

# 名动词多能性指数研究及词类标记的组合应用

周姣美

北京师范大学国际中文教育学院/  
北京市海淀区新街口外大街19号

100875

zhoujiaomei@outlook.com

杨丽姣

北京师范大学国际中文教育学院/  
北京市海淀区新街口外大街19号

100875

yanglijiao@bnu.edu.cn

肖航

教育部语言文字应用研究所/  
北京市朝阳区门内南小街51号

100010

exiaohang@163.com

## 摘要

名动词是汉语词类研究及词性标注的难点问题。过去五六十年以来，有关名动词概念与体系类属、鉴定标准、标注方法等方面的争议不断，但基于语料库资源，以名动词的动态分布以及量化研究为支撑的研究较为缺乏。本文以现代汉语名动词作为主要考察对象，基于语言学理论方法的反思，将信息论与语料库方法相结合，引入香农-维纳指数作为量化指标，从多能性指数的研究视角对名动词进行考察，结合《信息处理用现代汉语词类标记规范》的修订研究，分析了名动词类别属性判断在现有印欧系语法词类体系框架下的困境，探讨了名动词跨类属性、词类标记的组合处理及其对于语料库建设、词典编纂等应用领域词类信息标注的探索意义。

**关键词：** 名动词；词语多能性指数；词类；标准修订

## A study of nominal verb polyfunctionality index and the combined application of POS tag

Jiaomei Zhou

Chinese Language and Culture College Chinese Language and Culture College

Beijing Normal University/

No. 19, Xijiekou Outer Street

100875, Beijing

zhoujiaomei@outlook.com

Lijiao Yang

Chinese Language and Culture College

Beijing Normal University/

No. 19, Xijiekou Outer Street

100875, Beijing

yanglijiao@bnu.edu.cn

Hang Xiao

Institute of Applied Linguistics Ministry of Education/

No.51, South Small Street, Chaoyangmennei

100010, Beijing

exiaohang@163.com

## Abstract

Nominal verbs are a difficult problem in Chinese part-of-speech (POS) research and tagging. In the past fifty to sixty years, there have been many controversies about the concept and system, identification criteria and annotation methods of nominal verbs, but there is a lack of studies based on corpus resources, supported by the dynamic distribution of nominal verbs and quantitative studies. This paper takes modern Chinese nominal verbs as the main object of investigation, combines information theory with corpus methods based on the reflection of linguistic theoretical methods, introduces the Shannon-Wiener index as a quantitative index, and examines nominal verbs

from the perspective of polyfunctionality index, combines with the study of Revision for Standard of POS Tag of Contemporary Chinese for CIP, analyzes the dilemma of nominal verb attribute judgment in the framework of the existing Indo-European grammatical POS system, and discusses the combined treatment of nominal verb co-category attributes and part-of-speech tags and their implications for the exploration of part-of-speech information tagging in the fields of corpus construction and lexicon compilation.

**Keywords:** nominal verb, word polyfunctionality index, part-of-speech, revision for standard of POS tag

## 1 引言

现代汉语中有些动词具有名词的语法性质, 尽管在词典中被标记为动词, 在具体语境中往往难以判断其究竟是名词还是动词, 如“发展”“研究”等词。这类词并非个例, 其语法功能与普通动词或普通名词相比均有较大区别, 而规模却有扩大的趋势。从二十世纪五六十年代至今, 学界对这一现象争议不断, 出现了名动词、动词次类、动名兼类词、动词名物化、动词名用、动名漂移、动名跨类等诸多不同的观点。理论语言学上的争议与计算机中文信息处理的词性标注难点密切相关, 尽管多个语料库及分词工具的词类标注体系均对这一词汇聚合规定了特定的标记, 但并不能很好地解决实际操作中标注一致性差的难题。在本文中, 为讨论方便, 仍然把这类词汇聚合称为名动词。

《信息处理用现代汉语词类标记规范》是现代汉语信息处理系统中的一个重要参考词类标记集, 它吸收了众多语言学研究成果, 从现代汉语信息处理的实际需求出发, 提供了一套统一的现代汉语词类标记体系。2020年发布的“《信息处理用现代汉语词类标记规范》修订稿”专门讨论了对名动词现象的处理, 提出以组合标记“n\_v”的形式标注名动词(杨丽姣等, 2021)。

本文以现代汉语名动词作为主要研究对象, 基于语言学理论方法的反思, 将信息论方法、语料库语言学方法相结合, 引入香农-维纳指数作为量化指标, 衡量汉语名动词语法功能的灵活程度, 从多能性指数这一新的视角对名动词传统难题进行考察。在此基础上, 进一步讨论了面向信息处理用的词性标记组合设置以及名动词组合标记的可操作性问题。

## 2 相关研究

关于名动词现象, 一直存在概念与体系类属、鉴定标准、标注方法等诸多争议。

黎锦熙(1960)提出动词名物化的说法, 他认为一部分用作主宾语的动词失去动词的特点, 获得了名词的语法特点, 包括可以受定语修饰、可以用名词或代词复指、可以和名词组成联合结构等。朱德熙(1961; 1982)反对上述名物化说法, 他认为动词作主语或宾语的时候仍然是动词。他提出了名动词的说法, 将名动词看作动词的一个小类, 定义为兼有名词性质的动词。陆俭明(1994)、俞士汶(2005)、陆丙甫(2009)赞同名动词的说法, 不主张将其处理为动名兼类词, 并对名动词的数量、语法特点作了进一步补充。胡明扬(1995)、郭锐(2002; 2011)、黄昌宁(2009)则认为这种现象可以处理成动名兼类, 吕叔湘(1979)提出了动词名用的说法, 认为这类现象中词义无明显变化, 语法特点有所改变, 与动名兼类不同。沈家煊(2007; 2009; 2012)认为对于名动词的词性区别应该淡化处理。在大的词类体系上, 可以认为名词包含了动词。夏全胜等(2014)从心理语言学和神经语言学视角分析名动词语义加工机制, 认为名动词与动词或名词的加工机制有所不同。

对于名动词的鉴定, 概括起来主要有四种标准: 语法功能标准、指称性标准、意义标准、数量标准, 其中以功能标准为主, 部分标准是否有效仍存在争议。朱德熙(1961; 1982; 1985)、胡明扬(1995)、俞士汶(2002)采用若干条功能标准判断名动词, 如可作动词“有”的宾语、可充任“进行、加以”这一类动词的准谓词性宾语、可受名量词修饰等。郭锐(2002)结合数量标准和功能标准判断名动词, 将这类词在所有动词中所占比例作为一个考量因素。陆丙甫(2009)根据名动词作宾语时的指称化程度判断其词性。安华林(2005)引入意义标准, 如“研究”这类词没有语义转类, 没有明确的名词性义项, 可与“检查”等有明确的名词性义项的词区分开。

对于名动词的标注, 朱德熙提出NV标记, 并进一步将NV细分为NV<sub>t</sub>和NV<sub>i</sub>两个小类, NV<sub>t</sub>表示及物的名动词, NV<sub>i</sub>表示不及物的名动词。俞士汶(2002)在语料库标注规范中

将名动词标注为vn，黄昌宁(2009)认为应根据名动词所在语法位置标注其词性，出现在述语位置上标作动词(v)，出现在名词的语法位置上标作名词(n)。杨丽姣等(2019; 2021)在《信息处理用现代汉语词类标记规范》修订方案提出了组合标记，用组合标记(n\_v)来标注名动词，并将这类词处理为动名跨类词。

### 3 语料标注与词表提取

#### 3.1 名动词标注语料库构建

本研究从国家语委现代汉语通用平衡语料库中选取600篇文本，包含散文、新闻、公文等多种语体，总规模达100万字。在对语料进行观察之后，本文认为单独为名动词设定组合标记存在其合理性。在实际使用时，名动词既有动词的功能，又有名词的功能，然而在词典中却没有名词的词性标记。标注语料时如果完全依据一个词的语境和功能进行标注，将引起词类系统的混乱，因为某些词的用法并未固定下来，因此将组合标记作为一种过渡用标记具有可行性和实用性。本文依照“《信息处理用现代汉语词类标记》修订稿”中的现代汉语词类标记设置，结合名动词的鉴定标准，组织语言学及应用语言学专业研究生10人，采用两两对照的方式对600篇文章中的名动词进行精标注，最终得到现代汉语精标注语料库。标注规则如下：

a.对于名物化的动词，应标注为名动词，若在句中没有被名物化，则仍按动词标注；根据名动词的定义，对于在句子中根据句法、语用功能都不能明确为指称(reference)还是述谓(predication)的词，则标注为名动词。

b.参考以下四条依据，符合其中任意一条标准则考虑该动词是否应标注为名动词：能受名量词修饰，充当动词“有”以及部分形式动词（“进行”、“作”、“加以”、“予以”、“给予”、“给以”等）的宾语，进入“有没有X”、“N的V”框架，充当体词性短语的中心语。

c.名动词的标注符号为n\_v。

以下为名动词的标注示例：

- 东方/l 各国/n 随着/p 社会/n 经济/n 的/u 发展1/n\_v, /w 纷纷/d 探索/v 东方/l 文化/n 作为/v 独立/a 的/u 文化/n 生存/n\_v 体/n\_g 的/u 可能性/n 问题/n, /w 本/rd 民族/n 文化/n 逐渐/d 成为/v 社会/n 发展2/n\_v 的/u 精神支柱/n。

- 我们/rp 一定/d 要/vu 根据/p 现在/t 的/u 有利/a 条件/n 加速/v 发展3/v 生产力/n。

发展1/n\_v：“社会经济的发展”为“N的V”结构，此处“发展1”被名物化；

发展2/n\_v：“社会发展”为“精神支柱”的定语，“发展2”是名词性短语的中心语；

发展3/v：“发展生产力”为动宾结构，此处“发展3”是动词；

《现代汉语词典》(第7版)中，“发展”词条如下：【发展】fāzhǎn动①事物由小到大、由简单到复杂、由低级到高级的变化：事态还在～，社会～规律。②扩大(组织、规模等)：～党的组织，～轻纺工业。③为扩大组织而吸收新的成员：～新党员，～工会会员。词典中“发展”只有动词词性，没有名词词性，但在实际使用中有关名词的语义和语法特点。在标注过程中，我们发现类似“发展”的情况广泛存在于语料中。

#### 3.2 名动词词表构建

首先提取语料中所有具有名动词标记(n\_v)的词语，去重处理后得到初步词表；然后参考《现代汉语词典》(第7版)中的词性标记，删去已确定具有名词词性的词；最后，人工对筛选出的词语进行增删。

在词性标注的基础上，本研究对语料中的名动词进行初步提取，用Python和正则表达式提取出所有具有n\_v、n和v标记的词，得到三个列表，去重并计数，取n\_v列表和v列表的交集，此交集再与n列表取差集，得到初步的名动词表，共包含1002个词。

初步提取的词表中，有780词的n\_v标记数量在3次及以下，在100万字的语料中出现频次过低，因此将这一部分词删去，剩余223个词；再查阅《现代汉语词典》(第7版)，删去“全球化”、“用地”、“收费”3个未收录的词，对照词典中的词性标记，保留没有名词词性的词语，删去“认识”、“贷款”等37个动名兼类词；最后，人工筛查剩余词语，删去“国有”、“航天”2个人工标注错误的词、“惊喜”、“恐惧”、“烫伤”3个没有动词词性的词，所得词表如下表1所示，共178个词。

发展	合作	展览	审美	休息	申请	交往	补充	启迪
研究	交接	治疗	死亡	预测	胜利	考察	成长	上市
服务	进步	刺激	探索	增长	实施	控制	促进	设立
医疗	提高	防雷	形成	转换	束缚	利用	发挥	探讨
学习	应用	观测	安排	表达	损害	流行	繁衍	体现
演出	重视	欢迎	护理	结合	体验	批评	繁殖	伪造
调查	出现	竞争	注意	伤害	威胁	破坏	分布	向往
发现	观察	理解	追求	生存	指导	认知	奉献	消化
消费	危害	使用	创造	宣传	制作	失败	革新	写作
实践	支持	处理	到来	转变	帮助	统治	关怀	信任
开发	关注	仲裁	发行	成功	保健	喜爱	观赏	选举
努力	污染	抽查	交流	出版	表意	想象	函授	循环
培训	准备	调整	介绍	诞生	尝试	预防	合成	延伸
笑	表现	发射	考验	独立	冲突	愈合	呼唤	运算
保护	冲击	分析	浪费	防治	搭配	运输	恢复	增加
表演	训练	改变	试验	分化	反射	运行	计算	直播
发生	进展	感悟	思考	解释	分裂	装扮	解剖	制约
管理	了解	革命	统一	经营	关心	综合	描述	贮存
改革	生长	记载	突破	精炼	监测	尊敬	配合	
生产	选择	旅游	享受	考试	检测	尊重	普及	

表 1: 现代汉语名动词表

## 4 名动词多能性指数计算方法

### 4.1 “香农-维纳”指数

Shannon和Weaver(1948; 1998)提出香农-维纳指数 (Shannon-Wiener index), 奠定了信息论的基础。香农-维纳指数也被称作香农物种多样性、香农熵、香农信息指数或H等, 用于度量信息的不确定性, 是基于现实世界的概率描述下的一种信息量的度量方式。香农-维纳指数的计算方法描述如下: 对于一种信号源A, 其发出的信号U有n种, 每种信号对应出现的概率 $p_i$ , 那么信号源A的不确定性为单个信号出现的不确定性的统计平均值, 该值就是A的香农-维纳指数。为了方便观察和计算, 需要对香农-维纳指数进行标准化, 计算公式如下所示:

$$H_{norm} = \frac{-\sum_{i=1}^n (p_i \cdot \ln p_i)}{\ln n} \quad (1)$$

标准化后的香农-维纳指数取值区间为[0, 1], H值越小, 不确定性越小, H值越大, 不确定性越大。

### 4.2 名动词多能性指数计算方法

#### 4.2.1 词语多能性

Hieber(2020; 2021)提出用香农-维纳指数来计算英语和Nuuchahnulth语<sup>1</sup>的词语多能性。词语多能性 (lexical polyfunctionality) 指一个词项 (lexical item) 具有不止一个话语功能或句法功能, 如述谓 (predication)、指称 (reference) 或修饰 (modification), 与之相对应的则是动词、名词、形容词。

如果一个词具有述谓、指称、修饰三个功能, 并且在使用时的频率完全相等, 那么可以称之为完全多能词 (perfectly polyfunctional lexical item), 例如, 一个词在某个语料库中出现300次, 三种功能各有100次; 与之相反的是, 如果这个词在某个语料库所出现的300次中, 均作指称 (reference) 用法, 则为完全单功能词 (perfectly monofunctional lexical item)。对于在使用中具有多种句法功能的词, 需要一个可以衡量一个词汇究竟有多“多能 (polyfunctional)”的度量指标。

<sup>1</sup>加拿大西部的一种早期语言。



#### 4.2.2 名动词的多能性

本研究3.2中提取的名动词在词典中大部分只有动词词性，小部分兼有动词和形容词词性，具有动词的典型语法特征。然而在实际使用中，基本都有名词的功能，体现出了跨类的特征。

在3.1的标注语料库中，本文分别统计出名动词作名词用法（标记为n\_v）和作动词（标记为v）用法的次数。表2为随机抽取的5个在《现代汉语词典》（第7版）中只有动词词性的名动词。按词频由高到低依次排列，“发展”是语料库中最高频的名动词，“合成”为低频名动词。5个词的n\_v标记所占比例平均为32.3%，接近一半的“研究”被标注为名词词性，“发展”和“服务”两个词也有35%以上的名动词标记，词频较低的“分析”和“合成”二词的名动词标记比例相对高频词较低，分别为17.9%和21.1%。在3.2的名动词表中，有10个词在《现代汉语词典》中标

	n_v数量	v数量	合计	n_v比例
发展	216	356	572	37.9%
研究	206	244	450	46.0%
服务	45	70	115	39.1%
分析	10	46	56	17.9%
合成	4	15	19	21.1%

表 2: 随机抽取的名动词词性标注情况

注有动词和形容词两种词性，这10个词分别为“努力、进步、应用、统一、成功、独立、精炼、失败、尊敬、尊重”。在语料中，这些词不仅有动词和形容词的功能，还有名词的功能，即动形兼类名动词。表3为这些动形兼类词的标注情况。词频最低的词是“精炼”，在语料中出现了8次，其中2次被标注为名动词，1次被标注为动词，5次被标注为形容词；词频最高的词是“成功”，在语料中出现了92次，其中27次被标注为名动词，12次被标注为动词，53次被标注为形容词。从三种标记所占比例来看，10个词的n\_v标记、a标记、v标记平均比例分别为27.3%、26.5%、46.2%。因此，在语料中这类词将近一半的用法仍为动词词性，名动词标记和形容词标记比例相当，前者略高于后者。

名动词	n_v数量	v数量	a数量	合计	n_v比例	a比例	v比例
努力	30	43	1	74	40.5%	1.4%	58.1%
进步	31	18	14	63	49.2%	22.2%	28.6%
应用	22	53	12	87	25.3%	13.8%	60.9%
统一	8	21	17	46	17.4%	37.0%	45.7%
成功	27	12	53	92	29.3%	57.6%	13.0%
独立	6	27	16	49	12.2%	32.7%	55.1%
精炼	2	1	5	8	25.0%	62.5%	12.5%
失败	6	14	1	21	28.6%	4.8%	66.7%
尊敬	5	9	7	21	23.8%	33.3%	42.9%
尊重	5	18	0	23	21.7%	0.0%	78.3%

表 3: 名动词表中动形兼类词的标注情况

#### 4.2.3 名动词多能性指数计算方法

本文基于Hieber(2020; 2021)的词语多能性指数计算方法，运用香农-维纳指数来衡量现代汉语名动词的多能性，具体计算方法描述如下：

一个名动词在语料中有可能出现三种标记：n\_v、v、a，对应该词语在不同语境中的不同词性：名动词、动词、形容词，每个词性标记的频率为：

$$p_i = \frac{n_i}{N} \quad (2)$$

其中，N为该词在语料库中的词频， $n_i$ 为该词在语料中被标注为某一词类*i*的数量。

可以计算出该词的香农-维纳指数 (H) 如下式, H的取值区间为[0, 1], H值越小, 不确定性越小, H值越大, 不确定性越大。换言之, 名动词作为一个具有名词性质的动词, 该词的H值越小, 它的多功能性就越小, H值越大, 它的多功能性就越大。

$$H_{norm} = \frac{-\sum_{i=1}^n (p_i \cdot \ln p_i)}{\ln n} \quad (3)$$

其中n为名动词在语料库中被标注的词类种数, 对于名动词而言, n=3。除了动形兼类词外的名动词没有a标记, 这一部分词的概率等于0, 然而ln0不存在。针对这种情况有两种常用方法可以解决 (Stefan Th. Gries, 2009): 一种是把三种标记的数量都加1, 另一种方法是简单地认为ln0等于0。本研究采用前一种方法, 在计算H时将除动形兼类词外的名动词的三种标记数量各加1, 例如, 表 2中的“发展”的n\_v、v、a标记实际数量为215、355、0, 经过处理后为216、356、1。

### 5 名动词多能性指数分析

名动词的多能性指数如下表 4所示, 根据词频由高到低排列178个名动词, 计算出的多能性指数值越大, 表示这个词语的多能性越强, 其取值区间为[0,1]。例如“发展”的多能性指数约为0.61, “形成”的多能性指数约为0.22, 前者的多能性比后者更强。观察n\_v和v标记数量, “发展”标注为n\_v和v的次数分别为216次和356次, “形成”则是10次和186次, 在数量和比例上来看, “发展”比“形成”更灵活更多能, 与多能性指数得出的结论相符。在名动词表中, 有一部分

序号	词语	n_v数量	v数量	a数量	多能性指数 (H)
1	发展	216	356	0	0.614171
2	研究	206	244	0	0.64062
3	发现	31	265	0	0.329753
4	学习	39	222	0	0.409249
5	出现	16	225	0	0.254059
6	使用	12	207	0	0.228963
7	发生	19	199	0	0.30238
8	形成	10	186	0	0.223093
9	生产	18	167	0	0.32769
.....					
169	繁衍	4	6	0	0.81752
170	分化	6	3	0	0.808014
171	表意	6	3	0	0.808014
172	关怀	4	4	0	0.850864
173	函授	4	4	0	0.850864
174	束缚	6	2	0	0.782776
175	信任	4	3	0	0.858673
176	直播	4	3	0	0.858673
177	启迪	4	2	0	0.852792
178	精炼	2	1	5	0.819448

表 4: 名动词多能性指数

词有动词、形容词、名词三种词性的功能, 即动形兼类名动词。下表 5为10个动形兼类名动词的多能性指数, 由表可知, 这些词的H值均在0.6以上, 其中“尊敬”一词的H值高达0.975, 表明这是一个功能相当灵活的词语。从多能性指数的平均值来看 (表 6), 动形兼类词的H值平均为0.825, 附录中所有名动词的多能性指数平均值为0.585, 比前者低0.24。因此, 动形兼类词的平均多能性比名动词强。

为更直观地观察多能性指数 (H) 在每个区间的分布情况, 本文将计算所得的H值按照区间来统计其数量, 如在区间[0,0.1]的H有0个, 在区间(0.1,0.2]的H有2个, 依次类推, 将所有名动词与动形兼类的名动词的多能性指数分布情况统计如图 1和图 2所示, 横坐标为10个多能

序号	词语	n_v数量	v数量	a数量	多能性指数 (H)
1	努力	30	43	1	0.673248
2	进步	31	18	14	0.947666
3	应用	22	53	12	0.840001
4	统一	8	21	17	0.937593
5	成功	27	12	53	0.85852
6	独立	6	27	16	0.865649
7	精炼	2	1	5	0.819448
8	失败	6	14	1	0.703815
9	尊敬	5	9	7	0.974883
10	尊重	5	18	0	0.630712

表 5: 动形兼类名动词的多能性指数

	平均多能性指数
所有名动词	0.585
动形兼类的名动词	0.825

表 6: 词语多能性指数的平均值

性指数的区间, 如0.1对应区间(0,0.1], 1对应区间(0.9,1], 纵坐标为在该区间H的数量。名动词作为一个具有名词性质的动词, 该词的H值越小, 它的多功能性就越小, H值越大, 它的多功能性就越大。一般而言, 名动词的H值越大, 就说明其名词性程度越深。例如“利用”的H值约为0.16, “形成”的H值约为0.22, “生产”的H值约为0.33, “管理”的H值约为0.40, “演出”的H值约为0.59, “服务”的H值约为0.65, “实践”的H值约为0.70, “胜利”的H值约为0.81, 以上8个词的H值分别位于8个不同的区间, 它们的名词性逐渐增加, 多能性越来越强。

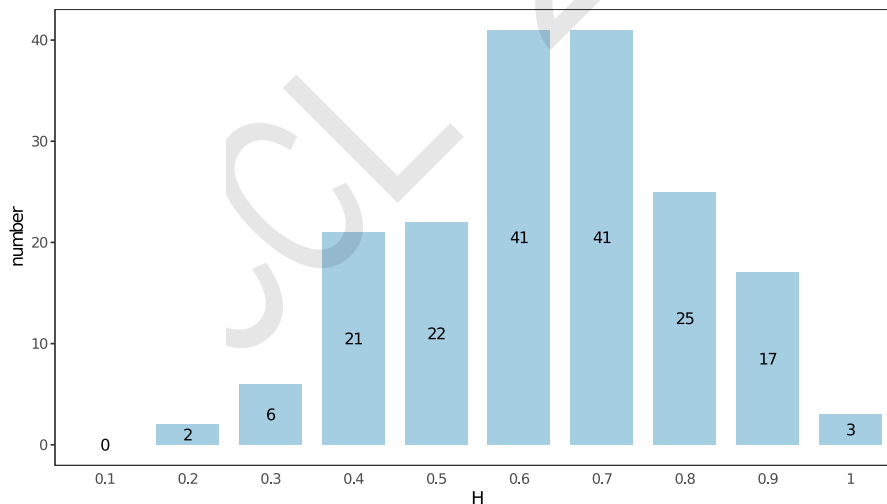


图 1: 名动词的多能性指数 (H) 数量分布

从图 1总体的分布趋势来看, H值在0.6以前时, 名动词的数量随着H的增大而增加, H值到0.6、0.7时数量最多, 0.7之后数量又逐渐减少; 从数据的峰值来看, 名动词的数量在(0.5,0.7)这个H值区间内最多, 低于0.3或高于0.9的名动词数量都比较少, 即多能性在中等偏高水平的名动词最多, 多能性较差或较好的名动词相对更少。因此, 名动词的总体多能性比较强, 功能特别灵活和功能特别不灵活的名动词较少。

如图 2所示, 动形兼类名动词的多能性指数数量分布与图 1不同, 集中在强多能性区域。没有一个动形兼类名动词的H值低于0.6, H值最多分布在0.8及以上, 这与表 6中计算的多能性

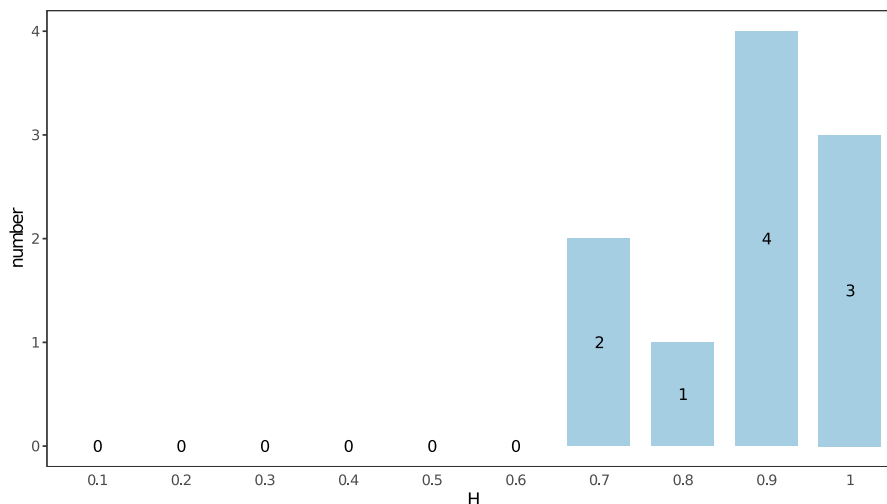


图 2: 动形兼类名动词的多能性指数 (H) 数量分布

指数平均值特点一致, 动形兼类的名动词比普通名动词的功能更灵活, 这主要是因为动形兼类名动词有动词、形容词和名词三种功能, 所以这类词的多能性总体更强。

多能性指数值越大, 表示该词语的多能性越强, 人们在不同语境中使用该词语的不同词性就越灵活。结合多能性指数对词表中的名动词进行分析, 本文发现这些名动词的性质存在差异, 某些名动词的多能性更强, 某些名动词的多能性较弱, 多能性指数可以较好地反映名动词在语料中的使用情况。此外, 多能性指数也可为名动词的辨别和处理提供参考。

## 6 临时活用、兼类、跨类的多能性指数对比分析

多能性指数可应用于其他现象的分析, 如临时活用和兼类词等, 为词类研究提供新视角。

《现代汉语词典》(第7版) 中存在大量动名兼类词, 即在词典中标注有动词和名词两种词性的词, 如“生活”、“工作”等。本小节计算前文精标注语料库动名兼类词n标记和v标记的数量, 根据4.2中的词语多能性计算方法, 计算语料中88个动名兼类词的多能性指数。表 7为词频最高的十个动名兼类词的多能性指数, H值最低约为0.07, 最高最约0.64。

动名兼类词	n数量	v数量	合计	多能性指数
生活	338	55	393	0.383708
工作	242	88	330	0.544968
教育	72	173	245	0.573052
作用	217	2	219	0.073539
要求	66	102	168	0.63926
关系	153	13	166	0.281812
活动	103	39	142	0.594252
组织	92	50	142	0.624316
领导	92	16	108	0.425816
运动	77	9	86	0.358718

表 7: 动名兼类词的多能性指数

从平均多能性指数来看, 名动词与动名兼类词分别约为0.585和0.584, 动名兼类词总体上没有比名动词多能性更强。此外, 两个词表都存在多能性很弱的词, 例如动名兼类词“作用”的H值为0.07, 语料库中标记为n的有217个, 标记为v的仅有2个。名动词“利用”的H值为0.16, 语料库中标注为v的有174个, 而标注为n.v的仅有5个。因此“作用”的动词用法和“利用”的动词用法可看作是一种临时活用现象。



对于那些多能性比较弱的名动词，如“增加”、“利用”等，我们认为不应看作是典型的名动词，即不应处理为动名跨类词；对于多能性较强的动名兼类词，如“工作”、“活动”、“报告”等，其性质与名动词接近，可将其处理为跨类词。

## 7 总结及讨论

### 7.1 总结

名动词是汉语词类研究及词性标注的难点问题，目前基于语料库资源，以名动词的动态分布以及量化研究为支撑的研究较为缺乏。本文构建了一个规模较大的名动词词表，将信息论与语料库方法相结合，引入香农-维纳指数作为量化指标，从多能性指数的研究视角对名动词进行考察，并探讨动词的临时活用、兼类、跨类情况。

首先，计算每个名动词的三种标记的概率分布，从而判断该词语的用法情况。研究发现，大部分名动词仍然以动词用法为主，小部分名动词的用法已经偏向于名词词性。然而，仍缺乏一个统一的标准来衡量这些词语究竟有多灵活，即多功能性有多强。因此，本文选用基于“香农-维纳”指数的词语多能性指数来衡量名动词的多能性，多能性指数值越大（最高值为1），表示该词多能性越强，功能越灵活；多能性指数越小（最小值为0），表示该词多能性越弱，功能越不灵活。根据词语多能性指数计算结果，名动词的平均多能性指数为0.585，多能性较强。动形兼类的名动词属于名动词中功能更灵活的一类，平均多能性指数为0.825，在语境中兼有动词、名词、形容词的不同功能。

除名动词外，本文运用多能性指数对动词的临时活用、兼类、跨类情况进行探讨，对于那些多能性比较弱的名动词，如“增加”、“利用”等，不应将其看作是典型的名动词，即不应处理为动名跨类词；对于多能性较强的动名兼类词，如“工作”、“活动”、“报告”等，其性质与名动词接近，可将其处理为跨类词。

### 7.2 词类范畴、词类标记的组合应用及可操作性问题

从根本上说，名动词理论上的莫衷一是、词性标注上的困境与汉语自身的特点及汉语词类分析所依凭的印欧系语法框架有关。

沈家煊(2012)认为，“名动词”对汉语来说某种程度上是个伪概念，朱德熙先生因英语“V-ing”与名动词对当而引入的这一特定概念，从大量汉语的事实出发是难以成立的。更严重的是，名动词现象对于现代汉语词类体系的内部一致性构成了挑战。

汉语词类问题早在上世纪五十年代就有过几次大的讨论，核心成果是认为词类主要是依据语法功能原则进行划分的，这已经成为学界的基本共识。然而在印欧语系语法中动词与名词这一基本的对立性范畴，在汉语这种非形态语言中始终存在种种扭曲，从语义表达功能上看，名词的典型功能在于指称，动词的典型功能在于陈述，可是在具体语境中，名动词可视为指称性的抽象陈述，或者既表陈述也表指称，在现有词类概念体系或印欧系语法话语体系下，名动词在应用层面的混淆与歧义，恐怕难以消除。以《现代汉语词典》（第7版）为例，在词性上标为动词，而配比的例句中表示指称性意义的例句比比皆是。如，爱戴：动词。敬爱并且拥护：受到人民群众的爱戴。

杨丽姣、肖航等(2021)认为名动词不是兼类词，语境信息不能完全消除歧义；不是名词或动词次类，其语法语义功能与其有较大的差别。在“《信息处理用现代汉语词类标记规范》修订稿”中，杨丽姣、肖航等提出用组合标记n.v来标注名动词，实际上将这类词处理为跨类词或者是共类词，n.v是为提升汉语语言信息处理用词类标记体系的内部一致性，在处理名动词这类特殊语言现象时一种阶段性尝试。名动词问题的根本解决需要突破现有理论语言学框架，一旦系统性的汉语理论语言学话语范式得到确立和被广泛采用，汉语词类体系以及词类概念必然进一步推陈出新。

多能性指数研究可为汉语词类研究提供新视角，为词类标记的组合应用提供量化指标。名动词多能性指数体现了词语功能的灵活性，在此基础上探讨名动词跨类属性和词类标记的组合处理具有可操作性，也有助于语料库和词表资源建设、词典编纂等应用领域的研究。

## 致谢

本文受到国家语委“十三五”科研规划2020年度省部级重点项目“面向国际中文教育的文本可读性智能评价方法研究及分析系统构建”（ZDI135-41）的资助。

## 参考文献

- Daniel W. Hieber. 2020. *Lexical flexibility: Expanding the empirical coverage*.  
<https://files.danielhieber.com/publications/ucsb-doctoral-colloquium/slides.pdf>.
- Daniel W. Hieber. 2021. *Lexical polyfunctionality in discourse: A quantitative corpus-based approach*.  
University of California, Santa Barbara.
- Stefan Th. Gries. 2009. *Statistics for Linguistics with R: A Practical Introduction*. De Gruyter Mouton,  
Berlin, New York.
- Shannon C. E. 1948. A mathematical theory of communication. *The Bell system technical journal*, 27(3):  
379-423.
- Shannon, C. E. and Weaver, W. 1998. *The mathematical theory of communication reprinted*. Urbana:  
University of Illinois Press.
2006. GB/T 20532-2006 信息处理用现代汉语词类标记规范. 中国标准出版社, 北京.
- 安华林. 2005. 从两种词表看名,动,形兼类的处理. *语言教学与研究*, (4): 31-39.
- 郭锐. 2002. 《现代汉语词类研究》. 商务印书馆, 北京.
- 郭锐. 2011. 朱德熙先生的汉语词类研究. *汉语学习*, (5): 14-26.
- 胡明扬. 1995. 动名兼类的计量考察. *语言研究*, (2): 91-99.
- 胡明扬. 1996. 词类问题考察. 北京语言学院出版社, 北京.
- 黄昌宁, 李玉梅, 靳光瑾. 2009. 动名兼类词及其词性标注规则. 张普、王铁琨主编《中国语言资源论丛》  
(一), 商务印书馆, 北京.
- 黎锦熙, 刘世儒. 1960. 语法再研讨——词类区分和名词问题. *中国语文*, (1): 5-8.
- 黎锦熙. 1992. 新著国语文法. 商务印书馆, 北京.
- 陆俭明. 1994. 关于词的兼类问题. *中国语文*, (1): 28-34.
- 陆丙甫. 2009. 基于宾语指称性强弱的及物动词分类. *外国语(上海外国语大学学报)*, (6): 18-26.
- 吕叔湘. 1979. 汉语语法分析问题. 商务印书馆, 北京.
- 沈家煊. 2007. 汉语里的名词和动词. *汉藏语学报*, (1): 27-47.
- 沈家煊. 2009. 我看汉语的词类. *语言科学*, 8(1): 1-12.
- 沈家煊. 2012. “名动词”的反思: 问题和对策. *世界汉语教学*, (1): 3-17.
- 夏全胜, 彭刚, 石锋. 2014. 汉语名词, 动词和动名兼类词语义加工的偏侧化现象——来自ERP的研究. *心理科学*, 37(6): 1333.
- 杨丽姣, 肖航, 刘智颖. 2019. 《信息处理用现代汉语词类标记规范》修订方案. *语言文字应用*, (3): 87-95.
- 杨丽姣, 肖航, 刘智颖. 2021. 《信息处理用现代汉语词类标记规范》修订研究. *语言文字应用*, 119(3):  
111-120.
- 俞士汶, 段慧明, 朱学锋. 2002. 北京大学现代汉语语料库基本加工规范. *中文信息学报*, 16(5): 51-66.
- 俞士汶, 段慧明, 朱学峰. 2005. 词语兼类暨动词向名词飘移现象的计量分析. 孙茂松、陈群秀主编《自然  
语言理解与大规模内容计算》, 清华大学出版社, 北京.
- 中国社会科学院语言研究所词典室. 2005. 现代汉语词典. 第7版. 商务印书馆, 北京.
- 朱德熙, 甲文, 马真. 1961. 关于动词形容词“名物化”的问题. *北京大学学报: 哲学社会科学版*, (4): 53-66.
- 朱德熙. 1982. 《语法讲义》. 商务印书馆, 北京.
- 朱德熙. 1985. 现代书面汉语里的虚化动词和名动词为第一届国际汉语教学讨论会而作. *北京大学学报:  
哲学社会科学版*, 世界汉语教学学会, (5): 3-8.