

Constructing ESAP Corpus Platform and Implementing Project in Local University

Zhao Yunxuan

Department of Foreign Languages, Liaocheng University, Liaocheng, Shandong, China
Zhaoyunxuan888@163.com

ABSTRACT

The academic English teaching in new era has been put on the agenda of serving national strategic development. However, facing the coming transition from English for General Purposes (EGP) teaching to English for Academic Purposes (EAP) teaching, most local universities have shown uncertainty and confusion in the implementation of curriculum reform. In collaboration with College English Learning campus network of the university, this research proposes to construct an open, ubiquitous and individualized ESAP online corpus platform and implement a coordinative, multi-dimensional and technology-enhanced Data Driven Learning(DDL) project in local university. It is hoped to establish a learning community of teachers and students and achieve human-computer cooperation. The present DDL teaching mode empowers students to address their discipline-specific language skill and academic literacy.

Keywords: local university, ESAP, online corpus platform, DDL, teaching mode

地方高校专业学术英语语料库平台建设及应用研究

赵蕴萱

聊城大学外国语学院, 聊城, 山东, 中国
zhaoyunxuan888@163.com

摘要

新时代的学术英语教学被提上服务国家战略发展的议程, 而地方高校面对从通过英语向学术英语的转型, 对改革实施过程表现出不确定性和模糊性。本研究依托已建成的校园网平台“大学英语教学网”, 自主开发开放式、泛在式、个性化的专业学术英语(ESAP)语料库在线平台。并在注意假说和建构主义学习理论的指导下, 提出建构具有协同开放性、多维共生性和智能增强性的跨学科数据驱动学习(DDL)教学模式, 以促成师生学习共同体和人机协作, 培养学生在特定学科或专业里的学术英语语言技能和学术素养。

关键词: 地方高校, 专业学术英语, 语料库在线平台, 数据驱动学习, 教学模式

1. 前言

2018年全国教育大会特别指出“公共外语教学改革要主动服务国家战略发展”, “要大力培养具有全球视野、通晓国际规则、熟练运动外语、精通中外谈判和沟通的国际化人才, 要有针对性地培养一带一路等对外战略建设亟需的懂外语的各类专业和管理人才等”。教育部高教司司长吴岩(2019)提出要培养“一精多会”、“一专多能”9的国际化复合人才。特别是在高等教育国际化和

“一带一路”倡议的背景下, “不同专业的人才必须具备较强的国际交流能力, 才能有效地参与国际交流、国际合作、国际竞争, 才能有效地在不同领域、不同行业发出中国声音、展现中国特色、建立世界标准”[1]。这对大学外语教学如何全面把握学术英语能力内涵, 如何有效地实施学术英语教学提出了更高的要求。

与清华大学、中山大学等将跨学科的通用学术英语或专业学术英语能力系统地纳入培养目标和课程体系的985高校相比, 学术英语能力培养问题相对而言仍是目前地方高校大学英语教学的“短

板”。具体表现在大学英语教学仍然以“提高学生英语水平和应对考试的语言教学”的通用英语为必修课[2]，所开设的通用学术英语(EGAP)一般安排在四级考试后的选修课程中，来自不同学科和不同专业的学生在同一课堂中学习，这类不区分特定学科和专业领域的EGAP课程有堕落为通用英语教学的危险，因为“如果不结合学生专业学习中的各种语类或各种典型场景的语言表达规律和表达方式专门的语言训练的，任何学术英语都是徒有其名”[3]。另外，地方高校学术英语教学资源相对贫乏，除了教材之外几乎没有其它辅助资料；课堂教学未实现技术优势与资源共享，与现代教育技术融合不足；大部分大学英语教师对跨学科开展语言教学缺乏信心，甚至产生焦虑情绪。集中表现为学生对用英语开展专业学习的需求与地方高校学术英语教学不能充分满足此要求的矛盾突出。大数据时代背景下，高校人才学术英语能力培养需要注入全新理念，大数据、智能、共享、碎片化、合作等关键词需要渗透到新型学习生态系统建构中。为此，本研究将研究视角转向地方高校，提出建构具有协同开放性、多维共生性和智能增强性的专业学术英语学习生态系统，打造全方位高校人才学术英语能力培养模式。

2. 研究背景

2.1. 国内外学术史梳理及研究动态

国外学术英语教学发展成熟，主要有三大流派，即“基于体裁的EAP教学”、“基于素养实践的EAP教学”和“基于语料库的EAP教学”。

体裁教学法模式下(Genre Analysis approach)有关语言教师借助体裁分析法帮助学生学习科研论文的篇章结构和语言特征的研究此起彼伏(如Cargill et al. 2017[4])。学术素养论(Academic Literacies Approach)主张扩大学术视野，关注社会政治情境因素，强调学科社会化的话语实践。Wingate(2015)认为将学术素养教育完全整合到课程中的包容性教学法可以更好地满足多样化学生群体的需求[5]。一个典型的例子是英国剑桥大学的Transkills项目。Boz介绍了如何使用来自导师和大学一年级学生的调查数据来确定从高中到大学学习的过渡过程中涉及到的群体问题，并将这些材料纳入了学生的学科写作资源中(见Charles&Pecorari, 2016[6])。然而，在教学法层面，实践型EAP很大程度上仍处于意识形态层面，尚未形成可操作的方法体系。数据驱动学习(Data Driven Learning)旨在将语料库语言学(CL)的工具和技术用于教学目的。由于不同学科所表达的知识各异，其阐述观点的语言修辞方式也有很大差异，DDL逐渐成为引领EAP教学和相关研究的新方向。例如，Anthony(2018)研究了EAP写作课堂中DDL的

使用工具和应用策略[7]；Wong(2019)介绍了她在香港大学所尝试实施的一项基于在线学术英语语料库查询平台的DDL跨专业研究生论文写作项目[8]。

国内相关研究自20世纪80年代起开始关注学术英语，但八九十年代学术英语教学研究较为罕见，进入21世纪以后相关研究热度才逐步提升。

2000年到2009年学术英语教学相关研究为萌发阶段，在争议中缓慢发展，成果主要集中在介绍国外、我国台湾以及中外合作办学高校等的学术英语课程建设经验和成果。2010年到2013年为突破阶段，研究者开始将目光转向清华大学、北京大学、医学院等985高校及专业性大学的学术英语课程建设的运行思路和实施实践。这一阶段“学术英语课程建设研究主要阐述了课程总体建设的可行性和取得的成效，但对课程建设的具体过程以及各环节存在的微观问题探讨不足”[9]。2014年至今为利用现代教育技术深度开发阶段，如何利用信息技术提升新时代大学生专业学术英语能力成为大学英语教学改革的热点。在国家政策与战略指引下，基于MOOC(慕课)、SPOC(私播课)、翻转课堂和在线语料库的学术英语教学新模式研究如雨后春笋般破土而出(如吴玲娟, 2015[10]；王奕凯, 刘兵, 2019[11]；刘萍等, 2018[12]，2020[13])，尝试用技术去创新学术英语教学，引领大学英语教学改革，以满足国家、社会和学生个人对专业学术英语应用能力提升的需求。

为此，本研究借鉴国内外EAP教学研究的先进理念和教学模式，将兼具用户友好性和可视化效果的语料库平台和强调学习者课堂内外的主动参与和发现式学习的数据驱动学习(DDL)与学术英语教学进行有效地融合，体现了学术英语领域最新的教学研究成果。

2.2. 核心概念

2.2.1. 学术英语

学术英语(English for Academic Purposes, EAP)是专门用途英语(English for Special Purposes, ESP)中最常见的一种语言教学。学术英语的对象一般为高校层次的大学生，重点是培养学习者的学术语言和学术素养，具体表现为培养学生阅读专业文献、撰写论文等学术语言能力，批评分析文本、综述信息等的学术认知能力和查找和梳理文献、处理数据等的学习策略。学术英语再细分为通用学术英语(English for General Academic Purposes, EGAP)与专业学术英语(English for Specialized Academic Purposes, ESAP)。通用学术英语强调各学科英语中的共性，课程倾向使用一般学术内容来实施，认为所教授的通用语言技能可以转移到学生自己的专业领域。专业学术英语与特定学科专业紧密相连，关注特定学科学生的特殊需求。

2.2.2. 数据驱动学习

数据驱动学习 (Data Driven Learning, DDL) 旨在将语料库语言学 (CL) 的工具和技术用于教学目的。该方法强调学习者课堂内外的主动参与和发现式学习, 引导学生借助语料库平台, 观察来自真实、自然语境中目标话语的聚合和共现关系, “发现词汇、语法范畴、语义和语用等不同语言层次之间的共选关系” [14]。当前学术界普遍达成共识, 语料库和语料库工具具有极大的教学潜力, 可以增强学习者的语言意识和元认知技能, 包括分析能力, 解决问题的能力, 协商能力和基于团队合作的发现学习能力语言学习, 促使其养成语言学习的主人翁意识。

3. 自建 ESAP 语料库在线平台的建设

3.1. 校园网平台的支持

校园网“大学英语教学网”平台提供网络平台支持, 在历经大学英语两轮教学改革后, 大部分地方高校基于校园网建有较成熟的大学英语网络平台, 学术英语语料库和相关语料库检索软件只需要加入此平台。另外, 网络机房提供硬件支持, 大部分院校都配备用于视听说学习的网络机房, 这为语料库驱动学习模式提供了硬件支持。教师可以在机房每台计算机上预先安装必要的软件和语料库, 学生端计算机可以共享教师端操作界面, 确保教师操作演示的可视化在受控练习和自由任务阶段, 教师鼓励学生分组合作学习, 互相获得技术帮助。基于上述平台和硬件的支持, 本研究依托已建成的校园网平台“大学英语教学网”, 自主开发开放式、泛在式、个性化的专业学术英语 (ESAP) 语料库在线平台, 实现资源共享, 突破时空的限制, 将教师和学生置于一个不受时间和地点限制的教学模式中。

3.2. 自建 ESAP 语料库的研制

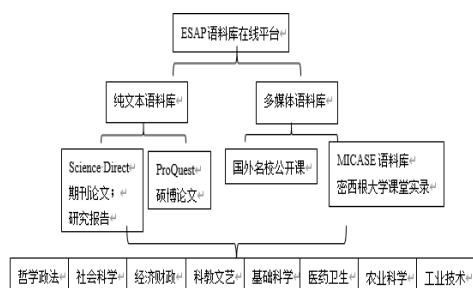


图 1 ESAP 语料库结构图

如图 1 所示, 该平台由纯文本语料库和包括音频、视频的多媒体语料库构成。纯文本学术英语语

料库由各院系的专业课教师推荐构成, 选取 2010 年以来 Science Direct Online 论文库中的期刊论文和研究报告以及 ProQuest 学位论文库中的硕博论文资源, 分为哲学政法、社会科学、经济财政、教科文艺、基础科学、医药卫生、农业科学和工业技术 8 个大类, 然后按照学校的 98 个本科专业和 29 个硕士学位点进行更精细的专业分类, 来自不同院系的学生可选择各自学科的语料库, 使学习更有针对性。学术英语多媒体语料库借鉴南华大学语料库平台的建库模式 [15], 同样按照学科分类选取 2000 年以来国外名校公开课和源自 Michigan Corpus of Academic Spoken English (MICASE) 语料库的密西根大学课堂实录的相关视频、音频等资料, 体现各学科专业近年来的研究热点。建构 ESAP 语料库在线平台将在很大程度上解决地方高校专业多、差异大、课程资源匮乏等一系列难题, 还可以实现资源共享, 突破时空的限制, 同时还可作为地方高校大学外语数字化教学平台协同建设机制的一部分。

专业学术英语 (ESAP) 语料库的建库思路如下: 第一, 学术语料收录的原则、方法和途径: 主要参照书面语料库 LOB, BNC 学术分库、SPACE 和口语语料库 MICASE, BASE 等大型学术英语语料库的收集方法和途径, 考虑语料收集的可行性以及语料的学术话语定位、代表性和实用性、时代性、规模大小和难度等元素。第二, 语料库元信息标注和论文结构标注: 原信息标注包括语料的产出时间、地点、作者、来源和所属学科等, 由于各学科论文结构的多样性, 结构标注包括 ABSTRACT, INTRODUCTION, LITERATURE REVIEW, METHODOLOGY, RESULTS 和 FINDINGS, DISCUSSION, CONCLUSION 各部分。第三, 文本转换软件和语料库分析工具的选择、培训和应用: 下载论文后使用 AntFileConverter 1.2.1 将 PDF 文档转换为纯文本形式; 拟使用一款绿色免费检索软件 AntConc 3.4.4w, 并进行相应的语料库检索技术培训, 该软件具有 Concordance, Collocates, Word List, File View, N-Gram, Keyword List 和 Part of Speech (POS) 等多样化的语料库检索功能。

4. 数据驱动学习 (DDL) 教学模式的建构

4.1. 理论基础

支撑数据驱动学习 (DDL) 的理论基础为注意假说和建构学习理论。根据注意假说 [16], 语言学习者需要注意输入的语言功能, 才能获得最大的习得机会, 而数据驱动学习通过提供真实语境中的学术语言来吸引学习者关注目标功能。建构主义学习理论认为知识并非通过被动接受获得, 而是学习者从自身经验世界出发, 通过“做中学 (Learning by doing)”和“发现学习 (Discovery learning)”去建构新知识。该理论指导下的教师角色是课堂活

动的组织者，任务情景的创设者，更是帮助学生建构意义的协助者和学生正确使用学习工具的指引者。

4.2. 协同开放性、多维共生性和智能增强性 DDL 教学模式

DDL 模式贯穿课堂内统一学习和课堂外自主实践活动全过程。课堂学习中，学习者在教师的指导下通过与语料库的互动进行归纳式的语言学习练习；课堂学习外，学生利用平台自主确认课堂讨论的结果或完成与专业相关的学术论文阅读和写作等。

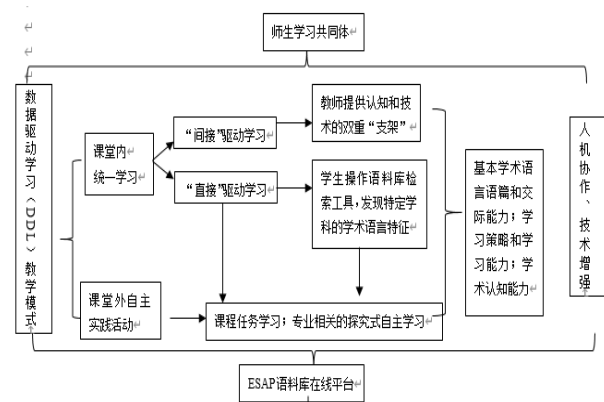


图2 数据驱动学习 (DDL) 教学模式

如图2所示，课堂教学以“间接”驱动学习开始，即教师预先准备好目标任务的语料库搜索结果以及相关的问题和练习，让学生使用教师准备好的讲义“间接”与语料库进行交互，并在教师的引导下对语料进行分类、归纳和提炼。“间接”驱动学习依赖教师提供的“支架”，提高学生对学术英语特征的认知或整体敏感性。课堂教学以“直接”驱动学习结束，并将延续到课堂学习外的任务学习和专业相关的探究式自主学习中，即教师指导学生“直接”通过检索语料库工具识别目标单词、词簇、语法和语篇在学术语境中的使用模式。“直接”驱动学习通过人机协作展开探究基础上的自主学习，并让多次数、多地点、多时段的学习成为可能。重要的是，“直接”和“间接”的语料库驱动学习都会促进有用的学习转移，帮助学习者发展分析能力，解决问题能力，谈判技巧和基于团队合作的发现学习技巧等有用的学术认知能力。特别需要注意的是，该模式下教师不需要成为目标学科和专业的专家，他们可以扮演“支持者”、“引导者”和“组织者”的角色，帮助学生在其专业领域学习中扮演“探究者”、“发现者”和“合作者”的角色。

数据驱动学习 (DDL) 将语料库检索软件 AntConc3.4.4w 所呈现的 Concordance, Collocates,

Word List, File View, N-Gram, Keyword List 和 Part of Speech (POS) 等多样的语料库检索功能用于教学目的，引导学生借助 ESAP 在线语料库平台，观察学术论文和学术会话语境中目标话语的聚合和共现关系。具体而言，引导学生利用技术优势分析这些专业内容是通过什么语言结构、修辞手段和元话语运用来建构和表达的，进而习得包括搭配、构词和含义在内的目标学术单词和词簇；通过语法-词汇相关性和语篇层次因素去习得语法结构；利用平台进行学术语篇体裁分析、话语分析和语用分析等，帮助学生掌握其专业领域所需要的语言和语篇能力。以下述范畴为例：

第一，词汇范畴。通过检索语料库工具中的 KWIC (语境中的关键词) 又称 Concordance (语境共现)、Collocates (搭配词)，学习者可以识别目标单词和短语在学术语境中的使用模式。例如，通过检索目标单词“data”的 KWIC (语境中的关键词)，学习者可以从软件呈现的结果中推断出“data”通常以复数形式出现，并且还可以作为“structure”，“method”和“type”等名词的前置修饰词。

第二，语法范畴。通过 KWIC (语境中的关键词) 又称 Concordance (语境共现)、Word List (词频表)、Collocates (搭配词) 等检索工具和通配符可以实现英语语法引导式数据驱动学习。例如，可以检索学术语料库中“have been *_v”的实际例子，归纳使用模式和用法；通过检索“that”索引并筛选后，生成转述动词词表和检索行，可以观察学术语料库中 show, illustrate, indicate, demonstrate, observe, predict, confirm, speculate 等多样性的转述词，并归纳转述句的用法。

第三，语篇范畴。通过检索语料库工具中的 Word List (词频表)、File View (浏览文件)、N-Gram (N 词词频表)、Keyword List (关键词表) 帮助学生关注学术语篇的语域和语体特征。通过使用 Keyword List (关键词表) 工具将目标语料库中的单词频率与参考语料库 (通常是通用语言语料库) 中的单词频率进行对比，可以识别目标语料库出现高频率的关键词，被视为目标语料库中的代表词或特征词。例如，通过 Keyword List 工具检索有关“cell adhesion (细胞粘附)”的目标英语学术论文，发现其代表词或特征词为 cell, integrin, adhesion, by, in, β , human 等。

第四，结构范畴。由于各学科论文结构的多样性，本研究分别标注 ABSTRACT, INTRODUCTION, LITERATURE REVIEW, METHODOLOGY, RESULTS 和 FINDINGS, DISCUSSION, CONCLUSION 各部分。基于论文结构的标注可以帮助检索者获取有关论文各个部分的关键特征。

4.3. “教”与“学”的双重挑战

基于 ESAP 语料库的数据驱动学习 (DDL) 教学模式在培养学生基本学术英语语篇和交际能力的

同时,帮助学生初步掌握特定学科特有的知识建构方式等学科研究技能,以及如何使用网络查找文献、数据处理、梳理文献等实践技能,提高学生自主学习、沟通和合作等学习能力。

该教学模式的建构是一个系统工作,面临着来自“教”和“学”的双重挑战,一方面,大部分大学英语教师没有接触过语料库语言学,不熟悉数据驱动学习(DDL)方法,如何帮助教师克服技术门槛,改造传统教学方式是本研究的一个难点。首先要让教师意识到该教学模式可以减轻教师负担,因为数据驱动学习更加专注于语言分析和解释的“方式”,可以减少为课程准备相关学科特定材料的备课量[7]。此外,数据驱动学习不需要教师成为目标专业学科的专家,他们可以扮演“辅助者”的角色,帮助学习者在其专业领域中扮演语言“研究者”的角色来完成任务。还可以组织教师定期进行语料库语言学培训和开展基于语料库的课堂活动的教研活动。另一方面,习惯了被动接受学习模式的学习者对探究基础上的发现式自主学习模式可能会出现抵触情绪。教师如何支持、引导和组织学习者克服特别是初始阶段的目标不明确、操作耗时长,效率低下等问题是本研究的另一难点。这需要借鉴超星学习通、云班课、QQ群等提供的技术沟通平台支持,通过师生、生生互动在线讨论并解决基于学术英语语料库的学习过程中遇到的检索技术问题,平台还可以聘请两到三位专长语料库操作的学生助教,随时关注并回答群里的提问。

5. 结论

该模式建构的实现需要解决以下两个难题,一是在线 ESAP 语料库平台的可用性,即语料的体裁、专业、规模和内容的代表性要满足师生的需求,不能呈现出过于狭窄或过于笼统的语料视图。二是提高该教学模式的可操作性,对 ESAP 语料的选择和分类,对分析软件的选择和培训,对如何引导学习者以专业内容为媒介,去“注意”和“发现”学术词汇、语法范畴、语义和语用等不同语言层次之间的共选关系以进一步“建构”新知识进行深入解析和具体探讨。总之,本研究尝试整合平台服务,实现学术英语资源的网络共享,突破时空的限制,将教师和学生置于一个不受时间和地点限制的教学模式中,并通过人机协作,实现技术优势,进而促进计算机智能与学术英语教学的深度融合。

项目基金

本文为山东省人文社会科学一般项目《基于大数据的当代美国学术期刊中孔子形象的变化与接受度研究》(19-ZZ-WX-12)的阶段性成果之一;聊城大学博士科研启

动基金项目《基于口语语料库的中国英语学习者话语标记语多层次对比研究》(321051905)的阶段性成果之一

REFERENCES

- [1] Wang, N. & Zhang, J.Y. (2018) Reflection upon In-Depth Integration of Information Technology and Foreign Language Teaching—Based on Technology Integrated College English Classroom Teaching Reform. *J. Technology Enhanced Foreign Language Education*, 05: 3-7.
- [2] Cai, J.G. (2019) A revisit to the EAP Needs Analysis of University Students of Non-English Major and Specialists under the Background of “Double First-Class” Initiative. *J. Foreign Language Education in China*, 02: 48-54+92.
- [3] Cai, J.G. (2019) Academic English Specificity Revisited: EAP or EGP? *J. Journal of Xi'an International Studies University*, 27: 7-11.
- [4] Cargill M et al. Scientists publishing research in English from Indonesia: Analysing outcomes of a training intervention to inform institutional action [A]. In Cargill M & Burgess S (eds.). *Publishing Research in English as an Additional Language: Practices, Pathways and Potentials* [C]. Adelaide: University of Adelaide Press, 2017. 169—186.
- [5] Wingate, U. (2015). *Academic Literacy and Student Diversity: The Case for Inclusive Practice*. Bristol, UK: Multilingual Matters.
- [6] Charles, M. & Pecorari, D. (2016). *Introducing English for Academic Purposes*. London and New York: Routledge.
- [7] Anthony, L. (2018). Introducing corpora and corpus tools into the technical writing classroom through Data-Driven Learning (DDL). In J. Flowerdew and T. Costley (Eds.), *Discipline Specific Writing* (162–180). Abingdon, UK: Routledge.
- [8] Hyland, K., & Wong L. L. C. (2019). *Specialised English: New Directions in ESP and EAP Research and Practice*. London: Routledge.
- [9] Zuo, X.Y. & Ning, Q. (2019) Review and Prospect of EAP Teaching Research in China since 2000. *J. Foreign Language World*, 191:73-81.

- [10] Wu, L.J. (2015) TED-Ed-Based Research on the Flipped Classroom of EGAP Listening and Speaking. *J. e-Education Research*, 36:81-87.
- [11] Wang, Y. K. & Liu, B. (2019) A Study on the Practice of a Blended Academic English Program for Postgraduates. *J. Foreign Languages and Their Teaching*, 05: 10-19+147.
- [12] Liu, P. & Liu, Z. X. (2018) A Study on Effectiveness of ESP Corpus-based Learning Method of Academic English Vocabulary. *J. Foreign Languages Research*, 35: 54-60.
- [13] Liu, P., Zeng, W. H., Liu, Z. X. (2020) Corpus-based Data-driven Learning of ESP Vocabulary for Doctoral Students. *J. Foreign Language Research*, 01: 64-69.
- [14] Xu, J. J. (2019) *Corpora and Discourse Studies*. Foreign Language Teaching and Research Press, Beijing.
- [15] Yuan, Y. L & Ge, L. L.(2019) A Corpus-Based Study of the Construction of Academic English Teaching Platform and its Teaching Model. *J. Journal of Xi'an International Studies University*, 27: 59-62.
- [16] Schmidt, R. (2001) Attention[A]. In P. Robinson (Eds.), *Cognition and Second Language Instruction[C]* (3-32). Cambridge: Cambridge University Press.