

基于语料库的英语近义词辨析*

——以suspend, halt, cease和pause为例

靳璐瑶 热孜亚·马力克 刘乐

(新疆农业大学外国语学院 新疆乌鲁木齐 830052)

摘要: 本文基于语料库检索,通过对词频、语域、搭配及语义韵四个方面分析suspend, halt, cease及pause这组近义词之间的差异。研究表明,这四个词语在语义韵方面差异不大,但是在其他三个方面有较为明显的差异。

关键词: 词频 语域 搭配 语义韵

中图分类号: H319.3 **文献标识码:** A

DOI: 10.12218/j.issn.2095-4743.2022.11.123

引言

现代语料库语言学是20世纪中后期兴起的一门语言研究科学。语料库是指在随机采样的基础上收集的、有代表性的真实语言材料的集合,是语言运用的样本(杨惠中, 2002)。语料库以其容量大、语料真实等优势在现代语言学研究 and 语言教育等领域发挥着越来越重要的作用。随着语料库语言学的兴起,人们对语料库的研究越来越多,国内学者主要使用语料库进行搭配、语义韵、语域和类连接等方面差异的检索,其中多数是对动词的辨析,对名词、副词以及形容词等辨析相对较少。也有研究者是用语料库对词进行分类研究,例如陶然(2019)研究分析了“恐惧类”词语,而肖志强(2018)研究的是一组海事英语近义词。这些研究表明,基于语料库的近义词辨析能够更客观、准确地发现传统方法所无法探究的近义词的某些特殊特征^[1]。

英语近义词因其意思相近或相同难以区别,学习者通常会借助词典辨析它们之间的差异。虽然英语词典对学习近义词有一定的帮助,但不能完全满足学习者的需求。本研究以语料库语言学理论为指导,尝试从频率、搭配、语义韵、语域四个方面对suspend, halt, cease和pause这组近义词进行辨析^[2]。

一、研究方法

1. 研究工具

英语国家语料库(British National Corpus, 简称BNC)最初是由牛津大学出版社在20世纪80年代至90年代初创建的,其特点是书面语与口语并存,词容量超过一亿,支持多种通用检索软件,并可直接进行在线检索,能检索搭配、词频、

语域等信息。

2. 研究设计

首先,在语料库中检索四个节点词suspend, cease, pause和halt,得出它们出现的频率以及在不同语域中的分布情况。其次,通过检索搭配获得四个近义词的常用搭配。然后,根据搭配分析四个词的语义韵。最后,分析得出四个词在词频,语域,搭配及语义韵方面的差异^[3-5]。

二、数据分析

1. 不同语域中的词频分布

频数是指某单词或短语出现的次数,而频率是频数标准化后的结果,是指在固定单词量中出现的频数(王慧&王倩文2020)。若要分析suspend, halt, cease和pause这组近义词在不同语域中的词频分布,先要了解它们整体的分布频率。这四个单词在BNC中的词频分布如下(见表1):

表1 suspend, halt, cease和pause在BNC中的词频分布

词汇	频数(Freq)	频率(per mil)
Suspend	401	4.01
Halt	1648	16.48
Cease	993	9.93
Pause	1800	18.00

通过表1我们发现, pause和halt两个词最为常用,使用最少的词是suspend。尽管这四个近义词的词义相近,但它们的内在含义、在不同的语域中的分布特征是不同的。人们在实际的语言活动中出于交际的需要会产生言语变体,这种言语变体被称为语域(register)(杨春霞, 2014)。通过统计分析它们在不同语域中的词频差异(见表2)有助于恰当地使用这些词汇^[6]。

*此文系新疆农业大学2021年大学生创新项目(校级)阶段性成果。

表2统计数据表明,出现频率最高的单词是pause,共1800次;频率最低的是单词suspend,仅401次。在小说语域中,单词pause出现了1046次, halt出现了533次, cease出现了111次,而suspend仅仅只出现了19次,这也说明pause和halt常用于小说等文学创作中。就suspend一词来看,它在非学术语域中出现的频率最高,共153次,表明其主要运用于非学术领域。在口语语域中,四个词的频率相对于其他7个语域来说整体的频率都比较低,说明这四个词不常用于口语。

2. 搭配分析

搭配 (collocation) 是指两个词语之间的“结伴关系”(Firth, 1957)。本文采用一种常见的计量方法MI值 (Mutual Information Score) 来对搭配进行统计研究。MI 值的大小表

明节点词和其搭配词的关系强弱, MI值越大代表搭配强度越高。使用BNC语料库的Collocates模板分别对suspend, halt, pause, cease进行检索, 跨距 - 4/+4检索其搭配词及MI值, 将与节点词共现频数排在前三十位的搭配词的MI值进行排序 (见表3)。

从表3数据中 (人工去除不具备搭配关系的词) 可见, 单词suspend和cease与名词搭配居多, 其中suspend常在表示“吊销执照”或“暂停……资格”时使用, 而cease常用于军事方面, 如在表示“停止敌对活动”或“停火”时, 采用动词cease, 而非suspend。halt大多与动词搭配, 且常为verb+to a halt的形式。pause与形容词搭配居多, 多用作名词词性。此外, halt, pause, cease在搭配词的词性方面具有多样性。

表2 pause, cease, suspend和halt在不同语域的词频分布

语域	Pause		Cease		Suspend		Halt	
	频数 (Freq)	频率 (per mil)	频数 (Freq)	频率 (per mil)	频数 (Freq)	频率 (per mil)	频数 (Freq)	频率 (per mil)
口语	100	10.04	49	4.92	24	2.41	43	4.32
小说	1046	65.75	111	6.98	19	1.19	533	33.5
杂志	62	8.54	36	4.96	24	3.3	87	11.98
报纸	63	6.02	75	7.17	51	4.87	279	26.66
非学术	158	9.58	241	14.61	153	9.28	321	19.46
学术	123	8.02	250	16.31	45	2.94	85	5.54
音乐	248	11.9	231	11.09	85	4.08	300	14.4
总和	1800	18.00	993	9.93	401	4.01	1648	16.48

表3 四个节点词在BNC语料库中的显著搭配词一览表

	suspend		halt		pause		cease	
	搭配词	MI值	搭配词	MI值	搭配词	MI值	搭配词	MI值
1	disbelief	9.47	skidded	10.29	momentary	8.68	amaze	9.58
2	pending	8.15	screeched	10.18	Ros	7.77	hostilities	8.73
3	ANC	7.57	grind	9.44	Guil	7.15	forthwith	8.21
4	judgement	6.8	abrupt	8.97	pause	6.48	exist	7.15
5	voted	6.59	grinding	8.56	lengthy	6.03	payable	5.66
6	licence	6.07	slide	6.33	pregnant	5.77	trading	5.38
7	armed	5.73	drew	5.01	awkward	5.75	feeding	5.02
8	operations	5.66	ground	4.76	brief	5.39	shall	4.84
9	struggle	5.62	came	4.71	slight	5.22	altogether	4.78
10	belief	5.58	call	4.71	breath	4.66	operations	4.49
11	decided	5.45	sudden	4.65	reflect	4.42	immediately	4.28
12	membership	5.35	decline	4.56	answered	4.35	payments	4.28
13	decision	4.79	brought	4.26	moment	4.28	apply	4.21
14	forced	4.77	campaign	3.92	long	3.98	function	3.97
15	agreed	4.69	attempt	3.72	after	3.7	production	3.83
16	action	4.07	called	3.55	replied	3.58	announced	3.47
17	policy	3.56	outside	3.44	added	3.45	never	3.44
18	power	3.53	car	3.43	followed	3.31	activities	3.26
19	until	3.3	finally	3.24	without	3.06	fire	3.18
20	court	3.1	front	3.2	there	3.02	would	3.04

3. 语义韵分析

具有相同语义特点的词项与关键词项在文本中高频共现,后者就会被“传染”上有关的语义特点,整个语境内就弥漫了某种语义氛围,这就是语义韵(Sinclair, J, 1991)。语义韵大体可分为积极语义韵、中性语义韵和消极语义韵三类(Stubbs, M, 1996)。通过语义韵分析,笔者对suspend, halt, cease及pause四个节点词的使用总结出以下规律:

1. 单词suspend所吸引的多数搭配词都属于中性词。

例1: If you want to suspend this policy you must send your certificate to us.

例2: The child will learn to suspend action in order to take further goals into account.

例1中的“suspend policy”(停止政策)及例2中的“suspend action”(暂停行动)均使suspend染上中性语义韵色彩。中性语义韵即节点词的搭配词项大都是中性语义的词汇所形成的中性语义氛围。因此, suspend在大多数情况下呈现中性语义韵的特征。

2. 单词halt多与搭配词构成固定搭配,所呈现的语义韵色彩为积极色彩或中性色彩。

例3: John brought in two years ago to halt the slide of the club.

例4: The cattle-truck came to a halt beside the knot of boys.

在例3中, halt与slide搭配构成短语“halt the slide”,表示“阻止...的下滑(阻止某种不乐观形势的继续)”,该短语在句中表示“阻止某俱乐部继续走下坡路”,呈现一种积极语义氛围。其他多数情况下, halt也呈现中性语义韵的特征,如例4中的“came to a halt”意为“停下来”,不具备积极或消极的语义韵色彩。

3) 与pause搭配的词项多为形容词,与搭配词形成消极或中性语义氛围。

例5: You are left with an awkward pause and this spoils the effect of what has gone before.

例6: There was a momentary pause during which Steve glanced at his wife.

在例5中, pause与表示“尴尬的”的词awkward搭配,具有消极语义色彩,使语境中弥漫消极语义氛围。在例6中, pause与表示停顿时间长短的形容词momentary搭配,既可营造积极的氛围,又可营造消极的氛围,具有中性语义韵。

4) 与cease搭配的多数词具有中性色彩。

例7: Calverton should cease production by November.

例8: The leaders were prepared to cease hostilities against their enemies.

在例7中, cease production意为“停止生产”,不具备积极或消极的语义韵色彩。但在例8中,当cease与hostilities搭配时意为“停止敌对行动”,具有积极语义韵色彩。

结语

本研究在BNC语料库的辅助下,通过对suspend, pause, halt和cease这组近义词在语域、词频、搭配词以及语义韵四个方面的辨析加深了对四个近义词的理解。研究表明, pause和halt使用频率最高。而在不同语域中,单词pause和halt常用于小说, suspend主要用于非学术领域, cease不常用于杂志。四个词在口语中均不常出现。从搭配方面来看, suspend和cease与名词搭配居多,其中suspend常在表示“吊销执照”或“暂停……资格”时使用,而cease常用于军事方面。pause与形容词搭配居多,而halt大多与动词搭配,常为verb+to a halt的形式。语义韵方面, halt和cease极少呈现积极语义韵的特征, pause在与awkward连用时呈现消极色彩。大多数情况下,四个词均呈现出较高的中性语义韵特征。研究表明,语料库有助于英语学习者更好地掌握词汇的真实含义、辨清近义词之间的细微差异,因此通过其丰富的数据来辨析近义词是一个比较可行的方法。

参考文献

- [1] Davies, Mark. (2004 —) BYU—BNC: The British National Corpus. Available online at <http://corpus.byu.edu/bnc> [EB/OL].
- [2] Firth, Papers in linguistics[C], 1957.
- [3] Sinclair, J. Corpus, Concordance, Collocation[M]. Oxford: Oxford University Press, 1991.
- [4] Stubbs, M. Text and Corpus Analysis[M]. Oxford: Blackwell Publishers, 1996.
- [5] 陶然. 基于COCA语料库的“恐惧类”同义词辨析——以“fright”“fear”“dread”“dismay”为例[J]. 现代交际, 2019, (17): 229+228.

作者简介

靳璐瑶(2002—)、刘乐(2000—): 新疆农业大学外国语学院19级学生。

热孜亚·马力克(1975—), 女, 新疆农业大学外国语学院讲师、硕士。