

大学英语四、六级听力测试命题质量历时研究

过君琰

(无锡学院, 江苏 无锡 214105)

摘要: 本研究以大学英语四、六级听力测试(2016-2024)为样本, 参照 Bachman & Palmer 的考试任务特征框架, 依据《大学英语教学指南(2020版)》、《全国大学英语四、六级考试大纲(2016年修订版)》和《中国英语能力等级量表》(2024版)中的听力理解能力总表, 从语篇输入和预期回答两个方面对其进行命题质量历时分析。研究发现四、六级听力测试基本符合《指南》、《考纲》和《量表》的要求, 内容效度较高, 但主题和考查技能维度比较集中、话轮数跨度较大、部分语篇难度过大、语速过快。基于研究结果, 文章为今后的四、六级听力测试命题和大学英语教学提出一些建议。

关键词: 听力测试; 语篇输入; 预期回答; 大学英语教学

基金项目: 十二批“中国外语教育基金”项目(项目编号: ZGWYJYJJ12A114)

DOI: doi.org/10.70693/rwsk.v2i2.202

一、引言

大学英语四、六级考试(College English Test, CET)从1987年开始实施, 至今已有将近四十年的历史。作为服务于大学英语课程体系的考试, 四、六级考试的发展历程体现了大学英语课程对学生的英语能力要求的变化过程。早期大学英语课程目标是培养学生具有较强的阅读能力、一定的听(和译)的能力、初步的写和说的能力^[1]。《大学英语教学指南(2020版)》指出大学英语课程的主要目的之一是“在高中英语教学的基础上进一步提高学生英语听、说、读、写、译等语言综合应用能力”^[2]。因此, 大学英语四、六级考试内容、题型和比例也发生了相应的变化, 从初期的阅读占最高比重(40%), 其它部分包括听力(20%)、写作(15%)、语言知识(15%)和综合能力(10%), 到现在听力和阅读比重一致, 各占35%, 翻译和写作各占15%^[3]。

听力是语言理解和交际能力的基础。良好的听力能力能够帮助我们在交际中迅速理解他人的言语内容, 准确理解其意图, 进而提高沟通的有效性。同时听力还是提升口语、阅读和写作等其它语言技能的重要基础。鉴于听力的重要性及其在四、六级考试中的突出比重, 本研究以大学英语四、六级考试2016年改革后的听力文本及问题进行命题质量分析, 以期对未来的考试命题提供一定的参考和建议。

二、研究框架

Bachman & Palmer^[4]的测试任务特征理论包括测试环境、测试说明、语篇输入、预期回答、以及语篇输入和预期回答的关系五个部分。国内研究者主要从语篇输入和预期回答这两个方面对国内听力测试进行分析。语篇输入是指考生要处理和回答的测试材料, 包括材料的呈现形式和语言特征。前者包括材料使用的语言、长度和输入材料的速度等, 后者包括材料的话题、结构和语篇特征。预期回答是指考生对测试任务做出的反应和输出。辜向东、李亚果^{[5][6]}的两项研究围绕2005年的CET改革, 分析了1996至2008年的四、六级听力理解真题的语篇输入和预期回答任务特征。除了CET听力以外研究者还运用该理论分析了高考英语全国卷^[7]、英语专业四级^[8]和法语专业八级听力测试的命题质量^[9]。

然而纵观现有CET听力试题的研究, 我们发现, 目前还没有对2016年最近一次改革过后的试题分析。本研究参考Bachman & Palmer^[4]的语篇输入和预期回答两方面的任务特征、结合《大学英语教学指南(2020版)》(以下简称《指南》)、《全国大学英语四、六级考试大纲(2016年修订版)》(以下简称《考纲》)和《中国英语能力等级量表》(2024)(以下简称《量表》)中的听力理解能力总表^[10], 对2016年6月至2024年6月间的大学英语四、六级听力测试真题进行综合分析, 研究其命题质量。

作者简介: 过君琰(1984—), 女, 博士, 研究方向外语教学、外语测试。

通讯作者: 过君琰

三、研究方法

(一) 研究数据库

研究者收集整理了从 2016 年 6 月到 2024 年 6 月共 35 套大学英语四、六级听力真题文本和问题，分别构建了四级和六级听力两个平行语料库。

(二) 分析框架与方法

本研究按照以下框架进行数据分析(表 1)。

表 1 大学英语四、六级命题质量分析框架

项目	内容	描述	涉及部分	分析方法
语篇输入	体裁	记叙文、说明文、议论文、描写文、应用文	除长对话以外的部分	ChatGPT 4o、人工
	主题	文化与艺术、科技与创新、环境与生态、教育与工作、历史与地理、健康与医学、社会与时事、经济与商业、生活与情感	所有部分	
	词汇	词汇丰富度 (形符数、类符数、复杂度、多样性)	所有部分	Vocabprofilers (type, token, frequency), TAALED 1.4.1 (MTLD), TAALES 2.2 (Academic Word List, AWL)
	难度	词汇难度、句法难度、文本难度、四六级考试大纲的覆盖率、超纲词汇词云图	所有部分	阅读分级指南针
	话轮数	长对话的话轮数	长对话	人工
	平均语速	平均每分钟朗读每则听力材料的词数	所有部分	人工
预期回答	考查技能	1. 理解明示的信息 (主旨大意、重要信息或特定细节、说话人明确表达的观点和态度等) 2. 理解隐含的信息 (推断隐含的意义、话语交际功能、说话人的观点和态度等) 3. 运用语言特征理解听力材料 (辨别语音特征、理解句间关系)	所有部分	Chat GPT 4o、人工

四、研究结果与讨论

本节按照表 1 的分析框架将结果进行归纳和讨论。

(一) 语篇输入特征分析

1. 体裁

从表 2 的统计结果可知，大学英语四、六级听力文本体裁较为多样，覆盖了记叙、说明和议论三种体裁。其中说明文占比最大。从各分项来看，除四级的新闻听力外，其余各分项也是说明文的比例最高，这一结果与之前的真题(1996-2008)分析结果一致^{[5][6]}。由此可见，说明文以解释、介绍和阐述为目的，加上客观性、条理性、简洁性和知识性等特征，有助于学生获取信息。与四级相比，六级记叙文比重明显降低，议论文比重加大。议论文旨在表达观点、论证主张，因此注重运用逻辑进行推理，思辨性更强，与其它体裁相比，对学生的听力理解要求更高。由此也可以看出，六级的考查难度要高于四级。六级的第三部分中，讲座最多，共 44 篇 (41.9%)，其次是报道 35 篇 (33.3%) 和讲话 26 篇 (24.8%)。

表 2 大学英语四、六级听力文本体裁统计

试卷类型	测试部分	体裁		
		记叙文	说明文	议论文
四级	短篇新闻	52.4%	45.7%	1.9%
	听力篇章	15.2%	66.7%	18.1%
	总计	33.8%	56.2%	10%
六级	听力篇章	10%	67.1%	22.9%
	讲话/报道/讲座	8.6%	61.9%	29.5%
	总计	9.1%	64%	26.9%

2. 主题

听力材料的题材多样,本研究将其归类为九大主题。从表 3 可知,生活与情感(四级 26.5%, 六级 21.6%)、社会与时事(四级 18.9%, 六级 19.2%)、教育与工作(四级 13.6%, 六级 17.6%)是四、六级听力文本最常考的三大主题,符合《指南》中“题材熟悉”的内容。从具体子部分来看,六级的三个部分均涵盖了九大话题,比四级的部分涵盖更广泛。研究者进一步将四级和六级类型相同的两个子部分(长对话、听力篇章)进行对比。四级的长对话主题中 88.6%是日常对话包含筹备派对、驾考、健身、旅行、网购等日常生活中常见的主题,11.4%是职场对话,主题包括公司业务、商业资讯等。就对话形式而言,其中有 3 个对话(4.3%)是以访谈形式进行,主题分别是剧场的特色与经营、英勇救人和家族企业的发展。六级长对话中访谈形式的对话比重达到 32.9%且受访者类型多样,如作家、教授、市长、退役篮球运动员、时尚杂志编辑、物理学家、植物学家等。51.4%是日常对话,主题包括参加婚礼、创业、推荐电视剧、写论文等,32.9%是职场对话,主题包括公司运营、合同修改、绩效管理等。这一点与《量表》中听力理解能力总表的六级要求能听懂正常语速的职场对话相符。就听力篇章的九大话题覆盖比例而言,六级比四级更均衡,而且同一话题切入更深。例如 2021 年 12 月四级第 1 套听力篇章第 1 篇提出了帮助孩子高效完成家庭作业的建议,而同年 6 月六级第 2 套听力篇章第 1 篇则更细致地讨论了家庭作业对学生的影响及争议。

表 3 大学英语四、六级听力文本主题统计

试卷类型	测试部分	主题								
		生活与情感	社会与时事	教育与工作	健康与医学	科技与创新	环境与生态	经济与商业	历史与地理	文化与艺术
四级	短篇新闻	18%	33.3%	3.8%	7.6%	10.5%	10.5%	3.8%	5.7%	2.9%
	长对话	40%	4.3%	25.7%	4.3%	2.9%	2.9%	12.8%	0%	8.6%
	听力篇章	25.7%	14.3%	15.2%	13.3%	9.5%	9.5%	4.8%	10.5%	0%
	总计	26.5%	18.9%	13.6%	8.9%	8.2%	8.2%	6.4%	6.1%	3.2%
六级	长对话	24.3%	12.9%	27.1%	2.9%	5.7%	4.3%	7.1%	1.4%	14.3%
	听力篇章	12.9%	15.7%	12.9%	17.1%	8.6%	7.1%	7.1%	10%	8.6%
	讲话/报道/讲座	25.7%	25.7%	14.2%	3.8%	7.6%	4.8%	4.8%	6.7%	6.7%
	总计	21.6%	19.2%	17.6%	7.3%	7.4%	5.3%	6.1%	6.1%	9.4%

3. 词汇

形符数(token)是文本中所有单词的数量,而类符数(type)是文本中不同单词的数量。例如 I can't say I like the toy.一句中有 8 个形符数(缩略词 can't 是 can 和 not 分开统计,是 2 个形符)和 7 个类符数(I 重复了一次)。表 4 汇总结果显示,四级听力文本中长对话的最多(分别为 274 个和 146 个),短篇新闻最少(173 个和 107 个)。六级中讲话/报道/讲座题型的形符数和类符数最多(405 和 207),听力篇章最少(262 和 153)。此外,无论是将缩略词分开还合并统计,四级三篇短篇新闻的总单词数均高于四、六级考试大纲的范围(表 5)。词汇的多样性指标以 MTLD (Measure of Textual Lexical Diversity)^[11]来计算,数值越高则说明越多样,我们可以从表 4 看到六级的平均值明显高于四级。词汇复杂度其中之一的指标学术英语词汇(Academic Word List, AWL)^[12]显示,除四级新闻听力外,六级各项的学术英语词汇也多于四级。

表 4 大学英语四、六级听力文本词汇统计

试卷类型	测试部分	词汇							
		形符数 (token)		类符数 (type)		多样性 (MTLD)		复杂度 (AWL)	
		平均	范围	平均	范围	平均	范围	平均	范围
四级	短篇新闻	173	120~208	107	71~134	70.9	39~128.8	0.048	0.005~0.161
	长对话	274	239~302	146	113~170	73.5	43.8~117.7	0.026	0~0.119
	听力篇章	241	211~287	136	112~167	70.7	40.2~116.4	0.055	0~0.149
	总计	224	120~302	128	71~170	71.5	39~128.8	0.045	0~0.161
六级	长对话	321	287~344	170	142~193	79	47.3~121.6	0.041	0.003~0.099
	听力篇章	262	248~292	153	116~179	84.8	39.8~141.2	0.066	0.008~0.15
	讲话/报道/讲座	405	347~464	207	162~243	79.1	42.6~132.3	0.067	0.01~0.137
	总计	340	248~464	181	116~243	80.7	39.8~141.2	0.059	0.003~0.15

表 5 大学英语四、六级听力文本单词数与考纲对比

试卷类型	测试部分	形符数 (缩写词分开计算)	单词数 (缩写词合并计算)	四、六级考试大纲
四级	短篇新闻 (总共)	520	507	总共 450-500
	长对话	274	263	每篇 240-280
	听力篇章	241	236	每篇 220-240
	长对话	321	311	每篇 280-320
六级	听力篇章	262	256	每篇 240-260
	讲话/报道/讲座 (总共)	1216	1190	总共 1200

研究者用 Vocabprofilers (<https://www.lex tutor.ca/vp/comp/>) 的 British National Corpus (BNC) 和 Corpus of Contemporary American English (COCA) 数据库统计四、六级听力文本词频。在听力理解中, 如果听者能够听懂其中 95% 和 98% 的词汇, 即 95% 和 98% 的词汇覆盖率, 则被认为是足够的和理想的^{[13][14][15]}。从表 6 我们可以得知, 总体而言, 最高频的 3,000 词可以覆盖 95% 的四、六级听力文本, 最高频的 4,000 词可以覆盖 98% 的听力文本。具体而言, 四级听力的三部分中, 最高频的 3,000 词和 4,000 词分别可以覆盖 95% 和 98% 的短篇新闻和听力篇章文本, 最高频的 2,000 词和 3,000 词分别可以覆盖 95% 和 98% 的长对话文本。六级听力的三部分中, 最高频的 3,000 词可以覆盖每个部分 95% 的文本, 最高频的 4,000 词可以覆盖长对话和讲话/报道/讲座 98% 的文本, 而听力篇章则需要前 5,000 词才可以覆盖 98% 的文本。

表 6 大学英语四、六级听力文本词频统计

试卷类型	试题部分	词频平均覆盖率				
		1,000	2,000	3,000	4,000	5,000
四级	短篇新闻	81.8%	92.4%	97.2%	98.5%	/
	长对话	90.5%	96.3%	99.2%	99.6%	/
	听力篇章	82.2%	92.7%	97.6%	98.6%	/
	总计	84.1%	93.5%	97.7%	98.7%	/
六级	长对话	86.7%	93.8%	97.4%	98.6%	/
	听力篇章	79%	90%	96.1%	97.9%	98.8%
	讲话/报道/讲座	80.5%	90.6%	96.8%	98.3%	/
	总计	81.8%	91.3%	96.8%	98.3%	/

四、六级听力文本难度通过“阅读分级指难针”(LanguageData.net/tester)进行衡量统计^[16]。“指难针”以2000年至2022年间的中考、高考、大学四、六级和考研英语阅读真题文本进行词汇、句法和篇章三维度12个核心特征进行难度计算,确定难度级别范围,并将其与《量表》中的三至七级衔接。我们将四、六级听力文本输入“指难针”,系统将分析文本与阅读真题文本的难度级别范围进行匹配,呈现其难度等级与难度数值。例如表7中,CET-4短篇新闻的词汇难度是5.15,整数为5,说明该文本难度大致处于“五级(CET-4)”水平¹,小数位为0.15,说明该文本在同等级别的阅读真题文本中难度排序百分位约为15%。

表7的统计结果显示,四级听力词汇与文本难度与阅读真题难度相比处于四级(高考)水平,句法难度处于三级(中考)水平。三个部分中短篇新闻的词汇难度和文本难度最高,有21篇(20%)词汇难度达到了六级(CET-6)水平,16篇(15.2%)文本难度达到了六级(CET-6)水平。这可能与新闻词汇的多样化和专业化有关。六级听力词汇与文本难度与阅读真题难度相比处于五级(CET-4)水平,句法难度处于四级(高考)水平。由此可见,六级听力三个维度的难度均高于四级。三个部分中,讲话/报道/讲座部分的三个维度的难度最高,其中分别有41篇(39%)、12篇(11.4%)和39篇(37.1%)的词汇难度、句法难度和文本难度达到了七级(考研)水平。值得注意的是,四、六级听力文本的句法难度比词汇和文本难度低,主要原因可能是听力要求听者在短时间内理解信息,复杂的句法会增加难度和记忆负担,同时听力文本也依赖语境和辅助信息来帮助听者理解信息。

研究者同时还运用“指难针”计算《大学英语课程教学要求》(2007)(以下简称《要求》)^[17]词汇和四、六级考试大纲词汇对四、六级听力文本的词汇覆盖比例。四级词汇可以覆盖四级听力99%以上的词汇。因为98%是理想的词汇覆盖率,所以掌握《要求》和《考纲》的四级词汇对四级听力理解已足够。然而,四级词汇对六级听力的覆盖率低于98%(尤其是第二部分听力篇章和第三部分讲话/报道/讲座),有11篇覆盖率甚至低于95%,因此四级词汇对六级听力考试是不够的,需要六级词汇才可以覆盖99%以上的听力文本词汇。但仍有一些文本过难,例如2017年6月第2套讲话/报道/讲座部分第3篇六级词汇覆盖率只有93.09%。

四级听力文本中,四级超纲词汇最高频的是own,包括它的派生词owner,ownership,共出现在22个语篇中,其次是participant,出现在12个语篇中。六级听力文本中,六级超纲词汇最高频effective包括其副词effectively和名词effectiveness,都是出现在16个语篇中,其次是participant和own(owner,ownership),都是出现在9个语篇中。

表7 大学英语四、六级听力文本难度

试卷类型	试题部分	难度				
		词汇难度	句法难度	文本难度	四级词汇覆盖率	六级词汇覆盖率
四级	短篇新闻	5.15	4.02	5.13	99.03%	99.13%
	长对话	4.15	3.59	4.11	99.66%	99.68%
	听力篇章	4.90	4.14	4.88	99.12%	99.18%
	总计	4.80	3.96	4.78	99.22%	99.29%
六级	长对话	4.68	3.96	4.68	98.31%	99.53%
	听力篇章	5.48	4.30	5.44	97.60%	99.12%
	讲话/报道/讲座	6.26	5.50	6.29	97.92%	99.10%
	总计	5.58	4.72	5.59	97.94%	99.23%

5.话轮数和平均语速

四级和六级长对话的平均话轮数都是6轮,总体与之前的5-8轮的研究结果相符^{[5][6]}。四级最少2轮,最长10轮,有20%少于5轮,8.6%多于8轮。六级最少是3轮,最长是13轮,21.4%少于5轮,10%多于8轮。由此可见,四、六级话轮数跨度较大。

根据表8的统计结果,四、六级总平均语速基本一致(140词每分钟左右),其中长对话相对来说语速较快约为150词每分钟,属于《量表》的听力理解总表中的语速较快范畴(约140~180词/分钟)。其它各项为每分钟135词左右,属于语速正常范围(100~140词/分钟)。与《考纲》要求(四级每分钟120~140词,六级每分钟140~160词)相比,四级真题的属于正常偏快语速,六级真题的语速基本符合。其中六级2016年6月、2020年12月听力篇章第2篇均超过每分钟167词。四级2016年12月第1篇长对话和六级2023年12月第2篇长对话

¹ 三级(中考):3.00-3.99,四级(高考):4.00-4.99,五级(CET-4):5.00-5.99,六级(CET-6):6.00-6.99,七级(考研):7.00-7.99。

的语速均超过每分钟 180 词。四级的录音时长为平均 94 秒，六级为 143 秒。

表 8 大学英语四、六级听力文本语速和时长统计

试卷类型	试题部分	语速(wpm)		时长(s)	
		平均	范围	平均	范围
四级	短篇新闻	135	115~167	75	52~94
	长对话	155	113~181	103	79~142
	听力篇章	134	113~160	106	84~128
	总计	140	113~181	94	52~142
六级	长对话	154	131~183	122	102~142
	听力篇章	134	108~167	116	92~145
	讲话/报道/讲座	136	106~158	175	139~222
	总计	141	106~183	143	92~222

(二) 预期回答特征分析

四、六级听力理解考查的技能虽覆盖了《考纲》的理解明示的信息和理解隐含的信息两大类，但集中在考查理解明示信息（四级 97.9%，六级 98.6%），尤其是考查重要的信息或特定的细节（四级 95.5%，六级 97.4%，见表 9）。这符合《指南》中基础目标的“获取要点和细节”^[2]、提高目标中的“抓住要点和相关细节”^[2]以及《量表》中听力理解能力总表五级的描述语“获取要点和细节”^[10]。

表 9 大学英语四、六级听力考查技能分析

试卷类型	试题部分	考查技能						语用语言特征 理解听力材料	
		理解明示的信息			理解隐含的信息				
		主旨大意	信息细节	观点态度	隐含意义	话语交际	观点态度	语音特征	句间关系
四级	短篇新闻	4.9%	93.9%		1.2%				
	长对话	0.7%	93.9%	1.1%	3.2%	1.1%			
	听力篇章	1.1%	97.7%		0.9%		0.3%		
	总计	2.1%	95.5%	0.3%	1.7%	0.3%	0.1%		
六级	长对话	1.4%	96.1%		2.5%				
	听力篇章	1.2%	98.4%		0.4%				
	讲话/报道/讲座	1.1%	97.8%		1.1%				
	总计	1.2%	97.4%		1.4%				

五、总结与建议

通过对 2016 年 6 月至 2024 年 6 月大学英语四、六级听力测试命题质量的实证研究，本文得出以下结论：首先，整体来说四、六级听力测试基本符合《指南》、《考纲》和《量表》的要求，内容效度较高。题材较为多样，主题比较广泛，难度基本适中。四级和六级也有一定的区分度。其次，具体而言测试也存在一定的不足。例如：主题方面比较集中于“生活与情感”、“社会与时事”和“教育与工作”这三个方面。难度方面，部分六级听力讲话/报道/讲座部分语篇的难度已经达到考研阅读真题水平。此外四级和六级都存在话轮数跨度较大、部分语篇语速过快等问题，且四六级总体平均语速区别不明显。考查技能方面过于集中在“理解明示的信息”，尤其是“重要信息或特定细节”的复现。

基于上述研究结果，本文尝试为大学英语四、六级测试以及大学英语教学提出以下建议：首先，在测试命题方面，应当丰富听力文本主题内容，稳定语篇难度和语速，拓展考查技能，提升全面性。其次，在教学中，应当注重拓展学生的听觉词汇知识和词汇深度知识（近义词、词性）。本研究结果显示最高频的 3,000 和 4,000 词可以覆盖 95%和 98%的四、六级听力文本，然而 Guo^[18]的研究显示，中国英语学习者只掌握了最高频的 1,000 词的

听觉词汇知识。同时教师还应培养学生的信息概括、总结和解释能力。

参考文献:

- [1] 金艳, 杨惠中. 走中国特色的语言测试道路:大学英语四、六级考试三十年的启示[J]. 外语界, 2018(2): 29-39.
- [2] 教育部高等学校大学外语教学指导委员会. 大学英语教学指南[M]. 北京: 高等教育出版社, 2020: 4、11、13.
- [3] 全国大学英语四、六级考试委员会. 全国大学英语四、六级考试大纲 (2016 年修订版) [EB/OL]. (2016) [2024-12-15]. <https://cet.neea.edu.cn/html1/folder/16113/1588-1.htm>
- [4] BACHMAN, L. F., PALMER, A. S. Language testing in practice: Designing and developing useful language tests[M]. Oxford: Oxford University Press, 1996: 49-60.
- [5] 辜向东, 李亚果. 改革后 CET 听力测试语篇输入与预期回答任务特征分析[J]. 西安外国语大学学报, 2010(4): 71-74+79.
- [6] 辜向东, 李亚果. CET 听力测试语篇输入和预期回答任务特征历时分析(1996-2007)[J]. 外语测试与教学, 2012(3): 17-26.
- [7] 彭莹莹. 2017 年高考全国卷英语听力测试语篇输入与预期回答任务特征分析[J]. 教育测量与评价, 2018(4): 11-16.
- [8] 孟凡娜. TEM-4 听力测试语篇输入及预期回答任务特征历时分析(2005-2014 年)[J]. 考试研究, 2015(6): 68-75.
- [9] 卢晓帆. TFS8 听力测试语篇输入及预期回答任务特征历时分析(2009-2019)[J]. 法语国家与地区研究, 2021(2): 34-46+92.
- [10] 中华人民共和国教育部、国家语言文字工作委员会. 中国英语能力等级量表[S]. 上海: 上海外语教育出版社, 2024: 9.
- [11] MCCARTHY, P. M. An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual, lexical diversity (MTLD)[D]. Memphis: The University of Memphis, 2005.
- [12] COXHEAD, A. A new academic word list[J]. TESOL Quarterly, 2000(2): 213-238.
- [13] DANG, T. N. Y., WEBB, S. The lexical profile of academic spoken English[J]. English for Specific Purposes, 2014(33): 66-76.
- [14] NATION, I. (2006). How large a vocabulary is needed for reading and listening?[J] The Canadian Modern Language Review, 2006(1): 59-82.
- [15] SCHMITT, N., COBB, T., HORST, M., & SCHMITT, D. How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), Nation (2006) and Cobb (2007)[J]. Language Teaching, 2017(2): 212-226.
- [16] 金檀, 陆小飞. 阅读分级指南针[EB/OL].(2023-03)[2024-12-24]. https://doi.org/10.1007/978-7-309-05825-1_110825
- [17] 教育部办公厅. 关于印发《大学英语课程教学要求》的通知 [EB/OL]. (2007-07-10)[2025-2-10]. http://www.moe.gov.cn/srcsite/A08/s7056/200707/t20070710_110825.html
- [18] GUO, J. Exploring the role of vocabulary in EFL learners' speaking proficiency: A mixed-methods study[D]. Auckland: The University of Auckland, 2003.

CET 4/6 Listening Test Design Quality: A Longitudinal Analysis

Junyan Guo

(Department of General Education, Wuxi University, Wuxi, China)

Abstract: Taking the CET-4 and CET-6 listening tests (2016-2024) as samples, this study analyzes the quality of test papers in terms of the characteristics of the input and the expected response based on Bachman & Palmer's framework for test task characteristics and in accordance with the "College English Teaching Guide (2020)", the "National College English Test CET-4 and CET-6 Syllabus (2016)", and the listening comprehension competency table in the "China Standards of English (2024)." The study finds that the CET-4 and CET-6 listening tests generally meet the requirements of the "Guide," "Syllabus," and "Standards", with relatively high content validity. However, the themes and skills assessed are somewhat concentrated. There is a large variation in the number of speech turns. Some of the materials are too difficult, and the speed of some of the speeches is too fast. Based on the research findings, the paper offers suggestions for future CET-4 and CET-6 listening test design and college English teaching.

Key words: listening test; input; expected response; college English teaching