

# 基于生成词库理论的现代汉语“黑”字义项的分布情况研究

王元

西南大学汉语言文献研究所, 重庆

收稿日期: 2023年7月20日; 录用日期: 2023年10月16日; 发布日期: 2023年10月25日

## 摘要

随着社会的发展、不同文化的交流碰撞, 现代汉语常用词的义项正在不断增加, 因此丰富并补充常用词义项是一项刻不容缓的工作。本文通过中英文词典释义对比与平衡语料技术发现了现代汉语“黑”字未被讨论到的义项, 并尝试结合生成词库理论与词典释义模型重新描写现代汉语中的“黑”字义项, 以期探索用新技术新理论丰富并描写常用词新义项的可行性, 为词典释义提供新思路。

## 关键词

黑, 生成词库理论, 平衡语料库, 描述性话语释义

## Distribution of the Meaning Items of the “Hei” Character in Modern Chinese Based on the Generative Lexicon Theory

Yuan Wang

Institute of Chinese Language and Document, Southwest University, Chongqing

Received: Jul. 20<sup>th</sup>, 2023; accepted: Oct. 16<sup>th</sup>, 2023; published: Oct. 25<sup>th</sup>, 2023

## Abstract

With the development of society and the collision of different cultures, the meanings of commonly used words in modern Chinese are constantly increasing. Therefore, enriching and supplementing the meanings of commonly used words is an urgent task. This paper discovers the unresolved meanings of the “Hei” character in modern Chinese through the comparison of definitions in Chi-

nese and English dictionaries and the use of balanced corpus technology. It attempts to combine generative vocabulary theory with dictionary interpretation models to re-describe the meanings of the “Hei” character in modern Chinese, in order to explore the feasibility of enriching and describing new meanings of commonly used words with new technologies and theories, and provide new ideas for dictionary interpretation.

## Keywords

Hei, Generative Lexicon Theory, Balanced Corpus, Descriptive Vocabulary Definitions

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

关于现代汉语“黑”字语义建构问题,张念歆、宋作艳(2015)考察了“黑+名”的物性结构修饰关系[1],王恩旭、袁毓林(2018)考察了词义中物性角色分布及其对“黑”释义的影响[2],王成(2021)考察了以词典用户为中心的现代汉语“黑”字释义[3]。但通过中英文释义对比与语料技术发现“黑”字仍有未被讨论到的义项。该字在北京大学中国语言学研究中心的CCL现代汉语语料库(以下简称语料库)中共有语料130,831条,本文依照李仕春(2016)介绍的抽样语料法[4],首先对“黑”字语料进行分类下载建立10,000条(包含人名、地名等无效语料1705条)共计51,798字的平衡语料库,并通过对比《牛津高阶双解词典》(第9版)(以下简称《牛高》)与分析语料找出《现代汉语词典》(第7版)(以下简称《现汉》)中漏收的“黑”字义项。然后根据生成词库理论,将释义模型[5]与描述性话语结合重新描写“黑”在现代汉语中的义项,以期探索用新技术新理论丰富并描写常用词新义项的可行性,为词典释义提供新思路。

## 2. “黑”的形容词性义项分布情况

形容词的语义具有依附性,释义应尽量呈现其修饰成分的物性角色,才能完整表达词义。因此运用物性结构理论描述形容词应先确定形容词所依附的成分,然后根据“物性角色+核心成分”的形式描写义项。《现汉》中“黑”字的形容词性义项描写见2.1~2.4,《现汉》中未收录但语料库中发现的“黑”字新形容词性义项描写见2.5~2.7。

### 2.1. 形像煤或墨的颜色(跟“白”相对)

该义项在“黑”字平衡语料库中出现频率最高,在语料库中共有5191条,占比约51.91%。“黑”的核心成分是一种颜色,人们日常生活和观念中黑色的认知典型便是煤或墨。但这一释义不能解释清楚该义项下包括的所有条目,语料显示“黑”也常表示一些近似黑色的深颜色。如《现汉》中该义项下的例子“脸都晒~了”,并不是表示脸晒成像煤或墨的颜色,而是表示脸的颜色晒得比以前更深(发黄或者发褐)了。该类在语料库中共有313条,占比约3.13%。

例1: 农民们只好背着香蕉去40公里外的武利镇兜售,卖不掉的眼看着发黑烂掉。(《人民日报》1994)

因此我们可以根据“形式角色+核心名词”的释义模式,凸显“黑”的形式角色,将这一义项描写

为：“形像煤或墨的颜色(跟“白”相对)：~头发 | 白纸~字；有时也泛指近似黑色的深色：~斑 | 脸都晒~了。”

## 2.2. 形黑暗

分析语料时发现这一义项下的“黑”有时是缺少光线而显得暗，该类在语料库中共有 1513 条，占比约 15.13%。

例 2：在平地教他们许多滑雪的基本动作，等轮到让他们自己练时天就快黑了。(《从普通女孩到银行家》)

有时是因为物体吸收了全部光线而呈现出来昏暗，如例 3 提到的黑洞，是时空曲率大到光都无法从其中逃脱的天体，黑是因为光线被全部吸收。该类在语料库中共有 128 条，占比约 1.28%。

例 3：英国牛津大学和南安普敦大学的天文学家发现的这个新黑洞位于狐狸座。(《人民日报》1995)

《现汉》只描写了黑暗的状态却没有交代成因，因此可能出现语料与义项匹配度不高的情况。我们可以根据“施成角色 + 核心名词”的释义模式，凸显“黑”的形式角色，或许可以描写为：“形由于缺少光线而看不清楚：天~；或由于吸收了全部光线和物质而显得黑暗：~洞”。

## 2.3. 隐秘的；非法的

该义项从物性结构角度看，突出了人们对“黑”所指事物的主观评价。但在考察语料库时发现并非所有语料的义项都同时符合“秘密”和“非法的”。有的只能对应“秘密、不公开”义，该类在语料库中共有 27 条，占比约 0.27%。

例 4：看见村里的土财主横行霸道，气愤不过，也只能画张黑帖儿，偷偷地贴到他家的门口。(孙犁《风云初记》)

有的符合“不道德的、违规的”，该类在语料库中共有 83 条，占比约 0.83%。

例 5：对于足球界的一些不正常现象，如假球、黑哨等。(新华社 2001)

有的符合“非法的”，该类在语料库中共有 172 条，占比约 1.72%。

例 6：人大代表呼吁严打黑社会的“保护伞”。(新华社 2003)

这一义项的释义可依次分为“秘密的、不公开的”“不道德的”“非法的”三个等级。因此我们可以根据“评价 + 成分”的释义模式，凸显“黑”的评价角色，推荐使用义项群释义：“a) 隐秘的，形容事情很难被注意到：~帖；b) 不道德的，形容事情不正当或不符合规定：~哨；c) 非法的，形容事物不符合规定或不符法律：~社会。”

## 2.4. 形坏；狠毒

该义项突出了人们对“黑”所指人的主观评价，形容人的道德败坏。该类在语料库中共有 160 条，占比约 1.6%。

例 7：古之所谓英雄豪杰者，不外面厚心黑而已！(李宗吾《厚黑学》)

因此我们可以根据“评价 + 成分”的释义模式，凸显“黑”的评价角色，建议释义改为：“形坏、狠毒；形容人的道德品性败坏：~心肠 | 这种人心太~。”

## 2.5. 因愤怒或怨恨而使脸色变得阴沉难看；因周围环境恶劣或生活艰难而心情沮丧

“颜色词会衍生出形容人情绪的形容词义项” [6]。《牛高》对这一义项的解释(without hope; very depressing [7])侧重“黑”的评价角色，描述出黑色带给人们看不到希望、消极沮丧的心理感受。通过语料我们发现这部分的施成角色主要有两种：第一种是表现在脸色上的“黑”，该类在语料库中共有 35 条，占比约 0.35%。

例 8: 范英明脸黑了几分。(柳建伟《突出重围》)

第二种是形容心情的“黑”，该类在语料库中共有 311 条，占比约 3.11%。

例 9: 中国围城国际象棋队遭遇“黑色星期五”，本轮战绩为 1 胜 2 和 3 负，是开赛以来最差的一轮。(新华社)

可用一个义项群描写该义项，采用“评价角色 + 核心名词”的释义模式，描写造成这种情绪或心理状态的原因 + 心理状态。

## 2.6. 形脏，因沾上泥土、灰尘等而变得不干净

《牛高》的“black”释义中有“脏”义(very dirty; covered with dirt [7])，从施成角色的角度看这个释义既写出脏的表现状态是“黑”，也指出造成“黑”这种状态的原因是“有污垢”。在语料库中也发现，汉语在形容水脏时也会用“黑”，如“黑河”“黑水”等。该类在语料库中共有 53 条，占比约 0.53%。

例 10: 脸干净了鼻子上有黑。(《人民日报》1993)

因此建议增加这一义项。可根据“施成角色 + 核心名词”的释义模式，凸显“黑”的施成角色。

## 2.7. 不添加其他东西的，特别是白面、白糖、牛奶等制品

《牛高》中的“black”中有形容词性义项“不添加牛奶的(without milk [7])”，而语料库中也发现了“黑咖啡”“黑面包”“黑巧克力”等，该类在语料库中共有 10 条，占比约 0.1%。这里的“黑”所依附的成分多为食物，表示在制作中不添加牛奶、白面等。因此建议《现汉》增加这一义项，可以根据“构成角色 + 核心名词”的释义模式，凸显“黑”的构成角色。

# 3. “黑”的名词性义项分布情况

“名词语义具有自主性，其义项的描写可以通过其自身蕴含的物性角色完整表达” [8]。因此可以先确定每个义项中的核心名词，然后根据“物性角色 + 核心名词”的形式把义项描写出来，以达到用描述性话语释义的目的。《现汉》中“黑”字的名词性义项描写见 3.1，《现汉》中未收录但语料库中发现的“黑”字新名词性义项描写见 3.2。

## 3.1. (~儿)夜晚；黑夜

《牛高》对时间词“night”的释义(the time between one day and the next when it is dark, when people usually sleep [7])由形式角色(形容白天到黑夜的一段时间)、构成角色(黑暗)和功用角色(人们通常在这段时间睡觉)构成。该类在语料库中共有 78 条，占比约 0.78%。

例 11: 我和方方走到停自己车的地方，摸黑坐进去，也很快开走了。(王朔)

根据物性结构理论，可从构成角色方面分析，即黑夜的组成成分和构成状态。考察语料发现“黑”在现代汉语这一义项中侧重突出“夜晚”“光线不足”等特点，功用角色“人们在这段时间睡觉”暂不

应纳入释义。故用描述性话语可以表示为“太阳落下后到天亮之前，光线不足、不明亮的一段时间。”

### 3.2. 错误，不正当或不道德的行为、观念、事情

《牛高》中对“wrong”的释义(not morally right or honest [7])由核心名词“行为”和评价角色“(这个行为不诚实或者不道德)”组成。该类在语料库中共有 124 条，占比约 1.24%。

例 12: 在布告中，日军混淆黑白，贼喊捉贼。(沈永兴、朱贵生《二战全景纪实》)

可从评价角色方面分析，确认“黑”的核心名词是一种行为，对它的评价是不正当或不道德。用描述性话语表示为“不正当或不道德的行为、观念、事情。”

## 4. “黑”的动词性义项分布情况

“运用论元结构理论描述动词义项的步骤是：先确定每个义项中的核心动词，然后根据核心动词所搭配的论元类型及论元数目，把该义项描述出来达到用描述性话语释义的目的。虽然‘黑’的动词性义项有自主性，但其他词语的参与更有助于完整表达动词的意义。其语义结构除了包含动作和必有的论元信息，还应包括意愿、工具、方式、依据、目的等”[9]。《现汉》中“黑”字的动词性义项描写如下。

### 4.1. 动暗中坑害、欺骗或攻击

该类在语料库中共有 11 条，占比约 0.11%。这里“黑”是二元动词，动作涉及到了收益性施事论元(“坑害”动作行为的发出者，通常是人)、受事论元(因施事“坑害、欺骗”行为而受到影响的对象)、方式论元(动作行为采用隐蔽地、不被人知道的方式)。方式论元具有非自立性和附庸性的语义特点，所表示的方式依附于核心动词“坑害、欺骗或攻击”所表动作行为之上，比较准确完整地描写了“黑”的动词性义项。

### 4.2. 动通过互联网非法侵入他人的计算机系统查看、更改、窃取保密数据或干扰计算机程序

该类在语料库中共有 86 条，占比约 0.86%。“黑”涉及了四种论元角色，分别是工具论元(互联网)、受事论元(计算机系统)、方式论元(非法)、目的论元(查看、更改、窃取保密数据或干扰计算机的程序)，它们与核心动词“侵入”一起比较准确完整地描写了这一动词性义项。

## 5. 结语

综上，我们将现代汉语“黑”字义项重新描写如下：① 形像煤或墨的颜色(跟“白”相对)：~头发 | 白纸~字；有时也泛指近似黑色的深色：~斑 | 脸都晒~了；② 形由于缺少光线而看不清楚：天~；或由于吸收了全部光线和物质而显得黑暗：~洞；③ (~儿)太阳落下后到天亮之前，光线不足、不明亮的一段时间：摸~起早贪~；④ a) 隐秘的，形容事情很难被注意到：~帖；b) 不道德的，形容事情不正当或不符合规定：~状；c) 非法的，形容事物不符合规定或不合法律：~社会；⑤ 形坏、狠毒；形容人的道德品性败坏：~心肠 | 这种人心太~；⑥ a) 因愤怒或怨恨而使脸色变得阴沉难看：脸~；b) 因周围环境恶劣或生活艰难而使心情沮丧，看不见希望：~色幽默；⑦ 形脏，因沾上泥土、灰尘等而变得不干净：~河 | ~水；⑧ 动暗中坑害、欺骗或攻击：被骗子~了两万块钱 | 昨夜被人~了一砖头；⑨ 动通过互联网非法侵入他人的计算机系统查看、更改、窃取保密数据或干扰计算机程序：~客；⑩ 错误，不正当或不道德的行为、观念、事情：混淆~白；⑪ 不添加其他东西，特别是白面、白糖、牛奶的制品：~咖啡；⑫ 姓。

## 致 谢

西南大学文学院李仕春教授、导师杨文全教授及《现代语言学》编辑部和匿名评审专家提出了宝贵的修改意见，致以最诚挚的谢意！

## 参考文献

- [1] 张念歆, 宋作艳. 汉语形名复合词的语义建构: 基于物性结构与概念整合理论[J]. 中文信息学报, 2015, 29(6): 38-45.
- [2] 王恩旭, 袁毓林. 词义中物性角色的分布及其对词语释义的影响——以“颜色语素 + 名物语素”复合词为例[J]. 外国语(上海外国语大学学报), 2018, 41(2): 31-41.
- [3] 王成. 以词典用户为中心的现代汉语颜色词释义研究[D]: [硕士学位论文]. 重庆: 西南大学, 2021.
- [4] 李仕春. 基于语料库的现代汉语“黄”字义项分布研究[J]. 中国海洋大学学报(社会科学版), 2016(5): 103-109.
- [5] 王恩旭, 袁毓林. 机器词典释义模版的建构和运用[J]. 中文信息学报, 2018, 32(1): 68-74.
- [6] 张祝祥, 杜凤兰. 英汉基本颜色词“黑”、“白”的认知语义分析[J]. 外语与外语教学, 2007(5): 47-50.
- [7] [英]霍恩比. 牛津高阶英语词典[M]. 第9版. 北京: 商务印书馆, 2016: 200-2497.
- [8] 袁毓林. 汉语名词物性结构的描写体系和运用案例[J]. 当代语言学, 2014, 16(1): 31-48.
- [9] 袁毓林. 论元角色的层级关系和语义特征[J]. 世界汉语教学, 2002(3): 10-22.