

doi:10.16104/j.issn.1673-1883.2018.04.012

ESP书面文本的句法复杂度分析

——从语体功能侧面展开的研究

卫志强, 卜 伟

(马鞍山职业技术学院应用外语系, 安徽 马鞍山 243031)

摘要:为研究ESP书面文本在句法复杂度方面具有的特点以及这些句法资源所要实现的文本功能,以自建语料库为基础,借助语体多维分析框架,对ESP和EGP两类文本做对比分析。结果显示:相较于EGP,ESP在句法复杂度和语体两方面特征明显;ESP范围内,句法复杂度对文本互动性、指称明晰性、显性劝说性和即席信息组织精密度具有显著的预示作用。研究证明不同语域文本在句法层面体现出的差异性有利于实现各自文本的特定交际目的。

关键词:句法复杂度;语体;文本功能

中图分类号:H319.3 **文献标志码:**A **文章编号:**1673-1883(2018)04-0062-04

A Study on the Syntactic Complexity of ESP Written Texts from the Perspective of Textual Functions

WEI Zhi-qiang, BO Wei

(Applied Foreign Languages Department, Ma'anshan Vocational College, Ma'anshan, Anhui 243031, China)

Abstract: To study the unique syntactic complexity features in the ESP written texts and the textual functions realized by those syntactic resources, An ESP corpus is built, based on which a contrastive analysis of ESP and EGP text is conducted by adopting the framework of multi-dimensional analysis. Results show that compared with EGP, ESP text is significantly different from EGP in syntactic complexity and textual dimensional features. In ESP texts, syntactic complexity can significantly predict the degree of textual interactivity, referential explicitness, overt persuasion and the online informational elaboration. This study demonstrates that differences on syntactic levels among texts from different registers are conducive to the realization of textual functions.

Keywords: syntactic complexity; style; text function

Hutchinson & Waters 认为 ESP 文本的语言描写是关乎 ESP 课程设计的重要因素之一^[1]。有关 ESP 语言形式的研究在数量上十分丰富,但多聚焦于 ESP 文本中的词汇。从句法维度对其做出的分析并不多见。因此从理论上来看,十分有必要填补这一空缺。

在我国,ESP 存在的理据已经得到充分的论证。ESP 课程也将成为今后一段时期高校外语教改的主要方向之一,胡壮麟等指出 ESP 与系统功能语言学中的语域理论相一致,他们建议对这种语言变体需要做出认真的研究与分析^[2]。在此背景下,我们首先收集和自建 ESP 语料库,从句法复杂度这一侧面对 ESP 文本做形式及功能上的探析。

一、理论背景

过往研究中对口语或书面语句法复杂度开展的测量主要为了反映学习者语言发展的水平和所处的阶段^[3];分析和调查诸如体裁、话题、母语背景、语言熟练程度等对变量句法复杂度施加的影响^[4-7];或从心理认知的角度探究课堂教学环境下任务难度和语言产出句法复杂度之间的关系^[8-9]。

这些研究无疑能够加深我们对个体的句法复杂度发展的认识,如 Ryshina-pankova 指出的那样,句法复杂度作为一项词汇语法资源,受到文本表达复杂内容和功能性需要的驱使,可用来构建不同的语义,实现不同的交际目的^[10]。从功能语言学的角

收稿日期:2018-06-24

基金项目:安徽省高校人文社科重点研究项目:专门用途英语书面文本的句法复杂度及语体特征研究(SK2018A1001)。

作者简介:卫志强(1983—),男,安徽合肥人,硕士,讲师,研究方向:语料库语言学。

度看,词汇语法系统居于语言的形式层,它体现的是语言的意义。不同的意义构型(概念、人际和语篇三种意义的组合方式)则形成了不同的语域。作为一种独特的语言功能变体,ESP文本在句法层面也应体现出有别于它类文本的特征。从目前情况来看,ESP句法复杂度方面的研究还较少。另一方面,已有的句法复杂度研究多半停留在对文本结构及形式特征的描写上。Biber认为,脱离语境对句法复杂度形式特征做出的研究不能揭示这一概念的本质^[1]。Ortega建议未来的研究中应该对句法复杂度的功能和交际价值给予足够重视^[2]。

而从语言功能角度对语言形式特征做自下而上的概括性分析的主流方法是由Biber所创立的语体多维分析法。在应用语料库对口笔语语言变异进行研究时,Biber对两类语料中的67种语言特征进行语法标注,基于某些特征的频繁共现,他使用因子分析法,共提取出具有代表性的6个主要语体维度^[3]。通过对每个维度所含语言特征的微观分析,确定它们所共享的语言功能,并根据它们为每个维度命名。Biber等认为语言特征之所以能够共现就是因为它们具有一致的交际价值,而多维度分析法就是从功能的角度对那些通过量化方法确定的语言模式做出解释的框架^[4]。鉴于此,本文将采用这一语体功能分析框架,通过对ESP及EGP语料的对比,在描写ESP文本句法复杂度特征的基础上,着力探讨该类文本中句法资源所要实现的意义和功能。

二、研究设计

(一)研究问题

1. ESP和EGP文本在语体框架的6个维度上是否具有显著差异?
2. 两类语料在句法复杂度方面有无显著差异?
3. ESP的句法复杂度与哪些语体维度的关系更为紧密?前者对后者具有怎样的预示效力?如何解释两者间的关系?

(二)语料简述

本研究所有ESP语料均来自近10年中国国内公开出版发行的大学阶段专门用途英语教材。涉及学科有机电、农林、财会、商贸、旅游、医药卫生、电子信息、轨道交通等。内容只包括阅读文章中的正文部分,但不包括图、表、注释、篇章及段落的标题等附属文字。通过随机抽样并用于研究的ESP文本为300份,语料容量为130 080词,平均每篇ESP文本长度为433.60词。用于参照的EGP语料源自

大学公共英语教材,文本数量也为300,语料容量为143 679词,平均每篇EGP语料的长度为478.93词。

(三)工具与方法

在语料制备过程中,我们使用办公扫描设备对纸质文本进行扫描,获取文本图像,使用ABBYY12.0文字识别软件将图像文件转换为纯文本格式,对文本识别出现的错误进行人工校对,确保语料真实有效。

研究使用L2SCA3.3.3句法复杂度分析器对语料进行处理,从而获取文本句法分析数据。据研发者陆小飞教授的报告,该工具与人工句法标注间的相关信度系数高达0.834以上^[7]。处理完毕后输出的结果会呈现出文本的单词、句子、动词短语、小句、T单位、从属小句、复杂T单位、并列短语及名词短语的数量,利用这些句法单位数据可计算出14种句法复杂度指标值。此外我们使用MAT1.3多维标注与分析工具对两类语料进行赋码,对所有语言特征做出统计,并计算两类语料各语体维度的分值。有关句法复杂度指标及其定义,语体维度名称及主要载荷因子详见表1、2。

最后,我们对上一步获取的句法复杂度及语体维度数据作出整理,并输入SPSS20.0社科统计软件包中。选择合适的统计方法对数据进行分析,并对研究结果做出恰当解释。

表1 句法复杂度指标名称及定义

指标	定义	指标	定义
MLC	平均小句长度	DC/T	T单位中从属小句比率
MLS	平均句子长度	CP/C	并列短语比率
MLT	平均T单位长度	CP/T	T单位中并列短语比率
VP/T	T单位中动词短语比率	T/S	并列句比率
C/T	T单位内部复杂度	CN/C	小句中名词短语比率
CT/T	复杂T单位比率	CN/T	T单位中名词短语比率
DC/C	从属小句比率		

表2 语体维度及代表性载荷因子

序号	维度名称	代表性载荷因子
1	交互性	私动词、that省略、缩略词、现在时、第二人称代词
2	叙事性	过去时动词、第三人称代词、完成体动词、公动词、句法否定
3	指称明晰性	WH型宾语从句、伴随介词结构、WH型主语从句、并列短语
4	显现劝说性	不定式、预测情态动词、劝说动词、条件从句、必要型情态动词
5	抽象性	连词、无主被动、过去分词状语从句、by型被动句
6	即席信息精密性	指示代词、that型宾语从句、宾位that型定语从句

注:限于篇幅,未能穷尽所有因子

三、结果与讨论

由表3可知,ESP与EGP文本在所有六个维度上的差异都达到显著水平($p < 0.05$)。除维度4效应

表3 ESP与EGP语体特征独立样本t检验

	ESP		EGP		t值	p值	效应量
	均值	标准差	均值	标准差			
维度1	-9.857	8.209	-0.411	9.953	12.681	0.000	-0.949
维度2	-3.442	2.375	-0.624	3.460	11.631	0.000	-0.814
维度3	6.971	4.556	2.441	3.681	-13.397	0.000	1.231
维度4	2.866	5.715	1.788	4.987	-2.463	0.014	0.216
维度5	2.765	3.736	0.828	3.204	-6.814	0.000	0.604
维度6	-1.589	1.451	-0.430	2.001	8.116	0.000	-0.579

量偏低外(*Glass's* $\Delta < 0.50$),其他维度的效应量绝对值都达到中高水平,反映了在这些维度上,两类文本间的差异切实存在。不论ESP或EGP,在维度1、2、6上的得分均为负值,且ESP得分更低,说明ESP文本的交互性、叙事性和即席信息组织精密度都不及EGP。在EGP中,交际者间的互动、情感和介入程度更高;ESP文本的静态描写及说明色彩更浓,信息整合度更高,积极的以事件为导向的记叙成分则较少。ESP在维度3和维度5上的得分高于EGP,说明前者的指称明确性和抽象度显著更高,语境依赖性更低。在维度4方面,由于t检验效益量较低,不能说明两类文本的显性劝说性差异特别明显。

再看表4,两类文本在所有句法复杂度指标上的差异都达到显著程度($p < 0.05$),除动词短语复杂度指标VP/T的效应量未达到中等水平外(*Cohen's* $d < 0.20$),其他指标的效应值均处于中高水平,表明两类文本在这些指标上具有切实差异。ESP的句子、T单位和子句显著更长(MLS; MLT; MLC)。但句内从属复杂度显著低于EGP(C/T; CT/T; DC/C; DC/T)。ESP的句间并列复杂度(T/S)也不及EGP。在并列短语(CP/C; CP/T)和名词短语(CN/C; CN/T)方

表4 ESP与EGP句法复杂度独立样本t检验

	ESP		EGP		t值	p值	效应量
	均值	标准差	均值	标准差			
MLS	18.834	3.803	16.737	3.792	-6.766	0.000	0.552
MLT	17.350	3.416	14.975	3.490	-8.422	0.000	0.688
MLC	12.030	2.500	9.579	2.185	-12.782	0.000	1.122
C/T	1.464	0.240	1.581	0.276	5.555	0.000	0.423
CT/T	0.376	0.149	0.418	0.146	3.516	0.000	0.284
DC/C	0.314	0.104	0.338	0.102	2.819	0.005	0.232
DC/T	0.481	0.227	0.557	0.247	3.957	0.000	0.320
CP/C	0.426	0.189	0.250	0.152	-12.501	0.000	1.158
CP/T	0.607	0.249	0.385	0.209	-11.808	0.000	1.062
T/S	1.089	0.099	1.124	0.113	4.019	0.000	0.329
CN/C	1.398	0.419	1.035	0.404	-10.808	0.000	0.882
CN/T	2.018	0.600	1.622	0.645	-7.786	0.000	0.635
VP/T	2.058	0.411	2.134	0.423	2.219	0.027	0.182

面,ESP复杂度更高。这些发现说明ESP的句法复杂度更多体现于短语层,而在句子内部和句子之间,EGP复杂度更高。

表5显示,与句法复杂度指标相关度较高的是维度1、3、4、6,维度2、5相关度较低($r < 0.20$),基本可以忽略。维度1与7个句法复杂度指标的相关达到中等以上水平($r > 0.40$)。与维度3、4、6的相关达到中等以上水平的指标分别为2个、2个和1个。这一结果说明句法复杂度的变化会带来文本互动性、指称明确性、显现劝说性及即席信息组织精密度的相应变化。

维度1与句法单位、名词短语和并列短语复杂度的7个指标呈负相关。句内从属、句间并列和动词短语复杂度与该维度呈正相关。这一发现与Biber等的研究结论部分一致,他们认为在高互动性的日常会话中,句内从属更加普遍,而信息性特点鲜明的学术写作中,复杂名词短语更为常见^[11]。在Biber看来,互动性文本的产出受制于严格的时间限制和情感及人际意义表达的需要,诸如模糊限制语、话语标记语、缩略及替代形式、非限定性代词和句间并列and等语言形式大量存在,致使内容的呈现是碎片化的,语言形式更为简略(reduced)。而信息性文本的内容需要保证具体、准确和集中(integrated)^{[13]129-135},因而有必要使用更多的名词、介词和并列短语。

维度3表征文本的指称明晰性和独立于语境的程度。张德禄认为情景语境和语篇的相关性控制着语篇对意义的选择和表达,显性意义由语言形式表达,隐含意义由非语言形式表达^[15]。文本的语境独立性越高,需要借助语言形式显化的隐含意义就越多。因此提高文本指称明晰性依赖大量的复杂的名词短语和并列短语。分析也发现,维度3之下的正向载荷因子中包含三类定语从句,它们可促进文本的指称明晰性,不过也会增加句法单位的长度,间接造成句法单位与维度3之间的低度正相关。

维度4与动词短语的相关度最高,达到中等水平。这是因为该维度所有7个正向载荷因子中,有4个和动词有关。同等篇幅文本中动词占比越高,名词占比则越低,故名词短语复杂度与该维度具有中低度负相关。此外,该维度与句内从属复杂度也具有低度正相关,出现此结果可能有两点原因:(1)该维度正向载荷因子中包含条件从句;(2)该维度表征文本显性劝说的程度,需要使用大量的劝说性动词(如command; demand; instruct等)来表明未来发生某事的可能和意向性。但这类动词往往都可以携

带宾语从句,所以句内从属复杂度得分越高,该维度得分也越高。

维度6与4个句内从属复杂度指标的相关均达到中低水平。不难发现,该维度下的4个具有最高因子载荷的语言特征中有3个和从句有关。Biber认为维度6的从属性语言特征是实时语言产出时信息阐述的标记。说话者一方面要聚焦于信息阐述,另外也要应对这种条件下语言产出时的种种限制因素^{[3][11]}。进行准确的词汇选择,并将信息“打包”放入含有更多实词和短语的少数语言结构中是不现实的。而附加更多的从句能够缓解产出时的时间压力,实现对信息的有效组织。从功能语言学角度看,这是词汇和语法间“协同效应”的典型例证。

表5 ESP句法复杂度及语体维度的相关系数

	维度1	维度2	维度3	维度4	维度5	维度6
MLS	-0.241**	0.000	0.259**	-0.023	0.189**	0.116*
MLT	-0.319**	-0.035	0.253**	-0.077	0.090	0.131*
MLC	-0.623**	-0.139*	0.323**	-0.249**	-0.043	-0.162**
C/T	0.418**	0.163**	-0.121*	0.227**	0.135*	0.376**
CT/T	0.368**	0.101	-0.125*	0.334**	0.102	0.364**
DC/C	0.440**	0.124*	-0.165**	0.359**	0.089	0.382**
DC/T	0.440**	0.143*	-0.161**	0.317**	0.097	0.404**
CP/C	-0.407**	-0.170**	0.593**	-0.199**	-0.059	-0.303**
CP/T	-0.283**	-0.145*	0.596**	-0.129*	-0.001	-0.197**
T/S	0.121*	0.056	0.023	0.103	0.195**	0.024
CN/C	-0.689**	-0.091	0.342**	-0.408**	-0.047	-0.052
CN/T	-0.483**	-0.024	0.294**	-0.291**	0.050	0.149**
VP/T	0.447**	0.195**	-0.055	0.465**	0.055	0.314**

注:*表示 $p \leq 0.05$; **表示 $p \leq 0.01$

为弄清句法复杂度对维度1、3、4、6的预示程度,我们以相关度较高的句法复杂度指标为自变量,以语体维度为应变量,采用逐步进入法进行多元回归。

选择维度1为应变量,自变量为MLC、DC/C、CP/C和CN/C。CN/C进入模型时能够解释的方差高达47.5%。 $F(1, 298) = 269.736; p = 0.000$ 。增加了DC/C后,模型2所能解释的方差为52%。 $F(2, 297) =$

160.899; $p = 0.000$ 。说明名词短语能够十分显著预示文本的互动性,且句内从属对维度1也具有一定的解释力。

当应变量为维度3,自变量为CP/T和CN/C时,CP/T进入模型,能够解释35.3%的方差, $F(1, 298) = 163.794; p = 0.000$ 。增加CN/C后,模型2可解释36.2%的方差。 $F(2, 297) = 85.850; p = 0.000$ 。可知并列短语是该维度最有效的预示变量。

应变量为维度4,自变量为DC/C、CN/C和VP/T时,VP/T进入模型1,能够解释21.6%的方差, $F(1, 298) = 82.175; p = 0.000$ 。增加CN/C后,模型2可以解释31.3%的方差。 $F(2, 297) = 67.608; p = 0.000$ 。动词及名词短语对该维度具有较强解释力。

当应变量为维度6,自变量为DC/T、CP/C和VP/T时,DC/T进入模型1,可解释16.3%的方差, $F(1, 298) = 58.155; p = 0.000$ 。增加CP/C后,模型2可解释19.3%的方差。 $F(2, 297) = 35.422; p = 0.000$ 。句内从属具有一定的预示作用。

四、结语

本研究发现ESP相较于EGP,在语体和句法复杂度方面都表现出鲜明特点,具体表现为:ESP文本的互动性、叙事性和即席信息精密度显著低于EGP;前者的指称明晰性、抽象性显著高于后者。ESP文本的句法单位、并列短语和名词短语的复杂度高于EGP,但句内从属和句间并列复杂度显著较低。ESP文本中,句法复杂度的变化能够显著影响文本的互动性、指称明晰性、显现劝说性和即席信息精密性,但不会影响它的叙事性和抽象程度。研究说明:不同语域文本因其传达的意义类型有别,在句法复杂度方面亦有差异。特定句法资源的使用是为了适应意义表达的需要,以实现不同的话语功能。本研究存在如下不足:选取的ESP样本数量不足,语料代表性有待商榷。其次,本研究只调查了句法方面的复杂度,对句法多样性未做研究。希望以后的研究能够从上述角度入手,获取对ESP文本更为科学而全面的认识。

参考文献:

- [1] HUTCHINSON T, WATERS A. English for specific purposes: A learning-centered approach[M]. Cambridge: Cambridge University Press, 1987: 21-23.
- [2] 胡壮麟,朱永生,张德禄,等.系统功能语言学概论[M].北京:北京大学出版社,2008: 346-348, 273-274.
- [3] 秦晓晴,文秋芳.中国大学生英语写作能力发展规律与特点研究[M].中国社会科学出版社,2007: 27-53
- [4] YOON H J, POLIO C. The linguistic development of students of English as a second language in two written genres[J]. Tesol Quarterly, 2016 (2):1-27.

据此可知,[ŋa]和[ɲia]恰是方言音留存古音的例证,南方方言音系,诸如闽南语对于古音的保存程度非常高,西昌话中留存这两个音而四川方言音系没有,恰好说明了方言彼此间相互接触和影响。

三、结语

西昌话中“我”和“你”表示领属,用在亲属称谓词前面时,有文白两读(文读:[ŋo]、[li],白读:[ŋa]、

[ɲia])。笔者通过归纳第一人称和第二人称文读、白读的使用情况,总结出两条语法规则:一是,文读使用范围较广,可作主语、宾语、定语和补语,并且其后可跟助动词“的”表示领属;而白读使用范围较窄,仅可作定语,其后不可跟助动词“的”表示领属,本身具有领属意义。二是,文白异读产生的原因是语言音系演变过程中一般规律和特殊规律的交叉,西昌方言音系与外来方言及少数民族语言的长期接触。

参考文献:

- [1] 林寿,王理嘉.语音学教程[M].北京:北京大学出版社,1992.
- [2] 殷科.西昌话探源——西昌话语近源方言音系的比较[D].成都:四川师范大学,2013.
- [3] 刘青.云南山区多民族杂居村落得语言接触[J].昆明学院学报,2012(4).
- [4] 项梦冰.客家话人称代词单数“领格”的语源读后[J].语文研究,2002(1).
- [5] 牟成刚.广南方言三个非三身代词的共时调查分析——koŋ³¹、nia⁴⁴和mei⁴⁴[J].文山师范高等专科学校学报,2009(2).
- [6] 黄尚军.成都话音系[J].四川师范大学学报,2006(1).

(上接第65页)

- [5] YANG W, LU X, WEIGLE S C. Different topics, different discourse: relationships among writing topic, measures of syntactic complexity, and judgments of writing quality[J]. Journal of Second Language Writing, 2015(28):53-67.
- [6] LU X, AI H. Syntactic complexity in college-level English writing: differences among writers with diverse L1 backgrounds [J]. Journal of Second Language Writing, 2015(29):16-27.
- [7] LU X. A Corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development[J]. TESOL Quarterly, 2011(1): 36-62.
- [8] ROBINSON P. Task complexity, task difficulty, and task production: Exploring interactions in a componential framework [J]. Applied Linguistics, 2001, 22(1): 27-57.
- [9] FREAR M W, BITCHENER J. The effects of cognitive task complexity on writing complexity[J]. Journal of Second Language Writing, 2015(30):45-57.
- [10] RYSHINA-PANKOVA M. A meaning-based approach to the study of complexity in L2 writing:The case of grammatical metaphor[J]. Journal of Second Language Writing, 2015(29):51-63.
- [11] BIBER D, GRAY B, Poonpon K. Should we use characteristics of conversation to measure grammatical complexity in L2 writing development?[J]. Tesol Quarterly, 2011, 45(1):5-35.
- [12] ORTEGA L. Syntactic complexity in L2 writing: progress and expansion[J]. Journal of Second Language Writing, 2015,(29): 82-94.
- [13] BIBER D. Variation across speech and writing[M]. Cambridge University Press, 1988:85-91; 129-135; 113.
- [14] BIBER D, CONRAD S. Register, genre, and style[J]. Cambridge University Press, 2009(2):23-224.
- [15] 张德禄.非语言特征的衔接作用[J].解放军外国语学院学报,2000,23(4):21-25.