

基于语料库的近义词用法比较研究

——以 *endless* 和 *enduring* 为例

李屹

天津外国语大学, 天津 300000

DOI:10.61369/EDTR.2025080041

摘要 : 本研究基于 COCA 语料库对 *endless* 和 *enduring* 的对比分析显示: (1) 词频分布存在显著差异; (2) 语域偏好呈现口语与书面语的分化; (3) 语义韵特征分别呈现中性和积极倾向。研究结果系统揭示了二者在语义语用层面的区别特征, 为近义词教学提供了实证依据, 对提升学习者词汇深度知识具有实践价值。

关键词 : 近义词; 语料库; 语域; 搭配词; 语义韵

A Comparative Study on the Usage of Synonyms based on Corpus: Taking Endless and Enduring as Examples

Li Yi

Tianjin Foreign Studies University, Tianjin 300000

Abstract : The comparative analysis of *endless* and *enduring* based on the COCA corpus in this study shows that: (1) there are significant differences in word frequency distribution; (2) Register preferences show a differentiation between spoken language and written language; (3) The semantic and rhyming features respectively show a neutral and positive tendency. The research results systematically reveal the distinct characteristics of the two at the semantic and pragmatic levels, providing empirical evidence for the teaching of synonyms and having practical value for enhancing learners' depth of vocabulary knowledge.

Keywords : **synonyms; corpus; register of language; collocation; semantic rhyme**

引言

作为全球通用语, 英语词汇体系复杂, 近义词辨析已成为以中国学习者为代表的 EFL 群体所面临的核心挑战。词汇是语言能力的基础, 近义词的准确运用直接影响语言表达能力的发展。

为此, 本研究选择以 COCA 为事实依据, 运用语料库语言学方法, 对形容词 *endless* 与 *enduring* 进行系统剖析。研究将逐层考察二者的词频分布、活跃的语域、典型的搭配行为及其语义韵倾向, 旨在通过厘清它们的差异, 为词汇教学提供扎实的实证依据。

一、文献综述

掌握近义词的精细差异, 是提升二语表达精准性与地道性的关键。根据 Lyons 的经典定义, 近义词是“意义上相近但不相同的词”^[4]。Firth 提出的“由词之结伴可知其词”^[2]奠定了搭配研究的理论基础, Xiao 和 McEnery 进一步指出“近义词的搭配及语义韵特征决定了其不可完全替代性”^[6]。

现有研究对近义词的辨析, 多倚重词典释义或研究者直觉, 二者均有其局限。词典例证有限, 在呈现词语全貌时难免捉襟见肘; 而直觉判断则深受个人语感左右, 如同雾里看花。以《牛津英语词典》对 *endless* 与 *enduring* 的释义为例, 其界定不仅抽象, 例证甚至出现交叉错配, 这无疑暴露了传统方法的固有

缺陷。

语料库语言学的兴起, 为近义词研究带来了范式革新。卫乃兴比较了动词 *commit*、*effect*、*cause* 在不同语料库中的语义韵特征^[8]、陆军系统考察了因果类动词在学习者语料中的语义韵差异^[7]与赵红艳借助 COCA 语料库对近义名词进行了多维度分析^[10]。这些研究均为本课题提供了重要的方法论参考。然而, 针对 *endless* 和 *enduring* 的实证研究尚显不足。

基于此, 本研究依托 COCA 语料库, 遵循数据驱动的研究路径, 通过考察词频特征、语域分布、搭配行为与语义韵倾向这四个维度, 对 *endless* 与 *enduring* 进行系统比对。重点探究以下问题: 首先, 二者在不同语域中的分布呈现何种规律? 其次, 它们的典型搭配行为如何刻画其用法特征? 最后, 二者的语义韵倾向

是否存在本质差异?

二、研究工具和研究步骤

本研究以美国当代英语语料库 (COCA) 为数据来源进行数据分析。该语料库因其规模宏大、结构平衡, 被学术界广泛采纳为研究当代英语使用特征的重要依据。^[9]

研究通过三个步骤展开分析: 第一步, 利用【Chart】模块绘制二者在不同语域中的分布图谱; 第二步, 采用【Compare】工具 (跨距设定为 -3 至 +3) 捕捉其高频搭配行为; 第三步, 结合

表1 endless 与 enduring 在各语域中的词汇频率分布

词项		博客	电影电视	学术期刊	口语	小说	网站	杂志	总数
endless	出现频次	2750	1005	1388	1010	3792	2493	2992	17189
	相对频率 (%)	21.38	7.85	11.59	8.01	32.05	20.06	23.73	17.31
enduring	出现频次	720	137	1766	579	491	1042	1534	7360
	相对频率 (%)	5.60	1.07	14.74	4.59	4.15	8.39	12.17	7.41

基于 COCA 语料库的语域分析表明, 形容词 endless (总频次 17189) 与 enduring (总频次 7360) 在语域分布上呈现显著的互补特征。具体而言, endless 高度集中于小说、杂志及口语等非正式语域, 其口语使用频率达 enduring 的两倍; 而 enduring 则显著偏向学术期刊等正式语域, 体现出鲜明的学术语体特征。这一分布差异从本质上揭示了两者在语体身份与交际功能上的根本区别。

研究表明, endless 呈现口语化特征, 应用范围更广;

表2 endless 与 enduring 的高频搭配词 (频率数与互信息值)

endless				enduring			
序号	搭配词	MI 值	MI Scores	序号	搭配词	频率	MI 值
1	an	3087	3.28	1	most	676	3.93
2	seemingly	948	8.92	2	freedom	183	5.65
3	possibilities	550	8.16	3	legacy	159	7.39
4	stream	318	6.97	4	operation	135	5.94
5	supply	312	6.00	5	popularity	74	6.90
6	hours	290	3.77	6	appeal	72	5.49
7	list	289	4.31	7	value	66	3.56
8	series	224	3.94	8	presence	55	4.23
9	cycle	203	5.93	9	relationship	51	3.07
10	wars	169	5.56	10	pain	50	3.55
11	summer	164	3.69	11	symbol	48	5.92
12	loop	160	6.92	12	mystery	48	5.27
13	seemed	152	3.34	13	image	46	3.43
14	variety	127	4.41	14	peace	44	3.45
15	parade	126	6.41	15	influence	43	3.75
16	source	104	3.41	16	fascination	34	7.24
17	array	102	5.99	17	friendship	30	5.15
18	debate	101	3.52	18	memory	30	3.33
19	opportunities	100	4.26	19	importance	29	3.66
20	sea	95	3.51	20	contribution	27	4.94

基于前 20 个高频搭配词的统计显示: endless 的搭配词中名词占 85% (如 possibilities、stream 等), 其他词类占 15%;

KWIC 索引行深入剖析其语义倾向。这一方法体系确保了多维度对比分析的完整性。

三、研究结果与讨论

(一) 词频及语域分布差异

语域作为功能性言语变体, 产生于特定交际情境、对象与话题的差异, 并体现为不同的语体风格^[9]。通过分析近义词在不同语域中的分布特征, 可为其用法辨析提供重要依据。基于此, 表 1 呈现了 endless 和 enduring 在 COCA 语料库中各语域的具体词频分布。

enduring 则体现书面语体色彩, 二者的语域分布差异为其语用辨析提供了实证依据。

(二) 搭配词差异

根据 Firth 的观点, “一个单词的完整意义总是和上下文分不开的”^[3]。搭配研究中, 互信息值越大, 意味着两词经常一起出现, 互为搭配词^[10]。基于此, 本研究通过 COCA 语料库, 在跨距 -3 ~ +3 范围内检索 endless 和 enduring 的搭配词, 以 MI 值为指标进行对比分析, 结果见表 2。

enduring 的搭配词中名词占 95% (如 freedom、legacy 等), 副词占 5%。

语义分析表明: enduring搭配词多具积极语义色彩(如 freedom、legacy)，呈现显著正面评价倾向；endless则主要与中性词(如 stream、series)及消极词(如 loop、debate)共现，呈现中性偏消极的语义特征。

(三) 语义韵分析

语义韵研究源于 Sinclair 的早期探索，后经 Louw 确立为学术概念，指“某些词项经常和具有某种语义特征的语言单位共现而产生的一种语义特征”^[4]。Stubbs 将其区分为积极、消极和中性语义韵三类^[5]。

前期分析显示，enduring 常与 freedom、legacy 等积极词共现，呈现积极语义韵；endless 则多与中性词及 loop、debate 等消极词搭配，呈现中性偏消极倾向。为验证此发现，本研究设定跨距 $-3 \sim +3$ ，提取 MI 值大于 3 的显著搭配词，并对 100 条随机索引行进行人工分类统计(见表 3)。

表 3 endless 与 enduring 的语义韵分布(前 100 个搭配词)

	positive	neutral	negative	total
endless	23	62	15	100
enduring	75	15	10	100

基于索引行的人工分类统计显示：(1) endless 搭配词以中性语义为主(62 个)，辅以少量积极(23 个)与消极(15 个)语义，呈现典型中性语义韵；(2) enduring 搭配词中积极语义占绝对优势(75 个)，中性(15 个)与消极(10 个)语义占比较低。数据分析证实，该词项具有显著的积极语义韵倾向。

四、结论与启示

本研究采用语料库语言学方法，以美国当代英语语料库(COCA)为数据来源，对近义词 endless 与 enduring 展开了系统的对比分析，主要考察其在词频分布、语域特征、搭配模式及语义韵倾向四个维度的异同。研究结果显示：首先，在词频方面，endless 的整体使用频率显著高于 enduring；其次，在语域分布上，endless 明显倾向于口语及非正式语境，而 enduring 则更多出现在书面语及正式文体中；此外，在语义韵方面，endless 常与中性词汇共现，呈现中性语义韵特征，而 enduring 则倾向于与积极词汇搭配，表现出正面的情感色彩。

基于上述发现，本研究从实证角度厘清了 endless 与 enduring 在语义内涵、语用功能及使用规律方面的系统性差异，不仅为英语近义词教学提供了理论参考，也为学习者词汇深度知识的构建与语用能力的提升提出了具有操作意义的建议。

参考文献

- [1] Church K, Hanks P. 1990. Word Association Norms, Mutual Information, and Lexicography[M]. Cambridge: MIT Press.
- [2] Firth J. 1957. Modes of Meaning: Papers in Linguistic 1934–1951 [M]. London: Oxford University Press.
- [3] Firth, J. 1957. Papers in Linguistics[M]. Oxford: OUP.
- [4] Louw B. 1993. Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies[A]. In M. Baker, G. Francis & E. Tognini – Bonelli (eds.). Text and Technology: In Honour of John Sinclair[C]. Amsterdam: John Benjamins.
- [5] Stubbs, M. 1996. Text and Corpus Linguistics. Oxford: Blackwell.
- [6] Xiao, R. & T. McEnergy. 2006: 103–129. Collocation, Semantic Prosody and Near Synonymy: A Cross-linguistic Perspective[J]. Applied Linguistics.
- [7] 陆军. 2010, 33(03): 276–286+329–330. 基于语料库的英语近义词搭配行为与语义韵研究 [J]. 现代外语.
- [8] 卫乃兴. 2006: 50–54. 基于语料库的学生英语语义韵对比研究 [J]. 外语学刊.
- [9] 吴琳, 蒋虹. 2016, 16(06). 基于 COCA 语料库的近义词辨析——以 choose 和 select 为例 [J]. 湖南工业职业技术学院学报.
- [10] 赵红艳. 2017, (09). 基于 COCA 语料库探讨近义名词的辨析——以 objective, aim, goal 和 purpose 为例 [J]. 英语广场.