

基于语料库的英语同义词 对比研究

——以Accident和Event为例

李晨阳, 梁梅红

北京工业大学, 北京

收稿日期: 2021年9月6日; 录用日期: 2021年10月11日; 发布日期: 2021年10月21日

摘要

本文以BNC语料库为基础, 从词的分布、搭配和语义韵三个角度对event和accident进行对比分析。研究发现, event比accident使用得更加频繁, 同时event和accident在书面语言中更常用。Event具有明显的积极语义倾向, 而accident具有消极语义倾向。通过这些分析, 本文不仅探索了event和accident在真实语料库中的使用差异和语用倾向, 而且为传统英语教学提供一个新的视角。

关键词

BNC语料库, 同义词, 语域分布, 搭配, 语义韵

A Corpus-Based Contrastive Study on English Synonyms

—The Case Study of “Accident” and “Event”

Chenyang Li, Meihong Liang

Beijing University of Technology, Beijing

Received: Sep. 6th, 2021; accepted: Oct. 11th, 2021; published: Oct. 21st, 2021

Abstract

Based on BNC corpus, this paper compares and analyzes event and accident from three aspects: word distribution, collocation and semantic prosody. Studies have found that event is used more frequently than accident. Event and accident are more commonly used in written language. Event

has obvious positive semantic tendency, while accident has negative semantic tendency. Through these analyses, this paper not only explores the usage difference and pragmatic tendency of event and accident with real corpus, but provides a new perspective for traditional English teaching.

Keywords

BNC Corpus, Synonyms, Register Distribution, Collocation, Semantic Prosody

Copyright © 2021 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

由于中英文语言的差异性, 中国英语学习者在习得英语同义词的过程中产生了很大的困难, 他们无法深刻地理解同义词的差异, 并准确的运用他们。BNC 是英国国家语料库的缩写, BNC 语料库在国际上享有很高的权威性。它的语料库包括 9000 万字的书面文本和 1000 万字的口语文本, 总容量超过 1 亿字。因此, 基于 BNC 语料库的词汇或短语研究近年来非常流行, 特别是基于 BNC 语料库对两个词的比较研究[1] [2], 如“capable”和“competent” [3], “accept”和“receive” [4]和“lead to”和“result in” [5]等。此外, 卫乃兴、李晓红研究了英汉语义偏好和语义韵在“实现两种语言对应词汇项的对等”中的作用[6] [7]。Richard Xiao 和 Tony Mcenery (2006)通过对英语近义词和汉语近义词的搭配行为和语义韵的比较, 分析了英语近义词和汉语近义词的三个案例[8]。LIU (2020)探讨了“seem”的语义韵与语义偏好作为判断其主要功能的指标[9], 通过对“seem”的语义韵与语义偏好的分析, 揭示出“seem”所体现的语义功能。因此, 基于语料库的同义词研究被越来越多的研究者所关注。然而, 对于 event 和 accident 这两个同义词的研究却很少涉及。同时, 虽然对其他同义词的搭配、语义韵等进行了研究, 但对其语用意义的研究却很少。

John I. Saeed 将同义词定义为“含义相同或相似但发音不同的单词”。他相信世界上没有两个词的意思完全相同[10]。Palmer (1981)认为, 同义词属于不同的语域, 因此有不同的搭配, 并代表了人们的不同态度[11]。基于语料库的同义词研究仍有广阔的发展前景。它们不仅可以探索词汇的搭配、组合, 而且更关注词汇的语义偏好, 帮助学习者把词汇融入语用内容。本文将基于 BNC 语料库为基础, 从搭配、语义韵和语义偏好等方面探讨 event 和 accident 的使用。虽然同义词有非常相似的意义, 但它们可能在搭配或语义韵行为上有所不同。本文探讨了积极、中性和消极三种语义韵, 以及 event 与 accident 之间的语义偏好。同时, 揭示了 event 与 accident 之间的语用意义。

2. 研究方法与设计

本文以 BNC 语料库为基础, 运用 BNC 语料库中的语言数据, 比较了同义词 event 和 accident 的用法, 揭示出它们的语用意义和使用规则。在 BNC 语料库中主要以搭配, 分布语域和共现性为关注点。本文从 BNC 语料库中分别对 event 和 accident 两词选取前 30 个例子, 并统计出这 30 例中与目标词共现性最高的前 10 例作为分析对象。通过对它们的搭配和语域的分析, 可以看出它们的类连接和语义韵, 如积极态度、消极态度和中性态度。本文主要利用 BNC 语料库中的互信息值(MI)进行分析。互信息值(MI)可以反映节点词与搭配词之间的相互吸引关系。互信息值越大, 两个词的搭配强度就越大。

本研究通过比较, 回答了三个问题:

- (1) BNC 中 event 和 accident 所使用的语用环境有什么不同;
- (2) 这两个词的搭配词有什么不同;
- (3) 它们的语义韵有什么不同。

3. 研究结果与讨论

3.1. Event 和 Accident 在不同语域中的分布

所谓的语域是指根据不同的场合和环境而区别的语言变体, 它是考察同义词差异的有效方法之一。语言可以应用在很多领域, 如新闻广播、报道、学术写作、广告语言、课堂语言、日常会话等。不同地区使用的语言会有不同的风格形式。同时由于同义词内部含义的差异, 它们在不同的语域往往表现出不同的分布特征。因此, 统计不同语域的频率差异可以帮助学习者有效地区分同义词。Rundell (1995) 使用 BNC 的口语语料库进行调查, 发现“began”在书面语中的出现频率几乎是“started”的两倍, 而“began”在口语中的出现频率远低于“started”。

对于 accident 和 event 的分布, BNC 语料库中显示这两个词之间的差异较小。人们习惯于使用这两组词来描述某件事情, 例如, 地震, 交通事故或总统选举。这些事情可能是一件平常小事, 也可能是国家的某件大事。考虑到人们话语谈论的方便程度, 这两个词也在一些口语材料或对话中使用, 但在书面材料中的适用程度高于口语材料。如表 1, 表 2 可知, BNC 语料库中 event 在口语材料中的运用仅占 4.02, 而 accident 仅占 31.89, 是各类语域中占比最少的量。这两个词在书面材料中的使用, 如报纸和小说中, 多用来描述最近发生的事故或事件。特别是“event”这个词在报纸上被广泛使用, 其数量达到了 180.4, 用来报道有关该地区和国家的一些重要事件。

在对这两组词的正式度进行研究的过程中, 主要将 BNC 语料库的书面语进行分类, 依据语域的正式度分为两类, 包括高正式度和低正式度。其中高正式度语域包括 newspapers, other published written material 和 academic prose, 低正式度语域包括 unpublished written material, fiction and verse, and non-academic prose and biography。通过观察表 1, 表 2 可知, 在 BNC 语料库中, “event”一词在书面材料中出现的频率要高于“accident”, 且因为“event”在一些学术文章或已发表书面材料中的使用频率较高, 所以 event 比 accident 有更高的正式度。因此, 在 BNC 语料库的协助下, event 和 accident 在不同语域的分布可以很明显地展现出来, 这有助于学习者更好地使用它们。同时, event 和 accident 之间的正式度差异也可以有所体现。

Table 1. Distribution of event in BNC registers

表 1. Event 在 BNC 语域中的分布

Category	Frequency per million words
Unpublished written material	199.93
Newspapers	180.4
Academic prose	147.29
Other published written material	123.13
Non-academic prose and biography	100.09
Other spoken material	47.93
Fiction and verse	27.38
Spoken conversation	4.02
Total	104.74

Table 2. Distribution of accident in BNC registers
表 2. Accident 在 BNC 语域中的分布

Category	Frequency per million words
Unpublished written material	250.75
Newspapers	105.18
Other published written material	70.13
Fiction and verse	62.93
Other spoken material	53.6
Non-academic prose and biography	37.35
Academic prose	34.61
Spoken conversation	31.89
Total	64.06

3.2. Event 和 Accident 的搭配与类连接分析

Firth 在 1957 年首次提出了搭配的概念。Firth 认为, 类连接是语法范畴之间的句法结构上的相互关系, 也就是说, 类连接代表了词语的搭配范畴, 是词汇与语法的结合。类连接是词语搭配研究中的一个重要概念。区分同义词的关键是了解许多典型的词类组合和搭配。event 的搭配主要包括实义词和虚词两类。在实义词中, 主要包括 infinitive, horizon, annual, organized, sporting 等。例如 sporting, 如 1) 所示。同时与 accident 所搭配的实义词主要包括 emergency, fatal, sickness, injury 等。例如 fatal, 如 2) 所示。在虚词中, 这两个词都与冠词搭配, 尤其是 a 和 an。例如 an, 如 3) 所示。此外, accident 常与介词 by 搭配, 形成固定结构 “by accident”, 与 suddenly 同义。英语学习者经常使用这个短语。如 4) 所示。

1) It could be the greatest English **sporting event** since we won the World Cup in 1966.

2) The **fatal accident** occurred on 5th December 1898 when a fall of earth trapped George Robinson against a tipping wagon, breaking his back.

3) But only BBC win when **an event** is screened simultaneously.

4) But I think I got into drama professionally in the first place **by accident**.

在 BNC 语料库中, event 和 accident 通常搭配的词类是形容词和冠词, 少量词也与介词搭配。这两个词通常所搭配的词性相同, 但是他们的具体搭配词汇是不同的。如表 3, 表 4 可知, event 通常与一些形容词搭配, 如 horizon, sporting, fundraising 等。然而, accident 通常搭配 sickness, fatal, horrific 等。因此, 这两个词的不同搭配使它们具有不同的语用意义。Event 可以用来指一件事是有组织的, 有计划的

Table 3. High frequency collocation words of event in BNC

表 3. Event 在 BNC 中的高频搭配词统计

搭配词	共现性	搭配词	共现性
Infinitive	125	Two-day	18
Horizon	54	Traumatic	17
Sporting	47	Denoted	16
Sponsored	42	Fund-raising	14
Three-day	23	Fundraising	11

Table 4. High frequency collocation words of accident in BNC
表 4. Accident 在 BNC 中的高频搭配词统计

搭配词	共现性	搭配词	共现性
Emergency	157	Chernobyl	46
Fatal	115	Tragic	40
Estimates	62	Freak	35
Prevention	57	Investigators	20
Sickness	47	Horrific	17

方式, 且该事件的发生规模较大, 多属于国家或机构所组织的重要事务活动。然而, 如果一件事情突然发生, 造成了严重的不良后果, 包括突发的不可预测自然灾害或交通事故等, 就可以用 accident (事故) 这个词。

3.3. Event 和 Accident 的语义韵和语义偏好探究

语义偏好和语义韵主要用来表达作者或说话人的态度。由词语搭配形成的语义韵可分为消极语义韵、中性语义韵和积极语义韵[12] [13]。通过观察 BNC 语料库的索引, 很多词语表面上似乎没有情感, 但当它们与某些特定词语形成搭配或许多联系时, 就会产生一些联想意义, 如喜恶、褒贬。

对于 event 和 accident 的使用, 根据 BNC 中 event 的搭配和 MI 值, 在节点词 event 后面有很多搭配。其中, horizon, sponsored 和 fund-raising 等有明显积极的语义倾向, sporting, three-day 和 denoted 等表现出中性语义倾向。数据表明, event 对态度积极和中性的形容词有明显的倾向。相反, 节点词 accident 后的搭配词如 fatal, sickness 和 emergency 等表现出消极的语义倾向。当形容词与 accident 搭配时, 其上下文通常是肯定和被动的。例如, 疾病的发生是不受控制的, 可以对人体造成严重的威胁。然而, event 指的是一些按照计划或组织发生的事情, 例如, 每年举行的运动会, 事情很重要, 规模也很大。因此, event 具有积极的语义偏好, 包括这个事件的内容通常是正式而重要的, 这也解释了为什么 event 的正式程度高于 accident。如果说说话人不能准确把握这两类词汇的语用含义, 那么将很容易造成词汇误用, 影响听众的理解, 产生误会。

通过 BNC 语料库, 可以比较这两个词的语用意义。学习者知道如何使用它们, 以及它们可以在什么语境中使用。同时, 也回答了如何与这两类词汇搭配的问题。

4. 教学启示

本研究运用 BNC 语料库, 分析了 event 和 accident 在词汇搭配、类连接和语义韵等方面的差异, 以帮助学习者更好地使用这组同义词。通过比较这组同义词, 增加了教学中对该组同义词用法的理解以及该组词汇上下文的运用。语料库提供的大量真实的语言数据可以使语言学习者和教师对同义词的语境进行比较和研究, 使语言学习更加可靠和有效[14]。在学习一门外语时, 我们不仅要学习语法和单词的表面意义, 还要学习单词的真正意义和单词在语境中的实际使用。语料库为语言教学和自主学习提供了真实的语境, 使学习者能够真正掌握和使用真实的语言。

因此, 在 event 和 accident 的同义词教学中, 教师不仅要讲解词语的结构, 发音和表层意思, 还要积极运用语料库指导词语的搭配和语义偏好, 为学生分析这两个词的分布差异, 培养学生的语用意识, 帮助学生在实际的语用环境中准确使用同义词, 把握同义词的深层含义和语用意义。同时, 还应鼓励学生自主使用在线语料库工具构建语料库, 使其最终成为英语教学中常用的工具。

5. 结语

运用语料库进行同义词辨析是一种实用准确的方法。它为同义词辨析提供了客观有力的依据, 也为今后的英语教学提供了新的思路。本文发现, *event* 和 *accident* 在词频分布、搭配、语义韵等方面存在差异。首先, *event* 经常出现在报纸等书面材料中, *event* 发生的总频率高于 *accident* 发生的频率。其次, *event* 的语用正式度高于 *accident*, 经常使用于高正式度语域中。最后, *event* 的搭配词具有积极的语义倾向, 而 *accident* 搭配词具有消极的语义倾向。对于 *event* 和 *accident* 两词的区分, 使教师和学生在今后的教学和学习过程中能够更好地运用它们, 为词汇的学习增加了准确性和信度。本研究选取 *event* 和 *accident* 作为目标词汇, 原因是尽管大多数的人们了解这两个词汇的表面含义, 也常会在书面语或口语中运用它们, 但对其语用意义和语义态度的深刻辨析仍然无法正确把握, 从而在生活和学习中产生严重的误解。本研究的不足之处在于语料选取有限, 在以后的研究中通过增加语料数量可以增强词汇分析的说理力, 并能获得大量的新发现。

参考文献

- [1] 庞杨, 杨玉晨. 语料库对同义词触发模式的解释: 以 *begin*、*start* 和 *commence* 为例[J]. 外语与外语教学, 2012(6): 21-25.
- [2] 孟庆楠, 钱景, 周晨煜. 海事英语中近义词的搭配分析及可视化呈现——以 *marine* 和 *maritime* 在世界英语中的用法为例[J]. 大连海事大学学报: 社会科学版, 2020(6): 112-118.
- [3] 冷雪莲. 基于 COCA 语料库辨析英语同义词 *Capable* 和 *Competent* [J]. 成都师范学院学报, 2015(2): 54-58.
- [4] 郑丹. 基于语料库的英语同义词研究——以 *accept* 和 *receive* 为例[J]. 黑龙江教育学院学报, 2018(12): 126-128.
- [5] Zhou, Q. and Zhang, L. (2020) Corpus-Based Approach to Explore the Semantic Prosody of Synonym: A Case Study of “lead to” and “result in”. *US-China Foreign Language*, **18**, 247-252. <https://doi.org/10.17265/1539-8080/2020.08.004>
- [6] 李晓红, 卫乃兴. 汉英对应词语单位的语义趋向及语义韵对比研究[J]. 外语教学与研究, 2012(1): 20-33.
- [7] Wei, N.X. and Li, X.H. (2014) Exploring Semantic Preference and Semantic Prosody across English and Chinese: Their Roles for Cross-Linguistic Equivalence. *Corpus Linguistics and Ling Theory*, **10**, 103-138. <https://doi.org/10.1515/cllt-2013-0018>
- [8] Xiao, R. and McEnery, T. (2006) Collocation, Semantic Prosody, and Near Synonymy: A Cross-Linguistic Perspective. *Applied Linguistics*, **27**, 103-129. <https://doi.org/10.1093/applin/ami045>
- [9] Liu, W.Q. (2020) A Corpus-Based Study of Semantic Prosody and Semantic Preference of “Seem”. *Open Journal of Social Sciences*, **8**, 282-291. <https://doi.org/10.4236/jss.2020.85020>
- [10] Saeed, J.I. (1997) *Semantics*. Blackwell Publishers Ltd., Oxford.
- [11] Palmer, R. (1981/1993) *Semantics: A New Outline*. 2nd Edition, Cambridge University Press, Cambridge.
- [12] 甄凤超, 杨枫. 再谈语义韵的短语及语用属性[J]. 外语教学理论与实践, 2019(3): 34-40, 49.
- [13] Stubbs, M. (1996) *Text and Corpus Analysis*. Black-Well Publishers, Oxford.
- [14] 王春艳. 基于语料库的中国学习者英语近义词区分探讨[J]. 外语与外语教学, 2009(6): 27-31.