

扩展意义单位分析的应用研究：以“Identity”为例

田靖潇

浙江工商大学外国语学院, 浙江 杭州

收稿日期: 2026年5月10日; 录用日期: 2026年6月17日; 发布日期: 2026年6月29日

摘要

本研究以“identity”一词为例, 基于John Sinclair扩展意义单位分析模型, 利用COCA语料库和WordSmith工具, 系统分析该词的语义特征及用法规律。本文梳理了扩展意义单位分析理论, 包括意核、搭配、类联接、语义倾向和语义韵等关键要素。然后在该理论指导下探索了“identity”的搭配结构, 通过分析其搭配、类联接、语义倾向和语义韵, 揭示了“identity”在不同语境中的多义性和使用特征, 结果显示, identity常与(身份的)构建或归属共现时多含积极语义韵, 在表明揭示或隐藏身份的语境中呈现消极语义韵。研究表明, 扩展意义单位分析为深入理解词汇意义和语用特征提供了有效路径。

关键词

Identity, 扩展意义单位分析, 搭配, COCA

An Applied Study of Extended Units of Meaning: Taking “Identity” as an Example

Jingxiao Tian

School of Foreign Languages, Zhejiang Gongshang University, Hangzhou Zhejiang

Received: May 10, 2026; accepted: June 17, 2026; published: June 29, 2026

Abstract

Taking “identity” as an example, this study systematically analyzes its semantic features and usage patterns based on John Sinclair’s Extended Units of Meaning (EUM), using the Corpus of Contemporary American English (COCA) and WordSmith tools. This paper first outlines the theory of EUM, including its key elements: core, collocation, colligation, semantic preference, and semantic prosody. Guided by this theoretical framework, this study then explores the collocational structures of

“identity”. By analyzing its collocations, colligations, semantic preferences and semantic prosodies, the study reveals how “identity” exhibits polysemy and distinctive usage patterns across different contexts. The results indicate that “identity” tends to carry a positive semantic prosody when co-occurring with notions of construction or belonging, whereas it exhibits a negative semantic prosody in contexts implying the revelation or concealment of identity. The study demonstrates that EUM provides an effective approach to understanding lexical meaning and pragmatic features.

Keywords

Identity, Extended Units of Meaning, Collocate, COCA

Copyright © 2026 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

词汇是语言研究中的重要单位，它承载着丰富的语义信息与话语功能。传统的词汇意义研究往往侧重于单词的字面释义，但词汇的意义往往依赖于其上下文、搭配模式和语法结构。因此，扩展意义单位(Extended Units of Meaning)分析理论为我们提供了一种更为系统的理论框架，能够全面地揭示词汇的语境意义。该方法不仅有助于深入理解词汇的内在语义特征，还能更精确地描述词语在实际语境中的使用规律。

在众多具有复杂语义内涵的词汇中，identity 一词具有多重含义和重要的社会文化功能。它不仅是社会学和心理学研究中的核心概念，也是语言学研究中的一个重要词汇。Identity 的语义特征在不同语境中展现出多义性，涵盖了从个体认同到集体归属等多个层面。因此，基于语料库的大规模定量分析对 identity 的意义探讨尤为重要，它能够提供更客观的证据，帮助揭示其在真实语料中的具体语义倾向和搭配规律。

尽管现有研究已经探讨了 identity 的多义性和社会语用功能，但大多数研究仍依赖小规模文本分析，且缺乏量化分析和语料库方法的支持。为填补这一研究空白，本文将结合扩展意义单位分析理论与 COCA (Corpus of Contemporary American English) 语料库，系统分析 identity 的搭配、类联接、语义倾向和语义韵。具体而言，本研究提出以下问题：

- 1) 在 COCA 语料库中，identity 的典型搭配和类联接型式是什么，呈现出何种特性？
- 2) 这些搭配与类联接体现了 identity 怎样的语义倾向与语义韵？
- 3) 在不同语境中，identity 的扩展意义单位分析是如何展现其多义性的？

基于此，本研究将运用扩展意义单位分析方法对 identity 的词汇意义进行分析和探讨。通过揭示其在不同文本、语境中的运用，探讨其意义的扩展路径、搭配运用等。本研究将首先系统梳理 identity 的语义特征，然后探讨 identity 在不同文本和语境中搭配规律、语义倾向及语义韵等，呈现这一词汇的丰富内涵和准确运用。

2. 文献综述

2.1. “Identity” 的概念与相关研究

“identity” 一词在学术讨论中承载着丰富的内涵，涉及到哲学上的“同一性”、社会心理学中的“身

份认同”及日常语言中的“身份特征”等多重意义项。这一概念横跨多个学科，其意义并非固定不变，而是在不同理论传统和研究方法中被不断重新界定，因此，梳理既有研究对该词的理解方式对本研究的实证分析定位理论坐标。

在社会心理学领域，Tajfel 提出的社会认同理论将身份界定为个体从群体归属中获得的自我认知，强调群体分类对身份形成的核心作用[1]。Giddens 在现代性背景下提出“自我认同”概念，突出身份的反思性与动态建构特征[2]。而在语言学领域，“identity”常被用作话语分析的考察对象，Wodak 等人通过话语-历时维度考察了身份在政治与媒体话语中的建构、协商与争夺[3]。Fairclough 指出，话语不仅是现实的反映，更是社会身份生成的积极场所，强调语言使用与身份建构的内在联系[4]。Oteíza 从系统功能语言学的评价理论出发，探讨语言如何建构说话者的身份定位及与受众的亲合关系[5]。这些研究为理解 identity 的话语功能提供了丰富的定性观察，但多数分析基于小规模文本或典型示例，其结论的可重复性仍有待进一步检验。

传统的语义学虽已经认识到语义学所存在的不足，却未能发展出系统描写语境意义的方法工具。这一方法论缺口为语料库语言学与扩展意义单位提供了连接窗口。国内学者也不断在 EUM 框架下开展新的方法论探索，卫乃兴系统探讨了语料库驱动的词语搭配研究方法[6]。李文中深入讨论了 EUM 中的关键问题与解决方案[7]；濮建忠对 EUM 的核心成分进行了重新审视，增强了该模型的可操作性[8]。然而，直接以 identity 为核心词进行的实证研究仍属少见，基于此，本研究尝试运用 EUM 框架，对 identity 在真实语篇中的搭配型式、语义倾向与语义韵进行描写。

2.2. 扩展意义单位

意义单位的界定是语义学与语料库语言学期长期关注的核心议题。传统词汇语义研究主要以“词”为基本意义单位，依赖词典释义来解释其语义特征。然而，由于词汇本身的多义性与语境依赖性，单纯依靠词典释义往往无法充分捕捉词语在实际语言使用中的动态意义[9] [10]。随着语料库语言学的发展，研究者逐渐认识到，词汇意义必须通过其典型语境、共现规律以及语法环境来系统呈现。

在这一背景下，Sinclair 提出的“共选理论”(co-selection)和“扩展意义单位”(Extended Unit of Meaning, EUM)框架，为词义研究提供了突破性的理论路径。Sinclair 指出，词义不是独立存在的，而是由其共现的语言特征共同形成。他进一步提出扩展意义单位包括五个核心要素：意核(core)、搭配(collocation)、类联接(colligation)、语义倾向(semantic preference)和语义韵(semantic prosody)。这一框架促使语义分析从静态走向动态，从词典语义走向语境化语义[11] [12]。

意核是扩展意义单位分析的核心和起点，是意义单位分析必不可少的成分。Sinclair 将意核定义为“意核是词项结构的必要成分之一，由一个或多个不变的词语或只受一定语法变化影响的词语所组成” [13]。濮建忠认为，从任何词出发都有可能找到可扩展的意义单位，因此以任何词为节点词寻找扩展意义单位都值得尝试[8]。

搭配是意核周边所形成的词汇型式，观察搭配是寻找扩展意义单位的第二步[8]。Firth 将搭配定义为“搭配是习惯性结伴的具体词语”，他在提出“欲知其词，先观其伴”(You shall know a word by the company it keeps)时就强调了搭配的重要性[14]。而后，Sinclair 将搭配定义为“搭配是两个或多个词在文本中短距内的共现” [11]。这一概念为更科学和精确的统计和研究提供了参考，是研究意义单位重要的一步，也为后续类联接研究做出重要铺垫。

类联接是意核周边所形成的语法型式，Firth 首先提出类联接的概念为“在句法结构中语法范畴间的相互关系”[14]。后 Sinclair 将其重新定义为“类联接是一个语法类或结构型式与其他语法类或结构型式，或其与某词或短语的共现” [13]。濮建忠建议，暂且将扩展意义单位的左边界定位在动词或形容词上，一

般不包括其前的情态动词、副词或系动词等；将右边界定位在与其密切结合的名词词组或介词词组的中心名词上[8]。

语义倾向是意核周边所形成的语义集，Sinclair 将其定义为“有时在短语结构中可能存在一种选择具有特定意义词语的清晰倾向。词类并不重要，任何具有适当意义的词语都可以”[13]。也就是说，在分析一个扩展意义单位时，一些实词的词性不同，但它们仍然可以在语义上归纳为这个节点词的同义语义倾向。因此语义倾向也就是在扩展意义单位的分析过程中，将节点词周围的词在意义上归为同一个语义集。

语义韵是整个意义单位(复合词项)所传达的态度意义，寻找语义韵是探索扩展意义单位分析的最后一步[15]。Sinclair 并没有对语义韵提出明确的定义，且学界针对语义韵持续存在着争议和讨论，但不可否认的是，语义韵并非从某个节点词或搭配词总结出来的，而是针对整个扩展意义单位或复合词项而言的，是整体的态度，可以是一般性的宏观概括，也可以是具体的态度归纳[7]。

3. 研究设计

本研究运用 COCA (Corpus of Contemporary American English) 语料库和 WordSmith8.0 工具，选择以 identity 为节点词来进行分析。选择该词的原因是该词是一个常用的单词，有丰富的使用形式和语义内涵。检索结果涵盖新闻、小说、杂志、学术等多种体裁，以确保语域分布的广度。根据索引结果，节点词共出现 54,504 次。由于 COCA 规模较大，为保证可操作性，本研究在分析总结过程中随机抽样 200 条节点词索引行进行分析，每条上下文包含±30 字符左右(按 COCA 默认设置)，确保样本量兼具代表性与可处理性。检索设置均为不区分大小写(case-insensitive)，并排除标题、元数据等非正文文本，以确保所取数据的语言结构准确反映自然使用。为揭示节点词 identity 的核心搭配特征，本研究设置搭配词跨度为±4 词，该跨度选择遵循 Sinclair (1991) 的建议，既能获取核心意义单位，又避免跨度过大导致搭配不强。

4. 结果和讨论

4.1. 语域分布

Table 1. The register distribution of identity in COCA
表 1. Identity 在 COCA 语料库中的语域分布

SECTION	BLOG	WEB	TV/M	SPO	FIC	MAG	NEWS	ACAD	ALL
FREQ	5309	7030	2149	3520	2244	5594	4592	24,066	54,504
WORDS (M)	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	993
PER MIL	41.28	56.58	16.78	27.91	18.97	44.36	37.72	200.90	54.89
PERC	9.74%	12.90%	3.94%	6.46%	4.12%	10.26%	8.43%	44.15%	100%

从 COCA 统计结果来看(见表 1)，identity 一词的语域分布差异较大。在学术用语中的标准化频率最高(PER MIL = 200.90)，远高于其他类型文本，这表明该词在社会科学和人文学科的理论化讨论中具有核心地位，常用于描述主体建构、文化身份、群体差异等抽象概念。相比之下，杂志、博客和新闻等媒体语域也呈现出较高的使用频率(约 37~56)，这表明 identity 在公共话语中同样重要，尤其在社会议题、文化评论与价值讨论中具有较强的能见度。

相比之下，小说、影视文本和口语等生活化语域中，identity 的频率明显较低(约 16~28)，这与其抽象名词性质相关：叙事和对话更倾向于通过具体情节和人物行为呈现身份相关内容，而较少直接使用概念性词汇。

4.2. 扩展意义单位分析

在 COCA 语料库中检索 *identity* 作为目标索引词来进行分析, 然后根据索引行来更进一步具体分析其搭配、类联接、语义倾向和语义韵, 分析探讨 *identity* 的扩展意义。

4.2.1. 搭配分析

如前文所述, 搭配词对于目标词研究的重要性不言而喻, 故在此不再赘述。首先在 COCA 中设置跨距(span)为+4/-4 对目标词 *identity* 进行检索, 根据搭配频率呈现高频搭配词前 20 位, 如表 2 所示:

Table 2. The top 20 high frequent collocates of *identity* in COCA
表 2. *Identity* 在 COCA 语料库中前 20 位高频搭配词

序号	搭配词	频数
1	national	2306
2	cultural	1586
3	theft	1464
4	social	1282
5	sense	1156
6	group	994
7	personal	787
8	crisis	750
9	true	729
10	reveal	697
11	development	597
12	orientation	580
13	collective	568
14	culture	554
15	protect	553
16	define	500
17	establish	478
18	individual	477
19	formation	458
20	mistaken	428

以上高频搭配词主要可按照词性大致分布在形容词、名词和动词。形容词主要有 *national*、*cultural*、*social*、*personal*、*collective*、*individual* 等表示身份归属与差异的搭配词, 传达出个人或群体在社会、文化等对身份与认同的定位。名词主要有 *theft*、*sense*、*group*、*culture* 等表示社会关系、个体行为以及群体与个人的动态关系, 都表明个体如何在群体、社会及文化中如何定位自己的。而动词则包含了 *reveal*、*protect*、*define*、*maintain*、*establish* 等表示构建、界定和维持身份等相关, 强调个体或群体如何塑造和维护自己的身份或认同等。

观察分析 *identity* 的高频搭配词可以发现, 不同词类的搭配词在语义功能上具有相互指向的特征。形

容词搭配主要表明身份类别区分和社会分类，反映了身份在民族、文化、性别等层面的分化。名词搭配主要涉及在社会实践和互动中所产生的身份差异，展示了身份与行为、群体关系等所产生的联系。动词搭配词呈现出了语篇中身份是通过揭示、构建、维护等而产生和动态存在的。这三类主要的高频搭配词表明，identity 一词并非单一的概念，而是一个涉及社会结构、文化归属、自我群体定位等多层面的复杂语义集合。

4.2.2. 类联接分析

卫乃兴指出，一旦具体的词形被选中，其语法结构也随之选定[16]。类联接是寻找意义单位的第三步，由于索引行篇幅数量过大，无法进行一一观察，因此本研究将随机从中选取 200 条索引行，人工进行读取与归纳总结类联接，并进行统计。由于此局限，因此本研究将节点词的位置搭配词复现频数阈值设置为 2，来对各位置上出现的词进行观察。对节点词所在索引行的左侧搭配进行观察后发现，identity 左侧所呈现的类联接有以下几种结构：

- of identity (22)
- the identity (17)
- Adj. {national| cultural| social}+ identity (59)
- (his| your| their| whose| my) + identity (24)
- an + identity (9)

通过观察发现，左一位置的两大强搭配型式为 of + identity 和 the + identity。先沿着 of identity 向左分析，可以看到，其左边位置的享有共同语法特征的主要结构有 the + NOUN + of identity 和 Adj. + NOUN + of identity 用来表示某事或某物的身份特征。而 the identity 继续向左延伸可发现有以下共有语法特征的主要结构：Prep. + the identity 和 Verb + the identity。

同时在 Adj. + identity 的结构中，享有共同语法特征的主要结构有 the + NOUN + of + Adj. identity, Prep. + Adj. identity, Poss. Adj + Adj. identity 和 Prep + a/an + Adj. Identity。另外一组 possessive Adj. + identity 的结构中，享有共同语法特征的主要结构有 Verb + Prep. + Poss. Adj identity 和 Verb + Poss. Adj identity。

至此，对 identity 在所建语料库的左一位置上的搭配词中的前 4 位进行了类联接分析，所有语法共有结构已总结完成，因此可以选择停止对左侧搭配型式的分析，一共有以下八种语言型式：

- Adj. + NOUN + of identity
- Prep. + the identity
- Verb + the identity
- the + NOUN + of + Adj. identity
- Possessive Adj. + Adj. identity
- Prep. + a/an + Adj. identity
- Verb. + Prep. + Possessive Adj. identity
- Verb. + Possessive Adj. identity

以上对 identity 左侧搭配的语法型式的描写体现出在实际运用中不同型式的语法结构对其搭配结构选择的不同作用，整体 identity 在语法上偏好与形容词短语结合，突出其作为“身份”或“认同”概念的具体性，同时与动词搭配时则更多强调身份的构建和变动。

下面对 identity 的右侧结构进行观察，在右一位置上，较强搭配型式是 identity and, identity of (17)

与 identity is (12)。通过对索引行进行观察可以发现, identity 右一位置上 and 出现频率高是 identity 位于句末, 下句的句首所用的 and 作为连接词, 因此 and 在此构不成语法上的共现, 故在此不再讨论 and。继续观察 identity of 的右侧搭配结构, 通过观察可发现共享语法特征的结构为 identity of + the + NOUN 来表示一个实体的身份特征。Identity is 的右侧搭配结构所能共享语法特征的结构为 identity is + DONE 和 identity is + Adj. 来表示身份的真假性和所遭受的不同变化。虽然以上三种是右一位置上出现频率较高的搭配词, 但观察下面的较低频次搭配词可发现这些搭配词中有一个共享的语法特征也就是 identity 后常跟介词短语, 故在此将 identity + Prep 作为一个类联接来考虑。

以上为 identity 的右侧搭配进行了类联接分析, 所有搭配已分析完毕, 在右侧的搭配型式主要有:

identity of + the + NOUN

identity is DONE

identity is + Adj.

identity + Prep.

尽管 identity 右侧具有可归纳的语法型式结构, 但笔者在对索引行观察时发现, 右侧的搭配结构受文本的语境影响较大, 类联接的普遍适用性较低, 这里仅作参考。通过以上分析和总结可知, identity 的主要搭配结构在其左侧, 右侧有较为浅显的语法结构。在 identity 的各类语法搭配结构中, identity 与名词、动词、介词等都有较为紧密的结合, 这体现了 identity 在语法结构中的动态性。

4.2.3. 语义倾向分析

要了解一个词的完整语义, 对该词的语义倾向分析尤为重要, 通过掌握其搭配实词的语义特征, 可以更加完整了解该词在具体运用中更完整和全面的意义[17]。对 identity 的语义倾向主要基于上述类联接的语法结构, 再结合具体的实词来概括其语义倾向集, 从而呈现该核心词的语法和意义的共现搭配, 以下是按照前文类联接对 identity 的共享语义的结构所产生的语义倾向进行的总结, 在描写其模式时, [A] 用来表示可选单词, 这表示 A 存在或者不存在。{Sempref: A} 表示一组语义倾向。{A|B} 表示 A 或 B 的绝对存在。[A|B] 表示这两种符号的结合, 表示一种可选的词组[7] [18]:

(1) [the] + {Sempref: Adj: CULTURAL} + {Sempref: N: CONCEPTUALISATION} + of + identity 属性和建构过程
Sempref: CULTURAL: cultural, different, male, false, ongoing

Sempref: CONCEPTUALISATION: conceptualisation, application, construction, confirmation, formation, transformation

(2) [Adj] + {Sempref: N: VIEW} + Prep. + the + identity 看法与观点

Sempref: VIEW: view, allegation, declaration, judge

(3) [BE] + {Sempref: V: SEE} + [Prep] + the + identity 呈现与验证

Sempref: SEE: see, confirm, produce, get, know, alter

(4) [the] + {Sempref: N: CONSTRUCTION} + [of] + {Sempref: Adj: COLLECTIVE} + identity 具体的身份构建

Sempref: CONSTRUCTION: construction, expression, preservation, dependence

Sempref: COLLECTIVE: collective, cultural, detailed, male, female, new, personal, social, teenage, true, false,

(5) [Prep] + {Sempref: Poss. THEIR} + {Sempref: Adj: COLLECTIVE} + identity 拥有者

Sempref: THEIR: their, his, her, my, our

(6) Prep + [V.] + a/an + {Sempref: Adj: STUDENT} + identity 个体化与社会定位

Sempref: STUDENT: student, consistent, defined, important, online, original

- (7) [BE] + {Sempref: V: KEEP} + [Prep] + {Sempref: Poss. YOUR} + **identity** 变动与认同
Sempref: V: KEEP: keep, want, have, prove, maintain, disguise, change, resume, keep, reveal
- (8) **identity** + of + the + {Sempref: N: VICTIM} 特定角色与类别
Sempref: VICTIM: victim, athlete, target, penitent, witness, person
- (9) **identity** + is + {Sempref: V: REVEAL} + [Prep] 显现与隐藏
Sempref: REVEAL: reveal, subsume, achieve, threaten, withhold, define

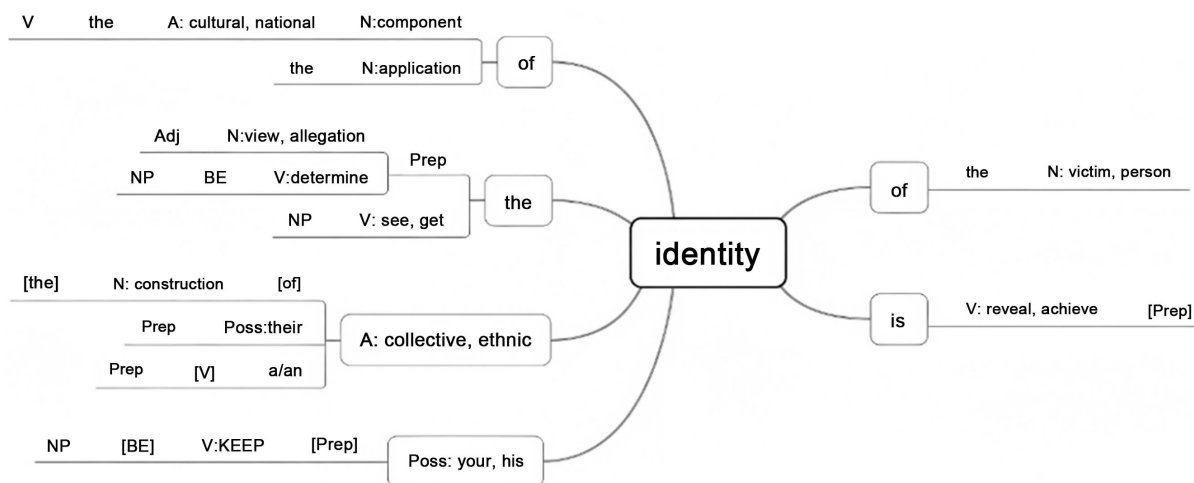


Figure 1. Path analysis diagram of “identity”
图 1. Identity 的路径分析图

以上型式可通过路径分析图直观呈现出来，从图中也可以看出，不同的搭配词选择对意义传达也有不同的作用(见图 1)。在 collective 等的搭配型式常常传达出(身份)的一致性，暗含多人群体的同一属性。而搭配其他名词如 component 等时则常常补充其身份特性，和代词搭配时常为指明具体的身份归属。总体来看，identity 在本语料中的语义倾向主要集中于一组围绕身份的建构、界定、实现与显现等过程的语义倾向，体现了身份在现代语境中所具有的可变性、社会嵌入性与互动性特征。

4.2.4. 语义韵分析

语义韵的分析并不单单是在“积极”或“消极”中做选择，而是不同的人对不同的文本会分析出不同的语义韵，且需要结合具体的文本来探索其传达出的态度或情感，本文将根据前述总结来探讨语料库中能体现态度的词语来进行总结，参考李文中的做法，通过抽取高评价性词语或特征来抽取确定语义韵 [7]。

(1) [the] + {Sempref. Adj: CULTURAL} + {Sempref: N: CONCEPTUALISATION} + of + **identity** 是非评价性或描述性的语义韵，表示身份的文化属性与认知建构，由“application”，“cultural”等词体现出。

(2) [BE] + {Sempref: V: SEE} + [Prep] + the + **identity** 是认知性的语义韵，表明对身份的获取，由“see”，“get”，“know”等词可以体现出。

(3) [the] + {Sempref: N: CONSTRUCTION} + [of] + {Sempref: Adj: COLLECTIVE} + **identity** 是积极的表示建设性与过程性的态度，表明对身份的建构与表达，由“construction”，“expression”等词可以体现出。

(4) [BE] + {Sempref: V: KEEP} + [Prep] + {Sempref: Poss. YOUR} + **identity** 是中性的包含稳态的态

度,表明身份现状的维持,由“keep”,“have”,“change”等词可以看出。

(5) **identity + of + the + {Sempref: N: VICTIM}**是中性的表示指称性的语义韵,表明特定人物身份,由“victim”,“athlete”等词可以看出。

(6) **identity + is + {Sempref: V: REVEAL} + [Prep]**是消极的态度,表明对特定身份的隐藏与威胁等,由“reveal”,“threaten”,“withhold”等词可以看出,由于在本语料中“reveal”多出现在身份暴露、隐私泄露等负面语境,因此呈现“负向语义韵”。

综上所述,“identity”在不同语境中呈现出多层次的语义韵分布,体现出该词在语义上的动态性与语境依赖性。从 cultural, conceptualisation of identity 等结构中主要呈现出其描述性的语义韵,强调身份作为一种文化与认知对象的呈现;从 see 与 know identity 等结构中体现出认知性的语义韵,指向身份的理解与获取;从 construction 与 expression of collective identity 的结构中体现出建构性的语义韵,强调身份的生成与表达过程;从 keep 与 maintain your identity 结构中体现出稳态的语义韵,表明身份的维持与自我管理;而 identity of the victim 或 athlete 结构中暗含指称性的语义韵,呈现对特定人物类别的界定;最后,带有 threaten, withhold, reveal 等高评价性动词的结构则呈现出明显的负向语义韵,指向身份被暴露、受威胁或受到外力干预的语境。

总体而言,identity 的语义韵并非单一方向,而是表明出认知、建构、指称与评价等多重语义,呈现出身份概念的多维性、过程性等,这也印证了身份研究中关于其动态性与社会嵌入性的理论共识。

5. 结语

本研究基于 John Sinclair 的扩展意义单位模型运用语料库方法对 COCA 语料库中 identity 一词进行扩展意义单位分析,揭示了抽象概念在真实的语篇中如何通过词汇搭配、语义倾向不断被建构、使用与再语境化。同时,本研究进一步验证了 Sinclair 的 EUM 框架在解释抽象名词意义生成机制方面的解释力,也验证了该框架在处理语义和语用功能时的适用性。

研究发现 identity 的意义在真实的语篇中呈现出高度的过程性与关系性特征:一方面,动词共现词显示它常被描写为可被“构建、转化、定义、确认”的动态对象;另一方面,与之搭配的主体类名词(如 person, athlete, victim 等)表明其用法常和社会角色、语境位置与互动关系共同实现。本研究也验证了语料库在语义、语用与社会文化研究中的方法优势,为今后开展更大规模、多语域或跨文化的身份研究提供了可复制的框架与启示。总体而言,本研究通过语料库驱动的方法,揭示了 identity 在现代语篇中的多维语义特征,为语言、文化与社会认同研究提供了新的证据与分析视角。

参考文献

- [1] Tajfel, H., Turner, J., Austin, W.G. and Worchel, S. (2001) An Integrative Theory of Intergroup Conflict. *Essential Readings*, 94-109.
- [2] Giddens, A. (2023) Modernity and Self-Identity. In: *Social Theory Re-Wired*, Routledge, 477-484. <https://doi.org/10.4324/9781003320609-62>
- [3] Wodak, R. (2001) The Discourse-Historical Approach. In: Wodak, R. and Meyer, M., *Methods of Critical Discourse Analysis*, Sage Publications, 63-94. <https://doi.org/10.4135/9780857028020.n4>
- [4] Fairclough, N. (2003) *Analysing Discourse: Textual Analysis for Social Research*. Routledge.
- [5] Oteíza, T. (2006) J.R. MARTIN y D. ROSE. 2003. Working with Discourse. Meaning beyond the Clause. *Estudios filológicos*, 41, 293-295. <https://doi.org/10.4067/s0071-17132006000100019>
- [6] 卫乃兴. 基于语料库和语料库驱动的词组搭配研究[J]. 当代语言学, 2002(2): 101-114+157.
- [7] 李文中. 接着做: 扩展意义单位分析[J]. 当代外语研究, 2021(6): 13-26+88+161.
- [8] 濮建忠. 扩展意义单位模型再解读[J]. 外语研究, 2020, 37(2): 1-8+112.

-
- [9] Cruse, D.A. (1986) *A Lexical Semantics*. Cambridge University Press, 1986.
- [10] Lyons, J. (1977) *Semantics*. Cambridge University Press.
- [11] Sinclair, J.M. (1991) *Corpus, Concordance, Collocation*. Oxford University Press.
- [12] Sinclair, J.M. (1996) The Search for Units of Meaning. *Textus*, 4, 75-106.
- [13] Sinclair, J.M. (2003) *Reading Concordances: An Introduction*. Pearson.
- [14] Firth, J.R. (1968) *Selected Papers of J. R. Firth, 1952-1959*. Indiana University Press.
- [15] 濮建忠. John Sinclair 的短语理论与意义研究[J]. 当代外语研究, 2021(6): 60-76+160.
- [16] 卫乃兴. 共选理论与语料库驱动的短语单位研究[J]. 解放军外国语学院学报, 2012, 35(1): 1-6+74+125.
- [17] Firth, J.R. (1957) *Papers in Linguistics 1934-1951*. Oxford University Press.
- [18] 李文中. 短语理论框架综论[J]. 外语教学与研究, 2018, 50(1): 49-59+160.