

# 词典中“操作”的词类标注问题

## ——一项基于语料库的个案研究

郝晨, 唐菊, 孟波

南开大学汉语言文化学院, 天津

收稿日期: 2023年9月7日; 录用日期: 2023年10月16日; 发布日期: 2023年10月25日

### 摘要

词类标注是词典的一项重要任务, 词类标注的准确性直接影响词典的编纂质量。典型名词和典型动词的标注较容易进行, 但对于兼类词的标注争议较大, 兼类词的词类标注一直是词典编纂的难点。本研究以语料库作为研究工具, 选取词类识别有争议的“操作”一词进行个案研究, 考察现有词典对“操作”一词的释义情况, 初步判断其词类。同时结合语料库中“操作”一词的使用模式, 通过搭配和类联接等方法进一步确定“操作”一词的词性处理, 试图对“操作”一词进行准确的词类标注。研究发现“操作”在现代汉语中应同时属于名词、动词, 词典应将其处理为名动兼类词。

### 关键词

语料库, 词类标注, 兼类词, 类联接, “操作”

# The Problem of Part-of-Speech Labeling for “Cao Zuo” in Dictionaries

## —A Case Study Based on Corpus

Chen Hao, Ju Tang, Bo Meng

College of Chinese Language and Culture, Nankai University, Tianjin

Received: Sep. 7<sup>th</sup>, 2023; accepted: Oct. 16<sup>th</sup>, 2023; published: Oct. 25<sup>th</sup>, 2023

### Abstract

Part-of-speech labeling is an important task in dictionaries, and the accuracy of part-of-speech labeling directly affects the quality of dictionary compilation. The labeling of typical nouns and verbs is relatively easy, but there is a significant controversy over the labeling of concurrent

words. The labeling of concurrent words has always been a difficulty in dictionary compilation. This study uses a corpus as a research tool and selects the controversial word “Cao Zuo” for word class recognition as a case study. It examines the interpretation of the word “Cao Zuo” in existing dictionaries and preliminarily determines its word class. At the same time, combining the usage patterns of the word “Cao Zuo” in the corpus, further determining the part-of-speech processing of the word “Cao Zuo” through methods such as collocation and class connections, attempting to accurately label the part of speech of the word “Cao Zuo”. Research has found that “Cao Zuo” should belong to both nouns and verbs in Chinese, and dictionaries should treat it as a noun-verb hybrid word.

## Keywords

Corpus, Part-of-Speech Labeling, Concurrent Words, Class Connections, “Cao Zuo”

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

词类是词典学和词汇语义学的要素[1], 对词类进行标注是词典学最重要的任务之一[2] [3]。词类标注的准确性反映了词典的可信度。在现代汉语中, 词类划分一直是语法研究中的难题[4]。汉语中存在大量的兼类词, 兼类词指“在一定的词类系统中兼属两个或两个以上不同的词类的词”[5]。在词典的词类标注中, 典型名词和典型动词的标注较容易进行, 但对于动名兼类词的标注争议较大。现有三部大型汉语词典对“操作”一词的词类标注和释义存在矛盾, 虽然给出了“操作”一词的名词用法, 但仍将“操作”一词标记为动词, 且“操作”一词在日常生活中使用频率较高, 有必要对“操作”一词做出准确的词类判断。本研究结合语料库的使用模式考察“操作”一词的词类, 具体考察了国家语委现代汉语语料库(以下简称为国家语委语料库)和北京语言大学 BCC 语料库(以下简称为 BCC 语料库), 以期对“操作”一词的词类标注做出补充。

## 2. 比较三部汉语词典对“操作”一词的词类标注和释义

选取《现代汉语词典(第七版)》、《新编汉语多功能词典》和《现代汉语规范词典》三部词典进行检索, 调查分析“操作”一词的词类标注和释义。现将三部汉语词典中“操作”一词的释义情况列表如下, 具体见表 1。从表 1 可以发现, 三部汉语词典均对“操作”一词明确标记为动词。其中《新编汉语多功能词典》和《现代汉语规范词典》给出了“操作”一词的名词用法, 如“作定语”、“带上定语活用为名词, 作主语”、“作宾语”, 这些均为典型名词的语法功能, 但三部词典将“操作”一词以上的名词用法视为词类活用, 并未标记其名词词性和义项。

由于汉语缺乏形态变化, 传统的汉语语言学在判定一个词是兼类词还是词类活用上较为棘手, 依据主观判断往往较难定夺, 以致对于同一个词形在使用中是兼类词还是词性活用, 各家说法不一。黄伯荣、廖序东[6]指出词类活用的标准: “词的活用是一种特殊的用法, 在特定的条件下, 为了表达上的需要, 临时借来一用。”曾德祥[7]综合了各种教材和研究的意见, 提到关于“词类活用”和“兼类”区分的标准: 词类活用最显著的是它的临时性, 即它只是在特定的语言环境中临时取得了另一个词的语法功能而被用作另一类词; 词的兼类则是某个词在长期的运用过程中因语言习惯而取得了不止一类词的语法功能, 换言之, 词的兼类具有约定俗成的稳定性。综上所述, 区分“活用”和“兼类”, 关键在于确定“临时

性”和“约定俗成的稳定性”。

“操作”一词确有名词用法，但三部词典均未对“操作”设立名词义项，其名词用法常常被处理为词类活用。“操作”一词的名词用法是已经形成规约化，还是仅仅只是作暂时的词类活用？“操作”一词是否属于兼类词？词典中对“操作”一词的词类标注是否合理？为了准确地回答以上问题，应当掌握充分的语料证据，调查“操作”一词在语料库中的实际使用情况。语料库可以提供真实的语言素材，可以利用大规模语料建设搭配库，从大数据的角度考察语言现象[8]。本研究使用语料库考察“操作”一词的词类标注及真实语料环境，总结“操作”一词在真实语料中的词类用法，得到更为全面和客观的结果。

**Table 1.** Comparison of part-of-speech labeling of the word “Cao Zuo” in Chinese dictionaries

**表 1.** “操作”一词在汉语词典中的词类标注对比

词典	释义及配例
《现代汉语词典(第七版)》	【操作】<动>按照一定的程序和技术要求进行活动或工作。 例：~方法 ~规程。
《新编汉语多功能词典》	【操作】<动>按照一定的程序和技术要求进行活动操作。 ① 作带宾、带补谓语。 例：他能熟练地~新机器了。 你们厂的工人~得真不错。 她边干边琢磨，不久便能独立~了。 ② 作定语，或构成词组作定语。 例：~规程。 每个企业都需要有一批~技能熟练的工人。 他们用手工~方法分析氨基酸排列顺序。 ③ 带上定语活用为名词，作主语。 例：豆腐新制法的~很简便。 现在“工业机器人”的~，可不受化学和气温等限制。 ④ 作宾语。 例：有的大学生到车间不久，就学会了~。 学习后期可参加实际维修~。
《现代汉语规范词典》	【操作】<动>按照一定的程序和技术要求进行活动。 ① 作定语：~技术。 ② 作宾语：学习~。 ③ 作谓语：~机器。

### 3. “操作”一词在语料库中的表现

#### 3.1. 国家语委语料库

我们使用国家语委语料库检索了“操作”一词的词类标注和使用情况，结果发现 559 条与“操作”有关的语料，得到表 2 所示的词类分布情况。名词用法共有 13 条，动词用法共有 546 条。

**Table 2.** The distribution and frequency of part-of-speech in the corpus of the National Language Commission for “Cao Zuo”

**表 2.** “操作”在国家语委语料库中的词类分布及频率

词类	频次	频率
动词	546	0.9764
名词	13	0.0232

《现代汉语》中名词的主要语法特征如下：

1) 经常作主语和宾语，例如：牛吃草；多数能作定语和带定语，例如：柳树梢头、河边柳树。不能作补语。

2) 名词前面一般能够加上表示物量的数量短语，一般不能加副词。

表 3 列举了所有“操作”具有名词功能的语料，根据《现代汉语》中列举的名词语法特征，我们对得到的语料进行客观分析，标记出词语在句子中所起的语法功能。从表 3 中可以看出，“操作”一词在句中常常具有以下名词性语法功能：作定语中心语、作定语、作介词宾语。

**Table 3.** The noun corpus and its grammatical function of the word “Cao Zuo” in the corpus of the National Language Commission

**表 3.** 国家语委语料库中“操作”一词的名词语料及其语法功能

序号	语料	语法功能
1、2	在许多工业企业里，手工 <b>操作</b> 和笨重体力劳动还占相当大的比重；即使在一些技术比较先进的工业企业里，在某些生产环节上，特别是辅助生产环节上，也还存在着不少手工 <b>操作</b> 和笨重的体力劳动。	作定语中心语
3	如果逐步地用机器来代替这些手工 <b>操作</b> ，提高生产过程的机械化和自动化水平，特别是尽快地实现笨重体力劳动的机械化，就可以大大地提高工业企业的劳动生产率。	作定语中心语
4	对于工业企业中已经采用机械 <b>操作</b> 的生产环节，根据科学技术的不断发展，和人们对生产规律认识的逐步深入，积极地改进现有的机器设备和生产工具，提高它们的自动化水平，充分发挥它们的生产潜力，也是提高劳动生产率的一个非常有效的方法。	作定语中心语
5	改进产品设计，采用先进的工艺和 <b>操作</b> 方法，也是提高劳动生产率的重要途径之一。	作定语
6	在进行工艺设计的时候，采用先进的工艺和 <b>操作</b> 方法，也能够有效地提高劳动生产率。	作定语
7	如果工人不具备必要的文化技术知识和 <b>操作</b> 技能，那末，即使有再好的机器设备和生产工具，再先进的生产工艺，也是不能提高劳动生产率的。	作定语
8	对非技工的训练，通常称为工作训练，即教授一种特定的 <b>操作</b> 方法。	作定语
9	这类操作(N)，多极简单，其训练有时仅须数小时，多者亦不过数周或数月。	作定语中心语
10	倘然一般国民知识程度极低，则虽极简单的机械 <b>操作</b> ，往往须经相当时间的训练，方能使其娴熟。	作定语中心语
11	工人学习了制胶理论后，已能在 <b>操作</b> 中掌握胶的物理性能，逐步解决了球鞋不耐磨的问题，到 1995 年年底，大底磨耗已由 8 月的 3.5 立方厘米，降到 2 立方厘米，围条掘折性由 9 月的一万转，提高到四万转以上。	作介词宾语
12	又根据永明油漆厂的经验，推行了“高温、急火、速冷”的炼油方法，改进了 <b>操作</b> ，贯彻了清洁制度，解决结皮问题等等。	作宾语
13	1958 年北京市轻工业的大跃进，是在克服原材料不足的情况下取得的，各行业通过多次改进技术 <b>操作</b> ，改进产品设计，改进配方等办法，千方百计地节约原材料，此外，还积极地试验代用原料，利用野生植物，从废料中提取化工原料，自制原料来开辟原料来源。	作定语中心语

此外，我们发现在国家语委语料库中，很多具有名词语法功能的“操作”被标记为动词。我们随机抽取了8条语料，分析“操作”一词在句子中所起的语法功能。如表4所示，可见，虽然“操作”在这些语料中被标记为动词，但实际上仍然具有名词的语法功能，如“作定语中心语”、“作定语”等。

**Table 4.** The verbal usage and grammatical function of the word “Cao Zuo” in the corpus of the National Language Commission

**表 4.** 国家语委语料库中“操作”一词的动词用法及语法功能

序号	语料	语法功能
1	这次《泰晤士报》老板要引进全套电子控制的印刷设备，这不但省人力，而且大大简化了印刷技术的 <b>操作</b> 。	作定语中心语
2	在这巨大的 <b>操作</b> 面前，对于刚离开学校的我来说，真感到要学习的东西太多了。	作定语中心语
3	我们知道大型高炉的 <b>操作</b> 是十分复杂的。	作定语中心语
4	他不忍见到国家财产遭受损失，于是他不断钻研，提高自己的高炉技术，就这样他班的 <b>操作</b> 成绩超过同时的大学生出身的工程技术人员，他被评为全厂的产品质量优良的技术员。	作定语
5	爆炸焊机适于水下两根管道对接的焊接，其 <b>操作</b> 大致为：在待焊两管外再加一套管，并在两管待接合处安放环状炸药和起爆装置。	作定语中心语
6	青年们刻苦钻研业务，每人都掌握了熟练的 <b>操作</b> 技术和应急检修机车的过硬本领。	作定语
7	这种用简便的 <b>操作</b> 而取得较为丰富的艺术效果，在装饰设计中是很可取的。	作定语中心语
8	吉原先生干得很不错，他的 <b>操作</b> 准确熟练，没有一个多余的动作。	作定语中心语

### 3.2. BCC 语料库

我们使用 BCC 语料库的词典功能，检索了“操作”一词，其词类标注及释义情况如下：

BCC 词典：【操作】<动>按照一定的程序和技术要求进行活动或工作：|~方法|~规程|

可见 BCC 语料库对“操作”的词性标注是动词，与其他三部汉语词典词类标注及释义相同。

弗斯在阐述词语搭配时提出了类联接这个概念。类联接是词语搭配研究中的一个重要概念，指的是文本中语法范畴间的结合[9]。类联接是词语搭配发生于其中的语法结构和框架。一个类联接代表了一个类别的词语搭配，可称为搭配类。从现代汉语词典的词类义项和国家语委语料库中可得出，“操作”一词确有名词用法。通过观察，我们发现在语料库中，“操作”一词作为名词用法的常见类联接及搭配有：

- 1) 名词 + 操作
- 2) 操作 + 名词
- 3) 介词短语 + 操作
- 4) “的” + 操作
- 5) 数量短语 + 操作
- 6) 操作 + “是”
- 7) 操作 + “要”
- 8) “等” + 操作

## 9) 名词 + 名词 + “等” + 操作

分别在 BCC 语料库中检索以上类联接的频次, 如表 5 所示。同时统计“操作”一词的常见类联接在 BCC 语料库中的总频次, 如表 6 所示。

**Table 5.** The frequency of common class connections of the word “Cao Zuo” in the BCC corpus

**表 5.** “操作”一词的常见类联接在 BCC 语料库中的频次

类联接	搭配实例	BCC 子语料库					总频次
		多领域	文学	报刊	对话	篇章检索	
名词 + 操作	机械操作	29,899	104	13,178	0	0	43,181
操作 + 名词	操作方法	37,165	163	21,336	0	0	58,664
介词 + 操作	在操作中	4701	44	1991	0	0	6736
“的” + 操作	最受用户喜爱的操作	11,896	112	6185	0	10,752	28,945
数量短语 + 操作	一次操作	2142	7	322	0	0	2471
操作 + “是”	操作是很规范的	496	0	122	0	440	1058
操作 + “要”	操作要仔细	969	1	269	0	923	2162
“等” + 操作	进行备份和恢复等操作	595	0	209	0	547	1351
名词 + 名词 + “等” + 操作	进行图文编辑等操作	28	0	13	0	0	41

**Table 6.** The total frequency and common grammatical functions of the common class connections of the word “Cao Zuo” in the BCC corpus

**表 6.** “操作”一词常见类联接在 BCC 语料库中的总频次及其常见语法功能

序号	类联接	总频次	常见语法功能
1	名词 + 操作	43,181	作主语, 作宾语, 作定语
2	操作 + 名词	58,664	作主语, 作宾语, 作定语中心语
3	介词 + 操作	658	作介词宾语
4	“的” + 操作	28,945	作主语, 作宾语, 作定语中心语
5	数量短语 + 操作	2471	作主语, 作宾语
6	操作 + “是”; 操作 + “要”	3220	作主语
7	名词 + 名词 + “等” + 操作; “等” + 操作	1392	作宾语

#### 4. 词类分析

通过对三部汉语词典和两个语料库的调查结果可发现, “操作”一词不止有动词用法, 还具有名词的语法功能。如语料库中的例句“这次《泰晤士报》老板要引进全套电子控制的印刷设备, 这不但省人力, 而且大大简化了印刷技术的操作。”“操作”在句中带定语, 作定语中心语, 是典型的名词用法。此外, 搭配频率可以较为清晰直观地反映“操作”一词经常所使用的结构类型。“操作 + 名词”“名词 + 操作”“‘的’ + 操作”这些结构也是较为典型的名词用法。BCC 语料库中的频次统计结果显示, 以上三个类联接的频次分别高达 58,664、43,181 和 28,946, 说明“操作”一词的名词用法是长期的、稳定的, 并不是临时的词类活用。

本研究以《现代汉语》中的名词语法特征作为标准, 以词语在句子中所起的语法功能作为依据对“操



作”的词性进行划分,通过观察“操作”在语料中的实际用法,总结出“操作”常见的名词性类联接用法,统计“操作”在句中的搭配类型频率。结果显示“操作”一词的名词用法相当高频。在我们调查的三部词典中,“操作”均被标注为动词,缺乏其名词释义和举例,显然违背了“操作”一词在汉语中有大量名词用法的事实。可见三部词典和语料库对于“操作”一词的词类标注都存在一定问题,只把“操作”一词标注为动词,忽略了“操作”一词作为名词使用的次数。根据语料库分析结果,可以得出“操作”不仅可以作动词,也可以做名词,是一个典型的兼类词。对于“操作”这类兼类词的词类标注,汉语词典的标注方式还有所欠缺。

## 5. 结语

本研究调查了三部汉语词典中“操作”一词的词类标注和释义,并通过搭配和类联接确定“操作”一词作为名词用法的常见结构,结合国家语委语料库和 BCC 语料库检索名词用法常见结构,发现“操作”一词的名词用法在汉语的使用中相当广泛。实证研究结果得出,“操作”一词具有大量名词用法,应属于兼类词。在汉语词典中,经常存在对兼类标注有分歧的词[10],应对“操作”此类兼类词进行修订,增加其作为名词的义项和释例,增加其名词词性。

如何解决汉语词典中多次出现类似于“操作”的词类标注缺失的问题?我们认为可以从以下两点入手:

第一,学界应当重新认识兼类词,词类并不是越单一越好。汉语词典均将“操作”一词的名词用法作为临时使用,而忽略其名词用法的稳定性和长期性,普遍存在人为减少兼类词的情况。这和长期以来学界认为兼类要尽可能少的“简略原则”有关[11]。实际上兼类词普遍存在于汉语中,并不是越少越好,人为减少兼类词违背了语言事实,不利于学习者利用词典进行学习。

第二,各类汉语词典在进行词类标注时,不能仅仅依靠权威词典的标注和释义。应使用 BCC 语料库、国家语委语料库等大型语料库对语料进行充分的分析和研究,应以语料库的实际语料用法为依据,全面客观地收录词语的用法,

## 参考文献

- [1] Beck, D. (2002) *The Typology of Parts of Speech Systems: The Markedness of Adjectives*. Routledge, New York.
- [2] 王仁强. 大数据视域下的汉英词典兼类表征策略研究——以《汉英词典》(第三版)为例[J]. 外语电化教学, 2020(6): 71-79.
- [3] 于屏方, 谭景春. 汉、英单语语文词典微观结构对比研究[M]. 北京: 科学出版社, 2021.
- [4] 陆俭明. 现代汉语语法研究教程[M]. 第四版. 北京: 北京大学出版社, 2013.
- [5] 胡明扬. 现代汉语词类问题考察[J]. 中国语文, 1995(5): 381-389.
- [6] 黄伯荣, 廖序东. 现代汉语(增订六版) [M]. 北京: 高等教育出版社, 2017.
- [7] 曾德祥. 词类“活用”与“兼类”的界定问题[J]. 成都大学学报(社会科学版), 1994(3): 99-103.
- [8] 邢丹, 饶高琦, 荀恩东. 基于大规模语料库的介词结构搭配库构建[J]. 中文信息学报, 2020, 34(11): 1-8.
- [9] Firth, J.R. (1957) *Papers in Linguistics 1934-1951*. Oxford University Press, London.
- [10] 徐艳华. 《现代汉语词典》与《现代汉语规范词典》兼类词标注差异研究[J]. 鲁东大学学报(哲学社会科学版), 2020, 37(5): 14-19+25.
- [11] 陆俭明. 关于词的兼类问题[J]. 中国语文, 1994(1): 28-34.