标记。如果作为标记，在域的划归上则可以处于第三域，与所区分的名词同属于一个域。

第四类，名词

在汉语中，有一种传统上认为是体词性谓语句的情况比较特殊，体词能够处于事件的中心。朱德熙的《语法讲义》把体词性谓语分为三类，第一类为名词、时间词；第二类为名词性偏正结构；第三类为数量词和数量名结构。体词做谓语是汉语中比较特殊的现象。体词处于事件的中心位置。可以支配名词、副词和助词成分。例如：

（136）今天刚星期三。

²⁸ 引自邵敬敏（2006）“把字句”“被字句”的认知解释，载邢福义主编《汉语被动表述问题研究新拓展》华中师范大学出版社。第298页。

(137) 这个人急性子。

(138) 每人刚好两本。 (语法讲义 102、103 页用例)

在有副词或助词修饰的情况下，体词作为事件中心的地位是比较明显的，其表达的是变化的事件。但是如果没有副词或助词，如：“今天星期三。”“我广州人。”“每人两本”的句子，我们则看做是类似于形容词处理方式的静态事件，是一个静态的描述事件。由于并不都是动词形容词充当谓语的 center，所以用动词域或谓词域的称呼有失妥当，我们可以称之为第一域。

原型的事件结构中，处于第一域的一般为动词。但是发生了事件的组合以后，第一域内的情况会比较复杂。以本文谈到的九类交叉依存现象来看，处于第一域的有动词、形容词和介词。

5.2.2 第二域

汉语中，处于名词与动词之间的部分是事件内副词。在组块的理论中，副词可以分为修饰整个句子，修饰短语或者修饰单个的词 of 若干类。在依存的理论体系中，则认为只是词语之间的依存关系。因而副词只是对单个词的修饰，不存在修饰短语的副词。相应地，我们对副词的看法也是不一样的。副词有若干种类，我们这里谈到的副词则是指跟动作行为关系比较直接的几种。根据袁毓林(2004)的观察，副词的排列次序如下：语气副词排在最前，时间副词排在次前，协同、重复、方式等紧靠谓语动词排最后，总括、限定论元结构中的从属成分的范围的范围副词居中。在袁的论述中，时间副词，协同、重复、方式类的副词基本上属于我们说指的事件内副词。它们在动词的外围，都又比名词在距离上离动词要近。

同时，部分副词也可以作为论述的起点，这样就出现了副词在名词前的句子。这样的副词多表示的是句子之间的连接关系，应该

属于连接域（本文划归为第四域）。如语气副词作为论述的起点的情况：

（139）其实正如许多女性在社会压力下不断追逐外貌百分百，许多雄孔雀也表示自己是不得不然。

（140）也许几年之间，也许还要几十年，以下的情况可能出现。

事件内的副词处于名词与动词之间的例子如本文研究的副词与动词分离类的交叉依存现象，“这个事务所已经打赢了十起案件”。这句话中，副词“已经”是放在动词外围的，名词“事务所”又在更外围。副词“已经”也可以放到动词的后面，为“这个事务所打赢了十起案件已经”。可以解释为本为事件内副词的“已经”移位到了事件外围的副词域的右部。方言中还有“杀多几个人”“我走先”的说法，有点类似于副词的后置。前置或后置在句法上应该是同质的。汉语中一般把前置的副词称为状语，后置的则称为补语，划分为两个不同的语法范畴。刘丹青（1995）则指出，状语和补语在句法性质上应该是一致的，其对立应该是属于语用方面的。这与我们的观点是相类似的。

副词“已经”位于第一域的外部，但同时其句法依存只跟第二个动词“赢”有关。其句法依存存在的原因符合认知上语义优先（semantic priming）的原则。语义优先的原则认为，词义的激活可以跨越许多个中间的链接，因而不必要求词语线性序列上有严格的相邻。同时也遵从了词序上域的原则。

在现代汉语中，副词处于动词后部的情况也较为常见，这种情况多发生在口语和方言中。例如：

（141）他走了已经。（口语用法）

(142) 我走先。 (方言用法)

这类的说法多出现在口语中，同时表达特殊的语用效果。一般来说，副词处于第二域的时候，第一域内的无时体标记的动词可以带自己的宾语。副词修饰语后置的时候，第一域内的动词则不可以带宾语。

由于形容词可以直接修饰动词，在与其修饰的动词共现的时候，这些形容词也应该划入第二域内，类似于副词的功能。例如“突然”。在“这件事突然发生了”中，“突然”即处于第二域。汉语中形容词可充当的句法成分比较灵活，其所处域的判断需要参照其他词类所处的位置。这即后面部分将谈到的域的相对性。

另外有一类现象是在第二域讨论的时候需要注意到的，即传统上认为做状语的形容词。即本文谈到的第二类名词与形容词分离类的现象。

(143) 为突出城市园林绿化的亚热带风光及民族特色，南宁市大量种植棕榈科植物和开花乔木、花灌木、彩叶植物以及地被植物、草坪等，市区的一百多条主要街道实现了绿化、彩化、果化、香化。

这些词语是不是都处于第二域是值得讨论的。从语义的角度来讲则是只跟名词发生联系的。因而不适于划归于第二域的范围，而是应该归并到第三域，跟名词的修饰语处于同一域内。

副词在域的划归中占有非常重要的作用。副词经常是域定界的参照。例如前面谈到的祈使动词和部分引介施事的介词，在有无副词出现的情况下其分析的结果是不同的。

(144) 最近一波波的关厂裁员潮让中年族人人自危。

(145) 最近一波波的关厂裁员潮，已让中年族人人自危。

(146) 这些产品被行家们鉴定为具有国际或国内先进水平的高科技产品。

(147) 这些产品已被行家们鉴定为具有国际或国内先进水平的高科技产品。

在无副词出现的(144)和(146)中,我们前面已经指出“叫”和“被”是一种角色区分的标记,与名词一起处在了第三域。但是,在有副词出现的(145)和(147)中,副词明显的定界作用激活了介词固有的动词属性,因而与动词的距离更近,划归为第一域。之所以说是激活,是因为用不用“被”跟句子的表现形式是紧密相关的。可以说“这些产品行家们已鉴定为具有国际或国内先进水平的高科技产品。”但是不可说,“这些产品被行家们已鉴定为具有国际或国内先进水平的高科技产品。”

对于汉语中一些特殊的用法,如词语活用的情况,域的分析能够规避开一些关于词性的讨论而取得正确的分析结果。例如,“功夫熊猫很中国”的句子,名词“中国”的功能有很多讨论,或者叫做名词的活用也好,或认为是词性发生了改变。在域的观点来看,副词“很”作为划界的标准,很容易确定其左边的“功夫熊猫”处于第三域,右边的“中国”处于第一域,因而是相对谓语。

5.2.3 第三域

处于副词域外部的域可以称之为名词域,或者为第三域。名词域内一般为与动词有直接依存关系的名词及其相关成分,在语义上讲则是动词支配的论元成分和相关的名词性成分。名词本身又可以带有自己的修饰语。

第一类,名词及其支配的词语;

本文研究的交叉依存现象中，有三类是由于同属于第三域的词语的分离造成的。分别是名词与量词分离类、名词与形容词分离类和名词与名词分离类。即下面三个原型的例句：

(148) 去年电话机生产一千八百一十一万部。

(149) 南宁市大量种植棕榈科植物。

(150) 张学良在建设中现代化意识较强。

第三域分为前后两个部分。前两句都是有句法依存关系的两个词语分处于第三域的两侧，第三句则是两个名词在同侧的分离。分处于两侧的词语之间是紧密相关的。词语之间的先后顺序也是有严格规则的制约，只有在不发生歧义的情况下可以允许部分的自由语序。

一般来说，名词会作为一个表述的起点。所以表述起点选择的不同，会影响到第三域内名词前后顺序的不同。同时，另外一个重要的制约因素为避免歧义的发生。为了区别不同的名词成分在事件中扮演的语义角色，有的语言采用给不同的成分加标的方式，形成了屈折的形式。由于区别的明显，可以允许比较自由的语序。如朱德熙（1985）曾指出的拉丁语中允许极端自由语序的用例。在汉语中，则是通过语序来区分不同的成分，一般把主语放在动词的前面，把宾语放在动词的后面。但是，在没有歧义发生的情况下，汉语中则可以允许比较灵活的语序。例如：“小明知道这件事”，可以改变语序为“这件事小明知道”。如果会发生歧义，则语序比较固定。例如：“妈妈喜欢小华”的句子。“妈妈”与“小华”都是指人的名词，如果放在一起或者前后颠倒就会引起歧义或者与原来的意思相违，所以其语序前后不能改变。我们称之为“无歧义相对自由原则”。这主要是针对书面语来说的，口语中由于有停顿或重音等手段可以加以区分，因而在使用上会更加灵活一些。

不仅每部分内部的相互位置是非随意的，分别处在两侧部分的词语之间的相对位置也是非随意的，有一定的限定条件。曾经引起过广泛争论的“王冕死了父亲”的句子，即是这种情况的一个代表。文炼、胡附（2000）曾指出，“王冕”与“父亲”之间的领属关系是根据不及物动词的特点获得的，这不能类推到动词为及物动词情况下的灵活使用现象。我们认为，避免歧义是这里最重要的原因。“死”作为一价动词，其直接相关的名词成分为“父亲”。“王冕”与“父亲”之间是一种修饰关系。虽然放置到了动词的两边，但是同属于一个域，他们之间的这种句法依存关系依然是人首先感知到的。同时，由于同属于第三域，其位置的先后必须考虑到避免歧义的发生。为了避免与同位结构“父亲王冕”引起歧义，“父亲”不能放在“王冕”的前面，只能放在“王冕”之后，形成“王冕父亲死了”和“王冕死了父亲”两个可能的选择。关于这个句子的主宾语的争论很多，如果我们看做是同属于第三域，加上无歧义相对自由原则的制约，则可以避免这些争论且作出较为整齐的解释。徐枢（1988）在论述“名_受+名_施+动”句式时，曾指出施事为表示遍指的疑问代词，受事为指人名词或人称代词时，不适用这种结构。因为理解为“名_受+名_施+动”和“名_施+名_受+动”两种可能。这基本类似于我们表述的“无歧义相对自由”的思想，只是徐枢限定在了极小的范围内。同时，受事为周遍性疑问代词时，只能放在动词的前部，这也是为了避免歧义的发生。例如“他哪儿都去过”中的“哪儿”不能放于动词后。如为“他都去过哪儿”则是正常的疑问表达。因此，两个位置的不同是功能分化的结果。周丽颖（2007）论述的否定形式的周遍性成分也不能后移的观点却是值得商榷的。极小量词表完全否定时，并不总是置于动词之前。既可以说“他一根烟没有吸过”“我一句话没说”，也可以说“他没有吸过一根烟”“我没说一句话”。

名词之间可以有的关系只有并列、偏正、无关三种可能，因而

在语序上也会反映着三种关系。从语义上来讲，比较复杂的为偏正关系，可以有所属、质材、用途、数量、时间、处所等关系。

有这样一个句子：“那本书小明看过了”，一般认为，这样的句子为主谓谓语句。“那本书”为大主语，“小明”为小主语。在这个句子中，“那本书”与“小明”的前后位置关系是不能够改变的。“那本书”与“小明”之间并没有所属的关系。如果说成是“小明那本书看完了”，则语义上可能为“小明”领有“那本书”。所以，为了表达确定的语义，两个名词的位置关系不能改变。这是因可能与所属关系产生歧义而位置固定的例子。还可能与同位关系产生歧义而不能改变语序的例子。如“李白玲你该有印象吧”不能改变语序说成是“你李白玲该有印象吧”。“你李白玲”之间可能形成同位的关系，因此两者之间的词序不能自由改变。同时，我们还有另外一个句子，如“他酒喝多了”，这个句子中的两个名词的位置却是比较自由的，同为主谓谓语句，但既可以说“他酒喝多了”，又可以说“酒他喝多了”。这是因为“他”与“酒”之间不会形成领有的偏正关系，在语义上无直接关系。所以其前后的位置在无歧义的情况下可以相对的自由，可以根据表述起点的不同自由选择。

另外，有依存关系的两个成分分处于第三域不同侧的时候，名词可以另外添加新的修饰语，使得不能整体移位到后部充当宾语性成分。如以下两个句子所示：

(151) 去年电话机生产一千六百一十一万部。

(152) 去年这种类型的电话机生产一千六百一十一万部。

语义联系的形成是由社会决定的，人们在社会现实中会感知这种联系，而且在人习得语言过程中也在不断的强化这种联系，因而两个词语的同现会激活他们之间的联系。用网络的观点来解释即为两个结点之间连通的权重较高。从上面的两个例子来看，不仅两个

词语的共现会激活词语之间的联系，而且连接的方向对激活也是有直接影响的。这与 Ninio (1994; 1998) 对其他语言的调查结果是一致的。

副词可以作为第一域与第三域的分界。一些貌似动作句的句子通过副词的位置判断，可以确定其身份的不同。例如：

(153) 这种药吃着很苦。

(154) 这个西瓜吃着很甜。

副词“很”作为划界的标记，属于第二域。那处于副词两边的分别为第一域和第三域。“这种药吃着、这个西瓜吃着”都是属于第三域的。可以推论，这里的“着”并不是一般意义上的时体标记，而是表示动作是一种手段。

在前文综述中也曾经论述过，主谓谓语句应该是一个形式的标准。但是现在对主谓谓语句范围的界定上都参照了语义。由于语义的复杂导致大家的观点不尽相同。形式的定义与语义的划分是相悖的，理论上即存在着不统一性，所以在主谓谓语句上始终不能达成一个共识。

名词域中还包括传统分类中的时间名词、地点名词，作为事件发生的时间地点的说明。这些成分在传统上也称之为非论元性成分。在汉语中，所谓时间名词、地点名词的分类其实是值得商榷的，在对于具体的词汇判断上有时候会处于两难的境地。例如，对“北京”这个词语是专名还是地点名词的判断上，是有争议的。我们认为，类似于“北京”这类的词语，只划归为普通的名词即可，至于其指称的对象作为地点的功能是由其社会属性决定的，在句子的上下文中得到激活。正如我们在名词之下没有再区分水果名词与动物名词作为平行的子类一样，另外指明其功用的子类的做法在分类系统性上是不够完善的。其在句子中功能应该是由现实决定的，其与动词

的关系种类才应该用地点或时间这样的分类标准来表示。这与胡裕树、张斌（1999）的观点在本质上是相通的。“常见的现代汉语语法书都在名词下边列举方位词、时间名词、处所名词。这几类词不是用同一个标准划分出来的，当然不宜并列。”

第二类，动词和形容词；

动词、形容词也可以处在第三域的范围，这不仅限于一部分通常所说的名动词和名形词。如“调查、困难（朱德熙 1985）”在下面两例的用法：

（155）调查很重要。

（156）我不怕困难。

一般认为，这些词语本身兼有动词和名词的性质，但也存在部分的争议。在英语中，动词必须转变为不定式的形式或者动名词的形式才能出现在主宾位置上，不定式和动名词这两种形式都属于是名词的范畴，这跟英语中主宾语都要求是名词性成分是一致的。而在汉语中，不带时体标记的动词放在第三域实际也是名词的用法，并不是词性的改变，或者转化为名词性成分，而是域的相对性决定了其语法功能。动词表达的事件作为一个指称处于第三域。如：

（157）建筑是开发浦东的一项主要经济活动，这些年有数百家建筑公司、四千余个建筑工地遍布在这片热土上。

另一种需要注意的现象则为下文中将详细论述到的词序交换后形成的名词性指称。语序的改变在这里承担了改变功能或语义的作用，我们可以看做是一种句法表达手段。例如：

(158) 上海浦东开发与法制建设同步

“浦东开发”是“开发浦东”这个动词短语交换词语的位置后形成的名词性短语。语序改变形成名词性短语主要是针对动词可以与名词形成动宾关系的词语之间。如果“V+N”式组合不会形成动宾关系，则允许“V+N”式的偏正名词短语，V与N之间是修饰关系。

第三类，关系标记；

第三域当中的特殊情况还有对“被、叫、让、由”几个介词和“叫、让、请”等祈使动词的处理，在前文已经有了较为详细的论述，即看做是一种关系的标记，与名词一起处在第三域中。

第三域中的名词大约包括了以往所说的主语和宾语。以往人们认为宾语名词与动词的语义距离一般要小于主语名词与动词的距离。其原因是主语往往可以省略，而宾语却不可以。我们认为，省略并不能作为判断语义距离远近的证据。恰恰是因为主语在人们的认知网络中凸显程度较高，认知较容易，所以才经常省略。而宾语往往是说明的重点才不能省略。这是交际的实际相关的要求，应该属于语用方面的原因。从依存的角度来看，两者并没有距离上的绝对远近关系，可以一并归入第三域内。

第四类，助词“的”；

名词域当中需要注意助词“的”形成的名词性结构。“的”作为一个助词，除了可以看做是偏正关系的标记外，还具有让不同的词转变为名词性结构的功能。这有点类似于 Tesnière 所论述的“转位”。朱德熙（1982）指出，助词“的”的作用是附加在名词、人称代词、形容词、动词以及各类词组之后组成名词性结构。“的”在不同的词类之后可以表示不同的具体关系，这需要根据现实的不

同作出不同的理解。

(159)浦东开发开放是一项振兴上海,建设现代化经济、贸易、金融中心的跨世纪工程,因此大量出现的是以前不曾遇到的新情况、新问题。

(160)但这种法制紧跟经济和社会活动的做法,受到了国内外投资者的好评。

5.2.4 第四域

在名词域的外围,还有一个连接域,也可称之为第四域。连接域内部的词语的作用是表示一个单句与前后单句之间的连接或者一个句子中前后分句之间的连接。第四域内的词语可以是连词、副词等。分句之间的关系是逻辑层面的,“不是词组和词组之间的关系²⁹”。

第一类, 连词

连词经常是成对的出现的。表达两个或两个以上事件之间的逻辑关系,有并列关系、先后关系、因果关系、条件关系等。一些句内的连词应用则是逻辑关系在事件枝节成分中的体现,如连词“和、或”等。

连词既可以处在连接域中,又有可以处于其他域的情况。例如:

(161)假新闻虽然为数甚少,但影响极坏。

(162)虽然假新闻为数甚少,但影响极坏。

两个“虽然”所处的位置不同,会引起句子表达作用的不同。

²⁹ 引自朱德熙《语法讲义》1982,北京:商务印书馆,页215。

在第一个句子中，“虽然”处于连接域，因而表达是两个分句之间的连接，其句子的主语可以不要求一致。在第二个句子中，“虽然”处于第二域，表达的则是事件内的逻辑关系，因而其主语（或者称之为表达的主体）必须是一致的。

我们可以通过分别添加不同的主语来验证句子的合法性。

(163) 假新闻虽然为数甚少，但影响极坏。 (√)

(164) 假新闻虽然为数甚少，但这件事影响极坏。 (×)

(165) 虽然假新闻为数甚少，但影响极坏。 (√)

(166) 虽然假新闻为数甚少，但这件事影响极坏。 (√)

第二类，副词

处于第四域内的副词也可以是事件内的副词，是同一个副词的不同投射结果。需要特别讨论的是，这些副词放到连接域中，与存在于其他域内的现象的作用的差异。例如：

(167) 忽然老陈大笑起来，大家都很惊愕。

(168) 这位公主性格不明，后来忽然变成为另一寡妇，又突然变为奸细刺客，实在交待不清。

第一句是构拟的句子，其中的“忽然”处在连接域，连接修饰的是两个单句。表示的是整个事件的突然。第二句是语料库中的用例。副词“忽然”处在第二域，其修饰的是动作的性状，是指“大笑”这个动作本身的突然性。

5.3 域的交错

由于汉语中语序的句法功能极其重要，所以其语序承担的语法

功能也会比形态丰富的屈折语更多也更加精细。汉语中可以通过域的交错来表达不同的句法功能。平面图中域的交错在线性序列上即表现为词语前后语序的颠倒。在这一节，我们将探讨汉语中相同的一组词语在词序颠倒情况下的句法功能差异，这是汉语中比较特殊的一种句法现象，而且在前人的研究中探讨的较少。

陆俭明（1980）曾经考察了汉语中各种易位现象。陆文论述中的易位与我们论述的语序改变有许多相似之处，但也有一些不同。

首先，论述的立足点是不同的。陆俭明（1980）的论述是以短语结构语法为立足点的，因此其讨论的范围是组块之间的易位。例如整个谓语与主语的易位等。而我们讨论的则只是词类之间的语序改变，不同于易位的所指。特别是一些包括句末语气词的句子主谓易位，更是在我们讨论的范围之外。出现了句末语气词的句子主谓易位现象，我们更倾向于看做是一种追补结构，排除在句子结构之外。

其次，陆俭明只是详细描述了各种可能发生易位的现象，并指出其与原句是基本相同的。我们的论述不仅指出其位置的不同，更侧重于指出词类语序改变引起的具体句法上的差异。

我们首先用一个简单的例子来说明同一组词语经过词序前后改变引起的语法意义差异。由动词“盗窃”与名词“汽车”构成的短语“盗窃汽车”是一个动词性短语，但是两个词语颠倒之后的“汽车盗窃”则是个名词性短语。这种格式是在日语的影响下大量出现的现象，但是已不能够理解为是日语的照搬，其已经融合在了汉语的语法系统之中。语料库中有大量这样的用例，如“浦东开发、法制建设”等。这里的“盗窃”是动词还是名词是有不同的看法的。姚汉铭（1980）认为是兼属动名两类，王冬梅（2003）则认为是动词转换为了名词。我们认为这样的处理是值得商榷的。“盗窃”在这里并没有改变词性，依然是动词。只是通过与“汽车”的语序颠倒实现了表达名词性指称意义。语序在这里充当着改变语法意义的

角色，“盗窃”是无争议的动词。Kwong Oi Yee (2003,2004) 通过语料库调查发现，香港和台湾的语料中有 15%的动词具有动名范畴的使用转换，北京的语料则有 18%的动词有动名范畴的转换使用。可见，动词用作名词是汉语中一种普遍的现象。

把句法关系附着于语序可以避免硬性改变词类范畴的尴尬。这样的处理方式对一些有争议的词类划分也有参照意义。如“研究方法、没有调查”中的“研究、调查”是不是动词的争论。朱德熙(1980)在《现代汉语语法研究》中谈到歧义分化时论述了使用词类或语法关系来分化的两种处理方法，但参照其《语法答问》中对“调查、研究”的论述，他依然认为这些词是名动词而不是绝对的动词。在 Tesinière 的依存理论中，曾经谈到了词语之间的“转位”关系。通过功能词可以改变一个词语的句法范畴。Tesinière 曾经把转位关系作为与“依存、并列”相并列的三种语法关系之一，只是在后来的依存理论中没有受到应有的重视。在汉语中由于形态的缺乏，很多时候是语序承担了类似转位的语法功能，还有一些功能词的使用如“的”等。这个重要的语法关系应该在依存理论中得到合理的继承和发展。我们知道，汉语语法学界曾经有过关于汉语中的词有没有词类分别的争论。后来还有很多关于兼类词问题的争论。无论是《马氏文通》中论述的“词无定类”的观点，还有后来对一些词语兼类问题上的争论，其实都有把一部分本来由词序表达的句法功能强加到了词本身上面去的因素。

并不是所有的词语交错之后都形成名词性的短语，经常有两可的状态存在。这是由汉语的特点所决定的。正如吕叔湘(1979)所指出的，“由于汉语缺少发达的形态，许多语法现象就是渐变而不是顿变，在语法分析上就容易遇到各种中间状态。”但是交错必然产生功能的差异。这里我们区别词序的两种语法意义：一种是动态的表述，一种是静态的指称和描述。这是汉语中语序承担的另外两种重要的语法功能。如前文谈到的比喻，这类似于摄像中一般场景

的拍摄和特写镜头的处理。一般的拍摄是对事件全景的交代，特写镜头关注的则是静态的局部，因而具有强调的意思。类似地，在静态的表述句中，在功能中也具有强调的意味。这是与人的认知习惯相同的。

前面的例子中，“盗窃汽车”即是较常用的动态表述，受动词“盗窃”支配的对象“汽车”放在动词之后的位置，是一个动作事件的表述。而在“汽车盗窃”中，受“盗窃”支配的“汽车”放在前面，词序的变化使得其指称的功能增加，因而成为了一个名词性的短语。值得注意的是，在汉语中两个名词之间的关系如果不能单纯靠语义给以明显的区别时，词序就比较固定。例如，“生物化学”与“化学生物”是两个截然不同的概念，两个词的前后顺序就不能随便颠倒。再比如，“鲁迅研究（某事）”与“研究鲁迅”中，“鲁迅”与“研究”的语义关系是不同的，不能通过语序的改变来表达相同事件的名词性指称作用。这里有个基本的原因就是要避免歧义的发生。在无歧义的情况下，在句子中词序可以相对的宽松，可以有不同的使用情况，以适应不同的表达需要。这里要注意区分句法与语义两个层面。句法层面是从宏观词类之间来谈的，而语义则是结合具体的词语来看的。因而，语义层面避免歧义的要求是从属于句法层面之下的。

我们按照论元成分与非论元成分的次序来分别讨论语序颠倒的语法意义的区别。首先讨论论元成分的语序改变，然后再讨论非论元成分的语序改变。

第一种，论元性成分的语序改变。论元性成分的改变包括两部分，一个是动词的主语相对于动词语序的改变，一个是动词的宾语相对于动词语序的改变情况。这里使用的主宾语是个比较宽泛的概念，因为不仅仅限于动词的施事或者受事。

前面的举例“汽车盗窃”即属于动词的宾语相对于动词语序改变的情况。“盗窃团伙”则属于动词的主语相对于动词语序的改变

的情况。如果施事受事都出现，其正常的表述结构应该为“团伙盗窃汽车”。为了表达不同的语法意义，通过语序的改变区别了这些不同的语法意义。以这个例子为参照，我们可以看下面这些比较整齐的现象：

A	B	C	D	E
汽车盗窃团伙	汽车盗窃	盗窃团伙	团伙盗窃汽车	盗窃汽车
语法学习者	语法学习	学习者	某人学习语法	学习语法
足球爱好者	足球爱好	爱好者	某人爱好足球	爱好足球
项目管理者	项目管理	管理者	某人管理项目	管理项目
节目主持人	节目主持	主持人	某人主持节目	主持节目

从 A、B、C 三组语料来看，都属于名词性的短语。其中的 C 组的后四个用例则干脆可以看做是单个的名词。其语序都是把常用的动态表述语序的前后换位。D 组的语料则是正常情况下的动态表述形式。E 组则是动词短语。通过几组语料的平行对比可以看出，这些词的词性并没有改变，仅仅是改变了词序，但是整个动词短语的功能变化成了名词性的短语。通过语序的调整就把动词短语与名词短语分化开来，把动态的事件表述与静态的名物指称分化开来。

需要注意的是，在 A、C 两组语料中，短语的中心语为表泛指“团伙”、“者”和“人”，而不能是定指的词语“某人”。前置的动词“学习、爱好、管理”等在这里起到了分类的作用，因而区别于动词短语中具体指称的名词。从这点看，这种情况下放在动词后面的施事并不是定指的。

可以简单小结，无论是施事还是受事，都可以通过与动词位置的改变来表达不同的语法意义。都具有动态表述与静态描述的两分的区别。但是，受事成分与动词位置的语序改变并不是完全随意的，还有其他的限制因素在里面。前文所举例的“汽车盗窃、足球爱好、

语法学习”等例之所以能够这样改变语序，还可能跟韵律有紧密的关系。单音节词构成的“偷车、爱球”类组合则不能够做类似的改变。鉴于本文研究的中心没有涉及韵律，有待于另文研究这个问题。

第二种语序改变为非论元性成分语序的改变。非论元性成分主要包括定语、状语。其语序也可以做类似于论元性成分语序的改变，其跟论元性成分的改变具有类似的语法功能。

同意定语后置的人所认同的句法格式一般认为有四种（申小龙1986）：

第一种即为数量词后置。如：“山上有镜面白石一块。”

第二种为动词性结构后置。如：“我有本事叫凤丫头弄出来咱们吃。”

第三种为形容词后置。如：“荷塘四面，长着许多树，蓊蓊郁郁的。”

第四种为的字结构后置。如“旧小说的另一优点，和前一优点有连带关系的，是典型人物塑造的精妙。”

我们把二三四类中用插入语形式出现的定语后置现象排除在我们讨论的范围之外，因为插入语可以是多种成分，不仅仅可以是定语性成分，也可以是其他的词类。这样就只剩下第一种数量词后置的现象为标准的定语后置的现象。

数量名结构是汉语中常用的结构，其一般的语序为“数+量+名”。同时，这种结构的语序在一定情况下也可以改变，同样也具有表述与描述的区别。“数+量+名”为动态的表述，“名+数+量”则为静态的指称与描述。这两种结构表达的逻辑语义是一样的，但是具体的句法和语用功能是不同的。我们在生活中经常会听到或看到这样的用法。

（169）白酒一斤

(170) 大枣十五枚

(171) 陈皮六克

(172) 柴胡两支

(169)句的说法在餐馆服务中经常听到,(170)、(171)、(172)的用法在医生开的药方中更是常见。静态的描述是对局部的放大,因而在语用上具有强调的作用。在实际的使用中我们也可以很清晰的感觉到这种差异。例如,在饭店里顾客会跟服务员说“一斤白酒”,而服务员在向柜台报告时会说“白酒一斤”。两者具有不同的强调点。

在本文研究的非投射性依存现象中,也有一类跟这种现象有密切的关系。“去年电话机生产一千六百一十一万部”的句子中,“电话机生产”为正常表述状态下的“生产电话机”的语序改变。“电话机一千六百一十一万部”为正常状态下“一千六百一十一万部电话机”的语序改变。例句中实现了描述的双重叠加,因而在语用上具有更强的强调作用。换一个角度,我们可以认为,交叉依存的现象中的前于词序的发生是由于其要表达静态描述的意义。这个情况与前面谈到的名词前置于数量词的情况有一点不同之处。其前置的名词位置没有紧邻数量词,而且移位到了动词之前。这样其实是实现两重描述的叠加的需要,即“电话机生产”与“电话机一千六百一十一万部”的叠加。句子静态描述的意味也更强,这跟我们的语感是相符的。传统的关于主谓谓语句的研究中,有论述到主谓谓语句具有强调作用的观点,这在本质上与我们的观点是相通的。

这些思考如果应用在机器翻译上,就给我们提供了一个可能的处理思路。例如,我们可以在表示出一个句子的逻辑语义的同时,也表示出其语序传递的描述功能的层级。描述层级越高,在相应的目的语中也需要相对应的强调句式出现。就像“去年电话机生产一千六百一十一万部”的例句,其逻辑语义我们可以表示为:

生产 {去年 电话机 一千六百一十一万部}

而两重叠加的静态描述语序也增加了强调的效果，我们需要标注作为翻译的重要参照指标。例如，我们可以标记为描述层级 <2>。

在翻译为英语时，就需要使用强调层级较高的强调句式。如 “It is 16,110,000 telephones that was produced last year.” 而不是翻译为一般句式的 “They produced 16,110,000 telephones last year.”

副词与形容词组合时的语序也是可以颠倒的，前后语序的改变同样有表述与描述的区别。例如：

(173) 很精彩 精彩得很

但是这样的用法的对比研究是关注比较少的。副词与动词的语序改变则相对普遍一些。特别是在口语中，副词后置于动词的现象更多。例如：

(174) 已经吃了 吃了已经

(175) 曾经去过 去过曾经

但是，副词前置于动词的情况依然占据主导地位。这是由于副词多是由动词演化而来的，经历了一个所谓的语法化过程。在前文中已经有论述。

由形容词与名词语序可改变的现实，对照前文论述到的动词与名词语序的改变，以及本文论述的其他词类语序调整的现实，可以看出这样一个趋势：即词法与语法是一个统一的整体，并没有截然不同的词法与语法构成的区别。这从一定程度上符合了 Hudson 关于语言网络的观点，由此来看，网络观可能是更加符合人类语言的本质的。

语序与语言的句法、语义、语用三个不同的层面都有紧密的联系。语序是这三个层面的表达所借助的共同工具。正如刘丹青(1995)在论述语义、句法、语用三者关系时指出的，“三者如相缠的三根老藤，难分难解，甚至如三团黄泥捏成的泥人，你中有我，我中有

你。”语序是三个层面互相沟通的共同工具，因此也是争论最多的成分。语序的改变必然在三个层面有相应地差异。语序改变引起的语义的不同是最容易察觉的。但也存在大量语序改变但语义关系基本不变的情况。我们用无歧义相对自由的原则来解释。严格来说，凡是语序的不同在语义上总是有差别的。正如朱德熙（1987）所指出的，“事实上不变的只是低层次上的语义关系，高层次上的语义关系由于句子的结构不同不可能没有差别。³⁰”语序改变引起的句法关系的不同则较语义层面难以察觉，以至于人们把部分句法关系的改变附加到词性之上。句法关系的改变是个渐进的过程，因此会存在很多中间的形式。我们区分了两种宽泛的句法差别--表述与描述。描述的极点可以固化为一种指称，成为名词的范畴，进而影响到具体的词性。至于语序的改变引起的语用差异，则是非常广泛的，且具有一定的不确定性。一些细微的语用差异，甚至只有部分人能够察觉。所以，用语用来解释的很多现象其实都在一定程度上具有主观性的。可以看出，语序在三个层面体现的广泛性是有差异的。语用最广，句法次之，语义最少。因而可以用不同的语序表达相同的语义。我们可以用图的形式来看语序的三个层级，语用是处于最上层的，然后分别是句法和语义。如下图所示：

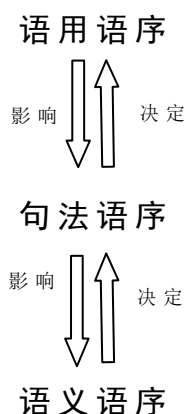


图 7 语序三个层面关系示意图

³⁰ 朱德熙，变换分析中的平行性原则，《句型和动词》语文出版社 1987 年版，页 7。

三个层面有自下而上的决定和自上而下的影响两个过程。

语义向上的决定过程是直接的，语序改变引起的语义关系不同，则句法关系必然不同，因此决定了在语用上不具有可比性。语序改变但语义关系基本不变的情况下，句法关系可以大致相同，但语用效果必然不同，或者句法关系也可以不同，则语用效果也不同。

反过来，自上而下的影响过程则是渐进的间接过程。语序改变引起的语用效果的不同，会逐渐影响到人们对句法关系的认知。如人们对于受事做主语的争论。句法关系的不同，也逐渐影响到语义的分化，如前面谈到的词类转化的问题，义项的分化与句法关系差异的固定也紧密相关。

Connolly (1991) 认为有时候语序改变不影响句法，而只影响语用。我们认为，总体来看是正确的，其指称的是我们谈到的中间状态。但是，严格来说语序的改变还是在句法上有细微的差异。在汉语中，语序的改变是有对应两个层面的，一个是句法的层面，一个是语用的层面。前面谈到的几个例子：

(176) 盗窃汽车 汽车盗窃

(177) 一千六百一十一万部电话机 电话机一千六百一十一万部

第一组中句法结构的改变是很容易察觉的，而第二组句法结构的改变不易为人察觉。我们发现，第二组中的两个短语可以添加的修饰成分是不同的。可以说“仅一千六百一十一万部电话机”，不可说“仅电话机一千六百一十一万部”。语序的改变在总是有差异的，即有句法层面的，也有语用层面的（在语义上可能会不变）。笼统地用语用来解释语序的改变往往掩盖了句法改变的重要性。

以往的研究中经常会提到，依存语法可允许较灵活的语序，并把它作为依存语法一个重要的特点来看待。对这个特点，我们需要

从几个层面来理解。第一，这个结论是来源于部分屈折语的，对汉语来说是否正确是需要慎重考虑的；第二，允许灵活语序最突出的表现是在句法关系表示的时候，在用依存形式表示句法关系的时候可以不考虑语序，这是因为语序表达的句法关系在依存表示中明确化了；第三，在依存分析的时候，这个标准则是从宽泛的角度来说的。对于汉语来说，“一千六百一十一万部电话机”与“电话机一千六百一十一万部”的句法结构认为是相同的，即是从宽泛意义上来说的。细微的句法关系是有差异的，但是还是处在演变之中。只有到了类似“盗窃汽车”与“汽车盗窃”的阶段才有了质的差异。本文根据汉语的现实指出这种句法差异是系统性存在的。

句法上语序的调整是与语用上的强调紧密结合在一起的。胡裕树、范晓（1985）很早就论述过句法与语用的关系是互相制约的。“语用离不开句法，任何语用上的东西，都是附丽在一定的句法结构上的。”这里其实暗示了语序是一些语用的前提，即只有句法上能够提供区别不同语用效果的形式，才能够达到区别效果的作用。这些形式在不同的语言中是不尽相同的。同样是表达强调，有些语言可能主要通过语序的变化，在其他语言中可能主要通过形态。如日语中“wa”是话题的标记。考察这些形式与语用之间的关系，可以加深对某种语言的理解。

语用的解释能力是需要加以限制的，需要区分不同的解释层面。不应该把所有句法、语义不能很好解释的现象都归咎于语用。有用一定的形式表达的固定的语用。例如强调，反问等。有与具体语境或文化相结合的语用。例如话语与话语者身份是否适当等。范开泰（1985）在《语用分析说略》一文中曾经论述到一些语序对表达内容的重点的影响。“一般来说，这个重点在句子的后面部分，尤其是动词后面的部分。”但其论述是按照心理结构分析中的焦点论来看的，且只涉及到动词相关的一部分，缺乏可操作性。没有把结构与语用的关系固定化。本文的论述则是从词序提供的分化可能性出

发，从相反的通道来看的，希望能对我们以后的研究有所启发。

5.4 域的性质

5.4.1 相对性

语序域中，不同的域对其域内词语的类别具有很强的选择性，但并不是绝对的标准。汉语语序中域的概念是相对的。这与汉语的词性特点有关系。在印欧语言中，其词类与句法功能之间具有对应的关系。作主语的一般为名词，做谓语的为动词，副词做状语，形容词作定语，清晰明了。这样，词性在句子中充当成分的相对位置也会是固定的，因而可以通过外表的词性来判断其所充当的句法成分。（朱德熙 1985:4）但是在汉语中，词性与句法成分之间的关系是极其复杂的，并没有类似的一一对应的关系。因此，通过单纯的词性来判断句法成分的思路就会带来极大的困难。换一种说法，在英语中，我们基本可以通过句法的分析来得到语义的组合方式。而在汉语中的分析则是相反的，往往是首先理解了语义之后才按照类似的方式来作出的句法分析。因此，单纯依靠单个词的词性来判断句法成分的做法在汉语中是值得商榷的。

在汉语中，句法关系的判断必须参照词性之间的相互关系（语义关系和位置关系）。例如，名词在汉语中可以做主语和宾语，还可以单独做谓语，也可以做其他词的修饰语。所以，在判断名词充当的句法功能时候就会有了一种可能。如果我们参照其与其他词的词序，则可以比较容易的判断其句法功能。

（178）今天星期二。

这个简单句子的句法关系是相对简单的，由于第一域的先在性，

“星期二”为谓语处于第一域，“今天”为主语处于第三域。判断的标准其实是按照词语之间的相对位置，而不是绝对的词性。如果加入其他的修饰语后，其判断则会更加明确。例如，今天刚星期二。由于副词“刚”的出现，使得“星期二”的谓语性能变得唯一。

用词序域的理论来解释，增加了副词后的句子格局会更加明晰。“刚”作为事件内副词应该处于第二域，“星期二”后又没有其他词语，所以名词“星期二”是必然处于第一域的。而“今天”则处于第三域。这样作为一个静态表述的基本结构就清晰地呈现出来。

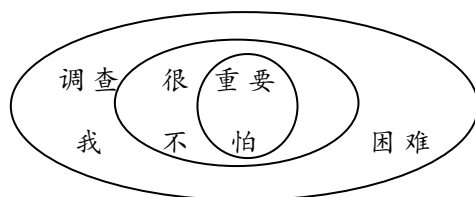
汉语中很多动词、形容词在与名词的划界上存在着争议，朱德熙（1985）曾分出了名动词与名形词的类，来解释这些争议。在我们看来，这种现象体现的正是域的相对性。“历史研究、精神准备、农村调查”中的“研究、准备、调查”与其划归为名动词，不如依然看作是动词，功能的变化是域的交错造成的，分别为“研究历史、准备精神、调查农村”的语序改变，语序承担了句法改变的功能，可以直接在关系表示上表明，而不再附加于动词本身。同理，“经济困难、生命危险、婚姻自由”中的“困难、危险、自由”也是域交错之后形成的句法关系改变，为“困难经济、危险生命、自由婚姻”。同时，划入不同的域，则可以表现出其句法功能的差异。朱德熙（1985）提出的两个句子：

（179）调查很重要。

（180）我不怕困难。

如果用域来讨论，则可以避开关于词性问题的争论，同时可以准确地表示其句法的功能。有了副词“很”的出现，那么动词“调查”就必然相对地处于第三域，形容词“重要”就处于了第一域。第二个句子中“不”的出现。明确了“怕”处于第一域，“困难”处于第三域。由于“怕”不带时体的标记，也可以把“怕困难”作

为整体处于第一域，但这样的处理不影响人们对其语义的理解。



域的相对性不仅可以回避掉一些词类争议的现象，而且可以用来解释一些原本类别明确的词语在句中可以产生功能变化的现象。最明显的例子为离合词的使用。例如，“洗澡”为动词，在“洗了个澡”中，“个”作为量词有明显的功能指向，则“澡”与“个”之间为数量关系。因此在分析中应该做类似于名词的处理，处于词序域的第三域。

同一个词在不同的句中可以归于不同的域，不同的归属反映的是对词语不同句法作用的想法。前例谈的是动词和形容词处于第三域的情况，形容词还有可以处于第二域的情况。例如形容词“突然”，在下面两个句子中即是处于不同的域。第一句是处于第一域，第二句中是处于第二域。

(181) 这件事很突然。

(182) 但是，从1992年起他突然从人们的视野中消失。

交叉依存现象中的几个名词与形容词分离类的例子，也是比较典型的例子。

(183) 南宁市大量种植棕榈科植物和开花乔木、花灌木、彩叶植物以及地被植物、草坪等。

(184) 这个著名的侨乡由于与港澳台地区长期保持着良好的经贸往来和市场联系

“大量、长期”一般认为是状语成分，修饰谓语动词。在域的视角下则应该划归为第三域，为修饰名词的形容词。而修饰动词的非谓形容词如“轻轻地、慢慢地”则应划归为第二域。

语气副词的位置是比较自由的，可以处在不同的域中。其对域的划分并不影响，因为其表达是语用功能。如文炼、胡附（2000）指出的，他们在句子中的地位类似于通常所说的插说语（独立成分），不能要求其句法成分对等，它们充当的是语用成分。

5.4.2 递归性

自然语言具有递归的特性。即可以使用有限的句法规则，通过不断地同形替换，以构成无限的复杂的句子。递归性是语言的本质属性之一，是区分人类语言与动物的交际符号系统的标志，这是大家的基本共识。

袁毓林（2008）指出，线性的语法模型不能够很好地处理语言中的嵌套结构，这是因为语言知识具有“结构依赖（structure-dependent）”的特点。平面域的视角为突破这种局限提供了一种可能。

递归性在转换生成语法中较多地谈到，同样也适用于词序域的理论。谓词与其支配的词语构成一个简单事件结构。事件结构的域之间也具有递归的特点。通过不同的递归方式可以构成无穷的句子实例。递归的方式大致可以分为三种。

第一种是嵌套的方式，即构成一个事件的各域可以整体处于另一个事件域内作为其组成部分。如图所示：

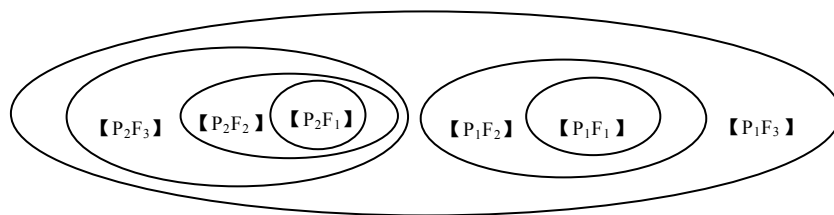


图 8 域的嵌套

两图都是事件结构整体处于第三域，分处于域的前部和后部。
在语料库中可以找到很多这样的用例，如：

(185) 国家统计局预测：全球经济发展将给中国带来很多机遇。

(186) 拥有一个终年不冻的天然深水良港，是水东开发区发展
乙烯后继加工项目的有利条件。

第一句中，是事件“全球经济发展将给中国带来很多机遇”处于另一事件“预测”的第三域，做事件“预测”的一个部分。第二句中，“拥有一个终年不冻的天然深水良港”处于事件“是”的第三域，做事件“是”的一个部分。

第二种递归方式为事件域的叠合。

叠合是对语言空间网络结构的扁平化处理，在线性表现上总是体现为先后的关系，这是受语言线性一维的物理特征所制约的。形式上的叠合，在具体的词语关系上则有若干种不同。可分为并列、拼合、交叉等。这里我们以典型的事件并列作为解释的基础。事件的叠合可以形式化表示为下图的形式：

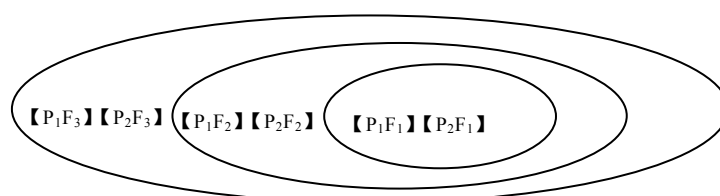


图 9 域的叠合

语料库中的例句如：

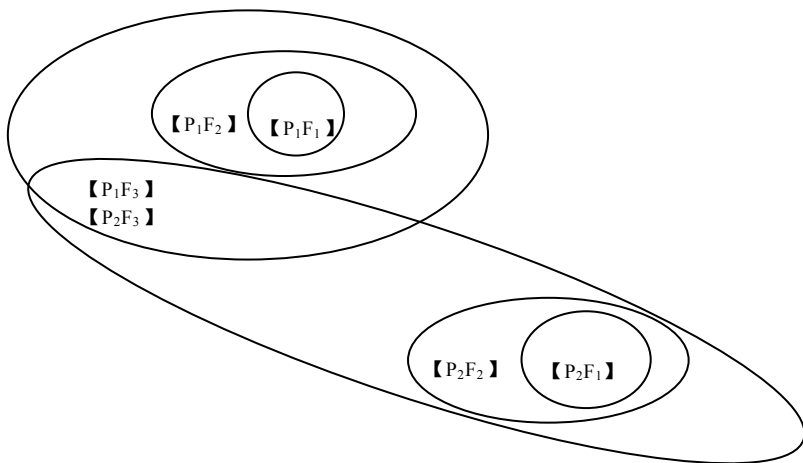
(187) 上海浦东近年来颁布实行了涉及经济、贸易、建设、规划、科技、文教等领域的七十一件法规性文件，确保了浦东开发的有序进行。

(188) 目前中国已发展形成整套规范的煤层气开发技术，有二十多项技术接近或达到国际先进水平。

严格来讲，完全并列的两个事件之间的先后顺序应该是可以调换的。但是，现实中绝对同时发生的事件几乎是没的，总会有哪怕是极小时间上的先后差异，但人们还是认为是同时发生的。这样，一些连续发生的事件也会使用并列结构的表达方式。在形式上称之为叠合。连用的谓词之间也逐渐确定按照时间顺序排列的习惯。能够使用并列组合方式的事件之间的支配成分往往是共享的，也可以说是合并为一体的。

第三种递归的方式为事件域的链接。

链接在形式上既可以表现在发生在单句内，也可以表现为发生在复句之内。都表示的是事件的链接。一种是在前部，一种是在后部。分别如下两图所示：



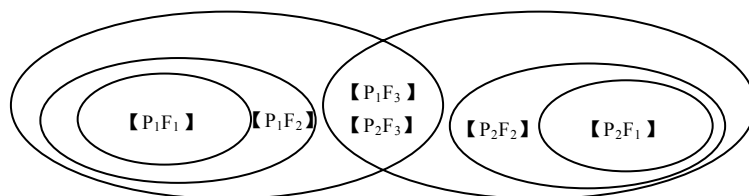


图 10 域的链接

第一种是两个事件的动词前论元的共享，但谓词和其他论元不共享。例如：

(189) 他已于当天傍晚乘船离开布拉柴维尔回国。

事件“乘”与“离开”共享了动词前的论元成分。

另外，还可以是第一事件的动词后论元与第二事件动词前论元共享的情况，但谓词和其他论元不共享。例如：

(190) 会议通过决定，任命姜兴长和沈德咏为最高人名法院副院长。

其中共享的部分为“姜兴长和沈德咏”，为事件“任命”的动词后论元与事件“为”的动词前论元。

5.4.3 拓扑等价性

拓扑这个词语来源于希腊语“ $\tau \acute{o} \pi \omicron \varsigma$ ”（地点）和“ $\lambda \acute{o} \gamma \omicron \varsigma$ ”（研究）的音译，其本意为地貌学。拓扑学主要研究拓扑空间在拓扑变换下的不变性质和不变量。通常的平面几何或立体几何研究的对象是点线面之间的位置关系和度量关系，拓扑学对于研究对象的

长短、大小、面积、体积等度量关系和数量关系都无关。拓扑学是数学中一个重要的、基础的分支。起初的拓扑学属于几何学的一个分支。现在已发展成为研究连续性现象的数学分支。

拓扑概念的一个基本性质即为“拓扑等价”。在一个球面上任选一些点用不相交的线把它们连接起来，这样球面就被这些线分成许多块。在拓扑变换下，点、线、块的数目仍和原来的数目一样，这就是拓扑等价。³¹ 直线上的点和线的结合关系、顺序关系，在拓扑变换下不变。

直线上的点和线的结合关系、顺序关系，在拓扑变换下不变，这是拓扑性质。句子中由词语及其依存关系形成的句法依存树与由词序关系构成的语序依存树之间可以构成拓扑等价的关系。即词语的点与词语之间联系的线段数目在拓扑变换下保持不变。这是德语中句法依存树与语序依存树之间拓扑等价的核心思想。

在计算机领域，“拓扑”还经常用来指称网络的结构类型。具体可以分为星形拓扑、环形拓扑、总线拓扑、分布式拓扑、树型拓扑、网状拓扑等类型。组网的方式可以不考虑结点的位置和结点之间线段的长度，只考虑结点之间的连接方式，是拓扑思想的具体应用。在本质上与由词语构成的句法依存树或依存网是相同的。

依存图中使用的拓扑等价概念，有异于德语中的使用情况。德语中指的是句法树和语序树之间是一种拓扑等价的关系。而这里指的是调整语序后的句法依存图与原图是拓扑等价的关系。以交叉依存的现象为例，我们来看词序域视角下句法依存结构变换与其原结构之间的拓扑等价关系。以交叉依存现象中的名词与量词分离类的句子“去年电话机生产一千六百一十一万部”为例来说明。我们用符号来代替具体的词语结点，可以表示为如下的形式：

³¹ 引自维基百科 <http://en.wikipedia.org/wiki/Topology>

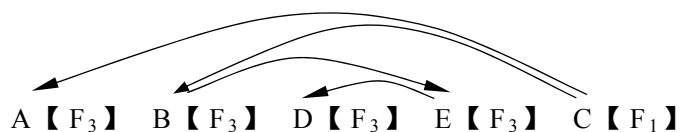
去年 电话机 生产 一千六百一十一万 部
 A B C D E

其句法依存的关系我们表示为如下的形式：



其语序的关系采用给词语加下标的方法标示各词所属的域，可以描述为：A **【F₃】** B **【F₃】** C **【F₁】** D **【F₃】** E **【F₃】**。其中带括号的**【F_x】**表示的是词序上所属的域的不同。

合并同一域的词语结点在同一侧，原有结点之间的句法依存关系保持不变，那么其图形可以转换为如下的形式：



调整语序后的句子的结点数及依存边的总数保持了不变，结点之间的结合顺序关系也没有改变，因此原图与变换后的图形是一种拓扑等价的关系。

就句法分析来说，可以归并相同域的词语与同一侧。这样可以规避开词序分离的情况，从而把一些投射性依存分析的算法和技术平移过来。这是针对单个事件来说的，多个事件构成的句子则需要分解为单个事件以后再归并。

就句法研究来说，也可以发现一些以往没有关注到的语言现象。例如动补结构后半部分与前部分之间的关系，在以往的研究中是很少涉及的。词序域内词语的归并，可以使人明确地察觉到这些句法上的联系。

5.5 小 结

本章论述了环形语序域原则，这是词语在线性位置上的制约原则。环形语序域是语言的三维立体网络在二维空间投射的结果，因而同一个网络允许不同的投射结果，而且语序在一维化过程中也具有一定的灵活性，但都符合环形域的原则。环形语序域原则与第四章的属性和谐原则共同作用才形成了交叉依存的现象。与属性和谐原则相类似，环形语序域原则的提出也是受交叉依存现象研究的启发。但是，此原则同样适用于其他非交叉类依存现象中的语序解释，因而也是一条普遍性的语序原则。

在环形语序域的视角下，我们重新对汉语中一些争议的问题进行了梳理。对主宾定义、汉语的语序类型、主宾互换等问题提出了语序域视角下的解释。结合交叉依存的现象，我们解释了汉语的四个语序域，并论述了域的相对性、递归性和拓扑等价性三个基本性质。

环形语序域原则制约了词语的线性位置，是分宏观和微观两个层面来看的。宏观层面上的词序域位置是固定的，而域内相对位置的微观词序则是比较灵活的。只要在没有歧义的情况下，可以允许一定程度上相对自由的语序，我们称之为“无歧义相对自由原则”。一般认为，支配词能够决定被支配词的线性位置，这是 Link Grammar 的一个基本假设。印欧语的大多情况也是符合这个原则的，但对小部分语序灵活的情况却不适用。从汉语的实际来看，语序承担了更多的语法功能。有依存关系的词语之间的前后位置不是绝对的，经常可以互换，而依存关系的种类基本不变。这只有在环形语序域的视角下才能得以解释。因此，支配词能够决定被支配词线性位置的说法是需要修正的。

环形语序域原则与前一章的属性和谐原则，两条原则总体上来说比较宏观，这是由信息处理的实际情况所决定的。句子的合法不

合法对于语法研究者来说可以取比较严格的标准，但是对于面向大规模真实文本的信息处理来说，则必须要取比较宽泛的标准，以便可以解释汉语中的方言、欧化句式以及灵活用法等许多现象。朱德熙（1988）曾论述到：“语料中包含的层次越是复杂，语料内部的一致性就越低，能够从中归纳出来的语法规律也就越概括，作为规律的约束力就越低。”因此，在遵守这两条大原则的基础之上，还需要根据不同的研究需要做细化的分类研究。

6 依存图库的构建

6.1 相关的理论探讨

本文考察了汉语中交叉依存现象存在的基本情况，归并为不同的两大类共九个小类。从本文的分析可以看到，这些句子的句法依存关系并不符合以往依存表示体系要求的独根节点、单一父节点、单纯结点、非交条件等若干黄金原则。

句法关系的表示目前主要有两种主要的形式，一种是短语结构树，一种为依存结构树。作为语言表示的一种工具和手段，无论是短语结构树或是依存结构树，其实都是不能够被证实或者证伪的。只能以是否有效，是否适合语言的现实来做评价的标准。语言表示的方法是为语法分析的理论服务的，“它的体系、体例、结构、形式，都要和语言分析理论和实践相适应（曹聪孙 1987:210）”。一般认为，基于词汇关系的依存树表示方式要优于基于组块理论的短语结构表示形式：其表示的颗粒度更加精细；表示的关系更为基础；对句法关系的表示更加明确等。而且，依存的表示方式可以用同一种规则来表示更多类型的语言，便于对不同的语言进行对比研究，因而具有更广泛的适用性。在基于机器学习的计算语言学研究中，树库作为机器学习的语言资源基础，近几年来也有从基于短语结构表示方式的构建思路向基于依存结构表示方式的构建思路转变的趋向。

虽然依存树的表示具有以上的这些优势，但是也有其不够精细的地方。其投射性依存的假设忽视了语言现实中由于语序灵活变化而形成的非投射性依存的现象。本文对交叉依存现象的研究也证明了实际的句法关系违反了以往的很多黄金标准。因而，需要一些修

正才能更好地符合语言的现实。对独根节点、单一父节点、单纯结点等标准的违反，会使得最后的呈现结果为图的形式。

使用树的表示方式一方面跟依存语法的研究传统有关，从 Tesnière 开始即使用的是树形的表示方式。其《句法结构基础》一书中使用了大量的篇幅来介绍树型的表示形式。另一方面则是跟计算科学上基于树的计算模式相适应的。随着理论研究的深入，基于图(网)的语法理论已经受到大家关注。同时图形计算技术的发展，也为我们修正基于树的表示方式提供了技术上实现的可能。

使用树形结构表示句法依存关系的句子称之为依存树，由树形表示的句法关系的句子构成的语料库叫做树库 (treebank)。树库作为标注了语法信息的语料库，可以同时作为基于统计和基于规则的两种语言学研究的重要基础。它不仅可以为基于规则的研究者提供尽量大的带有句法信息的研究数据以得到比较客观的语言规则以及验证其理论的适用性，也可以作为基于统计的语言学研究数据学习的基础和分析器评测的标准。因而，树库的构建成为计算语言学研究的一个重要方面。鉴于以往依存树库的标注中投射性假设、独根节点、单一父节点等原则的不足，我们需要采用新的标注形式，即允许多根节点，允许多重父节点，因此是图的形式，我们称之为“依存图”。相应地，基于这种图表示方式下构建的语料库应该称之为依存图库 (dependency-graph bank)。

依存图表示方式，是在依存树表示的基础上演化而来的。使用的是基于树但又不同于树的表示方式，能够更好地表示汉语中的语言现实。目前，虽然也有一些依存图理论的提出，但是主要是在句法分析器中使用的一种基于图的运算方法，例如尝试所有词语之间的组合可能的算法。实际的关于汉语依存图的语法理论和语料库的构建规则现在还没有看到。

现有的很多语料库中同时使用短语结构树与依存关系两种标注体系，如德语的 TIGER 树库和汉语的台湾中研院 SINICA 树库等。

构建者一般认为这种树库可以同时表示层次结构和依存关系，因而具有功能强大、处理灵活的特点。对这个观点我们需要做重新的认识。这实际上是针对短语结构语法体系和依存语法体系现有的不足而使用的权宜之计，并不应该看成是其优势之处。从理论的角度来看，短语结构表达的层次关系在依存结构中也可以很好的体现，而以往认为的依存对于无明确依存关系成分的处理不足的特点可以在我们的词序域中得以修正。因此，我们主张使用纯粹的依存标注形式，而不使用不同的理论体系相结合的方法。

6.1.1 从依存树到依存图

图结构是比树结构更为复杂的一种数据结构。因此，图型的表示方式与树型的表示方式也是不同的。根据汉语的语言现实，我们在现有依存树表示的基础上进行一些修改，来实现图的表示形式。这样的修改不仅是为了在形式上符合语言现实的需要，同时还有以下两点理论的考虑：

第一，从数据结构的角度来看，树型结构中的结点是按照层次连结的。只在上下两个结点之间才有信息交换，相邻结点或同层结点之间一般不进行数据的交换。这就使得结点之间信息共享的能力较低，可靠性不高。从句子分析的角度来看，即可能对上下文中提供的语言信息的利用存在不足。而图的数据结构则没有明显的层级，节点之间可以比较自由地信息共享，因而可以弥补树形结构表示的不足。

以往的依存分析技术，都是基于树的模型提出来的，其假定的几个黄金标准都是以树形计算模式为基础的，这是与树形数据结构的算法比较成熟有直接关系的。但是实际的语言数据并不都是符合这些理想状态下的树形结构的。同时，计算技术的发展已经突破了树形数据处理的局限，能够进行基于图的方式计算。因此，为依存

图的实现提供了技术的可能。同时，技术的发展为提供更符合语言实际的图库构建模式提供了可能。

第二，从语言处理的整体（如：机器翻译）来看，句法表示的实际必须要考虑到后续的语义对应。现有关于依存树表示的黄金标准，假设了语言的表示结构是独根节点、单一父节点，而且需要符合投射性依存的基本原则。虽然可以方便数据的运算，但是跟语言现实的情况是不符合的。因而根据现有的假设表示的树库作为习得数据的分析器分析得到的句法树是不完全正确的。特别是在从句法挂靠语义的过程中，必然会发生错误的叠加信息。

因此，我们可以针对这些特点来推论依存树与依存图之间的根本区别：第一，最直接的不同是在形状上的差异。依存树是树形的，而依存图则是网状的。第二，依存树按照树形数据结构假设了单一父节点、投射性依存等原则，而依存图中的任意两个结点之间都可能直接有直接的关系。所以允许多重父节点，允许非投射性依存，结点之间可以自由的信息共享。

依存图的标注有两个基本的思路：一个是在现有的树库基础上进行修改，把原来语料上标注错误的形式修正为真实的句法依存关系，并且增加漏标的依存关系，如一个词语结点有多个父节点的情况；另一个思路则是新建一个标注了非投射性依存现象的语料库。根据前文的研究可知，真实的句法依存关系将违背几乎所有以往标注中遵守的黄金标准（冯志伟描述为5个条件：独根结点条件、单一父结点条件、单纯结点条件、非交条件、互斥条件）。那么在原有语料库基础上的修改将会是一个巨大的工程。不仅原有的一系列标注规则需要改变，而且在实际操作中很可能还会受到原有标注标准的误导引起标注的错误，在人力物力上都不见得会好很多。因此，我们认为可能选择新建语料库会更为合理。或者是只使用原有的生语料，在原有语料上重新标注，或者是选择新语料新建语料库。

由于新建语料库需要考虑的因素很多，如语料的来源、题材、

文体、语体、领域，以及不同语料的数量比例等。而在这里我们主要是讨论句法标注的具体标准问题，语料本身是否重新选材对现有研究来说关系不大。因此，我们采用在原有语料上重新标注的折衷方法来解释具体的标注规则。

标注的语料，理论上我们可以根据需要的不同进行不同的选择。鉴于宾州汉语树库现在已经有短语结构、谓词论元结构、命题结构、特征结构等多种方式多个层面的标注成果，经常用来作为学习器学习和评测的标准，那么我们可以尝试对此语料库的语料进行依存图的标注。这样，就可以为横向的对比短语结构、特征结构、依存树、依存图几种标注方式提供了基础。而且，可以为基于统计的分析器提供统一的学习语料，然后作为横向比较分析器准确性的一个基础。

众所周知，语料库在现代语言学研究中有很重要的位置。因此，依存图库的构建相应地会有两个方面的意义：

（一）对语法研究的意义

语料库基础上的语言学研究是近年来一个重要的发展方向。语料库为理论语言学研究提供了一个由无穷衍生的语言事实抽样出来的代表性样本，可以作为语言研究的基础。对真实语言交际的各个层面，包括词汇、句法、语义、语用、语篇等进行全面的分析和把握。

语言学家通过观察语料库，可以较全面地把握语言事实，分析和研究语言系统的规律性。避免自我内省发现的规律的主观性和片面性。标注了词性、依存关系、词序域信息的依存图表示，可以为研究者提供更为详尽的语言信息。同时，也可以作为验证研究者提出的基于图的各种理论解释力强弱的基础。

（二）对统计句法分析器的意义

大规模的语料库可以作为使用统计语言模型方法处理自然语言的基础资源。经过分词、词性标注、句法标注、语序域标注等信息的语料库，可以做为句法分析器数据学习的基础材料。同时，也

可以做为对不同句法分析器进行评测的共同平台。依存图库的构建同样也可以为基于图的句法分析器提供一个数据学习的基础和评测的统一平台。

6.1.2 事件作为中间单位

对语言进行分析，就是分析语言的构成结构。分析的过程即是把一个句子分析为很多个较小的单位，这些单位又可分解为更小的单位的过程。短语结构的语法分析，从大到小依次把语言分为“句子”、“短语”和“词”三级单位。这是跟短语结构的语法体系和组块认知的心理学基础相符合的分解方式。在依存语法的分析中，以往的分析则是把整个句子看做是一棵树，是一个“词--句子”的两级构成体系，没有中间单位的存在。这样的处理思路有其优势，即在操作思路上比较简单。可以描述语序相对比较自由的语言。但是，也正是由于在词汇与句子之间没有过渡的单位，在对词语之间无依存关系但有语序制约的现象上就不能够很好地解释与处理。

汉语语法研究史上，曾出现过五种语法研究的本位，分别是马建忠的“词类本位”，黎锦熙的“句本位”，朱德熙的“词组本位”，徐通锵的“字本位”和邢福义的“小句本位”。各种本位都有不可忽视的理据性。目前来看，对词组本位的应用更为广泛一些。根据 Mel'čuk (1988) 的观点，词组只是对句法关系的体现，而不是对句法关系的表示。而且，对于语序自由的语言来说，很难说有一贯的、固定的词组形式。因此，我们采用依存的表示方法。依存可用来表示更多类型的语言。在依存图的思路下，我们则主张使用事件结构作为句子分析的中间单位，形成“词--事件--句子”由小到大的三级处理单位。这里的事件单位是句法的定义。理想状态下的事件结构是指一个谓词及其支配成分的集合。既包括必有成分，也包括可选的成分。每一个事件都可以用一棵小的依存树表示。

以事件结构为中心，向下可以进行事件构成单位----词语的研究，向上可以进行事件之间组合的单位----句子的研究。一个句子的整体则不再像以前认为的必须是一棵树的形式，而可能是一个事件的树或若干个不同的事件树的组合。这些事件通过不同的方式组合在一起，最后形成线性的表层结构。而且，事件在组合的过程中还会激活一些新的句法联系。新句法联系的激活既有一定的确定性，即已经习用的句法关系的激活；也有一定的随机性，即一些新用法和新的句法关系的产生。随机产生的新句法关系在达到一定的临界点后会以爆炸性的速度拓展，从而可能引起整个语法系统不可逆的改变。这既是语言最难把握的部分，同时也是语言最具魅力、最为奥妙的部分所在。

把事件结构作为语法研究的中间单位是根据汉语的实际来看的。印欧语中如果有两个或两个以上谓词的出现，一般都会有明显的形式标记来区分开主要谓词和非主要的谓词（如：在英语中，动词如果不是主要谓词的情况一般会使用不定式形式或动名词形式）。因此，印欧语中把一个单句作为一棵树来表示是有形式上基础的，即形式上的理据性。而汉语中不仅无这样的形式标记，而且可以把复句紧缩为一个单句的形式，较之印欧语会更加复杂。因此，我们主张把事件结构作为中间单位，可以起到以简驭繁的功能。

事件内部各词语之间的主从关系是比较明晰的。Tesnière（1959）称之为词语之间的层级，可以用图示的形式表示如下：

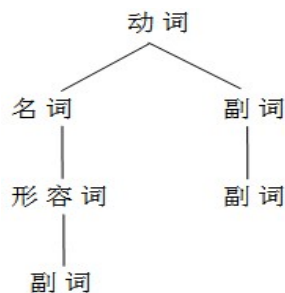


图 11 词类依存关系层级

而组成复杂句子的事件之间却没有明确的主从关系。事件的先后排列大致只是时间上的线性关系(还有很多事件的排列并不符合时间序列的原则),不是句法上的主从关系。因此,时间规则只能是一个大致的描述,而不能作为判断的准则。与此相类似的观点如汉语中关于焦点后置和强调前置的两种说法,也只能是大致的描述,不能作为判断的规则,其在逻辑上严格来说是矛盾的。再如对动补结构那个成分是中心的争议。“动补”这个术语本身即是认为动词是处于支配地位的,其后面的成分是其补充作用的。而也有学者(李临定 1984)通过动词成分可省略的现象来论证其实后项“补”的部分才是最重要的,应该是该结构的中心。所以,复杂的事件如果从不同的角度来看,特别是在不同的交际环境中,都可以充当中心的作用。因而对汉语这类无形式标记的语言区分事件的主从关系是无解的命题。比较合理的做法是把各个事件平等地看待,然后探讨其组合的方式与特点。

把事件作为研究的对象是近几十年来词汇语义学、逻辑语义学、句法学共同关注的一个重要话题。词语语义学的研究从原来主要关注动词本身的语义,向外扩展到关注动词语义涉及的各个部分在句中的分布情况;句法的研究也不再把事件作为一种不可再分的基本单位,开始向内关注事件结构从词汇语义到句法的映射关系;逻辑语义的研究也从关注动词的语义分解表示扩展到对句法与事件结构之间的映射关系表示。

Davis 和 Demirdache (2000) 以及 Ritter 和 Rosen (2000) 都曾经独立地提出,存在着跨语言不变的事件结构,不同的语言通过激活或者突出其不同的部分来对事件进行具体的实现。因而,事件结构是不同语言之间能够形成横向对比的共同基础。研究事件结构与句法关系之间的映射关系是目前众多学者努力的一个方向。对句法与事件结构之间映射关系的研究,目前大多采用的是短语结构的

句法体系。大家比较共识的一个观点即是，句子在短语结构层面的表现在很大程度上是由事件结构生成的。但是，这并不意味着句法结构与事件结构之间映射关系的研究必然限定在短语结构的形式上。既可以用短语结构的语法来研究，也可以使用其他的句法体系来研究。我们的研究则是从依存语法的角度来看的。

对事件这个概念的定义以往主要是对动词来说的，且在学者们中间存在很多的分歧。英语中的句子都是有动词做谓语的，因此，Vendler (1957) 根据时态把英语的动词分为 *state*、*activity*、*achievement* 和 *accomplishment* 四类，这也即是对事件的四种基本分类。Pustejovsky (1991) 则把事件分为 *state*、*processes* 和 *transitions* 三类。根据汉语的现实，不仅动词形容词可以做谓语，一些名词数词等也可以做相对谓语，因此我们把事件概念做宽泛意义上的使用。即：谓词与其所支配的成分共同构成一个事件单位。可以把事件分为两个大类，一类是行为性的事件，即以动词为核心的事件；一类是描述性的事件，是以形容词或者相对谓词为核心的事件。每一个事件都可以用一棵小的依存树来表示，树与树的结合则构成了图，即依存图的形式。事件与事件结合在一起，构成的是复杂事件。相应地，在事件的结合中，有三种基本方式：一种是行为性事件的结合，一种是描述性事件的结合，第三种为行为性事件与描述性事件的结合。这是数学上简单的组合关系。行为性事件与描述性事件结合时，行为性的动词倾向于跟描述性的谓词结合在一起，构成一个复合事件。复合事件是复杂事件中的一个子类，类似于事件语义研究中论述的动词分解为外动词和内动词的结合。这样的组合符合事件发展的过程，与汉语的时间理据性相符合。行为性动词与描述性谓词强制结合时，动词的论元和描述性谓词的论元的位置会做出调整。由于这样的调整，使得事件在结合时呈现出不同的类型和模式。

这样，我们可以把句子的标注分为两大类来看，一类是单个事件的标注，一类是复杂事件的标注。值得注意的是，单个事件与复

杂事件的分类标准不同于单句和复句的基于形式的划分。形式上的单句表达的既可以是单个事件，也可以是复杂事件。既存在由单个谓词构成的单句，也存在由两个或两个以上谓词及其支配成分构成的复杂单句。而复句的形式也可以表达的是单个事件。

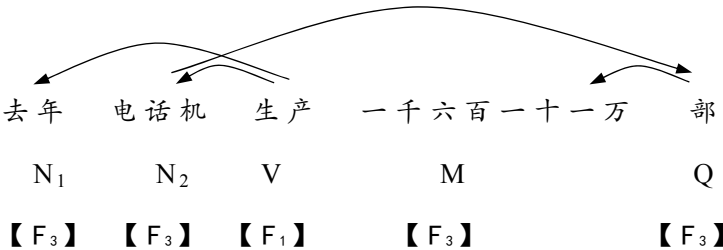
6.2 事件标注

6.2.1 单个事件的标注

对单个谓词构成的句子的标注，很多处理方式是与以往的处理相类似的，只需加标词语所属的不同词序域即可，因此这里不再赘述。但是需要注意词序域原则下词语在同一域分属不同侧的情况的标注，涉及到一些在词序上分离但有句法依存关系的情况。其中一些即是本文谈到的单个谓词下的交叉依存现象，另外还有一些其他的情况：

第一，名词与量词分离的现象

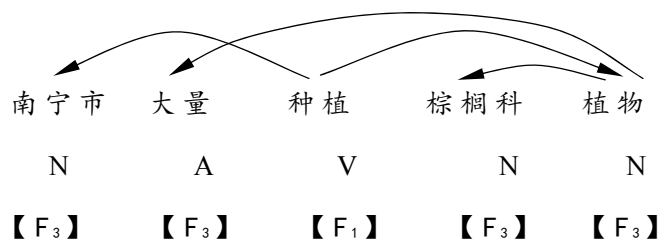
同属于第三域的名词与量词分列于动词的两边，但是有句法依存的关系。这是本文所论述的交叉依存现象。交叉依存的现象需要正确的标注，而不必限于以往作为黄金标准的投射性依存的假设，直接把有依存关系的两个成分表示出来。



第二，形容词和名词分离的情况

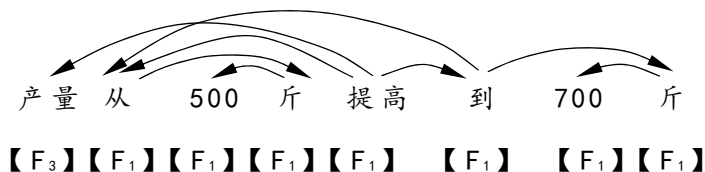
形容词与其修饰的名词同属于第三域，具有句法上的依存关系，

但在语序上分离，分属于第三域的不同侧。

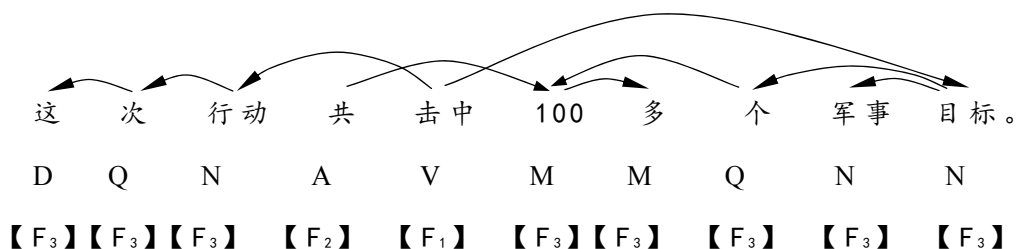


第三，介词连用的情况

以往分别把各个介词分别依存与动词做法割裂了框式介词之间的句法联系，因而在依存图标注中需要明确的表示出来。如本文谈到的“从...到...”结构中“从”与“到”之间的句法关系。与动词同属于第一域。



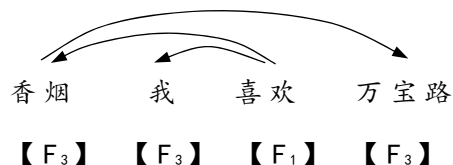
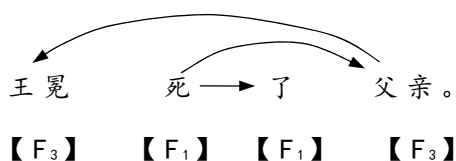
第四，副词和数词分离的情况；副词与其修饰的数词分属于不同的词序域。副词在第二域，数词与其修饰的名词处于第三域。



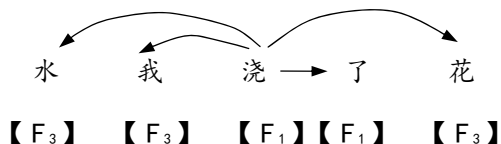
另外，还有一些不属于交叉依存现象，但是也是具有依存关系的词语在语序上分离的现象，其标注形式也不同于以往。或者表现为有边跨过最高结点，或者是依存关系的标注不同。

第五，名词与其修饰语名词分离的情况

又可以分为两种，一种是修饰语名词在前的情况，如：王冕死了父亲。一种是修饰语名词在后的情况，如：香烟我喜欢万宝路。这两类句子是曾在学者有过很多争论的句子，在域的角度来看，词语之间的修饰与被修饰关系是比较明确的。



第六，一些无形式标记的“话题--论述”结构；



这里要注意区分其与第五类两例的不同之处。前面两例中的“死”与“王冕”、“喜欢”与“万宝路”的关系不同于“浇”与“水”之间的关系。“水”与“浇”之间的关系是比较直接的，两者之间是动作和工具的关系，而“死”与“王冕”的关系则是间接的，需要通过“父亲”来过渡，因此不能与“王冕”表示为直接的依存关系。“万宝路”与“喜欢”之间的关系也是间接的，通过“香烟”来过渡。依存关系种类定义和描写我们将在下一节探讨，这里只是提醒其关系的不同。

由单个谓词形成的句子的标注与以往的依存树的标注是基本

相类似的。我们在新的标注中需要特别注意的是语序上分离的现象以及一些非论元性成分不借助助词出现情况。

6.2.2 复杂事件的标注

由单个事件构成一个句子是理想状态下的情况，实际的语料中这样的句子相对还是较少的，更多的句子表达的是复杂事件的集合。复杂事件是单个事件通过不同的方式组合而成的。可以是单句的形式，也可以是复句的形式。我们一般用形式上的单句或者复句作为划分句子的不同形式标记，但是对于事件结构来说，单复句的区别没有实际的意义，单句中可以是复杂的事件，复句也可以表示的是单个事件。

对复杂事件的标注，需要区分开不同事件的中心谓词及其分别支配的成分。我们使用加标的方法，对各个谓词按照次序的先后分别加标 P1、P2.....的符号。其直接支配的成分分别继承这些符号，而不直接支配的成分则用默认的继承，不做形式的标记。对于共享的成分则同时加标两个谓词及域的符号以示区别。

在前文域的递归性部分，我们指出了域的三种递归方式，分别为嵌套、叠合与链接。这是从形式递归的角度来说的，同一个递归形式之下，词语之间的组合关系是不尽相同的。这里我们说明的是事件组合时的几种不同的关系分类。

复杂事件的组合方式主要有以下几种基本的方式：

第一种：并列

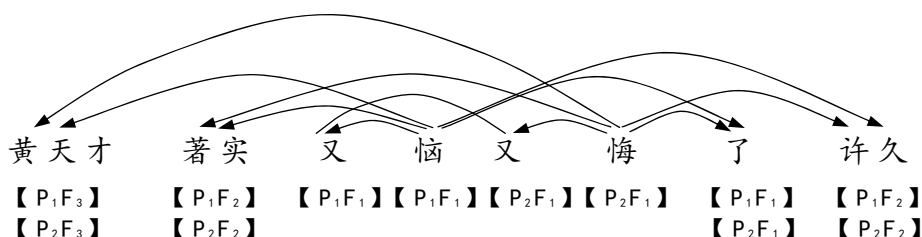
并列型组合方式是事件之间较为常见的结合方式。类似于树邻接语法（Tree Adjoining Grammar）中所讲的“相邻”的结合方式。

并列型的组合方式，其实是对语言空间网络结构的扁平化处理，这是受语言线性一维的物理特征所制约的。心理学上也认为，人们

交际的过程是多个信道同时进行的。对并列结构的处理在头脑中也是并行的。

如：“黄天才著实又恼又悔了许久。”由两个行为性事件构成，“黄天才著实恼了许久”与“黄天才著实悔了许久”两个事件并列组合在一起。支配成分共享，通过成对的副词“又。。。又。。。”的连接使事件得以并合。

(191) 黄天才著实又恼又悔了许久。



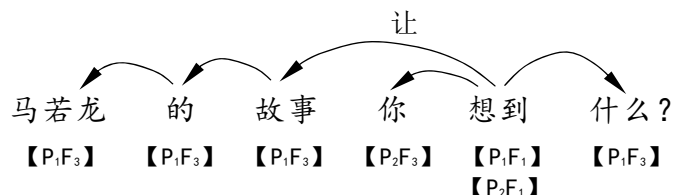
严格来讲，完全并列的两个事件之间的先后顺序应该是可以调换的。但是，现实中绝对同时发生的事件几乎是没的，总会有哪怕是极小时间上的先后差异，但人们还是认为是同时发生的。这样，一些连续发生的事件也会使用并列结构的表达方式。连用的谓词之间也逐渐确定按照时间顺序排列的习惯。能够使用并列组合方式的事件之间的支配成分往往是共享的，也可以说是合并为一体的。

并列型组合事件的共享可以有几种不同的形式。第一种是所有的论元共享。如上例的“黄天才著实又恼又悔了许久”。第二种是部分论元的共享。汉语中部分称之为兼语句的句子也是使用了并列的组合方式，但使用的是同一个动词，共享了谓词和宾语。

例如，“马若龙的故事让你想到什么”由两个事件构成，分别是“马若龙的故事让想到什么”和“你想到什么”。在前文中论述过，“让”在这里是一种关系的标记，表示一种致使的关系。同时也区分开“故事”与“你”对动词“想到”不同的角色。我们把“让”直接放在边上表示依存的关系。两个事件的动词和宾语共享。兼语

句都是采用的这种结合模式。

(192) 马若龙的故事让你想到什么。



“故事”与“你”之间有没有句法上的关系是值得研究的。传统的看法认为是祈使动词“让”连接了两个名词。而第二个名词“你”又同时充任第二谓词的主语。我们认为，正如主语和宾语之间的关系一样，很难说主语和宾语之间有直接的句法关系。因此，这里的“故事”和“你”之间也是没有句法上的关系的。

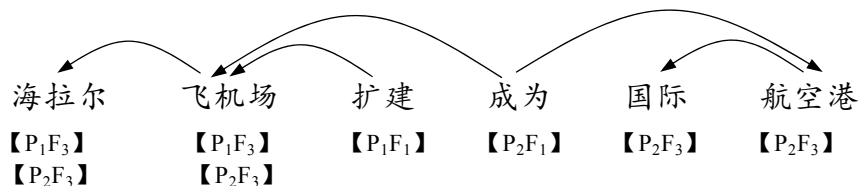
第二种：链接

链接是两个以上事件的依次出现。可以分为单句内和复句内两种情况。一般是第一论元的共享，但谓词和其他论元不共享。

A. 单句内的链接

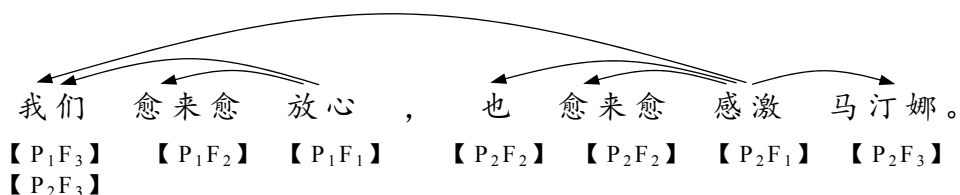
(193) 海拉尔飞机场扩建成为国际航空港。

由两个事件“扩建海拉尔飞机场”和“海拉尔飞机场成为国际航空港”组合而成。共享了第一论元“海拉尔飞机场”。



B. 复句内的链接

(194) 我们愈来愈放心，也愈来愈感激马汀娜。

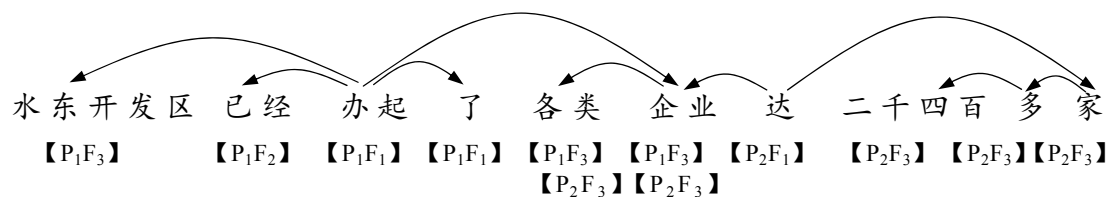


第三种：接合

接合的组合方式一般是与相对谓语的描述性事件相关的，是在一个单句范围内的接合，体现在其与不同的事件类型组合当中。又可以细分为三种，一种是“行为性事件+描述性事件”的方式，一种是“描述性事件+行为性事件”的方式，第三种则是“描述性事件+描述性事件”的方式。

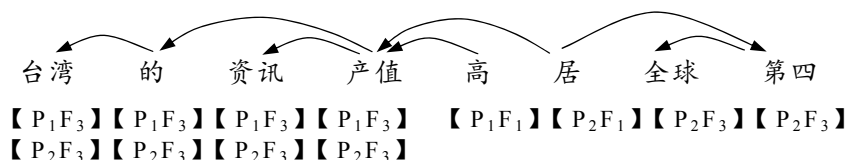
第一种例子如：“水东开发区已经办起了各类企业达二千四百多家”。是由行为性事件“水东开发区已经办起了各类企业”和描述性事件“各类企业达二千四百多家”组成的。两个事件接合在一起，第一谓词的受事与第二谓词的主体共享，构成线性的表层序列。这一类的描述性事件一般是相对谓词的情况，因而不能像拼合组合方式中那样把两个谓词整合在一起。如图所示：

(195) 水东开发区已经办起了各类企业达二千四百多家。



第二种的例子如“台湾的资讯产值高居全球第四。”是由描述性事件“台湾的资讯产值高”和行为性事件“台湾的资讯产值居全球第四”组成。两个事件接合在一起，共享了第一论元，构成线性的表层序列。

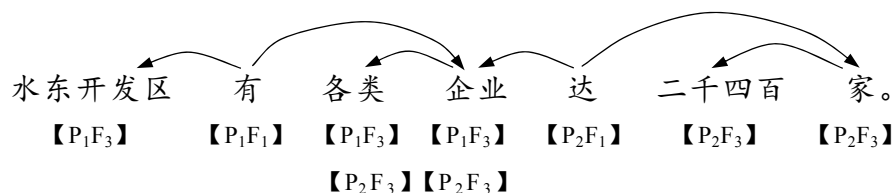
(196) 台湾的资讯产值高居全球第四。



“描述性事件+行为性事件”的组合方式为接合，两个谓词虽然放置在一起，但不同于拼合中两个谓词拼合的整合方式，两个谓词之间不能整合为复合事件，这是需要注意的。

第三种则是“描述性事件+描述性事件”的方式。两个描述性事件也可以使用接合的组合方式。在例句(195)的基础上，我们可以稍微变换构成一个新的句子为：“水东开发区有各类企业达二千四百家”。为两个描述性事件“水东开发区有各类企业”和“各类企业达二千四百家”的组合。当然，这样的句子还可以细分为若干种不同的类型。

(197) 水东开发区有各类企业达二千四百家。



第四种： 拼合

拼合是指两个简单事件组合成为一个复合事件。复合事件与复杂事件是两个不同的概念。复杂事件是对由两个以上事件组成的句子的统称，复合事件则是一个动作性谓词与描述性谓词构成事件的组合，类似于事件语义中论述的动词分解为外动词和内动词的结合。所以往往也认为是一个事件。在汉语中则多用动补结构来表达这种

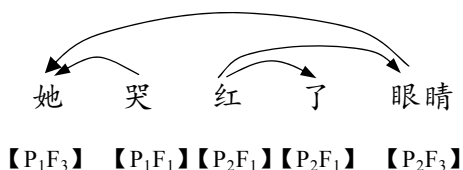
复合的事件。

拼合的组合方式有三种基本的类型，一种是无任何共享成分的组合，第二种为只共享主语，第三种则为只共享宾语。分别如下：

A 无任何共享成分的拼合

例如：“她哭红了眼睛。”由两个简单事件构成，分别为“她哭”和“眼睛红了”。“哭”与“红”都是一价的谓词，“红”为形容词，是描述性谓词。在与行为性事件结合为较大的复合事件时，倾向于跟前面的动词“哭”放在一起。因此，行为动词“哭”和描述性谓词“红”结合性较强，都放在了第一域的内部，且遵守时间序列的原则。第二个事件的论元“眼睛”在本例中放到了“红”的后面。

(198) 她哭红了眼睛。



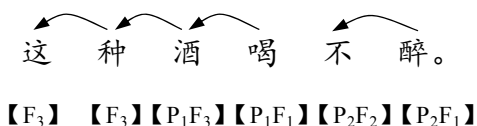
事件拼合以后，两个事件各自支配的成分之间在句子的域中会激活一些固有的句法关系。上例中，“眼睛”为一价的名词，要求有其他名词与之同现，“她”和“眼睛”虽然分属于第三域的两侧，但在拼合后领属关系实现了激活。因此需要表示出来。因此，“眼睛”的相对位置需要处于“她”的后部，或者处于第三域的同侧，或者分置于两侧。这是由“无歧义相对自由”的原则决定的。

当然，也可以形成“她眼睛哭红了”的句子，在逻辑语义上是等价的。不同的是使用的是“交叉”的组合方式。“哭”作为行为动词又属于“自动词”，它在跟行为性更强的“他动词”结合时，则会放到后面。一般来讲，谓词之间的排列是遵守的时间顺序原则，这跟事理发展的顺序也是相符的。

拼合中有一类特殊的现象需要注意，即如果有歧义发生的情况下，行为性动词和描述性谓词即使有很强的组合性也不能结合在一起，而采用并列的组合方式。如上例的：“李丹唱这首歌红了。”这句话由两个表述构成，“李丹唱这首歌”和“李丹红了”。动词“唱”与形容词“红”原来是倾向于结合在一起的，以形成一个复合事件。但是如果“唱”和“红”结合在一起，“李丹唱红了这首歌”中的“红”则指向为“这首歌”，这与原句所表达的意思不同。所以，为了避免歧义，在这句话中的“这首歌”不能后移。

也可以第一事件的受事出现，施事不出现，采用相同的拼合方式。例如“这种酒喝不醉。”的句子由两个事件构成，“喝这种酒”和“不醉”。两个事件涉及的施事成分则没有出现。否定副词“不”的位置比较灵活，其分布位置不受词序域的影响。

(199) 这种酒喝不醉的。



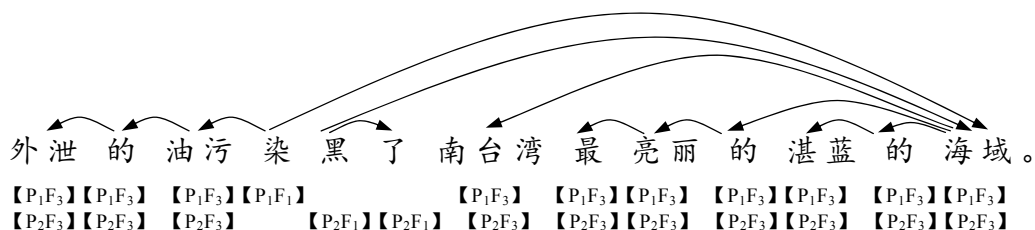
拼合的两个谓词之间是一种事理上的先后顺序，从句法的角度来讲虽然可以形成一个韵律的单位，但并没有强制共现的关系，因此很难说两者之间有句法依存的关系。如果从复杂事件的角度来看，则是第二个实现了时体标记的谓词更可能是事件的中心，因而更可能是句法上的中心词。

B 两个事件的受事共享。

例如：“外泄的油污染黑了南台湾最亮丽的湛蓝的海域。”由两个事件“外泄的油污染南台湾最亮丽的湛蓝的海域”和“南台湾最亮丽的湛蓝的海域黑了”构成。动词“染”和形容词“黑”倾向于整合为复合事件，使用拼合的组合方式。“黑”在语义上是指

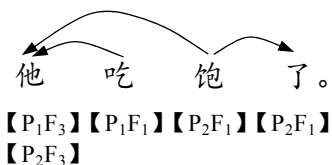
向“海域”的。

(200) 外泄的油污染黑了南台湾最亮丽的湛蓝的海域。



C 两个事件的施事共享。

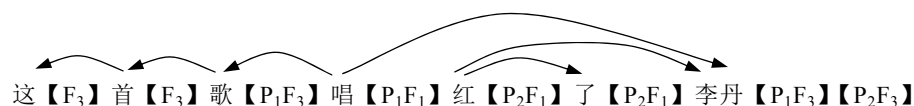
例如：“他吃饱了”。由两个事件“他吃”和“他饱了”组合而成。动词“吃”与形容词“饱”整合为复合事件。两个事件的施事共享，为下图的方式：



施事共享的情况还可以结合其他的组合方式一同使用，例如上例中的“吃”如果带有宾语，则是拼合与交叉的组合方式共同使用的情况。

施事共享也可以都处于后部，例如“这首歌唱红了李丹。”由两个表述构成，“李丹唱这首歌”和“李丹红了”。由于“唱”跟“红”之间有逻辑先后关系，所以“红”必须放在“唱”的后面。如果采用常规模式，则为“李丹唱红了这首歌”。由于这种模式的语义指向上，“红”与“这首歌”可以形成一个表述，且具有优先性，所以不能采用。而采用了把受事提前，施事后置的模式。使得后置的“李丹”与“红”构成一个事件。

(202) 这首歌唱红了李丹。

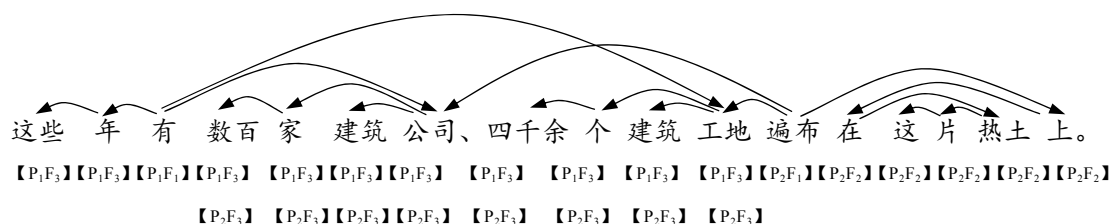


第五种： 嵌套

嵌套是指一个事件充当另一事件的一个成分的现象。嵌套组合方式的使用是在所有语言中普遍存在的现象。在以往依存树的论述中也多有见到，类似于树邻接语法（Tree Adjoin Grammar）中谈到的“接”的结合方式。能够用嵌套表示方式跟动词的性质有关。一些动词可以带一个句子作为其宾语。同时，大部分的动词都允许另一个事件作为其主语的形式。传统的依存语法中由于不承认有大于词的单位存在，在对由一个整体充当另一个事件中的一个成分的情况处理上有点不足。但是，如果把事件作为词和句子的中间单位，即可以解释和描述这些现象。正如对并列结构的处理中有人允许并列结构作为整体充当成分一样，我们也允许嵌套结构在依存表示中存在。一个事件作为另一个事件宾语的用例如下。

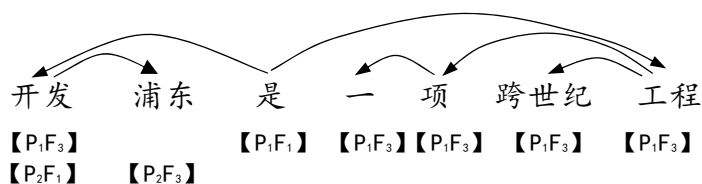
(203) 这些年有数百家建筑公司、四千余个建筑工地遍布在这片热土上。

事件“数百家建筑公司、四千余个建筑工地遍布在这片热土上”做动词“有”的宾语。



事件作为一个整体充当另一个事件的主语的情况也是广泛存在的，例如：“开发浦东是一项跨世纪工程”。其中的“开发浦东”即是一个事件充当主语的嵌套形式。

(204) 开发浦东是一项跨世纪工程。



第六种：交叉

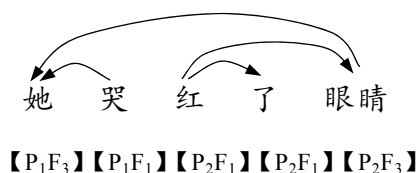
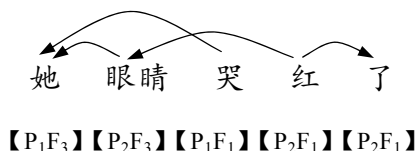
严格来讲，交叉并不能算是一种另外的组合方式，因为两个谓词之间基本是使用的拼合的组合方式，只是在各自支配的论元位置上出现了不同的选择，才形成了交叉的现象。为区别开与拼合的组合方式下的无交叉序列，我们把交叉作为一种单独的组合方式来说明。

树的交叉现象是汉语中是很普遍的。这也是体现依存图表示的优势所在之处。允许依存边的交叉，可以表示词汇之间复杂的联系，这是以往依存树表示不足的地方。事件交叉组合的方式又可以分为以下几类：

a. 两棵树的论元无共享的情况。但是两颗树的共现可以激活某些潜在的句法联系。如：“她眼睛哭红了”一句由两个事件构成，分别为“她哭”和“眼睛红了”。由于“哭”是行为动词，“红”为状态描述词，有合并为整体事件的倾向，所以要放在一起。其在逻辑上等价的表示还有有“她哭红了眼睛。”

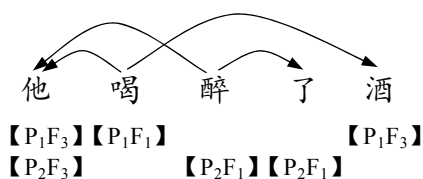
(205) 她眼睛哭红了。

“眼睛”作为一价名词，要求有名词充当其配价的成分，因此当“她”与“眼睛”共现的时候，两棵树的共现激活了词语之间的这种配价关系。所以，“眼睛”的相对语序必须处于“她”的后面，不能放在“她”的前面。对比下面两个句子：



b. 两个事件的施事共享，其组合为交叉的方式。如“他喝醉了酒。”由两个事件“他喝酒”和“他醉了”构成。施事共享，“喝”与“醉”结合为复合事件。“喝”与“酒”的边与“醉”和“他”的边形成了交叉。

(206) 他喝醉了酒。



类似的例子还有很多，如“孩子听烦了说教。”等。两个谓词的整合方式是受“无歧义相对自由”原则制约的。比较下面这三个句子：

(207) 他喝醉了酒。

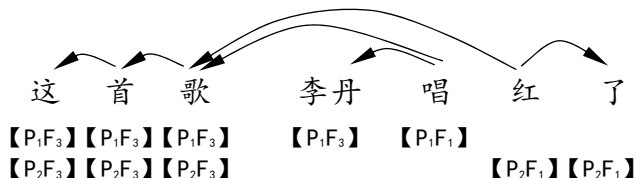
(208) 李丹唱这首歌红了。

(209) 李丹唱红了这首歌。

可以看出，三者之间是有细微的差别的。第一句中的“醉”由于在与“他”和“酒”之间的关系上比较明晰，不会因位置的不同而发生歧义，所以其位置的先后比较自由。既可以说“他喝酒醉了”，“他酒喝醉了”，也可以说“他喝醉了酒”。但是第二句与第三句中的“红”与“李丹、这首歌”之间的位置则比较固定，这是因为位置的不同会产生不同的意义。

如果第三句中的第二个事件“这首歌红了”的主语移位，则会出现有边跨过最高节点的情况，是另一种非投射性依存现象。例如“这首歌李丹唱红了。”这句话中的“红”则是只指向“这首歌”的。因为第一个表示中“唱”的施事出现在前应有的位置，所以“红”优先指向于移位的受事“这首歌”。如下图所示：

(210) 这首歌李丹唱红了。



6.3 依存图标注规则

依存图的标注有异于依存树的标注。具体来说，我们需要标注以下几个方面的信息。第一是词语本身的信息，包括词语的词性、义项和各个义项上的属性和纬度。第二是作为事件核心的谓词与它所支配的成分的依存关系，依存关系的方向以及其不同的依存关系种类的标注。第三，对复杂事件中不同子事件之间区分的标注以及组合后激活的新句法关系的标注；第四，则是语序信息的标注，即

语序上所属的域的范围的表示；

6.3.1 词汇信息标注

1. 词性标注

词性的标注在不同的语料库中使用的是不同的标准。宾州汉语树库的标注中就把许多词典中定义的动词标注为名词。如“开发、建设”等。这是将句法功能叠加在词性上的做法，我们认为应该把词序承担的句法功能与词性本身区别开。所以，对词语应该坚持使用词典中的词性标注。以“上海浦东开发和法制建设同步”一句为例，标注为如下的形式：

上	海	浦	东	开	发	与	法	制	建	设	同	步
N		N		V		C		N		V		V

这里需要把“开发”“建设”原有的“NN”标记修正为动词“V”的标记。

2. 词语属性的描述

在词语属性描述方面，我们可以采用词表的方法描写。对词语的不同义项的属性进行较为细致的描写，然后为句子的词语进行属性的链接。这样习得的数据在判断组合成立时候会排除掉很多干扰的因素。词语属性描述的示例可参照第四章的论述。

我们选取中国教育部公布的《现代汉语常用词表（草案）》中最高频的动词“生”（频率 398）³²为例子，按照前文论述的几个属性做一个标注的样例。

“生”作为动词有四个义项，

³² 此处只讨论“生”做动词用的情况，标注的为其四个义项的各个属性及纬度。

生

- ① 生育，出生：胎～ | 卵～ | ～孩子 | 优～优育 | ～于北京
- ② 生长：～根 | ～芽
- ⑧ 产生，发生：～病 | ～效 | 惹是～非
- ⑨ 使柴、煤等燃烧：～火 | ～炉子

义项②⑧⑨是从基本义项 ①生育、出生 分化出来的。其分化的维度为时间属性上的结束后“状态”维度。另外，前文已经指出，第一义项的“生育”与“出生”应该分化为两个不同的义项。在分化为新的义项后，又可以具有时间的四个维度。如下表所示：

属 性		义 项				
	纬度	① 生育； ～孩子	① 出生 ～ 于北京	② 生长 ～根 ～芽	⑧ 产 生 ； 发 生 ～病 ～效	⑨ 使柴、煤等 燃 烧 ～火 ～ 炉子
时间属性	开始	+		+	+	+
	持续	+		+	+	+
	终点	+		+	+	+
	结束后	+	+	+	+	+
情感属性	褒义					
	中性	+		+	+	+
	贬义					
位置属性	起点位置					
	经过位置					
	目的位置					
	发生位置	+		+	+	+

表 7 “生” 各义项属性维度标注示例

6.3.2 句法关系标注

第一，事件结构的标注；

分为单个事件标注和复杂事件的标注两种。对于单个事件，我们用【P₁】的符号来给每个成分加标。以便与组合的事件保持统一。对于组合的事件，我们把每棵表示一个事件的子树按照处于第一域内谓词的先后顺序分别加标 P₁、P₂、P₃……的次序。各个被支配的成分也各自标明其受到哪个谓词的支配。对共享的修饰成分，则分别标注两个谓词的号码。同时，对组合后激活的新的句法关系也加以标注。

单个事件的标注实例如：

(211) 这次行动共击中 100 多个军事目标。

这 次 行 动 共 击 中 100 多 个 军 事 目 标
【F₃】【F₃】 【F₃】 【F₂】【F₁】 【F₃】 【F₃】【F₃】 【F₃】 【F₃】

复杂事件的标注实例如：

(212) 妈妈眼睛哭瞎了。

妈 妈 眼 睛 哭 瞎 了
【P₁F₃】 【P₂F₃】 【P₁F₁】 【P₂F₁】 【P₂F₁】

第二，依存关系种类的标注

相较于短语结构的句法标注，依存关系的标注在操作上有一定的便捷性。依存关系的判断只需要考虑两个词语之间的关系，相对来说要比较简洁。而且可以回避开类似于短语结构标注中经常遇到

的判定短语的边界、多分组合的划分、同形的歧义结构标注等许多有争议的问题。

但是，就依存关系分类本身来说，也是存在很多的争议的。首先，对于 Tesnière 最早提出的并列、转位以及依存三种关系的归并上存在很多分歧。大部分理论都只承认依存关系，把并列关系也转化为依存的表示或者看做是一个整体。而对 Tesnière 作为基本关系之一的转位关系在大部分理论中没有体现。而且，对依存关系的分类、类别命名也都是具有争议的。

一般来讲，较多的体系使用较表层的句法依存关系，如 Subject、Object、Adverbial 等。比较特殊的如 Mel'čuk (1988) 的 MTM 模型使用数字来表示动词配价成分，用描述性的标签来表示可选的成分。而在面向实践的句法分析系统中，更多的使用的是无标记的句法依存结构。

针对汉语的依存关系，也有若干种不同的分类。哈工大的依存树库中对依存关系定义为 34 个类型。黄昌宁苑春法曾设计了 32 种依存关系。周明、黄昌宁 (1994) 划分了 44 中依存关系。最多的如张敏 (1993) 曾涉及了 106 种依存关系。

印欧语中句法依存关系的确定是有形式基础的。经常表现为词语间形态上的一致性。如英语中主语与动词之间在数上有一致关系，动词对宾语有格的选择等。抛开具体应用的实际，单就纯粹的句法关系本身来看，汉语中的句法关系的确定更多地要借助语义的信息。

依存关系种类的定义要求能够具有一定的适应性。既要能方便地加工处理大量的真实文本，也要能提供足够多的区别信息。严格来说，每两个词语之间的关系都是不尽相同的，但是分类过细就等于是没有分类。而且，“类分的过细，导致在标注时，要仔细推敲，降低了操作性，影响了标注的效率。(周强 黄昌宁 1994)”。而分类过于粗疏，则又会遗漏很多有用的区别信息。因此，我们提议使用一种层级式的句法关系描述体系。用有层次的分类体系来表示词

语之间的依存关系会有两个好处。首先，对于语料库开发来说，可以实现分步骤递进式开发的思路。首先标注上层的句法关系，可以满足经济性的原则，而且避免数据稀疏的问题。然后再具体标注各个次类，可以为语言研究者提供精细的关系分类。在标注过程中，随着标注语料的增多，还可以对一些没有遇到过的现象可以进行随机的拓展。第二，对于句法学习器来说，分层次的标注形式可以满足其对依存关系颗粒度大小不同的学习要求。

依存关系的分类，我们借鉴事件结构的顺序来描写，分别定义如下：

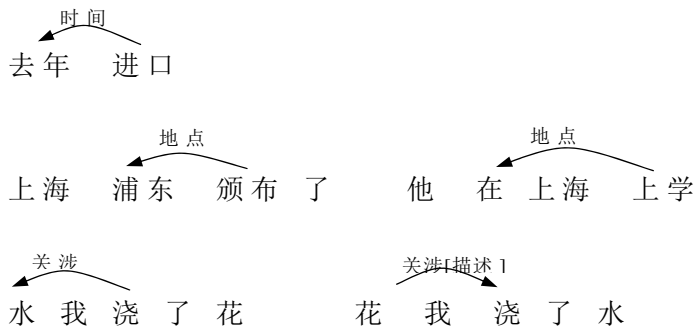
1. 谓词直接相关的依存关系

① 必有论元

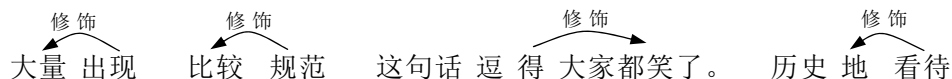
典型的必有论元是名词的形式，也可以有其他形式充当类似于名词的功能。同时，还有一些相对谓词与主语形成的必有论元关系。另外需要注意词序改变引起的句法关系的不同。



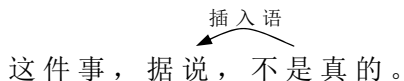
② 可有论元，可以具体有分为小类，如时间、地点、关涉等。



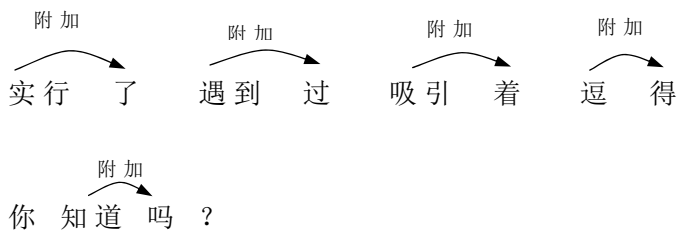
③ 修饰成分 可以是副词、形容词等



④ 插入语；现在对于插入语的研究有泛化的趋向，把很多句中的成分也可作是插入语。我们只限定在使用逗号分隔开的情况。

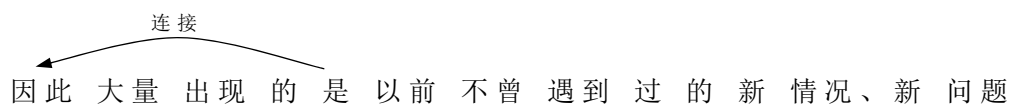


⑤ 附加成分；主要是“着、了、过”时体标记和助词类的得以及一些语气词。



⑥ 连接关系，连接关系表示连词与事件之间的关系，不同于连词之间相互的关系。每个连词与其连接的事件核心可以用连接来

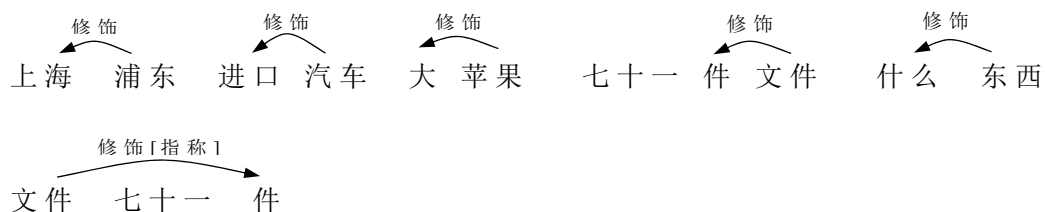
表示。



3. 名词相关

名词除与动词可能构成的上述几种关系外，其本身还可以有如下的关系：

① 修饰关系；修饰语可以是名词、形容词、动词、量词、数词等。



② 同位关系

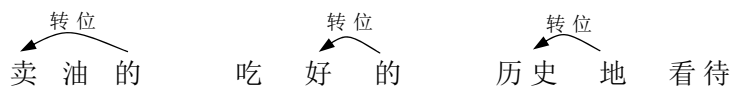


③ 附加关系

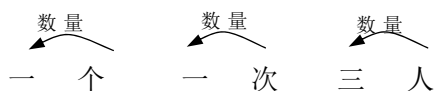


3. 其他关系

① 转位关系；转位关系也可以看做是依存关系中的一个子类。通过转位，实现不同的语法功能。汉语中的标记为“的、地”。



② 数量关系；是修饰类的一个子类。一般是指数词与量词之间，也有数词直接与名词有关。

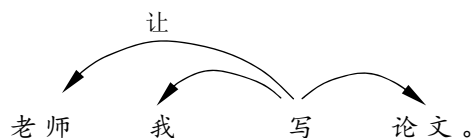


③ 逻辑关系；可以是事件之间的，也可以是连接成分之间的。主要表现为连词和成对使用的框式连词。可以根据具体逻辑关系的不同分为若干不同的小类。



④ 特殊关系

在武汉大学特征结构语料库中，曾认为一般对句法关系的定义都是使用了泛语言的标记，因而在很多时候引起人们的争论。因此其对关系的表示都是使用语言中本身的标记。其对属性、逻辑等关系的表示上是比较成功的，但是对一些基本关系的表示上却不能聚类出合适的词语。鉴于此，我们参照武汉大学特征结构语料库标注的部分思路，把部分不好定义的关系使用语言本身来表示，而对一些基本关系依旧使用泛语言的标记。例如前文中对部分兼语句的处理，“老师让我写论文。”“让”在这里是一种关系的标记，表示一种致使的关系。同时也区分开“老师”与“我”对动词“写”不同的角色。我们把“让”直接用作关系的标记，放在依存边之上。



第三，依存方向的标注

依存方向的判断标准是比较明确的。Mel'čuk (1988) 论述了判断句法依存的方向的两个基本标准。

① 是看其对句法结构所能充当的句法成分的影响力的大小，影响力大的词语作为这个句法结构的支配词。例如，数量短语可充当的句法配价的能力是由其名词所决定的，所以其句法依存的支配词为其中的名词。

② 是选择与结构外词语有形态依存关系的词语。

两个标准不是等同的，在第一规则不能奏效的情况之下才考虑第二规则。

依存方向的表示有两种选择，一种是从支配词指向被支配词，另一种则是从被支配词指向支配词。两种表示方式在不同的理论中都可以见到，但是不同的形式表达的是相同的内容。我们建议使用由支配词指向被支配词的方式。例如：



动词“吃”支配名词“饭”，那么依存的方向即是由动词指向名词。

6.3.3 语序信息标注

鉴于汉语中词类范畴与功能之间不存在简单的对应关系的现实，我们可以得出单靠词类范畴本身是不能够区分词语在不同位置上的功能的结论。因此，需要在标注词语之间句法关系的同时标注词语所属词序的信息。两种信息同时具备才能够提供足够多的真实信息供学习器习得。

我们用 F_1 、 F_2 、 F_3 、 F_4 四个符号分别表示词序域不同的范围，处于同一域内的词语加标同样的域标记。如上例：“妈妈【 P_1F_3 】眼睛【 P_2F_3 】哭【 P_1F_1 】瞎【 P_2F_1 】了【 P_2F_1 】”中的 F_3 、 F_1 分别表示处于第三域和第一域。如果一个域内没有词语，如上例中第二域中则没有出现词语，那么这个域的符号可以不用标记。

连用的介词、连词等本身在句法上构成一个框架结构，但是在语序的标注中我们依旧按照各个词语各自联系的成分标注其所属的辖域，然后用依存关系表示之间的联系。

对于一个事件充当另一个事件的一个部分的情况，在标注上需要有继承的关系。充当一个部分的事件的谓词加标其支配谓词符号和所属的词序域即，被支配的词语继承上位的词序域。在这个标记之后要加标事件本身的谓词符号和此事件的词序域。例如前文中事件嵌套的例子：

(213) 开发浦东是一项跨世纪工程。

可以标注为如下的形式：

开	发	浦	东	是	一	项	跨	世	纪	工	程
【 P_1F_3 】	【 P_1F_3 】	【 P_1F_1 】	【 P_1F_3 】	【 P_1F_3 】	【 P_1F_3 】	【 P_1F_3 】	【 P_1F_3 】	【 P_1F_3 】	【 P_1F_3 】	【 P_1F_3 】	【 P_1F_3 】
【 P_2F_1 】	【 P_2F_3 】										

其中“开发浦东”为充当部分的事件，因此继承第一事件的域标记为“开发【 P_1F_3 】浦东【 P_1F_3 】”，同时又带有自身事件的标记“开发【 P_2F_1 】浦东【 P_2F_3 】”。

6.4 小 结

本章介绍了依存图库的构建原则。根据语言现实和计算技术的

现状，把依存的表示从树的形式发展为图的形式。依存图的标注是一个全面的构想，既包括以往定义的投射性现象，也包括非投射性的现象。本文讨论的交叉依存类现象的标注也都包括在其中。

把事件作为标注的中间单位，区分单个事件与复杂事件的不同标注。而且，事件在组合后会激活新的句法关系。我们分别定义了依存图三个层面的标注标准，包括词汇信息、句法信息和语序信息。

依存图库的构建，不仅可以为语法研究提供真实的语料，而且可以为基于统计的分析器提供数据学习的基础和统一的评测平台。

7 结 论

7.1 本文研究工作的总结

现有的依存句法分析器基本上都是基于投射性依存的假设的。这是与计算机科学上树型数据结构相对应的处理方式。要达到句法分析器分析精度的进一步提高，必然要放弃投射性的完全假设，涉及到对非投射性依存现象的处理。因此，对非投射性依存现象的处理，是为实现大规模真实文本完全正确的句法分析这个战略目标努力的过程中必然要面临的一个课题。

现有对非投射性依存现象的研究中，对不同语言中的有边跨过最高结点类的现象有一定的提及，但极少涉及到交叉依存类的现象。因此，我们选取交叉依存现象作为研究的入手点。本文的主要目的即是对交叉依存现象进行系统的考察，通过对交叉依存现象的分析和比较，得到规律性的认识，为此类现象的认识与系统性处理提供理论基础。

通过对宾州汉语树库语料的调查，我们得到了汉语中九类交叉依存的现象。分别为：名词与量词分离类、名词与形容词分离类、名词与名词分离类、副词与数词分离类、框式介词类、把字相关类、动词与副词分离类、动词与数词分离类、框式副词类。然后探讨了这类现象有哪些共性，以及现象背后的两条共同制约原则。研究成果应用于指导语料库标注和句法分析，以期能为中文信息处理贡献一点力量。投射性与非投射性依存现象的划分，虽然是从计算语言学的角度来说的，但也给我们提供了一个新的语言研究的角度，这有助于我们发现以往没有注意到的语言规律和具体的规则。

具体而言，本文的贡献及创新可以归纳为以下两个方面：

● 首次提出了交叉依存现象中词语语序的使用是遵守的环形语序域原则。语言是个多维空间的网络结构，多维空间的语言网络在拓扑变换为一维线性序列时，经历了中间状态的二维环形平面图。具有依存关系的词语可以分属于不同的词序域，因此形成线性语序上的隔断。而且同一域的词语可以在一维化时处于分离的两端而形成线性序列上的隔断。这就造成了各种不同的依存边交叉的情况。环形语序域原则解释了语言网络到线性语言相对应的过程，突破了以往只是把语言网络作为一种隐喻使用的现状，把语序的研究从一维线性的视角扩大到二维的平面视角；这使我们可以重新认识汉语的语序特点，汉语在宏观语序上是比较固定的，也即域的总体划分是固定的。在域内的微观语序上则比较灵活。语序灵活的现象并不是完全自由，只有在不发生歧义的情况下才能允许较自由的语序。可以称之为“无歧义相对自由”原则。这个原则可以适用于许多存在争议的语言学问题上，如对主宾互换问题、主谓句中两主语之间的语序问题的讨论等；

● 首次明确提出把事件结构作为依存语法中在词语与句子两种语法单位之间存在的中间单位。以往的依存语法系统中只承认词语和句子两级的单位，在对整体句子结构的把握上有点松散。事件结构的使用可以兼顾句法关系和语序两个方面，而且可进行向上和向下两个层级的研究，对于句子分析和表示都具有重要意义。

另外，区分系统性语义特征和零散的语义特征，把系统性的语义特征（属性）作为句法层面的标准来使用。提出了句法依存关系的成立是由于词语组合上的属性和谐原则。词语在组合时要求义项具有的属性在取值上不相斥，组合关系才可能成立。组合关系的成立也不因词语间语序的隔断而中断。把属性作为句法层面的标准使用，是针对汉语现实提出的细化依存关系描写的方法，有助于句法分析正确结果的取得。同时，词语属性及其不同取值的描写也跟词语义项的分化有紧密的联系，用属性取值的分化作为义项分化的一

个标准，具有较强的操作性，有可能为面向信息处理的机读辞典的构建提供一个新的描写方式；

针对应用方面的需要，我们利用上述研究结果，提出了构建依存图的设想和具体的标注原则。现有的依存树库构建基本上都是基于投射性原则的。需要构建一个标注了非投射性依存关系的语料库作为基于统计的分析器学习和评测的基础。同时，依存图标注修改了独根结点、单一父结点、单纯结点的黄金原则，在标注原则上是一个新的尝试。

从理论的角度来说，面向人的语言学研究与面向机器的语言研究在本质上是相通的，两者可以互相借鉴与互相促进。机器处理遇到的问题可以给语言学研究提供一个新的观察角度，这些角度往往是面向人的语言研究中考虑不到的。新的视角的研究，很可能对语言学理论带来深刻的变革。虽然投射性与非投射性现象的划分是从最后的分析结果来看的，是一种现象的分类。其最直接的动力是提高语言信息处理的精度的需要，但其同时对依存语法研究本身来说也具有促进意义。现象的划分使我们看到了一些貌似无关的语法现象之间的自然聚类，给我们提供了一个很好的观察视角，可以有助于发现以往没有关注到的理论问题，即这些现象之间存在着怎样的共性，这些共性又是由哪些因素制约的。研究的结果将不仅会影响面向工程的语言处理系统，也有助于依存语法理论本身的深入和完善。

7.2 下一步的研究计划

一般认为，科学研究要经过从理论到技术再到工程三个不同的阶段。各个阶段之间先后相继，且各自承担着不同的研究任务。本文初步考察了汉语中交叉依存现象的分布及制约其形成的两条基本原则，总体上来说侧重于理论的，还有很多方面需要继续改进

与细化。具体而言，下一步我们还需要在以下两个方面继续深入的研究：

（1）在现有研究的基础上，深化对词语属性的考察，把系统性的语义特征概括为句法上的属性概念，并且优化各个属性之间的构架；

（2）考察非投射性依存现象中的第二类，即有边跨过最高结点类的情况，梳理这类现象在汉语中的分布，并尝试用现有的原则来做解释，同时修正原来的原则描述中不够合理的地方。

附录 宾州树库例句

附录一 名词与量词分离类

ID11444	目前国内提供高阶主管进修课程的公私立大学共有台大、交大、中山、大叶、元智等十馀所，为什么企管所特别受社会人士青睐？
ID11413	此次与会的学员多以加勒比海及中美洲地区的英语系邦交国为主，包括来自中美洲的贝里斯及东加勒比海的格瑞内达、圣文森、多米尼克、圣克里斯多福等五个英语系国家的高阶新闻官员、重要媒体负责人及新闻从业人员共十一位。
ID9954	而一九九五到一九九七年查出的猎杀藏羚羊就有三千只。
ID498	仅在民用的生物技术、信息技术、自动化技术、能源技术和新材料技术领域，已取得一千二百多项研究成果，其中达到国际先进水平的有五百六十项。
ID3661	据化工部介绍，一九七八年，中国化工进出口总额不到三亿美元，去年达到七十五亿美元，出口创汇超过一千万美元的化工企业已有三十家。
ID35	据初步统计，目前在中国境内承包工程的国外承包商已有一百三十七家，承包的工程达一百四十一项，其中最大规模的项目达二十七点七亿元；中外合资合作的建筑企业近二千家。
ID158	镇江市近年依托大运河建厂的外商独资企业就有八家，合资企业更多。
ID151	常州市沿运河两岸零点五平方公里范围内，大型厂矿企业有一百八十六家。
ID145	据初步统计，仅运量在万吨上的企业就有二千多家，小的乡镇企业则不计其数。
ID153	无锡市沿运河企业有五百多家，超亿元产值的大型企业有五十多家，成为无锡市经济发展的主骨架。
ID394	去年，电话机生产一千六百一十一万部，程控交换机生产三百三十四万线，微机生产十二万台，产量分别成倍增长。

(11 句 14 个用例)

附录二 名词与形容词分离类

ID9937	孙怀玉称，国际资本的大量涌入，为武汉经济连续多年保持两位数增长速度提供了有利条件。
ID954	农业在稳定粮食种植面积，提高单产的同时，着力调整、优化经济结构，热带水果、瓜菜和畜牧、水产等优势产业进一步获得发展，粮食、水果、瓜菜、干胶产量均达到或超过历史最好水平。
ID920	李铁映说，中国有决心、有信心实现到本世纪末初步建立社会主义市场经济体制的目标。
ID9091	法国大量出版“欧元书”
ID83	现在广西已初步形成了沿海开放城市、沿海经济开放区、边境开放城镇相结合，由沿海、沿边、沿江向腹地推进的多领域、多层次的对外开放总体格局。
ID806	而投资回报率始终高居全国之首，则是吸引外商不断追加投资的另一个重要原因。
ID785	为突出城市园林绿化的亚热带风光及民族特色，南宁市大量种植棕榈科植物和开

	花乔木、花灌木、彩叶植物以及地被植物、草坪等，市区的一百多条主要街道基本实现了绿化、彩化、果化、香化。
ID421	此间经济界人士认为，这个著名的侨乡由于与港澳台地区长期保持着良好的经贸往来和市场联系，借助金融融资、信息传递、产业转换等有利条件，福建发展高新技术产业将会有良好的前景。
ID13262	当时，秦国人口约两千万人，动员劳役的人数却多达两百多万，才打造完成这一座地下皇城。
ID127	由浙江医科院院长、中国科学院士毛江森主持在世界上率先研究成功，并具有国际先进水平的甲肝减毒活疫苗，去年经卫生部批准正式投入生产和使用，目前该区生产此疫苗的普康公司已形成年产五百万人份的生产规模，这对有效地控制甲肝流行具有重大意义。
ID112	进出口银行通过书面介绍、实地考察等形式向日本公社债研究所全面介绍了今年来中国金融体制改革的情况、银行成立的背景、银行管理和运营机制、银行业务发展现状以及以后的发展目标，使之对中国进出口银行有了较深的了解。
ID5	对此，浦东不是简单的采取“干一段时间，等积累了经验以后再制定法规条例”的做法，而是借鉴发达国家和深圳等特区的经验教训，聘请国内外有关专家学者，积极、及时地制定和推出法规性文件，使这些经济活动一出现就被纳入法制轨道。
ID6787	北方蜜蜂一到冬天，既不生产蜂蜜，又大量消耗蜜糖和王浆，还易受到鼠害，造成的损失严重。
ID4	浦东开发开放是一项振兴上海，建设现代化经济、贸易、金融中心的跨世纪工程，因此大量出现的是以前不曾遇到过的新情况、新问题。

(14 句 14 个用例)

附录三 名词与名词分离类

ID9973	吉林省长春市今年规划的首批经济适用房工程，目前已全面开工，并且有望在今年年底首期竣工六十五万平方米。
ID9349	长期依靠调进粮食解决吃饭问题的贵州省，由于农业连续六年获得丰收，今年粮食总产量再创历史最好水平，初步实现了农村人口粮食自给。
ID9312	民政部有关负责人表示，在“晚会”期间正式认捐且经核实准确无误的单位和个人，其认捐款物已全部到位。
ID8869	“博格坎普在世界杯结束前状态还不错，但是今年下半年，他的状态就差多了，”胡子长过头发的维阿里说。
ID7008	张学良在东北建设中，现代化意识较强，大胆开创。
ID3881	跳水选手已全部抵达罗马，并开始赛前训练。
ID3787	俄罗斯驻扎在波罗的海三国的西北军队集群的最后一批应撤离人员已于 3 1 日全部撤出拉脱维亚。
ID3360	科尔认为，俄罗斯军队最终全部撤离德国是“欧洲战后历史的终结”。
ID284	迄今，呼伦贝尔盟十三个旗市已全部进入全国电话直拨网，并成为中国最大的电话本地网。
ID2759	这部分资金将全部来源于国家金融机构，不对社会公众发行金融债券。
ID10702	《放生》这个集子里写了许多老人，其中固然有些像〈城仔落车〉里的阿婆那样无知艰困，不过更多的却是有如〈青番公〉那样满肚子不合时宜智慧的。
ID11309	戴胜通反问：「台湾如果环境不好、不好赚钱，大企业为什麼都把他们在国外的第二代叫回来接班，在大企业国际化潮流如此普及的今天，他们尽可以让年轻一代

	留在外面创业!」
ID5347	邝其志说, 香港除储备庞大外, 政府并无负债, 累积储备约为四百八十亿美元。
ID9762	三峡工程虽然投资巨大, 但在全国所占比重很小。
ID6138	据了解, 岭澳核电站四亿美元的股本金目前已全部安排到位, 其余资金由国家开发银行负责从国内外筹措。
ID7253	这些案件包括: 天津“二·一三”贩卖非法图书案、河北正定“四·一五”政治性非法出版物案、北京贩卖政治性非法图书案、湖北贩卖政治性非法图书案, 共涉及非法出版物及图书三万多册, 目前案件涉及的非法出版物已经全部收缴。

(16 句 16 个用例)

附录四 副词与数词分离类

ID9953	据估计, 十年来大约有两万只藏羚羊被杀。
ID9872	目前在站的博士后约有七百余入。
ID9871	截止今年九月, 中科院共招收博士后近两千两百人, 其中留学回国人员二百五十人。
ID9837	由于各大景区软、硬件建设不断加强, 全县旅游业得到迅猛发展, 年平均接待海内外游人一百万人次以上。
ID9820	中国是一个少林国家, 森林面积仅占世界森林总面积的百分之三点三。
ID9794	展交会共设二百多个摊位, 来自全国二十多个省、市的六百多家企业, 北大、清华、复旦等六所高校, 以及中科院、电子部、机械部等所属八十五所科研机构云集安徽省科技成果展示中心, 对二千多项专利、科研成果进行展示和交易。
ID9763	工程总投资仅占建设期间国民收入累计值的千分之一点二三, 占全国基本建设投资累计值的千分之九点八。
ID9708	今年上半年, 南通市共完成自营出口总额十点七三亿美元, 在江苏仅次于苏南的苏州和无锡
ID9650	共筹资三百零六亿港元, 折合约四十亿美元。
ID9600	据介绍, 到目前为止, 中国教育部仅批准了全国四所高校开展现代远程教育试点。
ID9499	集团全年共生产原油超过一亿吨, 天然气一百四十八亿立方米。
ID9476	据有关方面统计, 今年全年台湾同胞到祖国大陆约为一百七十万人次。
ID9444	但至今仅建成两座污水处理厂, 十四座正在施工, 其余十七座还处于“纸上谈兵”的境地。
ID9436	但他们指出, 这些企业的污水排放量只占全流域污水排放总量的四分之一左右, 污水的很大部分来源于城镇居民的生活污水。
ID9356	全省国土面积十七点六万平方公里, 喀斯特岩溶地貌占百分之七十三, 人均耕地只有零点八亩。
ID9309	截至目前, 捐款到位五点一四八四亿元, 捐物到位五点六八亿元, 款物共到位十点八二八四亿元, 占当初认捐数的百分之九十九。
ID9303	巴西本届甲级联赛的最佳射手为效力于桑托斯队的维奥拉, 他共射入 2 1 个球。
ID9263	这次展览共展出其中的五百多幅作品, 完全采用组照的形式, 分为人物服饰、民居建筑、生活方式、节日文化、饮食文化、宗教文化、人生礼仪、传统体育和民间游艺、综合等九个专题, 配有详细的中英文说明。
ID9262	评选委员共来自荷兰、日本、英国、瑞士、法国等七个国家, 也是国内历次摄影比赛中外国评委最多的一次。
ID9261	据悉, 这次是世界上第一次以民俗文化为主题的摄影大赛, 共收到来自世界四十

	多个国家和地区的一点六万多幅作品，是中国历来参赛范围最广、规格最高、文化内涵最丰富的摄影比赛之一。
ID9228	以往为堵塞漏洞，需要来往于银行和海关之间对纸制报关单进行二次核对，在本地核对至少需三天，异地核对就要近一个月，这在很大程度上影响了正常进出口业务的贸易效率。
ID9198	他在昨天的电话中说：“只有一种方式可以说明问题：让全体球员投票表决。”
ID9190	坏消息伴着袭来的寒流滚滚而来，整个联赛距离冻僵仅余 8 日。
ID9164	该团赴台五十天，共演出二十四场，尤其是在歌仔戏的发祥地宜兰文化中心与兰阳戏剧团联谊演出引起极大轰动。
ID916	据统计，目前中国共有国有企业三十多万家，其中国有大中型企业一万四千五百多家。
ID9067	欧洲央行只拥有 5 0 0 亿欧元的储备金，由各成员国央行根据本国在欧元区内的人口比例和国内生产总值的比例来提供。
ID9064	欧洲央行委员会的决策采取简单多数表决制，每个委员只有一票。
ID892	它所取用的水取自流经朱砂土质的赤水河，原料则选用当地特产高粱及优质小麦等，在端午至重阳节之间酿制，共需经过两次投料，七次蒸馏，八次发酵、九次蒸煮，历时十月酿成后，还须放入特制的陶土酒坛中密封贮存三年以上，最后由经验丰富的老酒师精心勾兑而成。
ID8890	通过对医疗照顾人员三万元以上支出的大额病例进行核查，有关部门共拒付违规金额四百七十多万元。
ID8887	北京市卫生局副局长金大鹏说，通过实施“总量控制，结构调整”的改革措施，按可比口径计算，今年前九个月共为国家和企业减少支出十点九亿元。
ID888	直到一九八七年年产量才达到一千二百吨。
ID8606	按照转变职能、政企分开的要求，通过改革，各部门转交给企业、社会中介组织和地方的职能共有二百多项，人员编制总数基本达到精简一半的预期目标，政府效率进一步提高。
ID8546	三十六个常设司级政府部门，至今以中文为母语的本地公务员担任司长的只有数位。
ID850	预计今年发展中国家经济增长百分之五点九，而明年仅增长百分之四点九，比 1 0 月份的预计低了一点三个百分点。
ID8470	夏普公司于今年 1 月推出的一种电视机的辅助设备能耗仅为 0 . 9 瓦。
ID8408	叶震在此间举行的新闻发布会上透露说，今年一到十一月，中国共完成房地产投资两千四百九十亿元，比去年同期增长百分之十八点三。
ID8397	增加投资、扩大内需政策的效果已经有所显现，尽管全年的社会消费品零售总额只比去年增长百分之六点八，考虑到今年市场物价仍在低位运行，消费品的物价指数是负增长，扣除物价指数后的市场销售实际增幅可能在百分之九左右。
ID839	城市北出口道路建设工程已与香港泰华公司达成建设协议，投资约需一点八亿元，南绕城快速路工程也有数家外商前来洽谈投资。
ID8376	在这批重点企业中目前只有十四家还在加紧施工。
ID8133	据中国最早几间采用 D N A 亲子鉴定的中山医科大学鉴定中心透露，今年该中心已做了五百多宗亲子鉴定，比去年多了二百多宗，比一九八四年整整多了十倍。
ID7954	日本大真空公司日前宣布研制出了据称是世界上最小的晶体振荡器，其体积只有 0 . 0 1 立方毫米。
ID7942	保体育十佳评选已有 4 1 年历史，参加今年评选的共有 1 9 7 名体育记者。
ID7740	而 1 9 9 5 年还只有南非和北部非洲五个因特网联接，这表明非洲科技已开始与国际接轨。
ID7700	到 1 9 9 8 年上半年，全世界共拥有电脑四亿台，全世界因特网使用者 1 . 4 亿，

	预计到 2 0 0 1 年因特网使用者将达三亿。
ID7637	法律共有十二章二百一十四条。
ID7532	还可以有另一种赛程：每周依然是 3 场比赛，共进行 4 2 场常规赛，这样可以避免球员受伤，保持比赛水准。
ID7530	本赛季如果 2 月 1 日开战，最多可安排 4 9 场常规比赛，如果 2 月 7 日揭幕，可安排 4 5 场赛事。
ID7496	石油工业是尼日利亚经济的支柱产业，其收入约占政府预算收入的百分之七十。
ID7495	据业内人士估计，今年全年石油收入可能只有 7 4 . 5 亿美元，而 1 9 9 6 年和 1 9 9 7 年则分别为 1 0 8 . 9 亿美元和 1 1 9 . 9 亿美元。
ID7494	今年头 1 0 个月石油总收入仅为 6 9 . 9 亿美元。
ID7478	据《中国青年报》报道，深圳姑娘嫁往香港最多的一九八八年是一千零六十九人，而今年一至十一月份仅有二百九十九人。
ID7469	巴黎 C A C 4 0 种蓝筹股平均价格指数比前一交易日仅微升了 0 . 6 8 点，以 3 8 7 3 . 1 0 点收盘。
ID7448	在技术股带动下，道一琼斯 3 0 种工业股票平均价格指数 2 8 日连续第七个交易日上升，但全天仅微升 8 . 7 6 点，报收于 9 2 2 6 . 7 5 点。
ID7390	次仁顿旦上任时，村里帐面上只有三百一十八元人民币，他就垫支四千元，购买电线杆，又组织村民采挖水晶矿增收四万元，修建了年久失修的老电站，带领村民兴建库容两万方的蓄水库和一座三十七千瓦的电站，不仅全村用上了电，而且解决了六百多亩农田的灌溉和村里的人畜饮水。
ID7374	据统计，截至目前，全省共有五十二家公司的股票在上海和深圳证券交易所上市交易，数量居中国第三，还有三家公司的股票在香港联交所上市。
ID7367	多米现在法国巴黎圣日耳曼队踢球，身价大约为 3 0 0 万英镑。
ID731	据了解，全世界每年草药的销售额约一百五十亿美元，中国中药出口只有六亿美元，而且百分之七十以上是未加附加值的药材。
ID7294	北大法律系的台湾学生杨惟婷告诉记者，一九九二年她随父亲在上海读书时，全校只有她一个台湾学生。
ID7281	而根据国家教育部的统计，到今年夏天，祖国大陆共有一百所大专院校招收台湾学生，在校台湾学生超过一千五百人。
ID7197	但两轮下来，中国只有两人进入八强，日本棋手也折损了三分之二；但韩国的 5 名棋手只有小将睦镇硕今天惜败于中国的常昊，其他 4 人全都顺利过关，这加重了他们争夺本次大赛桂冠的筹码。
ID7196	本届比赛中国按照东道主优先的惯例，选派了 9 名棋手，而日本和韩国分别只有 6 名棋手和 5 名棋手。
ID7188	比赛共进行五轮，三盘两胜制的决赛将于明年 6 月在上海和南京两地举行，冠军得主将获得 1 5 万美元的奖金。
ID7119	中巴车上共有九名乘客。
ID678	1 9 9 7 年发达国家经济形势的特点是美国增长强劲而日本经济疲软，美国经济增长率估计为百分之三点七，日本仅为百分之零点八。
ID67	目前，约有十五万家外商投资企业在中国银行开立帐户，其中二万多家获得中国银行的贷款支持。
ID6691	干相同的活，别人记十个工分，他却只能得七分；集体分东西，他分到的一份肯定最差；到了结婚年龄却没有姑娘愿意嫁给他。
ID6665	根据目前院方掌握的情况，中国在近几年有三例五胞胎出生，有两例全部死亡，另一例也只保住了一个。
ID6663	出生的当天，这个在 B 超上没露面的小宝宝给了医生一个措手不及，因为事先只准备了五个包孩子的婴儿被子，老六出生的时候和第一个出生的大哥在手术室里

	暂时“挤”在了一个急救台上。
ID6661	其他四个婴儿的情况相对稳定，为了让他们的胃肠功能保持正常，从周五开始，医生已经开始用一种新生儿专用的输液泵为孩子们“喂奶”了，每隔两小时一次，每次只有两毫升。
ID6656	第四个出生的女婴也只有六百一十克重，他们同时出现肺出血。
ID6654	第五个出生的男婴体重最轻，只有四百九十五克。
ID6603	丹顶鹤系国家一级保护珍禽，现全球仅存一千多只，每年来盐城滩涂越冬的丹顶鹤多达六百余只。
ID6565	一九九七年，中国的人均储蓄存款超过三千元，而二十年前每人每年平均才存二十元钱。
ID6558	当时，中国的电话普及率只有百分之零点三八，被称为“没有电话的国家。”
ID6555	一九七八年底，中国的外汇储备只有一点七亿美元；目前，中国的外汇储备已位居世界第二位，达一千四百多亿美元，增长了八百二十多倍。
ID6540	以寿光市洛城镇绿色食品蔬菜示范基地为例，该基地目前共有高标准蔬菜大棚二百五十个，引进推广了海内外近百个蔬菜品种。
ID6532	每个蔬菜大棚投入大约都要两、三万元，每年各项费用约三千到五千元。
ID6482	九十年代初期，中国的程控交换机大部分依赖进口，价格在每线三百美元左右，但随着国产程控交换机竞争力的增强，价格逐年下降，目前仅为每线九十美元。
ID6434	在六、七十年代，海内外约有一千万人到这里参观、学习。
ID6410	曾是哈尔滨工业大学高才生的这位“狗司令”，九年前开始建设当时只有四条警犬、七个士兵的基地。
ID6083	一九九六年，长飞公司共出口光纤产品二十一多万公里，实现创汇三千多万美元。
ID6053	目前全市上亿元的三资工业企业已达五十一家，而在一九九四年仅有二十三家。
ID5976	目前为止，中国只开放了五个经济特区和北京、上海、天津等六个城市作为试点。
ID5964	据统计，北京电信部门全年共投资五十亿元用于城市电信基础设施建设，新铺设光缆一千六百多条，光缆总长度达六千二百多公里；新开通各类电话交换机一百一十万门，使市内电话交换机总容量达到三百八十万门，移动电话总容量达七十六万门。
ID5911	这位发言人说，到目前为止，中国的有关部门和机构在美国市场共发行债券四十七点二亿美元，其中三分之二是财政部和国家开发银行发行的具有国家主权信誉和准主权信誉的债券，其余的商业性发行主体皆为中国实力较强的银行和非银行金融机构，具有良好的信誉。
ID5878	日本政府承诺的贷款约占外国政府向中国承诺贷款额的百分之四十以上，居首位。
ID5873	截至今年六月，中国共批准日本企业对华投资项目一万五千六百七十个，协议金额二百七十六亿美元，实际使用金额一百五十九亿美元，居中国吸收外资国别第二位。
ID5839	北京淄博工业产品展示暨投资贸易洽谈会规模较大，共有三十五个行业的六百多家企业参加；同时将推出二百二十个重点项目对外招商，主要涉及化工、医药、纺织、建材、机械、电子、陶瓷等行业。
ID5769	台经院研究六所所长董瑞斌则认为，目前金融风暴力道已减，各国如果一切顺利——政局持稳、没有暴乱、经济及金融改革措施也能贯彻实行——的话，「大约需要二、三年的复健期」。
ID5751	金融风暴发生后，大陆订单的确有需求减缓的现象，幅度约在二至三成之间，宜进受此拖累，营业额也衰退将近一成。
ID5657	一台三星制的摄影机不到台币一万元，一件帅气的防风夹克，则只合台币二百元。
ID5585	目标为本课程评估组副院长苏汉波今日（星期二）表示，教署迄今共接到来自二百五十间学校约二千一百名教师报名，申请九节研讨会的三千个名额。

ID549	在资本市场、保险业、连锁商业、批发销售、集装箱运输、房地产、信息服务等经济增长点的带动下，第三产业增加值预计年均增长百分之十一至十二，第三产业占国内生产总值的比重也将由一九九六年的百分之四十三，上升到二〇〇〇年的百分之四十八。
ID5434	在中心启用的高峰期，约有一百六十至二百名新闻主任、一百名行政主任和其他职系的工作人员在新闻及广播中心内提供支援服务。
ID543	而在一九九二年，这项数字仅为一千五百美元。
ID5389	此外，今年上半年共有十一万零九百八十份楼宇转让契约送交该处注册，较去年下半年的八万一千二百六十五份增加百分之三十六点六，较去年同期的七万八千二百五十五份，则增加百分之四十一一点八。
ID5387	一九九七年上半年送交该处注册的楼宇买卖合同（包括住宅及非住宅单位）共有十一万五千六百七十三份，较去年下半年增加百分之四十四点六，与去年同期比较，则增加百分之七十一一点五。
ID5365	沙田中文大学地区在二十四小时内共录得约八百毫米的雨量。
ID5362	土木工程署辖下土力工程处的紧急控制中心自七月二日上午六时二十五分发出山泥倾泻警告後，共接获逾一百四十宗山泥倾泻报告。
ID5339	「在过去五年，房屋委员会亦兴建了大量的资助自置居所（八万零二百七十五个），其中很多是被居於拆卸单位的住户所购买。此外，因向房屋委员会或私人机构购买自置居所的住户所腾空的出租单位亦已再分配予公共房屋轮候册上的申请人。平均来说，每年约有一万四千个在轮候册上的家庭获编配到租住单位。」
ID5326	截至今日（星期四）上午九时，本港各水塘之总存水量共为五亿四千五百八十万立方米，占总容量百分之九十三点一。
ID5309	他们於毕业後须在香港至少工作三年。
ID53	“八五”期间各项存款比“七五”（一九八六至一九九〇年）末净增五十亿元，年平均增长百分之二十七点四九。
ID512	设备从安装到投产只用了三个月时间，开发了企业自己的专利技术，为公司下一步对外输出氨纶丝生产技术奠定了基础。
ID494	宋健说：“如今，中国已能生产上万门数字电话程控交换机，而且这种交换机的总设计师只有二十八岁，武汉大学毕业的。”
ID489	中国国家气象局购买美国克雷公司的大型计算机，克雷公司只卖给我们两台处理器。
ID4829	共接待了来自泰国、新加坡、印尼、马来西亚、美国、加拿大、法国及港澳台等十三个国家和地区九千九百二十八人次。
ID4828	据统计，到目前为止，全市共接受华侨、港澳同胞捐赠近五亿元人民币，分别用于兴建学校、医院、民房、修路、造桥、建侨联会址等各项社会事业。
ID481	他们分为5个小组，共约见了约150位议员或议员助手以及政府官员，用他们在中国投资办企业的经验，阐述延长给中国贸易最惠国待遇的必要性。
ID4784	其中由美国市场有限公司联合四家美国公司共投资二亿美元，联合开发北海市涠洲岛旅游项目，美国纯品集团也投资二千万美元，在北海兴建一座大型果汁加工厂，而来自英国的西门子普莱赛机构有限公司则投资二千六百万美元，与北海市合作兴建北海机场候机楼。
ID471	自一九九一年投产以来，共完成技术改进和开发项目近千项，独立开发了包括大屏幕彩电玻壳在内的二十多个新产品、新品种，年产量超过设计能力百分之七十八，去年已达七百一十二万套；产品的国内市场占有率和品种覆盖率分别达到百分之七十和百分之九十，并返销日本，出口韩国和印度等国家。
ID4662	五百家企业共涉及四十三个行业，主要集中在电子、交通运输设备制造、电气机械制造和食品加工等行业，其中，资金、技术密集型的企业销售额占总销售额的

	百分之六十六点九，比一九九五年增长百分之二十七点四。
ID4660	全国最大的五百家外商投资企业的资产总额为四千九百一十五亿元，比一九九五年增长了百分之四十，其中企业资产额超过二十亿元的企业有四十六家；销售总额达四千六百九十八亿元，增长百分之二十七点八，约占全国已投产开业外商投资企业销售总额的四分之一；利润总额达三百六十亿元，增加五十八亿元，约百分之九十的企业盈利。
ID4557	而对中国来说，这一进口费用仅占其百分之二的出口。
ID4541	去年，舟山市共有二百六十艘渔轮在远洋和异国海域作业，捕捞量超过十万吨。
ID4533	去年，舟山市渔业产量达到一百零五万吨，相当于一九九〇年的两倍，在中国海水产品总产量中约占十分之一。
ID4450	西北地区企业与东部地区技术提供单位共签订技术合作合同二百八十七项，总金额达二十六点九亿元。
ID4423	在境外融资方面，迄今为止，工商银行已与日本、新加坡、英国、意大利、奥地利、瑞士、德国、美国等二十个国家签订了各类贷款协议一百五十多个，对外借款余额近八亿美元，共支持国内二百多个项目的建设。
ID4331	据这位官员介绍，目前中国共开发油田三百五十个，气田一百一十个。
ID433	天津作为中国北方重要的工业城市和有广阔腹地的对外贸易港，去年向俄罗斯出口的轻工业品和食品、畜产品占天津对俄出口额的百分之五十；而从俄罗斯进口的钢材和有色金属约占从俄进口总额的百分之六十。
ID4257	据了解，到去年底浦东仅有日本富士银行上海分行和泰华国际银行两家外资银行。
ID4189	迄今，约有一百五十多家跨国公司在这里投资。
ID4171	参赛的共有 1 2 支球队。
ID4074	截止今天的统计，正式报名参加本届运动会的共有 4 2 个国家和地区的 2 0 7 6 人，其中运动员 1 4 2 3 名；报名采访运动会的新闻记者有 1 3 6 0 人，其中境外记者 9 8 名。
ID4066	但令人满意的是，这次比赛涌现了一批十六七岁的新秀，年龄最小的只有 1 4 岁，其中不乏希望之星。
ID4064	但令人满意的是，这次比赛涌现了一批十六七岁的新秀，年龄最小的只有 1 4 岁，其中不乏希望之星。
ID4059	东道主河北队共夺得 9 枚金牌，列金牌总数第一。
ID4037	英国百米好手克里斯蒂在三天内两度失利，他在柏林的男子百米比赛中仅获得第三名，成绩为 1 0 秒 0 2 。
ID3966	今天共进行了男子自由体操、鞍马、吊环和女子跳马、高低杠个人单项赛的 5 个项目的决赛。
ID3965	本届锦标赛女子个人全能银牌得主罗马尼亚的米洛索维奇出现大失误，只获第六名。
ID3963	尽管她的动作潇洒自如，但难度无法与罗莉相比，只获得 9 • 8 7 5 分，夺得银牌。
ID3878	本届锦标赛共设游泳、跳水、花样游泳、水球和长距离游泳 5 个大项。
ID3877	组委会发言人说，“总共有近 3 0 0 0 人将为锦标赛的成功而努力。”
ID381	“八五”期间，中国共批准外商投资项目二十二点九八万个，实际利用外资额为一千一百四十七亿美元，分别比“七五”增长了九点一倍和六点八四倍，而且外资更多地流入国家鼓励投资类项目，基础产业和基础设施项目已成了外商投资热点。
ID373	据悉，在加工贸易台帐制度试点城市东莞市，去年十一月二十七日到今年三月十五日东莞地区海关共接受企业合同备案八千四百多份，比试点前略有上升，企业反应良好，普遍表示接受。

ID366	去年广东省整个加工贸易出口值达四百多亿美元，约占全省出口总值的百分之八十。
ID3520	据统计，到今年七月底，天津市共批准台商投资企业八百七十一家，协议投资额九亿多美元，其中台资占百分之七十三点八。
ID349	去年全年共完成进出口总额一百零九点九亿美元。
ID3398	在去年 1 1 月举行的援助越南巴黎国际会议上，越南共获得日本、法国、韩国和瑞典等国的 1 8 • 6 亿美元的援助。
ID3335	一九九三年一年，全省人事部门人才交流中心共登记求职者四十六万三千多人，办理调动和聘用手续的六万二千人，其中具有中高级专业技术职称的占百分之四十五点八，大学本科以上学历占百分之七十八。
ID3283	国内沿海省市参展企业增加，共有来自三十个省、市、自治区的 30 名企业代表到会。
ID3205	这次战争中包括德国在内的欧洲国家共死亡 4 0 0 万人。
ID3183	拉美一体化协会成立于 1 9 8 1 年，共有 1 1 个成员国，它已接纳 1 5 个国家和 5 个国际组织为观察员。
ID316	目前，大连市出口商品结构中，初级产品只占据不到三成，而工业制成品比重达七成以上。
ID3134	据博览会组织委员会介绍，这次博览会展出规模为历届之最，共展出图书九万种，其中海外原版图书三点五万种，内容涉及科技、经济、文化、教育和社会发展等众多领域。
ID3061	一百多年来，世界同时代的始祖鸟总共发现 8 块，且仅局限于这一地区。
ID2844	克拉伊什尼克在谈话中进一步解释说，“我们拒绝接受版图划分方案仅仅是拒绝波黑和平方案的一部分。
ID2777	为加快示范区的建设，山东省共投入资金六千二百多万元，建设面积达二十一点八万平方米，示范区基本实现了水、电、路、通讯四通。
ID2745	他否认了轨道站上食品储备只够维持 1 0 天的说法。
ID2720	因此，为这颗卫星的开发共投入了 4 1 5 亿日元。
ID2715	“菊花 6 号”卫星在 8 月 2 8 日下午被“H 2 — 2 ”号火箭送上椭圆形轨道后，原定要通过卫星上的远地点发动机 3 次发动使其逐步进入地球同步轨道，这中间大约需要一个月时间。
ID2667	私营工业、商业企业共有五千八百三十六家。
ID263	乡镇企业创造的国民生产总值约占福建省国民生产总值的三分之一。
ID2609	我仰头望去，好家伙，至少还有 1 0 0 0 多米高，陡峭的坡壁不下 6 0 多度，近处有几百米宽的平坡，在一排排绿树和枯树间裸露着一斑斑紫褐色的砂砾，这显然是 1 7 0 7 年最近的那次喷发的火山灰的残迹。
ID2604	富士山离东京只有 8 0 公里——这显然又是造化的特意安排：这个岛国最高的山与她的最大的城市相依为伴，互为辉映。
ID26	三年多来，这些城市社会经济发展迅速，地方经济实力明显增强；经济年平均增长百分之十七，高于全国年平均增长速度。
ID2559	《双向式英语》是学习英语的重大突破，只会 A B C 的人仅用几百个小时，配合教材反复听录音，说英语如同讲普通话一样流利通顺。
ID2498	在连续八天的劳动中，大学生们共清理六十六万平方米的冰层，占整个清冰工程的百分之七十。
ID2377	鲍恩曾到现场采访了这个被炸毁的防空洞，大约有 4 0 0 名平民丧生。
ID2343	目前苏联在波兰的驻军约为五万人。
ID2099	据广西壮族自治区交通部门透露，到目前为止，广西共利用世界银行、亚洲银行以及荷兰、韩国等国贷款一点五亿美元建设交通水运基础设施。

ID2083	目前，德国西门子公司在中国共设有三十四个合资公司、三家独资公司和十四家办事处。
ID2048	两家深圳企业近日先后在香港招股，市场反应空前热烈，共筹集近十亿港元巨资。
ID2029	据中国国家开发银行的官员透露，截止去年底，该银行共为二十四个国家重点项目进行外汇融资，其中已经完成签约的项目有十八个，对外签约总额约三十四亿美元。
ID1997	据了解，中国目前大约有六千六百万平方米商品房积压。
ID1941	一九九七年江西实际利用外资比去年增加五成左右；外商直接投资几乎增长了一倍。
ID1910	但是，在过去三年中，墨西哥贸易顺差额呈大幅递降趋势：1995年为78.89亿美元，1996年为64.24亿，而去年只有5.82亿。
ID1874	天津市台办主任曲耕莘说，尽管去年由于台湾当局的阻挠，台商在天津投资的增长幅度有所放缓，仅新批台资企业九十四家，协议投资总额达二点五亿美元，均比上年有所下降，但投资的质量和技术水平都有进一步提高。
ID1759	继青岛啤酒和山东华能分别在香港和美国成功发行H股和N股以来，去年又有山东晨鸣造纸集团、济南轻骑集团、烟台张裕集团、新华制药集团和鲁泰纺织集团等五家股份有限公司上市发行外资股，共吸纳外资二点六五亿美元。
ID1687	根据协定，该项目总投资额约为二千六百万美元，第一期投资一千万美元，在厦门市开办二十家示范便民连锁店和一个现代化的配送中心，然后根据市场拓展陆续在福建省九地市共开办三百家便民连锁店。
ID15069	澳洲政府每年编列预算，去年约有四千万澳币（相当於台币七亿元）接受大家申请，只要有出色企画案并找到片商和发行公司资助，就有拍片机会；至於投资金额则视题材而定。
ID15002	由於从采茶到做茶一气呵成至少要十多个小时，茶农每日枕著茶香睡，伴著茶香醒，这样的日子几乎得连续近个把月，至今不变。
ID15001	坪林人一年的生活作息更是跟著茶事走，尤其是春茶秋茶采收时节，家家户户几乎每天只睡不到四个小时。
ID1500	去年，纽约新上市的外国企业共有61家，创历年来最高纪录。
ID14788	「《破袜子》我只花一个半月就写完，但我更早的另一个长篇《离魂》却耗了一年半，」霜子表示，「写完」是她最大的压力源，为了纾解这样的压力，她总是在连载的一开始就言明「别问我进度，没再看见可能就是不写了」。
ID14747	阿里山，台湾第一处桧木砍伐区，伐木前的调查，有三十万棵以上的千年桧木，如今只剩几棵孤独老树，在外来的柳杉、樱花中，显得垂垂颓矣。
ID14711	中海拔雨大风小，终年云雾笼罩，又称为雾林带，稳定的湿乞孕育下，桧木长成东亚第一大针叶树，平均高三十至四十公尺以上。
ID14708	如今全球桧木只剩六种，海天各据一隅，分散在北美、日本与台湾。
ID14574	张绪中以自己待了十多年的中华电信为例，明年公司的人事预算比今年足足少了三千人，意味著有三千人将在一年内被「逼退」。
ID14496	根据国外研究，中年人忧思虽多，但会严重到需要求助於心理谘商或医药的，大约只占百分之十五。
ID14124	友邦萨尔瓦多发生强烈地震，位於首都附近的柯林纳斯社区至少有五百四十九人因山崩而罹难。
ID14001	在屏东县三十三个乡镇中，林边的面积倒数第二，只有十五.六平方公里，人口约二万三千多人。
ID1396	公司采用兼并、收购等形式，共利用和盘活社会资产一百多亿元；结合自己重点发展项目吸引海外资本，一汽先后与八个国家和地区建立了十八个合资企业。
ID13801	目前台湾主要的健身中心型态大致可分为「土」、「洋」两种。

ID1379	据悉，该行去年结售汇业务分别达到了九百一十亿七亿美元和四百七十一亿美元；全年共办理国际贸易结算业务一千三百九十八亿五千万美元。
ID13786	他喘著气说，进健身房半年来，原先不爱运动的他，生活有了极大的转变，不但整个人整整瘦了一圈，气色也变好了。
ID13605	「以日本为例，朝日新闻、读卖新闻发行量千万份，但只有一页分类广告，求职、转职者仰赖的是情报志，」李秉峰说，台湾的哈日风盛行，加上书籍销售通路也在改变，以前买书到书店，但现在便利商店的流通速度早已超越了传统书店，「大环境改变加上市场成熟，正是切入时机。」
ID13603	有多年报纸发行经验的副总经理李秉峰分析，目前的报纸分类广告遭遇很大瓶颈，从 ACNielsen 媒体大调查显示，为了看分类广告买报纸的人只占百分之三。
ID13594	去年十月台湾的失业率一举突破百分之三的关卡，估计约有三十多万人失去了职场的舞台。
ID13535	总人口约为二万八千多人。
ID1336	据统计，五年来天津港保税区共吸引了来自世界七十三个国家和地区及国内二十五个省、市、自治区的三千多家企业投资，累计协议投资总额达三十多亿美元，协议外资总额达二十多亿美元。
ID13318	当时其他画室一个月的学费大多在四十五元以上，台籍老画家的画室束口更要六、七十元，对于一个月军饷只有二十五元的夏阳而言，再怎么缩衣节食也凑不出学费来，除了一个月二十元的安东街画室。
ID13278	而两辆借自台湾收藏家黄任中和科学博物馆的铜马车复制品，总共由七千多个部件组装而成，尺寸为实物的二分之一。
ID13222	自去年六月到十二月十日止，全国各检察机关共受理二千多件与黑金相关的案件，起诉民选首长及各级代表共三十人，查获贪渎金额超过新台币二十五亿九千多万元。
ID13204	以推动人权教育为宗旨的人权纪念教育基金会执行长周碧瑟指出，这一年内又收集到简彰佑等一百八十位受难者名单，总共将有六百九十位政治受难者的姓名刻于纪念碑上。
ID13143	他认为，台湾量贩市场竞争也许激烈，但不见得饱和，数字也可以佐证，台湾量贩市场的营收只占零售业的百分之二十，量贩市场的潜力还大得很。
ID13129	好市多没开之前，一个月去四次传统市场，现在顶多去两次。」
ID13126	夜市仿冒的 CK 和 Levi's 等名牌牛仔裤，一千五百多元，但好市多只要一半价钱。
ID13105	去年十一月中象神台风过后，蔬菜价格飙涨，大润发推出周年庆特价，一颗高丽菜十五元，两天内全省分店卖出七万多颗；五元玉米条的销售数字更是惊人，两天约卖出十二万个。
ID13078	据估计，全台占地数千坪的量贩店约有百家。
ID13017	换句话说，只有这三个国家的「技术贸易」享有顺差，可以靠贩卖知识赚钱。
ID13016	然而综观全球，全世界一百多个国家中，只有美、日、英三个科技大国能够做到「技术输出」多于「技术输入」。
ID12933	本届残障奥运，亚洲国家中，南韩队成绩最优异，总共夺得五面金牌。
ID12927	在竞赛上，此次地主澳洲队一共拿到六十四面金牌，排名第一；而中华队也有所斩获，共夺得一金二银四铜。
ID12920	为期十二天的残障奥运可以说场场观众大爆满，总共卖出将近一百一十多万张票，显示这次残障奥运受到各界青睐。
ID12863	薛晓岚已经记不清楚在资讯人草创时期，有多少个夜晚是在办公室度过；他们也曾为了找经销商卖书被倒帐，让公司足足有两个月没进帐的受挫经验。
ID12804	区区二万六千人的下营小镇，只有四间小学、一所国中，左邻右舍不是同族就是同学、同事，那里才真有家人的感觉。

ID12786	「论精致、论多元，台北不会输，但论起大刀阔斧的建设，台北就差了一截，」陈威仁不讳言，像是台北卫生下水道吵嚷多年，至今铺设率只有四成九，表示台北的家庭污水有一半是直接流进了河川，先进国家首都是不会容许这种情形存在的。
ID12779	「三万名外佣中，每年到期遣返的约有一万人，如果这一万人离去时都怀著对台湾社会及雇主的仇恨，回去经过口耳相传，台湾的外交还有什麼希望？」
ID12734	庞大的市府支出若平均分配给二百六十万台北市民，每人足足可分得六万八千元，远超过高雄市的四万二、台南市的二万五或是台北县的一万七，难怪别的县市愤愤不平，直指台北市民是吃香喝辣、占尽便宜的「一等国民」。
ID12732	据估计，至少有七成五以上的中大型国际研讨会在台北举行，九成以上的跨国企业集团将台湾总部设在台北；至於全国一百大企业，尽管工厂遍及中南部，却有一半以上选择以台北为总公司所在地。
ID1268	过去五年间，土地开发商共投入四十亿港元用于电厂、区内主干道及地下管网、土地平整、邮电通讯等基础设施建设。
ID12382	我国目前共有二十九个邦交国，半数位於中南美洲及加勒比海地区，此区堪称我国外交重镇。
ID12327	郭东源指出，目前已知治疗疾病的药物标的只有五百个左右，人体基因解码後，推算标的会增加十倍。
ID12304	肠病毒共有六十八型，肠病毒检测晶片可以判别出是否感染肠病毒？是不是致命的肠病毒七十一型？是不是容易与肠病毒七十一型混淆的沙克奇 A16 病毒？
ID12201	那时仅有一百多支上市股票、涨跌都掌握在主力炒手手上的台湾股市，被视为是「吃人市场」，众人避之唯恐不及。
ID11918	目前在帛琉经营渔业公司的台商只有戴荣辉一人，获利颇丰。
ID11876	「彩虹的尽头」听起来仿佛遥不可及，但实际上，她是台湾目前距离最近的邦交国，直航只要三个半钟头的航程，或许「彩虹的尽头」对台湾人而言另外有两个更贴切的隐喻：不熟悉的美丽遥远国度，和对观光旅游或经贸投资都是商机处处的处女地。
ID11817	异乡讨赚的日子只持续了两年，吴建乐又回到故乡，那时海口一带才刚接上电线，入夜时的灯火还令乡民兴奋得难以入眠呢。
ID118	其中大专院校、大型科研院所和大中企业开办的企业约占百分之七十。
ID11545	由於远东企业集团捐建的组合屋只有五十户，过去租屋或住在违建中的弱势族群根本没申请资格，为此大隘社极力奔走，为老伙房倒塌、伤亡严重的慈云巷居民设法建组合屋。
ID11507	至於高学费政策，吴思华以政大为例，一年预算二十一亿台币，花在每位学生上的成本大约是二十万，如果真算成本，一个学分一万元还算是便宜。
ID11472	像李孟娜这样的女性经理人，在各校的在职生都是少数，大约只占五分之一比例。
ID11443	台大管理学院八十六学年度首次试办在职进修，招收四十四人，吸引了九百九十人报名；第二年增加录取名额到九十五人，结果来考的人更多，一共有千四百人。
ID11440	台大的情况则是，包括电机、三民主义、生物等九个所，去年首次开办在职专班，预计录取一六七人，但报名总人数却只有三三〇人。
ID1137	最近，“段和段律师事务所”已将总共一百八十五万五千美元的赔款，汇给这三家企业。
ID1128	去年四月，C R 公司开始了其破产程序，其所有资产的偿债率仅为百分之五。
ID11275	连战在西安出生，念的第一所小学叫「作秀小学」，因战局混乱，只念了一年就转学，当时作秀二字是「作育秀苗」的意思，不料到了台湾却变成在媒体前表演、主动推销自己的意思。
ID1125	中国企业状告美国第一国民银行赔偿一案，经上海市“段和段律师事务所”一年

	多的艰难诉讼，最近在美国新墨西哥州联邦地方法院达成调解协议，美国第一国民银行共赔偿三家中国企业一百八十五万五千美元。
ID11188	然而，台湾要防御中共的飞弹威胁，目前看来大概可以有两套做法，除了建构飞弹攻击系统外，另一种就是加入美国的区域飞弹防御系统（TMD）。
ID1109	这几年，全省共实施火炬计划项目七百八十四项，经省认定的高新技术企业达五百六十九家，在国家科委公布的第一、二批国家重点高新技术企业中，广东占了百分之二十。
ID11072	经过 921 大地震，日月潭受创严重，仅余两百八十二人的邵族更是雪上加霜、百废待举。
ID11038	叶宏灯说，目前东莞约有四百名台商子弟在大陆念书，百分之四十的孩子念大陆一般中小学，经济情况较好的台商则通常送孩子到香港子弟学校或外商子弟学校。
ID10984	去年八月，中共海关大力查缉走私，东莞总共有十二名台商被扣，这些台商因为报关手续不合规定，延伸出盗卖合同、产品内销未补税、空转、盗卖海关备案合同等，全被以走私名目扣押，现在还有七名未出狱，「一想起他们，我每天打坐、吃素，心里总不能平静，」叶宏灯说。
ID10971	谈起推车，「我们一开始想租香港机场的柜台，但香港机场六平方米的柜台，只租四分之一，月租就要二十多万港币，」张福台说。
ID10923	高文来指出，东莞位於广州与深圳中间，离香港三个小时车程，从虎门坐船则只要一个半小时，整个珠江三角洲都是它的腹地。
ID10918	如今台商已是东莞外资的主力部队，以当红的电子业来说，根据统计，东莞总共有二千八百家电讯资讯业，台港澳资占八成；整体资讯产值占五百亿人民币，台港澳资产值占五成。
ID10916	也不过就是十馀年前，东莞这个古籍里杨贵妃伫望的荔枝产地，还是个遍地稻黍的农村，经过十馀年的发展，现在已成了一处「没有围墙的加工区」，面积有二千五百二十平方公里，人口超过五百万的东莞，总共有三千多家台商。
ID10863	立法院第四届第二会期已於一月十五日结束，经过朝野立委长达一周的挑灯审议，共通过一百一十三件审议案，创下民国三十七年以来立法院审议案件最多的纪录。
ID10800	林文月回忆，小时候家里人口多，过一个年，至少要用大蒸笼蒸出两、三笼的萝卜糕才够全家上上下下享用。
ID10780	现代化浪潮之下，一年中能够感受到的「中国味道」，大概也只剩下一个春节了。
ID10731	「但因为编、导、演各有不同的政治立场，在开场的十几集做完时代背景的交代後，剧情就趋向传统以亲情、爱情为主轴的连续剧路线，」林玲玲笑著说，这是始料未及的，而因为受欢迎，类似的题材，徐进良共制作了三部。
ID10707	以台语为主要语言的乡土剧，共占总收视率将近四分之三，其中民视的《富贵在天》更是独占鳌头。
ID10650	他正式发表的第一篇小说〈城仔落车〉，写的就是一位带著外孙进城里去找女儿、新女婿的阿婆，她虽然「才五十岁」，然而「岁月和生活在她枯乾脸上，留下了很深的痕迹」，使得她「给人的印象大约有六十开外的光景」。
ID10526	去年中好不容易有了起色，却又遭逢九二一大地震及隨後的大停电，至少「震」掉了近百亿元台币的纺织品订单。
ID10413	「香港从一九七〇年代中期就开始培育本地人才，九九年澳门回归则迟至八〇年代中期才展开培训本土人才，准备的工作整整晚香港十年，」梁金泉说。
ID10392	一位留台大校友更明白表示，在台湾念书就只有四年，很难说台湾对他有什麼影响。
ID10381	尽管澳门政府行政暨公职司表示，现在澳门政府中仅两百人有台湾学历；但根据我驻外单位估计，目前澳门政府至少有一千名留台校友。
ID10380	谭惠珠仅是澳门政府留台大校友的例子之一。

ID10377	澳门与台湾，飞机的航程仅有一个半小时，是澳门大学教授黄汉强口中「早上起来坐第一班飞机到澳门後在威斯京酒店喝咖啡谈公务，下午到澳门离岛打一场高尔夫球，逛逛本岛闹区，晚上搭最後一班飞机离去，一天之内就可往返」的交流便利之地，但数十年来，两地之间的关系忽隐若现，在一九九九年十二月澳门主权移交之後，台澳关系会跨进怎样的进程？
ID10327	据统计，今年夏天发生在长江和嫩江流域的特大洪灾共冲毁房屋七百三十多万间，二点三亿多人受灾。
ID10248	刚刚结束徒步穿越世界第一大峡谷——雅鲁藏布大峡谷的中国科学家今天在此间透露，大峡谷比人们原先认为的更狭窄，谷底最窄处的宽度只有三十五米。
ID10244	自“闰秒”系统1972年实施以来，“协调世界时”已总共增加了21个“闰秒”。
ID10142	但中国选手获得了6枚银牌，其中短道速滑选手成绩突出，共得5枚银牌。
ID10110	改革开放以来，全国各类成人高校共培养毕业生一千万人。
ID10075	不过，在利物浦队队员中，参加下流活动的始终只有加拉切尔一人。
ID10004	美英在12月17日至20日对伊拉克的轮番轰炸中，动用了包括两艘航空母舰在内的约20艘战舰、包括B-52战略轰炸机、B-1战斗机、F-18“大黄蜂”战斗攻击机在内的约70架各种作战飞机，共向伊拉克发射了近500枚巡航导弹，超过了海湾战争期间多国部队使用的巡航导弹总和。
ID775	至今年十一月初，共有入区项目一百七十七个，投资总额六十四亿元，其中外资约占百分之五十，入区企业中有七家被认定为国家级高新技术企业。
ID8843	同样未加入欧元区的瑞典的通胀率为十五国最低，仅为百分之零点一。
ID9440	国家环境保护总局污染控制司副司长臧玉祥说，“今年底工业污染源的达标排放大约能消减太湖磷负荷量百分之十二，农业污染如能控制，则能再消减百分之十五。”
ID11608	虽然，石冈仔工作站的计画目前只有一年，但李杰颖说：「一年之後是不是离得开，我真的有点怀疑，」他表示，北埔和石冈之间的关注和协助应会一直持续下去，也希望大隘的社区经验能在每个乡镇开花结果，让台湾重现田园温馨。
ID13322	好不容易存了钱，也只够买上一管油画颜料，在存足其他颜色之前，只好不时打开来闻一闻那亚麻油的味道过过瘾。
ID999	据介绍，在这三十多万家中外合资企业中，中外合资经营企业十八万二千零五十九家，占六成；中外合作经营企业四万四千零九十四家，约占百分之十五；外商独资企业七万六千一百六十一家，约占百分之二十五。
ID6562	一九七八年以前，约有九成以上的城市居民就业靠国家安排，现在仅有不足四成。
ID13601	副总经理苏宏文表示，人事广告市场没有确切数字，因为以往独占分类广告市场的报纸不会透露实际营收；但根据报纸版位大小、售价试算的结果，可达一百亿台币，保守估计约有三、四十亿元，其中网路商机约占十亿元。
ID11439	台湾科技大学去年有五个研究所招收在职生，其中纤维所计画招收十名学生，结果只有二十人来报名，等於二人取一人；机械所计画录取十人，也只有三十人报考。
ID10789	在卜家两老的心目中，中国传统的农历新年，份量就是不一样，除了趁这个时候可以制作一些家乡口味的食品馈赠亲友，联络情感；他们也非常期盼全家团聚的热闹气氛，除夕的一顿年夜饭，就是最吃重的节目，全家三代共有十多口人，围炉时，「最少要十道年菜，」卜妈妈数说，鸡、鸭、鱼、肉之外，不可少的还有先生的安徽老家传统菜肴「炸圆子」（用绞肉混和切碎的葱、姜、蒜等材料及七分熟的面线一起搅拌後，搓成丸子，再在外层沾裹上太白粉，油炸而成），其他像珍珠丸、红烧猪脚、什锦菜等，平日两老就爱煮煮弄弄给家人吃，这时更要拿出全副看家本领，从除夕前三、四天就开始准备这顿大餐。
ID7253	这些案件包括：天津“二．一三”贩卖非法图书案、河北正定“四．一五”政治性非法出版物案、北京贩卖政治性非法图书案、湖北贩卖政治性非法图书案，共

	涉及非法出版物及图书三万多册，目前案件涉及的非法出版物已经全部收缴。
ID251	到目前为止，该省共有三十四个市、县被列为沿海经济开放市县，占全省总面积的百分之四十四，其中国家级的各种开发区七个，还有一批省级经济开发区和省级农业开发区。
ID4692	据介绍，小浪底大坝工程开工不到一年的时间里，业主就收到了一百四十七封索赔信，在C I P M的协助下，最终认定只有四十七封信函的索赔理由是站得住脚的。
ID4341	截至目前，中国海上石油工业累计与外国公司签订了一百一十多个石油合同，外商共投入资金约三十亿美元，占全国海上石油勘探开发总投资的百分之五十八。
ID4286	五年间共新建和扩建沿江港口一百多个，增加吨位三千万吨，其中对外开放口岸十六个，外轮从长江口进入长江，直驶可达洞庭湖畔的城陵矶港，开放水域达一千三百五十七公里。
ID3337	广州、深圳等十四个城市到北京参加全国首届人才技术交流会，共设招聘点一百个；一九九三年初，珠海、佛山等六市还组团到美国、加拿大招聘出国留学人员，有一千零六名留学人员填写了申请表，现有不少人已陆续回国就职。
ID1816	吴晓灵还披露，到目前为止，中国境内机构在国际资本市场上累计发行外币债务折合美元约为一百七十亿美元，有力地支持了国家经济快速、持续、健康发展。
ID13268	根据陶俑身上留下的刻文名字考证，当时约有八、九十位陶工，各自带著十几位助手投入陶俑的制作，总计应有近千名来自官方和民间的陶工合力完成这些需要塑、堆、捏、贴、刻、画的陶俑。
ID7976	到目前为止，该指数已比今年初上涨了百分之十七点九，比1 1月2 3日创造的最高记录9 3 7 4 . 2 7点也只有5 3点之遥。
ID10996	「我们在这儿赚钱，但也在台湾缴税，家庭或子女的日常消费也多在台湾，」张梅良会长也指出，台湾的九二一大地震，全大陆台商总共捐出六千万台币，东莞台商就捐了三千万，讲台商背弃台湾很不公平。
ID10137	1、1 2月在泰国曼谷举行的第十三届亚洲运动会上，中国选手参加3 6个项目中的3 2项比赛，共夺得1 2 9枚金牌，连续第五次获得亚运会金牌总数第一。
ID9929	至今年三月底，武汉市共吸引世界上五十多个国家和地区的投资，累计审批利用外资项目近五千个，协议外资额七十一亿美元，实际利用外资四十八亿美元；全市兴办“三资”企业三千七百四十家，协议外资额四十四亿美元，实际进资二十二亿美元，进资率达百分之五十，高于全国平均水平十七个百分点。
ID9756	根据世贸组织和贸发会议最近公布的一项调查报告，联合国确定的4 8个最不发达国家的总人口为5 . 5亿，约占全球人口的百分之十，但其贸易额仅占全球贸易额的百分之零点五，其人均年收入不足6 0 0美元。
ID7449	目前，该指数比年初升高百分之十六点七，距1 1月2 3日所创的历史最高点9 3 7 4 . 2 7点只差1 4 7点。
ID10005	据美英宣布，这次空袭行动共击中了1 0 0多个伊拉克军事目标，达到了预定目标。
ID9789	据福建省政府口岸办公室统计，至九月底，该省海港口岸共完成外贸货运量一千五百一十万吨，已提前一个季度完成省定全年一千五百万吨的生产计划。
ID9477	一至十一月，两岸交流项目共计一千九百多个，参加交流的共一万余人次，均比去年同期有大幅增长；两岸间接贸易额达一百八十二点九三亿美元，比去年同期增长百分之五点九；共批准台商投资项目二千六百三十个，协议利用台资金额二十四点五七亿美元，实际利用台资金额二十七点三三亿美元。
ID474	一是从一九九四年十一月十八日破土动工到锥炉点火成功，仅用了十三个月，工期比安玻一期缩短了十八个月，比世界建设最快的同类项目还提前五个月；二是实现了国内外人才、技术、设备、设计和施工等资源的最佳组合，无论硬件配置

	和软件应用都达到了当今国际同行业先进水平；三是与成套引进国外技术和设备相比，投资大大节省。
ID5368	他说：「截至目前为止，土力工程处共调派了五十名土力工程人员前往处理山泥倾泻事件。不论塌下的斜坡业权谁属，他们均向紧急救援队伍即场提供土力工程的专业意见和指导。他们就所需的紧急行动及修葺工程提供意见，以消除斜坡对市民所构成的即时危险，而拥有该等斜坡的业主，私人或政府部门，则需要负责进行上述斜坡维修工程。」
ID10062	截止一九九七年底，中国共建立各级各类自然保护区九百二十六个，面积达七千六百九十七万公顷，一大批极其重要的生态区域和珍稀濒危物种得到有效保护。
ID11444	目前国内提供高阶主管进修课程的公私立大学共有台大、交大、中山、大叶、元智等十馀所，为什么企管所特别受社会人士青睐？
ID6795	六十五岁的香港居民卢光辉五年前回广东探亲时，从深圳坐汽车到广州花费了四个小时；今年，他从深圳回到广州仅用了九十分钟。
ID5347	邝其志说，香港除储备庞大外，政府并无负债，累积储备约为四百八十亿美元。
ID8049	海外有舆论评论说：“这场空前的金融浩劫至少提前二十年把中国推上了影响全球金融的大国地位。”
ID1362	据浦东新区经贸局对浦东开发七年来引进的投资一千万美元以上的一百五十七个工业大项目跟踪调查，目前建成投产的有一百一十六个，投产率高达百分之七十三点九；已建成投产项目的建设时间，最长三十个月，短的只有十一个月，平均建设周期为十七点二个月。
ID4064	由于备战广岛亚运会，我国大部分田径好手没有报名参加此次比赛，国家队只有21名选手在短跑、接力、跳高、跳远、铅球等项目中露面，辽宁省女子中长跑队也没有前来。
ID5615	根据行政院公平交易委员会的统计，截至八十四年底，全国登记有案的多层次传销公司，共计有五百四十八家，然而因为竞争激烈，实际有活动的，大约只有二百一十家。
ID8763	在芬兰梅里塔银行、莱奥尼奥银行和合作股份中央银行等其他大的银行系统，至少有600名业务和计算机技术人员值班，对欧元的启动进行监视。
ID7572	据此间报纸报道，苏哈托家族在国内3200家公司中持有股份，至少拥有200亿美元的财产。
ID7440	协议还规定，双方每年总结执行情况和存在的问题，确保两国至少互相进口百分之三十五的有附加值的商品，以防止第三国产品在两国市场上泛滥。
ID11892	他认为，帛琉主要岛屿岩石多，并不适合农业发展，加上经济尚未完全开发，除了渔产，主要物资都得仰赖进口，所以物资条件并不优渥；另外，该国总人口只有一万八千多人，若要在该地设立工厂，将面临劳力短缺的问题，而事实上目前帛琉的劳力亦仰赖进口，外劳有六千多人。

(298 句， 307 个用例)

附录五 框式介词类

ID9354	刘方仁说：“贵州农村生产力已由主要解决‘吃饭’问题转向主要解决‘用钱’问题，实现了农业和农村经济的历史性跨越。”
ID9353	农村贫困人口由九十年代初期占农村总人口的百分之五十下降到目前的百分之十左右。
ID9158	随着两岸关系从紧张走向缓和，两岸文化交流蓬勃地发展起来。
ID8909	旅游大省陕西加快旅游产业发展，计划将旅游业总收入占全省生产总值从一九九

	七年的百分之六点三，提高到二〇一〇年的百分之十。
ID8888	用药报销范围的制定，使药品支出占总支出的比例由一九九六年的百分之七十六降到目前的百分之五十八，用药更趋合理。
ID8805	在国际社会的调解下，塞政府曾把导弹的交货期从今年7月推迟至11月，后来又推迟至年底。
ID872	资料表明，金川公司可回收的有价元素已由原来的七种增加到十四种，镍的选、冶回收率净增百分之二十六点五，贵金属的冶炼回收率提高十多倍。
ID8605	国务院组成部门由原有的四十个减少为二十九个。
ID8602	中共中央金融工作委员会成立；金融机构全面实行垂直领导；国家取消了对国有商业银行实行多年的贷款限额控制；人民银行对存款准备金制度实行改革，法定存款准备金率由过去的百分之十三下调到百分之八。
ID8496	另据欧盟统计，目前全球电子商务市场正在飞速发展，估计到2000年时的市场营业额将达到2000亿欧元，而上因特网的个人将从1996年底的8600万人增加到2.5亿人。
ID8469	三洋电机公司也将录像机的辅助设备能耗由4.5瓦降低到2.5瓦。
ID8468	索尼公司今年出售的所有音响设备的辅助设备的能耗已由原来的10瓦—15瓦降低到三瓦以下。
ID797	经济专家分析说，这表明不少跨国公司进一步看好中国乃至上海的政治经济形势，并开始全面调整对华经营战略，由早期分散的、缺乏系统性的单个投资逐步转向较大规模的、系统化的投资，由建立生产基地转向进行上、中、下游各阶段产品的综合开发，同时投资范围也从工业制造扩大到银行、保险、咨询、贸易等众多领域。
ID756	机电产品占出口总值由五年前的百分之十七点九提高到百分之三十一一点五。
ID576	目前，四川省对外商外资开放的领域不断扩大，现已由过去的一般加工业，扩大到基础工业、基础设施和一些贸易行业、外商直接投资、国际证券融资等等。
ID537	不断壮大的企业规模，已从流通领域拓展到生产领域、高新技术、房地产、建筑及投资建设大型基础设施等多行业多领域。
ID468	据介绍，安玻二期工程的建成投产，使安玻公司的彩色玻壳年生产能力由一期的六百万套扩大到一千一百万套，成为中国目前生产规模最大的彩色玻壳生产基地。
ID40	一九九五年九月建设部和外经贸部联合发布的《关于设立外商投资建筑业企业的若干规定》，使中国的建筑市场从允许境外企业到中国承包工程进入到允许境外企业到中国办合资建筑企业。
ID305	外商投资企业一至二月份进出口总值为一百六十五点九亿美元，占全国进出口总值的比重由去年同期的百分之三十七上升到百分之四十五点九。
ID2280	这份报告预测，到2010年，中国经济在世界经济中所占的比率将由目前的百分之二点五上升到百分之五或百分之六；中国出口在世界总出口中所占的比率也将由目前的百分之三扩大到百分之十。
ID2157	按原产地原则统计，产品的出口国就从这些国家和地区转移到了中国。
ID2143	与此同时，该公司营业额从过去的五千万美元增长到去年的二十亿美元，去年利润超过三点五亿美元，成为投资中国最成功的外商投资企业。
ID2051	“深业控股”上市保荐人为满足投资者需要，全面行使了超额配股权，公司募集资金由原来的四亿余港元增至五点三二亿港元，深圳高速公路股份有限公司的六点五亿股H股紧接其后推出，日前也大功告成，集资额超过四亿港元。
ID1968	三、一九九八年将纺织品出口退税率提高二个百分点，由目前的百分之九增至百分之十一，同时将输往欧美纺织品协议配额总量的百分之十五以上直接“戴帽”分配给纺织自营出口企业。
ID1911	特别值得指出的是，从去年7月份首次出现月度逆差后，去年下半年月月逆差，

	而且逆差额几乎成倍增长，由7月份的2000万美元猛增到12月份的7.22亿美元，使1997年全年的累积顺差额降到微不足道的数目。
ID1832	辽宁利用外资的领域已从过去的中小型加工项目，扩大到石化、冶金、机械、电子等老工业基地改造的四大支柱产业和农业、房地产、基础设施、能源、金融等行业。
ID1830	高国珠谈到，一九九七年辽宁外资企业的税收达到三十亿元，占全省工商税收的比重，已由一九九二年的百分之三上升到百分之十以上。
ID1828	一九九七年外资投资企业的年产值突破六百亿元，外资企业年产值占全省工业产值的比重，已由一九九二年的百分之五上升到百分之十七。
ID1696	报告说，该机构已将今年全球原油需求量的增长率从早先估计的百分之二点五降至百分之二点二，比最早预测的日均需求量减少160万桶。
ID1662	他预测今年世界经济实际增长率将从九七年的百分之三点九降至百分之三。
ID1550	改革开放以来，中国对外贸易迅速发展，进出口额由一九七八年的二百零六亿美元上升到一九九七年的三千二百五十亿美元；一九八三年以来进出口贸易一直持续增长，势头至今不衰。
ID1523	在告别了“短缺经济”年代后，商品供求关系已由“买方市场”转向整体性的“卖方市场”，竞争加剧。
ID15	去年实现进出口总值达一千零九十八点二亿美元，占全国进出口总值的比重由上年的百分之三十七提高到百分之三十九。
ID1488	芬兰对这一地区的出口占其全部出口的比例也从10年前的百分之一增至百分之六点二。
ID1443	优化资本结构试点城市由五十八个扩大到一百一十一个，国家集中抓的重点企业由三百户扩大到五百一十二户，企业集团试点由五十七户扩大到一百二十户。
ID14195	等到下回黄天才来取印的时候，壮老才告诉他那章子上刻著「天游化人」四个字，是康有为晚年的图章，让黄天才著实又恼又悔了许久。
ID13741	去年第三季开始，全球CD-R价格呈现崩盘走势，每片由年初的0.6美元跌至0.4美元，几乎逼近0.3美元的成本价，离五年前初上市时的十四美元更是天差地远。
ID13739	同样的三千名员工，月产量则从二十万片暴增到二亿片。
ID1308	去年全年，初级产品进口二百八十六点二亿美元，增长百分之十二点五；初级产品出口二百三十九点三亿美元，增长百分之九点一，占外贸出口总值的比重由上年的百分之十四点五下降到百分之十三点一。
ID1260	洋浦经济开发区管理局局长王永春说，开发区招商工作取得突破性进展，开发区建设已由土地开发迈向工业项目建设的新阶段。
ID1233	甘肃一九九七年全省国内生产总值达七百八十一亿点三亿元，同比增长百分之八点三；零售物价涨幅则从上年的百分之六点六下降到百分之一点八，居民消费价格涨幅由百分之十九点八下降到百分之三。
ID1228	目前，外商在陕西的投资区域已由关中平原辐射到陕南山区和陕北黄土高原，“三资”企业遍布全省十个地市，生产经营涉及电子、医药、食品、房地产和汽车制造等十多个行业，其中工业企业占三分之二。
ID12224	就如同已有二十九年历史的美国那斯达克指数，从去年十月中旬起，突然由二千八百点一路飙涨到今年三月中的五千多点，电信股、网路股、生化类股等轮番大涨，某些个股的本益比甚至高达五百（股价是获利的五百倍），以致被讥为「本梦比」。
ID11768	两岸问题在与会者就各种面向的剖析下可说是淋漓尽致，座谈会最後将焦点从两岸转到台湾内部，风格多元的作家和文化评论家平路，对国民党中央党部前的民众抗议事件提出了发人深省的看法。

ID11664	放眼所见，时下男模特儿也由肌肉男、酷哥转型为带有阴柔气质的生嫩学生、邻家男孩，或者看来稚气、缺乏经验、扁扁瘦瘦、接近中性的少男。
ID11301	「政府该回归到裁判角色，好好执行，」他认为，和大陆的经贸关系也是如此，因为只要有管制就会有投机，某些有势力的人就可以用光明正大的方式实施五鬼搬运，而政府不该制造这种让特定人得利的机会，一切该回归市场，让市场运作。
ID11170	去年十二月，副总统连战在三军大学主办的《跨世纪国家安全与军事战略》研讨会中，以「新世纪的国防战略与建设」为题的演讲中宣示，台湾将发展「远程飞弹」以反制中共的军事威胁，战略构想也将从「防卫固守、有效吓阻」，改为「积极防御、有效吓阻」，新一代兵力的整建将以「攻势防御」为原则。
ID10930	王勇铎九〇年到马来西亚考察准备设厂，但他发现，马来西亚的人口比大陆少得多，工资是大陆的两倍，语言不通，生活习惯不同，交通道路条件也未必好，他在与大陆的投资条件比较後，选择到大陆设厂，他的工厂主要生产印刷电路板，十年来员工从二、三十名扩充到两千名，现在还准备再扩厂。
ID10907	东莞市最热闹的附城区，台商经营的酒店，亮眼的灯饰从门口挂到屋顶，这样的闪灯跟酒店里的卡拉 OK 一样，不分季节，每天喧闹到天亮。
ID10868	一、「地方民意代表费用支给及村里长事务补助费补助条例」，将大幅提高县（市）、乡镇（市）长地方民代的支给费用，缩小北、高两市及其他县市、乡镇市长民代及村里长费用支给之差距，总计此法通过後国库增加的相关预算，将由每年四十四亿元增加到七十亿元台币。
ID10757	「他们称边拍边播的戏为『ON 档戏』，往往今天刚拍完，隔两天就播，」数届金马与亚太影后的电影红星陆小芬，最近加入八点档《富贵在天》的演出，却有满腹的牢骚。
ID10640	「兴票案」真相如何，短时间内可能很难解释清楚，但该案对宋楚瑜的选情已造成重大影响，根据中国时报在十二月二十日所作民调显示，支持宋楚瑜的比例已由案前的三成多下挫至二成三，陈水扁则以二成四超前，连战以二成二紧追在後，形成只有些微差距的情形。
ID10108	同期平均每万人口在校大学生人数，由八点九人增加到二十五点七人。
ID10000	新华社巴格达 1 2 月日电综述：1 9 9 8 年：美伊从危机走向战争
ID8837	欧洲统计局还说，欧洲联盟 1 1 月份的通胀率从一个月前的百分之一一点一降至百分之一一点零。
ID8194	2 月份交货的原油价格也从每桶 1 1 . 6 7 美元升至 1 1 . 9 0 美元。
ID8054	香港回归一年来的高度自治使所有臆测香港政局不稳的“预言”统统落空；“一国两制”的巨大成功使人们对九九澳门回归更加充满信心；以“汪辜会晤”为标志，海峡两岸关系也从“山重水复”走向“柳暗花明”。
ID7701	一些专家估计，信息高速公路建成后，知识对经济增长的贡献率将可能由本世纪初的百分之五至百分之二十上升到百分之九十。
ID7604	兵役法将义务兵现役期限从三年或四年统一降至二年，同时强化了公民的兵役义务。
ID7454	圣诞后新年前夕的华尔街交投清淡，纽约证交所上升股略少于下跌股，成交量从前一个完整交易日的 6 . 9 4 亿股减少为 5 . 2 6 亿股。
ID7452	据一家咨询公司预测，美国网上交易额有可能从今年的大约 1 3 0 亿美元增加到明年的 3 0 0 亿至 4 0 0 亿美元，这是导致技术股当天走强的主要因素。
ID7334	石油输出国组织（欧佩克秘书处 2 8 日宣布，该组织七种市场监督原油一揽子平均价上周继续下滑，从前一周的每桶 9 . 7 8 美元滑落到上周的每桶 9 . 5 0 美元。
ID7323	来自广东省有关部门的数据显示，一九七九年至一九九七年，广东国内生产总值平均每年递增百分之十四点二，比全国高四个百分点；而广东的城市化率也由百分之十一点七上升到百分之三十六，比全国高六个百分点。

ID7321	据悉，二十年来，珠三角地区有超过七成的原有农村劳动力实现了向非农产业的转移，常住非农业人口已在一九七八年的四百八十一万人增加到去年的一千三百多万人。
ID6875	从一九七九年算起，深圳的国内生产总值由一点九六亿元人民币发展到一九九七年的一千一百三十点一亿元，平均每年增长百分之三十二点二五。
ID6738	据捷克国家统计局日前公布的统计数字，今年头 1 1 个月，捷克进口额为 8 5 0 6 亿克朗（1 美元约合 3 0 克朗），出口额为 7 8 8 0 亿克朗，分别比去年同期增长百分之八点九和百分之十九点六，外贸逆差由 1 2 2 1 亿克朗降至 6 2 6 亿克朗，比去年同期减少了百分之四十八点七。
ID6636	在种种主张碰壁之余，中国人才找到了马克思列宁主义，才认定只有社会主义才能救中国。
ID6468	国际大公司的进一步加盟使上海港国际航线年内又开辟了十六条，每月的班轮数从去年四百九十二个增加到今年的六百三十个。
ID6442	七年后的今天，大寨已有制衣公司、水泥公司、酒业公司等七家企业，固定资产达五千四百多万元，村民的纯收入也由十年前的七百元升到二千七百多元。
ID6123	据江苏省计划与经济委员会介绍，江苏省计划到本世纪末时实现引进外资三百亿美元，争取达到四百亿美元，使全省外资实际投入占当年全社会固定资产投资总额的比例从一九九五年的四分之一提高到本世纪末的三分之一。
ID6035	首期工程竣工后，天津港的泊位将由原来的三泊位增加到六个，吞吐能力新增一百二十万吨，并进一步改善码头的作业环境，提高港口的综合通过能力。
ID5868	到一九九六年，双边贸易额已在一九七二年恢复邦交时的十亿美元增加到六百亿美元。
ID5709	目前值得庆幸的是，台湾内销市场并没有出现景气衰退现象，「光靠国内市场就可以养得活台塑了」，这对以内销为主的台塑来说，无疑吃了一颗定心丸。
ID5340	「由於推行多项措施的成果，轮候出租公屋的平均时间已逐渐由九年缩短至现时的六年半。我们亦已承诺会在未来的日子再进一步缩短这时间。」
ID4839	全省空港由最初的两个增至七个，空中航线发展到一百多条，形成了以省会杭州为中心，连接宁波、温州、黄岩、义乌、衢州和舟山，辐射海内外的航空运输网络。
ID4729	有中国“荷藕之乡”美誉的江苏省宝应县荷藕产品丰富多彩，由最初只生产盐渍藕发展到保鲜藕、速冻藕、藕汁饮料等五大系列五十多个品种。
ID4665	日本由前年的六十六家上升为七十一家，而美国由五十一家增至五十二家。
ID4656	五百家企业中，销售额超过十亿元人民币的企业由一九九五年的八十九家增加到去年的一百二十四家，五百家企业的平均销售额达到九点四亿元，比前年增加了二点零五亿元。
ID4623	许多企业经营范围已由液体化工和散装化工产品的仓储、灌包，扩大到散装水泥、集装箱、车客渡、散装化肥、沥青等各货种的装卸、储运、中转。
ID4505	据介绍，南非目前是中国在非洲最大的贸易伙伴，近年来中国与南非贸易发展较快，贸易额已在一九九一年的一千四百六十五万美元发展到一九九六年的十三点五亿美元，今年前九个月，双边贸易额达十一亿美元，比去年同期增加百分之十五。
ID4457	几年间，中国应用电子技术先后改造旧机床五万多台、关键生产设备和生产线五千条（台）；每万元国内生产总值能耗已由一九九 0 年的五点三吨标准煤下降到一九九五年的三点九四吨标准煤；机械行业仅推广应用楔横轧一项技术，创造直接经济效益四亿多元。
ID4208	他说，经过四年多的宏观调控，中国的通胀率已在一九九四年的百分之二十一点七降到一九九七年上半年的百分之一一点八，国民经济实现了“软着陆”，以控制通

	货膨胀为首要任务的宏观调控政策取得了巨大的成功。
ID3221	今日统一的德国正在从经济大国走向政治大国。
ID849	其中美国明年的经济增长率将从今年的百分之三点八下降到百分之二点四，日本明年的经济增长率只有百分之一一点一。
ID5596	在一九九一年至一九九六年间，男性的出生时平均预期寿命由七十五岁增加至七十六岁，而女性的出生时平均预期寿命则从八十一岁上升至八十二岁。
ID8835	欧元诞生前 3 6 小时，欧元区得到来自欧洲统计局的好消息：欧元区通货膨胀率再次稳步下降，从 1 0 月份的百分之一一点零降至 1 1 月份的百分之零点九。
ID8461	松下电器公司最近加强了节能技术开发，计划在 2 0 0 0 年将其电器产品的平均能耗标准由比 1 9 9 5 年降低百分之三十修改为降低百分之五十。
ID11091	民国五十六年，中研院人类学教授卫惠林请邓相扬带他到部落访问，他随著卫惠林挨家挨户找老人家访谈，因而学到不少田野调查技巧，之後，邓相扬只要一有空就往部落跑，找老人家闲聊。
ID9857	浙江省国内生产总值从一九七八年的二百三十二点九亿元上升到今年的五千余亿元，经济总量已跃升到中国第四位。
ID11016	卸下东莞台协会长职务後，他便接了东莞台商子弟学校筹备主任职务，从「一个小火坑，跳到一个大火坑，」他表示。
ID12204	当时国内蓄积多年的经济力开始爆发，连年钜幅的外贸出超，导致台币升值、热钱涌入，股市从七十四年的六百多点飙升到七十九年的一万二千多点，台北市的平均房价也从七十五年的每坪七．二万元飙到七十八年的二十八．五万，涨幅超过四倍。
ID12211	六年前张太太决定为罹患绝症的弟弟一家买栋房子，当时淡水新市镇的大楼房价从每坪二十二万元的高点滑落至十五万元，张太太一面庆幸自己买得便宜，一面坚信景气好转後，房市自然会涨回去，於是高高兴兴付了头期款，看著弟弟一家搬进新房。
ID10848	她认为，如果现代人对「年」的定义，从吃喝玩乐转移到内省与精神上的意义，能感悟到时光的消逝与工作告一段落；从而制订新希望与下一阶段的工作与成长目标，就不会无端而兴「一月今年始，一年前事空，凄凄百年事，应与一年同」之类韶华易逝的伤感了。
ID8196	美国权威的民间研究机构会议委员会 2 9 日发表报告说，美国 1 2 月份消费者信心指数从 1 1 月份的 1 2 6．4 点微降到 1 2 6．1 点，降幅仅为 0．3 点。
ID7112	男子跳远项目就是一个典型的例子，这个项目的参赛标准规定，只有在上个赛季室内成绩达到 8 米 1 5 和室外成绩达到 8 米 2 5 的选手才有参赛资格，而上个赛季全世界只有 3 名选手跳出了 8 米 1 5 的室内成绩，而达到室外跳远标准的选手也只有 1 3 人，这把像在欧洲田径锦标赛上获得第五名的瑞典选手马·桑勒伯恩这样的一批优秀运动员关在世界锦标赛的大门之外，这显然是不公正的。
ID6795	六十五岁的香港居民卢光辉五年前回广东探亲时，从深圳坐汽车到广州花费了四个小时；今年，他从深圳回到广州仅用了九十分钟。
ID5701	在石化业方面，以台塑的主力产品 PVC 来说，「目前远东区的 PVC 价格，已从七百美元一吨降至六百，比国际行情低了一截，」台塑企业总经理李志村指出，这虽是 PVC 原料——乙烯——去年持续跌价的结果，但韩、泰等国的低价倾销策略也是推波助澜的因素之一。
ID7289	如今，来祖国大陆求学的台湾学生已从福建扩展到北京、上海以及内地的各大院校，所学的科目也涉及经济、法律、物理、计算机等各门类和学科。

(98 句, 100 个用例)

附录六 名词与动词分离类

ID12206	许多人把房屋拿去抵押炒股票、再把股票拿去融资加码，埋下了日後暴跌时倾家荡产、债务缠身的恶因。
ID11890	许多台商将中帛建交视为到帛琉投资的利多，今年三月外交部的访问团中，就有包括观光业、纺织业、电子业、渔业等民间企业，随团前往考察当地投资环境，而当地台商也十分欣喜大使馆的设立，认为日後台商投资将更有保障。
ID9609	根据这份文件，韩国政府将先把第一银行的不良资产分离出来，然后把百分之五十一的股份卖给这一投资财团。
ID12338	相对的，处理型晶片则将实验室中的前处理步骤放到晶片上来做，使用者只要直接将血液滴到处理型晶片上，就可以得知结果。
ID9603	浙江大学还确定了今后的发展目标：在千家万户的计算机上普及安装具有一定接收系统的机顶盒，把远程教育送到城市、农村、单位、家庭和个人的计算机上，在普遍推行普通高等教育的基础上，推进中国后继教育（即大学后教育）和终生教育的发展。
ID7216	交流团的同学们纷纷表示，回到香港后，要尽快把江主席的勉励转告给周围的同学、老师和家人，让大家共享这一份喜悦，共同为香港的长期繁荣稳定做出贡献。
ID8982	在中美关系发展的关键时刻，江泽民主席提出“增进了解、扩大共识、发展合作、共创未来”的主张，呼吁中美双方从长远观点和战略高度审视和处理双边关系，把握两国人民的根本利益，着眼于把一个繁荣、稳定、和平的世界带入 2 1 世纪，从而为冷战后中美两国关系的发展指明了方向。
ID5827	去年初以来，台湾最具实力的汽车公司——裕隆集团中华汽车公司多次组团访问，就合作发展汽车工业进行了积极的磋商；福建方面则成立了由省委书记贾庆林、省长陈明义任正、副组长的省汽车工业领导小组，将发展本省汽车工业摆上重要议事日程。
ID8807	现在，塞政府终于做出让步，宣布了一个折中办法，即同意不把导弹部署在塞岛，但也不取消合同，而是准备把它们部署在希腊的克里特岛。
ID12739	不过，「台北快速起飞，把其他县市远远抛在後面，却是这十多年的事，」台大建筑与城乡所教授夏铸九分析。
ID12373	十三日，陈水扁总统在登机前发表简短谈话时表示，台湾要「站起来、走出去」，此次出国访问，即是希望将中华民国各项成就与世人分享，同时，也希望此行能对巩固我国邦谊有所帮助，把国际友谊带回来。
ID12225	「投资人不计较公司现在的亏损，却把希望寄托在遥远的未来，可是预期的事越是久远，失准率越高，」台大财务金融研究所教授邱显比指出。
ID4792	年初新构建的上海大旅游体制，正在将这个大都市的旅游、文体、商业、交通、现代工农业等融合成一个新兴支柱产业。
ID7144	《决定》总结了二十年农村改革的基本经验；确定了从现在起到二 0 一 0 年中国农村经济、政治、文化建设的目标和任务；阐述了社会主义初级阶段农村尤其不发达的主要特征；对中国农村基本经营制度作了更科学、更准确的新概括；提出坚持以公有制为主体、多种所有制经济共同发展的基本经济制度；提出坚持以市场为取向的改革，进一步明确了深化农村改革的方向和目标；把水利建设和科教兴农提到了突出位置；指出增加农民收入，减轻农民负担是农村实现小康的基本条件；系统地提出了农村基层民主建设的具体任务和指导原则；把党对农村工作的领导提到了一个新的高度。
ID14730	时下的台湾森林游乐区，全是过去的桧木林场，从阿里山铁路、到四大山脉里的绵长林道，目的都在将一段段园形巨木扁柏、红桧运送下山，「可以说与台湾林业

	相关的事情都是围绕在桧木身上发展出来的,」研究森林生态的静宜大学副教授杨国祯说,台湾林业史,正是一页桧木砍伐史。
ID13050	目前康定的实验室和生产线严禁开放参观,未来更不排除把实验室搬到国外去,以避免国内剽窃偷窥的眼光。
ID9950	从一九八四年至今,在青海太阳湖地区许多挖金的农民把拖拉机、卡车开进了无人区,一两万人的生活垃圾污染了环境,珍稀野生动物成了猎食对象,植被被大车压成了路面,淘金所挖的十几米的洞到处都是。
ID992	洽谈会组委会常务副主任、陕西省副省长贾治邦说,本次洽谈会将本着“互利、合作、发展,把东西部合作全面推向二十一世纪”的原则,以突出东西部地区的横向联合与合作为主要内容,精选一批符合国家产业政策的合作项目,面向国内外市场招商。
ID9892	拥有万亿美元可投资金、每年都要选择一个新兴市场作为投资机会的美国罗素二〇一一二〇协会,如今把目光转向中国的中西部。
ID9714	在扩大对外招商的同时,南通市还积极实施把当地“拳头”产品打出去的战略,提高了这些产品的国际市场竞争力。
ID9628	一九九五年开始,交通部把湘江大源渡航电枢纽工程列为“以电养航”的试点工程,工程总投资十八点九五亿元,其中世界银行贷款九千万美元。
ID9497	中国石油天然气集团公司在进行机构调整和企业重组的同时,把提高效益放在突出位置。
ID9343	湖北等受灾省市和贫困地区的商业部门,把节日市场供应的重点放在肉、菜、糖等“菜篮子”商品供应上,确保灾区人民买到称心的商品,过上欢乐的节日。
ID9320	“我已经打电话给斯特恩,要求他把我们的最后方案递交给资方劳工委员会,力争在1月7日他们定的最后期限内,能使双方达成协议,挽救赛季,”强硬的亨特今天口气相当缓和。
ID9117	今天出版的《人民日报》发表评论员文章指出,明年在深化改革、促进发展的过程中,必须把确保社会政治稳定放在突出位置。
ID9109	下雪后,北大湖滑雪场又调集人力将雪拍实,防止被风刮走,最近10来天又采取了把雪道附近的雪移到雪道上的备雪办法。
ID9012	中美双方需要珍惜两国关系发展来之不易的成果,进一步共同努力,为把一个繁荣、和平、稳定的世界带入21世纪作出新贡献。
ID8948	卢森博格非常欣赏本年度巴西最佳球员加马拉,认为这位巴拉圭籍后卫能够把球准确地从后场传给前锋是“十分不易”的。
ID8939	医院也将根据急救卡上的有关内容及时通知当地区县保险公司前往急救医院提供保险服务;保险公司对保险责任范围的医疗费用在保险金限额内将予以支持,并将在审核医院有关文件后,在十个工作日内,把应付款项划入医院帐户,医院需开出收据以作证明。
ID875	他们从世界各国引进新技术和先进设备,大量采用本国最新科研成果,建成了长六千五百米的亚洲最长的矿山主斜坡隧道;探索出胶结充填采矿法,每人每班平均采矿十七点六吨,达到世界先进水平;用二点六亿元人民币引进关键技术和设备,通过消化吸收,建成中国第一、亚洲最大、世界第五座镍闪速熔炼炉,把中国镍冶炼技术带入世界先进行列。
ID8736	即将到来的一九九九年是孔子诞辰二千五百五十周年,中国政府将此年定为孔子旅游年。
ID8671	国务院批准同意将雅鲁藏布江大拐弯处命名为雅鲁藏布大峡谷。
ID8572	基于以上认识,会议提出,要抓住机遇,调整思路,把经济增长切实转到以质量、效益为中心的轨道上来,为农业和农村经济发展开拓新领域,为农民收入增长开辟新来源,使农业和农村经济继续保持旺盛的活力。

ID857	如果改革措施不得力，信心危机依然存在，那么投资者就有可能把注意力转向其他新兴市场。
ID8521	同时，争取开工建设煤码头二期工程，尽快建设十至三十万吨矿石专用码头，把日照建设成为中国重要的煤炭、矿石进出口基地；尽快开工建设东港区三期工程，大力发展杂货和各种专用泊位；加快五万吨级液体化工专用码头及配套设施建设。
ID8452	1997年12月在日本京都召开的《联合国气候变化框架公约》缔约国第三次会议规定，日本等发达国家必须在2000年将二氧化碳等影响气候变化的温室气体排放量降低到1990年以下水平。
ID796	经过一段时期的“投石问路”，越来越多的外商象美国百时美施贵宝公司一样，看好上海的投资环境和市场潜力，纷纷追加投资扩大再生产，并逐步把上海建成其在中国甚至亚太地区的经营、销售中心。
ID793	如今，这家跨国集团已陆续在上海追加投资二千多万美元，并将上海施贵宝制药有限公司发展成为其亚太地区的生产、出口中心。
ID76	随着崇明海关办事处的设立，崇明县内的单位足不出岛就可以办理一切海关手续，这对进一步改善崇明县的投资环境，加快吸引外资，方便快捷地办理海关手续，把崇明建设成对外高度开放的大型贸易港口，带动出口加工、航运中转等外向型经济的发展，将起到积极的作用。
ID7581	黄黄路的建成，极大地改善了鄂东的投资环境；同时与沪宁、合宁、合（肥）九（江）、武黄高速路连通，将长江经济带连成一体，这对发展中国东西向交通、开发长江经济带、服务上海浦东和三峡工程将起到重要作用。
ID7407	一九八三年，中国地方志指导小组成立，要求到本世纪末将地方志修到县级，并出版志书五千种。
ID7123	约十分钟后，货车火势控制不住，有数十人一齐上前准备将客车推离货车。
ID7122	当时两车均起火，救援人员用水将已变形的客车火浇灭，集中力量抢救伤员。
ID7109	瑞典国家田径队领队肯·奥尔松说，国际业余田联决定提高参赛标准，目的是要把每个项目参赛的人数限制在14到16人之间，这显然是“错误的”，将不仅使世界锦标赛混同于一个普通的田径大奖赛，而且还将使比赛变得“不公正起来”。
ID7044	这就是说：一、北约要将其执行使命的范围扩大到世界任何地区；二、北约要撇开联合国自行执行使命。
ID6914	当地人几乎都把这一灌溉系统视为经济命脉，合理使用水和合理分配水也就一直是他们自觉遵守的“铁的法则”。
ID6887	特写：把滥用水的人送上法庭
ID6885	特写：把滥用水的人送上法庭
ID6863	此外，俄罗斯还加强了与中亚有关国家联手防御阿富汗冲突对独联体南部边界安全造成的威胁，进一步推进塔吉克斯坦的和平进程，并将塔吸收为俄、白、哈、吉4国海关联盟的新成员。
ID6788	怎样才能变冬闲为冬忙，让蜜蜂能追赶上花季？张世元把目光瞄向了南国。
ID6681	虽然东南亚金融危机使一些欧美公司对投资亚洲国家心有疑虑，但西门子仍把亚洲、特别是中国看成是其投资对象。
ID6583	要继续坚持和完善行业价格自律工作，引导企业把竞争转到提高产品质量，改善售后服务上来。
ID6572	盛华仁说，要切实把脱困任务和责任分解落实到每个企业、行业及各级政府。
ID6421	他把和“小狗”的合影放大，挂在墙上。
ID607	旨在将中国、俄罗斯、朝鲜三国交界的图们江三角洲建成世界性物流枢纽和经贸中心的图们江开发计划，是由丁士晟等中国专家率先提出、联合国开发计划署一九九一年确定支持并向全球公布的。
ID5942	通过二十五街区这个南美最主要的物资集散地，把浙江的优秀产品推入国际经济

	循环体中。
ID5680	「韩国贬值、韩国厂商为了抢美金而『跳楼大拍卖』，把台商也拖下水，」台湾区人造纤维加工丝工业同业公会理事长、以「海鸥牌」行销全球的宜进实业董事长詹正田忿忿指陈：原先一公斤一块八美金的加工丝，目前滑落至一块三，跌幅达到三成，「以韩国、东南亚等国对美金的贬值幅度来看，这种价钱可以让他们的利润比以前还丰厚，我们就吃亏了。」
ID5672	不过，长久以来都以出口为导向的东亚国家，此番贬值，也意味著可以用更低廉的价格把货品卖到以美元计价的国际市场去。
ID5575	办妥手续後，第二被告即将两本伪造的英国国民（海外）护照、香港身分证明书及香港身分证交与该两名中国人士。
ID5569	第三被告，三十岁工人於一月间在中国将他的身分证明书以二千元卖给一名男子。
ID5551	为了将旧措施失效和新措施尚未实施的间断时间减至最少，政府將於明日（星期三）把联合国制裁条例草案提交临时立法会。
ID4981	此时此刻，世界各国人民都把目光投向香港。
ID4945	在开发和使用沥青石方面，委内瑞拉居世界第一位，它今年制定的生产规划，计划将乳化油的年产量提高到两千万吨。
ID4921	为了将移民工作推向深入，中国政府近年来制定了一系列优惠政策，鼓励外商投资三峡库区。
ID4855	三是建立国家和地方重点工程投标单位一览表制度，将这些企业列入国家和地方重点工程投标单位一览表，在同等条件下优先中标。
ID4776	据透露，这家公司将视市场情况，把投资额逐渐增加到八千万美元。
ID4770	湖北省已规划三峡地区投资项目二百三十多个，总投资额达二百八十亿元，并将把宜昌市建设成为一百万至一百二十万人口规模、产业比例协调的中国最大水电中心、中国重要的重化工基地和世界级的旅游风景区等多功能的大都市。
ID4675	据中国银行信贷管理部门负责人王丽丽介绍，改革开放以来，中国银行始终把外商投资企业视作重点服务对象，积极配合国家对外开放和利用外资政策，采取多种灵活的融资方式扶持外商投资企业的建立和发展。
ID4658	外经贸部外国投资管理司官员吕建华介绍说，这次评选首次将非生产性行业列入了评比范围。
ID4448	这些企业还把三点六万项拥有自主知识产权的技术向其它企业、机构转移和扩散，创造收入四十四点三亿元。
ID4430	他呼吁有关各国汲取教训，总结经验，完善金融体系，将危机化为机会，并欢迎其它地区工商巨子前来投资，参与东亚地区的发展与振兴。
ID4429	他说，将这场金融危机看成是东亚各国经济由巅峰滑向谷底的转折点是大错特错的。
ID3895	联赛之后立即将集训队员拉到昆明，进行为期一个月的训练。
ID3874	意大利著名音乐家克·巴廖尼演唱了由他自己作词谱曲、慷慨激昂的会歌《水中水》，罗马交响乐团40多名演奏家的表演，令听众如痴如醉，而音乐喷泉、高保真音响效果、喷水屏幕和立体投影更把大家带进了一个声、光、色交融奇妙无比的境界。
ID3789	据此间新闻媒介报道，俄罗斯西北军队集群司令部机构人员31日将军用设施移交给了当地地方政府，并向拉脱维亚总统乌尔马尼斯递交了有关完成撤军的报告，随后乘飞机从里加撤回莫斯科。
ID3316	阿卡耶夫说，两国的经贸合作关系势头很好，中国目前已成为吉尔吉斯的重要贸易伙伴，吉尔吉斯希望将两国的经贸关系提高到一个新的水平，欢迎中国企业到吉尔吉斯投资。
ID3312	他表示，吉尔吉斯将永远把中国看作是中亚国家的友好邻邦和相互信赖的合作伙伴。

	伴。
ID3220	当今唯一超级大国美国近来则开始把德国视为“领导伙伴”，鼓励它在世界政治中“发挥领导作用”，其原因是显而易见的：美国要借重德国的力量。
ID3213	在西边，美国组织北约，扶植西德，以“遏制”苏联；在东边，苏联组建华约集团，把东德建成对付西方的桥头堡。
ID3161	中国已经向全世界显示，解决社会困苦问题的战略在于作出理智的规划，把力量集中到有合理成功希望的任务上。……
ID3150	威迪奥诺表示将把此意见转达给联合国秘书长。
ID3149	柬埔寨两位首相要求联合国秘书长将其派驻柬埔寨的代表的留驻时间延长为 6 个月或更长一些。
ID2982	我将同叶利钦总统和其他俄罗斯领导人一起，共同探讨建立面向二十一世纪的睦邻友好、互利合作的中俄关系，把两国在各个领域的合作提高到新的水平。”
ID2856	双方一致同意，两国将在平等互利的基础上，进一步发展两国、两军在各个领域里的合作，把两国间的友好合作关系提高到一个新水平。
ID283	黑山口岸联检部门将原来要二至三天办完的出入境手续改为一天办完。
ID282	商检部门将原来七至十天完成的产地检验时间缩短到一至三天。
ID2792	公报要求美国尊重古巴的领土完整，把关塔那摩军事基地还给古巴，并立即停止封锁。
ID2467	为了造福社会，王码电脑公司毅然放弃本来可以赚大钱的机会，把不加密的软盘奉献给用户，让他们自行拷贝。
ID2332	蚊子进食时就会把植物细胞毒素吸入体内。
ID2297	在世界经济一体化与日俱增的环境下，各国面对全球化带来的挑战，应通过持续推行健全的经济政策以及深化结构改革来从全球化进程中最大限度地受益并把负面影响减少到最小程度。
ID2206	专程前来参加交接仪式的中国机械工业部部长包叙定将这批总价值为 5 5 0 万美元的机械产品交给了拉里奥哈省长马萨，这些产品包括 1 8 0 0 多台拖拉机和筑路工程机械。
ID176	深圳市沙头角保税区今后五年将充分发挥保税区的区位优势和政策优势，以高新技术产业为先导，积极调整产品结构，实施以转口贸易和仓储业为辅助的经营战略，把沙头角保税区建成按国际惯例运作、国内领先的特殊综合经济区域，使其成为该市外向型经济的快速增长点。
ID1592	随着中国开放的不断深入和投资环境的改善，不少跨国公司已将中国列为其海外投资重点区域，制定了对华投资战略，有的跨国公司已经或正在研究将其地区总部移至中国。
ID149	运河整治后，苏钢在运河对岸建了新厂区，并自筹资金建设了一座跨运河大桥，把新厂、老厂连为一体。
ID149	运河整治后，苏钢在运河对岸建了新厂区，并自筹资金建设了一座跨运河大桥，把新厂、老厂连为一体。
ID1434	李鹏强调说，海关是管钱、管物的单位，是执法单位，这就要求全国海关干部职工提高自身的政治素质和业务素质以及职业道德水平，并且尽快普遍采用更加先进的现代监管手段，把全国的海关工作提高到一个新的水平。
ID13817	「台湾女人还是不太能接受美国女性将身体练成肌肉成块的做法，她们喜欢轻盈、匀称的体态，所以对舞蹈的热中大过重量训练；至於男性则大多强调上半身的运动，身材要有型，但也不喜欢太壮，」健身教练雷克斯觉得很有趣。
ID13806	目前台湾外商公司中最具规模、拥有三家大型连锁店的「加州健身中心」亚洲区总裁 EricLevine 表示，他不苟同本地业者将 SPA、芳香疗法加进健身中心的做法，那会让人「懒惰」。

ID13563	从传统市场卖大闸蟹、杨山梨的小贩到有政府管理，专对外籍人士开放的五星级俱乐部，全是一个样：把台商当成肥羊，一剥再剥。
ID13505	孰料，为了将大台北污水处理厂设在八里的东北岸，八里被设计为观光乡的蓝图昙花一现，政府以市地重划的方式将耕地划为住宅区、工业区等用途，以换取厂区用地。
ID13393	回到台湾，夏阳将中国民间神佛列入他毛毛人的创作领域里。
ID13121	李春材解释，好市多的策略是尽量将经营成本降到最低。
ID12972	重点不在知识的高下深浅，而在如何把知识化为商品，赚进利润。
ID12335	微阵列晶片则是将基因的片段或寡核酸、蛋白质、抗体冷生物分子点在晶片上，小小一片晶片上可以点上数万个点。
ID12334	微流体晶片是结合生物技术、微电子、微机械等等技术，将实验室中许多仪器的功能微小化到晶片上来处理。
ID12311	潘诏智解释，临床上发烧症状难以迅速判断原因，发烧晶片将常见、会引起发烧的病毒、细菌的 DNA 放在晶片上，用来检测发烧的原因以及感染源（细菌、病毒）的种类，可以作为医生辅助判断病情的工具。
ID12265	本来以为仗著课堂上教的那一套，赚钱不会太难，没想到他听信某位大牌投顾的预测，把所有资金转进金融股，接著阿扁当选时又误以为大势已去，匆忙挂单卖出，结果卖到最低点，离出国旅游的目标又远了一步，令他十分懊恼。
ID12255	张毅君还认识一位跑产业新闻的线上记者，一辈子不曾进出股市，两年前却鬼迷心窍被顺大裕放出的内线消息吸引住，不但抵押房子，还把一票亲友拖进来，至今难以翻身。
ID12195	简瑞璞指出，外商也将「做好投资组合」、「鸡蛋不要放在一个篮子里」等风险观念带入台湾，对台湾股市产生了相当正面的影响，也使得 e 世代的理财活动，相对於他们父母辈的菜篮族，有了更大的理性基础。
ID11823	撑了几年，等二弟蔡瑞庭（从父姓）结婚後也承袭哥哥的生意，经营饼店和饮料；为了避免兄弟打对台，蔡瑞庭把店开在离东势十多公里的麦寮。
ID11122	巴万更是将他视为「自己人」，「假如我们是泰雅族，他就是荣誉国民，新赛德克（泰雅族的一支）人！」
ID10829	他以自己的家庭为例，他的父母亲是来自大陆的客家人，由於客家传统重视宗族观念，即使儿女都各自成家，早年每次过年他的父亲都要把儿女孙辈招到老家围炉吃年夜饭。
ID10787	然後用甘蔗皮、木炭、橘子皮、稻壳等作为燃烧材料熏烤，要花上大半天的时间，把腊肉熏到香气浓郁、色泽澄黄。
ID10477	人类社会中刻意筑起的界线与杆格，是人为加上去的，最後也终需人们将它打破。
ID10230	文章最后指出，我们应当把学习证券法、贯彻证券法提上日程，以保证这部法律得到正确的实施。
ID10058	从一九九五年起，联合国将每年的十二月二十九日确定为“国际生物多样性日”。
ID10048	联合国发言人席尔瓦 26 日宣布，由于塞拉利昂局势日趋紧张，联合国已决定将大约 40 名工作人员撤到临近的几内亚。
ID13256	为了带给观众身临其境的真实感，史博馆特别将展场入口设计在二楼，制造出一种由地面下到坑底、回溯历史的情境。
ID7384	索康曾对他的母亲说：“他是我的财产，我愿意把他揉成团，装在口袋里，随我；我高兴把他拉成条，围在腰上，也由我。”
ID6037	天津港务局投资五千多万元，计划用十八个月将盐码头专用泊位改造为一座一点五万吨级散、杂货两用泊位，把两个万吨级驳船泊位改建成两个万吨级通用杂货泊位，并对区内道路、铁路、供电、供水、通讯等辅助配套设施给予改造。
ID214	这些建议分为六个部分，分别是：按照“一线放开，二线隔离”的原则加速洋浦

	保税区的开发建设；借鉴国际经验，尽快把三亚建设成国际性旅游城市；采取优惠政策扶持和鼓励海南热带农业及加工业的发展；努力争取实现琼台农业项下的自由贸易；鼓励和支持海南与外资合作，开发南海海洋资源；按照扩大对外开放的要求，允许和鼓励海南大胆进行各项改革试验等。
ID8049	海外有舆论评论说：“这场空前的金融浩劫至少提前二十年把中国推上了影响全球金融的大国地位。”

（122 句，126 个用例）

附录七 动词与副词类

ID993	同时，西部地区还将拿出优势资源，吸引东部投资兴建原材料基地、联办资源开发与深加工以及劳动密集型项目。
ID9911	资本结构的优化，使青岛一批企业迅速崛起，青啤公司、海尔集团和双星集团的产品品牌已先后被认定为中国驰名商标。
ID9874	现在博士后专项经费已增加到二百万元。
ID9836	距成都八十六公里，占地一千亩的花水湾温泉开发区，经过两年的开发，已发展成为以温泉、森林为特色，集旅游、休闲、娱乐、康疗、观光、会务为一体的大规模多功能休闲娱乐中心。
ID982	目前，宁波北仑区沿海二十公里黄金海岸的一百多平方公里的带状地段，已星罗棋布地矗立起投资一千万美元的大项目近七十个，形成一个以石化、钢铁为主导，电力、机械、电子、轻纺、建材、塑胶等多种产业辅助发展的新兴大工业经济带。
ID9779	已经建成并投入使用的第三煤气厂、纪庄子污水处理厂、东郊污水处理厂、新开河厂、杨柳青电厂、大港电厂、电话网扩建、肿瘤医院教学医疗设备等项目都是利用政府间贷款兴办成功的。
ID9699	目前，在广西一千五百九十五公里长的大陆海岸线上，已经建成了大中型港口三个，泊位和码头七十六个，其中万吨级泊位十个，年吞吐能力已达到一千万吨，并与世界上一百三十多个国家和地区建立了贸易往来关系。
ID9683	目前，在浦东注册营业的外资金融机构已迅速增至二十二家，占上海外资金融机构总数的百分之四十三。
ID9558	在新的一年里，澳门将回到祖国怀抱。
ID9428	今天子夜新年钟声响起时，太湖的历史将翻开新的一页。
ID9094	此外，一些关于欧元与管理、欧元与信息技术、欧元与法律等内容的书籍也已列在新书预告的宣传单上。
ID9048	该计划的初衷是为俄罗斯战略火箭部队服务，现在已转为军民两用，并可为国外客户服务。
ID8879	调查人员将从野马的尸体中取出子弹，以帮助缉拿凶手。
ID8506	为此，欧洲央行将根据各国的经济发展情况，对欧元区的经济形势作出判断，以便找到适合整个欧元区经济发展的利率水平。
ID8503	为此，欧洲中央银行将担负起掌管欧元区 11 国货币政策、确定利率、管理货币储备及决定货币发行量等职能。
ID8417	今年 8 月，美国大学体育协会已经就限制使用金属棒作出决定。
ID8290	他说，自《告台湾同胞书》提出“三通”倡议以来，在“和平统一、一国两制”基本方针和各项具体政策的推动下，两岸经贸交流从无到有，从小到大，已结出丰硕的果实。
ID8230	事后，舟山受害方在有关各方大力帮助下，终于调查到了肇事船船东。

ID8071	西非维和部队 2 9 日在弗里敦发表声明，承认维和部队已撤出马克尼，但称这一行动是“战术性的”。
ID80	数年前，北海还是北部湾一个默默无闻的小渔村，然而三五年时间北海已建成了一个现代化都市的框架，街上客流如潮，楼房拔地而起。
ID7886	樱内义雄说，1 9 9 9 年中国将迎来建国 5 0 周年。
ID787	据悉，南宁市已规划制定出较完整的园林绿化体系，力争到二 0 0 0 年，全市绿化覆盖率、绿地率、人均公共绿地面积分别达到百分之三十七、百分之三十二和十平方米。
ID7798	通过与店中雇员攀谈才知道，其实商店中的收银和财物电脑系统早已装好了欧元记帐软件，并且已经开始在每张付款小条上同时打出里拉和欧元的价格。
ID7736	巴西教育经费 1 9 9 3 年占当年国内生产总值的百分之三点五，1 9 9 7 年已增加到百分之四点六，1 0 年之后即 2 0 0 7 年将进一步增加到百分之六。
ID7708	现在利用基因技术培育的转基因食品已经摆上了普通百姓的餐桌，基因疗法已经开始挽救患者的生命，更引人注目的是克隆技术的重大突破，使人们复制动物成为可能。
ID7703	在信息技术领域，被一些专家认为是第三次信息革命的突破口的人工智能技术，目前已被美国、日本和欧洲等列为科研计划的重中之重。
ID7681	同他们的前辈相比，这一代年轻人的目光正在投向整个欧洲大陆。
ID7348	自 1 9 7 9 年第一段 2 5 公里的干线投入运行以来，地铁运行里程已扩建至 1 2 0 公里，年客运量达 1 . 5 亿人次。
ID7314	其实，不仅仅在水口村，整个珠江三角洲地区的农村都已透露出“都市”的气息。
ID7282	最早招收台湾学生的福建中医学院如今已发展为海峡两岸中医理论教学和学术交流的一个重要基地。
ID719	中国中医研究院研究人员今天说，他们已成功地从草药中提取出对糖尿病有特殊疗效的新药“糖脉康”。
ID6948	此次活动在广州赛马会举行起步仪式，它将揭开一九九九中国（广州）生态环境游活动的序幕，同时也作为广州市迎接澳门回归祖国系列活动中的首项活动。
ID6920	责任者将同样受到该法庭的审判。
ID6880	十八年后，深圳的产业结构已调整为一点三九比四十八点七九比四十九点八二，工业和第三产业成为拉动经济增长的主要力量。
ID6798	随着一个覆盖全地区的高速网络的逐步形成，珠江三角洲已迎来了以“分钟”而不是“公里”度量距离的“高速公路时代”。
ID6722	根据国务院领导人指示，国家发展计划委员会已对昆禄公路建设质量问题作出处理决定：
ID6448	仅一年多的时间，这个公司生产的“大寨春”已售出了近二千吨。
ID6430	大寨，这方昔日以粮为纲的偏僻黄土高原，如今已走上了农工商并举的发展之路。
ID6382	现已建成包钢稀土一、二、三厂等稀土生产厂家，其中有世界最大的稀土精矿、稀土合金生产厂。
ID6375	内蒙古包头市依托雄厚的稀土资源，靠着以包钢稀土研究院为主的世界一流科研力量和几十年的开发利用，已经建成中国最大的稀土科研生产基地。
ID6317	目前它已被列为国家重点建设的四大综合性、多功能国际深水中转港之一。
ID6313	根据规划，面积六百四十三平方公里的湄州湾将划分为南北岸两个各具特色的区域进行建设。
ID6218	如今，这家企业已发展成为一个拥有四个子公司、十多个羊绒生产基地的企业集团。
ID6173	金融、技术、房地产、劳动力市场等要素市场初步形成，经济运行中大部分资源已经开始通过市场进行配置，社会保障制度和收入分配制度的改革已迈出重要步

	伐。
ID6160	甘肃白银、兰州、金昌、西成四大有色金属工业基地的建设正加紧进行；青海扩建的全国最大的电解铝基地指日可待；新疆在二〇〇〇年将建成以石油天然气为主的综合性化工基地；宁夏正在着力建设一千万千瓦的火电基地及商品电基地；陕西力争在本世纪末使全省煤炭生产能力达到一亿吨，发电装机容量达到一千万千瓦，天然气开采规模达到三十亿立方米。
ID6158	甘肃省境内黄河干流上的刘家峡、盐锅峡、八盘峡三座大型水电站建成后，已建成了有色金属企业和矿山三十多个，电解铜、电解镍、电解铝、铅锌等五大高耗能生产系列，年产量在全国名列前茅。
ID6157	据不完全统计，这四个省区已先后建成铝、镍、铜、铅等有色金属加工企业三十多家，累计产值达六百多亿元，成为中国重要的有色金属冶炼基地。
ID6005	据介绍，江苏省去年公布了“九五”期间首批重点招商项目，其中二十个项目已寻求到合作伙伴，另外三十个项目继续对外招商。
ID5992	他说，该项目的实施将推动中叙两国经贸关系进一步发展，标志着中国纺织机械成套设备制造水平已提高到一个新的水平，具备了参与国际竞争的能力。
ID5933	金卡工程已建立起省市县银行互联网络系统，发行各种信用卡、储蓄卡六百多万张，作为金关工程的广东省无纸贸易中心已正式运行，省国际贸易通关网络中心正在筹建；金桥工程广东已和北京、上海等地联网调试成功，实现视频图像、数字、语音传送及与国际互联网联接等功能。
ID5718	生产成本约三、四美元一颗的 16MbD-RAM，在最低时只能卖到一．七美元，等於是「做越多赔越多」。
ID5695	再拨拨算盘，风暴後利率高涨，南韩目前利率已飙至百分之二十，外币贷款部份更是负担倍增，又有多少企业能不被拖垮？
ID5668	以受害最严重的印尼为例，去年六月底，一美元可以兑换二千四百元印尼盾，这个数字在今年二月中旬已贬破九千，贬值幅度高达百分之七十五（指其货币购买力跌掉四分之三，只剩下原来的四分之一）。
ID5534	较早前劳顾会已研究过这五项法例，认为需要更多时间考虑其中的条款，不宜原封不动提交立法局，仓卒通过，其中以雇员代表权、谘询权和集体谈判权等影响深远的事宜为然。
ID550	上海市经委的研究人员在提交的报告中指出，经过大量投入和积极培育，汽车、电子通信设备等六大支柱产业在未来三年内仍将保持良好增长势头，占全市工业的比重将上升到百分之五十五左右，并推动上海第二产业增加值年均增长百分之九至百分之十。
ID5440	发言人说，新闻及广播中心原定七月十日关闭，但由於所有租用工作间的传媒机构已离开中心或在七月九日前已做好船运安排，故中心可较原定时间略为提早关闭。
ID5369	麦隆礼说：「『岩土指南第五册』内已列出有关修葺斜坡工程的技术标准，该指南并已分发各有关人士。
ID5096	总额约一千七百亿港元的土地基金资产今日（星期二）移交香港特别行政区政府後，特区的财政储备已增至约三千五百亿港元。
ID5050	香港，经历了一百五十六年的漫漫长路，終於重新跨进祖国温暖的家门。
ID4997	所有国家和地区在香港的正当经济利益将受到法律保护。
ID4938	他说，委内瑞拉沥青石储量丰富，达到四百亿吨，而中国已经制定了一个规模宏大的能源发展计划；奥历若克乳化油作为一种新型燃料，已经受到越来越多的电力企业的青睐。
ID4926	目前，宜昌三峡机场已经建成；武汉至宜昌的高速公路已通车。
ID4906	目前，这家外企服务总公司的人才资源服务已跨入海外市场。

ID4903	现在公司已派出雇员二万人，服务于四十八个国家和地区的近五千家外商驻京机构和三资企业。
ID4837	此前，这个省已经购进一架“空中客车”A 3 2 0 客机，并已于近日投入营运。
ID4809	与此同时，西藏自治区有关部门已为区内公路建设制定出中长期规划。
ID4804	目前西藏一百二十万平方公里土地上长达二万二千三百余公里已建成的公路，连接了百分之七十七的乡，构成了全区公路交通网的雏形。
ID4753	施祥鹏说，一九八九年我在这里投资了一千多万美元，现在资产已发展到七亿元人民币，所有的收益我都没有收回去，我要继续投资在这里。
ID4643	宝成复线、进藏铁路等十条铁路，有的在修建电气化复线，有的正在紧张施工，有的即将开工，有的已列入国家计划。
ID461	到去年底，水东开发区已经办起了各类企业达二千四百多家，工业总产值达八点八四亿元。
ID451	这样，包括国内二十多个省市在内，目前水东开发区内签订了协议或合同的客商已经达到七百五十多家，总投资将突破一百五十亿元，实际到位资金已有三十多亿元，已上马建设的项目达一百六十多个。
ID4495	自一九五八年加入 I E C、一九七八年加入 I S O、一九七二年加入 I T U 以来，I E C、I S O 现行一万四千二百多项国际标准，中国国家标准中已采用六千四百六十三项，占国际标准总数的百分之四十五点五；I T U 制定的国际标准建议，凡适合中国国情的已基本转化为中国标准；中国现有一万七千七百二十项国家标准，其中百分之四十点五采用了国际标准和国外先进标准，有些重点行业如电工、电子、汽车、化工、冶金、邮电、船舶等领域，国家标准采用国际标准的比例已达到百分之七十以上；大约百分之五十的重点工农业产品已按国际标准和国外先进标准组织生产。
ID4462	他表示，“一国两制”是中国政府为推进祖国统一大业进程提出的指导性原则，它已载入中国宪法，成为一项长期不变的基本国策，香港将继续保持国际金融中心地位。
ID4403	目前中国县以上的包装企业已发展到一万五千家，每年完成三万亿元的工农产品包装任务，包装工业体系正在逐步完善，并且与世界五十多个国家和地区在包装技术方面开展合作交流。
ID4402	中国生产的木浆和纸包装材料、购物袋、塑料编织袋、复合软包装、金属大桶等制品已出口到欧洲、东南亚、香港等国家和地区，并大力发展防伪包装，绿色包装。
ID4399	全国包装科技工作会议提供的情况表明，中国包装工业的发展有长足进步，包装工业总产值占社会总产值的百分之一点五，在国民经济四十个主要行业中，包装行业已上升到第十六位。
ID4354	据宁波市外经贸委主任戴龙岳介绍，宁波市的综合经济实力已跨入中国大中城市前茅，全方位对外开放格局业已形成。
ID4280	据国家计委的一位官员介绍，自中国政府大力实施长江开发开放以来的五年间，长江两岸已建成西陵（宜昌）、武汉二桥、九江、铜陵、扬中等五座长江大桥，总投资超过百亿元。
ID4250	报道说，迄今，可经营人民币业务的外资银行已扩展到九家，其中通过有关检验于一九九七年三月正式营业的六家外资银行，到六月底人民币存款逾二亿，贷款逾一亿。
ID3984	本届锦标赛到今天为止，已决出 7 枚金牌。
ID3455	沈国放说，港英的三级政制架构，即港英最后一届立法局、两个市政局、区议会，只能存在到一九九七年六月三十日，这是英国对香港的管治届时终结的必然结果。
ID3189	苏梅克—列维九号彗星是美国天文学家在一九九三年三月发现的，现已分裂成二

	十一颗彗核，正高速向木星运行，将于七月份与木星碰撞爆炸。
ID3049	经过多年建设，国家重点实验室已建成一批基础性研究的实验研究基地，其中既涉及基础研究领域和新兴科技领域，也有与国民经济发展息息相关的工程项目。
ID2835	经过多年探索，在反复试验中终于研制出“治肝胶囊”。
ID278	目前，全盟已撤掉各类堵卡站二十多个。
ID2776	目前，十个农业高新技术开发示范区，已划出了十八万亩土地作为中心示范区。
ID2774	据了解，目前，山东已建立起十个象莱芜市这样的农业高新技术开发示范区。
ID2606	不知不觉间浓荫几乎完全封死了马路的上空，原来马路已经变窄了，并且有了坡度
ID2477b	经党中央批准，在安徽泾县兴建的皖南事变殉难烈士陵园已建成。
ID2339	他还呼吁各国团结一致对付外债、贩毒以及暴力活动，本届安第斯议会会议将延续到十六日。
ID2117	韩国世运工业株式会社与威海经济技术开发区共同出资二千九百八十万美元兴办的“黑豹”农用车项目，总设计能力达年产六十万辆，目前年产十万辆农用车的一期工程已建成投产。
ID2062	香港著名财团长江实业、百富勤作为战略性投资者已购入了“深业控股”百分之十二的股权，充分反映出投资者的信心。
ID1999	去年，中国建设银行和工商银行的个人住房抵押贷款规模已增加到一百亿元。
ID1960	石万鹏说，纺织工业摆脱困境的条件已基本具备：中国对纺织工业扭亏解困，出台了一系列强有力的政策和措施；纺织行业广大干部职工已经意识到面临困难的严重性、扭亏解困的紧迫性，不仅事关自身生存和发展，而且对整个国有经济发展和稳定的全局具有深远影响。
ID1837	大连瓦轴集团B股上市发行股票已融资达三点九亿美元。
ID1834	利用外资的方式辽宁已呈现出多元化发展趋势。
ID1795	恩戈兰在法国——科特迪瓦商人俱乐部举办的经济人士座谈会上说，西非经济和货币联盟各成员国的平均年经济增长率1996年已恢复到百分之五点九，1997年增长到百分之六点三，而1994年这一增长率仅为百分之二点六。
ID168	经济专家预计，今年中韩两国贸易额将增至二百五十亿美元。
ID1637	库场面积达六十多万平方米，年中转六十万吨散装水泥和五十万吨粮食的中转库已经建成，是全国四大水泥出口基地和十大接粮口岸之一。
ID1576	早些时候，赣北九江港已建成五千吨级的集装箱码头，南昌港成为江西继九江港之外的又一个“出海口”。
ID15129	到底她有什么魅力？她的成功是否代表了东南亚歌手已在台湾抢滩成功？
ID13620	一九九九年，该公司的营收超过新台币上亿元，去年更达到三亿台币；九六年创站时不到五名员工，四年内员工人数已扩增到一百九十人。
ID13466	摊开地图，台湾的三〇九个乡镇是一个又一个陌生的符号，还是你曾经驻足停留过的美好？
ID1331	天津港保税区投入运行五年来，已建成了中国第一货物分拨中心，具备了口岸关的功能，开通了天津港保税区经西安、兰州到新疆阿拉山口口岸的铁路专用线；建立了一批集仓储、运输、销售于一体的大型物流配给中心，开办了铁路和国际集装箱多式联运，月接卸集装箱能力达六千标准箱；形成了七千门程控电话的装机能力，供电能力达二点五万千瓦、日供水能力一万吨。
ID13054	同理，神圣不可侵犯的麦当劳「金色双拱门」商标谁不会画？永保明亮欢乐的店面设计和促销手法也未必有多大学问，然而当这么简单的商品融合了专利注册、品牌实力，以及严格的管理品质後，竟可以吃遍全世界，让全球二万多家加盟店替它壮大事事业版图，还无条件地永远奉上百分之五营业额做为服务费！
ID12963	台湾已经准备好了吗？

ID129	曾荣获国家重大发明奖的以氮作保护气直拉硅单晶技术，已由该区浙大半导体厂转化为国内外同类产品中的佼佼者，目前硅单晶年产已达二十多吨，居全国第一，其销售市场已从国内和东南亚扩展欧美许多国家。
ID1289	海现已建成四个万吨级码头，今后又将有四个万吨级泊位投入兴建。
ID1282	防城港市现已发展成为全国重要枢纽港之一，拥有生产性码头二十个，其中万吨级以上泊位十个，铁路和高速公路直达港口。
ID12315	李锺熙指出，根据美国市场研究的估计，目前全球的微阵列生物晶片用量约二十五万片，两年後，预估成长到一百二十万片，以一片一千美元的市价估算，到西元二〇〇三年，会有十二亿美元的规模，而全球生物晶片市场十年後将可成长到四百亿美元的规模。
ID12254	前阵子台风飙到每股二百多元时，靠著内线消息大赚一笔的朋友多麼得意，不到两年，台风已被打到下市、成为废纸一堆。
ID12104	台湾的命运更加坎坷，曾经先後受到强权的欺凌和殖民政权的统治。
ID11789	为了欢迎来客，他们早已在餐厅订好一大桌海鲜大餐，高龄阿嬷及兄弟、妯娌、侄子等都从附近乡镇过来团聚，展现强韧的家族凝聚力。
ID1171	首期开发的一平方公里的东片区，已建成厂房面积二十五万平方米，在建厂房四十幢，电力开关站、保税区、科技宾馆等一批配套设施已相继建成。
ID11698	中老年男子追求美貌，某一程度也就是从健康角度出发，只是现代的养生观念已延伸到美容保养领域。
ID1163	据汕头市科委主任张英强介绍，目前，汕头市已经初步形成了一批高新技术产业和骨干企业，其中，已被广东省科委认定的高新技术企业二十家，超声、海洋两家集团公司已被国家科委认定为“全国首批国家重点高新技术企业集团”。
ID1139	近年来，这个律师事务所已在国际上打赢了十余起涉外经济案件，为国内企业和公民追回赔款近千万美元。
ID1133	段祺华律师和他的同事们为此又进行了艰苦的调查取证，终于取得了文件递送的快件号码和第一国民银行签收的证据。
ID1118	经济技术开发区做为当地新的经济增长点，加工贸易发展迅速，目前已有二百二十家从事加工贸易的外资企业，保税工厂已占到企业总数的百分之四十一，逐步形成了涉及电子、化工、纺织、通讯以及汽车等行业的保税工厂体系，摩托罗拉、三星电子、雅马哈等都是其中的一员。
ID1085	根据今天签定的泄洪大坝及相关工程保险协议，若工程发生保险责任范围内的自然灾害或意外事故造成的经济损失，中保财产保险公司将按规定进行赔偿。
ID10828	「最怕的是，中国人强固的宗族观念，常常造成父母和孩子之间永远也纠葛不清的那种痛苦而亲密的感情，反而形成享受欢乐的阻碍，」以擅写家庭故事的作家小野表示，年节往往就是一个具体的冲突源。
ID10719	八点档一向被当作电视台的招牌和赚钱与否的指标，除了得有商业卖点，过去还需符合政治正确性，台语曾经一度被视为「不雅」，根本就与八点档绝缘，三台都有默契，台语不能在八点「阖家观赏」的时段出现，只能於中午时段或六点半档演出。
ID10694	早期的小说中，有一篇〈鱼〉曾经收入在国中国文课本里，而广为人知。
ID10679	黄春明曾经在许多场合提到自己由祖父母带大、与祖父母格外亲近的身世。
ID10313	家住湖北省新洲县的叶章荣在今年罕见的大洪水中失去了家园，不过在元旦到来之前，他们一家已经住进一幢崭新的二层小楼里
ID7558	火灾发生后，消防队员立即赶到出事现场，经过两个多小时的努力，终于将大火扑灭，并从灰烬中找到7具被烧焦的尸体。
ID9665	首期投资五百万元的中国近代海军博物馆日前已进入封顶阶段，将于明年一月建成。

ID6330	建设者们为这片开发区设计的战略步骤是：现在至二〇〇〇年的起步阶段，将建成以秀屿、霞塘、忠门为中心的三大组团式区域，总面积达二十九平方公里，港口年吞吐量达一千万吨；二〇〇〇年至二〇一〇年的成形阶段，将建成城区面积五十八平方公里，港口吞吐量达两千万吨；二〇一〇年至二〇二〇年的高速发展阶段，城区建成面积在一百平方公里以内，港口吞吐量达五千万吨。
ID1362	据浦东新区经贸局对浦东开发七年来引进的投资一千万美元以上的一百五十七个工业大项目跟踪调查，目前建成投产的有一百一十六个，投产率高达百分之七十三点九；已建成投产项目的建设时间，最长三十个月，短的只有十一个月，平均建设周期为十七点二个月。
ID7428	据介绍，这家公司的反电脑病毒专家们已编制出可检测和清除“远程探险者”的程序。
ID4927	万县五桥机场、宜昌至广州的铁路等基础设施项目以及一批水电、火电项目都已建成或在建，库区二十一个县（市、区）的程控电话和国内国际直拨电话也已开通。
ID4768	目前，全长三百五十公里的宜（昌）黄（石）高速公路、宜昌市六万门程控电话工程和三峡机场均已建成并投入使用。
ID6258	他们指出，香港回归后，内地与香港的经贸合作必将上升到新的水平，拓展到新的领域，也必将为中国明天的富强、统一和文明缔造坚实的基础。
ID1747	目前连接三市的泉厦、厦漳高速公路已建成或基本建成，港口、机场、铁路、通信等一大批重点基础设施也相继完工。
ID8291	截至今年十一月底，两岸贸易已累计达一千三百四十七亿美元，其中对台出口二百一十六亿美元，自台进口一千一百三十一亿美元。
ID1404	自一九八〇年新中国最早的一家外商投资企业——北京国际配餐公司成立以来，中国已累计批准了三十万家外商投资企业。
ID4692	据介绍，小浪底大坝工程开工不到一年的时间里，业主就收到了一百四十七封索赔信，在C I P M的协助下，最终认定只有四十七封信函的索赔理由是站得住脚的。
ID8807	现在，塞政府终于做出让步，宣布了一个折中办法，即同意不把导弹部署在塞岛，但也不取消合同，而是准备把它们部署在希腊的克里特岛。
ID9857	浙江省国内生产总值从一九七八年的二百三十二点九亿元上升到今年的五千余亿元，经济总量已跃升到中国第四位。
ID10024	巴格达上空硝烟虽已暂时散去，但美伊之间的对抗并不因美英动武而结束。

（140句，142个用例）

附录八 动词与数词分离类

ID8291	截至今年十一月底，两岸贸易已累计达一千三百四十七亿美元，其中对台出口二百一十六亿美元，自台进口一千一百三十一亿美元。
ID1404	自一九八〇年新中国最早的一家外商投资企业——北京国际配餐公司成立以来，中国已累计批准了三十万家外商投资企业。
ID9725	日照市政府最新公布的数字表明，截止目前，全市已累计批准外资企业四百四十二家，合同利用外资四点九亿美元，实际利用外资一点八二亿美元。
ID9482	文化部的统计显示，一至十二月中旬，两岸文化交流项目总计达二百七十三项、参加人数达两千六百七十三人次。
ID1816	吴晓灵还披露，到目前为止，中国境内机构在国际资本市场上累计发行外币债务折合美元约为一百七十亿美元，有力地支持了国家经济快速、持续、健康发展。

ID13268	根据陶俑身上留下的刻文名字考证，当时约有八、九十位陶工，各自帶著十几位助手投入陶俑的制作，总计应有近千名来自官方和民间的陶工合力完成这些需要塑、堆、捏、贴、刻、画的陶俑。
ID10359	五家企业技术力量雄厚，设备精良，但普遍负债大、冗员多、包袱重，经济效益持续下滑，累计亏损五点七二亿元，资不抵债。
ID8332	从一九九〇年《福建省台湾同胞投资企业登记管理办法》和《福建省台湾同胞投资企业劳动管理规定》制定以来，福建省已累积制定地方性涉台法规八部，初步形成了具有福建地方特色的涉台法规体系。
ID9929	至今年三月底，武汉市共吸引世界上五十多个国家和地区的投资，累计审批利用外资项目近五千个，协议外资金额七十一亿美元，实际利用外资四十八亿美元；全市兴办“三资”企业三千七百四十家，协议外资金额四十四亿美元，实际进资二十二亿美元，进资率达百分之五十，高于全国平均水平十七个百分点。
ID998	来自外经贸部的最新消息说，截止今年十一月底，中国累计批准设立外商投资企业三十万零二千四百六十四家。
ID9790	其中，进口九百五十八点二万吨，出口五百五十二点八万吨，分别比去年同期增长了百分之五十八点一和百分之三十八点二，国际海运集装箱吞吐量累计达五十七点六万标准箱，比去年同期增长了百分之三十四点五。
ID96	如今，甘肃省的外资企业已不再为投资风险担忧，为确保对外开放的顺利进行，“八五”期间（一九九一至一九九五年），甘肃省涉外保险业发展迅速，累计已经提供六百一十一亿元的风险保障，承保范围包括财产、责任、信用、人身四大类主、副六十多个险种。
ID953	预计全省乡及乡以上工业累计完成销售产值二百三十亿元，比去年增长百分之十六，工业经济总体效益回升，独立核算工业企业实现利润增加。
ID9502	吨油成本比去年平均降低三十九元，行政费支出比去年实际水平降低百分之十。
ID943	目前，汕头市累计批准利用外资合同一万二千五百二十二宗，实际利用外资五十一亿多美元，有四千七百五十六家三资企业投产、开业。
ID8931	四月四日，县政府办公室用此款购回桑塔纳 2 0 0 0 型轿车一辆，合计开支十九点四二万元，此车相对固定给崔广华使用。
ID870	据统计，累计进行五百九十九项专题研究，取得一百八十六项重大科研成果，其中一百三十九项成果荣获中国省、部级以上奖励，十一项成果达到国际先进水平。
ID8680	国家财政从一九九八年起五年内累计投入二十五亿元用于支持该项工作。
ID8010	塔拉修克在当天举行的记者招待会上说，1998年乌克兰共签署了180个双边和国际条约，21个多边协议。
ID776	今年，全区累计完成工业产值三十七亿元，技工贸总收入四十五点五亿元，出口贸易额一万二千七百四十七万美元，实现税利三点六亿元。
ID774	目前，汕头高新技术开发区已累计投资十八亿元，建成厂房三十七万平方米，区内水、电、通讯、保税仓库等基础配套设施已建成投入使用。
ID7257	截止十二月二十七日，天津港完成集装箱吞吐量一百零八十三万标准箱，提前四天超额完成一百万标准箱的全年目标；累计完成货物吞吐量六千七百四十八万吨，提前四十二天完成六千万吨的全年生产计划。
ID6984	截止到十二月二十日，中国累计收购皮棉六千三百六十三万担，已超过去年全年棉花收购量。
ID66	据统计，到去年底，中国银行向外商投资企业累计发放的外汇现汇贷款和人民币贷款达到二百一十亿美元和二千五百九十三亿元，人民币贷款余额已近四百五十四亿元。
ID6537	该市场始建于一九八四年，占地三十六点六公顷，累计投资近二亿元，并建成了直通北京的绿色通道，每天有二百辆卡车向首都运输新鲜蔬菜。

ID65	仅去年中国银行就累计向外商投资企业提供了六百九十多亿元的人民币贷款，另外还向外商投资企业发放外汇现汇贷款四十多亿美元，这些贷款重点支持基础原材料、化工、机械等行业。
ID6232	到去年底，进出保税区的货物总值累计已超过五十四亿美元，国际集装箱进出区量已达每月六千个标箱。
ID6228	已累计开发土地五平方公里，基础设施投资十一亿元，达到“七通一平”的投资要求。
ID61	据统计，迄今中国已累计批准外商投资项目近二十五点九万个，实际使用外资金额一千三百五十四亿美元。
ID6077	据了解，自一九九二年底武汉长飞光纤光缆公司与美国签订中国第一份光纤产品外销合同以来，中国已累计向国外出口光纤四十多万公里，从而结束了中国光纤光缆长期依赖进口的局面。
ID6043	截至今年三月，福建省福州市累计批准外资企业近五千家，合同外资一百二十八亿美元，实际利用外资四十六亿美元。
ID5837	目前已累计批准三资企业一千四百三十二家，利用外资十四点七亿美元。
ID553	经专家测算，今后三年上海固定资产投资总额预计为六千亿元至七千亿元，社会消费品零售总额预计年均增长百分之十三至十四，外贸出口预计年均增长百分之十至十二，这些规模与国内生产总值年均增长百分之十至十一是相适应的。
ID552	上海市计委主任韩正说，今后三年上海社会总需求的适度增长，也可以支撑国内生产总值年均增长百分之十至十一。
ID546	上海统计局公布的数字显示，一九九二年至一九九六年上海国内生产总值年均增长百分之十四点二，今年的增长速度也将达到百分之十三以上
ID510	连云港钟山氨纶丝有限公司是九十年代初引进日本东洋纺织技术和设备的一家苏港合资企业，前两期工程建成并投产以来，已累计实现利润一点四七亿元，跻身于国家二类大型企业行列。
ID509	这个采用差别化氨纶丝生产技术改造的项目，总投资七千万元，累计年产氨纶丝一千五百吨。
ID4891	据大连市政府官员介绍，近几年来，外商对大连市农渔业的投资逐年增加，到一九九六年，全市农村有外商投资的合资、合作、独资企业达一千六百多家，累计吸引外资金额十五点一亿美元，年实现出口供货额一百二十多亿元人民币。
ID4882	从一九九〇年至今年七月底，兵团累计承担对周边国家工程承包、劳务合作项目二十一项，合同总金额一千多万美元；境外投资项目四十八项，合同金额两千多万美元。
ID4840	随着基础设施不断改善，浙江省的航空运输总量以每年百分之二十五的速度递增，五年累计运送旅客一千八百多万人次，货物吞吐量约三十万吨。
ID4803	此后四十年来，中央政府累计投资四十亿元人民币，先后建成了（四）川（西）藏公路、青（海）（西）藏公路（西藏段）等干线公路。
ID4723	至此，合肥外商投资企业累计达一千一百多户，总投资额近四十二亿美元。
ID472	投产以来累计实现综合经济效益十点六亿元、出口创汇四千八百五十六万美元，相当于收回了一期工程的静态总投资。
ID4605	截止今年九月底，去年九月建成投产的企业，已累计生产柴油机四万台，创产值八千五百万元，为库区安置移民近三百人。
ID4522	一九九二年至一九九六年累计实际利用外资三十点七亿美元。
ID4357	在投资说明会上，宁波市向与会客商推出了七十八个招商项目，合计总投资四十六点七八亿美元。
ID4355	截止去年底，宁波市累计批准“三资”企业三千五百八十五家，总投资九十三亿美元，实际利用外资十八亿美元。

ID4340	目前，中国陆上石油对外开放地区总面积已达二百五十万平方公里，累计与三十五个外国公司签订了三十个石油合同，合同金额达七点七亿美元。
ID4314	“八五”（一九九一至一九九五年）期间，全省利用外国政府贷款项目累计达七十七个，贷款金额近十亿美元。
ID4184	建区以来，累计签约三资企业项目二千五百多项，协议外资额六十多亿美元。
ID378	据外经贸部提供的资料，“八五”期间中国的进出口总额累计达一点零一万亿美元，比“七五”时期（一九八六年至一九九〇年）增长一倍以上，在世界贸易中的排位由一九九〇年的第十五位上升至第十一位。
ID3667	十年来，中国化工行业累计吸收外资达八十五亿美元，用于化肥、石化的建设和技术改造。
ID3262	十五年来，在全省尚有四十多个贫困县接受政府财政补助的情况下，已累计支付二点六亿元专为难民解决生产生活困难。
ID28	三年来，这些城市累计完成固定资产投资一百二十亿元，昔日边境城市的“楼不高，路不平、灯不明、水不清、通讯不畅”的状况已得到了改变。
ID2798	截止八月三十日，该市日投劳力十五点八三万个，累计投工九十三点六五万工，已完成土石方七十三万立方米。
ID264	“八五”期间（一九九一至一九九五年），福建省乡镇企业累计上缴了一百八十五点六亿元人民币的税金，完成出口交货值一千零五十五亿元人民币。
ID225	到去年十二月底，区内已累计设立企业一千六百一十四家，总投资达十二亿美元，其中外商投资企业二百六十家，实际利用外资一点一三亿美元。
ID2232	国际物流和转口贸易规模不断扩大，进出区货物总值累计已超过六十亿美元，国际集装箱进出区量达到每月六千多个标准箱。
ID222	据统计，至去年年底，宁波保税区累计完成进出口贸易额八点一二亿美元，仅去年一年通过保税区海关的进出口贸易额就达三点六五亿美元。
ID2112	全市韩国投资企业累计达五百六十六家，去年对韩国进出口总额十二点三亿美元，占当年全市进出口总额的近六成。
ID2057	目前港商在深圳实际投入累计已近七十亿美元，占深圳实际利用外资总额的百分之六十五。
ID2032	到去年末，贷款的本息回收达到百分之百，累计实现利润一千万美元。
ID1870	据天津市政府台办负责人介绍，截止到一九九七年底，天津批准设立的台资企业累计已达一千二百七十一家，协议投资额近二十二点八五亿美元，协议台资额十八点五亿美元。
ID186	八年来，先后有十多个国家和地区的客商进保税区投资，到去年十二月底，已累计批准外商投资企业二百一十六家。
ID1827	六年间，辽宁已有近六千家外资企业开业投产，累计实现产值近二千七百亿元。
ID1780	据介绍，近几年来，油田加大了对油气资源的勘探和开发投资力度，投入勘探资金近十亿元，探明油气储量年均增长四千多万吨。
ID1779	自一九五四年勘探开发以来，已累计生产原油一千五百五十多万吨，加工原油近八百万吨。
ID1753	全年新批外商投资企业接近一千六百家，累计已达二万余家。
ID1741	从一九九三至一九九七年十月，该市合同利用外资累计达四十八点八亿美元，实际到资十九点一亿美元，相当于改革开放十八年来总和的百分之九十。
ID1735	三角地区的泉州市，过去五年间国民生产总值年均递增百分之四十二点九，辖下的晋江、石狮、南安三市跻身全国经济百强县。
ID1603	仅中国陆上三大天然气最富集的四川盆地，近四十多年来，已累计生产一千六百三十三亿立方米天然气，基本上解决了成都、重庆等一批大中城市的民用燃料，并形成以天然气为原料的中国最大的维尼龙生产线——四川维尼龙厂。

ID1601	到一九九七年底累计探明储量五千一百亿立方米，其中重庆地区二千六百亿立方米；共获气田八十六个。
ID1552	一九九六年外商投资企业进出口额一千三百七十一亿美元，是一九九二年的三点一三倍，年均增长百分之三十三点零五；占全国进出口总额的比重从一九九二年的百分之二十六点四三上升到百分之四十七点三。
ID13596	据估计，一般人一生长达三、四十年的劳动时间，平均会换七次工作。
ID1316	十省市合计出口一千六百二十一点九亿美元，占全国的百分之八十八点八。
ID1210	近年来，台湾从祖国大陆获得的贸易顺差累计已达八百亿美元，与台湾的外汇储备相差无几。
ID1174	高新区建立三年来，技工贸总收入累计达三十三亿元，工业总产值为二十四亿元，出口创汇达六千多万美元。
ID1094	到十一月底，全国企业欠缴的工商税款累计达到了四百一十九亿元，其中欠缴国内增值税和消费税就达三百零七亿元。
ID1073	有鉴于此，中国进出口银行在成立之初即开始了与国有商业银行的合作，并最先与中国银行签署了《委托代理协议》，三年来两行间的委托代理业务开展顺利，在中国进出口银行累计为出口企业提供的三百二十六亿元人民币出口卖方信贷中，近百分之七十的贷款是通过中国银行总行及其分支机构拨付到企业的。
ID1062	今年预计全省乡及乡以上工业累计完成销售产值二百三十亿元，比去年增长百分之十六。
ID1054	目前，汕头市累计批准利用外资合同一万二千多宗，实际利用外资五十一亿多美元。
ID10394	据侨委会的统计，近十年来澳门同学累计已有六千人左右来台念书。
ID1035	据统计，到目前为止中国已累计批准外商投资企业逾三十万家，实际使用外资约二千亿美元。
ID84	统计资料显示，过去五年广西对外贸易和利用外资规模迅速扩大，进出口贸易额累计达到一百亿美元，其中出口六十八点七亿美元，分别比“七五”时期（一九八六至一九九〇年）增长一点七八倍和一点四三倍；实际利用外资累计达到三十三点二四亿美元，占改革开放以来累计总额三十八点三九亿美元的百分之八十四点四；边贸成交额一百二十四亿元。
ID14	海关统计表明，“八五”期间（一九九〇年—一九九五年），中国外商投资企业的进出口呈直线上升之势，出口年均增长百分之四十三点二，进口年均增长百分之三十八点六。
ID289	据不完全统计，一九九四年以来，呼盟新建三资企业四十二家，累计引进国外资金四千多万美元、国内资金六千多万元，引进国内外先进技术三十多项，与东北、华东、华南、北京等地结成合作对子一百多个。
ID5615	根据行政院公平交易委员会的统计，截至八十四年底，全国登记有案的多层次传销公司，共计有五百四十八家，然而因为竞争激烈，实际有活动的，大约只有二百一十家。

（87句，88个用例）

附录九 框式副词类

ID12199	贾坚一回忆，民国五十年代的台湾，经济困窘，民众根本无财可理；又为了怕打仗，一攒到钱就去买黄金、换黑市美钞。
ID474	一是从一九九四年十一月十八日破土动工到锥炉点火成功，仅用了十三个月，工期比安玻一期缩短了十八个月，比世界建设最快的同类项目还提前五个月；二是

	实现了国内外人才、技术、设备、设计和施工等资源的最佳组合，无论硬件配置和软件应用都达到了当今国际同行业先进水平；三是与成套引进国外技术和设备相比，投资大大节省。
ID10518	「二十年前，台湾孩子穿的太子龙制服一下水就硬梆梆，棉衣裤一洗就缩水，尼龙衣穿在身上又红又痒，丝袜一刮就破，好一点的衣服都要花钱送去乾洗……，」中国纺织工业研究中心总经理姚兴川回忆，现在台湾的衣服越来越轻柔、保暖、防皱防脏、色泽光鲜而经久不褪，这可不是靠灰姑娘故事中的仙女魔棒，而是纺织产业一点一滴的研发升级成果所致。
ID9282	对于有足球王国之称的巴西来说，1998年是一个失望多于兴奋，胜利与失败并存，成绩既令人高兴但又让人感到缺憾的一年。
ID9031	货币名称欧元既用拉丁字母也用希腊字母标识。
ID8813	塞岛导弹问题一开始就引起了美英等国的关注和不安。
ID8364	“一到我们让她参加比赛的时候，她不是说手有伤就是说背部有伤，如果那些地方没伤，她就是手腕有伤，总之是不能比赛。
ID6	去年初浦东新区诞生的中国第一家医疗机构药品采购服务中心，正因为一开始就比较规范，运转至今，成交药品一亿多元，没有发现一例回扣。
ID5622	一般正规的传销公司都有坚强的生产实力，对自己的产品有满意保证，并且会鼓励加入者先从使用产品开始，销售时只要将产品的使用心得平实地说出来就好，不必加油添醋。
ID1660	他在回顾了九七年世界经济金融形势后说，由于包括日本在内的亚洲经济增长进一步放慢，即使现在经济运行较好的美国和其它国家的经济增长也将受到很大影响。
ID15157	当孙燕姿在隆冬穿上皮海大衣拍摄MTV，兴奋、新鲜地向大家炫耀她第一次上身的冬衣时，台湾好像与终年常夏的新加坡失去了距离，我们也感染到一个女孩飘洋过海在既熟悉又陌生的异国被接纳後的幸福感。
ID15131	孙燕姿去年六月在市场上现「声」的首支单曲「天黑黑」，一推出就引起热烈的回响。
ID14679	「念头一转，海阔天空，」邱加利指出，自我的调适很重要，人生没有「非要怎样不可」的，只要你能接纳自己就好。
ID14541	目前姜小姐和一位比她小十岁的男友交往，「没有拘束，没有强迫，只有真正想在一起的时候才在一起，」加上前年父亲病危时，三不五时地打电话来向儿女乞讨关爱，令她又同情又厌憎。
ID14267	作出来的人面陶罐，连刘益昌都赞不绝口。
ID12984	许文星指出，以往人类资讯都必须寄存於物质上，只有少数人才能拥有把玩；一旦寄身的物质毁损，资讯也随之佚失。
ID12881	例如资讯人的内部网路即提供有聊天室、请假单、工作守则、出差单，甚至连员工之间的二手交流区，也都一应俱全。
ID12306	王献煌指出，传统的细菌、病毒培养要七到十天，用分子生物方式检测也需要三天才知道结果，目前晶宇发展出缩短检体前处理时间的方法，再搭配肠病毒检测晶片，只要六个小时就可以得知结果。
ID12292	它将有如身份证，找工作、买保险、甚至找对象、结婚都得用到它。
ID12290	身体不舒服时，只要到药房买片检验的生物晶片，滴上一滴尿，就可以知道哪里有毛病，身体检查如同现在验孕一样容易。
ID11844	对于他的专业——社会研究、田野调查等，父母既没兴趣也不了解，然而他们从来不代表反对；尤其当弟弟也步上後尘，从企管系转攻社会学时，父母依旧未置一词。
ID11811	「她爸爸（吴建乐）二十二岁就结婚了，女孩子怎麼能拖？」阿嬷又爱又怜地说。

ID10756	连续剧卖座，电视台就没完没了地拍下去，边拍边播就成了台湾八点档的特色，品质当然也无法太讲究。
ID10698	黄春明的小说写到辛酸处，真的可以教人边读边噙著泪水。
ID14291	「这个『香菇头』，卑南遗址有挖出来，」郑永国说：「考古学家不知道是什麽，但我们做陶的人，一看就知道是什麽用途。」

（25 句，28 个用例）

参考文献

中文书目

专 著

- 丁声树等 《现代汉语语法讲话》（北京：商务印书馆，1961）。
- 王 力 《汉语史稿（第二册）》（北京：科学出版社，1958）。
- 王 力 《中国现代语法》（北京：商务印书馆，1985）。
- 王 惠 《现代汉语名词词义组合分析》（北京：北京大学出版社，2004）。
- 太田辰夫《中国语历史文法》蒋绍愚、徐昌华译，（北京：北京大学出版社，1958/1987）。
- 石毓智 《现代汉语语法系统的建立—动补结构的产生极其影响》（北京：北京语言大学出版社，2003）。
- 冯志伟 《机器翻译研究》（北京：中国对外翻译出版公司，2004）。
- 冯志伟 《现代语言学流派》（西安：陕西人民出版社，1999）。
- 朱德熙 《语法讲义》（北京：商务印书馆，1982）。
- 朱德熙 《语法答问》（北京：商务印书馆，1985）。
- 朱德熙 《现代汉语语法研究》（北京：商务印书馆，1980）。
- 吕叔湘 《汉语语法分析问题》（北京：商务印书馆，1979）。
- 吕叔湘 《现代汉语八百词》（北京：商务印书馆，1980）。
- 刘春卉 《现代汉语属性范畴研究》（成都：巴蜀书社，2008）。
- 刘丹青 《语序类型学与介词理论》（北京：商务印书馆，2003）。
- 刘涌泉 乔 毅 《应用语言学》（上海：上海外语教育出版社，1991）。
- 刘海涛 《依存语法的理论与实践》（北京：科学出版社，2009）。
- 邢福义 吴振国 《语言学概论》（武汉：华中师范大学出版社，2002）。
- 汤廷池 《国语语法研究论集》（第73页）（台北：台北海园书局，1978）。

- 沈家煊 《不对称和标记论》(南昌:江西教育出版社,1999)。
- 汤廷池 《国语变形语法研究:第一集 移形变位》(台北:台湾学生书局,1977)。
- 沈阳 郑定欧 《现代汉语配价语法研究》(北京:北京大学出版社,1995)。
- 志村良治 《中国中世语法史研究》江蓝生、白维国译,(北京:中华书局,1984/1995)。
- 张志公 《汉语语法常识》(北京:中国青年出版社,1953)。
- 张敏 《认知语言学与汉语名词短语》(北京:中国社会科学出版社,1998)。
- 陈望道 《文法简论》(上海:上海教育出版社,1978)。
- 陈昌来 《现代汉语动词的句法语义属性研究》(上海:学林出版社,2002)。
- 赵元任 《汉语口语语法》吕叔湘译 (北京:商务印书馆,1979)。
- 胡裕树 《现代汉语(重订本)》(上海:上海教育出版社,1995)。
- 施春宏 《汉语动结式的句法语义研究》(北京:北京语言大学出版社,2008)。
- 袁毓林 《汉语配价语法研究》(北京:商务出版社,2010)。
- 钱乃荣 《现代汉语》(北京:高等教育出版社,1990)。
- 黄伯荣 廖序东 《现代汉语(增订版)》(北京:高等教育出版社,1991)。
- 黎锦熙 《新著国语文法》(北京:商务印书馆,1924)。
- 张敏 《语料库,知识获取与汉语依存分析》,清华大学计算机系硕士论文,1993。
- 彭艳 《短语结构语法与依存语法的心理现实性研究》,上海外国语大学博士学位论文,2007。
- 胡憻 《基于多维特征属性描写的现代汉语概念语义网的建构研究》,武汉大学博士论文,2007。
- 周丽颖 《跟语序有关的几种句法现象研究》,苏州大学博士论文,2007。

论 文

- 马真、陆俭明 <形容词作结果补语情况考察>,见《汉语学习》1997年,

第 1、4、6 期。

马 真 <修饰数量词的副词>，见《语言教学与研究》1981 年第 1 期，页 53-60。

马庆株 <自主动词和非自主动词>，见《中国语言学报》1988 年第 3 期，
又载马庆株编《二十世纪现代汉语语法论文精选》，商务印书馆 2005 年版，页 224-245。

王红旗 <动结式述补结构品类家研究>，载沈阳、郑定欧主编《现代汉语
配价语法研究》，（北京：北京大学出版社，1995）。

王希杰 <施受、词序、主宾语>，见《语法研究和探索》四，北京大学出版社，1988，页 130-140。

王冬梅 <动名互转的不对称现象及成因>，见《语法研究和探索》十二，
商务印书馆，2003，页 274-291。

王跃龙 <“N+N”同位结构辨析及其依存图标注>，见《词义与计算----第
九届汉语词汇语义学研讨会论文集》，COLIPS PUBLICATION，
SINGAPORE，页 333-340。

方 梅 <汉语对比焦点的句法表现手段>，见《中国语文》1995 年第 4 期，
页 279-289。

方 梅 <宾语与动量词语的次序问题>，见《中国语文》1993 年第 1 期。

文 炼、胡 附 <汉语语序研究中的几个问题>，见《中国语文》1984 年
第 3 期。

文 炼 <词语之间的搭配关系>，见《中国语文》1982 年第 1 期。

文 炼 袁 杰 <谈谈动词的“向”>，见《汉语论丛》，华东师大出版社
1990。

文 炼 胡 附 <词类划分中的几个问题>，见《中国语文》2000 年第 4 期，
页 298-302。

申小龙（晓珑）<汉语语序研究述评>，见《语文导报》1986 年第 9 期。

申小龙 <汉语动词分类研究综述>，见朱一之 王正刚选编《现代汉语语法
研究的现状和回顾》（北京：语文出版社，1987）页 76。

- 司马翎、沈阳 <汉语动结式的句法构造和补语小句的内部结构>，见《语法研究和探索》第十四辑。(北京：商务印书馆，2008)。页 22-37。
- 史存直 <从汉语语序看分布理论>，见《语言教学与研究》1982 年第 1 期，页 4-21。
- 石定栩 <动词的名词化和名物化>，载《语法研究和探索（十二）》，北京：商务印书馆。2003 年 12 月第一版，页 255-273。
- 冯胜利 <论汉语的韵律结构及其对句法构造的制约>，见《语言研究》1996 年第 1 期，页 108-123。
- 冯胜利 <汉语动补结构来源的句法分析>，见《语言学论丛》第二十六辑，（商务印书馆，2002），页 178-208。
- 冯志伟 <特斯尼耶尔的从属关系语法>，见《国外语言学》1983 年第 1 期。
- 吕叔湘 <漫谈语法研究>，《中国语文》1978 年第 1 期，又载马庆株编《语法研究入门》，（北京：商务出版社，1999）。1999 年 2 月第一版，页 2-12。
- 吕叔湘 <主谓谓语句举例>，载《吕叔湘文集》第三卷，1986，页 531-544。
- 朱德熙 <现代汉语语法研究的对象是什么？>，见《语法研究和探索》四，北京大学出版社，1988，页 11-27。
- 任 鹰 <主宾可换位动结式述语结构分析>，见《中国语文》2001 年第 4 期，页 320-327。
- 刘子瑜 <汉语动结式述补结构的历史发展>，见《语言学论丛》第三十辑。（北京：商务印书馆，2004），页 188-264。
- 刘街生 <动结式组构的成分以其关系探讨>，见《语言研究》2006 年 6 月第 2 期。页 63-68。
- 刘知远 <汉语词同现网络的小世界效应和无标度特性>，见《中文信息学报》2007 年第 6 期，页 52-58。
- 刘海涛 <汉语语义网络的统计特性>，见《科学通讯》2009 年第 14 期，页 2060-2064。
- 刘海涛 冯志伟 <自然语言处理的概率配价模式理论>，见《语言科学》2007

- 年第3期, 页32-41。
- 刘海涛 <语言网络: 隐喻, 还是利器? >, 见《浙江大学学报(人文社会科学版)》2010年12月, 页1-12。
- 刘丹青 <形名同现及形容词的向>, 见《南京师大学报》1987年第3期, 页56-61。
- 刘丹青 <语义优先还是语用优先----汉语语法体系建设断想>, 见《语文研究》1995年第2期(总第55期), 页10-15。
- 刘丹青 <试论汉语语法的宏观研究>, 见邵敬敏 刘大为主编《九十年代的语法思考》(北京: 北京语言学院出版社, 1994) 页11。
- 邢 欣 <主谓谓语句的范围>, 见《新疆师范大学学报(哲学社会科学版)》1995年第1期, 页76-80。
- 李临定 <动补格句式>, 《中国语文》1980年第2期, 又载马庆株编《二十世纪现代汉语语法论文精选》, 商务印书馆2005年版, 页575-586。
- 李 敏 <“大主语隐含介词的主谓谓语句”再分析>, 见《语言教学与研究》1996年第3期, 页81-92。
- 李 敏 <现代汉语主宾可互易句的考察>, 见《语言教学与研究》1998年第4期, 页50-58。
- 李胜梅 <现代汉语中“被”的词缀化倾向>, 载邢福义主编《汉语被动表述问题研究新拓展》, 武汉: 华中师范大学出版社。2006年5月第1版, P186-196。
- 李芳杰 <定语易位问题刍议>, 见《语文研究》1983年第3期, 页24-31。
- 李芳杰 <大主语或小主语是“的”字结构的主谓谓语句>, 见《武汉大学学报(哲学社会科学版)》1998年第4期(总第237期)。
- 李英哲 陆俭明译 <汉语语义单位的排列次序>, 见《国外语言学》1983年第3期, 页33-39。
- 李宇明 <存现结构中的主宾互易现象研究>, 见《语言研究》1987年第2期。
- 吴福祥 <试论现代汉语述补结构的来源>, 见《汉语现状与历史的研究》,

(北京: 中国社会科学出版社, 1999)。

吴为善、吴怀成 <双音述宾结果补语“动结式”初探—兼论韵律运作、词语整合与动结式的生成>, 见《中国语文》2008年第6期(总第327期), 页498-503。

沈阳 <现代汉语句法结构中名次短语部分成分移位现象初探>, 见《语言教学与研究》1996年第1期。

沈阳 <名词短语分裂移位与非直接论元句首成分>, 见《语言研究》2001年第3期, 页12-28。

沈阳 <数量词在名词短语移位结构中的作用与特点>, 见《世界汉语教学》1995年第1期(总第31期), 页14-20。

宋文辉 <关于宾语必须前置的动结式>, 见《汉语学报》2006年第4期。

寿永明 朱绍秦 <领属关系主谓谓语句分析>, 见《浙江大学学报(人文社会科学版)》, 2002年3月, 第32卷第2期。

范继淹 <多项NP句>, 见《中国语文》1984年第1期, 又载《范继淹语言学论文集》, 语文出版社1986年版, 页239-251。

范开泰 <语用分析说略>, 见《中国语文》1985年第6期, 又载马庆株编《二十世纪现代汉语语法论文精选》, 商务印书馆2005年版, 页59-68。

范晓 <关于汉语的语序问题(一)>, 见《汉语学习》, 2001年10月第5期, 页1-12。

范晓 <关于汉语的语序问题(二)>, 见《汉语学习》, 2001年12月第6期, 页18-28。

范晓 <动词的“价”分类>, 见《语法研究和探索》五, 语文出版社1991。

陈平 <汉语双项名词句与话题—陈述结构>, 见《中国语文》2004年第6期, 页493-507。

陈平 <试论汉语中三种句子成分与语义成分的配位原则>, 见《中国语文》1994年第3期。

陈淑梅 <鄂东方言“把得”被动句>, 载邢福义主编《汉语被动表述问题研究新拓展》, 武汉: 华中师范大学出版社。2006年5月第1版, P33-45。

周迟明 <汉语的使动型复合动词>, 见《文史哲》第四期, 1958年, 页 175-226。

周 明 黄昌宁 <面向语料库标注的汉语依存体系的探讨>, 见《中文信息学报》1994年第3期, 页 35-52。

陆俭明 <汉语口语句法里的易位现象>, 见《中国语文》1980年第1期, 页 28-41。

陆俭明 <动词后趋向补语和宾语的位置问题>, 见《世界汉语教学》, 2002年第1期, 页 5-17。

陆俭明 <现代汉语中的数量词的作用>, 见《语法研究和探索》四, 北京大学出版社, 1988, 页 172-186。

孟维智 <主谓谓语句的范围>, 见《语法研究和探索(二)》1984年第2期。

胡裕树 <试论汉语句首的名词性成分>, 见《语言教学与研究》1982年第4期。

胡裕树, 张斌 <汉语的结构特点和语法研究>, 见[日]《中国语》1981年第8期, 又见《淮北煤师院学报(社会科学版)》, 1983年第2期, 页 89-94。

胡裕树, 张斌 <划分与切分>, 见《中国语文》1999年第10期, 页 243-245。

胡裕树, 范晓 <试论语法研究的三个平面>, 见《新疆师范大学学报》1985年第2期。又载马庆株编《二十世纪现代汉语语法论文精选》2005年, 页 1-11。

胡明扬 <语义语法范畴>, 载马庆株编《语法研究入门》, 北京: 商务出版社。1999年2月第一版, 页 195-199。

洪笃仁 <从现代汉语的词序看所谓“倒装”>, 见《厦门大学学报》1955年第4期, 页 119-134。

姚汉铭 <试探复合动词转化成名词的界限>, 见《河南师大学报》1980年第5期, 页 81-86。

袁毓林 <述结式配价的控制—还原分析>, 见《中国语文》2001年第5期, 页 399-410。

袁毓林 <定语顺序的认知解释及其理论蕴涵>, 见《中国社会科学》1999

年第2期, 页185-200。

袁毓林 <一价名词的认知研究>, 见《中国语文》1994年第4期, 页241-253。

袁毓林 <现代汉语名词的配价研究>, 见《中国社会科学》1992年第3期, 页205-223。

袁毓林 <基于统计的语言处理模型的有用性和局限性>, 载袁毓林著《基于认知的汉语计算语言学研究》2008年7月第1版, 北京: 北京大学出版社。P33-80。

曹广顺 <试论汉语动态助词的形成过程>, 见《汉语史研究集刊》2, (成都: 巴蜀书社, 1999)。页74-89。

曹聪孙 <语法图解法的比较研究>, 见朱一之 王正刚选编《现代汉语语法研究的现状和回顾》(北京: 语文出版社, 1987) 页210。

梅祖麟 <从汉代的“动杀”、“动死”来看述补结构的发展>, 见《语言学论丛》第十六辑, (北京: 商务印书馆, 1991)。

郭 锐 <述结式的配价结构和成分的整合>, 载沈阳、郑定欧主编《现代汉语配价语法研究》, (北京: 北京大学出版社, 1995)。

郭 锐 <汉语动词的过程结构>, 见《中国语文》1993年第6期。又载马庆株编《二十世纪现代汉语语法论文精选》(北京: 商务印书馆, 2005), 页287-298。

郭 锐 <述结式的论元结构>, 载徐烈炯、邵敬敏主编《汉语语法研究的新拓展(一)——21世纪首届现代汉语语法国际研讨会论文集》(杭州: 浙江教育出版社, 2002)。

徐 枢 <从语法、语义和语用角度谈“名_受+名_施+动”句式>, 载《语法研究和探索(四)》, (北京: 北京大学出版社, 1988)。

黄昌宁 张小凤 <自然语言处理技术的三个里程碑>, 见《外语教学与研究》2002年5月第3期, 页180-187。

陶红印 <相互动词及相互动词句>, 见《句型和动词》语文出版社, 1987, 页344-382。

蒋绍愚 <汉语动结式产生的时代>, 见《国学研究》第6卷, (北京: 北京

大学出版社，1999)。

蔡 莉 <NP₁+V²+NP₂ 句式 中 NP₂ 分裂前移初探>，见《语文学刊》2008 年第 3 期，页 110-112。

潘允中 <汉语动补结构的发展>，见《中国语文》1980 年第 1 期，页 53-60。

薛 红 <动补结构研究概述>，见朱一之、王正刚选编《现代汉语语法研究的现状和回顾》(北京：语文出版社，1987)，页 154-165。

英文书目

Books

Bennett, David. *Stratificational Grammar*. In Asher, Ronald (ed.) *Encyclopedia of Language and Linguistics*. Oxford, 4351-6.

Chao Yuen Ren. *A Grammar of Spoken Chinese*. University of California Press, 1968.

Givón, T. *Syntax : a functional-typological introduction*. Amsterdam ; Philadelphia : J. Benjamins Pub. Co, 1984.

Goldberg, Adele. *Constructions: A Construction-Grammar Approach to Argument Structure*. Chicago: University of Chicago Press, 1995.

Halliday, Michael. *An Introduction to Functional Grammar*. London, 1985.

Hudson, R. A. *Word Grammar*. Blackwell, 1984.

Hudson, R. A. *English Word Grammar*. Blackwell, 1990.

Hudson, R. A. *Language Networks: The new Word Grammar*. Oxford University Press, 2007.

Jackendoff, Ray. *Semantic Interpretation in Generative Grammar*. Cambridge, Mass: MIT Press, 1972.

- Karlsson, Fred, Atro Voutilainen, Juha Heikkilä, and Arto Anttila, editors. *Constraint Grammar: A Language-Independent System for Parsing Unrestricted Text*. Natural Language Processing, No 4. Mouton de Gruyter, Berlin and New York. ISBN 3-11-014179-5, 1995.
- Kruijff, G.-J. M. *A Categorical-Modal Logical Architecture of Informativity: Dependency Grammar Logic and Information Structure*. PhD thesis, Charles University, 2001.
- Lamb, Sidney. *An Outline of Stratificational Grammar*. Washington, DC. 1966.
- Wood, M. M. *Categorical Grammars*. Routledge Publisher, 1993.
- Mel'čuk, I. A. *Dependency syntax: theory and practice*. Albany: State University Press of New York, 1988.
- Schröder, I. *Natural Language Parsing with Graded Constraints*. PhD thesis, Hamburg University, 2002.
- Sgall, P., Hajičová, E. and Panevová, J. *The Meaning of the Sentence in Its Pragmatic Aspects*. Reidel, 1986.
- Starosta, S. *The Case for Lexicase: An Outline of Lexicase Grammatical Theory*. Pinter Publishers, 1988.
- Tesnière, L. *Éléments de la syntaxe structurale*. Paris: Klincksieck, 1959.
- Tenny, C. & Pustejovsky J. *Events as Grammatical Objects: The converging perspectives of lexical semantics and syntax*. Edited by Carol tenny & Jammes pustejovsky. CSLI publications. Stanford, California, 2000.
- Tang, Ting-chi (汤廷池) *A Case Grammar of Spoken Chinese*. Taipei: Hai-Guo Book. Co, 1972.
- Tsao, Feng-fu. *A Functional Study of Topic in Chinese: The First Step Towards Discourse Analysis*. Taipei: Student Book Co, 1979. (中文版《主題在漢語中的功能研究-邁向語段分析的第一步》. 謝天蔚譯. 北京: 語文出版社. 1994年初版. 1996年第二版)

Articles

- Amaral, L. A. N., Scala, A., Barthélemy, M. & Stanley, H. E. "Classes of behaviour of small-world networks", *Proc. Natl Acad. Sci. USA* 97, 2000, pp.11 149--11 152.
- Bach, E. "On Time, Tense, and Aspect: An Essay in English Metaphysics", In Peter Cole (ed.), *Radical Pragmatic*, (Academic Press, New York. 1981).
- Bach, E.; Brown, C.; and Marslen-Wilson, W. "Crossed and nested dependencies in German and Dutch: A psycholinguistic study", *Language and Cognitive Processes*, 1(4), 1986, pp. 249-262.
- Bartha, Csongor and Spiegelhauer, Tilmann and Dormeyer, Ricarda and Fischer, Ingrid. "Word order and discontinuities in dependency grammar", in *Acta Cybernetica*, Volume 17, Number 3, 2006.
- Broker, N. "Separation surface order and syntactic relations in a dependency grammar", In *Proceeding of the 36th Annual Meeting of the ACL and 17th International Conference on Computational Linguistics*, Montrea, 1998.
- Carroll, Glenn and Eugene Charniak. "Learning Probabilistic Dependency Grammars from Labelled Text", *Proceedings from AAAI Fall Symposium on Probabilistic Approaches to Natural Language*, AAAI Press, Cambridge, MA 1992.
- Cancho & Solé. "The Small World of Human Language", *Proceedings: Biological Sciences*, Vol. 268, No. 1482, Nov. 7, 2001, pp.2261-2265.
- Convington, M.A."A fundamental algorithm for dependency parsing", in *Proceedings of the 39th Annual ACM Southeast Conference*, 2001, pp. 95-102.
- Davis, H and Demirdache, H. "On lexical verb meaning: evidence from Salish", In *Events as Grammar Objects*, Carol Tenny and James

- Pustejovsky (eds). (CSLI Publication, 2000).
- Debusmann, R. and Duchier, D. And Kruijff, G.-J. M. "Extensible dependency grammar: A new methodology", *Proceeding of the Workshop on Recent Advances in Dependency Grammar*, 2004, pp. 78-85.
- Dikovsky, Alexander. "Polarized non-projective dependency grammars", in *Lecture notes in computer science*, (Springer, 2001), Volume 2099, pp.139-157.
- Dorogovtsev, S. N. & Mendes, J.F.F. "Accelerated growth of networks", In: *Handbook of graphs and networks. From the genome to the Internet*, Bornholdt, S and Schuster, G. H. (eds.) Winheim, Wiley-VCH, 2003, pp. 318-339.
- Duchier, D. "Axiomatizing dependency parsing using set constraints", *Proceedings of the Sixth Meeting on Mathematics of Language*, 1999, pp. 115-126.
- Duchier, D. "Configuration of labeled trees under lexicalized constraints and principles" , *Research on Language and Computation 1*. 2003, pp. 307-336.
- Duchier, D. and Debusmann, R. "Topological dependency trees: A constraint-based account of linear precedence", *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2003, pp. 180-187.
- Duchier, D. and Debusmann, R. "Topological dependency trees: A constraint-based account of linear precedence", in *39th Annual Meeting of the Association for Computational Linguistics (COLING)*, 2001, pp. 340-345.
- Eisner, Jason "Bilexical grammars and their cubic-time parsing algorithms", In Harry Bunt and Anton Nijholt (eds.), *Advances in Probabilistic and Other Parsing Technologies*, Kluwer Academic Publishers, October 2000,

pp. 29-62.

Eisner, J. M. "An empirical comparison of probability models for dependency grammar", *Technical Report IRCS-1996-11*, Institute for Research in Cognitive Science, University of Pennsylvania.

Eisner, J. M. "Three new probabilistic models for dependency parsing: An exploration", *Proceedings of the 16th International Conference on Computational Linguistics (COLING)*, 1996, pp. 340 – 345.

Gan Kok-Wee, Lua Kim-Teng. "Chinese character classification using an adaptive resonance network." *Pattern Recognition*. 1992. 25(8): 877-882.

Gaifman, H. "Dependency systems and phrase-structure systems". *Information and Control* 8, 1965, pp.304 – 337.

Gerdes, K. and Kahane, S. "Word Order in German: A Formal Dependency Grammar using a Topological Hierarohy". In *proceeding of ACL'2001*. PP.220-227.

Greenberg, J. "Some universals of grammar with particular reference to the order of meaningful elements", in J. Greenberg (ed.) *University of Language*, (Cambridge: MIT Press, 1966).

Greenberg, J.H. "The diachronic typological approach to language." In Shibatani, M and T. Bynon (eds.) *Approach to language typology*. Oxford: Clarendon. 1995, PP.143-166.

Harper, M. P. and Helzerman, R. A. "Extersions to constraint dependency parsing for spoken language processing", *Computer Speech and Language* 9:1995, pp.187-234.

Hays, David G. "Dependency theory: A formalism and some observations", *Language* 40. 1964, pp.511 – 525.

Hellwig, P. "Dependency unification grammar", *Proceedings of the 11th International Conference on Computational Linguistics (COLING)*, 1986,

- pp. 195-198.
- Hellwig, P. "Dependency unification grammar", In Agel, V., Eichinger, L. M., Eroms, H.-W., Hellwig, P., Heringer, H. J. and Lobin, H. (eds), *Dependency and Valency*, Walter de Gruyter, 2003, pp. 593-635.
- Hudson, R. "Language as a cognitive network", in Hanne Gram Simonsen and Rolf Theil Endresen (eds) *A cognitive Approach to the Verb: Morphological and Constructional Perspectives*. (Mouton de Gruyter Press, 2001).pp. 49-72.
- Huang, C.-T. James. Historical syntax meets phrase structure theory: Two notes on the development of verb-complement constructions. Paper presented at *the Joint Meeting of the 4th International Conference on Chinese Linguistics and 7th North American Conference on Chinese Linguistics*, The University of Wisconsin-Madison.1995.
- Järvinen, T. and Tapanainen, P. "Towards an implementable dependency grammar", In Kahane, S. and Polguere, A. (eds), *Proceeding of the Workshop on Processing of Dependency-Based Grammars*, 1998, pp. 1-10.
- Ji Donghong. "Semantic Annotation of Chinese Phrase using Recursive Graphs", In *Proceedings of ACL'2000 workshop on Chinese Information Processing*, Hong Kong. 2000, Pp. 101-108.
- Ji Donghong. "Design and Construction of Conceptual Graph Bank", *The 3th Workshop on Chinese Lexical Semantics*, May 2002, Tokyo.
- Ji Donghong et al. "Building a Conceptual Graph Bank for Chinese Language", in *Proceeding of International Conference on Lexical Resource and Evaluation*, 2004a.
- Johnson, R. King, M and Tombe (1985) "EUROTRA: A multilingual system under development", *Computational Lingusitics 11*, PP.155-169.
- Karlsson, F. "Constraint grammar as a framework for parsing running text",

- In Karlgren, H. (ed.), *Papers presented to the 13th International Conference on Computational Linguistics (CONLING)*, 1990, pp.168-173.
- Kwong, O.Y. and Tsou, B.K. "Categorial Fluidity in Chinese and its Implications for Part-of-speech Tagging". In *proceeding of the Research Note Session of the 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL-2003)*, Budapest, Hungary. Pp.115-118.
- Kwong, O.Y. and Tsou, B.K. "A Synchronous Corpus-Based Study of Verb-None Fluidity in Chinese". *Journal of Chinese Language and Computing*, 2004, 13(3):227-278.
- Kruijff, Geert-Jan M. "Formal and Computational Aspects of Dependency Grammar", Lecture notes for ESSLLI-2002.
- Lecerf, Y. Programme des conflits, modele des conflits, *Traduction automatique*, 1, No.4, 1960.
- Lecomte, Alain. "Proof-nets and dependencies", in *Proceeding of COLING-1992*, Nantes, pp394-401.
- Li, C., & Thompson, S. A. "Chinese", In B. Comrie (Ed.), *The world's major languages*. (London: Croom Helm, 1990), pp. 811-833.
- Li-li Chang, Keh-jiann Chen and Chu-Ren Huang. A Lexical-Semantic Analysis of Mandarin Chinese Verbs: Representation and Methodology. *Computational Linguistics and Chinese Language Processing*. Vol.5, No.1, February 2000, pp.1-18.
- Maruyama, H. "Structural disambiguation with constraint propagation", *Proceedings of the 28th Meeting of the Association for Computational Linguistics (ACL)*, Pittsburgh, PA, 1990, pp. 31-38.
- McConnell-Ginet, S. "Adverbs and logical form. A linguistically realistic theory", *Language* 58: I 1982, pp.44-84.

- McDonald, R., Crammer, K. and Pereira, F. "Online large-margin training of dependency parsers". *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL)*2005, pp. 91-98.
- Mel'čuk, I. A. *Word order in Russian*. In I. Boguslavskij, L. Iomdin, L. Krysin, eds, *Slovo i jazyk (Sbornik statej k vos'midesjatiletiju akademika Ju.D. Apresjana)*, pp. 499-525. Jazyki slavjanskix kul'tur, Moskva. 2011.
- Menzel, W. and Schröder, I. "Decision procedures for dependency parsing using graded constraints", In Kahane, S. and Polguère, A. (eds), *Proceedings of the Workshop on Processing of Dependency-Based Grammars*, 1998, pp. 78-87.
- Miller, G. A., Beckwith, R. Fellbaum, C., Gross, D & Miller, K. J. " WordNet: an on-line lexical database", *International Journal of Lexicography* 3, 1990 , pp.235-244.
- Miller, G. A. "WordNet: a lexical database for English", *Communications of the ACM* 38, 1995, pp.39-41.
- Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S., Ayzenshtat, I., Sheffer, M. & Alon, U. "Superfamilies of evolved and designed networks", *Science* 303, 2004, pp.1538-1542.
- Nivre, Joakim and Nilsson, J. "Pseudo-projective dependency parsing", *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics (ACL)*. 2005, pp.99-106.
- Nivre, Joakim. "Sorting Out Dependency Parsing", In *Proceeding of the 6th International Conference on Natural Language Processing*. 2008, pp. 16-27.
- Nivre, Joakim. "Constraints on Non-Projective Dependency Parsing", In *11th Conference of the European Chapter of the Association for Computational Linguistics*. (EACL). 2006. PP.73-80.

- Ninio, A. "Predicting the order of acquisition of three-word constructions by the complexity of their dependency structure", *First Language*, 14, 1994, pp.119-152.
- Ninio, A. "Acquiring a dependency grammar: The first three stages in the acquisition of multiword combinations in Hebrew-speaking children". In G. Makiello-Jarza, J. Kaiser & M. Smolczynska (Eds.), *Language acquisition and developmental psychology*. (Cracow: Universitas. 1998).
- Pustejovsky, J. "The syntax of event structure", In *Cognition* 41 (Elsevier 1991), pp. 47-81.
- Reape, Mike. "Domain union and word order variation in German", In Nerbonne et al. (Nerbonne et al., 1994), pp. 151 - 197.
- Ritter, E. and Rosen, S. "Event structure and ergativity", In *Events as Grammar Objects*, Carol Tenny and James Pustejovsky (eds). (CSLI Publication, 2000).
- Robinson, Jane J. "Dependency structures and transformational rules", *Language* 46. 1970, pp. 259 - 285.
- Samuelsson, C. A statistical theory of dependency syntax. *Proceedings of the 18th International Conference on Computational Linguistics (COLING)*, 2000.
- Sleator, D. and Temperley, D. "Parsing English with a link grammar", *Technical Report CMU-CS-91-196*, Carnegie Mellon University, Computer Science.1991.
- Sleator, D. and Temperley, D. "Parsing English with a link grammar", *Third Internatioanl Workshop on Parsing Technologies (IWPT)*, 1993, pp. 277-292.
- Strogatz, S. H. Exploring complex networks. *Nature* 410, 2001, pp. 268-276.
- Swinney, D. "Lexical access during sentence comprehension: (Re-)consideration of context effects", *Journal of Verbal Learning and*

- Verbal Behavior* 5, 1979, pp. 219-27.
- Sylvain Kahane, Alexis Nasr, and Owen Rambow. "Pseudo-projectivity: A polynomially parsable non-projective dependency grammar", In *36th Annual Meeting of the Association for Computational Linguistics and 18th International Conference on Computational Linguistics (COLING-ACL)*, 1998 , pp. 646-652.
- Tai H-Y, James. (戴浩一) "Temporal sequence and Chinese word order", In John Haiman (ed.), *Typological Studies in Language: Iconicity in Syntax*, (John Benjamins Publishing Company Press, Amsterdam/Philadelphia, 1985). Pp. 49-69.
- Tai H-Y, James. (戴浩一) "Chinese as a SOV Language", Papers from the Ninth Regional Meeting Chicago Linguistic Society, April 13-15, 1973, ed. By Claudia Corum, T. Cedric Smith-Stark and Ann Weiser, Chicago Linguistic Society.
- Tai H-Y, James. (戴浩一) "On the change from SVO to SOV in Chinese", Papers from the Parasession on Diachronic Syntax, April 22, 1976, ed. By Stanford B. Steever, Carol A. Walker, and Salikoko S. Mufwene, Chicago: Chicago Linguistic Society.
- Tang Li and Ji Donghong. "A Semantic Analysis Model for Chinese Sentence", In *Proceedings of Workshop on Chinese Information Processing(ACL)*, 2004b.
- Tapanainen, Pasi & Jarvinen, Timo. "A non-projective dependency parser", In *Proceeding of the Fifth Conference on Applied Natural Language Processing (ANLP)*. 1997, pp.64-71.
- Taku kudo, Yuji Matsumoto "Japanese Dependency Analysis Based on Support Vector Machines", EMNLP/VLC 2000.
- Vendler, Zeno. "Verbs and Times", *The Philosophical Review*, Vol.66, No.2 (April, 1957), pp. 143-160.

- Wang Yuelong and Ji Donghong. "A study on Annotation of Chinese Sentences Using Dependency Graph", in *Recent Advance of Chinese Computing Technologies*. COLIPS Publications, 2007, Singapore. Pp.177-182. (ISTP)
- Watts, D.J. & Strogatz. S. H. Collective dynamics of 'small-world' networks. *Nature* 393, 1998, pp. 440-442.