

名词短语：定语、中心语的语义类别及定中搭配*

张卫国

中国人民大学语言文字研究所

wgzhang@sina.com.cn

提要： 考察现代汉语中以名词做中心语的名词短语里中心语和定语的语义类别以及语义类别与定中搭配的关系，为计算机对名词短语的处理提供语义方面的基础。研究以《现代汉语辞海》和《同义词词林》提供的材料为基础在计算机上进行。介绍了研究的步骤、得到的有关数据及结论：与语义类别相同的名词搭配的词语的语义类别相同或基本相同；与语义类别相近的名词搭配的词语的语义类别有同有异；与语义类别不同的名词搭配的词语的语义类别不同或者相差很远。

关键字： 名词短语 定中搭配 语义类别

Noun Phrase: Semantic Classification of Modifiers and Head in Modifier-head Constructions

Zhang Weiguo

Institute of Chinese Language and Writing, Renmin University of China

Abstract: An introduction to the research in the semantic classes of modifiers and heads in the noun phrases as well as the relationship of the semantic classes and modifier-head construction. The research is performed in computer based on books *Word Corpus of Modern Chinese* and *The Thesaurus of Chinese*, and can be used to support the computer's processing of noun phrases. Brief explanations of the steps, data of the research is given. The main conclusion of the study is that: the words modifying nouns with identical semantic classes are identical in semantic classes, the words modifying nouns with close semantic classes are same or near in semantic classes, and the words modifying nouns with different semantic classes are different in semantic classes.

Keywords: noun phrases, modifier-head construction, semantic classes

1 题解

在汉语信息处理中，句子的核心是动词（包括形容词），以动词为核心的动词组合框架的研究已经有了很好的成果，为计算机对汉语的分析和处理提供了较好的基础。而在以动词为核心的组合框架中，动词以外最重要的部分是名词短语。在这种组合框架中，名词短语是

* 本研究受国家教委人文社科规划资助，项目号 96JAQ740018。

作为动词的论元出现的。如果对论元内部的句法、语义特点没有充分的认识 and 了解，在对整个动词框架乃至整个句子处理分析时也会出现问题。所以，对名词短语内部的搭配的研究，即名词短语内部的句法、语义问题的研究，对计算机的汉语分析和处理来说是十分必要的。

我们曾对名词短语中的句法问题进行专题研究，对名词短语中中心语前面的定语进行了深入细致的考察，把定语分为“区别性定语”、“描写性定语”和“限定性定语”三类，并对各类定语的构成、性质、作用做了比较深入的分析。^[1] 这篇文章为计算机对名词短语的处理提供了句法方面的基础。

作为项目 96JAQ740018 的内容的一部分，我们对名词短语中的词语的语义类别进行了研究。本文所介绍的，就是这项研究的内容的一部分。

名词短语中包括中心语和定语，中心语可以是名词性的，也可以是谓词性的，定语可以由名词性的、谓词性的词语充当，也可以由名词性的、谓词性的短语充当。我们这里把研究范围限制在中心语是名词性的、定语由词语充当的名词短语。

本文考察以名词为中心语的名词短语里与中心语搭配的词语的语义问题，研究的目的是，考察现代汉语中以名词做中心语的名词短语里与作为中心语的名词搭配的词语的语义，具体说就是考察名词短语中做中心语的名词和做定语的词语的语义类别以及它们的语义类别与定中搭配的关系，从而为计算机对名词短语的处理提供语义方面的基础。

2 研究依据的材料和原则

我们以两本著作提供的材料作为研究的基础：《现代汉语辞海》和《同义词词林》。

《现代汉语辞海》，^[2] 是中国人民大学语言文字研究所多年进行汉语词语搭配研究的一个集成，该书对 7700 多个现代汉语词在各种句法结构里的搭配进行了较全面和详尽的描写。其中，名词的搭配框架主要内容如下：

名词 p/ny/n 〈名〉释义。

〔在前〕

①并列	(~+名)	(~+数量)	(~+动)
(~+名)	(~+动)	③定中	(~+形)
②主谓	(~+形)	(~+名)	④其他

〔在后〕

①并列	(动+~)	(名+~)	(数量+~)	(介宾+~)
(名+~)	(形+~)	(动+~)	(时间+~)	④其他
②动宾	③定中	(形+~)	(处所+~)	

这些框架里，〔在后〕③定中这一框架为我们提供了名词做中心语、前面带有定语，也就是名词短语的搭配材料。

《同义词词林》(简称《词林》)，^[3] 是现有的对汉语词语进行语义分类的著作中比较好、使用较广的著作，中文信息处理或计算语言学研究中，多以它作为词语语义分类的依据。

研究侧重点是名词短语搭配中的语义问题，亦即以搭配词语的语义为研究重点。考察两

个词语搭配，是考察它们语义上是否可以搭配，对句法上的形式及变化不加以讨论，也不涉及词语搭配中的语用问题和具体句子。由于“介宾+~”是短语与名词的搭配，所以不予以考虑。“数量+~”也是短语与名词的搭配，但可反映名词对量词的选择，所以予以保留。

3 研究的方法、步骤

本研究的主要材料都存放在计算机里，材料、数据的分类、检索等处理，我们都编制程序，让计算机运行程序完成。

整个研究按以下步骤进行。

1、检索名词

从《现代汉语辞海》里面检索出词性是名词(N)的词条(ENTRY)。例如，

CDA002 爱情 ài qín g 〈名〉男女相爱的感情。

CDQ099 趣味 qùwèi 1 〈名〉让人愉快、感到有意思或有吸引力的特性。

词条中，〈名〉表示该词条的词性是名词，“CDA002”表示该词条在《现代汉语辞海》中的顺序，是拼音以A打头的词条中的第002条。

2、检索搭配释例

从这些词条 ENTRY 中，检索出名词 N 做中心语的定中搭配释例 P，也就是“(在后)③定中”这一框架为我们提供的名词做中心语、前面带有定语的名词短语中的搭配材料。

例如，上述词条的搭配释例如下：

CDA002 爱情 ài qín g 〈名〉男女相爱的感情。

(名+~) 姑娘的~ 战士的~ 夫妻的~ 主人公的~ 老人的~; (动+~) 萌发的~ 遗忘的~ 倾诉的~ 珍藏的~ 描写的~ 骗取的~ 获得的~ 得到的~ (形+~) 幸福的~ 纯洁的~ 老式的~ 忠贞的~ 真挚的~ 真正的~ 健康的~ 甜蜜的~; 虚假的~ 荒唐的~ 虚伪的~ 庸俗的~ 低级的~; 动人的~ 美好的~ 迷人的~ 美妙的~ 奇特的~; 短暂的~ 永恒的~ 长久的~ 永久的~; 宝贵的~ 崇高的~ 高尚的~ 伟大的~ 难得的~ (数量+~) 一种~ (时间+~) 往日的~ 过去的~ 现在的~ 将来的~ 未来的~ (处所+~) 心中的~ 心底的~ 内心的~

CDQ099 趣味 qùwèi 1 〈名〉让人愉快、感到有意思或有吸引力的特性。

(名+~) 艺术~ 文学~ 语言~ 绘画~; 音乐~ 电影的~ 电视的~ 歌舞的~; 科学~ 数学~ 物理~ 化学~ 医学~ (动+~) 生活~ 工作~ 欣赏~ 运动~ 演唱~ 创作~ 写作~ 表现~ (形+~) 无穷的~ 浓厚的~ 高雅的~ 健康的~ 低级的~ 卑劣的~ 庸俗的~ 荒诞的~ (数量+~) 一种~ 一些~ 一点儿~

这些是以名词 N 词条“爱情、趣味”作为中心语的搭配释例。其中，(名+~)、(动+~)、(形+~)、(数量+~)、(时间+~)、(处所+~)等是名词、动词、形容词、数量词、时间词、处所词等作为定语与中心语搭配的框架，~代替该词条。

3、检索搭配词语

从 2 的结果中,也就是从 ENTRY 是名词 N 的搭配释例 P 中检索出与 N 定中搭配的词语 M。根据(名+~)、(动+~)等框架,可以确定这些词语的词性。

例如,与名词 N“爱情”、“趣味”定中搭配的词语 M 如下:

CDA002 爱情 à i q í n g 〈名〉男女相爱的感情。

(名)姑娘 战士 夫妻 主人公 老人 (动)萌发 遗忘 倾诉 珍藏 描写 骗取 获得 得到 (形)幸福 纯洁 老式 忠贞 真挚 真正 健康 甜蜜; 虚假 荒唐 虚伪 庸俗 低级; 动人 美好 迷人 美妙 奇特; 短暂 永恒 长久 永久; 宝贵 崇高 高尚 伟大 难得 (数量)一种 (时间)往日 过去 现在 将来 未来 (处所)心中 心底 内心

CDQ099 趣味 q ù w è i l 〈名〉让人愉快、感到有意思或有吸引力的特性。

(名)艺术 文学 语言 绘画; 音乐 电影的 电视的 歌舞; 科学 数学 物理 化学 医学 (动)生活 工作 欣赏 运动 演唱 创作 写作 表现 (形)无穷 浓厚 高雅 健康 低级 卑劣 庸俗 荒诞 (数量)一种 一些 一点儿

这分别是《汉代汉语辞海》中检索出来的与“爱情”、“趣味”定中搭配的词语,与 2 中的内容对照可以知道,名词 N“爱情”下面的“(名)姑娘”表示,名词“姑娘”可以与“爱情”搭配成为“姑娘的爱情”。

4、确定语义类别

在 3 结果的基础上,从《词林》的数据中检索出各词条 ENTRY 里的与名词 N 搭配的词语 M 的语义类别 L,并把各词语转换为相应的语义类别。例如,对词条“爱情”、“趣味”的检索结果如下。

Df05 爱情 à i q í n g 〈名〉男女相爱的感情。

(名):Ah06 Ae10 Ah08 Aj05 Ai02 (动)Ig01 Gb07 Hi12 Hj40 Hi13 骗取 Je12 Je12 (形)Da09 Ee43 Eb29 Ee06 Ee05 Ka30 Ed43 甜蜜 Ed01 Ee28 Ee05 Ee42 Ed06 Ed40 Ed23 Ed40 Ed03 Ed43 Eb25 Ed14 Eb24 Eb24 Ed27 Ee42 Ee42 Ed20 Ed27 (数量)Di04 (时间)Ca11 Ca11 Ca10 Ca12 Da07 (处所)Df02 Df02 Cb23

Df13 趣味 q ù w è i l 〈名〉让人愉快、感到有意思或有吸引力的特性。

(名)Db09 Dk25 Dk06 Hg19 Dk28 Dk29 电视 Hh04 Dk03 数学 Db06 Bm17 Dk03 (动)Di19 Di19 Hj47 Id14 Hh01 Hg16 Hg16 Di20 (形)Ed38 Ed16 高雅 Ed43 Ed06 Ee42 Ee42 Ee28 (数量)Di04 Dn05 Eb01

其中,形式为“Ah06”这样的串是《词林》对词语进行语义分类的类别符号,意思是第 A 大类的第 h 小类的第 06 个子类。作为词条的名词前这样的串是该名词在《词林》中的语义类别,如“趣味”前的“Df13”。有些词条前面(如“哀思”)没有语义类别的符号,是因为在《词林》里检索不到这些词语。在(名)、(动)等搭配框架中,与词条中各搭配释例中的词语 M 相对应的是它们在《词林》里的类别。例如,“爱情”词条中,“Ah06”与“姑娘”相对应,意思是与“爱情”搭配构成“姑娘的爱情”的“姑娘”这个词的语义类别是“Ah06”。有些词语《词林》里没有,则没有相应的表明类别的符号。

这些类别符号不仅说明名词 N 和与之搭配的词语 M 的语义类别 L,而且说明,与名词

N 搭配的有哪些语义类别。例如, 类别 Ah06 中的词语可以与名词“爱情”定中搭配。

5、分析搭配特点

分析各词条 ENTRY 里可以与名词 N 搭配的词语 M 的类别 L, 可以考察和归纳出各名词 N 做名词短语中心语时搭配上的特点, 探讨名词短语中的搭配规律。

4 研究的结果及分析

4.1 研究得到的数据

计算机运行程序对研究材料处理, 得到以下四种重要数据:

1、名词 N 做名词短语 P 中心语时搭配释例表。这是从《现代汉语辞海》中检索出的所有名词词条以及各词条做中心语的定中搭配释例(即第 3 节 1、2、3 讨论的内容)。

2、名词 N 做名词短语 P 中心语时搭配词语语义类别 L 表。这是从《现代汉语辞海》中检索出的所有名词词条搭配的词语及其语义类别, 列出了各词条下与名词中心语 N 搭配的词语 M 的语义类别 L(如第 3 节 4 讨论的内容)。

3、名词短语中与名词搭配的词语语义类别表。列出可与《现代汉语辞海》中的 2094 个名词词条定中搭配的 1274 个语义类别及各类中的词语。部分内容如下面所示。

Aa01——百姓; 大伙儿; 大众; 等; 干群; 公民; 公众; 国民; 货; 军民; 口; 劳资; 民众; 群众; 人口 人类; 人们; 人民; 人群; 人手; 人物; 人员; 师生; 师徒; 手; 愚民;

Ae01——办事员; 店员; 干部; 干事; 公务员; 雇员; 会计; 会计师; 秘书; 售货员; 售票员; 书记; 文书 营业员; 职工; 职员;

Ae10——兵; 工兵; 官兵; 海军; 尖兵; 将帅; 军长; 军官; 军人; 空军; 连长; 陆军; 民兵; 女兵; 炮兵 骑兵; 上校; 士兵; 团长; 卫兵; 营长; 元帅; 战士; 侦察员; 政委; 指导员; 指挥员;

Ca11——从前; 过去; 近年; 旧日; 旧时; 前面; 事前; 往年; 往日; 往昔; 昔; 昔日; 先前; 以前; 以往 原先; 早年; 最近;

4、名词短语中做定语的词语 M 的语义类别 L 表。此表列出可与《现代汉语辞海》中的名词词条定中搭配的 1274 个语义类别。部分内容如下面所示。

Aa01;Aa02;Aa03;Aa04;Aa05;Ab01;Ab02;Ab03;Ab04;Ac02;Ac03;Ad01;Ad02;Ad03;Ae01;Ae02;Ae03;Ae04;Ae05;Ae06;Ae07;Ae08;Ae09;Ae10;Ae11;Ae12;Ae13;Ae14;Ae15;Ae16;Ae17;Ae18;Af01;Af02;Af03;Af04;Af05;Af06;Af08;Af09;Af10;Af11;Ag01;Ag02;Ag03;Ag04;Ag08;Ag09;Ag10;Ah01;Ah02;Ah03;Ah04;Ah05;Ah06;Ah07;Ah08;Ah09;Ah10;Ah12;Ah14;Ah15;Ai01;Ai02;Ai03;Aj01;Aj02;Aj03;Aj04;Aj05;Aj06;Aj07;Aj08;Aj09;Aj10;Aj11;Aj12;Aj13;Aj14;Aj15;Aj16;Aj17;Ak01;Ak02;Ak03;Ak04;Ak05;Ak06;Ak09;Ak10;Ak11;Al01;Al02;Al03;Al04;Al05;Al06;Am01;Am02;Am03;An01;An02;An03;An04;An05;An06;An07;Ba01;Ba02;Ba04;Ba05;Ba06;Ba07;Ba08;Ba09;Bb01;Bb02;Bb04;Bb05;Bb06;Bc01;B

4.2 对所得结果的分析

名词短语中的搭配，在语义上是有规律可循的。例如，下面列出了六个名词以及与之定中搭配的词语的语义类别。

Da01 机密 jī mǐ 2. 〈名〉机密的事。

(名) Di02 党内 Dm01 军事 Di08 Da05 (动) If23 Je01 谈论 Je12 Hn03 Ee39 (形) Eb02 Jd04 Ee35 Ed28 Ed28 Ed28 Ed27 (数量) Eb02 Dn08 (时间) Ca11 Ca11 Ca10

Da01 秘密 mìmì 2. 〈名〉秘密的事情。

(名) Di02 Di10 军事 Di23 Eb02 Aa02 Ah06 Ah14 (动) Hi31 Fb01 If19 Jd08 Ee39 Hi39 Hi20 Hc10 Je12 Gb08 Hi14 Fc04 Hi24 (形) Ed28 Eb01 (数量) Eb02 Dn08 (时间) Ca11 Ca10 (处所) Bp26 Dk26

Aj03 伴侣 bàn lǚ 〈名〉同在一起生活、工作或旅行的人。

(名) Ca13 Ca13 Di19 Di19 Hg08 Aj17 Aj10 年青人 老年人 (动) Id18 Hi01 Ga08 Jd08 Gb06 Hj40 Hi07 Jc05 Hi16 Ga06 Hi05 Hf07 Jd08 Hh04 Hh04 (形) Eb30 Ed03 Eb28 Ee43 Ee43 Ee06 Ee05 Eb32 Ee42 Ee05 Ga01 Ga01 Ef05 合适 Ef05 Ef04 Ed04 Ed12 Je10 Jb03 Ed03 (数量) Eb02 Ed12 一大群 (时间) Ca10 Ca11 Ca11 (处所) Cb27 Di19

Aj17 朋友 péng yǒu 〈名〉彼此有交情的人。特指恋爱的对象。

(名) Ab03 Ca17 Ae07 Ae02 异性 Cb09 各国 各界 Aa01 Di02 Di02 Aj06 观众 (动) Kb02 Hi01 Ga08 来访 Hi12 Hi16 Hi16 Hi16 Id18 Hi03 Hi05 Hi08 Hi08 (形) Ah14 Eb35 Ka01 Ah08 Je10 Ka12 Ed51 Ka06 Ih05 唯一 Ka30 Ed32 Ed32 Ee33 Ee05 Ee05 Ee06 Df02 Ee33 Ed32 Ed50 Eb36 Eb36 Ed03 可敬 可信 Ed27 Ee14 Ee14 (数量) Eb02 Di17 Hc16 Dn08 (时间) Ca11 Ca23 (处所) Cb11 身边

Bg07 声音 shēng yīn 〈名〉由物体振动而发生的波通过听觉产生的印象。

(名) Ab02 Bi18 小孩 Dc03 Ae09 Bp13 Ak11 变压器 Bo21 Bo22 Hi47 Di02 Aa01 Aa01 Aa01 (动) Hi40 Ic01 Fc09 Fc10 Hb13 Hj48 Hh04 Hi12 Hh04 喝采 Ic01 Fc09 Ic01 Hi42 Fc10 Eb25 Hi13 问话 Hi40 Hc14 Fb01 Fb01 Fc08 Ic08 写字 (形) Eb16 Eb01 Ed41 Eb16 Eb19 Ed17 Eb16 Ed17 Ed17 很细 婉转 Eb17 Ed12 难忘 Eb20 Ed41 Ib19 Ed41 Eb13 Ib19 Ed32 Ee33 Ee08 Ga01 Ee05 Eb25 Eb13 Ef06 Eb23 Ga02 (数量) Eb02 Di04 Dd07 Ee09 Di10 Eb24 十分钟 (时间) Ca31 小时候 Ca12 (处所) 北京 Dm05 Cb02

Bg07 响声 xiǎng shēng 〈名〉声音。

(名) Bo21 Bo21 Bo22 Bo22 Bo11 Bo04 Bo27 Bo01 Bp36 Bp13 电话 Bp31 Dk27 Ab02 Bi01 Bf01 Bo04 海浪 Al04 Bm01 Bg01 Bp13 音叉 Bm04 Be05 大河 Be06 大街 Bn03 Dm07 Bn03 Cb28 Di09 (动) Hg08 伐木 铲土 Bk09 赶车 敲击 Fa01 Hh04 Hd04 Hd16 Hg06 Jd07 Jd07 Ih13 Fc05 Gb08 Ga13 Gb10 (形) Eb01 Ed19 Ed41 Ed15 Ga05 Eb17 Eb17 Ed46 Eb16 Ga02 (数量) Dd07 Ee09 (时间) Ca31 Ka12 (处所) Bn03 Cb07 大街上 Dm07

从可以与这些名词搭配的语义类别的分析、比较可以看出以下几点：

1、对于语义类别相同的名词 N，与它们搭配的词语 M 的语义类别 L 相同或基本相同。例如，在“机密”和“秘密”、“声音”和“响声”的相同搭配框架下出现的语义类别，很多是共同的。而且，不是共同出现的语义类别，几乎都可以互换。这说明，语义类别相同的名词，与它们搭配的语义类别基本相同。

2、对于语义类别相近的名词 N，与它们搭配的词语 M 的语义类别 L 有同有异。例如，“伴侣”和“朋友”的语义类别相近，相同搭配框架下的语义类别就有同有异。有同，是因为名词 N 语义上有共同的地方；有异，是因为名词 N 语义上有差异。

3、出现在一个名词搭配框架中的词语，是否也能与和这个名词语义类别相同的名次搭配，除了受语义类别的约束外，还受社会习惯、逻辑等其他因素的制约。例如，和“朋友”搭配的“异性”也可以和“伴侣”搭配，而“同性”（《现代汉语词典》释义为“性别相同”）可以和“伴侣”或“彼此有交情的人”意义的“朋友”搭配，但和“特指恋爱的对象”意义的“朋友”搭配，就要收到社会习惯的限制。

4、对于语义类别不同的名词 N，与它们搭配的词语 M 的语义类别 L 不同或者相差很远。例如，“机密”、“伴侣”、“声音”三个名词下的相同搭配框架中的语义类别很少有共同的，而且几乎不能互换。

前面的分析说明，词语的语义类别和名词短语中的搭配，是有直接关系的。它不仅能描写出一个名词作为名词性短语的中心语时，可以受哪些词语的修饰，而且更重要的是它能反映一种搭配的趋势。能够加以描写的搭配毕竟是有限的，而可能出现的搭配则是无限的。如果在词库中说明词语的搭配趋势，就能以简驭繁，处理时就可以举一反三，以有限应付无限。

当然，我们的研究，只是这种研究的初步尝试。由于条件和时间的原因，有不少问题需要进一步研究分析和改进，这主要表现在以下方面：

1、目前的语义分类还是很粗疏的，因而靠目前的语义分类所反映出来的名词短语中的搭配特点，也还是很粗疏的。随着汉语词语新的、更科学的语义分类成果的出现，词语搭配语义上的特点和规律将更明显、准确。

2、《现代汉语词海》中的搭配释例虽然十分丰富，但是仍然未能穷尽地列举出每个名词的每个搭配框架下能出现的所有词语，所以，目前每个名词的各搭配框架下列出的语义类别，还是不充分的。

3、对已经得到的研究结果的分析和研究，还只是初步的。

参考文献

- [1] 《三种定语、三类意义和三个槽位》，张卫国，刊于《中国人民大学学报》，1996年第4期。
[2] 《现代汉语辞海》，张卫国、冀小军主编，林杏光审定，人民中国出版社1994年6月第一版。
[3] 《同义词词林》，梅家驹等编著，上海辞书出版社1983年10月第一版。

(100872)