

基于语料库的英汉近义词语义韵对比研究

——以“Cause”，“Arouse”和“导致”，“促使”为例

陶雯馨

上海海事大学外国语学院，上海

收稿日期：2023年7月4日；录用日期：2023年8月2日；发布日期：2023年8月15日

摘要

本研究采用 BROWN, LOB, FROWN, FLOB, CROWN 和 CLOB 等六个布朗家族英语语料库以及兰卡斯特现代汉语语料库(LCMC)和北京大学 CCL 语料库两大汉语语料库,通过 AntConc 语料库软件对 cause, arouse 和导致、促使两对英汉近义词“导致”类动词进行索引行分析。研究发现,英语“导致”类动词 cause 呈现明显的消极语义韵,而 arouse 则表现为积极语义韵;汉语中“导致”在语料中只能体现出消极语义韵,“促使”显示为积极语义韵。此外,为增加可对比性,笔者还参考了以汉语作为二语学习对象的学习者语料库,分析其反映的语义特征,发现了留学生和中文母语者中存在的差异,并对中文教学提供一些启示。

关键词

语料库, 语义韵, 英汉对比

A Corpus-Based Contrastive Study on the Semantic Prosody between English and Chinese Synonyms

—Taking “Cause”, “Arouse” and “导致”, “促使” as Examples

Wenxin Tao

College of Foreign Languages, Shanghai Maritime University, Shanghai

Received: Jul. 4th, 2023; accepted: Aug. 2nd, 2023; published: Aug. 15th, 2023

Abstract

In this study, six Brown family English corpora, including BROWN, LOB, FROWN, FLOB, CROWN

and CLOB, as well as two Chinese corpora, LCMC and the CCL Corpus of Peking University, are used to analyze the English and Chinese synonyms “cause, arouse” and “导致, 促使” by using the AntConc corpus software to collect the concordance. It is found that the English verb CAUSE turns out an obvious negative semantic prosody, while “arouse” shows a positive semantic prosody. The Chinese word “导致” only reveals the negative meaning in the corpus, while “促使” is displayed in a positive. Additionally, to enhance comparability, the author drew upon a learner corpus of Chinese as a second language, thoroughly analyzing its reflected semantic characteristics. As a result, notable dissimilarities in language usage between international students and native Chinese speakers were pinpointed, subsequently offering valuable insights for the field of Chinese language instruction.

Keywords

Corpus, Semantic Prosody, Contrastive Study between English and Chinese

Copyright © 2023 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

语义韵研究一直是语言学领域的焦点话题之一, 语义韵(Semantic Prosody)起源于语料库语言学, 其最早出现在 Sinclair [1]关于语料库的开创性著作中。然而, 语义韵的定义很难确定, Louw [2]强调语义韵背后隐藏的“说话人的用意”, 认为语义韵研究能发掘话语背后的讽刺意义或说话人对某一事件的预设立场。Morley & Partington [3]认为语义韵归根结底是对事物好坏的评价, 他们在此基础上提出语义韵极性: 消极/中性/积极。

本文在语义韵极性的理论上以语料库为媒介手段, 在英语汉语中各选取两个典型的“导致”类动词进行深入分析, 它们分别为英语“导致”类动词: cause, arouse 以及汉语“导致”类动词: 导致、促使。研究采用 BROWN, LOB, FROWN, FLOB, CROWN 和 CLOB 等六个布朗家族英语语料库以及兰卡斯特现代汉语语料库(LCMC)和北京大学 CCL 语料库两个汉语语料库, 通过 Antconc 语料库软件对上述动词进行索引行分析, 研究各个“导致”类动词的语义韵特征。

2. 文献综述

Stubbs [4]是最早研究 cause 语义韵的学者, 他坚持认为 cause 非中性, 而且在绝大多数场合表达了负面的语气, 其对 cause 的语义表征研究为后人的研究奠定了良好的基础。基于 Stubbs 的研究, Hunston [5]通过识别因果关系进一步区分了有关疾病、症状或情绪的被动实体的相关过程, 她在对于 cause 的实证研究中批判性地指出只有当人类或其他生命体明显地参与活动中时, cause 才意味着不受欢迎的事物。除此之外, 还有一些研究集中在汉语语义韵的历时性变化研究上, Stubbs 和 Bublitz 提出有关 cause 负面意义的实例在几十年或几百年后会增加[6], 然而因假设缺乏充分的实证支撑, 该观点颇受争议。受到 Stubbs 和 Bublitz 的假设的启发, Smith 和 Nordquis [7]对比了早期现代英语(Early Modern English period)和现代英语(present-day English)的标记, 对“导致”类动词的负面意义动词开展了历时研究, 在语料库的帮助下他们发现动词 cause 的语义表征随着时间的推移发生了巨大的变化。

随着语料库语言学在中国的蓬勃发展, 语义表征的研究越来越受到中国学者的关注, 尤其是作为经典词条的 CAUSE, 其语义表征得到了广泛的研究[8] [9] [10]。上述研究系统而详细地讨论了“导致”类动词的语义表征, 并对其在外语教学和词典编纂中的应用进行了说明。然而, 只有少数研究者对“导致”类动词语义表征的历时性变化进行了深入研究, 此外国内对“导致”类动词的英汉对比分析工作也做得很少。本论文将采用定量和定性的方法延伸前人的研究成果, 对“导致”类动词在英语和汉语中的语义表征进行仔细观察, 尝试探索其语义韵特征和一定的历时性变化, 在此基础上通过对比分析, 讨论研究在汉语语言教学中的实际应用。

3. 研究设计

随着计算机技术和语料库语言学的日益发展, 语料库在二语习得、翻译、词典编纂等理论和实践领域得到了广泛的应用。本研究主要基于英汉双语语料库, 其中英文语料主要来自于 BROWN, LOB, FROWN, FLOB, CROWN 和 CLOB 等六个布朗家族英语语料库, 这些英语语料库约包含 100 万个单词, 规模较大, 并且在一定的时间间隔内具有相似的采样框架, 可以作为描述语言变化的良好资源。中文语料则取自兰卡斯特现代汉语语料库(LCMC)和北京大学 CCL 语料库, 涵盖了大约 15 个子语料库和 100 万单词。如表 1 所示, 在上述语料库的帮助下, 研究团队检索到选定的英汉“导致”类动词的词语索引行(concordance line), 进而通过定量和定性相结合的方法, 揭示所选节点词的语义特征。

Table 1. Brief introduction to selected corpora (information sourced from Wikipedia)

表 1. 所选语料库简介(信息来源维基百科)

语料库	创建时间	语言	语料中词的总数
LOB	1961	英式英语	1,018,785
BROWN	1961	美式英语	1,027,021
FLOB	1991	英式英语	1,027,021
FROWN	1991	美式英语	1,027,021
CLOB	2009	英式英语	1,023,466
CROWN	2009	美式英语	1,026,226
LCMC	1990s	汉语	1,021,046
CCL	2000s	汉语	783,463,175

4. 分析与结果

4.1. Cause 语义韵分析

在布朗语系的语料库中(即 BROWN、LOB、FROWN、FLOB、CROWN 和 CLOB), 所有的动词形式(动词原形、现在分词和过去分词)都将被探讨, 本研究选择的第一个节点词是“导致”类动词 cause, 将对 cause 的所有动词形式和搭配词进行详细研究。考虑到其搭配的特点和语料库的标签, 重要的搭配可能超过所选节点词右侧的 5 个词, 因此跨度被设定为节点词左边的 5 个词和右边的 5 个词, 表 2 中列出了一些删除标签的 cause 词语索引行。

在上述截取 cause 的搭配关系中, 如 inconvenience, damage, fail, goiter 和 illness 等都有明显的负面含义, 由此作者推断 cause 具有负面的语义倾向。为了进一步揭示节点词的语义特征, 表 3、表 4 和表 5 给出了不同语料库中 cause 的语义分布情况。

Table 2. The concordance lines of cause in native English corpora
表 2. 英语语料库中 cause 词语索引行

序号	左边 5 个词	节点词	右边 5 个词	语料库
1	for the inconvenience he had	caused	the coffee bar management in	LOB-A
2	(inclusive of car parks) without	causing	damage or disruption and without	LOB-F
3	her that cramp is often	caused	by having lost salt through	LOB -F
4	impact with the utility pole	caused	a brief power failure in	BROWN-A
5	long period of time will	cause	goiter, owing to inhibition of	BROWN-J
6	shirked that duty even when	causing	(at least) short-term political	CLOB-F
7	perceived—something that can	cause	damage. Unfortunately this is	CLOB-E
8	that researchers predicted would	cause	the most illness this season	CROWN-F

Table 3. Distribution of cause in BROWN & LOB
表 3. cause 在 BROWN & LOB 中的分布情况

语料库	极性	出现次数	比例	语料库	极性	出现次数	比例
BROWN	Negative	118	61.78%	LOB	Negative	124	77.99%
	Positive	12	6.28%		Positive	12	7.55%
	Neutral	61	31.94%		Neutral	23	14.47%

Table 4. Distribution of cause in FROWN & FLOB
表 4. cause 在 FROWN & FLOB 中的分布情况

语料库	极性	出现次数	比例	语料库	极性	出现次数	比例
FROWN	Negative	100	74.1%	FLOB	Negative	124	79.1%
	Positive	5	3.7%		Positive	12	2.3%
	Neutral	30	22.2%		Neutral	23	18.6%

Table 5. Distribution of cause in CROWN & CLOB
表 5. cause 在 CROWN & CLOB 中的分布情况

语料库	极性	出现次数	比例	语料库	极性	出现次数	比例
CROWN	Negative	175	75.4%	CLOB	Negative	146	85.4%
	Positive	10	4.3%		Positive	5	2.9%
	Neutral	47	20.3%		Neutral	20	11.7%

参照上述表格可发现 cause 表现出了绝大部分的负面语义(61.78%, 74.1%, 75.4%, 77.99%, 79.1%, 85.4%), 这与之前提取 cause 索引行所表现出的负面上下文含义的结果相呼应。此外, “导致”类动词 cause 的语义特征随着时间的推移呈现出变化。根据布朗家族语料库的数据, 负面搭配的比例在一定程度上有所增加, 而且由于时间跨度和地区差异, 中性搭配的分布情况也各不相同: 在美式英语中, cause 的中性搭配的比例自 1961 年以来减少了近 10%, 然后保持相对稳定(31.94%, 22.2%, 20.3%); 而在英式英语中, 这一数字因时间而异, 从 1961 年到 1991 年增加了 4%, 然后从那时起又减少了近 7% (14.47%, 18.6%, 11.7%)。

为了更好地理解 cause 的语义韵, 作者从不同语料库中抽取的搭配如下表 6 所示。可以看出, 与 cause 相关的典型名词搭配包括了 death, effect, trouble, pain, problems, warming, women, people, system, concern 和 accident 等大量具有负面意义的名词, 因此我们可以得出结论, cause 明显表现出负面的语义情境。

Table 6. The collocates of cause and their frequencies in English corpora

表 6. 英语语料库中 cause 的搭配及其频率

Cause	常用搭配
BROWN	effect(9); death(7); trouble(6); State(5); system(4); pressure(4); change(4); action(4); something(3); radiation(3); power(3); order(3); interference(3); increase(3); failure(3); difficulty(3); conditions(3);
LOB	death(10); damage(8); trouble(6); accident(6); concern(5); delay(4); confusion(4); alarm(4); work(3); place(3); people(3); peace(3); pain(3); harm(3); group(3); effect(3); crisis(3);
FROWN	problems(7); disease(7); death(6); damage(5); women(4); people(4); effect(4); concern(4); blood(4); years(3); year(3); symptoms(3); stir(3); problem(3); pollution(3); pain(3); idea(3); human(3); effects(3); condition(3); argument(3); argued(3); ability(3)
FLOB	damage(13); land(12); degradation(12); concern(10); death(9); problems(8); extinction(7); behavior(7); reason(6); pain(6); events(6); trouble(4); time(4); things(4); terms(4); surprise(4); stress(4); failure(4); distress(4); understanding(3); system(3); result(3); reasons(3); pressure(3); population(3); parents(3); light(3); insufficiency(3); increased(3); effects(3); great(3); good(3); disruption(3); delay(3); decline(3); considerable(3); conditions(3); anxiety(3)
CROWN	problems(11); warming(8); people(8); something(6); death(6); damage(6); cancer(6); smoking(5); human(5); women(4); thing(4); inability(4); disease(4); concern(4); cases(4); behavior(4); system(3); student(3); particles(3); necessity(3); movement(3); investment(3); influence(3); illness(3); heart(3); failure(3); earth(3); door(3); diseases(3); cracks(3); better(3); advance(3);
CLOB	damage(14); flooding(8); change(8); problems(7); blindness(7); performance(6); development(6); illness(5); understanding(4); harm(4); effects(3); trouble(3); symptom(3); risk(3); incident(3); growth(3); flu(3); feel(3); disease(3); disability(3); depression(3); deaths(3); climate(3).

4.2. Arouse 语义韵分析

在这一部分中, 作者将详细分析“导致”类动词 arouse。下表 7~9 显示自 1930 年以来, arouse 的总频率急剧下降(30, 29, 11, 9, 5, 3), 从近 30 次下降到仅低于 10 次。在 19 世纪 60 年代正面搭配占总频率的最大比例, 相比之下 20 世纪 90 年代的频率很低, 而且 arouse 的语义分布没有明显的语义偏好。考虑到所选语料库中 arouse 的频率有限, 我们将每个语料库中的频率结合起来, 以计算出 AROUSE 的语义表征。结果发现, arouse 的正面搭配占 63.75%, 负面搭配占 26.25%, 10% 为中性搭配。因此, 在原始英语写作样本中, 作者倾向于采用 arouse 的正面语义语体, 下表也反映出美式英语语料库中的负面搭配比英式英语语料库中的负面搭配更受青睐。

Table 7. Distribution of arouse in BROWN & LOB

表 7. arouse 在 BROWN & LOB 中的分布情况

语料库	极性	出现次数	比例	语料库	极性	出现次数	比例
BROWN	Negative	8	27.59%	LOB	Negative	9	30.00%
	Positive	15	51.72%		Positive	20	66.67%
	Neutral	6	20.69%		Neutral	1	3.33%

Table 8. Distribution of arouse in FROWN & FLOB
表 8. arouse 在 FROWN & FLOB 中的分布情况

语料库	极性	出现次数	比例	语料库	极性	出现次数	比例
FROWN	Negative	2	22.22%	FLOB	Negative	1	9.09%
	Positive	4	44.44%		Positive	8	72.73%
	Neutral	3	33.33%		Neutral	2	18.18%

Table 9. Distribution of arouse in CROWN & CLOB
表 9. arouse 在 CROWN & CLOB 中的分布情况

语料库	极性	出现次数	比例	语料库	极性	出现次数	比例
CROWN	Positive	1	33.33%	CLOB	Negative	0	0%
	Negative	2	66.67%		Positive	2	40.00%
	Neutral	0	0%		Neutral	3	60.00%

4.3. “导致”语义韵分析

在本项研究中，针对汉语“导致”类动词我们将基于 LCMC 的语料库，主要介绍两个汉语语义体：导致、促使。“导致”一词的语义韵在 LCMC 中的分布情况如下表 10、表 11 所示。

Table 10. Distribution of “导致” in LCMC
表 10. “导致”在 LCMC 中的分布情况

	消极	积极	中性	总频率
导致(LCMC)	63 (79.7%)	7 (8.86%)	9 (11.39%)	79

表中的数据根据极性分别显示了 LCMC 中三类搭配的比例：“导致”的负面搭配占 79.7%，在 LCMC 中最常见的搭配是观察、结构、平衡、发展、减少、重要、贸易、结果、经济、疾病、死亡。为了更深入地研究导致的语义前景，我们还从 LCMC 观察中提取了一些索引行，结果发现在下述从 LCMC 中随机抽取的索引行中，“导致”的情感含义均是负面的，由此可以推断在很大程度上，“导致”一词的搭配具有明显的消极意义。

Table 11. Concordance lines of “导致” in LCMC
表 11. 导致在 LCMC 中的词语索引行

序号	左边 5 个词	节点词	右边 5 个词	语料库	语义韵
1	化及心内膜硬化，	导致	心脏泵效率下降，使	LCMC-J	消极
2	各种心律失常，接至	导致	碎死。50 岁以上	LCMC-F	消极
3	容易；一当通胀	导致	严重经济扭曲时，要	LCMC-J	消极
4	是自杀，而是性窒息	导致	意外死亡。在	LCMC-L	消极
5	激反应后，最终可以	导致	疾病。沃尔夫推测	LCMC-J	消极

4.4. “促使”语义韵分析

表 12 给出了“促使”在 LCMC 中不同语义体分布情况的一般描述，其显示在 LCMC 中 81.82% 的相

关搭配是积极的,中性搭配占整个选定数据的 11.36%,而负面仅占 6.82%。大体上可以得出结论“促使”作为一个节点词主要显示出积极语义韵,最常见的搭配包括:促使、平衡、人们、经济、社会、因素和发展。

Table 12. Distribution of “促使” in LCMC

表 12. 促使在 LCMC 中的分布情况

	消极	积极	中性	总频率
导致(LCMC)	3 (6.82%)	36 (81.82%)	5 (11.36%)	44

综上基于语料库的对比研究结果发现,虽英语与汉语虽隶属于不同语系,其各自的词汇也呈现出不同的语义表征,但“导致”类动词却都具相似语义韵特征。CAUSE 在英语的语境中多呈现出消极的语义韵,这与其对应的汉译“导致”刚好吻合;而 AROUSE 多使用在包含积极语义韵的上下文中,所以将其和在中文中同样具有积极含义的“促使”对应更加妥帖。

5. 讨论

在这一部分中,我们将对汉语母语者和以汉语为二语习得对象的留学生在语义情景方面进行对比研究。根据 Antconc 检索到的数据,对北京大学 CCL 语料库——留学生汉语作文语料中的“导致”、“促使”的搭配进行了研究,试图挖掘出外籍汉语学习者对汉语“导致”类动词语义表征的把握能力,并对提高学生的语义表征意识给出了一些实际意义。由于篇幅有限,现根据取自北京大学 CCL 语料库的数据,发现“导致”一词在语义表征方面的用法,并对不恰当的语义表征进行了详细分析,下表列出了导致在三个类别中的分布情况。

Table 13. Distribution of “导致” in CCL

表 13. “导致”在北京大学 CCL 语料库中的分布情况

导致	极性	次数	占比
北京大学和 CCL 语料库	消极	301	74.14%
	积极	77	18.97%
	中性	28	6.90%

表 13 显示了“导致”在北京大学 CCL 语料库中的语义前景特征的分布情况,研究团队发现“导致”的重要名词搭配有:伤害、问题、后果、污染、浪费、消耗、死亡、发展、短缺、原因、变化和进步等,在这些最常见的搭配中“伤害、后果、污染、消耗、死亡、短缺”表现出明显的消极含义,结果表明“导致”在留学生汉语作文语料中倾向于承担消极的语义前景。与以汉语为母语的人相比,虽然“导致”明显表现出消极的语义前景,但留学生们也倾向于使用积极的搭配,同时也普遍存在对中性搭配使用不足的情况。总而言之,对“导致”类动词的各种语义表征的研究,有利于促进留学生对语言使用和写作作品更好的理解,并提高中文学习者的写作能力。在揭示“导致”类动词在语义表征方面的语用功能的过程中,可以发现并纠正不恰当的语义表征,因此通过指出语义准绳的明显过度使用或不足来突出语料库数据中的重要搭配,为自然语言处理提供了一个完整的理论背景[11]。在今后的中文教学过程中,教学团队可以更加注重语义韵的相关知识讲解,帮助外籍汉语学习者更快、更好、更高效地掌握相关知识并灵活运用。

6. 结语

本文中通过定量和定性相结合的方法,对英汉近义词“导致”类动词 cause, arouse 和导致、促使的

语义表征进行了基于语料库的对比研究。研究发现, 英语“导致”类动词 cause 呈现明显的消极语义韵, 而 arouse 则表现为积极语义韵; 汉语中“导致”在语料中只能体现出消极语义韵, “促使”显示为积极语义韵。此外, 参考可比的以汉语为二语习得对象的学习者语料库中所反映的语义特征, 笔者发现了留学生和中文母语者中存在的使用差异, 并对中文教学提供一些启示。

参考文献

- [1] Sinclair, J. (1991) *Corpus, Concordance, Collocation*. 2nd Edition, Oxford University Press, Oxford.
- [2] Louw, B. (1993) Irony in the Text or Insincerity in the Writer? In: Baker, M., Francis, G. and Tognini-Bonelli, T., Eds., *Text and Technology: In Honour of John Sinclair*, John Benjamins, Amsterdam.
- [3] Morley, J. and Partington, A. (2009) A Few Frequently Asked Questions about Semantic or Evaluative Prosody. *International Journal of Corpus Linguistics*, **14**, 139-158. <https://doi.org/10.1075/ijcl.14.2.01mor>
- [4] Stubbs, M. (2009) The Search for Units of Meaning: Sinclair on Empirical Semantics. *Applied Linguistics*, **30**, 115-137. <https://doi.org/10.1093/applin/amn052>
- [5] Susan, H. (2002) *Corpora in Applied Linguistics*. Cambridge University Press, Cambridge.
- [6] Dominic, S. (2010) *Semantic Prosody*. Routledge, London.
- [7] Smith, A. and Nordquist, D. (2012) A Critical and Historical Investigation into Semantic Prosody. *Journal of Historical Pragmatics*, **13**, 291-312. <https://doi.org/10.1075/jhp.13.2.05smi>
- [8] 程铖, 潘璠. 基于语料库的 CAUSE 语义韵对比研究[J]. 高等工程教育研究, 2008(S1): 108-113.
- [9] 王海华, 王同顺. CAUSE 的语义韵对比研究[J]. 现代外语, 2005, 28(3): 297-307.
- [10] 顾鲜静. 基于语义韵的同义词搭配个案研究[D]: [博士学位论文]. 上海: 上海外国语大学, 2012.
- [11] 卫乃兴. 词语学要义[M]. 上海: 上海外语教育出版社, 2011.