

英语颜色词的语料库研究

胡 刚¹, 李 慧¹, 田传茂²

(1. 武汉大学 外语系, 湖北 武汉 430072; 2. 荆州师范学院 外语系, 湖北 荆州 434100)

摘要: 利用现代语料库手段统计英语简单颜色词和合成颜色词在 Collins CoBuild 语料库中的分布情况, 可以发现简单颜色词在普通文本 (General Texts) 和科技文本 (Scientific Texts) 中的分布不平衡。合成颜色词在语料库中的使用频率极低, 其功能意义为简单颜色词词组所取代。简单颜色词词组的搭配存在一定的规律。这一发现验证了某些西方语言学者的一个观点, 即泛称词逐渐取代特称词。

关键词: 英语颜色词; 语料库; 分布

中图分类号: H313 **文献标识码:** A

随着计算机技术的迅猛发展, 现代语料库 (计算机语料库) 的建设与发展也日新月异。各种语料库软件工具如索引软件 (concordancer)、词语搭配软件 (collocator)、词汇标记语法软件 (lexical parser)、词形归类软件 (lemmatiser)、短语查找软件 (phrase finder) 等被不断地开发出来^[1]。语料库软件的大量涌现提高了语料库的利用价值, 从而帮助语言研究者超出传统的个案研究的范围, 从宏观的角度再现人类语言的真实面貌。

Mike Scott 和 Tim Johns 研制的索引软件 MicroConcord (1993) 属于 COBUILD 语料库软件^[2]。该软件具有四种基本功能: (1) 查找并显示关键词 (key word) 及索引行 (concordance lines); (2) 显示关键词所在的语段 (context); (3) 定长 (span), 即定义关键词左右的搭配词数目, 并按字母或频率高低顺序显示字长范围内搭配词的频率; (4) 对索引行进行排序 (sorting)。笔者利用这一软件工具对英语颜色词在 COBUILD 语料库中的分布情况进行了尝试性分析。

一、英语简单颜色词在 COBUILD 语料库中的分布

语言的词汇系统, 由千万个大大小小的语义场或词汇集构成。在许多语义场里, 都有一个词群。每一个词群中的个体, 又以其独特的价值分布在语义场里。英语颜色词从简单颜色词到合成颜色词数量有很多, 构成了一个庞大的语义场^[3]。red, blue, yellow, green, black, white 是这个语义场中的主要成员, 分别表示客观世界

中的一种基本颜色, 我们称之为简单颜色词。那么, 这些简单颜色词在语言的实际使用中是如何分布的呢? 它们的使用频率情况如何呢? 笔者根据 MicroConcord 进行了粗略统计 (见表 1)。

表 1 六种简单颜色词的频率 (以每百万个单词为单位)

语域	简单颜色词					
	red	blue	green	yellow	black	white
一般文本 (General Texts)	85	53	158	21	144	190
科技文本 (Scientific Texts)	86	69	59	26	100	138
全 库	86	57	115	23	126	169

注: 以上及文中所有的统计数字来源于华中科技大学外语系的 COBUILD 语料库 (经过压缩)。库总容量为 7 301 775 个单词。该库分为一般文本 (General Texts) 和科技文本 (Scientific Texts) 两大部分; 前者为除科技文献以外的一切书面语, 容量为 4 348 703 个单词; 后者主要是科技文献, 容量为 2 953 070 个单词^[2]。

表 1 中的统计数字为标准化数字, 即表示关键词 (简单颜色词) 在每百万字中出现的频率。从表中可以看出, red, green, black, white 的使用频率明显高于 blue 和 yellow, 其中 white 的使用频率最高, 每百万字中出现 190 次, yellow 的使用频率最低, 每百万字中只出现 21 次。从使用的语域来看, red 的分布较为平均, 在一般文本和科技文本中出现的频率大致相当 (85 和 86), 其余五种颜色词, 特别是 green, black 和 white 在两种语

域里分布不平衡, green, black, white 在一般文本中的分布高于科技文本, blue 和 yellow 则正好相反, 在科技文本中的使用频率要高于一般文本。

二、合成颜色词在 COBUILD 语料库中的分布

构成颜色词这个大语义场的核心词 red, blue,

yellow, green, black, white 也拥有各自的小语义场。从语义学角度看, 这六种简单颜色词在自己的语义场中属于泛称词 (general terms), 分别拥有一个由不同合成颜色词组成的词群。合成颜色词指的是那些处于各种简单颜色词介面并以某种简单颜色为主色的特称词 (specific terms)。根据 The Random House Thesaurus (College Edition, 1989) 同义词典分别列出六种简单颜色词的同义词 (特称词) (见表 2)。

表 2 简单颜色词 (泛称词) 与合成颜色词 (特称词)

简单颜色词	合成颜色词 (复合词类排除在外)
red	ruby 深红、红宝石色、crimson 紫红、深红、cherry 樱桃红、鲜红、scarlet 鲜红、sanguine 血红、红润的、cerise 清亮的红、maroon 茶色、rosy 玫瑰色的、coral 珊瑚色
blue	azure 天蓝的、深蓝的、cerulean 天蓝、sapphire 蓝宝石色、lapis lazuli 天青石色、鲜蓝的、indigo 深紫蓝、turquoise 蓝绿的、湖蓝的、天蓝的
yellow	lemon 淡黄色、柠檬色、canary 鲜黄色、gold 金黄色、ocher 赭黄色、saffron 橘黄、flaxen (头发) 浅黄的、blond (头发) 金黄的
green	chartreuse 黄绿色的、lime 酸橙色、暗黄绿色的、olive 橄榄色、verdant 翠绿的、青翠的、jade 翡翠色、emerald 祖母绿
black	jet 黑玉色、raven 乌黑、乌亮、ebony 乌木色的、黑檀木色的、inky 墨黑的、sable 黑色的、黑貂皮色、swarthy (皮肤) 黝黑的、murky 漆黑的
white	ivory 象牙色、snowy 雪白、silvery 银白、milky 乳白色、ashy / ashen 灰白、pale / pasty 苍白的、grizzled / hoary (头发) 灰白的、bleached 漂白的、leadен 深灰色、alabaster 雪白、pearly (牙齿) 珍珠白的

表 2 中列举出的合成颜色词均是词典中的独立词项, 复合类颜色词如 purplish-blue 不在分析之列。那么, 这些合成颜色词在 COBUILD 词料库中的分布情形

如何呢? 我们以构成 blue 语义场的词群为例, 看看六个合成蓝色词 azure, cerulean, sapphire, lapis lazuli, indigo, turquoise 在库中的频率分布 (见表 3)。

表 3 六个合成蓝色词在 COBUILD 语料库中的频率 (以每百万个单词为单位)

语 域	合成蓝色词					
	azure 天蓝、深蓝	cerulean 天蓝	sapphire 蓝宝石色	lapis lazuli 鲜蓝的	indigo 深紫蓝	turquoise 湖蓝、蓝绿
一般文本	0.7	0.2	0.2	0.2	0	0.2
科技文本	0	0	0	0	0.33	0
全 库	0.4	0.13	0.13	0.13	0.13	0.13

注: 只统计关键词即六个合成蓝色词作为颜色意义的频率。

从表 3 可以看到, 合成蓝色词使用频率非常低, 每百万个单词中出现的频率不到一次, 而且几乎不在科技文献中出现。其他颜色的合成颜色词在库中的分布情形与合成蓝色词大致相似, 如合成红色词 rudy, 虽然在库中出现 43 次, 却只有一次用作颜色意义, 合成绿色词 verdant 也只一般在一般文本中出现了两次, 合成黄色词 canary 在库中出现了 17 次, 却只有一次作颜色意义

使用。这就提出了一个问题: 大千世界如此五光十色, 人们如何来表达眼睛所看到的各种合成颜色呢? 笔者对每个简单颜色词 (关键词) 的索引行逐一进行了检索, 发现它们之前经常出现程度修饰词, 如 light, deep, dark, bright, pure 等。鉴于此, 笔者将所有带有这类修饰词的 blue 索引行进行了排序 (sorting)。表 4 是经过排序后的 blue 索引行的一部分实录。

表 4 排序后的关键词 blue 索引行部分实录

1. coloured flowers, and, in the	clear blue	sky above, larks spilling silvery cal
2. deed and, beyond them, the	clear dark blue	of spare, it is just possible, at th
3. highly coloured for example, the	deep blue	of tetramine copper(II), I(u(M-1))
4. great big brow mouth and those	deep blue	eyes, who will change from child to
5. statuesque creature with	very grey pale blue	eyes, and a great star in those day
6. in with the one basic item—the	navy blue	jeans—and the latter because it w
7. s sight—the sea and sky a	brilliant blue	against the white of the breakers a
8. itty little Easter bonnet of	Prussian blue	crepe—de—chine and ostrich feathers
9. sways as he answers. Dressed in	faded blue	trousers, open—hecked white shirt a

表 4 告诉我们,各种不同的合成蓝色在实际语言运用中可以通过利用诸如 clear, deep, dark, prussian, faded, navy, brilliant 等修饰语与 blue 搭配来表达。不仅合成蓝色如此,其他合成颜色也是一样。笔者通过查找 red, blue, green, yellow, black, white 六个关键词,分析其索引行,发现简单颜色词除了表示本色外,

大都可与各种修饰词搭配构成词组来表示简单颜色的细微变化即合成颜色,如深红 deep red, 淡蓝 light blue, 鲜绿 bright green, 灰黄 pale yellow, 深黑 very black 等。表 5 是六个简单颜色词与各种修饰词搭配表示合成颜色的统计表。

表 5 COBUILD 语料库中表示合成颜色的简单颜色词词组

简单颜色词	表示不同合成颜色的简单颜色词词组
red	blood red, orange red, sharp cherry-red, bright red, deep red, dark red, pale red, spotty red
blue	light blue, dark blue, navy blue, clear blue, clear dark blue, pale blue, bright blue, sheer blue, purest blue, Prussian blue, brilliant blue, heavenly blue
green	pale green, dull green, iridescent green, bright green, fairly green, dark green, quite green, fresh green
yellow	pale yellow, brilliant golden yellow, bright yellow
black	very black, pitch black, deep black
white	

简单颜色词词组中的修饰语可分为三类:(1)表示程度的形容词,如 light, deep, bright, purest, sheer 等;(2)形容词性颜色词,如 pale, dark;(3)具有某种固定颜色的事物,如 orange, blood, sky, pitch, sea 等。通过修饰词的帮助,简单颜色词可以表达由浅至深(light, pale→sheer, pure→deep, dark)几乎所有的合成颜色。以红色为例,可用 blood red 表示 sanguine, dark red 表示 crimson 或 camine, deep red 表示 ruby 或 crimson, bright red 表示 cherry 或 scarlet 等。简单颜色词表示合成颜色除了利用词组形式外,也可采用复合词形式,如 cherry-red, sky-bule, olive-green, yellow-green, all-white, blue-on-white, black-and-white 等。这些复合词的构词方式主要有两种:(1)由表示具有某种固定颜色的事物

的词和该种简单颜色词组成,如 olive-green, sky-blue, emerald-green;(2)简单颜色词之间的搭配,如 blue-to-green, white-on-blue, grey-white, pink-and-white。

三、简单颜色词词组主要结构方式

从上面的分析中我们可以看到,简单颜色词词组的搭配存在着一定的规律。除了某些事物名词如 blood, orange, navy, pitch 只能与某一个简单颜色词搭配外,其他用于六个简单颜色词前的修饰语,从其共同的词频来看,bright, pale, dark, light 以及 deep 使用得比较普遍。表 6 列出了简单颜色词与主要前置修饰语的搭配情况。

表6 COBUILD 语料库中简单颜色词词组的搭配规律

简单颜色词	修饰语								
	light	pale	dull	dark	bring	brilliant	deep	quite /fairly	pure /sheer /clear
red		✓	✓	✓	✓				
blue	✓	✓		✓	✓	✓			✓
green		✓		✓	✓		✓	✓	
yellow		✓			✓				
black			✓				✓		
white									

表6显示, pale, dark, bright 三个词在简单颜色词的搭配中最为活跃, 可用来修饰 red, blue, green, yellow 四种颜色。

四、英语简单颜色词和合成颜色词语料库分布情况及其成因分析

通过比较表1和表3, 我们发现英语简单颜色词的使用频率要比合成颜色词高得多, 各种合成颜色词表示的意义往往可以通过简单颜色词与各种修饰词搭配来表达。在英语颜色词语义场中, 存在着泛称词取代特称词的趋势。词汇逐渐抽象化反映了词汇发展的一种趋势, 也符合人类的认识由具体到抽象的辩证法的认识观。

史料证明, 泛称词和特称词显现过三种时间次序: (1)特称词先于泛称词; (2)泛称词先于特称词; (3)泛称词、特称词、泛称词相继出现。关于第一种情况, 著名语言学家叶斯泊森(Jespersen O.)在《语言: 它的性质、发展和起源》(1922)中谈到中古英语时说, 在形容羊毛或鹅、马、牛、人的头发或某些动物的毛时, 要用不同的表示灰色的特称词。现代英语一般只用一个泛称词 grey。关于第二种情况, 在英语里, 15世纪以前只有泛称词 mathematics(数学), 17世纪以后又产生了特称词 topology(拓扑)、calculus(微积分)、cybernetics(控制论), 等等。关于第三种情况, 以汉语为例, 在甲骨文中, 先有泛称词“马”, 后有特称词“驳”等, 到了后来的典籍中“马”的特称竟达到100多个, 它们用来指明马的性别、年龄、高度、优劣、长相、毛色、产地、用途, 等等。这些形形色色的马的特称词, 后来绝大多数都消失了, 最常用的只剩下泛称词“马”。

从泛称词和特称词的三种发展情况看, 词汇发展的主导趋势是逐渐抽象化, 其决定性因素是客观需要的变化, 最重要的是社会历史时代的需要发生了变化。语言是特殊社会现象, 它的主要变化, 特别是词汇的变

化, 是受社会历史时代制约的。正如法国哲学家列维·布留尔(Lucien Levy-Bruhl, 1859—1939)所说, 原始人的语言“特别注意表现那些为我们的语言所省略或者不予表现的具体细节”^[4]。后来, 社会发展了, 语言也随之发展, 由于旧语义场中的特称词数量过多, 妨碍了语言交际功能的发挥, 于是特称词逐渐减少, 以致被泛称词取代。

合成颜色词既然仍在语言中存在, 表明它具有自身的价值。像 cherry(樱桃红), ruby(红宝石色), carmine(胭脂红), rosy(玫瑰色的), coral(珊瑚色), 它们表示颜色时, 能使人产生联想。如果用简单颜色词词组代替, 如用 bright red 代替 cherry, 红樱桃的形象就消失了。这些合成颜色词在文学作品中, 特别在诗歌语言中占据着重要的地位, 是简单颜色词无法取代的。

通过分析英语颜色词在 COBUILD 语料库中的分布情况, 我们发现: (1)英语简单颜色词 red, blue, green, yellow, black, white 在语料库以及两种语域中分布不平衡; (2)独立的合成颜色词的使用频率很低, 其功能为简单颜色词词组所取代; (3)简单颜色词词组的搭配存在一定的规律。通过语料库分析, 可以帮助英语学习者较好的掌握英语颜色词的全貌, 使他们能准确地使用各种颜色词或词组去描绘他们眼中五光十色、多姿多彩的世界。

参考文献:

- [1] 王建新. 语料库语言学发展史上的几个重要阶段[J]. 外语教学与研究, 1998, (4): 52—58.
- [2] SCOTT Mike, JOANS Tim. MicroConcord[M]. Oxford University Press, 1993.
- [3] 李侠. 从颜色词看中诗英译风格、文体和文化的转化[J]. 中国翻译, 1995, (2): 10—14.
- [4] 张志颜, 张庆云. 词和词典[M]. 北京: 中国广播电视出版社, 1994.

(责任编辑 涂文迁)