

# 信息处理用现代汉语词类体系

胡 明 扬

## 1 理论原则和具体要求

信息处理用的现代汉语词类的分类标准应该和母语教学用的现代汉语词类的分类标准保持一致，唯一的分类标准只能是句法功能。因为计算机处理自然语言只能根据词类序列，连同词类属性提供的信息，来判定句法结构，所以划分词类的标准只能是句法功能。句法功能主要是指句子成分功能，但是也包括在短语中的组合功能。在划分词类时，句子成分功能具有普遍性，但不具有排他性；短语组合功能，或所谓鉴定词和鉴定格式具有排他性，但不具有普遍性。这一点经过近年来的研究各家意见已经渐渐趋于一致，尽管在具体操作过程中还有不同意见。

信息处理用的现代汉语词类体系的分类肯定要比母语教学用的现代汉语词类体系的分类分得更细。因为掌握母语的人会自动补充很多必要的句法信息，而计算机却无法自动补充任何必要的语法信息，因此，分类分得细一些，实际上就是给计算机提供更多的句法信息，便于计算机自动进行句法分析。

信息处理用的现代汉语词类体系，不仅应该考虑尽可能满足计算机自动进行句法分析的需要，也应该考虑句法分析以后作出相应的语义解释的需要，此外，还应该尽可能为计算机辨认词库中没有的新词语创造条件。按理，信息处理用的现代汉语词类体系还应该考虑生成现代汉语语句的需要，但是鉴于目前的条件，本项目暂不考虑生成的需要。

词语的句法功能自有相应的语义基础，但是语义的判定有极大

的主观随意性，因而无法作为分类的标准，而且由于同样的理由，也无法作为判定词语的同一性的标准，更何况计算机根本无法像人那样自动辨认语义，所以当词语出现多功能现象时，分类的标准仍然只能是句法功能；任何偏离句法功能标准的做法势必破坏聚合关系和组合关系之间的有机联系，从而使词类和句法脱钩，最终使词类本身失去存在的意义。

信息处理用的现代汉语词类体系在某些具体问题的处理上可以和母语教学用的现代汉语词类体系有所不同，例如，固定短语（成语、俗语、谚语等）等都可以作为语项（lexical item）存入词库，如像“和/跟……一样/一般”“如/如果/假如……的话”，甚至“……分之……”等等都可以作为特殊的“词语”放在词库里。这些特殊成分一旦放在词库里，按理，也应该给出“词性”。这样做可能并不十分困难，例如，固定短语一般是一句句子，或者是名词性的、形容词性的、动词性的等等，可以分别标注相应的词性，“像……一样”等等可以注介词性的，“如果……的话”可以注连词性的，“……分之……”也许可以认为是数词性的。最后，一些近乎词缀的成分，如“非~，反~，超~”等等，“~性，~化，~度”等等可以作为特殊的“词类”来处理。

## 2 再分类讨论

### 2.1 名词

名词可以考虑把时间名词、处所名词、方位名词这三个附类分出去。这三类分出去以后沿用原名，还是称作“时间词”“处所词”“方位词”都可以，反正到了计算机里面都是代码。

这里“时间词”没有什么问题，“处所词”和“方位词”都有些小问题。表示处所的“处所名词”和“时间名词”不一样，不能直接用作状语，用作状语时必须在前面加介词，和/或后面加方位名词。这和一般名词没有什么不同，但是出现在主题位置上介词在一般情况下不出现，表示处所的专有名词后面和用作补语的“介词

+ 处所名词”后面不出现方位名词。这就和一般名词不太一样。不过，表示处所的名词，不加介词，又不跟方位名词，是不能用作状语的，所以从总的来看，和一般名词没有太大差别，似乎可以不予考虑。有一类表示处所的“不能单用的单音节‘字’+单音节方位名词”的单位应该说是“词”，如“国内”“国外”（甚至还有“国内外”）“校内（外）”等等，因为“国”和“校”等等平时都是不单用的。这一类“词”一般词典不收，看来应该补收。“国外”“校内”不在句首用作状语还需要在前面加介词，这也就和一般名词没有太大差别。但是，“我的笔记本放在桌子”绝对不行，必须说“放在桌子上”。“我的雨伞放在办公室/办公室里”似乎两可，而“他的孩子放在北京”完全正确，而相反，“放在北京+方位名词”绝对不行，可见表示处所的专有名词和可以兼表处所的单位名称，如“办公室”还是有些特别，这个问题跟“动+介”结构是不是算“词”也有关，很值得进行专题研究。

方位名词很明显可以分为两类。一类是不能单用的单音节方位名词，如“里”“外”“上”“下”等等，以及沿用文言的双音节“介词+处所名词”结构，如“以东”“之中”等等。这一类一般是不能单用的，只有在文言语体和近乎成语的单位中可以单用，如“上有老，下有小”“铁路横贯东西，以北是一片苇塘”等等。另一类是双音节的方位名词。这后一类可以单用，和一般名词相同，在某些情况下还可以直接用作状语，这就跟第一类不太一样。是不是可以把第一类和第二类分开？第一类可以称为“方位助词”（localizers）或者别的什么，第二类称为“方位名词”或“方位词”（locatives）。不过这个问题需要考察一下。

如果是这样，名词就可以分化为：

- |                               |               |
|-------------------------------|---------------|
| (1) 名词（含代名词）                  | 用作主语、宾语、定语    |
| (2) 时间词                       | 用作主语、宾语、定语、状语 |
| (3) 处所词                       | 用作主语、宾语、状语、定语 |
| (4) 方位词（含代方位词，<br>如“这里”“哪儿”等） | 用作主语、宾语、状语、定语 |

### (5) 方位助词

不能单用，似乎可以划归“助词”或称为“后置词”

如果处所词限于“单名+方位助词”，那就已经不是传统意义上的处所名词，这样就和方位词在句法功能上一致，仅仅是语义上不同，究竟怎么处理也要考虑，也许可以合并，成为“时间处所词”或“时处词”。

## 2.2 数词

数词中的序数词是通过加“第”来表示的，但是在很多场合往往不加，在口语中是通过直接“读数”来表示的，如“二楼”（不读“两楼”）“三楼”“么零九号”等等来和基数词区别开来的。这就是说，序数词是通过直接和名词结合这种方式来跟基数词相区别的，因为基数词是必须先和量词结合再修饰名词的。但是在书面语中如果不出现“第”字就没有区别，只能根据语境来区别是基数词还是序数词。因此就书面语而言究竟要不要，能不能分基数词和序数词需要进一步研究。序列词“甲乙丙丁”“子丑寅卯”等可以单列。

这样，数词可以分化为：

- (1) 数词            用作定语
- (2) 序列词        不作句子成分
- (3) 基数词
- (4) 序数词

## 2.3 量词

个体量词和其他名量词，如度量量词、集合量词等虽然性质不同，但是实际上都是单个儿的“词”，而且不单用，总跟数词和/或名词连用，所以可以不再细分。分数、小数、“人次”等复合量词可以单列。临时量词可以通过“数+量(+名)”结构规则去解决。

这样量词可以分化为：

- (1) 名量词

(2) 动量词

(3) 复合量词

## 2.4 形容词

现在的所谓性质形容词内容太杂。里面有一小部分表示时间、范围、方式等的形容词在语义上并不表示“性质”，在句法功能上可以同时用作状语，这部分形容词应该分出去，那样，既可以“正名”，也可以减少兼类。这部分形容词可以称为“情状形容词”或别的什么。还有一部分形容词不能直接修饰名词，必须加“的”才能修饰名词，如“美丽”等。如果不考虑生成问题，就可以不管；如果要考虑生成问题，这部分形容词要特殊处理。

非谓形容词分化出来可以确定不是谓语。唯谓形容词分化出来可以确定是谓语。这对计算机自动分析句法很有用处。

现在的所谓“状态形容词”在句法功能上和所谓“性质形容词”没有太大区别，也是可以用作谓语，加“的”可以用作定语（注意，现在所谓的“性质形容词”有一部分不能直接用作定语，必须加“的”才能用作定语），不同的只是不受程度副词的修饰，而这一点并不影响句法分析。现有的词类体系强调在短语组合中的特点，只是用“很”作为鉴定字来划分，不考虑句子成分功能，所以才这么分。就信息处理用的词类体系而言，所谓“状态形容词”可以并入性质形容词，作为性质形容词的派生形式（张斌就是这样处理的），在词库中只列出常用形式，另外规定派生的规则。但是，如果认为这些是独立的一类“词”，那就应该在词库里面作为词条一一列出。这一类派生的形容词究竟怎么处理还可以讨论。

这样，形容词可以分化为：

- |           |                      |
|-----------|----------------------|
| (1) 性质形容词 | 用作定语、谓语              |
| (2) 情状形容词 | 用作定语、谓语、状语（早、晚、快、慢等） |
| (3) 非谓形容词 | 用作定语                 |
| (4) 唯谓形容词 | 用作谓语                 |

## 2.5 动词

动词很重要，又很复杂。可以先把公认的附类分化出去，然后再来考虑其他问题。

助动词的问题是范围问题，究竟确定哪些是助动词，但是这一类该分出去没有问题。

趋向动词可以分立，但是其中的一部分，如“起来”“下去”等应和“动态助词”兼类；另一部分趋向动词跟动词结合得很紧，如“说开”“说来”等等应该整体作为一个“词”来处理；剩下的真正表示“趋向”的“动+趋”结构是否可以用语法规则去处理？

趋向动词也很复杂，可以分为3小类：(1)“动+趋”是一个词，(2)表示起始体、继续体等的动态助词（据房玉清统计，占全部用例的70%），(3)表示趋向，等于补语。这个问题需要研究。

从句法分析和语义解释来考虑，“联系动词”或“系词”应该分化出来。

从句法功能来考虑，还应该分出及物动词和不及物动词两类。这对于确定是否有省略和移位现象，以及采用配价理论和格语法都是不可或缺的前提。

及物动词还要不要再分下位小类，要看分了有多大好处。使令动词如果分化出来，有助于确定兼语结构，而双宾动词、小句宾语动词分化出来也有好处。

形式动词“加以、给予”等等后面接动词，跟所谓“谓宾动词”相同，但是目前所谓的“谓宾动词”中有很大一部分是“小句宾动词”。这个问题怎么处理要研究一下。

动词“有”（连同“有”字句）很有特点，要不要分化出来应该研究一下。“有+名”短语的用法往往和单个形容词相同，应该放在词库里，否则会带来很多麻烦。

“动词名用”有两类。一类是由句式决定的，近乎“引语”。另一类是动名词用法。这些可以由语法规则去解决。

这样，动词可以分化为：

- |             |                |
|-------------|----------------|
| (1) 不及物动词   | 用作谓语 不带宾语      |
| (2) 及物动词    | 用作谓语 带宾语       |
| (3) 使令动词    | 用作谓语 带兼语       |
| (4) 小句宾动词   | 用作谓语 带小句宾语     |
| (5) 联系动词或系词 | 用作名词谓语句的谓语动词   |
| (6) 助动词     | 用作谓语动词的动词短语的前项 |
| (7) 趋向动词    | 用作补语           |

(8) [谓宾 / 动宾动词, 形式动词] 用作谓语, 带动名词宾语补语问题研究得不够, 因而跟补语有关的分类问题也一时难以解决。心理动词是不是要分化出来要研究一下。

“动词 + 有”如“拥有”“占有”等等, 乃至“挂有”“写有”等等算什么? 离合动词怎么办? 这些都需要进一步研究。

## 2.6 代词

代词独立成为一类对母语教学而言比较方便。就印欧语而言, 代词是一个自成体系的封闭类, 但是就信息处理用的词类体系而言, 没有理由单立, 可以根据句法功能划归相应的词类。

## 2.7 副词

副词就句法功能而言是用作状语, 比较单纯。可是从构词和语义角度来看却十分复杂。

就信息处理而言, 可以只考虑句法功能, 而不考虑构词和语义问题。临时加“地”构成的副词可以由“~地”去处理。常用的带“~地”的副词应放在词库里, 如“历史地”“人为地”等等。由名词短语转化来的欧化副词, 如“大规模”“高速”等等也应该放在词库里。

程度副词有利于确定兼类词是否是形容词, 可以分化出来。语气副词表示命题以外的情态应该分化出来, 叫“情态副词”还是

“语气副词”倒是次要的，因为“语气词”也一样是表示命题以外的“情态”的。关联副词既用作状语，又用作连接词语，应该分化出来。

这样，副词可以分化为：

- (1) 普通副词            用作状语
- (2) 程度副词           用作状语，修饰性质形容词和心理动词
- (3) 语气副词           用作状语，表示情态
- (4) 关联副词           既在本句中用作状语，又在本句或跨句的篇章中起连接作用

## 2.8 介词

介词可以分为纯介词和动介词两类，这有利于一下子把介词跟可能是动词的分开。“从”“自”“于”等是纯介词；“在”“到”等是动介词。说“动介词”跟说兼类没有太大差别，特别是代码可能完全相同，如 *prep/v*。不过，分开纯介词和动介词对自动句法分析也许方便一些。这可以由计算机专家来决定。

如果要分，可分化为：

- (1) 介词
- (2) 动介词

## 2.9 连词

连词应该分化为并立连词和从属连词两类。这样分有利于分清独立分句和从属分句，而从属分句等于一个句子成分。近年来取消并立连词和从属连词的分界是不可取的。

长期以来“关联词语”一直逍遥法外，没人管，这一次应该“收编”，一一列入词库，并且标明是并列关联词语，还是从属关联词语。

这样，连词可以分化为：

- (1) 并立连词           连接独立分句和并立词语
- (2) 从属连词           连接从属分句

## 2.10 助词

助词的范围一直不清楚，这样就造成在整个语汇范围内有不少“散兵游勇”，这一次能不能清理一下。但是“散兵游勇”有多少还不清楚，要研究。

动态助词分出来有助于确定谓语动词，结构助词分出来有助于确定一些句法结构。

语气助词不参加句法结构可以分出来，这样有助于分清句子的界限。

这样，助词至少可以分化为：

(1) 动态助词

(2) 结构助词

助词是否一一单列？单标“助词”用处不大，不如单列，那样似乎更有利于句法分析。这当然还要看助词有多少，怎么处理更好。也许一部分助词可以标注小类，而一部分助词应该单列，不搞一刀切。

语气词是单列还是标小类还需要研究。

## 2.11 叹词

叹词不再分类。

## 2.12 象声词

象声词可以用作谓语，还用作相当于数量词的同位短语的前项，如“轰隆一声”等等，整个短语用作定语、状语。象声词可以不再分类。

# 3 尚待解决的问题

## 3.1 处所词的定义和范围

根据意义区分出来的传统的处所词能不能成立？像“校内、国

外”等等这样一些新定义的处所词能不能成立？处所词和时间词在句法功能上有没有区别，有哪些区别？“位移动词+表示处所的名词”=“动词+处所宾语”，凭这一点，传统的“处所名词”是不是需要保留？

### 3.2

“动+介”算“词”还是算“动词+介词短语”？

(1) 现代汉语：坐在了地上，走到了门口，等等。

(2) 文言：苦于，乐于，有利于，等等。

张金平的文章《也谈动词、形容词后“在、到、自”等的性质问题》（载《语言文化教学研究集刊》第二辑，北京，华语教学出版社，1998年，第260~268页）似乎已经解决了这个问题。

### 3.3

方位名词和方位助词的分合，分好还是合好？

### 3.4

就自动句法分析而言，基数词和序数词需要不需要分？

### 3.5

性质形容词大都可以直接修饰名词，有一部分性质形容词和所谓状态形容词不能直接修饰名词，必须加“的”以后才能修饰名词。这说明这部分形容词有较强的谓语句性，加“的”就是要去掉这种谓语句性（另一种说法是加“的”就成了“体词”性结构。两种说法角度不同，结论是相似的）。这样，加“的”不加“的”很不一样。吕叔湘着重谈到过这一点。任何一种词类的成员加“的”都能修饰名词，因此加“的”才能修饰名词的这部分性质形容词和所谓状态形容词应该说是“不能修饰名词”的，或者说是不能用作定语的，跟能直接修饰名词的形容词在句法功能上很不一样。加“的”不加“的”对形容词是不是要开特例，这个问题一直没解决，需要

研究。如果要考虑生成，不能直接修饰名词的那部分性质形容词就必须分出来。

### 3.6

所谓状态形容词如果认为形容词加“的”不加“的”没有区别，那么能用作定语和谓语，跟性质形容词没有区别；如果认为加“的”不加“的”是不同的结构，不同的句法关系，那么状态形容词只能用作谓语，不能用作定语，又跟唯谓形容词没有区别了。现在所谓状态形容词包括两类：一类是“碧绿、漆黑”等等，是凝固了的短语，另一类是“黑乎乎、亮晶晶、老实巴交、糊里糊涂”等等，是性质形容词的派生形式。怎么处理也需要再研究一下。

### 3.7

系词或联系动词的范围要确定下来。

### 3.8

“动词+趋向动词”有三种情况：

(1)“动+趋”是一个“词”，如“说开”跑去「除掉」”等等

(2)“动+趋”是“动+补”，后面的趋向动词实际上是修饰前面动词的成分

(3)“动+趋”是“动+动态助词”

这些不同情况也要分别处理。

### 3.9

小句宾语动词的范围，要和“谓宾动词”区分开，但允许兼类。

### 3.10

所谓“谓宾动词”的范围。“形式动词”也是一种谓宾动词。两者需要合，还是分？

### 3.11

“动 + 有”算什么，如“领有、占有、刻有”等等？

### 3.12

“有 + 名”如“有价值、有意义、有益、有利”等等？算一般短语，还是固定短语，还是“准词”？

### 3.13

使令动词的范围，应和兼语式联系起来考虑。

### 3.14

助动词的范围。

### 3.15

关联词语的范围和分类。

### 3.16

象声词的句法功能。

### 3.17

助词和语气词单列还是标小类？

# 信息处理用现代汉语词类的兼类问题

胡 明 扬

前人和时贤关于兼类问题的各种意见，我们在《兼类问题》（胡明扬，1996）一文中已经比较详尽地综述过，这里不再赘述。词类只能以句法功能或者说分布特征来区分，但是汉语没有足以区分词类的形态，因而词的多功能现象十分突出，给汉语词类的划分工作带来了几乎难以逾越的困难。为了解决这个难题，前人尝试过广泛随文转类的办法，结果是“类有定职，词无定类”。此路不通！从20世纪50年代起语法学界曾致力于“词有定类”的尝试，并且力求“定于一类”，但是由于多功能现象，同样有“词无定类”的危险。为了维护“词有定类”的目标，当时就想根据词义有无明显的变化来确定词的同源性，凡是词义有明显不同的就算是不同的词，那就不存在兼类问题，有人把这种观点概括为“词义不变，词性不变”。这当然也是一种办法，但是这么做就要扩大词的功能范围，而一旦扩大到无所不能，结果就是“词有定类，类无定职”，取消了词类在句法分析中的作用，也等于没有词类。所以张志公到80年代从另一个角度出发又认为汉语实词不能分类。看来此路也不通！那怎么办？剩下还有两条出路，兼类是一条出路，详尽的属性标注也是一条出路。属性标注的好处是比较灵活，不说兼类，而实际上把兼类的属性全部包括进来了，所以和兼类基本上内容相同，不过给出的第一个属性“定于一类”，计算机在扫描时，第一次得到的还是定死了一类，如“分发学习文件”中的“学习文件”这样的组合，“学习”定为“动词”，那么第一次扫描分析只能是“动宾关系”。当然，在定为动宾关系后通过全句和上下文的分析经过极其复杂的运算检验以后认为不符合原意，可以查另外的

属性，重新分析，可是每查一个属性都要反复进行排列组合的运算，还要一次次反复验证，而即使查到了“可以作定语”这一项，经过句法分析和语义解释，认为也是一种可能的句法结构，但是无法确定这样的分析就一定正确，而仅仅是一种可能的分析。这样大规模的反复运算，即使是计算机也难以承受。因此，1995年在北京语言大学的计算机工作站上用属性词典和配套的语法体系对军事语料中的一个句子进行分析，居然得出了一百多种可能的句法结构，就是得不出一种确定的句法结构，因此分析没有成功。兼类是另一条出路，注明兼类，第一次分析不成功，就可以改用兼类的词性进行分析，如“学习”注  $v/n$ ，那么在“动宾”结构无解的时候，就可以分析为“定语+名词”的结构，运算就简化了。从理论上来说，这两种办法都可以，可是初步实践，用兼类的办法似乎还省事一些，管用一些。可是很多人怕兼类太多就等于取消了分类，不同意采用兼类的办法。从逻辑上说，如果甲类词多数兼乙类词，甚至全部兼乙类词，那就根本不必分甲类和乙类，这是绝对正确的。所以最后就要看，如果允许兼类，在现有的词类体系中，特别是动名兼类、形名兼类的比例究竟是多少。有人认为不会少于三分之二，也就是 66% 左右，更有人估计高达 100%。如果的确是那样，兼类这条路就是死胡同，同样走不通。持这种观点的人当然坚决反对兼类，并且完全有正当的理由，因为如果兼类这么多，不但在教学上行不通，而且还会在理论上把整个词类体系搅混了，搅黄了。可是“认为”和“估计”不等于实际情况，所以还是有人主张走兼类的路子。到了 80 年代中期有人下了笨工夫对几万字的语料进行了统计。在那以后，又有人对几千个形容词和动词进行了统计，最多的统计了七千多个动词，而统计的结果惊人地接近，动名兼类都在百分之二十几，不超过百分之三十，形名兼类的比例更低。这样看来，兼类的路子还可以走一走看。因此到 90 年代大家的意见逐步接近，也就没人坚决反对兼类了。但是兼类这条路却不是容易走的，即使理论问题解决了，具体问题也会烦死人。特别是为信息处理用的词类肯定要比给母语教学用的词类分得细，而且还

要再分类，分小类，因为计算机没法自动提供额外的句法功能信息，分类分得细一些，提供的信息量就大一些；可是那样一来，不仅词类大类之间有兼类问题，小类之间也有兼类问题，要一一考察，一一标注，真要烦死人，而且这类苦差使、笨工夫现在很少人愿意去做。但是不做又不行！不仅如此，就口语而言，《水浒传》的“传”和“传球”的“传”读音不同，是两个根本不同的词，不存在兼类问题，可是书面文字中，两个“传”却是一种写法，因此就计算机处理汉字文献而言，“传<sub>1</sub>”和“传<sub>2</sub>”也存在一个兼类问题。这就是说“同形”的汉字也有兼类问题。当然，分类的标准确定了，用兼类的办法处理多功能的原则确定了，余下的工作再琐碎、再烦人，毕竟容易解决，最多就是多花一点时间和钱就是了。不过最终还得看上了机器的效果如何。主观的论辩不是客观的结论，最终还得由客观实践来检验，而计算机既没有先入之见，也没有偏见，最公正，只有通过这一关，才能最终弄清楚，兼类这种办法究竟能不能解决现代汉语词语突出的多功能的现象这个棘手问题。

#### 参考文献

胡明扬 1996 兼类问题，《词类问题考察》，胡明扬主编，北京语言学院出版社，北京。

# 及物动词与不及物动词分类考察

王俊毅

## 1 分类的意义

把动词分成及物动词和不及物动词两个类是对汉语动词的传统分类，是《马氏文通》以来几乎所有的汉语语法著作都要涉及的问题。但是由于它是一个“界限不清的类”（吕叔湘，1979），因而能不能这么分类、要不要这么分类，也一直是一个争议很大的问题。

王力（1944）基于汉语介词的缺乏，认为“及物不及物的分别，在中国语法里，并不是重要的”。<sup>①</sup> 陆志韦（1956）认为汉语的动词“不容易分‘及物’跟‘不及物’”，原因一是汉语动词不容易分清“主动”和“受动”，如“太太扶着‘丫头’”和“‘丫头’扶着太太”，“简直不知道谁扶着谁”；二是“形容词老是用得像动词似的”；三是有些动词“‘及物’只及到主人翁本身”。袁毓林（1998）指出把汉语的动词分为及物、不及物两类在语法描写上的作用“从目前的研究来看似乎并不明朗”。

那么汉语动词分及物不及物是否纯是为了照搬印欧语语法的模板呢？恐怕不能简单地这样讲。首先词类的依据或者说基础是意义，早期的语法学著作把动作的外射和内凝作为划分及物不及物的标准的确意义不大，但是这确实反映了两类动词的根本不同，无论是印欧语还是汉语，这种意义上的差别都是存在的，而意义上的差别必然会在语法上有所体现，这正是划分及物不及物的基础。

王力先生所说的“介词”是指像英语那样能置于动词之后介引名词成分的词，如 *arrived at Hangzhou* 中的 *at*。

其次，汉语动词在具体的句子中带不带宾语确实是比较自由，在具体语境中宾语常常不出现，但是不出现并不等于没有。例如：

A：你看电视了吗？

B：看了。

B的回答没有出现宾语，但是我们在理解这句话的时候没有宾语是不行的，我们会翻回头去在上文中或者在语境中找到它的宾语“电视”。这个过程在我们的理解过程中是不自觉的，然而在计算机理解自然语言时显得格外重要。我们必须告诉计算机“看”这个动词是个及物动词，如果它后面没有出现宾语，必须要在上下文中去找宾语，然后才能理解整个句子的意思。

第三，划分及物动词和不及物动词依据意义显得不那么可靠，依据功能也有种种麻烦，但是从总体上讲，有争议的毕竟是少数，大部分动词的归类还是意见一致的，这一点我们下文中再作讨论。

我们认为，把动词分为及物不及物两类是有意义的，尤其对于计算机处理自然语言，这一分类显得格外重要。划分及物不及物也是可行的，并不是分不清楚，至少大部分动词是能分清楚的。

## 2 研究概况

语法学家们划分及物和不及物的标准大体有三种情况：

### 2.1 意义标准

意义标准是早期汉语语法研究中使用的标准。如马建忠（1898）和黎锦熙（1924）。这一标准完全是从动词本身的意义出发的，如马建忠（1898）说：“一其动而仍止乎内也，曰内动字。一其动而直接乎外也，曰外动字。”黎锦熙（1924）对外动词的定义是“动作外射，及于他物”，内动词是“动作内凝，止乎自身”。这种单纯以意义为标准的分类，在操作上有困难，因为意义标准不好把握，分类结果也得不到形式上的验证，因而这样的分类标准实际用途不大。但是我们不能抹煞这种分类的意义，因为它从根源上讨论了两类动词的不同，讨论了存在及物与不及物这两种动词的语义