

基于语料库的英语近义动词搭配特征对比研究和教学

李瑶, 廖顺林

(中南大学 外国语学院, 湖南 长沙 410083)

摘要:文章对比分析了 achieve 和 gain, arise 和 emerge 两组近义动词在中国学生英语笔语语料库和英语国家语料库中的搭配特征和语义韵的异同,探讨了4个动词在中国学生英语笔语语料库中的异常搭配和原因。结果发现:相比本族语者,中国学生过度使用 achieve 和 gain,过少使用 arise 和 emerge;gain 的宾语搭配词比 achieve 丰富,但两者的宾语搭配词均不如本族语者丰富和具体,没有消极语义韵;arise 和 emerge 的主语搭配词和语义韵单一,而且被误用为及物动词;中国学生混淆了两组近义动词的部分搭配和形似词的用法。错误原因主要是母语干扰、过度概括和忽略上下文。文章最后提出了英语近义词辨析和教学可采用词块和语料库驱动的教学模式。

关键词:语料库;近义词;动词;搭配;语义韵

中图分类号:G642

文献标志码:A

文章编号:1674-5884(2019)03-0125-06

同义词指“在同一种语言中意义相同或相近的单词或表达”^[1]。Palmer 注意到同义词在大量参数上如文体和搭配上有不同的分布^[2]。一些学者将同义词分成绝对同义词(Absolute Synonym)和相对同义词(Relative Synonym)即近义词(Near-synonym)两类。Lyons 认为近义词是意义相似但并不完全相同的表达^[3]。Tylor 认为完全近义词(Perfect Synonym 或 Absolute Synonym)非常稀少^[4]。尽管意义相同或相近,同义词在风格、语法或使用上可能存在差异。

语义韵(Semantic Prosody)概念第一次由 Louw 提出。作为一种词语搭配的特殊行为,语义韵是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^[5]。语义韵一般分为消极语义韵(Negative Prosody)、中性语义韵(Neutral Prosody)和积极语义韵(Positive Prosody)三类^[6]。

近年国内外学者对近义词的搭配行为和语义韵的研究广泛关注,取得了一些成果。Sinclair 发

现,happen 和 set in 习惯性与表示不愉快事件的词搭配,表现出明显的消极语义选择趋向^[7]。Xiao 和 Mcenery 使用英汉两种语言的数据,对搭配、语义韵律和近义词进行了跨语言分析^[8]。卫乃兴介绍和讨论了语料库证据支持的词语搭配研究的基本方法和主要原则^[9]。陆军通过研究发现学习者对近义词搭配行为和语义韵知识掌握欠缺^[10]。

动词是句子的核心。几乎对于任何一种语言来说,与动词相关的部分是最难的^[11]。对于中国学生来说,英语中的近义动词非常难以辨析和掌握。因此,对近义动词在本族语和非本族语不同语料库中的搭配特征进行研究具有重要意义,对教学指导具有实用价值。

1 研究设计

1.1 研究对象和问题

本文选取了两组近义动词为研究对象,一组为 achieve 和 gain,表示“获得、得到”的意思,另一组为 arise 和 emerge,意思是“出现”。这四个动词

收稿日期:20181213

基金项目:2017年湖南省和中南大学教改项目“大学英语多元立体交互教学模式研究与构建”(湘教通[2017]452号);2018年中南大学研究生自主探索创新项目(2018zzts052)

作者简介:李瑶(1972-),女,湖南双峰人,硕士,教授,主要从事应用语言学和语料库语言学研究。

均是中国学生需要重点掌握的词汇。

中国学生的语料来自中国学生英语口语语料库 (Spoken and Written English Corpus of Chinese Learners 简称 SWECCCL) 中的书面语子库 (WECCL)。WECCL 包含全国各种层次的高等院校英语专业及部分非英语专业学生的限时和非限时作文共 4 950 篇, 作文文体包含议论文和说明文, 以议论文为主, 形符数为 1 248 476。语料来源广泛, 可以较为准确地反映学生作文的真实情况^[12]。作为参照的本族语语料库是英语国家语料库 (British English Corpus, 简称 BNC), 库容量 112 181 015 形符, 语域广泛, 话题丰富。

本研究将回答以下问题:

1) 中国英语学习者和本族语者对于两组近义词的使用频数是什么? 是过度使用还是使用不足?

2) 在这两组近义词的名词搭配和搭配词的语义韵方面, 中国英语学习者和本族语者是否存在相同或相异?

3) 两组近义词是否在 WECCL 中存在异常搭配? 如果存在, 可能的原因是什么?

1.2 研究工具和方法

本研究使用的工具为 AntConc 3.3.0w 和 Sketch Engine。AntConc 是一个多功能的语料库

检索和统计工具。Sketch Engine 是由语料库专家 Adam Kilgariff 设计的语料库在线检索程序, 它能够将同义词的常用搭配和习惯用法进行对比。

本研究采用基于语料库的对比研究方法, 结合定量和定性分析。首先使用 AntConc 3.3.0w 和 Sketch Engine 分别检索出这两组动词在 WECCL 和 BNC 中的使用频数, 进行卡方检验; 然后, 分别检索出这四个动词在 WECCL 和 BNC 中的名词搭配, 对比分析它们的语义韵特点; 最后, 对比两个语料库中两组近义词的搭配词汇的异同, 做出进一步分析和解释。

2 结果和讨论

2.1 两组近义词在不同语料库中的使用频数比较

表 1 说明, 中国英语学习者对近义词 achieve 和 gain 的使用频数 (Frequency) 与本族语者相比均存在显著差异 (Sig. < 0.05), 中国学生过度使用这两个动词; 中国英语学习者对近义词 arise 和 emerge 的使用频率与本族语者相比均存在显著差异 (Sig. < 0.05), 中国学生较少使用这两个动词。但是, 相对本族语者, 中国学习者对 achieve 和 gain 过度使用, 而对 arise 和 merge 过少使用。

表 1 四个动词在两个语料库中的卡方检验结果

WECCL—BNC	Freq. 1	Norm Freq. 1	Freq. 2	Norm Freq. 2	Chi-square	Significance
achieve	260	0.020 8%	16 710	0.014 9%	28.629 520 22	0.000 000 088
gain	555	0.044 5%	8 727	0.007 8%	2 025.208 019	0.000 000 000
arise	35	0.002 8%	9 523	0.008 5%	-63.885 777 03	0.000 000 000
emerge	20	0.001 6%	7 575	0.006 8%	-70.517 354 61	0.000 000 000

(Freq = Frequency 频数 Chi-square 卡方检验 Significance 显著性)

2.2 两组近义词在 WECCL 中的名词搭配和语义韵分析

2.2.1 achieve 和 gain 在 WECCL 中的动宾搭配比较和语义韵分析

使用 Antconc 检索 WECCL 中 achieve 和 gain 两个动词的搭配, 设检索条件为跨距 4L 到 4R, MI (Mutual Information) 互信息值为 3, 最小搭配频数为 2。检索结果中前 10 个搭配宾语如表 2:

结果显示, achieve 在 WECCL 中出现 2 次及以上的宾语近 30 个。搭配词中含有目标梦想 goal(s), purpose, aim(s), dreams; 成功发展声望 success, progress, development, fame; 工作等级 career, job, work, level; 还有读写能力和分数 literacy, scores, mark 等。这些搭配词大部分呈现

出积极语义韵。WECCL 中符合检索条件的 gain 的搭配宾语多达 50 个, 除了知识教育能力和发展成功等词, gain 的搭配词还包含了自由机会和平 freedom, opportunity, peace; 薪水利润金钱 salaries, profit(s), wealth, money; 利益优势 benefits, advantage; 以及大量情感类词汇 respect, pleasure, confidence, happiness, satisfaction, friendship, cooperation 等。这些搭配词大部分也呈现出积极语义韵。

两个近义词在 WECCL 的显著搭配宾词包括 success, skills, knowledge, job, fame。这显然受中文“获得成功、知识、技能、工作、声望”等表达的影响, 与母语保持了一致, 说明母语迁移发生了作用。与 achieve 相比, gain 的搭配词汇更丰富。仅发展成功类词汇, gain 比 achieve 多了

popularity, accomplishment(s), achievement 等搭配 词。这些搭配词呈现出明显的积极语义韵。

表2 achieve 和 gain 在 WECCL 中的前 10 个搭配宾语

Achieve				Gain			
Rank (排序)	Frequency (频数)	Salience (显著性)	Collocate (搭配词)	Rank (排序)	Frequency (频数)	Salience (显著性)	Collocate (搭配词)
1	46	13.87	success	1	110	6.99	knowledge
2	29	15.07	goal	2	40	4.07	education
3	21	15.41	goals	3	27	7.36	experience
4	12	9.45	knowledge	4	21	7.54	fame
5	9	10.48	something	5	20	6.37	skills
6	8	9.10	things	6	18	6.14	information
7	6	10.29	skills	7	14	4.85	job
8	5	11.66	purpose	8	14	5.04	ability
9	5	11.39	career	9	13	4.30	money
10	4	12.53	dreams	10	8	6.12	benefit

2.2.2 arise 和 emerge 在 WECCL 中的主谓搭配比较和语义韵分析

经检索发现,arise 各时态形式在 WECCL 中出现 35 次,其中 problem(s) 和 conflicts 这两个表示问题与冲突的词汇共出现 16 次,其余是 viruses,friction,debate,opinions,issue,question 等病毒、摩擦、意见问题类词汇和 thing(s) 事物等词,呈现出明显的消极语义韵。

WECCL 中与 emerge 搭配超过两次的主语为 things, problems, 出现一次的消极韵主语为 disease, disadvantage, 与学习和校园有关的词汇有 knowledge, information, examination system, schools,

learners, 以及其他词汇 news, types, styles, products, functions。这些主语词汇大部分呈现出中性语义韵。

2.3 两组近义动词在 BNC 中的名词搭配和语义韵分析

2.3.1 achieve 和 gain 在 BNC 中的动宾搭配比较和语义韵分析

使用 Sketch Engine 检索 BNC 中 achieve 和 gain 两个动词的宾语搭配,考虑到 BNC 的容量达 1 亿词汇,设检索条件为最小频数为 4,显著值 (Salience) 为 3。检索结果中前 15 个搭配宾语如表 3:

表3 achieve 和 gain 在 BNC 中的前 15 个搭配宾语

Achieve				Gain			
Rank (排序)	Frequency (频数)	Salience (显著性)	Collocate (搭配词)	Rank (排序)	Frequency (频数)	Salience (显著性)	Collocate (搭配词)
1	494	9.48	objective	1	350	8.93	access
2	407	7.87	result	2	332	7.86	experience
3	362	8.69	goal	3	175	6.75	control
4	280	8.06	success	4	168	6.89	support
5	238	8.51	aim	5	137	8.7	insight
6	204	6.38	level	6	131	7.59	advantage
7	185	6.57	effect	7	123	7.81	confidence
8	175	6.26	end	8	116	7.04	ground
9	132	7.47	balance	9	107	7.39	understanding
10	126	6.67	standard	10	105	6.72	knowledge
11	121	7.76	independence	11	103	5.59	information
12	114	7.23	target	12	92	6.51	benefit
13	110	6.99	status	13	89	7.72	reputation
14	85	5.15	change	14	88	7.31	entry
15	83	6.91	reduction	15	81	7.33	recognition

BNC 中与 achieve 搭配的词汇非常丰富,符合检索条件的宾语近 300 个,有目标结果效果类 objective, goal, aim, target, result, effect; 成功胜利进步类 success, victory, progress, growth, increase, improvement, development; 声望认可功绩类 fame, recognition, feat, reputation 等。而且, achieve 搭配了很多政治和经济类词汇,如 status, independence, equality, stability, reform, peace, harmony, reconciliation, freedom, order, justice, security, autonomy, union, ceasefire 等。achieve 绝大部分搭配词是抽象名词,含有积极和中性语义韵。但是, achieve 也搭配了部分消极语义韵词汇,如 reduction, notoriety (声名狼藉), separation, loss, hegemony (霸权), inflation, disease, eradication (消灭), elimination (消除), removal, deficit (赤字)。

符合检索条件与 gain 搭配的宾语近 200 个,有进入立足点类 access, entry, ground, foothold; 理

解认可支持类 understanding, recognition, acceptance, support; 优势资质晋升类 advantage, benefit, qualification, promotion; 胜利成功荣誉类 victory, success, scholarship, reputation, honour; 知识信息经历类 experience, knowledge, information; 还有 control, seat, weight 等等。gain 还搭配消极语义韵词汇如 revenge, notoriety, concession, reprieve (缓刑), respite (缓刑), monopoly (垄断), conviction (定罪)。BNC 中 gain 的宾语呈现出错综语义韵。

2.3.2 arise 和 emerge 在 BNC 中的主谓搭配比较和语义韵分析

使用 Sketch Engine 检索 BNC 中 arise 和 emerge 两个动词的主语搭配,考虑到 BNC 的容量达 1 亿词汇,设检索条件为最小频数为 4,显著值 (Saliency) 为 3。检索结果中前 15 个搭配主语如表 4:

表 4 arise 和 emerge 在 BNC 中的前 15 个搭配主语

Achieve				Gain			
Rank (排序)	Frequency (频数)	Saliency (显著性)	Collocate (搭配词)	Rank (排序)	Frequency (频数)	Saliency (显著性)	Collocate (搭配词)
1	107	8.05	dispute	1	126	7.19	picture
2	664	7.92	problem	2	18	6.46	consensus
3	180	7.92	difficulty	3	79	6.29	pattern
4	340	7.62	question	4	30	6.22	theme
5	50	7.29	confusion	5	6	5.45	ferret
6	162	7.16	situation	6	15	5.23	truth
7	168	7.1	matter	7	42	5.2	evidence
8	60	7.09	liability	8	6	5.18	butterfly
9	29	7.03	vacancy	9	16	5.18	conclusion
10	30	6.94	complication	10	5	5.03	paradigm
11	35	6.92	controversy	11	12	5.0	trend
12	108	6.9	income	12	29	4.87	difference
13	62	6.85	conflict	13	10	4.85	finding
14	102	6.68	opportunity	14	6	4.83	insect
15	40	6.57	obligation	15	5	4.67	compromise

BNC 中 arise 的主语搭配词汇包括问题类 problem, question, confusion, matter, issue, vacancy, 争议冲突类 dispute, difficulty, complication, controversy, conflict, misunderstanding, 还有情景机会收入职责类 situation, opportunity, income, obligation, 呈现出消极和中性语义韵。emerge 在 BNC 中的主语搭配词汇中有图片图案类 picture, pattern, 生物动物类 butterfly, insect, ferret, creature, 共识结

论发现类 consensus, theme, truth, evidence, conclusion, finding, 以及区别妥协类 difference, compromise, 呈现明显的中性语义韵。

检索还得出, arise 在 BNC 中的唯一主语包括 dispute, confusion, matter, controversy, complication, conflict 等争议和麻烦类词汇, emerge 的唯一主语包含 picture, creature, butterfly, insect, consensus, truth, trend, finding, compromise, opportunity 等图

片、生物、事实、趋势类和机会类词汇。

2.4 两组近义词在语料库中的异常搭配和错误分析

Corder 率先提出“错误分析”(Error Analysis)理论^[13]。错误主要分为语际错误(Interlingual Error)和语内错误(Intralingual Error)两类。前者指语言迁移和学习者母语干扰引起的错误。后者指学习者没有完全掌握目的语规则而导致的错误。错误分析提供了调查学习者语言的方法论^[14]。

通过以上分析可得出,BNC 中四个动词的搭配词汇比 WECCL 中对应的动词的搭配词汇更丰富、广泛和具体。例如,WECCL 中没有出现 BNC 与 achieve 搭配频率很高的 objective, results, balance, standard, independence 等词汇,也没有出现 BNC 中与 gain 搭配的 access, control, support, insight, advantage, confidence 等显著搭配词。

WECCL 中四个动词的部分搭配词汇在 BNC 相应的搭配中没有出现或频率非常少,显著值差异很大。例如,BNC 中 gain 没有与 progress, chances, accomplishment, treasure, spirit 等搭配。emerge 在 BNC 中没有 knowledge, function, learners, disadvantage 等主语名词的搭配。通过检索,BNC 中与 progress 搭配表示“取得进步”的动词是 make 或 achieve,与 happiness 和 quality 搭配的最佳动词也是 achieve,“获得机会(chances)”的最佳动词是 get, take 或 win。BNC 中 gain 与 opportunity 搭配了 9 次,但没有与 chance 搭配。这说明 chance 相对 opportunity 来说比较口语化,所以选用的动词也没有 gain 正式。

基于语料库辅助的错误分析不同于传统的错误分析,原因在于计算机处理和调查学习者语言的能力非常强大^[15]。

WECCL 中 achieve 与不定代词 something 搭配了 9 次,而在 BNC 中没有这样的搭配。这与母语迁移不无关系。中文表达“如果你想达到或实现什么”,在 WECCL 中可以找到逐字翻译的英文句子,如例(1)~例(3):

1) If you want to *achieve something* or intend to fulfill one of your ambitions, you must work hard, make efforts and get prepared. WARG1810.txt

这种口语体不适合正式的笔语写作,上句用具体的“achieve your dreams / goals / objectives”更佳。

2) A man who has a great amount of money cannot gain treasure and happiness. WARG0227.txt

英语中 treasure 表示“财富”时,偏重于“金银财宝”,而 wealth 含义较广,既可指大笔的钱财或物质财富,也可指抽象的精神财富。上句是想表示抽象的财富,所以用 gain wealth 更佳,而不是 gain treasure。

3) *Team spirit* is one thing they *gain*. Everyone should bear it in their minds... WARG1818.txt

team spirit 是团体精神,应该慢慢培养而成,而不是赢得,所以用 foster 更好。

同样,WECCL 中 arise 多次被中国学生误用。如例(4)~例(6):

4) Then, it's really important to arise the awareness of citizens of the urgent situation ... WARG0568.txt

5) ... receive higher education is a good thing for it can arise the quality of the whole society. WARG0557.txt

6) Advertisements of fast food arise the children's interests... WEXP0061.txt

arise 是不及物动词,以上三句均被误用为及物动词,应分别改为 raise the awareness(提高意识)、improve the quality(提高品质)、arouse the children's interest(激发孩子的兴趣)。

不及物动词 emerge 在 WECCL 中也被误用为及物动词。如例(7):

7) ... the present examination system is emerging severe drawbacks... WARG2262.txt

上句谓语部分应改为 has severe drawbacks,整句中文意思是“目前的考试体系存在严重弊端”。

以上中介语搭配错误受母语干扰、过度概括和忽略上下文等原因所致。

3 结语

相比 BNC, achieve 和 gain 两词在中国学生语料库 WECCL 中被过度使用,而 arise 和 emerge 两词被中国学习者过少使用。但是,中国学生使用的这两组动词的名词搭配词汇远不如本族语者的丰富和具体,而且搭配词的语义韵单一, achieve 和 gain 两词在 WECCL 没有消极语义韵词汇的搭配, arise 和 emerge 被误用为及物动词。中国学生混淆近义词与意义或词形接近词的用法。

在英语近义词辨析教学中,建议引入数据驱动的词汇学习方法(data-driven lexical approach)和充分利用 Sketch Engine 等检索程序。Sketch Engine 功能强大,能完整呈现近义词的常用搭配

和习惯用法。培养学生自主利用网络学习和思辨能力,利用语料库分析方法自主探索语言规律、发现问题和更正错误。

引入词块理论和语义韵概念到词汇教学尤其是近义词教学中,能够帮助学生更好辨析近义词的区别,全面掌握词汇的使用特点。尽量不要将母语迁移带入目标语学习中,杜绝逐字翻译的习惯。

参考文献:

- [1] Oxford Advanced Learner's English-Chinese Dictionary. 北京:商务印书馆,2014.
- [2] Palmer F R. Semantics [M]. Cambridge: Cambridge University Press, 1981.
- [3] Lyons J. Linguistic Semantics: An Introduction [M]. Cambridge: Cambridge University Press, 1995.
- [4] Taylor J R. Linguistic Categorization: Prototypes in Linguistic Theory [M]. Oxford: Oxford University Press, 1995.
- [5] Louw B. 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, 1993.
- [6] Stubbs M. Text and Corpus Analysis [M]. Oxford: Blackwell Publishers Ltd, 1996.
- [7] Sinclair J. Corpus, Concordance, Collocation [M]. Oxford: Oxford University Press, 1991.
- [8] Xiao R, McEnery T. Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective [J]. Applied Linguistics, 2006(1): 103-129.
- [9] 卫乃兴.基于语料库和语料库驱动的词词语搭配研究[J].当代语言学,2002(2):101-114.
- [10] 陆军.基于语料库的学习者英语近义词搭配行为与语义韵研究[J].现代外语,2010(3):276-286.
- [11] Palmer F R. The English Verb [M]. London: Longman Group Limited, 1974.
- [12] 文秋芳,梁茂成,晏小琴.Spoken and Written English Corpus of Chinese Learners 2.0 [M].北京:外语教学与研究出版社,2008.
- [13] Corder S P. The significance of learner's errors [J]. International Review of Applied Linguistics, 1967(4): 161-170.
- [14] Ellis Rod. The Study of Second Language Acquisition [M]. Oxford: Oxford University Press, 1994.
- [15] Leech G. Learner corpora: what they are and what can be done with them [A]. In S. Granger (ed.). Learner English on Computer. London: Addison Wesley Longman, 1998.

Corpus-based Comparative Study and Teaching of the Collocation Features of English Near-synonymous Verbs

LI Yao, LIAO Shunlin

(School of Foreign Languages, Central South University, Changsha 410083, China)

Abstract: Based on WECCL (Written English Corpus Chinese Learners) and BNC (British English Corpus), this paper makes a comparative analysis of the similarities and differences of collocation characteristics and their semantic prosody of two pairs of near-synonymous verbs. One pair of the verbs are “achieve” and “gain”, and the other pair of verbs are “arise” and “emerge”. It discusses the abnormal collocation and reasons of the four verbs in WECCL. It is found that compared with native speakers, the former pair are overused and the latter pair is less used by Chinese learners; the object collocation words of “gain” are more abundant than “achieve”, but the collocation words of both verbs are not as rich and specific as those of native speakers in BNC, and show no negative semantic prosody; the subject collocation words of “arise” and “emerge” are simple, so is their semantic prosody, and they are misused as transitive verbs. Errors are caused by the first language interference, overgeneralization and ignorance of context. In conclusion, lexical chunks and corpus-driven teaching mode is put forward for teaching and learning near-synonymous verbs.

Key words: corpus; near-synonym; verb; collocation; prosody

(责任校对 蒋云霞)