

中国英语学习者的口语语言特征研究

王婷婷 周丹丹

【摘要】口语语言特征是口语表达中语言知识、语篇知识和语用知识的集中体现。本研究以《中国英语能力等级量表》为参照,对英语专业四级口试(简称TEM4口试)中120名不同水平考生的口语表现进行分析,以探究我国英语学习者的区分性口语语言特征。研究发现:1)区分不同水平英语学习者的口语语言特征包含词汇语法准确性、语音语调准确性、词汇多样性、语义丰富度和流利性;2)不同语言特征对不同水平考生的口语表现的区分程度存在差异,某些特征在某些水平上比其他特征更有区分力。研究进一步探讨了研究结果对《量表》口语语言特征量表的验证和细化、口语测试评分的优化和我国英语口语教学实践的启示。

【关键词】《中国英语能力等级量表》;口语语言特征;TEM4口试

【作者简介】王婷婷,南京大学;周丹丹,南京大学外国语学院,E-mail:ddzhouj@nju.edu.cn(南京 210023)。

【原文出处】《现代外语》(广州),2023.6.829~840

1. 引言

口语区分性特征研究能为口语能力构念的探析和拓展、评分量表的构建、考生口语能力的诊断以及考试效度的验证提供重要依据(Gu & Hsieh 2019)。以往研究由于参考的语言能力理论模型或语言标准不同,所囊括的口语特征维度各有侧重(如Brown et al. 2005;兰春寿2008),对各个特征指标的操作化定义和测量方式也不统一,由此得出的结果较为零散。且多数研究仅粗略地将英语学习者分为高低水平(如Crosthwaite & Raquel 2019),颗粒度较低,难以揭示相邻水平学习者的差异性特征。鉴于此,本研究基于《量表》中对口语语言特征的界定,对全国英语专业四级口试(简称TEM4口试)中考生的表现进行分析,试图为我国英语学习者口语区分性特征提供基于统一度量衡的科学描述,同时为《量表》和口试评分标准的验证和优化提供自下而上的、基于考生表现的实证证据。

2. 文献综述

2.1 口语区分性特征研究

近年来,越来越多的研究采用自下而上、基于学习者实际表现的分析方法探究区分不同水平英语学习者的口语特征。根据对口语能力表征方式的不同,此类研究大体可以分为两类。一类从语言系统的不同层级来界定口语特征,主要涉及语音语调、流利性、词汇、语法、内容等维度(如Sato 2012;王华等2018)。每个维度又下含更具体的特征指标。语音维度主要包括音段特征和超音段特征两个层面,前者指单个声音的发音特征,常用指标包括音素错误率和音节结构错误率;后者指单个声音的发音以外的特征,包括重音、语调、节奏等(Isaacs & Trofimovich 2012)。流利性特征主要体现在速度性流利(speed fluency)、中断性流利(breakdown fluency)和修复性流利(repair fluency)三方面(Bosker et al. 2013)。词汇和语法特征主要以语言要素使用的准确性和复杂

性为表征。内容维度的特征主要涉及口语产出的数量和质量,前者主要指口头产出内容中覆盖的关键信息数和内容的丰富度等,后者则涉及信息的准确性、内容的相关度等(Brown et al. 2005)。另一类研究则从更具概括性的“表达效果”的角度展开分析,采用最多的便是复杂度、准确度和流利度这三个指标(如 Révész et al. 2014)。

现有研究在探究口语区别性特征时往往根据各自研究目的,有所侧重地选用不同维度的特征对考生的口语表现进行标注分析(如 Brown et al. 2005; 兰春寿 2008),缺乏统一的理论框架,致使同类研究较为零散、缺乏可比性。此外,以往研究仅粗略地将英语学习者分为高低水平(如 Crosthwaite & Raquel 2019),颗粒度较低,难以揭示相邻水平学习者的差异性特征,无法为口语教学提供更有针对性的启示。因此,有必要在统一的标准参照下对口语区分性特征进行更为细化的分析。

2.2 口语语言区分性特征分析框架

《量表》参考了国内外影响较大的语言能力理论和量表,在实证研究的基础上提出了口语交际能力等级描述(金艳、揭薇 2020),为我国英语教学与测评提供了参考。在此基础上,揭薇、金艳(2022)进一步构建了口头表达语言特征量表,将口语语言特征分为7类:准确性、丰富度、连贯性、流利度、得体性、灵活性和互动性。具体描述语见表1。

表1 口语语言特征描述语(揭薇、金艳 2022:6)

准确性	说话人的语音、语调以及词汇、语法、句法结构等语言要素运用符合规范的程度
丰富度	说话人运用词汇、语法等语言要素的范围大小以及语法结构的复杂程度,同时包括交际过程中参与或发表谈话的充分程度
连贯性	连续话语在内容和形式上组织连接的条理性或逻辑性
流利度	连续话语在语速、语调、停顿等方面的自然程度
得体性	说话人在一定交际场合选用恰当话语、采用合适交际手段以达到交际目的的能力
灵活性	说话人在交际过程中采用恰当交际策略促使交际活动顺利进行的能力
互动性	说话人在交际过程中的对话组织、话轮转换等能力

该量表为口语特征研究提供了具体的、可操作化的评价指标。本研究将以此为分析框架,探究我国英语学习者口语产出中的区分性语言特征。上述七个维度中,准确性、丰富度、连贯性、流利度这四个维度主要是关于语言运用能力和话语组织能力,而得体性、灵活性和互动性这三个维度主要涉及社会语言能力,包含说话人在交际场合中为达到交际目的而采用的交际策略等(揭薇、金艳 2022),多体现于互动任务中。而此类任务涉及多名考生的轮流谈话,产出的为非连续性语篇,难以进行流利性、连贯性等特征的分析。基于上述考虑,本研究仅聚焦于准确性、丰富度、连贯性、流利度这四个维度。

3. 研究方法

3.1 研究问题

本研究旨在回答以下两个研究问题:1)《量表》中界定的准确性、丰富度、连贯性、流利度四个维度的口语语言特征中,哪些能有效区分不同水平的英语学习者? 2)各相邻水平的英语学习者在以上四个维度的口语特征上存在何种差异?

3.2 研究对象

本研究的语料来自参加某年TEM4口试的考生录音,所有考生的答题录音均由英语专业口试中心统一评分。评分流程如下:所有录音会被随机分到包含30至35名考生的组,评分员以组为单位,从内容、语音语调及语法与词汇三方面对考生表现按百分制打分。每份录音由两位评分员评判,取二者的平均分对考生进行组内排序。最后,再按照正态分布的理论常模,将考生最后成绩评定为“优秀”、“良好”、“合格”与“不合格”4个等级。

为保证样本代表性,本研究采用分层随机抽样的方法,从各组中随机选取报道等级为优秀、良好、合格、不合格的考生样本各30份,分别代表高水平组、中高水平组、中低水平组和低水平组。单因子方差分析发现四组考生总分组间均值差异显著($F(3, 116)=2197.322, p < .001, \eta^2=0.983$)。Bonferroni事后检验表明,相邻组别的总分均有显著差异($p < .001$)。

3.3 研究工具

TEM4口试包含复述故事、即席讲话和交谈三项任务。复述任务中考生主要是对原文进行重复或改述,产出的语篇难以完全体现其语言特征;而交谈任务涉及两名考生的轮流谈话,产出的为非连续性语篇,难以进行流利性特征的分析。因此,本研究选择考生在即席讲话这一任务上的录音作为研究语料。该任务要求考生基于个人经历就某一给定话题进行口头表述,3分钟准备,3分钟答题。

3.4 研究过程

研究者将120名考生的录音进行转写,然后从准确性、丰富度、连贯性和流利度四个维度进行编码标注。各个维度的描述语、具体指标和编码工具见下页表2。

准确性由语音语调和词汇语法无错误小句比两个指标体现,具体错误在BFSU qualitative coder中进行人工标注。语音语调错误参照徐鹰、章雅青(2021)的标准,包含吞音、加音、音素错误、单词发音错误、重音错误以及语调不规范等。词汇语法错误参考Crosthwaite & Raquel(2019)的编码体系,包含单词使用错误、搭配错误、句法结构错误、时态错误、主谓一致性错误等。

丰富度和连贯性通过Coh-Metrix来分析。Coh-Metrix能从词汇、句法、语义等角度分析文本的表层和深层特征(杜慧颖、蔡金亭2013)。其中词汇复杂度、句法复杂度和潜语义分析分别反映了词汇、句法和语义的丰富度,可作为丰富度指标。连贯性特征则从指称衔接、连词和潜语义三方面来分析(McNamara et al. 2014)。

流利度由速度流利性、中断流利性和修复流利性三方面体现。速度流利性和中断流利性的计算采用了De Jong & Wembe(2009)编写的Praat程序进行自动分析。对修复流利性的计算采用了人工分析,在BFSU qualitative coder中标注出考生的重复和自我修正行为。

另外,针对语音语调错误、词汇语法错误、考生重复和自我修正行为这四方面人工标注的指标,本研究邀请另一名应用语言学博士生进行了独立标注,标注前对其进行了培训。用于标注者间一致性检验的重复标注比例通常占总样本量的10-25%(O'Connor & Joffe 2020),考虑到口语表达的语言更零散(Martinkova 2013),为提

升标注效度,本研究中重复标注比例设为30%(共40份,每个等级各10份)。分析结果显示,标注者间的一致性系数(Cohen's Kappa)为0.84,具有较高信度。剩余数据由第一作者独立标注。

表2 口语语言特征分析框架

语言特征	描述语	编码工具	指标
准确性	说话人的语音、语调以及词汇、语法、句法结构等语言要素运用符合规范的程度。	人工编码	<ul style="list-style-type: none"> √ 语音语调无错误小句比 √ 词汇语法无错误小句比
丰富度	说话人运用词汇、语法等语言要素的范围大小以及语法结构的复杂程度。	Coh-Metrix	<ul style="list-style-type: none"> √ 词汇复杂度 <ul style="list-style-type: none"> ·词汇多样性D测量 ·内容词具体度、多义度 ·名词、动词上义度 √ 句法复杂度 <ul style="list-style-type: none"> ·主要动词前所含平均单词数量 ·名词词组前的平均修饰语数量 ·所有句子句法相似度均值 √ 语义丰富度 <ul style="list-style-type: none"> ·相邻句子语义相似度平均数 ·所有句子语义相似度平均数 ·句间新旧信息语义平均相似度
连贯性	连续话语在内容和形式上组织连接的条理性或逻辑性。	Coh-Metrix	<ul style="list-style-type: none"> √ 指称衔接 <ul style="list-style-type: none"> ·相邻句子名词重叠、论元重叠、词干重叠平均数 √ 连词发生率 √ 潜在语义分析 <ul style="list-style-type: none"> ·相邻句子语义相似度平均数 ·所有句子语义相似度平均数 ·句间新旧信息语义平均相似度
流利度	连续话语在语速、语调、停顿等方面的自然程度。	Praat 自动分析、人工编码	<ul style="list-style-type: none"> √ 速度流利性 <ul style="list-style-type: none"> ·平均音节长 √ 中断流利性 <ul style="list-style-type: none"> ·每分钟无声停顿数、有声停顿数 ·平均无声停顿时长 √ 修复流利性 <ul style="list-style-type: none"> ·每分钟重复次数 ·每分钟修正次数

4. 研究结果

4.1 区分不同水平英语学习者的口语语言特征

针对研究问题1,本研究对每个特征以口语水平为组间变量进行了单因素方差分析。结果显示(见表3),不同水平的英语学习者在词汇语法无错误小句比($F(3, 116)=19.08, p < .001, \eta^2=0.157^{***}$)、语音语调无错误小句比($F(3, 116)=7.371, p < .001, \eta^2=0.331^{***}$)、词汇多样性D测量($F(3, 116)=4.167, p < .01, \eta^2=0.097^{**}$)、所有句子语义相似度平均数($F(3, 116)=3.594, p < .05, \eta^2=0.085^*$)、平均音节长($F(3, 116)=9.245, p < .001, \eta^2=0.19^{***}$)、每分钟停顿次数($F(3, 116)=3.568, p < .05, \eta^2=0.084^*$)和每分钟重复次数($F(3, 116)=3.688, p < .05, \eta^2=0.087^*$)等指标上

存在显著差异。这些指标分别属于口语语言特征量表中的准确性(词汇语法准确性、语音语调准确性)、丰富度(词汇丰富度、语义丰富度)和流利度三大特征维度。连贯性这一维度并未体现。

准确性 随着口语水平的提升,考生在语音语调和词汇语法准确性上整体呈上升趋势。虽然中高水平组语音语调准确性的平均值略低于中低水平组,但事后 Bonferroni 检验表明,两者无显著差异($p=.11$)。

丰富度 考生的词汇丰富度随口语水平的提高而提升。语义丰富度通过句子语义相似度指标反映,语义重叠度越小,内容越丰富(McNamara et al. 2014)。除低水平组外,其他组口语水平越高,语义相似度越低,即内容越丰富。

流利度 平均音节长是速度流利性的指标,值越大,流利度越低(Bosker et al. 2013)。各水平考生在该指标并未呈现线性增长或下降的趋势。仅中低水平组的流利度显著低于中高水平组。每分钟停顿次数基本呈现口语水平越高,停顿次数越少的趋势。中高水平组的停顿次数少于高水平组,但较后者,前者平均停顿时长更长。在每分钟重复次数上,除低水平组外,其他各组口语水平越高,重复频率越低。

4.2 各相邻水平间的区分性特征

针对研究问题2,本研究进一步对存在主效应的语言特征进行了事后两两比较,以探究区分相邻水平考生的语言特征。下页表3呈现了相邻水平间的对比结果,包含效应量偏 η^2 和Cohen's d ,以及在 $p < .05$, $p < .01$ 和 $p < .001$ 三个水平上的显著性(由*标注)。结果显示,除词汇语法准确性能区分各个水平的英语学习者外,其他语言特征仅对特定水平的学习者具有区分度。高水平和中高水平学习者在平均每分钟重复次数($p < .01$, $d=2.09^{**}$)、所有句子语义相似度平均数($p < .05$, $d=0.53^*$)和词汇多样性D测量($p < .05$, $d=0.69^*$)上存在显著差异。中高水平和中低水平学习者在平均音节长($p < .001$, $d=0.90^{***}$)和动词上义度($p < .05$, $d=0.53^*$)上差异显著。中低水平和低水平组的区分性语言特征为语音语调无错误小句比($p < .01$, $d=0.82^{**}$)、平均每分钟修正次数($p < .05$, $d=0.62^*$)、每分钟停顿次数($p < .05$, $d=0.61^*$)和相邻句子论元重叠平均数($p < .05$, $d=0.51^*$)。

5. 讨论

本研究发现区分不同水平英语学习者的口语语言特征包括:词汇语法准确性、语音语调准确性、词汇多样性、语义丰富度和流利性,部分验证了揭薇、金艳(2022)构建的口语语言特征量表,涉及了口语语言的准确性、丰富度和流利度特征,但连贯性特征并未体现。这可能与TEM4口试评分标准的设置有关,其中并未包含“连贯性”等描述语,因此,评分员也并未运用这一特征来评判考生表现。另一个可能的原因是相较于书面表达,口头产出结构更零散,句式更不规整(Martinkova 2013),而Coh-matrix对连贯性的测量主要是基于对指称衔接、语义重叠和连接词的分析,对连接手段使用的有效性则不加以区分。如很多考生过度使用“and”、“so”等对话篇连贯性无实际贡献的连接词在Coh-matrix中也会被算作是连贯性的正向指标。此外,Coh-matrix对隐形衔接手段(如“With that being said,...”,“Instead of scolding me for doing that,...”)的分析尚不成熟,这也会影响分析结果。

通过对各相邻水平学习者的表现进行对比,我们发现,并非每个特征对所有水平的考生都有区分度。在

表3 总体方差分析及相邻水平间对比结果

特征	指标		总体方差分析	高水平 vs. 中高水平	中高水平 vs. 中低水平	中低水平 vs. 低水平
			η^2	Cohen's <i>d</i>		
准确性	语音语调无错误小句比		0.157***	0.42	0.42	0.82**
	词汇语法无错误小句比		0.331***	1.24**	0.55*	0.63*
主丰富度	词汇复杂度	词汇多样性D测量	0.097**	0.69*	0.13	0.32
		内容词具体度	0.050	0.42	0.36	0.13
		内容词多义度	0.038	0.45	0.04	0.23
		名词上义度	0.024	0.33	0.03	0.15
		动词上义度	0.047	0.01	0.53*	0.02
	句法复杂度	主要动词前所含平均单词数	0.022	0.19	0.20	0.33
		名词词组前的平均修饰语数	0.026	0.01	0.05	0.01
		所有句子句法相似度均值	0.021	0.08	0.28	0.37
	语义丰富度	相邻句子语义相似度平均数	0.042	0.15	0.17	0.23
		所有句子语义相似度平均数	0.085*	0.53*	0.31	0.01
		句间新旧信息语义平均相似度	0.005	0.14	0.09	0.16
连贯性	指称衔接	相邻句子名词重叠平均数	0.022	0.06	0.37	0.09
		相邻句子论元重叠平均数	0.055	0.52	0.13	0.51*
		相邻句子词干重叠平均数	0.015	0.01	0.16	0.13
	潜语义分析	相邻句子语义相似度平均数	0.042	0.15	0.17	0.23
		所有句子语义相似度平均数	0.085*	0.53*	0.31	0.01
		句间新旧信息语义平均相似度	0.005	0.14	0.09	0.16
	连词	所有连词发生率	0.011	0.04	0.17	0.06
流利度	速度流利性	平均音节长	0.19***	0.02	0.90***	0.23
	中断流利性	每分钟停顿次数	0.084*	0.63	0.21	0.61*
		平均停顿时长	0.009	0.04	0.17	0.01
	修复流利性	平均每分钟重复次数	0.087*	2.09**	0.23	0.21*
平均每分钟修正次数		0.053	0.20	0.38	0.62*	

注: * $p < .05$, ** $p < .01$, *** $p < .001$.

以上语言特征中,除词汇语法准确性能区分各个水平的英语学习者外,其他特征仅能区分特定水平的学习者。这表明随着口语水平的提升,口语的各个特征并不是呈现线性发展的(Gu et al. 2019)。

语音语调准确性和语篇连贯性仅能区分低水平和中低水平学习者。发音的准确性对低水平和中低水平英语学习者的区分力在前人研究中已得到证实(Higgs & Clifford 1982)。但这一特征在其他相邻水平间无显著差异,并未有相关研究佐证。为深究其原因,笔者进一步对其他各水平的考生录音进行了分析,发现随着语言水平提升,语音语调准确性的区分力逐渐降低,发音清晰和语调自然逐渐成为差异性特征。对于语篇连贯性差异仅体现于低水平和中低水平学习者间这一结果,通过进一步分析各水平考生录音,发现原因可能在于,相较于低水平学习者,中低水平学习者会有意识地使用简单的显性衔接手段(如“however”、“ultimately”等副词)来实现句际间的连贯,这更容易被 Coh-metrix 标记为连贯性特征。随着语言水平的提升,除了显性衔接手段,较高水平的学习者也会使用一些隐性衔接手段来提升口头表达的条理性和连贯性,然而这些隐形衔接手段未能

很好地被 Coh-Metrix 识别。

速度流利性和动词的上义度仅能区分中低水平和中高水平学习者。中低水平学习者的速度流利性显著低于中高水平,这可能由于这一水平的考生二语资源缺乏、二语知识程序化程度较低,只能通过放慢语速来释放注意力资源以使得口语产出顺利进行(Ahmadian & Tavakoli 2011)。动词的上义度反映了动词的具体程度,值越大,动词越具体(Crossley et al. 2009)。本研究发现,随着学习者口语水平由低水平往高水平过渡,他们倾向于使用更具体、更精确的动词来进行口头表达。这一结果与先前研究中发现的“随着语言水平的提升,学习者倾向于使用低频词、抽象词和下义词”(Eguchi & Kyle 2020)这一结果相左。这可能与任务特征有关,本研究选择的任务是要求考生口头描述自身的一段经历,较高水平学习者通常能运用具体、精确的表达详细地展开叙述,而较低水平者只能用有限的语言来表达。

修复流利性、语义丰富度和词汇多样性是区分中高水平和高水平学习者的口语语言特征,三个指标均随口语水平提升而提升。在修复流利性上,中高水平学习者会通过更多的重复或修正以产出更为复杂和准确的内容,随着口语水平进一步提高,流利度也可以与复杂度和准确度等共同提升(李茶、隋铭才 2017),这一阶段学习者自我修复的频率也会降低。语义丰富度反映了内容丰富度。高水平组显著高于中高水平组,这表明内容的丰富度和创意性是这两个水平学习者特有的差异性特征。在词汇多样性上,前人研究发现该指标最能有效预测二语学习者口语水平(Crossley et al. 2009),本研究进一步证明,该特征是中高水平往高水平学习者突破的重要标志。

本研究为我们提供如下启示:第一,在口语教学中,针对不同水平学习者的特点,可以采取分层教学干预措施。如,对于中高水平英语学习者,应聚焦于词汇多样性的拓展,结合阅读和听力训练增加内容性知识的输入,避免口头表达过于空洞单一。针对中低水平的学习者,应增加口头产出的频率,以提升表达的自动化程度。对于低水平学习者,则应更多关注基础的发音训练,夯实基础语法知识。第二,口试评分标准的描述语应对标最新的口语能力构念和语言能力标准,尽量涵盖口语区分性特征的核心指标。例如,TEM4 口试可增加“使用显性或隐性衔接手段处理口头表达的条理、层次和逻辑”等与连贯性相关的描述语。第三,《量表》除了涵盖不同维度的口语语言特征之外,仍需进一步细化不同能力级别的英语学习者的差异性特征,从而为不同阶段学生英语口语能力的提升提供更具体的参考。

结合本研究发现,我们进一步细化了揭薇、金艳(2022)构建的口语语言特征量表,凸显了各维度相邻水平的显性差异性特征(见下页表4)。

在语音语调准确性上,从低水平到中低水平,仅要求无发音错误。从中低水平往上,不仅要求语音语调准确,还强调发音清晰、语调自然。中低水平到高水平学习者落于“发音清晰”和“语调自然”这两个连续体上的不同位置,但其中的显性差异性指标尚不明晰,仍需未来进一步探究。在词汇语法准确性上,从低水平到中低水平,仅须尽可能减少语言错误。从中低水平到中高水平,在无错误的基础上,还囊括了正确使用复杂形式的词汇和语法结构。例如,一个中高水平考生在“My ruler was very old and ugly, with some scratches on it.”句子中正确运用介词短语做伴随状语。从中高水平到高水平,除以上能力要求,还进一步强调了措辞的准确性和信

表4 各相邻水平英语学习者区分性特征

	低水平-中低水平	中低水平-中高水平	中高水平-高水平
准确性	语音语调符合规范 能正确使用简单形式的词汇、语法、句法结构等语言要素	语音语调符合规范;发音清晰,语调自然 能正确使用复杂形式的语言要素	语音语调符合规范;发音清晰,语调自然能准确运用复杂多样的语言要素
流利性	能在较少停顿情况下,语流顺畅地完成简短的口头表达	能运用自然的语速产出较为充实的内容	能自如、流利地,充实、详细地进行表达
复杂性	能使用基础的表达进行口语产出	能使用多样化的表达进行口语产出	能用多样化、复杂的表达产出丰富的内容
连贯性	能用简单词句等显性衔接手段	能结合显性和隐形衔接手段提升条理性 and 连贯性	

息承载力,即不仅要正确使用、还要准确运用复杂且多样化的语言表达,譬如,一名高水平考生使用“cast a glance towards”动词短语精准描述了“瞟了一眼”这一动作。

在流利度上,从低水平到中低水平,只要求能在较少停顿的情况下完成简短的口头表达(对话速不做要求)。从中低水平到中高水平,对话速提出要求,要听起来自然,同时内容的充实度需要提升。从中高水平到高水平,则进一步要求他们能在速度流利性和修复流利性上同时提升,且口头表达的内容要更加充实和详细。

在复杂性上,对低水平学习者不做硬性要求,仅要求他们能够使用简单的语言形式进行口头表达。从中低水平到中高水平,词法和句法应达到一定的多样性(diversity),包括词义和词形的基本变化、较长句和复合句的使用等。从中高水平到高水平,不仅包含多样性,还涉及复杂性(sophistication),包括使用信息承载力更高(more informative)的表达、采用多种修辞手法、论证方式等。

在连贯性上,从低水平到中低水平,仅要求能使用简单的显性衔接手段。对于中低水平以上的英语学习者,除了显性衔接手段外,还要能够使用一些隐性衔接手段来提升口头表达的条理性和连贯性。

6. 结语

本研究以《量表》为参照,通过分析TEM4口试即席谈话任务中考生的表现,探究了区分不同水平的英语学习者的口语语言特征。研究发现:1)区分不同水平英语学习者的口语语言特征包含词汇语法准确性、语音语调准确性、词汇多样性、语义丰富度和流利性。这基本与《量表》所界定的口语语言特征相吻合,但不同水平的英语学习者在连贯性这一指标上未体现出差异性。2)不同语言特征对不同水平的口语表现的区分程度存在差异,除词汇语法准确性能区分各个水平的考生外,其他语言特征仅对特定水平的学习者具有区分度。

本研究为基于语言标准的口语区分性特征研究提供了一定的思路,研究结果对我国英语口语教学、口语考试评分的优化,以及《量表》口语语言特征描述的验证和细化具有参考价值。需要指出的是,本研究只是从语言特征维度探究了影响考生表现的区分性特征,未涉及内容维度。内容维度口语考试构念中不可或缺的因素,在对口语能力的评判中占有很大的权重(Sato 2012)。未来研究可以结合语篇分析、话语分析等方法探究考生口语表现的内容区分性特征。

 参考文献:

- [1]Ahmadian, M. & M. Tavakoli. 2011. The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research* 285: 373-376.
- [2]Bosker, H. R., A.-F. Pinget, H. Quené, T. Sanders & N. H. de Jong. 2013. What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing* 30(2): 159-175.
- [3]Brown, A., N. Iwashita & T. McNamara. 2005. An examination of rater orientations and test-taker performance on English-for-academic-purposes speaking tasks. *ETS Research Report Series*(1): i-157.
- [4]Crossley, S. A., T. Salsbury & D. S. McNamara. 2009. Measuring L2 lexical growth using hypernymic relationships. *Language Learning* 59: 307-334.
- [5]Crosthwaite, P. R. & M. Raquel. 2019. Validating an L2 Academic group oral assessment: Insights from a spoken learner corpus. *Language Assessment Quarterly* 16(1): 39-63.
- [6]De Jong, N. H. & T. Wempe. 2009. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods* 41(2): 385-390.
- [7]Du, Huiying(杜慧颖) & Cai Jinting(蔡金亭).2013.A Coh-Metrix-based model of linguistic features predicting argumentative writing quality of EFL learners. *Modern Foreign Languages*(3): 293-300, 331.[2013, 基于 Coh-Metrix 的中国英语学习者议论文写作质量预测模型研究.《现代外语》第3期.]
- [8]Eguchi, M. & K. Kyle. 2020. Continuing to explore the multidimensional nature of lexical sophistication: The case of oral proficiency interviews. *The Modern Language Journal* 104(2): 381-400.
- [9]Gu, L. & C.-N. Hsieh. 2019. Distinguishing features of young English language learners' oral performance. *Language Assessment Quarterly* 16(2): 180-195.
- [10]Higgs, T. & R. Clifford. 1982. The push towards communication. In T. V. Higgs(ed) *Curriculum, Competence, and the Foreign Language Teacher*. Lincolnwood, IL: National Textbook Company, 57-79.
- [11]Isaacs, T. & P. Trofimovich. 2012. Deconstructing comprehensibility: Identifying the linguistic influence on listener's L2 comprehensibility ratings. *Studies in Second Language Acquisition*(3): 475-505.
- [12]Jie, Wei(揭薇) & Jin, Yan(金艳).2022. CSE Typical Activities Scales and Language Features Scales of Oral Expression. *Foreign Language World*(6): 2-10.[2022,《中国英语能力等级量表》口头表达典型活动和语言特征量表研究.《外语界》第6期.]
- [13]Jin, Yan(金艳) & Jie, Wei(揭薇).2020. *China's Standards of English Language Ability: CSE Speaking Scales*. Beijing: Higher Education Press.[2020,《中国英语能力等级量表——口语力量表研究》.北京:高等教育出版社.]
- [14]Lan, Chunshou(兰春寿).2008. The effect of test takers' linguistic features on their oral test scores. *Journal of Tianjin Foreign Studies University*(4): 20-26.[2008, 口语产生的语言特征对英语口语评分的影响.《天津外国语学院学报》第4期.]
- [15]Li, Cha(李茶) & Sui, Mingcai(隋铭才).2017.A study of complexity, accuracy and fluency in English learners' oral English development in light of Complexity Theory. *Foreign Language Teaching and Research*(3): 392-404, 480-481.[2017, 基于复杂理论的英语

学习者口语复杂度、准确度、流利度发展研究.《外语教学与研究》第3期.]

[16]Martinková, P. 2013. Means of Coherence and cohesion in spoken and written discourse. *Linguistics*: 167–181.

[17]McNamara, D., A. Graesser, P. McCarthy & Z. Cai. 2014. *Automated Evaluation of Text and Discourse with Coh-Matrix*.

Cambridge: Cambridge University Press.

[18]O’ Connor, C. & H. Joffe. 2020. Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods* 19: 160940691989922.

[19]Révész, A., M. Ekiert & E. Torgersen. 2014. The effects of complexity, accuracy, and fluency on communicative adequacy in oral task performance. *Applied Linguistics* 37: 828–848.

[20]Sato, T. 2012. The contribution of test-takers’ speech content to scores on an English oral proficiency test. *Language Testing* 29 (2): 223–241.

[21]Wang, Hua(王华), Fan, Tingting(范婷婷) & Zeng, Yongqiang(曾用强).2018. Investigating the construct of speaking proficiency under the listening-to-speak integrated task. *Modern Foreign Languages*(3): 413–424.[2018,听说综合性任务下口语能力本质研究.《现代外语》第3期.]

[22]Xu, Ying(徐鹰) & Zhang, Yaqing(章雅青).2021. Distinctive pronunciation features in integrated listening-to-speak tasks. *Foreign Language Testing and Teaching*(3939–48).[2021,综合性听说结合任务构念中的语音特征研究.《外语测试与教学》第3期.]

Analyzing the Distinguishing Language Features of Chinese

EFL Learners’ Oral Expression

Wang Tingting Zhou Dandan

Abstract: Language features of oral expression are the converged embodiment of linguistic, discourse and pragmatic knowledge. With reference to the China’s Standards of English Language Ability (CSE), this study examined 120 test takers’ performances across four proficiency levels in TEM4–Oral (Test for English Majors Band 4 Oral Test), aiming to delineate the distinguishing language features of Chinese EFL learners’ oral expression. Results indicated that:(1)seven features, namely lexical and grammatical accuracy, pronunciation and intonation accuracy, lexical diversity, semantic richness, and fluency can significantly discriminate L2 learners across levels; and(2)not all features differed between all pairs of adjacent levels across the proficiency scale. Some features had more differentiating power than others at certain levels. Based on such findings, suggestions were provided for validation and refinement of CSE’s Language Features Scales of Oral Expression, optimization of rating for English speaking assessment, as well as oral English teaching and learning practice in China.

Key words: China’s Standards of English Language Ability(CSE); language features of oral expression; TEM4–Oral