

# 数据驱动学习模式在《新职业英语》教学中的运用探究

刘 静

成都职业技术学院，国际教育学院，四川 成都  
Email: 48087577@qq.com

收稿日期：2020年10月7日；录用日期：2020年10月20日；发布日期：2020年10月27日

---

## 摘 要

数据驱动的学习理念符合大数据时代背景，语料库技术的发展又使数据驱动学习的实现途径更加丰富。本文探究如何用数据驱动学习的理论框架指导《新职业英语》的教学实践，该方法以词项为核心，主张语言是由结构功能一体的各种词项联结而成。通过将词项置于真实的语言材料中，学生可以自主观察并归纳词项的词义、搭配、类联接、语义韵以及语法型式。

## 关键词

数据驱动学习，词项，索引行，扩展意义单位

---

## Using DDL in the Course of *English for Careers*

Jing Liu

School of International Education, Chengdu Polytechnic, Chengdu Sichuan  
Email: 48087577@qq.com

Received: Oct. 7<sup>th</sup>, 2020; accepted: Oct. 20<sup>th</sup>, 2020; published: Oct. 27<sup>th</sup>, 2020

---

## Abstract

DDL (Data Driven Learning) is in harmony with the age of big data. The development in corpus technology enriches the approaches to implementing DDL in teaching and learning practices. This study explores ways to formalize the ideas of DDL in the course of *English for Careers*. Lexis lays at the core of DDL, which assumes that the powerhouse of language is lexicogrammar, a unified stra-

tum of forms and functions embodied by various lexical items. Through observing lexis embedded in authentic language materials, students generalize meanings in terms of collocation, colligation, semantic prosody and grammatical patterns.

## Keywords

Data Driven Learning, Lexis, Concordance Lines, Extended Units of Meaning

Copyright © 2020 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

《新职业英语》(外语教学与研究出版社)是笔者所在的高职高专院校非英语专业学生必修的一门通识平台课程,该课程培养学生具备在本行业领域内运用英语进行基本交流的能力,并能为后续行业专业英语课程学习奠定基础。大数据时代,数据驱动学习模式顺应了高职高专公共英语教学发展的改革趋势,而语料库技术的发展又使数据驱动学习的实现途径得以丰富。本文探究如何用数据驱动学习的理论框架指导《新职业英语》的教学实践,研究所采用的语料是大学英语教材语料库(COLEN)。该语料库由广州中医药大学薛学彦老师创建,语料来源为大学英语教材中的课文,共计 930 个文本,库容 758,754 词次,对该库语料数据的挖掘有助于呈现高等教育阶段在校英语学习需掌握的词项及其延伸结构。

## 2. 数据驱动学习理论及教学理念

伯明翰大学的 Tim Johns 在上世纪 90 年代提出了数据驱动学习的理念(Data driven learning, DDL),鼓励学习者充当语言的研究者,通过观察真实的语言材料,结合上下文概括和归纳语言使用规律,该方法也被称为索引课程(Concordancing Program) [1]。DDL 认为有效的语言学习本身也是一种语言研究,对索引行的观察能促使学习者掌握归纳学习的策略,尤其是辨析异同,推断假设和验证的策略[2]。DDL 主张消解词汇与语法的二元对立,尽可能扩大语言意义单位,根据这种语言观念,词项的概念远远大于词汇,从节点词到搭配和类联接,再到语义倾向和语义韵,是从具体的词汇层面到抽象的语法层面乃至到语用层面的一种扩展,这一整套分析意义的模型被 Sinclair 称为扩展意义单位(Extended Units of Meaning) [3]。

## 3. 通过索引行探究词项

笔者按照数据驱动学习的思路,在《新职业英语》的教学中引导学生自下而上地探究真实语料中的扩展意义单位。该方法的基本思路是教师挑选为教学目的服务的索引行并创建易于观察的视觉界面,学生利用索引行呈现的语境共现信息对词项进行发现式探究。这里所指的词项被视作语言的基本单位,形式上较为灵活和开放,可以是单词,也可以是连续的词语或词块,还可以是非连续的结构或构式[4], P. 106。如图 1,本文从搭配和类联接,语义韵及语法型式三个方面引导学生对词项进行分析。

### 3.1. 搭配和类联接

在讲《新职业综合英语 1》第 3 单元 Reading A 的新单词 conduct 时,教师可以给学生提供相关的语料数据,供其自主挖掘 conduct 不同词性的常见搭配及类联接。教师可以使用北京外国语大学语料库团队

开发的 BFSU Colligator，将经过词性赋码的 COLEN 语料加载其中，再用正则表达式 `conductS*_VV(G|P|Z)*` 匹配出 `conduct` 作动词时的原形，第三人称单数和现在进行时的屈折形式，教师将节点词右边的跨距设置为 4，如图 2，学生可以据此观察 `conduct` 作动词时的宾语名词或名词短语。被动语态需要在节点词左边的跨距寻找动词宾语，增加了观察的难度，为了保持观察界面的简洁清晰，此处检索中排除了屈折形式 `conducted`。

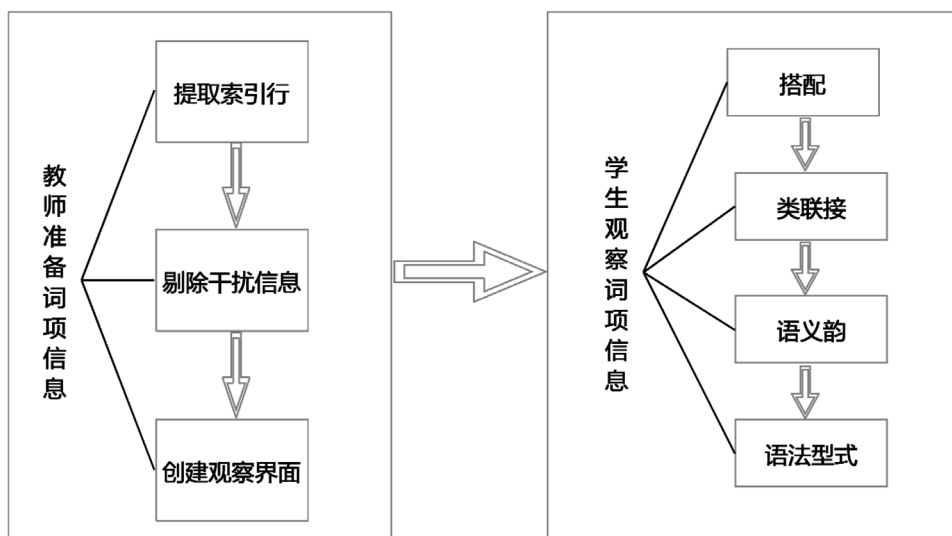


Figure 1. Lexis-based DDL model  
图 1. 以词项为中心的数据驱动学习模式

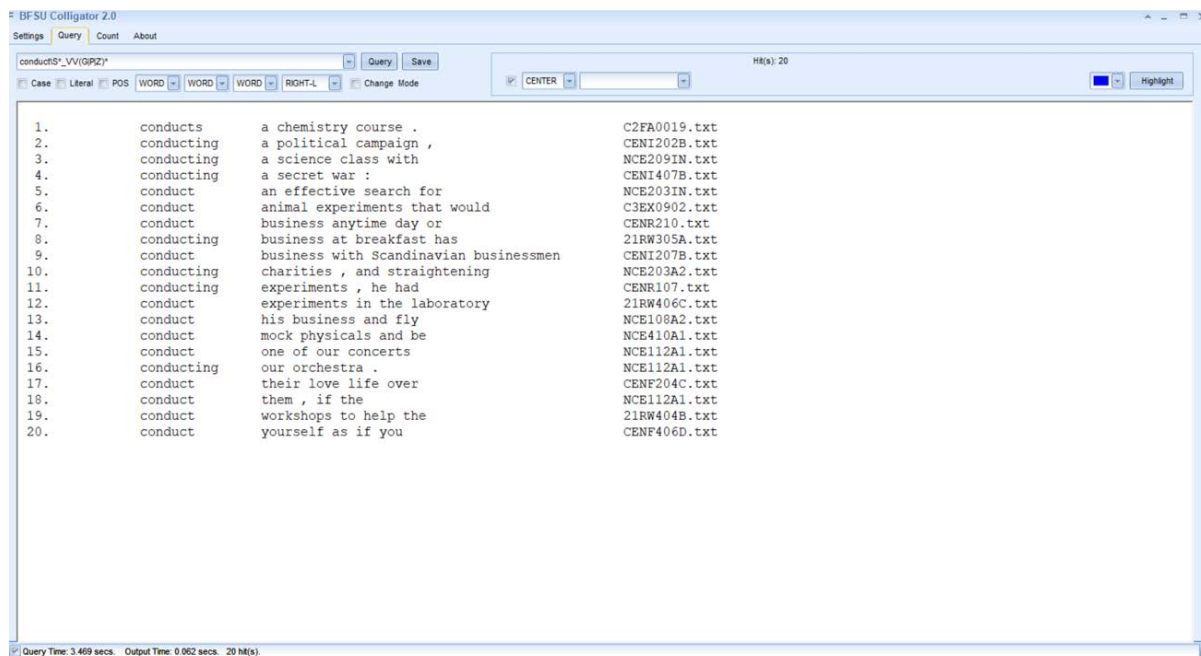
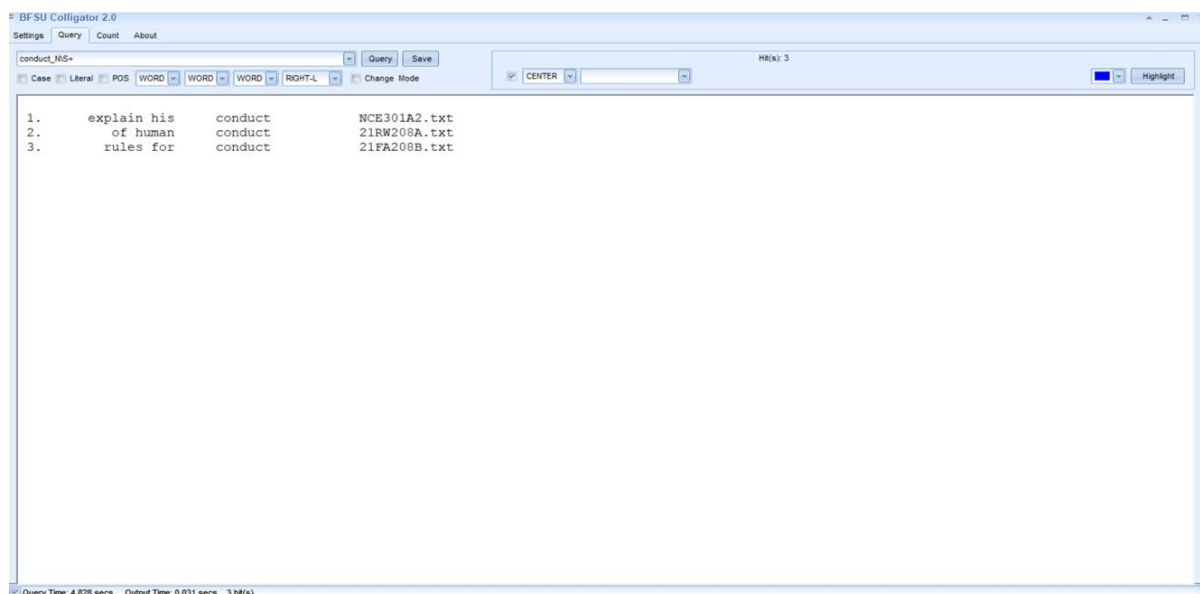


Figure 2. Collocations of conduct as a verb in COLEN  
图 2. COLEN 中 `conduct` 作动词的搭配词

同理，教师用正则表达式 `conduct_N\S+` 匹配到 COLEN 语料库中 `conduct` 作名词的索引行，如图 3：



**Figure 3.** Collocations of conduct as a noun in COLEN

**图 3.** COLEN 中 conduct 作名词的搭配词

学生通过观察图 2 和图 3，可以看出 conduct 既可用作动词，又可用作名词，其中作动词的用法在 COLEN 中更为常见。如表 1，教师可以进一步引导学生归纳出 conduct 两种词性的词义、常用搭配和类联接的典型型式。

**Table 1.** Meaning, collocation and colligation of conduct in COLEN

**表 1.** Conduct 在 COLEN 中的词义、搭配和类联接

| 节点词     | 词性    | 搭配        |             | 数量            | 词义           | 类联接                       |
|---------|-------|-----------|-------------|---------------|--------------|---------------------------|
|         |       | 左搭配       | 右搭配         |               |              |                           |
| conduct | 动词    |           | business    | 4             |              | V+n                       |
|         |       |           | experiments | 3             |              | V(+art)+adj+n             |
|         |       |           | course      | 1             |              | V(+art)+n+n               |
|         |       |           | campaign    | 1             |              |                           |
|         |       |           | class       | 1             | 组织,<br>安排,   | V+poss.adj(+n)+n          |
|         |       |           | war         | 1             | 实施,          |                           |
|         |       |           | search      | 1             | 执行           |                           |
|         |       |           | charities   | 1             |              | V+poss.pron+n             |
|         |       |           | physicals   | 1             |              |                           |
|         |       |           | life        | 1             |              |                           |
|         |       |           | workshop    | 1             |              |                           |
|         |       |           | concerts    | 1             |              | V+card.+of + poss.adj + n |
|         |       | orchestra | 1           | 指挥            | V+poss.adj+n |                           |
|         |       | them      | 1           |               |              |                           |
|         |       | yourself  |             | 举止,<br>表现     |              |                           |
| 名词      | his   | 1         | 行为,         | v +poss.adj+N |              |                           |
|         | human | 1         | 举止          | prep +adj+N   |              |                           |
|         | for   | 1         |             | n +prep+N     |              |                           |

这是一个简单的索引行分析案例，教师利用语料库资源向学生呈现语言事实，鼓励其对某个单词进行深度挖掘，自己动手发现和解决问题，这种数据驱动学习的基本理念符合“以学习者为中心”的思想 [4], P. 109。

### 3.2. 语义韵

英语中很多词看起来不带有明显的感情色彩，但当它们与某些特定词汇为伴构成搭配或类联接时，会产生一些联想意义(如褒贬、好恶等) [4], P. 16, 这种联想意义被称为语义韵。词语搭配形成的语义韵主要分为三类：消极语义韵(negative prosody)、中性语义韵(neutral prosody)和积极语义韵(positive prosody) [5]。在讲《新职业综合英语 1》Unit 6 Reading A 中 adapt 一词时，教师可以使用通配符定义检索项 adapt|adapts|adapted|adapting，然后在 AntConc 中进行检索，得到 19 条结果，如图 4：

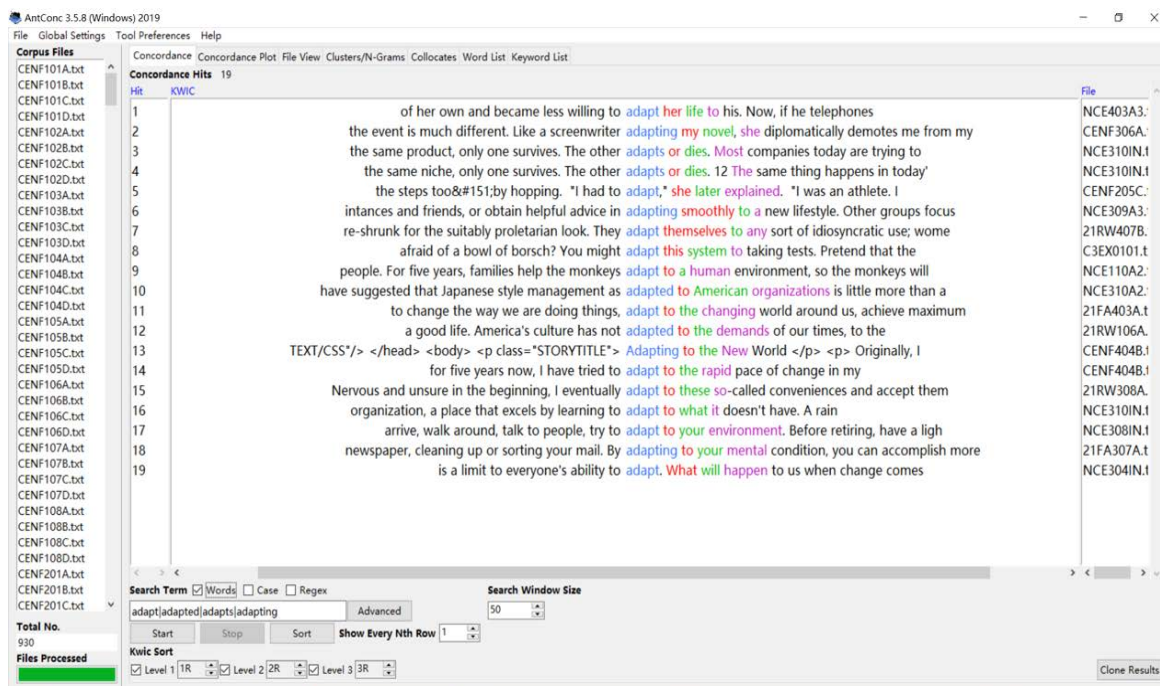


Figure 4. Exploring the semantic prosody of adapt  
图 4. Adapt 的语义韵分析

通过观察，学生可以发现 adapt 后面跟的宾语或者介词宾语为 life (生活)，novel (小说)，new lifestyle (新的生活方式)，idiosyncratic use (新奇的用法)，tests (测试)，environment (环境)，organizations (组织)，changing world (变化的世界)，rapid pace of change (迅速的变化)，demands of our times (我们时代的需求)，the New World (新世界)，conveniences (便利设施)，mental conditions (精神状态)，what it doesn't have (所缺少的条件)。这些宾语有变化、环境、情况、条件等语义倾向，主语不得不出改变与之相适应，否则就面临淘汰或灭亡的境遇，这一点从第 3 句中的 adapts or dies (适者生存)中也能得到证实。由此可见，adapt 一词蕴含着某人或某物具有可塑性、适应性和灵活性的联想意义，具有积极的语义韵。

### 3.3. 语法型式

数据驱动学习同样适用于《新职业综合英语》的语法教学。比如可以运用于《新职业综合英语 1》第 3 单元 Self-study Room 的教学，该模块旨在让学生掌握被动语态的构成和用法。被动语态的时态变化

是教学难点,如果用“先告知语法公式再辅以例句”这种自上而下的方式,难免会让学生觉得概念抽象,难以理解。为了避免学生陷入套用公式,死记硬背的怪圈,可以尝试让学生通过观察和归类语料文本,自主归纳语言现象。在讲到被动语态的过去完成体时,教师可以将经过自动词性赋码的 COLEN 语料库加载到 AntConc 中,然后利用正则表达式  $\backslash S+_{VHD}\backslash(S+_{[RX]}\backslash w+\backslash s)^*\backslash S+_{VBN}\backslash S+_{VVN}$  匹配到该库中所有包含被动语态过去完成体的索引行,共命中检索结果 228 行。如图 5,通过索引软件的排序和彩色显示功能,学生很容易观察到过去完成体被动语态的构成型式是 had been done。同时,学生还能在观察中发现 had 和 been 之间是可以有副词的,比如 never, actually, already, always 等,这是课本中没有提及但却是十分常见的语言现象。学生通过自主观察发现了语言规律,不仅印象深刻还能收获成就感。

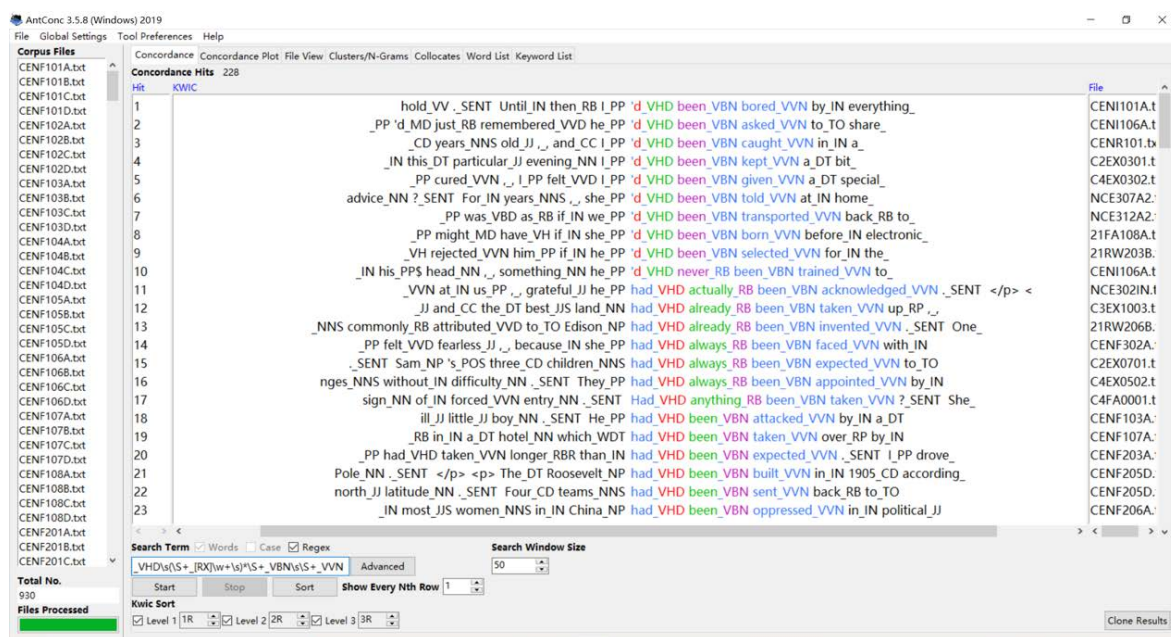


Figure 5. Past perfective aspect of passive voice  
图 5. COLEN 中被动语态的过去完成体

#### 4. 结语

身处大数据时代,教师可及的语料素材呈指数级增长,随之也带来信息过载的问题。如何从海量数据中挖掘出既满足学生职业需要又贴合课程标准的资源从而实现所教即所需,是教师应该思考、探索和解决的问题。数据驱动的学习理念符合大数据时代背景,语料库技术的发展又使数据驱动学习的实现途径更加丰富。本文用数据驱动学习的理论框架指导《新职业英语》的教学实践,通过将词项置于真实的语言材料中,引导学生自主观察并归纳词项的词义、搭配、类联接、语义韵以及语法型式。但囿于笔者自身的语料库技术水平,对语料库和检索工具的使用尚处于初浅阶段。如何使用更复杂的检索技术深度挖掘数据以及如何建立更适合高职高专学生需求的语料库从而使数据驱动理念的实践更加务实,是需要进一步探究的问题。

#### 参考文献

- [1] Tribble, C. and Johns, G. (1997) *Concordances in the Classroom: A Resource Guide for Teachers*. Athelstan Publications, Houston, 37-58.
- [2] Johns, T. (1994). *From Printout to Handout: Grammar and Vocabulary Teaching in the Context of Data-Driven Learn-*

ing. In: Odlin, T., Ed., *Perspectives on Pedagogical Grammar*, Cambridge University Press, Cambridge, 293-313. <https://doi.org/10.1017/CBO9781139524605.014>

- [3] Sinclair, J. (1996) The Search for Units of Meaning. *Textus: English Studies in Italy*, **9**, 75-106.
- [4] 梁茂成, 李文中, 许家金. 语料库应用教程[M]. 北京: 外语教学与研究出版社, 2010: 106, 109, 16.
- [5] Stubbs, M. (1995) Collocations and Semantic Profiles: On the Cause of Trouble with Quantitative Studies. *Functions of Language*, **2**, 23-55. <https://doi.org/10.1075/fof.2.1.03stu>