

英语学习者水平、口语情感表现与 词汇复杂度关系研究

鲍贵 陈长亮

(南京工业大学 语言文化学院, 江苏 南京 211816)

摘要:鲜有研究考察英语学习者口语情感表现如何影响词汇复杂度。本研究利用166名英语专业学生独白数据,探究英语水平、情感词频率和情感词效价对词汇复杂度(含词汇多样性和词汇复杂性)的预测效应。路径分析表明,英语水平对词汇复杂性有直接影响,对词汇多样性则没有,但是它通过情感词频率的中介对之产生间接影响。情感词频率对词汇多样性和词汇复杂性均产生直接影响,对前者的作用更大,而情感词效价对这两个测量均没有影响。整体上,英语水平和情感词频率对词汇多样性的预测力比对词汇复杂性的预测力更大。这些结果表明,英语口语教学中应重视情感词的使用和情感层次的区分,发挥语言水平对词汇复杂度的积极作用。

关键词:英语水平;情感表现;情感词频率;情感词效价;词汇复杂度

中图分类号:H319 文献标志码:A 文章编号:1674-6414(2026)02-0178-14

0 引言

情感表现是人类的一个基本特征,也是跨文化交际的重要技能,影响着人际关系构建(Dewaele, 2005)。成功的跨文化交际不仅要求外语学习者有扎实的语言基础,而且还要求他们具备恰当的情感表达能力。如何在交际中准确表达自己的想法并把措辞蕴含的情感传递给交际方尤为重要。前期情感研究主要包括两个方向。一是情感因素与二语习得之间的关系,二是探究二语学习者的情感词记忆及其在二语产出中的情感表现。前者受到传统心理学与积极心理学的影响,表现为对情感个体的考察(董良峰等, 2025; 范开芳等, 2023),而后者则与社会文化视角相关,关注个体因素与社会文化因素之间的情感互动(秦

收稿日期:2025-11-13

基金项目:教育部人文社会科学研究基金项目“基于文本挖掘技术的英语学习者任务型口语产出机制研究”(24YJA740002)和江苏省研究

生科研与实践创新计划项目“英语学习者与母语者口语产出细颗粒句法复杂度对比研究”(KYCX25_1691)的阶段性研究成果

作者简介:鲍贵,男,南京工业大学语言文化学院教授,博士,硕士研究生导师,主要从事应用语言学、自然语言处理研究。

陈长亮,男,南京工业大学语言文化学院硕士研究生,主要从事二语习得、自然语言处理研究。

引用格式:鲍贵,陈长亮. 英语学习者水平、口语情感表现与词汇复杂度关系研究[J]. 外国语文, 2026(2): 178-191.

丽莉 等,2022b)。由情感互动衍生的相关情感理论主张将认知与情感相结合,如社会文化理论中的“主观情感经历”概念(Perezhivanie)(秦丽莉 等,2022a)。

目前,国内在二语习得领域基于“主观情感经历”概念开展的实证研究难得一见(秦丽莉 等,2019;秦丽莉 等,2022a),认知与情感的整合研究亟需更多实证支持。近年来,伴随着二语习得领域中的“社会文化转向”(Johnson, 2006)与“情感转向”(Pavlenko, 2013; Prior, 2019),出现了一些探究二语学习者在任务条件下情感表现的研究,如考察二语学习者在作文中的情感表现以及语言水平因素对情感表现的作用(Abdi Tabari et al., 2024; Wang, 2020; 毛文伟, 2022),但是尚无研究考察学习者在口语产出中的情感表现如何影响二语表现的不同维度(如语言准确性、流利度和复杂度)。本文运用情感分析技术,利用路径分析率先考察英语专业学习者的英语水平和情感表现对词汇复杂度的直接与间接效应,希冀提升情感表现在二语教学和发展中的积极作用,促进社会文化视角下的情感理论在二语习得领域的应用,拓展“主观情感经历”概念在我国社会文化语境中的研究,加深对认知与情感表现之间关系的认识。

1 文献综述

在二语习得领域,情感表现研究主要集中于情感因素与二语水平之间的关系(Li et al., 2020; Teimouri, 2018; 李成陈 等, 2022)、情感词效应和语言产出中的情感表现。前者关注学习者的外在情感状态,后者则与情感词(emotion words)的理解、记忆与应用有关。情感词记忆研究是一个活跃的研究领域,研究的主要发现是学习对情感词有记忆效应(Arriagada-Möding et al., 2022; Ferré et al., 2013; Naranowicz et al., 2023)。例如,马尔钦·纳拉诺维茨等(Naranowicz et al., 2023)考察了情绪和性别对非平衡双语者(波兰语和英语)情感词加工反应时和准确性的影响。该研究操纵电影片段使样本量相同的波兰男、女大学生处于积极和消极情绪中,对240个情感词(每种语言中积极、消极和中性抽象形容词各40个)做出情感判断(积极、消极、中性)。该研究发现,参与者在积极情绪中比在消极情绪中对二语积极词反应更快,但是只在积极情绪中,参与者对母语(波兰语)中性词的反应比二语(英语)中性词的反应更快;女生在积极情绪中对情感词的反应比在消极情绪中更快,而男生则没有这种情绪效应;无论性别和使用的语言是什么,参与者对积极词和消极词的反应都比对中性词的反应更快、更准确。以上研究表明二语学习对情感词的记忆与运用效果会受到文化背景和情感词效价的影响。

既然情感词效应在二语编码任务中普遍存在,由此引发的一个议题是:情感词效应是否会迁移到学习者产出性任务中?一系列研究考察学习者写作中的情感表现(Abdi Tabari et al., 2021; Abdi Tabari et al., 2022; Abdi Tabari et al., 2024; El-Dakhs et al., 2023; Uzun et

al., 2022; Wang, 2020; 毛文伟, 2022), 发现二语学习者的认知与情感在二语写作任务中表现出关联性。例如, 王一舟(Wang, 2020)基于中国大学生作文语料库发现学习者倾向带着积极的情感写作, 写作提示的积极情感值与作文情感分(情感词效价测量)呈中度正相关($r = 0.35, p < 0.001$)。这项研究还发现写作提示的情感和作文中的情感表现影响学习者句法复杂度指标的变化。马哈茂德·阿卜迪·塔巴里等(Abdi Tabari et al., 2021)也认为更复杂的语言使用与作文情感趋于中性有关。相较于情感表现与句法复杂度关系的研究, 对词汇复杂度的探究进一步揭示了二语学习者在任务完成中的情感词效应。马哈茂德·阿卜迪·塔巴里等(Abdi Tabari et al., 2022)通过比较英语中高水平的留学生在两种作文提示(诱发积极和消极情感的新冠疫情提示和诱发中性情感的产品推销广告提示)下的议论文, 发现有疫情提示的作文比无疫情提示的作文导致学生文本中的词汇多样性(lexical diversity)更丰富, 但是无疫情提示下的作文在词汇习得年龄(lexical AoA, 词汇复杂性的一种测量)方面比在有疫情提示下的词汇习得年龄更大(即更复杂的词汇使用)。当积极和消极情感分作为协变量被控制时, 作文情感提示对词汇习得年龄不再有显著影响, 对词汇多样性的影响只有边际显著性($p = 0.079 < 0.1$), 显示情感表现在作文提示对词汇复杂度方面所发挥的中介作用。与马哈茂德·阿卜迪·塔巴里等(Tabari et al., 2022)相比, 马哈茂德·阿卜迪·塔巴里等(Tabari et al., 2024)扩大了对词汇复杂度的考察范围, 增加了对平均对数词频与词汇熟悉度得分的测量。该研究发现, 在简单写作任务中, 负面情感值与词汇多样性呈正相关关系, 但与词汇习得年龄呈负相关关系, 作文情感分与词汇习得年龄呈正相关关系。在复杂写作任务中, 积极情感值和作文情感分与词汇多样性均呈负相关关系。但是, 马哈茂德·阿卜迪·塔巴里等(Tabari et al., 2021)却得出不同结论: 积极情感值与消极情感值都与词汇多样性呈负相关关系。这或许表明任务复杂度引起的认知差异会影响二语学习者对情感词的使用(Tabari et al., 2024)。产出模态的差异是否也会影响情感词的使用呢? 二语学习者在口语任务中的情感表现又如何影响词汇复杂度呢?

在情感词效应和情感表现方面, 同书面语研究相比, 口语研究相对稀缺(Dewaele et al., 2002; Graham et al., 2001; Pavlenko et al., 2007)。让·马克·德韦勒等(Dewaele et al., 2002)通过对话和电影短片复述, 调查语言水平和性别等因素对二语口语情感词使用的影响。该研究发现, 二语情感词的使用与语言水平、语言材料类型、性格外向性有关, 有些情况下也与说话者的性别有关。阿内塔·帕夫连科等(Pavlenko et al., 2007)克服让·马克·德韦勒等(Dewaele et al., 2002)实验设计存在的局限, 如把参与电影短片复述的大学生人数从34人扩大到129人, 同时控制性别和社会教育背景等因素, 对比分析了俄语母语者($n = 49$)、美语母语者($n = 50$)和美国高水平俄语学习者($n = 30$)口语复述中的情感词使用情况。该研究发现, 美语母语者偏好使用形容词模式描述情感, 而俄语母语者倾

向使用动词模式;高水平俄语学习者的情感描述方式从美语中的形容词模式转向俄语中的动词模式;在情感词汇多样性方面,较之于美语母语者,俄语母语者文本类符—形符比更高;与俄语母语者相比,俄语学习者使用更少的情感词,但是两者在情感词汇多样性(类符—形符比)方面大致相同,都比美语母语者使用更丰富的情感词。这些研究揭示了情感词使用的跨语言和跨文化影响。

概而言之,前期研究对二语学习者语言产出情感表现的深度挖掘不足,对口语产出文本的挖掘更是如此。在词汇复杂度测量方面,前期研究主要考量的是词汇多样性,对词汇复杂性的关注度不足,尚无研究回答二语学习者在口语任务中的情感表现如何影响词汇复杂度。词汇多样性与词汇复杂性测量易受文本长度的影响,但是前期研究缺乏有效控制。尽管让·马克·德韦勒等(Dewaele et al., 2002)考察了英语水平对情感词使用的影响,但是用学习者文本中的形态词汇准确率作为英语水平的替代测量存在局限,把替代测量转化为三个类别水平还会导致统计效力的降低。

作为对二语习得研究“情感转向”的回应(Pavlenko, 2013; Prior, 2019),本研究以英语水平为来源变量,把情感表现(包括情感词频率和情感词效价)作为中介变量,考察它们对英语专业学生口语产出词汇复杂度(包括词汇多样性和词汇复杂性)的直接和间接影响。

2 研究方法

2.1 研究问题

本研究回答以下两个问题:

(1)英语水平、情感词频率和情感词效价对学习者口语词汇多样性产生怎样的预测作用?

(2)英语水平、情感词频率和情感词效价对学习者口语词汇复杂性产生怎样的预测作用?

本研究把英语水平作为自变量,把情感表现作为中介变量,词汇复杂度视作因变量。情感表现包括情感词频率和情感词效价,体现情感表现的不同维度。本研究把英语水平作为自变量或预测变量,主要因为语言水平在二语发展中的关键作用(Dewaele et al., 2002; El-Dakhs et al., 2023; Graham et al., 2001; 鲍贵, 2010)。二语水平是二语学习者整体上的发展水平,指学习者对二语的了解程度,即学习者在不同交际情境中如何能够应用二语,但是对二语水平的操作定义(如课程水平或标准化语言测试)通常因研究而异,考察语言特征与二语水平之间的关系或对二语水平区分度的研究也是如此(Bulté et al., 2020)。例如,鲍贵(2010)把英语水平操作定义为标准化的英语专四成绩分布,研究发现在口语词汇多样性测量方面,低、中等英语水平组之间没有显著差异,但是高水平组则显著好于低、中

等水平组。把情感表现作为中介变量的理由在于罗莎·M·希梅内斯·卡塔兰等(Catalán et al., 2017)、马哈茂德·阿卜迪·塔巴里等(Tabari et al., 2022)和马哈茂德·阿卜迪·塔巴里等(Tabari et al., 2024)发现文本情感表现能显著预测词汇复杂度。全国英语专四成绩所代表的英语水平能够有效影响二语学习者口语词汇复杂度表现(Bao, 2024; 鲍贵, 2010),且英语水平对情感词的使用频率也存在显著影响(Dewaele et al., 2002)。上述研究结果表明英语水平可能会通过情感表现这一中介变量对词汇复杂度产生影响。词汇复杂度包括词汇多样性和词汇复杂性两个测量。在中介变量中,本研究认为情感词频率影响情感词效价,因为前者是后者的来源。本研究采用路径分析回答研究问题。路径分析相对于多元回归分析的一个优势是它能够检验一个预测变量是否会通过另一个预测变量的中介作用间接影响因变量。在本研究中,路径分析用于探究英语水平、情感词频率和情感词效价是否会对词汇多样性和词汇复杂性分别产生直接和间接影响。

2.2 参与者

本研究的参与者为来自华东地区四所高校的166名英语专业大二学生,其中男生12人,女生154人,年龄在18~21岁之间,正式学习英语时间在八年左右,且都没有去英语国家旅游或学习的经历。这批学生在完成本研究任务之前参加了全国英专四级笔试,笔试成绩在约半年后由考试委员会向考生所在学校发布。全国英语专业四级考试社会认可度高,能够有效测量学习者的英语水平,且信度高,因此本研究利用英语专业四级考试成绩衡量学习者英语水平。这批学生的专业四级笔试成绩分布在51~91分之间。

2.3 任务设计与实施

本研究的任务设计采用全国英语专四口语测试中的独白任务。所有参与者于参加全国英语专四笔试半个月之后的同一时间自愿在本校的语言实验室完成独白任务。在口语测试前一周,研究者与各高校四位口语教师确认了参与学生在此次独白话题的熟悉程度。独白任务要求参与者结合自身讲述一次成功经历,难度适中。测试过程中,研究者要求参与者根据录音设备上的时间提示,就主题准备两分钟。准备期间允许参与者做笔记,然后即席讲话三分钟。所有参与者的独白录音保存为MP3或WAV格式的音频文件。

2.4 数据转录

本研究中,参与者独白的最小时长为58.49秒,最大时长为182.98秒,平均时长为160.71秒($SD = 22.58$)。四名经过培训的硕士研究生对166份音频文件进行了转录,并对转录结果进行交叉核对。在转录过程中,对错误的开头、自我修正、填充语和机械重复等做出标注,转录结果保存为Word文档。研究者参照音频文件对提交的转录文档再次核对,剔除错误的开头和自我修正等标注,并把Word文档转化为文本文件,以备后续统计分析。参与者独白文本长度分布在106~389个词之间($M = 237.34, SD = 56.82$)。

2.5 变量测量

本研究设计的预测变量和中介变量分别是英语水平和情感表现(包括情感词频率和情感词效价),因变量是词汇复杂度(包括词汇多样性和词汇复杂性)。英语水平操作定义为参与者的专四笔试成绩。情感表现的两个测量利用情感分析。

情感分析调用 Python 库 vader Sentiment, 其中的 vader 或 VADER 指 Valence Aware Dictionary for Sentiment Reasoning(情感推理的效价感知词典)。VADER 把条目特征与多条概括性规则结合在一起,利用黄金标准词库(包括 7000 多个条目)对条目特征的情感效价或强度进行评分,评分等级介于-4(极端消极)和 4(极端积极)之间,评分信度很好(Hutto et al., 2014)。本研究对学生文本情感词频率的测量只考虑类符(types),不考虑可能重复使用的形符(tokens),如“love”(情感分为 3.2)和“loved”(情感分为 2.9)是不同类符。对每个学生文本中出现的消极词和积极词类符数的计算调用 Python 库 NLTK 和 vaderSentiment。根据消极词的情感分为负值、积极词的情感分为正值,从文本中提取这两种类符并计数。鉴于每个学生文本长度不一,本研究采用情感词频率比较不同文本中的情感词使用,操作定义为每个文本中消极和积极情感类符累计数与该文本形符总数的比率再乘以 100,表示每 100 个词中情感词类符出现的频率。

本研究对情感词效价的测量利用文本综合情感分。综合情感分的计算以文本包含的各个句子为单位,把所有句子的综合情感分的平均值作为文本情感分,以消除不同文本使用的句子数对综合情感分的影响。综合情感分的计算较为复杂,这里举个例子。在话语“I raised my hands bravely and I got good grades”中,“bravely”和“good”这两个词的积极情感分别为 2.3 和 1.9,其总和为 4.2。其他词都是中性情感词,情感分均为 0。因此,在这个句子中,中性情感分计数为 8,积极情感分之总和为 $2.3 + 1 + 1.9 + 1 = 6.2$ (1 是常数),情感分总和为 $8 + 6.2 = 14.2$ 。归一化的积极情感分为积极情感分之总和与情感分总和的比率,即 $6.2/14.2 \approx 0.44$,归一化的中性情感分为 $8/14.2 \approx 0.56$,归一化的消极情感分为 0。综合情感分为积极情感分总和除以积极情感分总和的平方与归一化参数(经验值为 15)之和平方根得到的比率,即 $4.2/\sqrt{(4.2^2 + 15)} \approx 0.74$ 。本研究对学生文本中综合情感分的统计调用 Python 库 vaderSentiment 中的类 SentimentIntensityAnalyzer,对每个文本中的每个句子统计综合情感分,再把每个句子的综合情感分累计在一起,最后计算平均值。

在因变量中,词汇多样性利用文本词汇多样性测量(MTLD)。该测量逐词计算类符—形符比,把类符—形符比小于 0.72(因子阈限值)的连续词串作为一个因子。对于不构成完整因子的文末连续词串,部分因子值为 1 和类符—形符比差异与 1 和因子阈限值差异的比率(McCarthy et al., 2010; 鲍贵等, 2023)。这一方法有效控制了文本长度对传统的类符—形符比测量的影响(Kyle et al., 2021)。词汇复杂性测量的操作定义为每 100 个文本

词中包含的低频词类符数,参照移动平均类符—形符比(MATTR)(Covington et al.,2010),使用移动平均词汇复杂性测量(moving average lexical sophistication)以控制文本长度的影响。该测量从文本开头以100个形符(从第1个形符到第100个形符,不含标点符号)为一个字符串计算低频类符数,然后顺延一个词,即从第2个词到101个词构成的字符串计算低频类符数,依次类推,直至计算在最后100个词构成的字符串中的低频类符数。移动平均词汇复杂性值为计算得到的所有低频类符数平均值。高频词常用1000词词表采用伊恩·斯蒂芬·保罗·内申(Nation,2016)开发的BNC/COCA词表中英美口笔语最常用1000词条。常用1000词词表以外的词视作低频词。该测量调用Stanza库自编Python代码计算不包含在常用1000词词表中的类符数。

2.6 统计分析

本研究对多个观测变量之间的关系采用路径分析。Shapiro-Wilk W 正态性检验发现,在英语水平、情感词频率、情感词效价以及词汇多样性和词汇复杂性测量数据中,只有词汇多样性数据违反正态分布假设($W = 0.98, p = 0.001$)。对该测量数据开展平方根转化,转化后数据满足正态分布假设($W = 0.96, p = 0.066 > 0.05$)。本研究调用R 4.4.3开展路径分析,调用的数据包是lavaan(Rosseel,2012),绘图数据包是semPlot(Epskamp,2022)。

3 研究结果

本节分三小节报告统计分析结果。第一小节报告各个测量描述性统计分析结果,第二、第三小节分别报告英语水平、情感词频率和情感词效价对词汇多样性和词汇复杂性影响的路径分析结果。

3.1 英语水平、情感词频率、情感词效价、词汇多样性和复杂性之间的相关性

表1报告各个测量变量数据的相关系数矩阵,包括平均数(M)和标准差(SD)。

表 1 各个测量相关矩阵

		1	2	3	4	5
1	英语水平	1				
2	情感词频率	0.15 *	1			
3	情感词效价	-0.03	0.30 **	1		
4	词汇多样性	0.12	0.44 **	0.01	1	
5	词汇复杂性	0.20 *	0.23 **	0.05	0.28 **	1
	M	74.83	7.14	0.23	6.90	9.74
	SD	8.06	1.92	0.13	0.69	2.95

注:* 和 ** 分别表示在 0.05 和 0.01 水平上呈显著相关。

表 1 显示,各个变量之间关系的强度处于低、中等水平。英语水平与情感词频率、词汇复杂性均呈显著正相关(前者: $r = 0.15, p < 0.05$;后者: $r = 0.20, p < 0.05$),与其他两个变量均没有显著相关($p > 0.05$)。词汇多样性和词汇复杂性呈中度正相关($r = 0.28, p < 0.01$)。这两个测量与情感词频率均呈显著正相关,只是情感词频率与词汇多样性的相关度要高于它与词汇复杂性的相关度($r = 0.44$ 相对于 $r = 0.23$),但是都与情感词效价没有显著相关($p > 0.05$)。情感词频率与情感词效价的相关度只处于中等水平($r = 0.30, p < 0.01$),说明它们可以作为情感表现的两个不同测量指标。

3.2 英语水平、情感词频率和情感词效价对词汇多样性影响的路径分析

对词汇多样性和词汇复杂性两个测量分别开展路径分析,分析模型为刚好识别模型。与词汇多样性相关的标准化路径分析结果如图 1 所示。图中单向直线箭头表示直接影响,其中的长线箭头显示路径系数,短线箭头表示残差方差,星号 * 和 * * * 分别表示在 0.05 和 0.001 水平上有显著预测力。

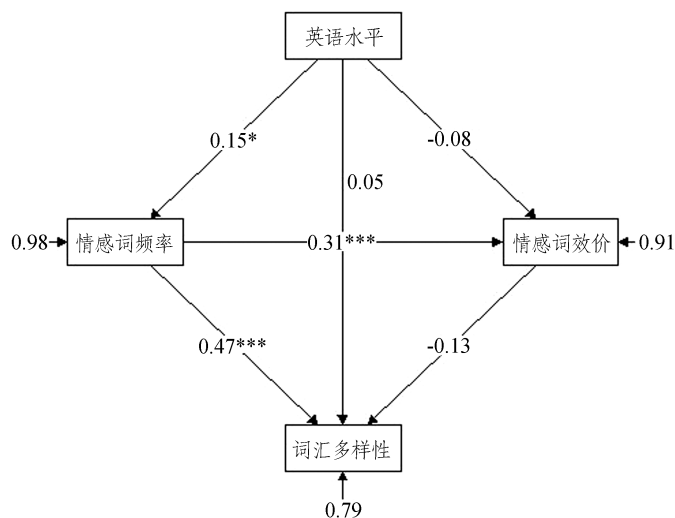


图 1 英语水平、情感词频率和情感词效价对词汇多样性影响的路径

图 1 显示,英语水平对词汇多样性没有产生直接影响,且效应小($z = 0.68, p = 0.498 > 0.05, \beta = 0.05$)。在三个间接路径中,英语水平经由情感词频率中介对词汇多样性产生显著的间接影响($z = 1.92, p = 0.05, \beta = 0.15 \times 0.47 \approx 0.07$),但是它经由情感词效价的中介对词汇多样性没有间接影响($z = 0.88, p = 0.378, \beta = (-0.08) \times (-0.13) \approx 0.01$),也没有通过情感词频率和情感词效价的中介间接影响词汇多样性($z = -1.27, p = 0.203 > 0.05, \beta = 0.15 \times 0.31 \times (-0.13) \approx -0.01$)。情感词效价对词汇多样性没有直接影响($z = -1.79, p = 0.073 > 0.05, \beta = -0.13$)。相比之下,情感词频率对词汇多样性产生直接影响,且效应很大($z = 6.44, p < 0.001, \beta = 0.47$)。虽然能够显著预测情感词效价($z = 4.17, p < 0.001, \beta = 0.31$),但是情感词频率没有对词汇多样性产生间接影响($z = -1.65, p = 0.100 > 0.05, \beta =$

$0.31 \times (-0.13) \approx -0.04$)。概而言之,以上路径模型能够预测约 20% 的词汇多样性变异,主要预测变量是情感词频率和英语水平。

3.3 英语水平、情感词频率和情感词效价对词汇复杂性影响的路径分析

英语水平、情感词频率和情感词效价对词汇复杂性影响的标准路径分析结果如图 2 所示。图中单向直线箭头表示直接影响,其中的长线箭头显示路径系数,短线箭头表示残差方差,星号 *、* * 和 * * * 分别表示在 0.05、0.01 和 0.001 水平上有显著预测力。

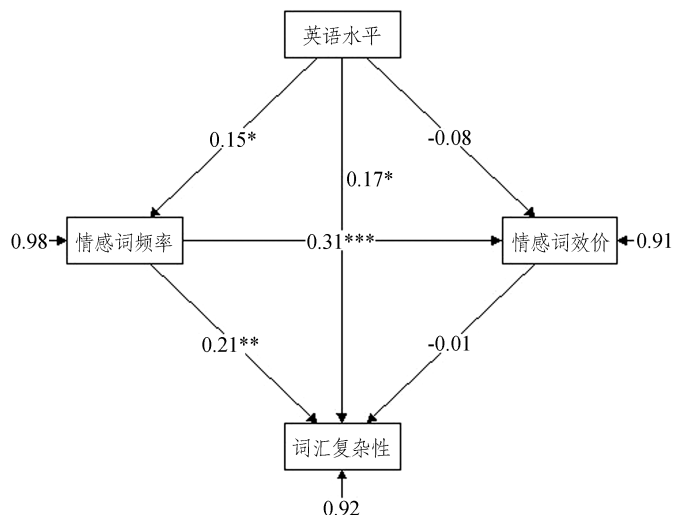


图 2 英语水平、情感词频率和情感词效价对词汇复杂性影响的路径

图 2 显示,英语水平对词汇复杂性产生显著性直接影响,效应达到中等偏上 ($z = 2.18, p = 0.029 < 0.05, \beta = 0.17$)。虽然英语水平对情感词频率产生直接影响 ($z = 2.01, p = 0.045 < 0.05, \beta = 0.15$),但是它没有借此间接影响词汇复杂性 ($z = 1.60, p = 0.109, \beta = 0.15 \times 0.21 \approx 0.03$),也没有经由情感词效价的中介间接影响词汇复杂性 ($z = 0.17, p = 0.863 > 0.05, \beta = (-0.08) \times (-0.01) \approx 0$),同样也没有通过情感词频率和情感词效价间接影响词汇复杂性 ($z = -0.17, p = 0.862 > 0.05, \beta = 0.15 \times 0.31 \times (-0.01) \approx 0$)。情感词效价对词汇复杂性没有直接影响 ($z = -0.18, p = 0.861 > 0.05, \beta = -0.01$)。情感词频率直接影响词汇复杂性 ($z = 2.67, p = 0.008 < 0.01, \beta = 0.21$),但是它没有通过情感词效价的中介对词汇复杂性产生间接影响 ($z = -0.18, p = 0.861 > 0.05, \beta = 0.31 \times (-0.01) \approx 0$)。概而言之,以上路径模型能够预测约 8% 的词汇复杂性变异,主要预测变量是情感词频率和英语水平。

4 讨论

本研究得到两个主要发现。其一,在直接效应方面,英语水平对词汇多样性没有直接影响,但是对词汇复杂性有直接影响;情感词频率对词汇多样性和词汇复杂性均有直接影

响,且对词汇多样性的作用更大,情感词效价则对这两个测量均没有直接效应。其二,在间接效应方面,只有英语水平经由情感词频率的中介间接影响词汇多样性。

以上结果表明,英语水平促进情感词的有效使用,对词汇复杂度有直接或间接作用,对情感词效价没有预测力,体现出语言水平和情感表现之间的复杂关系。这些发现得到前期研究的部分支持(Dewaele et al., 2002; Abdi Tabari et al., 2021; Abdi Tabari et al., 2022; El-Dakhs et al., 2023)。让·马克·德韦勒等(Dewaele et al., 2002)在一项实验中发现不同语言水平组对以荷兰语为母语的法语学习者口语文本情感词形符频率有显著预测力($\beta = 0.33, p = 0.032 < 0.05$),同本研究发现的高水平学习者比低水平学习者使用更多情感词的结果趋于一致。与本研究不同的是,马哈茂德·阿卜迪·塔巴里等(Abdi Tabari et al., 2021)发现积极情感值、消极情感值均与词类符—形符比(词汇多样性测量)呈显著负相关(前者: $r_s = -0.24, p = 0.006$;后者: $r_s = -0.25, p = 0.004$)。造成差异的主要原因可能是马哈茂德·阿卜迪·塔巴里等(Abdi Tabari et al., 2021)研究得到的文本积极情感值和消极情感值与本研究把每个文本中消极和积极情感词类符累计数得到的频率测量不同。除此之外,在计算情感分时,马哈茂德·阿卜迪·塔巴里等(Abdi Tabari et al., 2022)发现在无疫情情感提示和有疫情情感提示下词汇多样性和情感分均没有显著相关性(前者: $r = 0.002, p = 0.991 > 0.05$;后者: $r = -0.19, p = 0.180 > 0.05$),词频范围(一种词汇复杂性测量,与本研究使用的测量不同)也与情感分没有显著关联(前者: $r = -0.06, p = 0.667 > 0.05$;后者: $r = 0.18, p = 0.205 > 0.05$)。在马哈茂德·阿卜迪·塔巴里等(Abdi Tabari et al., 2022)研究中,情感分的计算等同于文本积极情感分与消极情感分的差值,而本研究使用的综合情感分测量更复杂。尽管本研究使用的情感分析软件 VADER 与马哈茂德·阿卜迪·塔巴里等(Abdi Tabari et al., 2022)使用的软件 SentimentAnalysis 采用的评分方法不同,情感分计算方法也不一致,但是本研究同样发现情感词效价与词汇多样性、词汇复杂性均没有显著相关性。这或许表明,学习者在情感词使用的范围和情感词的情感度方面有趋同性。迪娜·阿卜杜勒·萨拉姆·埃尔—达赫斯等(El-Dakhs et al., 2023)以 64 名以阿拉伯语为母语的英语学习者为被试,让其用母语和二语书面描述在快乐和悲伤情境下的自身经历,研究发现文本情感词频率与二语水平没有显著相关性,与本研究结果不同。主要原因可能是迪娜·阿卜杜勒·萨拉姆·埃尔—达赫斯等(El-Dakhs et al., 2023)在计算时似乎只考虑了情感词形符,而本研究则考察情感词类符。然而,同本研究一样,迪娜·阿卜杜勒·萨拉姆·埃尔—达赫斯等(El-Dakhs et al., 2023)也发现二语水平不能显著预测情感词效价。迪娜·阿卜杜勒·萨拉姆·埃尔—达赫斯等(El-Dakhs et al., 2023)的以下发现或许能够提供佐证:二语(外语)产出比母语产出更具有情感极性。即,相对于母语作文,二语作文总是降低情感词效价,在悲伤情境下比在快乐情境下差异更

明显。迪娜·阿卜杜勒·萨拉姆·埃尔-达赫斯等(El-Dakhs et al., 2023)同本研究一样,参与者都是英语作为外语的学习者,除课堂口语教学外,缺乏必要的二语社交情感经历,不能有效表达复杂的情感层次,从而导致极性的情感表达方式。因此,外语教师在口语教学过程中应有意识地引导学习者利用情感词汇表达自己的情感,同时也可开展以情感事件为主题的词汇教学与情境练习,从而提升学习者对情感词的感知力,增强他们恰当使用情感词的意识,促进语言能力的全面发展。

本研究的发现还可以从其他前期研究和定性分析中得到部分佐证。C. 雷·格雷厄姆等(Graham et al., 2001)考察了文化背景对情感的感知能力,54名母语为日语、38名母语为西班牙语的英语学习者和85名英语母语者被要求识别独白录音中表现的八种情绪(如气愤和喜悦),研究发现学习者语言水平与情感判断没有关系。在写作研究中,毛文伟(2022)也发现,在以积极情感为主线和以消极情感为主线的两类命题写作任务中,不同水平日语学习者产出的积极情感句占比和消极情感句占比都没有显著差异,学习者倾向于使用中性表达句,与本研究得出的结果(情感词效价均值仅为0.23)趋同。造成语言水平对情感分或情感词效价无预测力的一个主要原因或许在于,本研究中不同水平学习者处于同一个年级水平,口语训练主要限于课堂教学,课外练习机会少,因而口语词汇使用差别可能不是很大,在口语测试时间压力条件下更有可能偏向使用高频情感词,这些情感词在情感强度方面具有同质性,从而降低了不同强度情感词使用差异,使综合情感分趋同。下面使用一个案例来说明。一名高水平女生(专四成绩为80分)和一名低水平女生(专四成绩为69分)的口语文本均包括19个情感词,文本长度大致相同(分别为268和272个词),情感词效价相同(均为0.32)。两名学生开头的句子(词数大致相同)如下:

高水平学生文本:

I'm a very shy girl and I don't dare to speak in the public. Once upon one time, I was forced to take part in a speech competition. I don't want to feel embarrassed on the stage and say nothing. So I had to try my best to practice it. At first, I practice before my mirror and record it and watch the video every night.

低水平学生文本:

What left me deep impression was a special experience in my high school. As a competitor in a public speaking competition, I needed to stand on the stage in front of a large audience in the hall to deliver my speech clearly. In the following days I was nervous and anxious. I was so afraid that I might not able to say a word because of nervous.

以上文本中,高水平学生使用五个情感词,包括两个积极情感词(“want”“best”)和三个消极情感词(“shy”“forced”“embarrassed”),而低水平学生也使用五个情感词,包括3

个积极情感词(“impression” “special” “clearly”)和两个消极情感词(“nervous” “anxious”)。两名学生的情感词使用量相同,从而导致两个文本的情感词效价均接近0值,即高水平学生文本综合情感分为-0.09,低水平学生文本综合情感分为0.05。高水平学生文本情感词中只使用了一个常用1000词以外的情感词(“embarrassed”),而低水平学生文本情感词中则使用了三个常用1000词以外的词(“impression” “nervous” “anxious”),说明学习者对低频情感词使用具有独立于语言水平的波动性,也随之降低了英语水平对情感词效价和词汇复杂性的区分度。

5 结语

前期口语任务研究聚焦于任务复杂度,鲜见从社会文化角度对学习者的情感表现进行考察。本研究则关注言语表现的情感维度,包括情感词频率和情感效价,把口语产出的语言维度与社会情感维度结合在一起,利用中国语境下的166名英语专业学生的独白文本,综合分析英语水平、情感词频率和情感词效价对口语产出词汇多样性和词汇复杂性的影响。研究发现,英语水平直接影响词汇复杂性,能够通过情感词频率的中介间接影响词汇多样性。另外,情感词频率直接影响词汇多样性和词汇复杂性,情感词效价则对它们没有直接影响。这些发现凸显情感表现的文本内容对词汇复杂度的积极作用,增加情感维度丰富了词汇复杂度影响因素的研究,也为任务型研究的深入开展奠定了基础。

以上研究结果对外语教学具有参考价值。在大学英语口语教学中应重视情感表现,提高学生的跨文化情感交流能力,特别要增加他们对情感词情感梯度的意识。建议教学中提供英语母语者不同语言交际场景中的情感表现方法和措辞方式,使学生切身体会到情感交流的跨文化差异和不同情感词使用的恰当性。教师可以利用多媒体资源,展示英语母语者的真实情感表达。这些资源不仅可以丰富课堂教学内容,还可以激发学生的学习兴趣 and 动机,使他们更好地体会和模仿英语中的情感表达。在口语教学中,教师还应注重情感表现的多样化,如通过模拟不同的现实情景(如模拟购物和餐厅点餐),让学生在具体的情境中练习使用情感词汇,提升情感词表达的丰富性。这不仅有助于他们理解和掌握情感词的使用特点和多样性,还可以提高他们的口语词汇复杂度。

本研究仅限于探讨学习者口语中熟悉的独白任务,在其他熟悉程度不同的口语任务(如复述和描述任务)中学习者的情感表现如何影响词汇复杂度,仍需后续研究。另外,本研究只从横向考察了不同水平二语学习者的口语情感表现,未来研究可以通过纵向设计,考察不同年级学生的口语情感表现,从而更清晰地揭示二语学习者词汇复杂度的产出机制。

参考文献:

- Abdi Tabari, M. , G. Bui & Y. Wang. 2021. The Effects of Topic Familiarity on Emotionality and Linguistic Complexity in EAP Writing[J]. *Language Teaching Research* (4): 1616-1634.
- Abdi Tabari, M. & Y. Wang. 2022. The Effects of Prompt Types on L2 Learners' Textual Emotionality and Lexical Complexity[J]. *Journal of Second Language Studies* (1): 34-57.
- Abdi Tabari, M. , X. Lu & Y. Wang. 2024. Mapping the Associations Among Task Complexity, Emotionality, and Linguistic Complexity in L2 Writing[J]. *Language Awareness* (2): 345-372.
- Arriagada-MöDinger, F. & R. A. Ferreira. 2022. The Effect of Emotional Valence on Auditory Word Recognition Memory in English as a Foreign Language[J]. *Journal of Psycholinguistic Research* (2): 309-322.
- Bao, G. 2024. Differential Effects of EFL Proficiency and Mode on Noun Phrase Use[J]. *System* (2): 1-13.
- Bulté, B. & H. Roothoof. 2020. Investigating the Interrelationship Between Rated L2 Proficiency and Linguistic Complexity in L2 Speech[J]. *System* (4): 1-16.
- Catalán, R. M. J. & J. M. Dewaele. 2017. Lexical Availability of Young Spanish EFL Learners: Emotion Words Versus Non-Emotion Words[J]. *Language, Culture and Curriculum* (3): 283-299.
- Covington, M. A. & J. D. Mcfall. 2010. Cutting the Gordian Knot: The Moving-Average Type-Token Ratio (MATTR)[J]. *Journal of Quantitative Linguistics* (2): 94-100.
- Dewaele, J. M. & A. Pavlenko. 2002. Emotion Vocabulary in Interlanguage[J]. *Language Learning* (2): 263-322.
- Dewaele, J. M. 2005. Investigating the Psychological and Emotional Dimensions in Instructed Language Learning: Obstacles and Possibilities[J]. *The Modern Language Journal* (3): 367-380.
- El-Dakhs, D. A. S. , Ahmed, M. M. , Altarriba, J. et al. 2023. Differential Emotional Expression in Autobiographical Narratives: the Case of Arabic-English Bilinguals[J]. *International Journal of Bilingualism* (5): 962-979.
- Epskamp, S. 2022. Semplot: Path Diagrams and Visual Analysis of Various SEM Packages' Output[M]. <https://CRAN.R-project.org/package=Semplot>.
- Ferré, P. , Sanchez-Casas, R. & I. Fraga. 2013. Memory for Emotional Words in the First and the Second Language: Effects of the Encoding Task[J]. *Bilingualism: Language and Cognition* (3): 495-507.
- Graham, C. R. , Hamblin A. W. & S. Feldstein. 2001. Recognition of Emotion in English Voices by Speakers of Japanese, Spanish and English[J]. *International Review of Applied Linguistics in Language Teaching* (1): 19-37.
- Hutto, C. & E. Gilbert. 2014. VADER: a Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text[C] // Eytan Adar & Paul Resnick. *Proceedings of the International AAAI Conference on Web and Social Media*. Washington, DC: AAAI Press(1): 216-225.
- Johnson, K. E. 2006. The Sociocultural Turn and its Challenges for Second Language Teacher Education [J]. *TESOL QUARTERLY* (1): 235-257.
- Kyle, K. , Crossley, S. A. & S. Jarvis. 2021. Assessing the Validity of Lexical Diversity Indices Using Direct Judgements[J]. *Language Assessment Quarterly* (2): 154-170.
- Li, C. , Dewaele, J. M. & G. Jiang. 2020. The Complex Relationship Between Classroom Emotions and EFL Achievement in China[J]. *Applied Linguistics Review* (3): 485-510.
- Mccarthy, P. M. & S. Jarvis. 2010. MTLT, VocD-D, and HD-D: a Validation Study of Sophisticated Approaches to Lexical Diversity Assessment[J]. *Behavior Research Methods* (2): 381-392.
- Naranowicz, M. , Jankowiak, K. & K. Bromberek-Dyzman. 2023. Mood and Gender Effects in Emotional Word Processing in Unbalanced Bilinguals[J]. *International Journal of Bilingualism* (1): 39-60.
- Nation, I. S. P. 2016. *Making and Using Word Lists for Language Learning and Testing* [M]. Amsterdam: John Benjamins.
- Pavlenko, A. & V. Driagina. 2007. Russian Emotion Vocabulary in American Learners' Narratives[J]. *The Modern Language Journal* (2): 213-234.
- Pavlenko, A. 2013. The Affective Turn in SLA: From "Affective Factors" To "Language Desire" And "Commodification of Affect"[J]. *The Affective Dimension in Second Language Acquisition* (3): 3-28

- Prior, M. T. 2019. Elephants in the Room: An “Affective Turn,” Or Just Feeling Our Way? [J]. *The Modern Language Journal* (2): 516-527.
- Rosseel, Y. 2012. Lavaan: an R Package for Structural Equation Modeling[J]. *Journal of Statistical Software* (2): 1-36.
- Teimouri, Y. 2018. Differential Roles of Shame and Guilt in L2 Learning: How Bad is Bad? [J]. *The Modern Language Journal* (4): 632-652.
- Uzun, K. & Ö. G. Ulum. 2022. Sentiment and Sentence Similarity as Predictors of Integrated and Independent L2 Writing Performance[J]. *Acuity: Journal of English Language Pedagogy, Literature and Culture* (1): 1-18.
- Wang, Y. 2020. Emotion and Syntactic Complexity in L2 Writing: a Corpus-Based Study on Chinese College-Level Students' English Writing[J]. *The Asian Journal of Applied Linguistics* (1): 1-17.
- 鲍贵. 2010. 英语学习者语言复杂性变化对比研究[J]. *现代外语* (2): 166-176.
- 鲍贵, 莫俊华, 陆俊花. 2023. R 自然语言处理导论[M]. 南京: 东南大学出版社.
- 董良峰, 朱思达, 张跃. 2025. 中亚汉语学习者二语学习动机、外语焦虑和交际意愿之关系研究——一项基于结构方程模型的实证研究[J]. *外国语文* (2): 175-190.
- 范开芳, 王纯磊. 2023. 外语情绪、外语投入与线上外语学习成效的动态关系研究[J]. *外国语文* (3): 140-148.
- 李成陈, 韩晔. 2022. 外语愉悦、焦虑及无聊情绪对网课学习成效的预测作用[J]. *现代外语* (2): 207-219.
- 毛文伟. 2022. 情感分析工具的性能对比及其在二语习得研究中的应用[J]. *外语电化教学* (5): 94-100.
- 秦丽莉, 何艳华, 欧阳西贝. 2019. 新手教师情感对认知发展影响的叙事研究[J]. *现代外语* (6): 818-829.
- 秦丽莉, 姚澜, 牛宝贵, 等. 2022a. 社会文化理论框架下 *Perezhivanie* 的概念内涵与研究综述[J]. *北京第二外国语学院学报* (1): 11-24.
- 秦丽莉, 姚澜, 牛宝贵. 2022b. 第二语言学习者情感研究述评[J]. *外语教育研究前沿* (2): 51-58.

EFL Proficiency, Emotional Expression and Lexical Complexity in EFL Learners' Oral Production

BAO Gui CHEN Changliang

Abstract: There has been a paucity of research on how EFL learners' emotional expression affects their lexical complexity in spoken production. To this end, the present study explores the extent to which EFL proficiency, the frequency of emotion words, and the valence of emotion words predict lexical complexity, including lexical diversity and lexical sophistication, by analyzing the monologues of 166 English major students. A path analysis revealed that EFL proficiency directly affected lexical sophistication but not lexical diversity; however, an indirect effect of EFL proficiency on lexical diversity was mediated by the frequency of emotion words. The frequency of emotion words directly influenced both lexical sophistication and particularly lexical diversity, whereas the valence of emotion words showed no significant direct effect on either of them. Overall, EFL proficiency and the frequency of emotion words contributed more to lexical diversity than to lexical sophistication. These findings suggest that a pedagogical emphasis should be placed on the use of emotion words and distinctions in emotional gradations, leveraging EFL proficiency to enhance lexical complexity.

Key words: EFL proficiency; emotional expression; frequency of emotion words; valence of emotion words; lexical complexity

责任编辑: 朱晓云