

扩展单位意义模型下基于COCA语料库的 近义词辨析

——以Keep和Retain为例

刘鑫, 赵祎

河北大学外国语学院, 河北 保定

收稿日期: 2025年6月24日; 录用日期: 2025年7月25日; 发布日期: 2025年8月6日

摘要

本研究基于扩展单位意义模型EUM (Extended Unit of Meaning)理论, 结合美国当代英语语料库COCA (Corpus of Contemporary American English)的丰富数据, 系统辨析了近义词“keep”与“retain”在实际语言运用中的差异。通过定量分析与定性分析相结合的研究方法, 本研究深入探讨了两者在语域分布、搭配模式、语义倾向及语义韵等方面的异同点。研究发现, 尽管两者在广义上均表达“保持”或“保留”的概念, 但在具体语境中却展现出独特的词汇特征和偏好。本研究不仅为近义词辨析提供了实证基础, 也为英语教学、翻译实践、词典编纂及跨文化交际等领域提供了参考。

关键词

扩展单位意义模型, 语料库, 近义词辨析

Analysis of Synonyms Based on COCA Corpus under the Extended Unit of Meaning

—Taking Keep and Retain as Examples

Xin Liu, Yi Zhao

School of Foreign Studies, Hebei University, Baoding Hebei

Received: Jun. 24th, 2025; accepted: Jul. 25th, 2025; published: Aug. 6th, 2025

Abstract

Based on the theory of Extended Unit of Meaning (EUM) and the rich data from the Corpus of

Contemporary American English (COCA), this study systematically analyzes the differences between the synonyms “keep” and “retain” in actual language usage. Through a combination of quantitative and qualitative analysis, this study examines the similarities and differences between the two in terms of register distribution, collocation patterns, semantic tendency and semantic rhyme. It is found that although both of them express the concept of “keep” or “retain” in a broad sense, they show unique lexical features and preferences in specific contexts. This study not only provides an empirical basis for the identification of synonyms, but also provides references for the fields of English teaching, translation practice, lexicography and cross-cultural communication.

Keywords

Extended Unit of Meaning, Corpus, Synonym Analysis

Copyright © 2025 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

近义词是自然语言处理中一个重要的研究领域，它不仅涉及词汇语义学的核心问题，也是语言理解和生成的关键环节。在英语学习过程中，正确理解和运用近义词对提高语言表达的准确性和丰富性至关重要。然而，由于英语近义词在形式上相似，但使用场合、语域分布及搭配关系等方面可能存在细微差别，这对学习者构成了挑战。因此，如何系统地研究近义词之间的差异，并将研究成果应用于教学实践中，成为了一个值得深入探讨的话题。

在这一背景下，扩展单位意义模型(Extended Unit of Meaning, EUM)作为一种理论框架被提出，该模型强调从更宽泛的语言单位出发来考察词语的意义，超越了传统意义上对单个词汇项的关注，而是考虑到了词汇在具体语境中的使用情况以及与其他语言成分的相互作用[1]。EUM认为，理解一个词语的意义需要考虑其在句子乃至更大范围内的功能和角色，强调词汇意义不仅限于其词典定义，还包括其在实际语言使用中的搭配、语义倾向和语义韵等复杂特征。这为近义词的研究提供了新的视角。

随着计算语言学的发展，大规模语料库的应用为语言研究带来了前所未有的机遇。美国当代英语语料库(Corpus of Contemporary American English, COCA)作为目前最大的免费在线英语语料库之一，包含了从1990年代至今的各类文本类型，包括报纸杂志、书籍、学术论文等，为语言研究提供了丰富的数据支持[2]。利用COCA进行近义词辨析，可以从真实语境中获取大量实例，进而实现对词语用法的精确分析。

鉴于上述背景，本研究选择“keep”与“retain”这一对常见近义词作为研究对象，采用语料库的方法，结合COCA语料库以及定性、定量分析手段，旨在通过扩展单位意义模型的理论框架，结合COCA语料库的丰富数据，深入辨析这两个词在实际语言使用中的细微差异。

本研究将围绕以下几个核心问题展开：1) 在扩展单位意义模型的框架下，“keep”与“retain”在语域分布上有何异同？2) 这两个词在搭配词方面有哪些显著区别？3) “keep”与“retain”在语义倾向和语义韵上分别呈现出怎样的特征？通过解决这些问题，本研究将全面揭示“keep”与“retain”在实际语言使用中的差异，为英语教学和语言研究提供有价值的参考。

2. 文献综述

2.1. 近义词辨析研究概述

近义词辨析一直是语言学的重要课题之一。它不仅是语言学习的基础,也是语言理解和表达的关键。传统的近义词辨析方法主要依赖于词典释义和语法规则,通过比较词语的定义、词性、语义范围等方面来区分近义词[3]。早期的近义词研究主要关注词汇的字面含义和基本语义特征,如 Ogden & Richards 提出的“基本词汇表”。然而,随着语言学理论的发展,学者们开始意识到单纯依靠字典定义无法完全解释近义词之间的差异。例如, Firth (1957)提出了“上下文原则”,强调词汇的意义是由其使用的语境决定的[4]。

近年来,随着计算机技术的进步和语料库语言学的发展,越来越多的研究者开始借助大规模语料库来探究近义词的使用差异。其中, Sinclair 等人提出的“语料库驱动的词汇语义学”方法尤为突出。这种方法主张通过观察词汇在真实语境中的使用情况来确定其意义和用法,从而弥补了传统词典定义的不足。这种方法通过对大量真实语料的分析,考察近义词在不同语境下的使用频率、搭配情况、语义韵等特征,从而更加准确地揭示近义词之间的差异。

2.2. 扩展单位意义模型(EUM)的理论背景与应用

扩展单位意义模型(Extended Unit of Meaning, EUM)作为词汇语义学的重要理论框架,由 Sinclair 等学者提出并发展[5]。该模型强调词汇的意义不仅仅局限于其词典中的定义,而是包含了一系列在实际语言使用中与词汇共现的语言特征,如搭配词、语义倾向和语义韵等。这些特征共同构成了词汇的扩展单位意义,为理解词汇的复杂性和多样性提供了更为全面的视角。

近年来, EUM 在词汇教学、词典编纂及自然语言处理等领域得到了广泛应用。在词汇教学方面, EUM 帮助教师和学生认识到词汇学习不仅仅是记忆单词的中文意思,更重要的是掌握词汇在不同语境中的具体用法和搭配[6]。在词典编纂方面, EUM 促使编纂者更加注重词汇的语境信息和实际用法,以提高词典的实用性和准确性[7]。

2.3. COCA 语料库在近义词辨析中的应用

美国当代英语语料库(Corpus of Contemporary American English, COCA)作为世界上最大的平衡语料库之一,它是由美国 Brigham Young University 的 Mark Davies 教授开发的高达 5.2 亿词汇的美国最新当代英语语料库,是当今世界上最大的英语平衡语料库[8]。它包含了从 1990 年至今的多种类型的美国英语文本,为近义词辨析提供了丰富的数据资源。COCA 语料库不仅规模庞大,而且数据更新及时,能够反映当代美国英语的实际使用情况。

近年来,越来越多的学者开始利用 COCA 语料库进行近义词辨析研究。他们通过检索和分析近义词在语料库中的使用实例,结合定量和定性的方法,深入剖析近义词在语域分布、搭配词、语义倾向和语义韵等方面的异同[9]。赖康生(2022)从词频、语域分布、搭配、语义韵 4 方面辨析了 abolish、cancel 和 repeal [10], 迟红丹(2023)对 extend 和 expand 从语域、搭配和语义韵三方面进行了对比研究[11]。这些研究不仅揭示了近义词之间的细微差别,还为语言教学和词典编纂提供了实证支持。

2.4. “Keep”与“Retain”的研究现状

尽管“keep”和“retain”在英语中被广泛认为是近义词,但它们在实际使用中存在一定的差异。已有研究多关注于这两个词汇的基本含义和用法,如 Kirkpatrick (2011)对“keep”的多义性进行了探讨[12], 而 Biber, D., Johansson, S., Leech, G., et al. (2002)在 COBUILD 语料库的基础上分析了“retain”的常见搭

配[13]。然而, 这些研究并未充分考虑这两个词汇在不同语境下的细微差别, 特别是从扩展单位意义模型的角度出发进行系统性的对比分析。

2.5. 现有研究的不足与本研究的切入点

尽管已有学者在 EUM 框架下利用语料库进行近义词辨析研究, 但针对“keep”与“retain”这一对近义词的系统性辨析仍显不足。特别是缺乏在 EUM 理论框架下, 结合 COCA 语料库的丰富数据, 对这两个词在语域分布、搭配词、语义倾向和语义韵等方面进行全面而深入的探讨。因此, 本研究旨在填补这一空白, 通过定量分析与定性分析相结合的方法, 为“keep”与“retain”的辨析提供更为全面和准确的实证支持。

3. 研究方法

3.1. 研究设计

本研究采用混合研究方法, 即定量分析与定性分析相结合的方式, 以期全面地探讨“keep”与“retain”这两个近义词在实际语言使用中的差异。定量分析侧重于统计两词在 COCA 语料库中的出现频率、搭配词以及语域分布等量化指标; 定性分析则通过具体语境下的例子来深入理解两者在语义倾向、语义韵等方面的细微差别。

3.2. 研究框架

本研究基于扩展单位意义模型(EUM)的理论框架, 结合美国当代英语语料库(COCA)的数据资源, 对近义词“keep”与“retain”进行深入辨析。研究框架主要包括数据收集、数据处理、数据分析和结果解读四个步骤。通过这一框架, 本研究旨在系统地探讨“keep”与“retain”在语域分布、搭配词、语义倾向和语义韵等方面的异同。

3.2.1. 数据收集

本研究的数据主要来源于美国当代英语语料库(COCA)。COCA 是一个大型的平衡语料库, 包含了从 1990 年至今的多种类型的美国英语文本, 如小说、新闻、杂志、学术文章等, 为近义词辨析提供了丰富的语境信息。在数据收集过程中, 作者分别从 List (列表)、Chart (图表)、Collocates (搭配)、Compare (比较)和 KWIC (语境关键词)检索了“keep”和“retain”两个词在 COCA 语料库中的全部使用实例, 并下载了相关文本数据。

3.2.2. 数据处理

在数据处理阶段, 为了确保分析的有效性和准确性, 研究首先对数据进行清洗和标注, 去除重复的文本片段、不完整的句子以及与研究目标无关的内容。鉴于 COCA 语料库的庞大规模, 为了提高分析效率并确保结果的代表性, 本研究采用了随机抽样的方法, 从“keep”和“retain”的广泛使用实例中分别抽取了具有语域和文体代表性的样本进行分析。最后, 根据扩展单位意义模型(EUM)的核心理念, 本研究构建了以“keep”和“retain”为核心的分析单元, 深入探索了它们的搭配词等关键语言特征。

3.2.3. 数据分析

在数据分析阶段, 本研究采用定量分析与定性分析相结合的方法。在定量分析方面, 本研究主要运用统计软件对收集到的数据进行处理和分析, 包括计算词频、共现频率、搭配强度等指标。在定性分析方面, 本研究则通过人工阅读和分析例句, 结合语言学理论和常识判断, 对定量分析结果进行验证和补充。这种定量与定性相结合的分析方法, 既能够确保分析结果的客观性和准确性, 又能够深入挖掘词汇意义的细微差别和内在规律。

3.2.4. 结果解读

在数据分析的基础上, 本研究将对“keep”与“retain”在语域分布、搭配词、语义倾向和语义韵等方面的异同进行解读。通过对比两者的异同点, 揭示其在实际语言使用中的微妙差异, 为英语教学提供实证支持。同时, 本研究还将讨论研究结果对词汇教学、词典编纂等领域的启示和应用价值。

4. Keep 和 Retain 的词义辨析

4.1. Keep 和 Retain 在不同语域的词频分布

语域是指语言使用的场合或领域的总称, 是语言变体按照使用的情况进行划分的结果。它代表了一种具有某种具体用途的语言变体, 这些变体因社会情景的不同而有所区别。近义词因内在语义差异而在不同语域中频率各异, 统计其分布差异有助于为英语学习者选择恰当词汇提供宏观指导, 促进语言教学的精准性[11]。在 COCA 语料库的 Chart 模块分别检索, 得出 keep 和 retain 这 2 个词在不同语域的词频分布。

如表 1 中统计结果显示: 在 COCA 语料库的数据统计中, keep 一词的出现频次显著高于 retain, 具体表现为 keep 的频数达到 400,910 次, 而 retain 则仅为 11,800 次。这一数据差异揭示了在实际语言应用中, keep 相较于 retain 享有更高的使用偏好, 表明语言使用者更倾向于选择 keep 来表达相关概念或语境需求[11]。其次, keep 一词在八大体裁中, 在电视电影当中的占比最高, 占 21.60%; 在博客中占 15.05%, 可见该词在口语和网络用词当中非常常见, 说明这个单词的使用场合比较随意; 再次, 学术文章当中的使用频次在八大体裁中占比最低, 说明该词在书面语体裁当中的使用频次较低; 最后, keep 一词在新闻当中的占比较低, 占 9.78%, 再次佐证了该词多使用于非正式语体。

Retain 一词在八大体裁中, 在学术文章当中的占比最高, 占 24.67%, 说明该词在书面语体裁当中的使用频次较高, 这反映出该词在表达保留、维持特定状态或权利时的专业性和精确性; 其次是在杂志、网站、新闻中, 分别为 17.26%、17.10%、16.02%, 这可能与网络环境下信息传播的准确性和专业性要求有关。再次, 口语、小说、电视电影当中的使用频次在八大体裁中占比较低, 分别为 4.66%、4.01%、2.19%。再次佐证了该词多使用于正式语体; 为英语学习者在不同文体写作当中使用该词提供了一定的参考。

Table 1. The frequency distribution of “Keep” and “Retain” in different registers of COCA

表 1. Keep 和 Retain 在 COCA 不同语域中的词频分布

Section	Keep		Retain	
	FREQ	PCT	FREQ	PCT
博客	60,335	15.05%	1663	14.09%
网站	52,239	13.03%	2018	17.10%
电视/电影	86,524	21.60%	258	2.19%
口语	46,859	11.69%	550	4.66%
小说	48,797	12.17%	473	4.01%
杂志	52,874	13.18%	2037	17.26%
新闻	39,236	9.78%	1890	16.02%
学术文章	14,046	3.50%	2911	24.67%
All	400,910	100%	11,800	100%

如表 2 所示, 从历时发展的角度来看, 单词 keep 随时间的发展, 其词频 1990~2009 呈稳步上升状态, 在 2010~2019 年呈现下降趋势。retain 的使用频率在 1990 年至 2019 年间整体呈现下降趋势。

Table 2. Changes in the annual usage data of “Keep” and “Retain” over the years

表 2. Keep 和 Retain 历年使用数据变化情况

Section	Keep		Retain	
	FREQ	PER MIL	FREQ	PER MIL
1990~1994	44,063	363.84	1585	13.09
1995~1999	48,694	388.90	1468	11.72
2000~2004	49,166	394.51	1417	11.37
2005~2009	50,368	409.33	1254	10.19
2010~2014	49,825	403.94	1222	9.91
2015~2019	46,220	376.53	1173	9.56
All	400,910	403.73	11,800	11.88

4.2. Keep 和 Retain 在搭配词方面的异同

研究高频搭配词的语义特征是考察语义韵的重要依据。首先, 在 COCA 语料库的 Word 界面分别输入两个同义词 keep 和 retain, COCA 网站自动选取跨距-4/+4 之间的搭配词, 按照搭配频率列出前 15 个显著搭配词, 进一步考察两个同义词在搭配词方面的异同。

按照词性 keep 的搭配词进行分类(如表 3 所示), 可细分为动词性搭配, 如 try, shut 等; 名词性搭配,

Table 3. The top 15 collocations of “Keep” and “Retain” in the COCA corpus

表 3. Keep 和 Retain 在 COCA 语料库中前 15 位搭配词

S/N	Keep		Retain	
	Word	FREQ	Word	FREQ
1	try	17,839	still	857
2	eye	16,982	control	715
3	mind	15,456	right	606
4	alive	8642	power	533
5	track	7211	message	515
6	safe	6140	attract	384
7	secret	5990	medical	341
8	quiet	3629	ability	333
9	mouth	3470	information	324
10	warm	3260	original	258
11	shut	3237	recruit	229
12	promise	2950	moisture	221
13	busy	2896	position	217
14	clean	2820	memory	216
15	pace	2779	employee	215

包括普通名词搭配, 如 eye, mind, track, secret, mouth, pace 等; 抽象名词搭配: promise。形容词性搭配如 alive, safe, quiet, warm, busy, clean; 通过具体分析发现, eye, mind, mouth, pace 等词都与人的身体部位或是行动类有关, 搭配词义偏主观性。而其中 alive, safe, quiet, warm, busy 等形容词性的搭配也和人的状态有关, 这与系动词所搭配的表语表达主语的身份、状态与特征这一说法相一致。另外, 通过搭配分析可知与 keep 相搭配的词汇大多属于具体词汇, 虽仍存在 promise 这类抽象词汇, 但总体上具体词汇多于抽象词汇。

按照词性对 retain 的搭配词进行分类(如表 3 所示), 可细分为动词性搭配, 如 attract, recruit 等; 名词性搭配, 包括普通名词搭配, 如 moisture, employee 等; 抽象名词搭配: control, right, power, message, ability, information, position, memory。形容词性搭配如 medical, original; 通过具体分析发现, 与 keep 相搭配的名词性词汇大多属于抽象词汇, 且词汇大多数比较正式。

总体来说, “keep” 多与具体名词搭配, 而 “retain” 更倾向于与一些抽象的名词搭配, 如 “retain one’s memory” 等。“keep” 的搭配词更加丰富多样, 而 “retain” 的搭配词相对较少这种差异表明, “keep” 在日常表达中用途更广, 而 “retain” 则更专注于描述抽象概念的保持或维护。

4.3. Keep 和 Retain 的语义倾向分析

语义倾向指的是一个词汇或短语在特定语境下所呈现出的主要特征, 这些特征可能包括褒义、贬义或中性等不同的情感色彩或评价倾向。在语言学中, 语义倾向通常被视为度量语言的重要标尺, 它反映了词汇或短语与概念之间主要内涵的偏离程度。在语言使用中, 研究者需要根据语义倾向来选择合适的词汇和短语, 以达到表达意思和情感的目的。本研究分别选取 keep 和 retain 在 COCA 语料库中前 5 位搭配词进行语义倾向分析。

如表 4 和表 5 所示, 通过分析 keep 一词的形容词、副词、动词名词搭配词可知, 该词的语义倾向有积极、中性和消极三种类型, 其中大部分为中性词, 占 65%, 其次是积极词, 占 20%, 消极词占 15%。

Table 4. Semantic tendencies of the top 5 combinatorial words of “Keep” in the COCA corpus—ADJ ADV

表 4. Keep 在 COCA 语料库中前 5 位搭配词的语义倾向表——形容词、副词

节点词	+ADJ	FREQ	语义倾向	+ADV	FREQ	语义倾向
Keep	alive	8672	积极	straight	1620	中性
	safe	6140	积极	apart	888	消极
	quiet	3629	中性	tight	541	中性
	warm	3260	积极	low	477	中性
	busy	2896	消极	upright	286	中性

Table 5. Semantic tendencies of the top 5 combinatorial words of “Keep” in the COCA corpus—VERB NOUN

表 5. Keep 在 COCA 语料库中前 5 位搭配词的语义倾向表——动词、名词

节点词	+VERB	FREQ	语义倾向	+NOUN	FREQ	语义倾向
Keep	try	17,839	积极	eye	16,982	中性
	shut	3237	中性	mind	15,456	中性
	inform	1447	中性	track	7211	中性
	struggle	1372	积极	secret	5990	中性
	lock	1099	中性	mouth	3470	中性

因此 keep 一词倾向于与中性词共现, 来保持主语状态和特征。在与形容词搭配时, 也倾向于与积极词共现, 同时也存在一些形容词如 busy 等表达出消极的语义倾向。

如表 6 和表 7 所示, 通过分析 retain 一词的形容词、副词、动词名词搭配词可知, 该词的语义倾向有积极、中性和消极三种类型, 其中绝大部分为中性词, 占 85%, 其次是积极词, 占 10%, 消极词占 5%。因此 retain 语义倾向较为中性或积极。

Table 6. Semantic tendencies of the top 5 combinatorial words of “Retain” in the COCA corpus—ADJ ADV

表 6. Retain 在 COCA 语料库中前 5 位搭配词的语义倾向表——形容词、副词

节点词	+ADJ	FREQ	语义倾向	+ADV	FREQ	语义倾向
Retain	medical	341	中性	still	857	中性
	original	258	中性	nevertheless	32	中性
	traditional	95	中性	nonetheless	31	中性
	existing	75	中性	indefinitely	16	消极
	qualified	54	积极	thereby	14	中性

Table 7. Semantic tendencies of the top 5 combinatorial words of “retain” in the COCA corpus—VERB NOUN

表 7. Retain 在 COCA 语料库中前 5 位搭配词的语义倾向表——动词、名词

节点词	+VERB	FREQ	语义倾向	+NOUN	FREQ	语义倾向
Retain	attract	384	中性	control	715	中性
	recruit	229	中性	right	606	中性
	manage	139	中性	power	533	中性
	hire	77	中性	message	515	中性
	ensure	63	积极	ability	333	中性

总体来说, 在语义倾向方面, “keep”与“retain”均表现出一定的中立性, 但在特定语境下, 它们的语义倾向也会有所不同。例如, “keep”在表达保持健康、保持联系等积极含义时, 常带有正面的语义倾向; 而在表示保持警惕、保持沉默、保持忙碌等情境下, 则可能带有中性或略带负面的意味。相比之下, “retain”由于其专业性和精确性, 在表达保留权利、保留记忆等含义时, 往往带有较为中性的语义倾向, 较少涉及情感色彩。尤其是在法律文件或学术报告中, 它常用来描述一种稳定的状态或持久的权利拥有。

4.4. Keep 和 Retain 的语义韵分析

语义韵作为语料库语言学领域近年来聚焦的核心议题之一, 其概念最初由英国语言学家 Sinclair 在 1987 年引入。语义韵指的是一种由具有相似或共通特性的表达模式重复性地与特定关键词共现所营造出的独特语义环境或氛围。这一概念依据其蕴含的语义倾向性, 可细化为四大类别: 积极语义韵、中性语义韵、消极语义韵及混合语义韵。具体而言, 当某一节点词(即核心词)的共现词汇几乎全部展现出消极语义特征时, 该节点词即被判定为具有消极语义韵; 相反, 若其共现词汇普遍呈现积极语义特征, 则归为积极语义韵。若节点词的搭配词汇在语义上既不显著偏向积极也不偏向消极, 则表现为中性语义韵。而混合语义韵则是一种更为复杂的情形, 其中节点词的搭配既包含中性词汇, 又显著融合了具有明确积极意涵与强烈消极意涵的词汇, 体现了其语义韵的多元性与复杂性[14]。本文试图借助 COCA 语料库推测

判断 KEEP 和 RETAIN 的语义韵, 并对其进行对比分析。

从高频搭配来看, 与 keep 搭配频次排名前 20 的词中, “中性词” 13 个, 占 65%, “积极词” 4 个, 占 20%, “消极词” 3 个, 占 15%; 与 retain 搭配频次排名前 20 的词中, “积极词” 2 个, 占 10%, “消极词” 1 个, 占 5% “消极词” 和 “积极词” 各 1 个, 各占 5%, “中性词” 17 个, 占 90%。

在语义韵层面, 通过分析发现, “keep” 总体来说多为中性语义韵, 在不同语境下可表现出丰富的语义韵特征。例如, 在表达保持忠诚、保持诚实等积极品质时, “keep” 呈现出褒义的语义韵; 而在表达保持沉默以隐瞒真相、保持现状以逃避改变等消极行为时, 则可能带有贬义的语义韵。相比之下, “retain” 由于其专业性和客观性, 更多地呈现出中立的语义韵特征, 较少涉及情感色彩强烈的褒贬评价。

5. 启示

5.1. 语言教育学的视角

从语言教育学的视角审视, 本研究对近义词 “keep” 与 “retain” 的深入辨析为现代词汇教学开辟了新的路径, 并提供了坚实的实证基础。本研究强调, 近义词辨析不仅是语言习得过程中的关键环节, 更是提升学生语言精准度与表达能力的有效策略。通过引入扩展单位意义模型(EUM), 结合语料库语言学方法(基于 COCA 语料库的分析), 本研究为教育工作者构建了一个系统化、数据驱动的词义教学框架。

在此框架下, 教师可设计多维度的教学活动, 包括但不限于基于真实语料的案例分析、语境敏感的词义辨析练习以及语义倾向与语义韵的探讨。这些活动旨在引导学生深入理解 “keep” 与 “retain” 在不同语域中的细微差别, 进而促进其词汇网络的丰富化与精细化构建, 提升其在复杂语境中的语言运用能力与适应性。

此外, 本研究呼吁教育界重视语料库资源在词汇教学中的应用, 倡导以数据为驱动的教学模式, 以增强学生的自主学习能力、语言分析能力和批判性思维能力, 为培养具有国际视野和跨文化交际能力的高素质人才奠定坚实基础。

5.2. 翻译学的视角

翻译作为跨语言、跨文化的交际行为, 其核心在于准确传达原文的意义与情感。本研究通过对这两个近义词在不同语境下的使用情况进行详尽分析, 揭示了其在新闻、杂志、博客、口语、学术文章等不同语域中的特定用法与搭配习惯, 为翻译工作者提供了精准选词、确保译文准确性与专业性的科学依据。

5.3. 词典编纂者的视角

对于词典编纂者而言, 本研究提供了关于 “keep” 和 “retain” 等近义词在实际使用中具体差异的详尽分析。这种基于大规模语料库的实证研究能够帮助编纂者更准确地定义词汇的含义、用法和语境限制, 从而在词典中给出更加精细和准确的释义和例句。这不仅可以提高词典的实用性和权威性, 还能帮助读者更好地理解和使用这些词汇。

5.4. 跨文化交际的视角

在跨文化交际领域, 本研究通过对 “keep” 与 “retain” 的辨析, 揭示了不同文化背景下词汇使用的差异性与共性, 为跨文化交际者提供了深刻的洞见与启示。跨文化交际的成功与否, 很大程度上取决于交际双方能否准确理解并尊重对方的文化背景与语言习惯。本研究指出, 在跨文化交际过程中, 交际者应具备高度的文化敏感性 with 语言意识, 避免因词汇使用不当而引起的误解或冲突。

6. 结语

6.1. 研究成果与讨论

本文在扩展单位意义模型(EUM)的理论框架下,借助美国当代英语语料库(COCA)的丰富数据,对近义词“keep”与“retain”进行了详尽而系统的辨析。通过多维度的分析,包括语域分布、搭配模式、语义倾向及语义韵等方面,本文揭示了两者在实际语言运用中的细微差别和独特特征。

在语域分布上,“keep”的使用频率远高于“retain”。“keep”在口头语和非正式语体中的高频率使用,而“retain”则更多见于学术语域等正式场合。

从搭配词角度来看,“keep”和“retain”都能与表示状态、品质或物品名词搭配,但是“keep”的搭配范围更广,涵盖了更多具体的名词。而“retain”则更多与抽象名词相匹配。

就语义倾向和语义韵而言,“keep”的含义较为宽泛,既可表达积极意义和中立意义,也可带有消极色彩。相比之下,“retain”的语义更为正式和专业,倾向于中性或积极意义,较少带有消极情感和贬义语义韵。这一发现对于翻译实践和跨文化交际尤为重要,因为准确理解和传达词汇的细微语义差别,是确保信息传递准确性和有效性的关键。

6.2. 研究意义与局限性

本研究通过对“keep”与“retain”的全面分析,不仅加深了对这两个近义词的理解,也为英语教学提供了实证支持。然而,由于COCA语料库本身的特点及其数据收集的时间跨度,本研究的结果可能存在一定的局限性。例如,语料库中的数据可能无法完全反映所有英语变体中的词汇使用情况。因此,未来的研究可以通过扩大语料库规模或引入其他语料库来进一步验证这些发现。

6.3. 总结

综上所述,本研究不仅为近义词辨析领域提供了新的视角和实证基础,也为英语教学、翻译实践、词典编纂及跨文化交际等多个领域提供了有价值的参考。未来研究可以进一步探索其他近义词在类似框架下的辨析,以丰富研究者对语言细微差异的理解和应用。同时,随着语料库技术和计算语言学的不断发展,期待未来有更多创新性的方法和技术手段应用于近义词辨析研究中,以推动该领域的持续进步。

参考文献

- [1] Sinclair, J. (1991) *Corpus, Concordance, Collocation*. Oxford University Press, 57.
- [2] Davies, M. (2024) *The Corpus of Contemporary American English: 450 Million Words, 1990-Present*. <http://corpus.byu.edu/coca/>
- [3] 王力. 中国现代语法[M]. 北京: 北京联合出版公司, 2019: 25.
- [4] Firth, J.R. (1957) *Syncretism and Ambiguity*. In: *Selected Writings*, Oxford University Press, 13.
- [5] Stubbs, M. (2001) *Words and Phrases: Corpus Studies of Lexical Semantics*. Blackwell Publishing, 43.
- [6] Nation, I.S.P. (2001) *Learning Vocabulary in Another Language*. Cambridge University Press. <https://doi.org/10.1017/cbo9781139524759>
- [7] Hunston, S. and Francis, G. (2000) *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. John Benjamins. <https://doi.org/10.1075/sc1.4>
- [8] 赵红艳. 基于 COCA 语料库探讨近义词的辨析——以 Objective, Aim, Goal 和 Purpose 为例[J]. 英语广场, 2017(9): 19-22.
- [9] Biber, D., Conrad, S. and Reppen, R. (1998) *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511804489>
- [10] 赖康生. 基于 COCA 语料库的英语近义词辨析: 以 Abolish、Cancel 和 Repeal 为例[J]. 萍乡学院学报, 2022(1):

84-88.

- [11] 迟红丹. 基于语料库 COCA 的英语近义词辨析研究[J]. 对外经贸, 2023(4): 64-68.
- [12] Kirkpatrick, A. (2011) *English as an Asian Language: Asian English and Asian English Language Education*. Routledge.
- [13] Biber, D., Johansson, S., Leech, G., *et al.* (2002) The Longman Corpus Network: An Overview. *International Journal of Corpus Linguistics*, **7**, 177-219.
- [14] 王瑞, 姜雪. 基于语料库的 Complete、Finish、Perform 语义韵对比研究[J]. 外语电化教学, 2016(5): 29-33.