
英语能力人工智能分级测评实施规范

Implementation Specifications for Artificial Intelligence-Based
Grading and Assessment of English Proficiency

(报批稿)

【日期】 发布

【日期】 实

施

中国智慧工程研究会 发布

目 录

1. 范围	1
2. 规范性引用文件	1
3. 术语和定义	1
4. 缩略语	4
5. 英语能力人工智能分级测评等级划分	4
5.1 英语能力人工智能分级测评框架	4
5.2 英语能力人工智能分级测评阶段划分	7
5.3 英语能力人工智能分级测评评价框架	8
5.4 英语能力人工智能分级测评分项能力评价	10
6. 英语能力人工智能分级测评题型	19
6.1 听力能力	19
6.2 口语能力	22
6.3 阅读能力	23
6.4 写作能力	28
7. 英语能力人工智能分级测评系统建设	31
7.1 系统定义	31
7.2 系统架构	31
7.3 考务管理系统	31
8. 英语能力人工智能分级测评题库建设	35
8.1 系统构架	35
8.2 模块	36

8.3 功能.....	37
8.4 数据汇总	38
8.5 题目来源	39
9. 英语能力人工智能分级测评系统考务	40
9.1 考前准备阶段	41
9.2 考试阶段	41
9.3 考试结束阶段	42
10. 英语能力人工智能分级测评评分系统.....	42
10.1 人工评分	43
10.2 智能评分	44
11. 英语能力人工智能分级测评报告	44
11.1 报告模式	44
11.2 报告类型	45
11.3 报告模块	46

前 言

本文件按照GB/T 1.1-2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国智慧工程研究会提出并归口。

本文件起草单位: 苏州驰声信息科技有限公司、通师前沿（吉林省）教育科技有限公司、驰声教育考试研究院。

本文件主要起草人: 林远东、刘云童、丁俊华、高淑君、甄文超、刘盼宁、潘宁波、刘文、吴靖、徐冬燕、李冬菊、王新、陈雅琴等。

英语能力人工智能分级测评实施规范

1. 范围

本文件规定了计算机辅助英语能力测评的分级体系、测评维度、系统要求及实施规范。

本文件适用于各级各类学校的英语能力计算机测评，可为英语教学、学习及相关研究提供参考。

2. 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对于的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GA/T 1048.1-2013 标准汉译英要求

GB/T 36632-2018 信息技术学习分析技术规范

GB/T 6865-2009 语种熟练程度和外语考试等级代码

GB/T 15528-2013 语音通信用传声器和耳机测量方法

ISO/IEC 25010:2011 系统与软件质量模型

GB/T 25000.62-2014 软件工程 软件产品质量要求与评价 (SQuaRE) 易用性测试报告行业通用格式 (CIF)

GB/T 35273-2020 个人信息安全规范

3. 术语和定义

下列术语和定义适用于本文件。

3.1 英语能力人工智能等级体系 (Level System of AI - Based English Proficiency)

此为基于人工智能技术的英语测评，构建的一套针对不同层次英语学习者能力进行精细区分与阐释的评价体系。它深度融合《中国英语能力等级量表》，借助人工智能的多元分析模型，基于 CEFR（欧洲共同语言参考框架）标准，从语音产出（pronunciation）、语义理解（semantic comprehension）、语用能力（pragmatic competence）及读写技能（literacy skills）四个维度构建多模态评估体系，通过标准化测试与动态诊断相结合的方式，实现学习者英语综合能力的精准分层定位。

3.2 智能人机交互 (Intelligent AI - Human Interactive)

依托前沿的人工智能与计算机网络技术，革新考试的实施与管理模式。它运用自然语言处理技术，依据考试设计需求，通过智能算法精准命题与智能组卷，该系统支持英语听说读写综合测试全流程智能化处理，可实时采集并解析考生的语音输入与文本输入，动态生成并呈现适配听、说、读、写能力的分级测评任务，同步完成多模态答案采集、基于标准的智能评分、多维度学习数据分析，最终输出个性化能力诊断报告与精准结果解读。

3.3 多模态测评 (AI - Based Multimodal Assessment)

运用人工智能语音测评技术、图像识别技术、情感分析技术以及大数据语料库，采用线上多模态测试手段。

3.4 核心维度 (Core Dimensions)

为科学构建英语听说读写能力的多维评价体系，本研究聚焦于全面刻画学习者综合语言能力的发展特征。本评价体系立足实际应用导向，围绕以下核心能力维度展开系统设计：

- 语音韵律能力 (含听力辨音准确性与口语发音规范性)
- 语境理解能力 (涵盖阅读文本推理与听力信息提取加工)
- 逻辑表达能力 (包括观点结构化组织与语言连贯性衔接)
- 跨文化交际能力 (涉及情感精准传递与文化差异适应性调适)

依托人工智能技术体系，我们将通过智能学习模型对各能力维度开展多模态数据采集与深度分析，打通“能力描述—科学量化评估”的闭环链路，最终为学习者生成动态追踪、精准诊断的能力发展画像。

3.5 语音韵律表现能力 (Phonetic and Prosodic Performance Ability)

学习者在人工智能辅助下，于口语输出 (如对话、朗读、演讲) 及听力理解中展现的综合语音素养。具体包括：词汇发音的精准度 (音素、音节)、语音语调的自然度 (重音、声调)、句子及语篇朗读/表达的流畅韵律感 (连读、弱读、节奏把控)，以及根据不同语境 (如日常交流、学术陈述、跨文化对话) 动态调整语音韵律的适应性。

3.6 语境理解与运用能力 (Context Comprehension and Application Ability)

学习者借助人工智能提供的多模态语境案例库（含口语对话、听力材料、阅读文本）及智能分析工具，在听说读写全场景中精准解析语境信息的能力。具体表现为：听力中捕捉隐含语境线索（如语气、停顿），口语中把握对话背景与交际目的，阅读中分析文本语境逻辑（如上下文关联、文化隐喻）；并能基于语境灵活选择词汇、语法结构及交流策略（口语回应或书面表达），有效完成信息传递与任务目标。

3.7 逻辑表达与创新思维能力 (Logical Expression and Innovative Thinking Ability)

学习者在人工智能启发下，针对不同语言任务（口语阐述、书面写作、观点辩论）的核心需求，围绕特定话题组织语言的能力。具体包括：口头或书面表达中逻辑结构的清晰性（如因果、对比、递进），信息传递的条理性；同时能结合听读输入（如听力材料观点、阅读文本素材）提出新颖见解，展现批判性思维与创造性表达的能力。

3.8 跨文化交际情感交流与文化感知能力 (Emotional Communication and Cultural Perception Ability)

学习者在语言实践中（口语互动、文本阅读、写作输出）综合运用语言与非语言要素传递与感知情感的能力，以及跨文化语境中的适应能力。具体表现为：通过语音语调（口语）、标点语气（写作）或文本情感倾向（阅读）传递情感信息；借助人工智能的文化背景解析工具，理解不同文化语境下的语言习惯（如礼貌表达、禁忌用语）与价值观念，并在口语输出或书面表达中体现文化敏感性（如调整措辞以适配目标文化）。

3.9 人工智能英语测评生态系统 (AI - Based Ecosystem for Standard English Assessment)

整合智能化线上考试平台、精准智能语音打分系统、动态更新的多模态测试题库、深度大数据分析系统、个性化学习推荐系统以及智能教学辅助系统等。各子系统相互协作，形成一个完整的生态闭环，为学习者提供全方位的口语学习和测评服务。

3.10 智慧考场 (Smart Assessment Room)

按照先进的计算机软硬件指标和智能化设施标准打造的考场。配备高清摄像头、麦克风阵列、智能互动屏幕等设备，实现多模态数据采集。同时，采用智能监考系统，运用人脸识别、行为分析等技术确保考试公平公正。

3.11 综合报告 (Comprehensive Assessment Report)

学习者完成线上英语能力人工智能考试后,将获得覆盖听力理解、口语表达、阅读理解、写作能力四大核心维度的详细数据统计与深入分析结果,具体包含语音韵律(口语)、语境分析(阅读)、逻辑连贯性(说写)、情感传递(说写)等多维度表现评估。

3.12 动态题库 (Dynamic Assessment Database)

根据人工智能英语测评的发展需求和学习者的能力变化,运用大数据分析和人工智能算法建设的题库。实时更新试题内容,涵盖不同难度层次、话题领域和文化背景的题目。

3.13 安全与隐私保障 (Security and Privacy Protection)

在测评过程中,采用先进的加密技术、访问控制技术和数据脱敏处理,确保考生的个人信息、考试数据和测评结果的安全。

4. 缩略语

下列缩略语适用于本文件。

CNSET 英语能力等级计算机测评 (Chivox National Standard English Test)

5. 英语能力人工智能分级测评等级划分

5.1 英语能力人工智能分级测评框架

英语能力评价围绕听力理解、口头表达、阅读理解、书面表达四大核心维度展开,各维度从“基础识别与输出”到“高阶批判与适配”分层构建,覆盖从小学到高端外语人才的全学段能力进阶,通过明确核心评估要素与分级指标,实现对考生英语能力的精准测评。

5.1.1 听力理解维度:

核心评估要素:

聚焦考生对语音信号的解码、信息提炼及语境洞察能力,包含三层递进逻辑:

- 语音识别: 精准捕捉发音细节(音素区分、重音位置)、语调模式(陈述降调、疑问升

调)、语篇韵律(意群停顿、逻辑重音),构建“单词-句子-语篇”的语音感知体系。

- 信息提取:从对话/文本中定位关键信息(数字、地点、人物关系),梳理时序、因果等逻辑细节,提炼主旨(对话主题、段落核心观点),实现“碎片信息-关联整合-全局概括”的信息处理。
- 语境推断:挖掘语言隐含语义(言外之意、文化隐喻),适配场景差异(日常/学术/跨文化),识别交际信号(礼貌表达、语气适配),达成“字面意义-场景关联-文化理解”的语境穿透。

分级评价逻辑:

- Level 1-2(小学阶段):从“识别基础词汇发音、提取显性关键词”,进阶到“辨别多音节重音、理解校园场景隐含语义”,聚焦语音与信息的基础感知。
- Level 3-6(中学/大学基础):拓展至“捕捉学术讲座韵律、梳理专业文本逻辑链”,强化信息深度提取与跨场景语境适配。
- Level 7-9(专业/高端人才):要求“识别口音变体、理解跨文化隐喻(如学术典故)”,实现对复杂语音信号的精准解码与多元语境的深度融合。

5.1.2 口头表达维度

核心评估要素:

围绕语音质量、语言组织与交际实效,构建“输出-结构-互动”的能力闭环:

- 语音输出质量:包括发音准确度(如元音/辅音标准性)、韵律掌控(如连读、弱读、节奏流畅度)及语调适配(如情感表达、语境语气调整)。
- 语言组织能力:涵盖词汇精准度(如专业术语使用、同义替换多样性)、语法正确性(如时态一致、从句嵌套)及逻辑连贯性(如观点分层、连接词衔接)。
- 交际有效性:评估话题契合度(如内容完整性、主题相关性)、互动适应性(如话轮转换、反馈回应)及文化得体性(如礼貌表达、跨文化语用规则)。

分级评价逻辑:

- Level 1-2(小学阶段):从“单词语音清晰、简单句交流”,进阶到“复合句运用、课堂互动参与”,夯实表达基础。
- Level 3-6(中学/大学基础):拓展至“学术语调适配、专业词汇运用”,强化复杂场景下的表达逻辑与文化得体。

- Level 7 - 9 (专业 / 高端人才): 要求 “口音自然度、跨文化异议回应”, 达成专业领域 (如学术研讨、国际谈判) 的高效交际。

5.1.3 阅读理解维度:

核心评估要素:

基于文本解码、信息处理与批判分析, 构建 “理解 - 提炼 - 反思” 的认知链条:

- 文本解码: 突破词汇语义 (多义词、熟词生义)、句法结构 (长难句拆分)、标点表意 (破折号解释、分号逻辑), 实现 “词汇 - 句法 - 标点” 的文本解构。
- 信息处理: 精准定位细节 (数据、事实核查), 推导隐含逻辑 (因果、隐含信息), 归纳主旨 (段落 / 篇章主题), 完成 “细节 - 推理 - 概括” 的信息挖掘。
- 文本批判: 审视逻辑严谨性 (论点有效性、论据充分性), 关联跨文本信息 (多源对比), 分析文化语境 (隐喻、背景影响), 达成 “逻辑 - 关联 - 文化” 的深度批判。

分级评价逻辑:

- Level 1 - 2 (小学阶段): 从 “识别基础词汇、提取显性细节”, 进阶到 “辨析多义词、推导简单隐含逻辑”, 构建基础阅读认知。
- Level 3 - 6 (中学 / 大学基础): 拓展至 “解析专业文本长难句、批判论据充分性”, 强化学术阅读的信息处理与逻辑批判。
- Level 7 - 9 (专业 / 高端人才): 要求 “解构跨文化隐喻、整合多源学术观点”, 实现对复杂文本 (如国际期刊论文) 的深度解码与学术批判。

5.1.4 书面表达维度:

核心评估要素:

围绕语言建构、内容组织与文体适配, 打造 “形式 - 逻辑 - 风格” 的写作能力体系

- 语言建构: 提升词汇丰富度 (学术词汇占比、搭配准确), 强化语法复杂度 (非谓语、复杂从句), 保障书写规范性 (标点、格式正确), 奠定 “词汇 - 语法 - 书写” 的语言基础。
- 内容组织: 明确主题 (论点清晰、主题句突出), 严密逻辑 (递进 / 对比结构、论证链条), 完整信息 (论据支撑、结论呼应), 搭建 “主题 - 逻辑 - 信息” 的内容框架。
- 文体适配: 区分文体差异 (学术写作 vs 日常书信), 优化修辞效果 (比喻、排比运用), 遵循跨文化表达 (引用规范、商务礼仪), 实现 “文体 - 修辞 - 文化” 的风

格适配。

分级评价逻辑：

- Level 1 - 2（小学阶段）：从“基础词汇书写、单段主题明确”，进阶到“复合句运用、文体初步区分”，夯实写作基础。
- Level 3 - 6（中学 / 大学基础）：拓展至“学术词汇运用、多维度论证逻辑”，强化复杂文本（如议论文、报告）的内容组织与文体适配。
- Level 7 - 9（专业 / 高端人才）：要求“跨文化文体精准适配、学术写作创新表达”，达成专业领域（如国际学术论文、高端商务文书）的写作标杆。

5.2 英语能力人工智能分级测评阶段划分

表1 英语能力等级阶段划分

等级		参考学段
Level 1	1C	学前、一、二年级
	1B	小学三年级
	1A	小学四年级
Level 2	2B	小学五年级
	2A	小学六年级
Level 3	3C	初中一年级
	3B	初中二年级
	3A	初中三年级
Level 4	4C	高中一年级/中职
	4B	高中二年级/中职
	4A	高中三年级/高职高专
Level 5	5B	大学英语(基础)
	5A	大学英语(发展)
Level 6	6	大学英语(提高)
Level 7	7B	英语专业
	7A	
Level 8	8B	高端外语人才
	8A	
Level 9	9B	

	9A	
--	----	--

5.3 英语能力人工智能分级测评评价框架

表2 英语能力人工智能描述及评价框架

等 级 (Level)	能 力 项 目 (Competence Item)
Level 1	<p>听: 能识别基础英文字母、300 个常见词汇 (如 “cat” “red”) 的标准发音, 听懂简单日常用语 (如 “Hello!” “Goodbye!”)、单句指令 (如 “Sit down.”), 关键信息捕捉准确率\geq85%。</p> <p>说: 能准确拼读基础字母和词汇, 使用简单主谓宾结构句子 (如 “I have a book.”) 进行自我介绍、物品描述等日常应答, 发音清晰, 语流较为连贯。</p> <p>读: 能辨认基础字母、词汇的印刷体, 读懂简单标语、菜单、便签等文字内容, 图文对应理解准确率\geq80%。</p> <p>写: 能规范书写 26 个英文字母, 正确拼写 300 个基础词汇, 书写简单短句 (如 “My name is Tom.”), 字母书写格式和拼写正确率\geq80%。</p>
Level 2	<p>听: 可识别 800 个常用词汇及简单短语 (如 “a cup of.”) 的发音, 听清日常对话、简单故事中的问答内容, 理解由 2 - 3 个分句组成的简单句子, 关键信息理解准确率\geq80%。</p> <p>说: 熟练拼读 800 个常用词汇及短语, 能用并列句 (如 “I like apples, and he likes bananas.”) 在日常交流中表达观点、描述经历, 语法基本正确, 表达较为流畅。</p> <p>读: 能识别 800 个常用词汇及短语, 理解日常便签、短故事、简单通知等文本, 准确提取其中的时间、地点、人物等关键信息, 段落主旨理解准确率\geq70%。</p> <p>写: 正确拼写 800 个常用词汇及短语, 书写包含 2 - 3 个分句的段落, 描述日常生活事件, 使用 “first” “then” 等基础连接词, 逻辑较为清晰, 语法错误率\leq12%。</p>
Level 3	<p>听: 能识别 1500 个日常词汇、短语及简单复合句的发音, 在校园活动、家庭聚会等场景对话中, 理解稍长的语音内容和事件过程描述, 对简单因果、并列等逻辑关系的语音表达理解准确率\geq85%。</p> <p>说: 准确拼读 1500 个日常词汇、短语及句子, 在模拟较复杂场景对话中, 清晰阐述个人见解, 介绍人物、活动等, 表达逻辑连贯, 语言流畅度较高。</p> <p>读: 识别 1500 个日常词汇、短语及句子, 读懂校园通知、故事短文、简易说明书等文本, 理解其中的人物关系、事件发展脉络, 要素提取完整率\geq85%, 复杂句理解准确率\geq80%。</p> <p>写: 正确拼写 1500 个日常词汇、短语及句子, 在模拟场景中完成文字交流, 书写个人及亲友详细信息、活动过程描述, 合理运用 “because” “so” 等连接词, 表意准确, 逻辑合理。</p>

Level 4	<p>听：识别 2500 个词汇、较长句子及语篇的发音，在生活场景对话中，理解人物关系描述、事件经历阐述及评价建议，对复杂语义的理解深度达标率$\geq 70\%$，关键信息遗漏率$\leq 10\%$。</p> <p>说：熟练朗读 2500 个词汇、句子及长语篇，运用连读、重读等语音技巧，在真实场景对话中围绕生活话题深入讨论、交换见解，复杂句使用率$\geq 30\%$，表达逻辑清晰。</p> <p>读：识别 2500 个词汇、句子及长语篇文字，掌握略读、扫读等阅读技巧，理解生活场景对话文本中的复杂情节、观点论述，主旨归纳准确率$\geq 80\%$，细节信息提取完整度$\geq 85\%$。</p> <p>写：正确拼写 2500 个词汇、句子及长语篇，运用分段、过渡句等写作技巧，在模拟真实场景文字对话中阐述观点、描述事件，内容有深度，书写规范。</p>
Level 5	<p>听：识别 3500 个词汇、复杂句子及语篇的发音，在复杂社会场景对话中，听清各类话题信息、观点讨论，捕捉表达修正、话题转换等细节，语义理解丰富度$\geq 85\%$，逻辑关系把握准确率$\geq 90\%$。</p> <p>说：精准朗读 3500 个词汇、复杂句子及语篇，根据不同语境调整发音风格，在复杂社会场景交流中深入探讨话题、灵活表达观点，使用 15% 以上的专业词汇，内容丰富，逻辑紧密。</p> <p>读：识别 3500 个词汇、复杂句子及语篇文字，流畅阅读复杂社会场景文本，理解其中的隐含信息、复杂逻辑关系，结合文化背景知识理解文本内容，准确率$\geq 85\%$。</p> <p>写：正确拼写 3500 个词汇、复杂句子及语篇，在复杂社会场景文字交流中发表见解、展开论述，内容详实，逻辑严谨，书写规范且富有文采。</p>
Level 6	<p>听：识别各类句子、专业语篇的发音，在多元文化场景对话中，听清复杂话题讨论、专业见解发表，捕捉文化交流要点、特殊语言结构，理解内容条理清晰，能从多角度深入理解。</p> <p>说：清晰、有韵律地朗读各类句子、专业语篇，在多元文化场景对话中参与复杂话题讨论，发表深刻见解，灵活运用丰富词汇和多样句式，逻辑严密，跨文化表达恰当。</p> <p>读：识别各类句子、专业语篇文字，阅读精准且有节奏，在多元文化场景文本中理解复杂观点论述、文化隐喻，能结合跨文化知识进行批判性阅读。</p> <p>写：书写各类句子、专业语篇，在多元文化场景文字交流中深入探讨复杂话题、发表独到见解，文章逻辑紧密，词汇句式丰富多样，书写规范且具文学性。</p>
Level 7	<p>听：识别各类复杂句子、高难度专业语篇的发音，在学术讲座、专业研讨等场景对话中，听清复杂话题主导发言、深度交流内容、专业见解发表，捕捉精准语言表达和清晰逻辑思路，理解复杂学术概念，把握论证逻辑。</p> <p>说：朗读各类复杂句子、高难度专业语篇，语音纯正，语调富有感染力，在高难度学术场景对话中主导话题讨论、深入交流观点、系统发表见解，专业术语运用准确，逻辑严谨，论证充分。</p> <p>读：识别各类复杂句子、高难度专业语篇文字，阅读流畅且逻辑清晰，掌握精读、分析等高级阅读技巧，理解高难度学术文本中的复杂理论阐述、多层论证结构。</p> <p>写：书写各类复杂句子、高难度专业语篇，在学术场景文字交流中主导话题研讨、深入分析问题、系统阐述观点，文章逻辑严谨，论证充分，专业表达规范，符合学术写作要求。</p>
Level 8	<p>听：识别各类复杂专业语篇的发音，在跨领域国际会议、高端商务谈判等场景对话中，听清复杂</p>

	<p>话题主导讨论、高效交流内容、观点总结陈述，捕捉丰富流畅的专业表达和跨文化交际特色，理解前沿专业观点，把握逻辑严密性。</p> <p>说：完美朗读各类复杂专业语篇，语音语调精准，接近母语者水平，在跨领域高端场景对话中主导或深度参与复杂话题讨论，高效交流观点、总结核心内容，语言表达丰富流畅，跨文化交际能力出色，专业表达具权威性。</p> <p>读：识别各类复杂专业语篇文字，阅读精准且富有韵律，能根据语境灵活调整理解方式，理解跨领域高端文本中的复杂专业论述、行业前沿观点，综合运用专业知识。</p> <p>写：书写各类复杂专业语篇，在跨领域高端场景文字交流中主导话题研讨、深度分析问题、系统总结观点，文章逻辑严密，语言精准，提出创新性专业见解，符合国际专业文档写作标准。</p>
Level 9	<p>听：精准识别各类复杂专业语篇（如学术专著、国际条约解读）的发音，在国际学术论坛、顶尖外交磋商等场景对话中，听清复杂话题深度探讨、多轮专业辩论、权威观点阐释，捕捉跨文化语用细微特征，理解复杂学术理论和创新方案，还原逻辑体系。</p> <p>说：以母语者专家级水平朗读各类复杂专业语篇，语音语调纯正自然，在国际顶尖学术、外交场景中主导核心话题研讨，系统阐述权威观点，提出创新性专业方案，语言表达极具说服力，跨文化交际能力卓越。</p> <p>读：精准理解各类复杂专业语篇（如 SCI 论文、国际组织政策文件）文字，阅读能力达到母语资深专家水准，深度剖析复杂专业论述、前沿研究成果，准确判断学术观点创新性，构建完整知识体系。</p> <p>写：撰写各类复杂专业语篇，在国际顶尖学术、外交场景文字交流中产出具有权威性、指导性的专业文档，文章逻辑严密，论证充分，表达精准，提出的创新性专业方案具有国际影响力，符合国际顶级学术期刊发表要求。</p>

5.4 英语能力人工智能分级测评分项能力评价

5.4.1 听力 listening

听力能力是指考生通过听觉精准识别英语音素与音节的发音特征，判断单词及句子发音是否符合标准规范（如连读、弱读等语音现象），并在此基础上理解语义、快速抓取关键信息的综合语言技能。其核心要求考生通过听觉对语音信号进行解码，同时结合语境逻辑完成信息提取，避免因发音辨识偏差导致理解障碍。

具体体现为：

- 发音辨识能力：准确分辨元音、辅音等音素特征，识别连读、弱读、重音等语音现象，判断发音是否符合英语标准规范；
- 语义理解能力：在准确辨识发音的基础上，理解单词、句子及语篇的语义内容；
- 信息提取能力：从语音信号中快速捕捉关键信息（如数字、地点、人物关系等），梳理语境逻辑，形成完整的语义理解，确保信息提取的准确性和完整性。

表3 听力能力描述

能力等级	能力描述
Level 1	能准确识别 26 个英文字母及 300 个核心基础词汇（如 “cat” “book” “red” “one”）的国际音标（IPA）发音，清晰分辨 “Good morning.” “This is a pencil.” 等单句型日常场景语音。对单元音（如 /i:/ 与 /ɪ/）、辅音（如 /b/ 与 /p/）的识别准确率达 92% 以上，对易混淆音素（如 /v/ 与 /w/）的区分准确率≥85%。能捕捉 “Open the door.” “I like apples.” 等单句核心语义，满足小学 1-4 年级基础听力需求，误听率≤18%。通过与教育部《英语语音标准数据库》对比，生成包含音素偏差坐标图的《基础发音评估报告》，标注舌位、爆破音等具体偏差。
Level 2	可识别 800 个常用词汇（如 “banana” “family” “computer” “happy”）及 “a cup of” “look at” 等简单短语发音，精准捕捉 “not at all” 连读（/no tæ t l/）、“he is” 弱读（/hi z/）等基础语音变现象。在课堂互动（如 “Open your book and take out a pencil.”）、家庭对话（如 “Have you eaten breakfast?”）等场景中，理解含 2-3 个分句的简单对话，对陈述句降调 / 疑问句升调的语调识别准确率达 88%。生成包含语流切分标记的《日常语流理解分析报告》，信息遗漏率≤15%，符合小学 5-6 年级《英语课程标准》要求。
Level 3	精准识别 1500 个日常词汇及 “in order to” “because of” “turn on” 等常用短语发音，熟练解析 “tomato” 重音位置（/t me to /）、“He is reading” 中 “is” 弱读（/ z/→/s/）等发音技巧。在校园活动（如 “Our class won the basketball game.”）、社区交流（如 “Where is the nearest post office?”）等场景中，理解含时间状语从句的复合句（如 “When I get home, I do my homework.”），对意群停顿的识别准确率达 82%。生成带重音热力图的《语篇韵律评估报告》，关键信息捕捉率≥85%，满足初中阶段（7-9 年级）听力需求。
Level 4	在超市购物（如 “Can I have two kilos of apples?”）、旅行规划（如 “We will go to the park next Saturday.”）等日常场景中，准确识别 2500 个词汇及 5-8 句连贯语篇的发音，精准处理 “yesterday” “school” 等关键词重读和逗号 / 句号处的意群停顿。理解含 3-5 个分句的叙述性语篇（如 “Last weekend, I went to the zoo and saw many animals, including pandas and tigers.”），通过语调变化区分句子重点，对话篇逻辑重音的识别准确率达 78%。生成包含语义框架标注的《日常语篇听力理解报告》，逻辑关系误判率≤12%，符合高中阶段（10-12 年级）要求。
Level 5	准确识别 3500 个词汇及复杂句式（如含定语从句的 “The girl who wears a red hat is my sister.”）的发音，精准定位 “Although it rains, we still go out.” 等句子的逻辑重读位置（“rains” “still”），熟练解析连读（如 “would you” →/w d u/）、同化（如 “nice to meet you” →/na s t mi t ju/）、失去爆破（如 “not bad” →/no t bæd/）等进阶技巧。在商务预约（如 “May I make an appointment for 3 pm tomorrow?”）、文化介绍（如 “Spring Festival is a traditional Chinese holiday.”）等场景中，理解符合母语者习惯的语流，生成带语音特征参

	数表的《复杂语篇发音技巧评估报告》，语音特征识别准确率 $\geq 90\%$ ，满足大学基础及发展阶段需求。
Level 6	精准识别 3000 + 核心词汇及 500 + 科技 / 文化领域基础术语（如 “artificial intelligence” “cultural heritage”）的发音，根据语篇逻辑（如因果论证的语调变化）动态调整理解策略。在高考听力材料、学术讲座摘要（如 “Science Fair Introduction”）等场景中，理解含复合句的论述文，生成包含关键词标注的《学术语篇听力分析报告》，专业术语识别准确率 $\geq 75\%$ ，符合大学提高阶段需求。
Level 7	在医学（如 “diagnosis” “pharmaceutics”）、法律（如 “legislation” “jurisdiction”）等专业领域及国际论坛（如 “World Health Assembly”）场景中，精准识别复杂语篇发音，细腻把握学术报告逻辑重音（如论点关键词重读）、商务谈判委婉语气（如 “May I suggest an alternative approach?” 的升调）等语用特征。运用连读、停顿等技巧解析语篇韵律（如段落间 2 秒停顿），生成带语用策略标注的《跨领域听力能力评估报告》，跨领域语义理解准确率 $\geq 92\%$ ，满足英语专业阶段需求。
Level 8	通过跨文化语音处理平台在国际会议（如 “UN Climate Conference”）、学术研讨（如 “Global Economics Forum”）等场景中，精准识别英音 / 美音 / 澳音等多口音复杂语篇发音，根据语境动态调整重读（如提案陈述中 “key points” 重读）、弱读技巧（如功能词弱读）。通过语调变化解析隐喻（如 “a Pandora’s box” 的降调暗示负面含义）和文化梗（如 “Alice in Wonderland” 的联想语义），生成包含跨文化语用规则的《跨文化语用听力评估报告》，口音适应准确率 $\geq 95\%$ ，满足高端外语人才基础层需求。
Level 9	借助全维度语义监控系统精准识别学术报告（如 “Nature 期刊论文朗读”）、国际协议（如 “WTO trade agreement”）等专业文档中密集术语（如 “global governance” “sustainable development” “quantum entanglement”）和复杂句式（如含多重定语从句的长难句）发音。熟练运用重音、弱读、停顿等技巧解析外交辞令隐含语义（如 “With respect to your proposal...” 的委婉否定），在高端学术论坛（如 “Nobel Prize Symposium”）、国际商务洽谈（如 “Cross-border M&A negotiation”）等场景中，生成包含论证逻辑树的《专业级听力能力评估报告》。专业语义深度理解准确率 $\geq 98\%$ ，达到高端外语人才（同传译员及专家）水平。

5.4.2 口语 speaking

口语能力是指考生在英语表达中展现的语流流畅度与语音韵律掌控能力，核心在于通过协调发音技巧与交际需求，实现自然连贯且富有节奏感的语言输出，避免因停顿、韵律偏差导致的沟通障碍。

具体体现为：

- **流畅性表现能力：**自然流利地输出内容，减少停顿、重复或卡顿，连贯阐述观点而不断。

运用逻辑连接词衔接句子，确保表达的连续性和完整性。

● 语音韵律掌控能力：把握重音、连读、弱读等语音技巧，如通过重音突出关键信息、正确处理 “not at all” 等连读现象。

通过语调变化区分陈述与疑问（如疑问句升调），使表达兼具节奏感与交际功能。

● 交际有效性实现能力：根据场景调整表达风格（如日常对话 / 学术演讲），确保语用得体的；通过语音韵律传递隐含语义（如委婉建议的语调），提升信息传达效果。

表4 口语能力描述

能力等级	能力描述
Level 1	能准确拼读 26 个英文字母及 300 个核心基础词汇（如 “cat” “book” “red” “one” “apple”），模仿 “Good morning.” “Open the door.” 等单句指令发音，单元音（如 /i:/ 与 / /）、辅音（如 /b/ 与 /p/）发音准确率≥92%，易混淆音素（如 /v/ 与 /w/）区分率≥85%。能用 “yes/no” 回答基础提问，用 “I like...” “This is...” 等 5 词以内简单句表达，语流卡顿≤3 次 / 分钟，满足小学 1-4 年级基础需求，语法错误率≤20%。通过与教育部《英语语音标准数据库》对比，生成包含发音波形图、舌位坐标的《基础语音评估报告》，标注 /t/ 爆破音缺失、/æ/ 口型偏差等具体问题。
Level 2	可拼读 800 个常用词汇（如 “banana” “family” “computer” “happy” “teacher”）及 “a cup of” “look at” 等短语，准确输出 “not at all” 连读（/no tæ t l/）、“he is” 弱读（/hi z/）等语流音变，语流自然度≥80%。能使用 “I want to...” “She has...” 等简单陈述句描述日常活动，用 “and” “but” 连接 2-3 个分句（如 “I like apples, but he likes bananas.”），逻辑连接词使用率≥30%，满足小学 5-6 年级需求，信息完整度≥85%。生成带语流切分标记（如 “Have you eaten breakfast?”）的《日常表达流畅度报告》，标注连读缺失、重音错位等问题，语法错误率≤15%。
Level 3	精准拼读 1500 个日常词汇及 “in order to” “because of” “turn on” 等短语，熟练处理 “tomato” 重音（/t me to /）、“is” 弱读（/ z/→/s/）等发音技巧，语音特征识别率≥85%。能运用 “When I get home, I do my homework.” 等时间状语从句叙述事件，使用 “first” “then” 等顺序词组织信息，逻辑链条完整率≥80%。在校园场景中主动发起对话（如 “Where is the basketball court?” “Can I join you?”），满足初中阶段需求，语用错误（如称呼不当）≤1 处 / 分钟。生成带重音热力图的《语篇韵律评估报告》，标注逻辑重音偏移问题。
Level 4	准确拼读 2500 个词汇及 “would you like” “how about” 等复杂短语，精准输出连读（如 “would you” →/w d u/）、失去爆破（如 “not bad” →/no t bæd/）等技巧，语音技巧正确率≥85%。能使用 “The girl who wears red is my friend.” 等定语从句论述观点，插入 “for example” “in fact” 等衔接词，专业术语（如 “environment” “technology”）使用率≥10%。在超市购物（如 “Can I have two kilos of apples?”）、旅行规划等场景中完成 5-8 句连贯交流，满足高中阶段需求，逻辑断层率≤10%。生成包含语义框架标注的《日常交际能力报告》。

Level 5	<p>熟练拼读 3500 个词汇及学术短语解析 “Although it rains, we still go out.” 等复杂句式逻辑重音定位准确率$\geq 90\%$。能运用让步状语从句 + 定语从句(如 “Although he is young, the boy who studies hard can achieve success.”) 进行主题演讲，使用 “furthermore” “nevertheless” 等高阶连接词，专业术语使用率$\geq 20\%$（如 “innovation” “sustainability”）。在商务预约（如 “May I make an appointment for 3 pm tomorrow?”）、文化介绍场景中完成跨场景表达，生成带语音特征参数（基频、语速）的《复杂表达评估报告》，跨场景适配率$\geq 85\%$。</p>
Level 6	<p>精准拼读 3000 + 核心词汇及 500 + 科技 / 文化术语（如 “artificial intelligence” “quantum computing” “renaissance” “cultural heritage”），根据语篇逻辑动态调整重读策略（如论点关键词 “key points” 重读），术语发音准确率$\geq 88\%$。能就 “Environmental Protection” 等学术话题进行 5 分钟连贯论述，使用 “hypothesis” “conclusion” 等学术词汇，采用 “问题 - 对策” 结构（如 “First, we analyze the problem... Then, we propose solutions...”），生成含逻辑框架标注的《学术口语分析报告》，论证充分度$\geq 80\%$，符合大学公共英语要求。</p>
Level 7	<p>熟练掌握 5000 + 专业词汇（如医学 “diagnosis” “pharmaceutics”、法律 “legislation” “jurisdiction”），精准处理专业领域委婉表达（如 “May I suggest an alternative approach?” 的升调暗示建议），语用特征识别率$\geq 92\%$。能在国际论坛模拟场景中主导话题讨论，使用 “in accordance with” “pursuant to” 等法律术语或 “pharmacodynamics” “pathogenesis” 等医学术语，专业术语使用率$\geq 40\%$。生成带语用策略标注（如 “委婉拒绝策略: I appreciate your idea, but...”）的《跨领域口语评估报告》，跨文化表达得体率$\geq 90\%$，满足英语专业阶段需求。</p>
Level 8	<p>精通 8000 + 跨领域术语，准确识别英音 / 美音 / 澳音等多口音发音特征（如英音 /r/ 不发音、澳音 /æ/ 变音），口音适应准确率$\geq 95\%$。能解析 “a Pandora’s box”（降调暗示负面）、“Alice in Wonderland”（升调引发联想）等文化隐喻的语调暗示，在 UN Climate Conference 等国际会议中完成议题辩论，使用 “carbon footprint” “renewable energy” 等前沿术语，占比$\geq 15\%$。生成包含跨文化语用规则（如西方直接表达 vs 东方委婉表达）的《同传预备报告》，信息完整度$\geq 90\%$，满足高端外语人才基础层需求。</p>
Level 9	<p>掌握 10000 + 前沿术语（如 “global governance” “sustainable development” “quantum entanglement”），精准运用重音、停顿解析外交辞令隐含语义（如 “With respect to your proposal...” 的长停顿暗示保留意见），专业语义传达准确率$\geq 98\%$。能在 Nobel Prize Symposium 等高端论坛中进行主旨发言，同步处理多模态信息（如结合手势、PPT 术语），使用 “epistemology” “ontological” 等哲学学术词。生成带论证逻辑树（论点 - 论据 - 结论可视化）的《同传级评估报告》，达到国际同传译员标准，实现 “语义 - 语用 - 文化” 三重精准传达。</p>

5.4.3 阅读 reading

阅读能力是指考生对英文文本的信息解码与理解深度的综合技能，核心在于通过协调词汇识别、句式解析与文本韵律感知，实现快速准确的语义提取，避免因文本复杂度导致的理解断层。

具体体现为：

● **流畅性表现能力：**快速连贯地识别文本信息，减少因词汇障碍或长难句导致的停顿、回读，实现语义自然衔接（如连续阅读 5-8 句不中断）。

运用逻辑连接词（如 “however” “therefore”）梳理上下文关联，确保段落内信息的连续性和完整性。

● **文本韵律掌控能力：**把握文本逻辑重音，通过转折词（如 “but” “although”）定位语义重点，通过强调词（如 “key” “critical”）突出关键信息。

解析篇章节奏，通过段落首句 / 尾句快速定位主旨，借助标点符号（破折号、分号）梳理信息层次，使阅读兼具速度与理解有效性。

● **理解深度实现能力：**根据场景需求调整阅读策略（如浏览新闻时快速提取要点，精读学术文献时分析论证逻辑），确保理解适配性；通过文本韵律特征（如段落间隔、标点停顿）传递隐含语义（如讽刺语气的文本标记），提升信息解读的准确性。

表5 阅读能力描述

能力等级	能力描述
Level 1	能精准识别 26 个字母及 300 个核心基础词汇（如 “cat” “book” “red” “one” “apple”），理解单句型文本（如 “Good morning.” “Open the door.”），识别句号 / 问号等基础标点的表意功能，单词形式匹配准确率 $\geq 92\%$ 。能将 “apple” 等词汇与实物图片对应，提取 “Close the window.” 等指令句的核心动作语义，满足小学低年级（1-2 年级）基础阅读需求，单句语义误读率 $\leq 18\%$ 。通过与教育部《基础英语词库》对比，生成包含词汇识别偏差坐标图的《初级阅读能力评估报告》，标注 “b/p” 形义混淆、“red” 拼写识别错误等具体问题。
Level 2	可识别 800 个常用词汇（如 “banana” “family” “computer” “happy” “teacher”）及 “a cup of” “look at” 等简单短语，理解含 2-3 个分句的文本（如 “I like apples, and he likes bananas. Have you eaten?”），通过句号 / 问号区分陈述 / 疑问功能，信息完整提取率 $\geq 85\%$ 。能从 “Today is Monday. I go to school.” 中提取时间 / 事件等显性信息，生成带语义切分标记（如 “时间→事件”）的《日常阅读流畅度分析报告》，标注 “family” 词义误解、短语搭配错误等问题，符合小学中年级（3-4 年级）《英语课程标准》要求。
Level 3	精准识别 1500 个日常词汇及 “in order to” “because of” “turn on” 等常用短语，解析含时间状语从句的复合句（如 “When I get home, I do my homework.”），通过逗号 / 句号定位意群停顿逻辑，关键信息（时间、地点、人物）捕捉率 $\geq 85\%$ 。能归纳 “Our class won the basketball

	game.”等校园活动段落的核心事件，生成带文本重音标注的《基础语篇理解报告》，分析“game”等关键词的语义权重，满足小学高年级（5-6 年级）阅读需求。
Level 4	在超市购物清单(如“two kilos of apples.”)、旅行计划(如“Go to the park next Saturday.”)等日常文本中，准确识别 2500 个词汇及 5-8 句连贯语篇，通过破折号 / 括号解析补充信息（如“Last weekend—I went to the zoo—saw pandas”），逻辑重音词（如“yesterday”“zoo”）识别率 $\geq 78\%$ 。能推导隐含信息（如“umbrella missing \rightarrow rained yesterday”），生成包含“时间 - 事件 - 因果”语义框架的《日常语篇阅读报告》，逻辑关系误判率 $\leq 12\%$ ，符合初中阶段（7-9 年级）要求。
Level 5	准确识别 3500 个词汇及含定语从句的复杂句式（如“The girl who wears a red hat is my sister.”），解析“Although it rains, we still go out.”的让步逻辑关系，通过冒号 / 引号划分论点（如“Key points: 1. Rainy day 2. Outdoor activity”），专业术语（如“environment”“technology”）识别率 $\geq 10\%$ 。能对比“News report”与“Personal diary”对同一事件的不同描述，生成带观点对比矩阵的《复杂文本阅读分析报告》，满足高中阶段（10-12 年级）需求。
Level 6	精准识别 3000 + 核心词汇及 500 + 科技 / 文化领域基础术语（如“artificial intelligence”“quantum computing”“renaissance”“cultural heritage”），理解“Science Fair Introduction”等学术摘要文本，通过分号 / 句号划分“问题 - 假设 - 方法”论证层次，术语语境理解准确率 $\geq 75\%$ 。能批判“Data sample size: 10 people”等论据的充分性，生成包含“论点 - 论据 - 漏洞”标注的《学术阅读批判性报告》，符合大学基础阶段（公共英语）需求。
Level 7	熟练掌握 5000 + 专业词汇（如医学“diagnosis”“pharmaceutics”、法律“legislation”“jurisdiction”），解析“World Health Assembly”等国际论坛报告，通过破折号 / 斜体识别术语定义（如“Vaccination—process of administering vaccines”），跨领域语义（如医学术语在公共政策文本中的引申义）整合率 $\geq 92\%$ 。能构建“Symptom \rightarrow Diagnosis \rightarrow Treatment”等术语语义网络，生成带概念关联图谱的《专业领域阅读评估报告》，满足英语专业阶段需求。
Level 8	精通 8000 + 跨领域术语，解析英音 / 美音等多口音学术文献（如《Nature》不同版本论文），通过方括号 / 脚注处理文化注释（如“[British English: autumn vs American English: fall]”），“a Pandora’s box”等文化隐喻的理解准确率 $\geq 95\%$ 。能对比中西节日文本（如“Spring Festival vs Christmas”）的价值观差异，生成包含“文化维度 - 语义偏差”分析的《跨文化阅读评估报告》，满足高端外语人才基础层需求。
Level 9	掌握 10000 + 前沿术语（如“global governance”“sustainable development”“quantum entanglement”），解析诺贝尔奖级学术文献（如《Nature》量子物理论文），通过多重标点（如括号内嵌套破折号）处理复杂逻辑（如“Entanglement—(a quantum phenomenon)—affects particle behavior”），专业语义深度理解率 $\geq 98\%$ 。能生成包含“假设 - 实验 - 结论”的论证逻辑树，标注“hypothesis validation”等学术推理链条，达到《翻译专业能力标准》同传译员及专家级阅读水平。

5.4.4 写作 writing

写作能力是指考生在英文文本创作中展现的逻辑连贯性与语言规范性的综合技能，核心在于通过协调段落组织、语法应用与韵律表达，实现结构严谨且富有节奏的内容输出，避免因逻辑断层或语法错误导致的交际障碍。

具体体现为：

● 流畅性表现能力：

自然组织英文段落，避免句式断裂或内容跳跃，通过过渡句（如“Furthermore”“On the contrary”）衔接观点，实现段落间逻辑连贯（如连续 3-5 段论述不中断）。

运用逻辑连接词（如“however”“therefore”）梳理上下文关系，确保段落内信息的连续性和完整性。

● 语言韵律掌控能力：

精准运用语法结构，通过复合句（如时间状语从句、同位语从句）丰富表达层次，避免简单句堆砌（如每段至少 1 个复杂句）。

规范使用标点符号（如破折号表解释、分号表并列），通过标题加粗、关键词斜体等方式突出核心信息，使文本兼具逻辑节奏与视觉韵律。

● 交际有效性实现能力：

根据场景需求调整写作风格（如日常书信的亲切语气、学术论文的正式表达），确保语用得得体；通过段落结构（如总分总框架）和韵律特征（如主题句前置、结论句强调）传递隐含语义（如通过对比结构暗示观点倾向），提升信息传达效果。

表6 写作能力描述

能力等级	能力描述
Level 1	能规范书写 26 个字母及 300 个核心基础词汇（如“cat”“book”“red”“one”“apple”），组成 5 词以内简单句（如“I like apples.”“This is a red book.”），单词拼写准确率 $\geq 92\%$ ，大小写 / 标点（句号 / 问号）使用正确率 $\geq 85\%$ 。能仿写“Open the door.”等单句型指令，完成图文匹配书写（如绘制苹果图并标注“apple”），满足小学低年级（1-2 年级）需求，语法错误率 $\leq 20\%$ 。通过与教育部《基础英语书写标准库》对比，生成包含字母笔顺偏差、单词间距错误的《初级写作评估报告》，标注“b/d”书写混淆、“red”拼写缺漏等问题。
Level 2	可书写 800 个常用词汇（如“banana”“family”“computer”“happy”“teacher”）及“a cup of”“look at”等简单短语，组成含“and/but”的 2-3 分句复合句（如“I like apples, but he likes bananas. Have you eaten?”），逻辑连接词使用率 $\geq 30\%$ ，信息完整度 $\geq 85\%$ 。能撰写 5-8 句日常活动日记（如“Today is Monday. I went to school. I had math class.”），使用时间词（“today”“yesterday”）排序事件，符合小学中年级（3-4 年级）要求。生成带语义切分标记

	<p>(如“时间→事件→感受”)的《日常写作流畅度报告》,标注“family”拼写错误、时态混淆等问题。</p>
Level 3	<p>精准书写 1500 个日常词汇及 “in order to”“because of”“turn on” 等常用短语,构建含时间状语从句的复合句(如 “When I get home, I do my homework.”),段落首句主旨句准确率≥85%。能撰写 10-15 句校园活动记叙文(如 “Our class won the basketball game. Everyone was excited. We celebrated after school.”),使用 “first”“then”“finally” 等顺序词组织事件,满足小学高年级 (5-6 年级) 需求。生成带逻辑重音标注的《语篇写作报告》,分析 “game”“excited” 等关键词的语义权重,标点 (逗号 / 句号) 使用正确率≥82%。</p>
Level 4	<p>在超市购物清单 (如 “Buy two kilos of apples, three bananas”)、旅行计划 (如 “Next Saturday, we will go to the park. Bring snacks and a blanket.”) 等日常文本中,准确书写 2500 个词汇及 5-8 句连贯语篇,使用破折号 / 括号补充信息 (如 “Last weekend—I went to the zoo—saw pandas, tigers and elephants”),逻辑链条完整率≥78%。能推导隐含信息并写入文本 (如 “Forgot umbrella→Rained yesterday”),符合初中阶段 (7-9 年级) 要求。生成包含 “时间 - 事件 - 因果” 语义框架的《日常写作报告》,逻辑关系误判率≤12%,段落衔接词 (如 “meanwhile”“furthermore”) 使用率≥20%。</p>
Level 5	<p>熟练书写 3500 个词汇及含定语从句的复杂句式 (如 “The girl who wears a red hat is my sister.”),解析 “Although it rains, we still go out.” 的让步逻辑并应用于写作,专业术语 (如 “environment”“technology”) 使用率≥10%。能撰写对比类短文 (如 “News Report vs Personal Diary: News is factual, while diary is emotional.”),使用 “on the contrary”“similarly” 等对比连接词,满足高中阶段 (10-12 年级) 需求。生成带观点对比矩阵的《复杂写作报告》,论证结构 (论点 - 论据) 完整率≥85%,标点 (冒号 / 引号) 功能准确率≥80%。</p>
Level 6	<p>精准书写 3000 + 核心词汇及 500 + 科技 / 文化领域基础术语 (如 “artificial intelligence”“quantum computing”“renaissance”“cultural heritage”),遵循学术摘要结构 (背景 - 方法 - 结论) 撰写 200 词短文 (如 “Science Fair Introduction: This project explores... Methods include... Results show...”),通过分号划分论证层次,术语使用准确率≥75%。能批判论据充分性 (如 “Sample size too small→results unreliable”),符合大学基础阶段 (公共英语) 需求。生成包含 “论点 - 论据 - 漏洞” 标注的《学术写作批判性报告》,学术连接词 (如 “hypothesis”“conclusion”) 使用率≥30%。</p>
Level 7	<p>熟练掌握 5000 + 专业词汇 (如医学 “diagnosis”“pharmaceutics”、法律 “legislation”“jurisdiction”),遵循国际论坛报告结构 (现状 - 挑战 - 建议) 撰写 300 词分析文 (如 “World Health Assembly Report: Current pandemic status... Challenges include... Proposals are...”),通过破折号定义术语 (如 “Vaccination—process of administering vaccines”),跨领域语义整合率≥92%。能构建 “Symptom→Diagnosis→</p>

	Treatment”等术语语义网络，满足英语专业阶段需求。生成带概念关联图谱的《专业写作评估报告》，专业论证逻辑（如医学因果推理）准确率 $\geq 88\%$ 。
Level 8	精通 8000 + 跨领域术语，遵循多口音学术文献格式（如英 / 美拼写差异标注）撰写 500 词对比文（如 “Spring Festival vs Christmas: Cultural meanings differ—Spring Festival emphasizes family reunion, while Christmas focuses on religious celebration”），通过脚注处理文化注释（如 “[British English: autumn vs American English: fall]”），隐喻表达（如 “a Pandora’s box”）准确率 $\geq 95\%$ 。能对比中西写作风格差异（如东方委婉 vs 西方直接），满足高端外语人才基础层需求。生成包含 “文化维度 - 语义偏差” 分析的《跨文化写作评估报告》，跨文化表达得体率 $\geq 90\%$ 。
Level 9	掌握 10000 + 前沿术语（如 “global governance” “sustainable development” “quantum entanglement”），遵循诺贝尔奖级文献结构（假设 - 实验 - 结论）撰写 800 词研究论文（如 “Quantum Entanglement: Hypothesis—particles affect each other; Experiments—Bell test; Conclusion—non-local realism”），通过多重标点（括号内嵌套破折号）处理复杂逻辑（如 “Entanglement—(a quantum phenomenon)—affects particle behavior regardless of distance”），专业语义深度表达率 $\geq 98\%$ 。能生成包含 “假设 - 验证 - 推论” 的论证逻辑树，达到《翻译专业能力标准》同传译员及专家级写作水平。生成带三维论证模型的《同传级写作评估报告》，学术创新点识别率 $\geq 95\%$ 。

6. 英语能力人工智能分级测评题型

人工智能英语测试借助先进技术，能全面考查考生的听说读写能力。以下从题型、语料选用、考查点、评分维度等方面介绍常见测试题型：

6.1 听力能力

6.1.1 基础识别类

6.1.1.1 题型

包括图片选择、词语匹配、正误判断等形式。图片选择题要求考生根据听到的内容从提供的图片中选择正确的一项；词语匹配题要求考生将听到的词语与对应的解释或图片进行匹配；正误判断题要求考生判断听到的内容与题目描述是否一致。

6.1.1.2 语料选用

语料以简单、常见的词汇和表达方式为主，包括日常生活中常用的名词、动词、形容词

等；句子多为简单句，结构清晰，易于理解；话题涉及个人信息、家庭、日常用品、熟悉的环境等贴近生活的内容。

6.1.1.3 考查点

重点考查考生对基本语言信息的识别能力，包括对单词、短语和简单句子的理解；考查考生辨别细节差异的能力，如不同事物、动作、状态的区分；同时也考查考生对基本语言形式的感知能力。

6.1.1.4 评分维度

识别准确性：评估考生对听到的信息与题目选项的匹配准确程度；

细节辨别能力：考查考生对相似信息或细微差别的区分能力；

反应速度：根据考生完成题目所需的平均时间判断其对基础信息的反应快慢。

6.1.2 理解归纳类

6.1.2.1 题型

包括信息填空、要点选择、内容排序、主旨归纳等题型。信息填空题要求考生根据听到的内容填写空缺的信息；要点选择题要求考生从多个选项中选出听到内容的关键要点；内容排序题要求考生根据听到的内容将事件或步骤按正确顺序排列；主旨归纳题要求考生概括听到内容的主要意思。

6.1.2.2 语料选用

语料涵盖简短对话、故事、通知、讲座片段等多种形式；话题范围扩展到学校生活、工作场景、社会现象、科普知识等；句子结构包括简单句和复合句，适当出现一些常见的语法结构和表达方式；词汇既包括基础词汇，也涉及一些常用的主题词汇。

6.1.2.3 考查点

考查考生对听到内容的基本理解能力，包括对句子意思和语篇主要内容的把握；考查考生归纳和提取关键信息的能力，如时间、地点、人物、事件等；同时考查考生对语篇结构和逻辑顺序的理解能力。

6.1.2.4 评分维度

理解准确性：评估考生对听到内容的理解是否准确；

信息完整性：考查考生获取的信息是否完整涵盖关键要点；

归纳合理性：判断考生对信息的归纳是否恰当、简洁；

逻辑连贯性：针对内容排序题，考查考生对事件发展顺序或逻辑关系的把握程度。

6.1.3 分析推理类

6.1.3.1 题型

包括细节问答、关系判断、意图推断、逻辑分析等题型。细节问答题要求考生根据听到的内容回答关于细节信息的问题；关系判断题要求考生判断听到内容中人物、事物之间的关系；意图推断题要求考生推断说话者的言外之意或真实意图；逻辑分析题要求考生分析听到内容的逻辑结构或因果关系。

6.1.3.2 语料选用

语料形式多样，包括对话、访谈、新闻报道、讲座等；话题涉及社会生活、文化习俗、科技发展、职场交流等多个领域；语料中包含一定数量的复杂句和多样化的语法结构；词汇不仅有常用词汇，还包括一些与特定话题相关的专业词汇和表达。

6.1.3.3 考查点

考查考生对细节信息的捕捉和理解能力；考查考生分析语篇中各种逻辑关系的能力，如因果关系、转折关系、递进关系等；考查考生根据听到的信息进行推理和判断的能力，包括对说话者态度、意图、情感的推断；同时考查考生对较为复杂语篇结构的理解能力。

6.1.3.4 评分维度

细节准确性：评估考生对细节信息的理解和回答的准确程度；

逻辑分析能力：考查考生对语篇中逻辑关系的分析是否正确；

推理合理性：判断考生的推断是否基于听到的信息且符合逻辑；

信息关联能力：考查考生将不同信息点联系起来理解的能力。

6.1.4 评价创造类

6.1.4.1 题型

包括观点评析、内容重构、拓展回答、综合评价等题型。观点评析题要求考生对听到的观点进行分析和评价；内容重构题要求考生用不同的方式重新组织听到的内容；拓展回答题要求考生根据听到的内容进行拓展和延伸，发表自己的看法；综合评价题要求考生对听到的多个信息进行综合分析和评价。

6.1.4.2 语料选用

语料多为论述性、议论性的文本，如演讲、辩论、评论、学术讲座等；话题涵盖社会热

点问题、文化现象、科技前沿、伦理道德等具有一定深度和广度的内容；语料中包含复杂的句子结构、丰富的词汇和多样化的修辞手法；涉及不同立场和观点的表达，具有一定的争议性和思辨性。

6.1.4.3 考查点

考查考生对复杂语篇的深层理解能力，包括对观点、论据、论证过程的把握；考查考生的批判性思维能力，如对观点的合理性、论据的充分性、论证的严密性的评价；考查考生的创造性思维能力，如对听到的内容进行重构、拓展和创新表达；同时考查考生综合运用语言知识和文化背景知识进行分析和评价的能力。

6.1.4.4 评分维度

评价深刻性：评估考生对听到内容的评价是否深入、有见地；

逻辑严密性：考查考生的分析和论证是否具有逻辑性和严密性；

创新性：判断考生在内容重构或拓展回答中是否有新颖的想法和表达方式；

综合运用能力：考查考生综合运用语言知识、背景知识和思维能力解决问题的能力。

6.2 口语能力

6.2.1 基础表达类

6.2.1.1 题型

包括单词朗读、句子模仿、简短自述等形式。单词朗读要求考生清晰读出给定的基础词汇；句子模仿要求考生仿照示例句子的语调、节奏进行重复表达；简短自述要求考生用 3-5 句话介绍个人基础信息，如年龄、籍贯、兴趣爱好等。

6.2.1.2 语料选用

语料以中小学阶段核心词汇和基础句型为主，涵盖日常高频名词（如 school、family、food）、动词（如 study、play、like）、形容词（如 happy、easy、new）等；句子多为“主谓宾”“主系表”等简单结构，长度控制在 5-8 个单词；话题聚焦个人基本情况、日常作息、身边事物等无复杂背景的内容。

6.2.1.3 考查点

重点考查考生对单词发音的准确性，包括音标、重音位置等；考查考生对基础句型的口头运用能力，能否正确把握句子节奏；同时考查考生用简单语言传递基础信息的能力。

6.2.1.4 评分维度

发音规范性：评估单词发音的标准程度及句子语调的合理性；

表达连贯性：考查口头表达时的流畅度，是否存在过多停顿或重复；

信息完整性：根据简短自述包含的关键信息，判断基础信息传递的完整度。

6.2.2 情景交际类

6.2.2.1 题型

包括对话模拟、情景应答、角色扮演等形式。对话模拟要求考生与考官围绕指定日常场景完成 3-5 轮对话，如餐厅点餐、商店购物等；情景应答要求考生针对突发简单情景作出即时回应，如他人问路时提供方向指引；角色扮演要求考生代入特定身份（如游客、店员）完成预设交际任务。

6.2.2.2 语料选用

语料以生活场景高频交际用语为主，包含服务类词汇（如 menu、price、discount）、社交短语（如 excuse me、could you...、I'm sorry）等；句子以简单句和常用复合句为主，出现“if”“because”等基础连词；话题涉及衣食住行、社交礼仪、公共服务等贴近现实生活的场景。

6.2.2.3 考查点

重点考查考生在真实语境中理解他人意图的能力；考查考生运用恰当语言进行即时回应的能力；同时考查考生根据身份调整表达语气的社交语言能力。

6.2.2.4 评分维度

语境适配度：评估语言表达与场景的匹配程度，是否符合交际规范；

回应及时性：考查对对方话语的反应速度，是否存在过长停顿；

互动协调性：根据对话中的眼神交流、话题承接情况，判断交际配合能力。

6.2.3 综合论述类

6.2.3.1 题型

包括话题阐述、观点辩论、故事复述等形式。话题阐述要求考生围绕指定主题（如“环境保护”“网络利弊”）进行 1-2 分钟论述；观点辩论要求考生针对简单命题（如“学生是否该带手机进校园”）表达立场并给出 2-3 个理由；故事复述要求考生听完一段简短叙事，用自己的语言重述核心情节。

6.2.3.2 语料选用

语料涵盖社会热点、文化习俗、哲理故事等中等难度内容，包含抽象名词（如 environment、technology、happiness）、逻辑连接词（如 however、therefore、in addition）等；句子出现“定语从句”“状语从句”等复杂结构；话题具有普遍性和争议性，适合展开多角度表达。

6.2.3.3 考查点

重点考查考生对复杂话题的逻辑梳理能力；考查考生运用丰富词汇和多样句型的表达能力；同时考查考生传递深层观点和情感的能力。

6.2.3.4 评分维度

逻辑清晰度：评估论述的层次结构，观点与论据的关联程度；

词汇丰富度：考查是否使用不同词性词汇及同义替换表达；

语法复杂度：根据复合句使用频率和准确性，判断语言运用水平；

内容深度：根据观点的独特性和论据的充分性，评估思维表达能力。

6.3 阅读能力

6.3.1 细节理解类

6.3.1.1 题型

包括事实信息查找、细节匹配、数据提取等。要求考生从文章中快速找到与题目相关的具体信息，如人物、时间、地点、事件等；在细节匹配题中，需将题目中的描述与文章中的对应内容进行匹配；数据提取题则要求准确获取文章中的数字、比例等数据信息。

6.3.1.2 语料选用

语料涵盖各类说明性、记叙性文本，如产品说明书、日程安排表、新闻报道、故事片段等；文本中包含明确的事实性信息，如具体的时间节点、地点名称、事件经过、数据参数等；难度分布广泛，简单文本语言简洁明了，复杂文本可能存在一定的信息干扰或表述转换。

6.3.1.3 考查点

重点考查考生对文章细节信息的定位和识别能力；

考查考生对具体信息的准确理解，包括对同义表述、句式转换的识别；

同时考查考生快速浏览和筛选信息的能力，看能否在大量文字中迅速锁定关键细节。

6.3.1.4 评分维度

信息准确性：评估找到的细节信息与题目要求的匹配程度；

定位速度：根据考生找到目标信息所需的时间判断其信息检索效率；

细节完整性（针对多信息题）：考查获取的细节是否涵盖题目要求的所有相关内容；

理解准确性：关注对细节信息的理解是否正确，有无因表述转换而产生误解。

6.3.2 推理判断类

6.3.2.1 题型

包括隐含信息推断、逻辑关系推理、作者意图推测等。要求考生根据文章表层信息推断出隐含的深层含义；在逻辑关系推理题中，需分析句与句、段与段之间的因果、递进、转折等关系；作者意图推测题则要判断作者写作的目的、想要传达的潜在观点等。

6.3.2.2 语料选用

语料涉及议论文、评论性文章、文学作品、科普说明文等；
文本中包含间接表达的观点、隐藏的逻辑线索、带有情感倾向的词汇等；
题材涵盖社会现象分析、科学研究探讨、文化内涵解读等，部分文本存在信息留白或表述委婉的特点。

6.3.2.3 考查点

重点考查考生对文章隐含信息的挖掘能力；
考查考生对文本逻辑关系的分析与梳理能力；
同时考查考生结合上下文语境进行合理推测的能力，看能否透过表面文字把握深层意义。

6.3.2.4 评分维度

推断合理性：评估推断结果与文章逻辑的契合程度；
依据充分性：根据推断所依据的文本信息是否充足判断；
逻辑准确性（针对逻辑推理）：考查对句段间逻辑关系的判断是否正确；
意图把握度（针对意图推测）：关注对作者潜在写作目的的理解是否到位。

6.3.3 词义猜测类

6.3.3.1 题型

包括生词含义推测、熟词僻义判断、短语语义理解等。要求考生结合上下文语境、构词法知识等推测出生词的含义；在熟词僻义判断题中，需确定常见词汇在特定语境下的特殊含义；短语语义理解题则要准确把握文中固定短语、习语的整体意义。

6.3.3.2 语料选用

语料涵盖各类文本，如科普文章、人文社科论著、文学作品等；
文本中包含一定比例的超纲词汇、多义词及习惯用语；
上下文提供可利用的线索，如定义解释、同义替换、反义对比、举例说明等；
题材涉及专业领域知识、跨文化交际内容等，词汇难度呈梯度分布。

6.3.3.3 考查点

重点考查考生利用语境线索猜测词义的能力；

考查考生对构词法（词根、词缀等）的运用能力；

同时考查考生对词汇在具体语境中含义的灵活理解能力，而非机械记忆词典释义。

6.3.3.4 评分维度

词义准确性：评估猜测的词义与该词在文中实际含义的吻合程度；

线索利用率：根据对上下文线索的挖掘和运用情况判断；

语境适配性：关注推测的词义是否符合所在句子及段落的语境；

灵活度（针对多义词）：考查对词汇在特定语境下含义的灵活把握程度。

6.3.4 主旨大意类

6.3.4.1 题型

包括段落主旨概括、全文主旨归纳、最佳标题选择等。要求考生用简洁的语言提炼出段落的核心内容；在全文主旨归纳题中，需概括出文章的中心思想；标题选择题则要从选项中选出最能涵盖文章主要内容的标题。

6.3.4.2 语料选用

语料涵盖议论文、说明文、记叙文、散文等多种文体；

文本结构清晰，或有明确的中心句，或通过逻辑串联呈现主旨；

题材广泛，涉及社会热点、文化现象、科技发展、人生感悟等，篇幅长短不一。

6.3.4.3 考查点

重点考查考生对文章整体内容的把握能力，能否区分主次信息；

考查考生对文章结构的分析与梳理能力，进而归纳出主旨；

同时考查考生的语言概括能力，看能否用精炼的语言准确表达核心思想。

6.3.4.4 评分维度

主旨准确性：评估概括的主旨与文章核心内容的匹配程度；

概括精炼度：根据表述的简洁性、有无冗余信息判断；

全面性（针对全文主旨）：考查归纳的主旨是否涵盖文章各部分关键内容；

标题适配度（针对标题选择）：关注所选标题对文章内容的涵盖与提炼程度。

6.3.5 观点态度类

6.3.5.1 题型

包括作者观点识别、态度倾向判断、不同观点比较等。要求考生从文章中找出作者明确或隐含的观点；在态度倾向判断题中，需分析作者对某一事物的情感态度（支持、反对、中立等）；不同观点比较题则要对比文章中出现的多种观点及其差异。

6.3.5.2 语料选用

语料以议论文、评论文章、访谈录、社论等为主；

文本中包含表达观点的词汇、体现态度的语气词及论证过程；

题材涉及社会争议话题、学术讨论、文化评价等，具有较强的论辩性。

6.3.5.3 考查点

重点考查考生对作者观点的准确识别与理解能力；

考查考生对作者态度倾向的感知能力，包括对语气、措辞的敏感度；

同时考查考生对不同观点的分析与比较能力，看能否理清观点间的联系与区别。

6.3.5.4 评分维度

观点识别准确性：评估识别的作者观点与文章实际表述的一致程度；

态度判断精准度：根据对作者情感倾向的判断是否准确判断；

比较清晰度（针对观点比较）：考查对不同观点差异与联系的分析是否清晰；

论证理解度：关注对作者支持观点所采用论据的理解程度。

6.3.6 篇章结构类

6.3.6.1 题型

包括段落作用分析、文章结构梳理、过渡句补全等。要求考生分析某一段落在文章中的作用（如引出话题、举例论证、总结升华等）；在文章结构梳理题中，需理清文章的整体框架及各部分的逻辑关系；过渡句补全题则要选出能使段落间衔接自然的句子。

6.3.6.2 语料选用

语料涵盖结构完整的各类文章，如说明文、议论文、记叙文等；

文本采用常见的篇章结构，如总分总、并列式、递进式、对比式等；

部分文本结构存在一定复杂性，如包含插叙、倒叙或多层面逻辑关系。

6.3.6.3 考查点

重点考查考生对文章整体结构的把握能力；
考查考生对段落功能及句间逻辑关系的理解能力；
同时考查考生对文章行文思路的分析能力，看能否理清作者的写作脉络。

6.3.6.4 评分维度

结构分析准确性：评估对文章结构及段落作用的分析是否正确；
逻辑梳理清晰度：根据对文章各部分逻辑关系的梳理情况判断；
衔接恰当性（针对过渡句补全）：考查所选句子使段落衔接的自然程度；
完整性（针对结构梳理）：关注梳理的结构是否涵盖文章主要部分及逻辑关系。

6.4 写作能力

6.4.1 应用文写作类

6.4.1.1 题型

包括书信写作（如感谢信、道歉信、邀请信等）、邮件撰写、通知拟写、便条创作等。要求考生根据特定的情境和写作目的，按照相应的格式规范完成写作任务；在书信写作中，需注意称呼、正文、结尾、署名等要素的完整性；通知拟写则要确保信息明确、条理清晰。

6.4.1.2 语料选用

提供的情境提示涵盖生活、学习、工作等多个领域，如朋友间的往来、学校事务的处理、职场中的沟通等；

包含具体的写作背景信息，如事件缘由、沟通目的、期望达成的效果等；

格式要求因文体而异，有的需要严格遵循传统格式，有的则可采用简洁的现代格式。

6.4.1.3 考查点

重点考查考生对不同应用文文体格式的掌握程度；

考查考生根据情境准确传达信息的能力，包括内容的完整性和针对性；

同时考查考生运用恰当语言进行表达的能力，看是否符合不同应用文的语气和风格要求。

6.4.1.4 评分维度

格式规范性：评估写作内容在文体格式上的符合程度；

信息准确性：根据传达的信息与情境要求的匹配度判断；

内容完整性：考查写作内容是否涵盖所有必要信息；

语言得体性：关注语言表达的语气、措辞是否符合应用文的特点和情境需求。

6.4.2 记叙文写作类

6.4.2.1 题型

包括事件叙述、人物描写、经历分享等。要求考生围绕给定的主题或情境，完整地叙述一件事情的发生、发展和结果。在人物描写中，需通过具体事例展现人物的性格、外貌、行为等特征；经历分享则要真实地描述个人的一段经历及感受。

6.4.2.2 语料选用

主题涉及日常生活中的趣事、难忘的经历、印象深刻的人物等；

提供的提示可能是一个关键词、一幅图片、一段简短的文字描述等；

有的会对写作的字数、情节的详略等提出一定要求。

6.4.2.3 考查点

重点考查考生对叙事要素（时间、地点、人物、事件、原因、结果）的把握能力；

考查考生组织故事情节的能力，看是否有条理、有逻辑性；

同时考查考生运用生动形象的语言进行描写的能力，包括对细节的刻画和情感的表达。

6.4.2.4 评分维度

叙事完整性：评估故事或经历的叙述是否完整，要素是否齐全；

情节连贯性：根据事件发展的逻辑顺序和过渡自然程度判断；

描写生动性：考查对人物、场景等的描写是否形象、具体；

情感真实性：关注表达的情感是否真挚，能否引起读者共鸣。

6.4.3 议论文写作类

6.4.3.1 题型

包括观点论述、问题分析、利弊讨论、辩论性写作、问题解决方案提出等。要求考生针对给定的话题或有争议的议题，明确表达自己的观点并运用充分的论据进行论证；在问题分析中，需剖析问题产生的原因、带来的影响等；利弊讨论要全面分析某一事物的优点和缺点；辩论性写作中，需同时考虑正反两方的观点并表明自己的立场；问题解决方案提出则要分析问题症结并给出切实可行的办法。

6.4.3.2 语料选用

话题涵盖社会热点、教育问题、科技发展、文化现象、具有争议性的社会问题、学术观点、伦理道德等多个领域；

有的会提供相关的背景信息、不同的观点、研究数据、专家观点、案例等作为参考；对文章的结构（如引言、论点、论据、结论）有一定的潜在要求，部分对论证深度和广度有要求。

6.4.3.3 考查点

重点考查考生明确表达观点和构建论证体系的能力；

考查考生收集、筛选和运用论据支持观点的能力；

考查考生的批判性思维能力，能否全面、客观地看待问题；

同时考查考生的逻辑思维能力，看能否使文章的论证过程严谨、有条理，以及提出创新性解决方案的能力。

6.4.3.4 评分维度

观点明确性：评估作者的观点是否清晰、鲜明；

论据充分性：根据所选用的论据对观点的支持程度判断；

逻辑严密性：考查文章的论证过程是否符合逻辑，条理是否清晰；

结论合理性：关注得出的结论是否与前文的论证一致，是否具有说服力；

论证深度：评估对观点的论证是否深入、有说服力；

思维全面性：根据是否考虑到问题的多个方面判断；

方案可行性（针对解决方案提出）：考查提出的解决办法是否切实可行；

语言逻辑性：关注文章的语言表达是否逻辑清晰，论证过程是否顺畅。

6.4.4 说明文写作类

6.4.4.1 题型

包括事物介绍、流程说明、原理阐释等。要求考生对某一事物的特征、性质、功能等进行清晰的介绍；在流程说明中，需按先后顺序详细描述做某事的步骤；原理阐释则要通俗易懂地解释某一现象或技术的工作原理。

6.4.4.2 语料选用

说明对象涉及自然现象、科技产品、文化习俗、生活常识等；

提供的提示可能是一个说明对象的名称、一幅相关的图片或一段简要的背景介绍；

有的会对说明的详略程度和侧重点提出要求。

6.4.4.3 考查点

重点考查考生对说明对象的理解和掌握程度；

考查考生运用恰当的说明方法（如举例、列数字、作比较等）进行表达的能力；

同时考查考生组织说明内容的能力，看能否使文章条理清晰、易于理解。

6.4.4.4 评分维度

内容准确性：评估对说明对象的介绍是否准确无误；

方法恰当性：根据所采用的说明方法与说明对象的匹配度判断；

条理清晰度：考查文章的结构和内容安排是否有条理；

语言简洁性：关注语言表达是否简洁明了，避免冗余和晦涩。

7. 英语能力人工智能分级测评系统建设

7.1 系统定义

人工智能英语系统是实现英语综合能力测评从命题、制卷到机器阅卷、自动评分的全过程操作，辅助开展大规模、常态化的英语测评的软、硬件集合。

7.2 系统架构

系统包括考务管理系统、命题制卷系统、阅卷评分系统和考试系统。

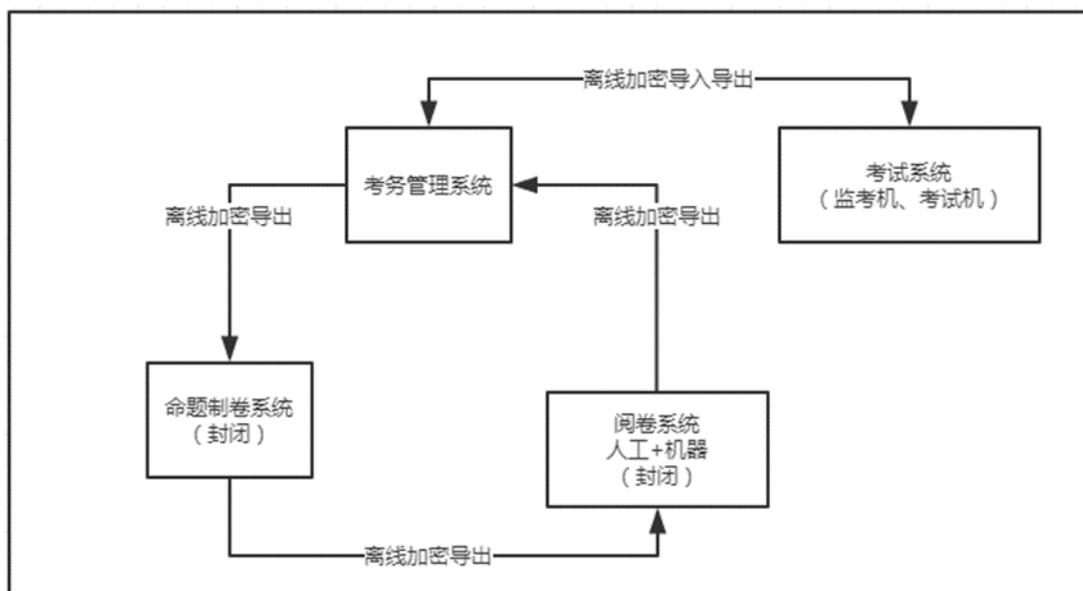


图1 人工智能英语系统架构图

7.3 考务管理系统

7.3.1 构架

7.3.1.1 C/S 和 B/S 混合架构

系统通常采用 C/S（客户端/服务器）和 B/S（浏览器/服务器）相结合的架构。

C/S 架构用于实现一些对性能和稳定性要求较高的功能，如考试客户端的程序，能更好地与本地硬件设备（如耳麦、声卡）交互，确保考试过程中音频的稳定采集和处理。

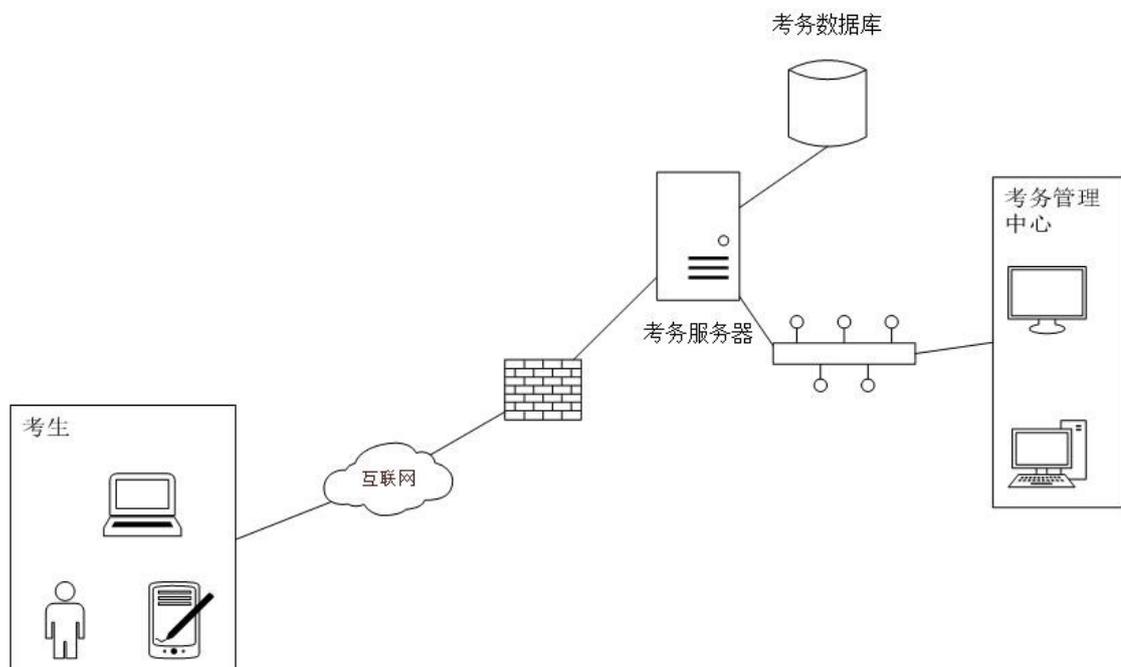
B/S 架构则用于考务管理的一些 Web 端功能，如管理员通过浏览器进行考务安排、成绩查询与统计等操作，方便不同地点的管理人员随时登录系统进行管理，无需安装专门的客户端软件。

7.3.2 功能

7.3.2.1 考务编排

考务管理系统包括：考生信息管理、考试资料和设备管理、考试流程标准化、监考和作弊防控、试卷和成绩管理、档案和追溯管理等功能。

考务管理系统采用B/S（浏览器/服务器）的体系结构，拓扑结构如下图所示：



支持手动和自动两种考务编排方式，引入考点群编排概念，可根据需要对考场、场次以及座位进行合理编排，最大限度地合理利用机房和电脑等资源，同时防止考试舞弊现象。

7.3.2.2 试题管理

试题类型丰富，客观题支持单选题、多选题、填空题、判断题、匹配题等；主观题中，针对“听”设置听力理解、听力填空、听力复述等题型，针对“说”涵盖文本朗读、视频朗读、角色扮演、口头作文、话题论述等形式，针对“读”包含阅读理解、段落排序、词义猜测、摘要概括等题型，针对“写”则有句子仿写、段落写作、应用文写作、议论文写作等题型，并可自定义每种题型的步骤，满足考试的各种题型需求。

还支持根据双向细目表制定组卷策略，保证同种试卷策略下的多份试卷等值和所考察知识点的均衡，支持手动组卷与自动组卷两种方式，可对卷面进行编辑与修改，同时可生成 Word 文档以供打印。

7.3.2.3 监考与异常处理

监考系统具有强大的监控功能，可对所有考试端进行登录、开考及收卷的控制，还具备断点续考、重考、转移考试、备考、单机收卷、强制收卷、备份收卷等异常处理机制，确保考试顺利进行。

7.3.2.4 自动识别与评分

我们构建了覆盖口语、听力、阅读、写作四大模块的智能语言能力测评体系，依托先进的多模态语言处理技术实现全维度精准评估。

听力测评基于深度神经网络的语音解码技术与语义场景建模算法，支持噪声环境下的语音分离与关键信息提取，重点评估考生对对话/独白的语义理解准确率、核心信息抓取能力及逻辑推理水平，系统可标记漏听信息点并分析干扰因素（如背景音混淆、语速适应不足等）。

口语测评采用梅尔频率倒谱参数与隐马尔可夫模型为核心的语音识别技术，结合自创的拒识汉语方言干扰过滤算法与“韵律-内容-情感”三因素绑定模型，从句子韵律自然度、内容完整性、发音准确性等多维度进行动态评判，显著提升复杂语境下的语音识别精度，并可生成可视化卷面报告，精准定位考生口语问题（如方言口音过重、语调起伏失衡、关键信息遗漏等）。

阅读测评运用 Transformer 架构的语义表征模型与知识图谱关联技术，从词汇运用精准度、长难句解析能力、篇章逻辑把握度三个层面进行量化分析，不仅能评估基础阅读理解水平，更能诊断深层语言障碍（如熟词僻义误判、段落衔接断裂、主旨归纳偏差等）。

写作测评融合语法树解析、风格迁移模型与内容相关性评估算法，可对作文的语法正确性、词汇丰富度、逻辑连贯性及内容切题度进行多级评分，并通过错误类型标注（如主谓不一致、修辞不当、论点支撑缺失等）提供针对性改进建议。

四大模块通过统一的知识图谱实现数据互通，最终生成涵盖“输入-处理-输出”全流程的综合评估报告，为学习者提供从语言感知到表达应用的能力画像。

7.3.3 成绩统计与卷面评析

具备考生试卷还原与详尽的考卷评析功能，便于考生查卷。

具有图表分布图，对考试结果可进行难度、标准差、试卷信度及试题区分度的统计与分析。

可统计平均分、及格率、各小题得分率、区分度、难易度等信息，并提供详细的成绩分析统计报表。

7.3.4 数据校验

考试数据在流转过程中采用签名机制和双向校验机制，确保数据包完整性和准确性。

同时系统在回收数据过程中还可自动进行缺考、丢包、坏包的核查与统计，确保所回收数据的可用性与对异常数据的及时处理。

7.3.5 系统维护与更新

考试服务器安装完毕后，考生端只需在 IE 地址栏输入服务器的 IP 地址即可实现自动安装，当有新版本时，考生端能自动进行识别与更新升级。此外系统还支持无盘环境部署，安装完成后，可对考试环境进行自动测试，如耳麦、声卡、键盘、鼠标测试。

7.3.6 系统软硬件配置环境

7.3.6.1 硬件环境

服务器通常采用高性能的专用服务器：

CPU 为 Xeon 系列以上、频率 2.4G 或以上

内存 16G 或以上

配备 146G 以上的双硬盘（其中一个硬盘作为磁盘镜像）

双电源备份

1000M 网卡

配备 UPS 电源以保证供电稳定

考试计算机采用主流品牌计算机：

处理器 I3 及以上

内存 4G 及以上

硬盘 160G 以上

100M 或以上网卡

2 个以上 USB 接口，用于连接考试专用耳机

考试专用耳机为包耳式设计，USB2.0 接口、线长在 1.5 米以上，适合不同头型佩戴，整体无任何线控或按钮调节，耳机集成声卡，麦克风为单指向降噪麦克风。

7.3.6.2 软件环境

服务器操作系统一般为 Windows Server 2008 R2、Server 2012 或 Server 2016 等。

考试计算机操作系统为 Windows 7 或以上。系统还需安装专业的英语考试软件，包括考试客户端程序、监考端程序、考务管理端程序等，以及相关的语音识别引擎和数据库管理系统。

7.3.7 系统性能

7.3.7.1 高准确性

智能测评技术依托多模态技术能力，构建起覆盖听说读写全维度的精准评估体系。语音识别技术通过高精度声学建模与语义理解，可完整捕捉口语回答的流利度、准确性与表达逻辑；OCR 扫描与图文识别技术则针对手写内容（如填空题、作文等书面作答）实现快速转写与结构化解析，确保非结构化数据的有效转化。特别在写作评测环节，大语言模型的深度语义理解与内容生成能力被深度赋能，可从立意深度、逻辑架构、语言规范、创新表达等多元维度开展精细化评估。经大规模测试验证，该技术在听力理解、口语表达、阅读作答及写作输出等场景的评分信度与人工评分高度吻合，充分体现了技术驱动下评测的客观性与科学性。

7.3.7.2 高稳定性

系统在考试过程中能够稳定运行，具备强大的容错能力和异常处理机制，能够应对各种突发情况，如网络故障、设备故障等，保证考试的顺利进行，减少因系统问题导致的考试中断或数据丢失等情况的发生。

7.3.7.3 高效性

支持大规模的并发考试，能够同时处理多个考场、多个考生的考试请求，快速完成语音识别、评分、数据存储等操作，在考试结束后能够及时生成成绩统计报表和卷面评析报告，为考务管理提供高效的支持。

7.3.7.4 可扩展性

系统具有良好的可扩展性，能够根据考试规模的扩大、题型的增加或功能的升级等需求进行灵活扩展和升级，无需对系统进行大规模的重新开发。

8. 英语能力人工智能分级测评题库建设

8.1 系统构架

8.1.1 数据层

负责存储题库中的各类数据，包括题目文本、音频、视频等资源，以及题目相关的元数据，如难度系数、知识点、适用场景等。

通常使用数据库管理系统来高效组织和管理这些数据，确保数据的安全性和可靠性。

8.1.2 逻辑层

是题库系统的核心处理部分，包含题目生成、难度评估、信度和效度分析等算法和逻辑。

它根据一定的规则和策略，对数据层中的数据进行处理和运算，实现题库的各种功能，如根据用户需求生成符合特定难度和知识点的试卷。

8.1.3 应用层

为用户提供与题库交互的界面，包括出题者用于管理题目、设置参数的后台管理界面，以及考生在测评过程中访问题目、进行答题的前端界面。

应用层通过与逻辑层的交互，实现题目展示、答题记录、结果反馈等功能。

8.2 模块

8.2.1 题目管理模块

负责对题目进行添加、编辑、删除等操作。

出题者可以在此模块中输入题目内容，包括文字描述、语音提示、相关图片或视频等信息，并设置题目所属的知识点、题型、难度等属性。

8.2.2 难度评估模块

运用特定的算法和模型，根据题目的语言复杂度、话题熟悉度、任务要求等因素，对题目进行难度系数的评估和调整。

该模块还可以结合考生的实际答题数据，动态优化题目难度评估。

8.2.3 信度和效度分析模块

通过统计分析方法，对题目进行信度和效度检验。信度分析主要关注题目在多次测试中的稳定性和一致性，效度分析则侧重于题目是否能够准确测量出考生的英语能力。

根据分析结果，对题目进行筛选和优化，确保题库中题目的质量。

8.2.4 试卷生成模块

根据测评的要求和参数，如考试时间、题型分布、难度比例等，从题库中随机抽取或按照一定规则组合题目，生成符合要求的试卷。

试卷生成过程中要确保题目之间的关联性和整体性，避免出现重复或相互干扰的情况。

8.2.5 数据汇总模块

收集和整理考生的答题数据，包括答题时间、答案内容、得分情况等。

对这些数据进行分类统计和分析，为题目难度调整、信效度分析以及教学反馈提供依据。

同时，数据汇总模块还可以生成各种报表，直观展示题库的使用情况和考生的整体表现。

8.3 功能

8.3.1 题目录入与编辑

支持多种题型的题目录入，如对话填空、朗读短文、话题陈述、图片描述、听后选择、读后续写等。

出题者可以方便地输入题目文本、录制语音、上传相关多媒体素材，并对题目进行格式调整和内容修改。

8.3.2 题目审核与发布

对录入的题目进行审核，确保题目内容准确无误、语言规范、难度适中且符合测评目标。

审核通过的题目可以发布到题库中，供测评使用；未通过审核的题目则返回给出题者进行修改。

8.3.3 题目检索与查询

能够根据题目编号、关键词、知识点、难度系数等多种条件对题库中的题目进行快速检索和查询，方便出题者和管理人员查找和管理题目。

8.3.4 试卷生成与导出

按照预设的规则和参数生成试卷，并提供试卷导出功能，可将试卷以 Word、PDF等格式导出，以便打印或在线发布。

系统还应支持试卷的预览功能，让出题者在生成试卷后能够提前查看试卷的整体布局和内容。

8.3.5 试卷检验

命题专家完成试卷命制后，通过试卷预览模式进行完整性、正确性检查，包括：试卷的基本信息、试卷的总分、听力部分总分、口语部分总分、阅读部分总分、写作部分总分、试卷的总时长、试卷的题目数量、试卷内的题型的完整性、试题内容信息、试题内容的拼写和语法错误等。

命题专家和命题负责人可以按照考生的实际操作方式，模拟完成每套试题的答题过程，以便检验试卷的实际完整性。

命题专家制作完成试卷后，提交给命题负责人，命题负责人审核确认后，试卷的校验工作才能确认完成。

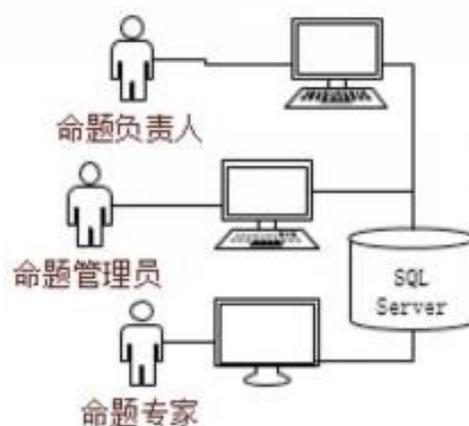
8.3.6 数据统计与分析

对考生的答题数据进行详细的统计和分析，包括题目得分率、难度系数变化趋势、信度和效度指标等。

通过可视化图表和报表的形式展示分析结果，帮助出题者和教学人员了解题库的质量和考生的学习情况，为教学改进和题库优化提供参考。

8.3.7 用户管理

命题制卷环境（外部隔离）



命题制卷系统包括三个用户角色：命题系统管理员、命题负责人、命题专家；

命题制卷系统初始化时，通过默认的命题系统管理员账户登录。该命题系统管理员可以创建新用户，定义角色；

命题系统管理员和命题负责人有角色管理的权限，可以创建新角色，对角色进行权限分配和成员维护；

命题系统管理员和命题负责人可以对系统用户进行信息管理和账户密码管理；

命题制卷系统存储所有用户的各种操作记录。命题负责人和命题系统管理员可以查看用

户的日志信息（含操作对象、操作时间、操作功能等）。

8.4 数据汇总

8.4.1 数据收集

在考生进行英语测评过程中，系统自动记录考生的各项答题数据，包括每个题目对应的答题时间、考生的语音回答内容、系统自动评分结果等。

同时，还收集与考生相关的背景信息，如年龄、性别、学习阶段等，以便进行更全面的数据分析。

8.4.2 分类整理

将收集到的数据按照不同的维度进行分类整理，如按照题目类型、知识点、难度系数、考试批次等进行分类。这样可以方便对不同类别数据进行单独分析，也便于发现数据之间的潜在关系和规律。

8.4.3 统计分析

运用统计学方法对整理后的数据进行分析，计算各种统计指标，如平均分、标准差、最高分、最低分、得分率等，以描述考生群体在各个题目和整体测评中的表现。

此外，还可以进行相关性分析、差异性分析等，探究不同因素（如性别、学习阶段等）与考生成绩之间的关系，为题库优化和教学改进提供依据。

8.4.4 结果呈现

将数据汇总和分析的结果以直观的方式呈现出来，通常采用图表（如柱状图、折线图、饼图等）和报表相结合的形式。例如，通过柱状图展示不同难度系数题目上考生的得分情况，用折线图呈现同一知识点在不同考试批次中的难度变化趋势，以报表形式详细列出各题型的平均得分、标准差等统计数据。这些可视化的结果能够帮助教育工作者和题库管理人员快速了解题库的使用效果和考生的学习状况，从而做出相应的决策。

8.5 题目来源

8.5.1 专业出题团队编写

由具有丰富英语教学经验和专业知识的教师、专家组成出题团队，根据英语测评的目标和要求，按照一定的规范和标准编写题目。

这些题目通常具有较高的质量和针对性，能够准确考查考生的英语能力。

8.5.2 改编自权威教材和资料

从国内外权威的英语教材、英语考试真题、英语语言学习资料等中选取合适的内容进行改编。

这些素材经过了长期的教学实践检验，具有较高的权威性和实用性。通过对其进行适当的改编，使其符合人工智能英语测评的题型和要求，能够有效地利用已有的优质资源，丰富题库内容。

8.5.3 收集真实语言场景素材

通过收集现实生活中的真实语言场景素材，如英语电影、电视剧、广播节目、英语角活动等，从中提取有价值的语言片段作为题目素材。

这些素材具有真实、自然的特点，能够更好地考查考生在实际语言环境中的运用能力。自适应题库建设。

9. 英语能力人工智能分级测评系统考务

考试系统包含三个阶段：考前准备阶段、考试阶段、考试结束阶段，考试系统对硬件环境的要求如图：

场景	设备	性能要求
	监考机	CPU: Intel i5 (或者同等性能) 内存: 2G 硬盘: 100G 接口: USB×1 操作系统: Windows XP SP3, Windows 7或以上
	考试机	CPU: Intel i3 (或者同等性能) 内存: 2G 硬盘: 100G 接口: USB×1 操作系统: Windows XP SP3, Windows 7或以上
	网络带宽	100M以上

<p>考试现场</p>	<p>耳麦</p>	<p>接口：USB</p> <p>声卡：集成声</p> <p>卡类型：头戴</p> <p>护耳线长：2.8米</p> <p>喇叭频率范围：20Hz～20KHz喇叭灵敏度：98dB±3dB</p> <p>喇叭阻抗：32</p> <p>15% 麦克风类型：单指向</p> <p>麦克风频率范围：20Hz～16KHz麦克风灵敏度：-42±3dB</p> <p>麦克风阻抗：0.68KΩ</p>
-------------	-----------	---

9.1 考前准备阶段

考务管理人员在组织完成考生报名与编排后将本次考试的报名与编排信息提交到考试管理程序进行校验、生成考试任务。

考点校管理员登录监考机程序，根据解密私钥和授权的时钟锁，使用加密试卷包U盘，将当次考试的试题导入到监考机程序；考前1周左右，启动考场环境检测，检查考场内的所有考试机状态是否正常，进入“准备考试”阶段，考试机进入用户登录界面并可自动完成考试流程，对所有考试机进行考前检测。

9.2 考试阶段

在备考室内完成身份认证的考生，按照监考老师的指令依次进入考场机位，使用准考证号登录考试机程序。随后，考生需依次完成耳机、麦克风、键盘等相关设备的检测，根据提示调节合适的音量与输入设备参数，并进行试音和键盘打字测试。系统自动对考生试录音质量、键盘输入响应进行智能判断，当音质未达标或键盘输入异常时进行智能提示，考生调整后继续重新测试，测试完毕后进入准备就绪状态，等待监考老师的指令准备进入正式考试。

监考老师在监考机程序内监测考生状态，所有考生准备就绪后，按照监考老师的指令，启动当场考试。监考机程序自动将当场考试包含听力、阅读、写作、口语的试题发送到考试机程序。考生先进入听力答题环节，佩戴耳机听取听力材料并通过考试机作答；听力考试结束后，自动跳转至阅读答题环节，考生在屏幕上阅读文本并完成相应题目；接着进入写作答

题环节，考生使用键盘在考试机上完成写作任务；最后进入口语答题环节，考生根据题目要求进行口语表达。全部答题结束后，考生提交答卷，系统后台自动对考生本次听力作答完整性、阅读答题记录、写作文本规范性以及口语录音质量进行智能判断。对于听力漏答、阅读未完成题目、写作字数不足或格式错误、口语录音质量不符合评分要求的考生进行检出，考务管理人员可根据考务要求合理安排考生进行重考。在考生正常答题完成后按照监考员指令离开考场。

同一场的考生都完成答题后，监考老师结束当场考试，并开始进入下一场的考试实施阶段。

9.3 考试结束阶段

所有场次的考试全部完成后，考点校系统管理员进行数据回收，将考生答题数据和考试过程数据进行统一交给上一级机构。

考试系统在英语口语考试机考考场的局域网内运行，采用客户机与服务器方式（C/S）体系结构。不同考场内安装的计算机考试系统应相互独立，互不干扰。

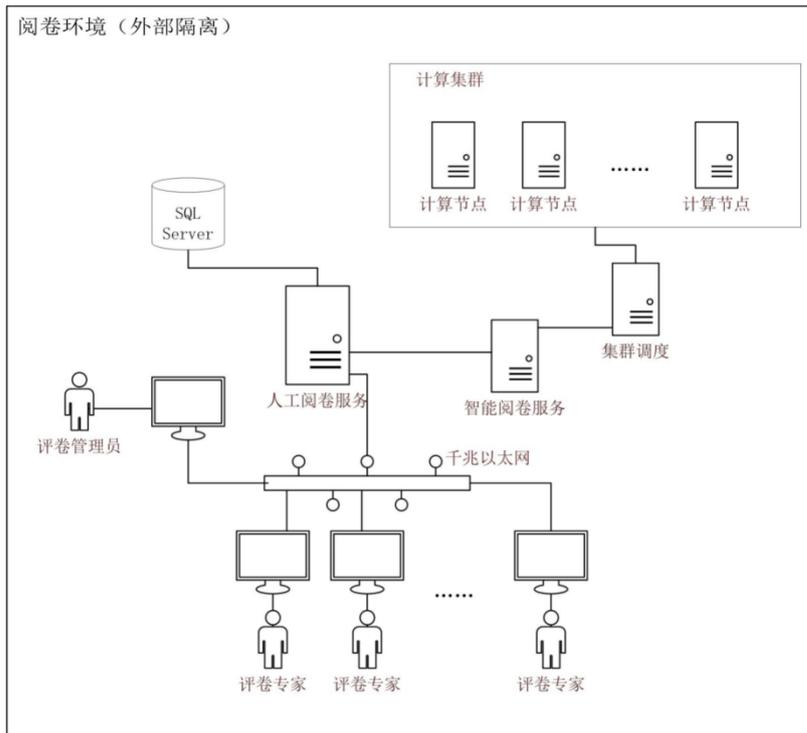
考试系统的监考机读取考前制作的解密私钥，通过密钥解析考试任务，根据考试任务里的时间信息控制电子试卷包的启封和下载时间。

考生的答题数据通过加密算法进行压缩保存，避免数据泄露，保证考生语音文件的完整。答题过程中，考试机程序在考试机本地硬盘保存答题数据时，同时在监考机上保存一份副本，解决考生在答题过程中出现异常，可以切换到备用机上继续答题。监考老师可以根据考试实际情况安排考生重考。

考试结束后，考试机和监考机要封闭保留到成绩发布后，以备阅卷期间的数据复查。

10. 英语能力人工智能分级测评评分系统

评分系统由基础功能、人工评分和智能评分三大部分构成，全面涵盖英语听说读写能力的评估，确保对考生英语综合能力进行精准测评。



10.1 人工评分

参与人工评分的评卷专家需要通过考核,确保每一位专家对评分标准的理解是一致的。

人工评分完成两方面的工作:人工定标、特殊音频评分。

人工定标阶段采用双人评分+专家仲裁的模式实现,及由两组老师分别对样卷音频进行打分,同一个音频两次打分的分差大于总分的15%时,由专家进行仲裁打分;分差在总分的15%以内的音频分数取两次打分的平均值。在对样卷评估时,不仅关注口语表达,还会综合考量听力理解(如对听力材料的理解准确程度)、阅读理解(如对阅读文本的分析深度)、书面表达(如写作的逻辑结构、语法词汇运用)等方面,全面校准评分标准。对于特殊音频评分阶段,除了处理口语相关特殊情况外,若涉及听力材料的特殊情况,也会交由专家从听力评分维度进行打分。

场景	设备	性能要求
人工评分	服务器	CPU: Intel E5-2670 (或者同等能) 内存: 8G 硬盘: 100G 接口: USB1 操作系统: Windows 2008 或以上

客户端	客户端	CPU: Intel i3 (或者同等性能) 内存: 2G 硬盘: 100G 接口: USB×1 操作系统: Windows XP SP3, Windows 7 或以上
	耳麦	接口: USB 声卡: 集成声卡类型: 头戴护耳线长: 2.8 米 喇叭频率范围: 20Hz~20KHz 喇叭灵敏度: 98dB ±3dB 喇叭阻抗: 32 15% 麦克风类型: 单指向 麦克风频率范围: 20Hz~16KHz 麦克风灵敏度: -42±3dB 麦克风阻抗: 0.68KΩ
	网络	100 兆

10.2 智能评分

智能评分阶段包含机器学习、评分模型校验、机器评分三个阶段。

机器学习的数据来源于人工定标阶段输出的数据，通过文本训练、特征提取、自适应训练三个步骤生成评分模型。

在生成评分模型过程中，充分考虑听说读写各部分数据特征。例如，在处理阅读相关数据时，提取文本理解、主旨概括等特征；在处理写作数据时，分析语法准确性、词汇丰富度、内容连贯性等要素，以构建更完善的评分模型。

机器评分阶段，利用服务器资源对听说读写各项能力进行全面评分，智能评分系统支持分布式工作、动态负载均衡，同时监控各部分评分任务的服务器资源使用情况，如 CPU 使用率、内存使用情况等，确保评分高效准确。

场景	设备	性能要求
	分发节点	CPU: Intel i5 (或者同等性能) 内存: 4G 硬盘: 500G 操作系统: CentOS或以上

智能评分	计算节点	CPU: Intel E5-2670 (或者同等性能) 内存: 8G 硬盘: 500G 操作系统: CentOS或以上
	监控节点	CPU: Intel i5 (或者同等性能) 内存: 4G 硬盘: 500G 接口: USB1 操作系统: Windows 7或以上
	网络	1000M或以上

11. 英语能力人工智能分级测评报告

英语能力人工智能分级测评的报告功能是对考生英语听说读写综合能力进行全面、细致评估后生成详细报告的重要环节，以下是对其报告模式、类型、模块等方面的描述：

11.1 报告模式

11.1.1 在线报告

考生和相关人员（如教师、家长等）可以通过登录系统，在指定的界面查看测评报告。这种模式方便快捷，无需等待纸质报告的邮寄，能够及时获取测评结果。

在线报告可以具备交互功能，例如可以点击链接查看具体的语音样本、答题记录等，方便用户深入了解测评细节。

11.1.2 纸质报告

以纸质文档的形式呈现测评报告，通常会采用精美的排版和印刷，使报告看起来更加正式和专业。

纸质报告适合需要存档或需要更直观展示的场所，如学校的教学档案、学生的留学申请材料等。

系统可以提供报告下载功能，用户可自行打印，也可以由考务机构统一打印并发放。

11.2 报告类型

11.2.1 考生个人报告

针对每一位考生生成的个性化报告,详细展示该考生在英语听说读写测评中的各项表现。包括整体得分、各个维度的具体得分、与同批次考生的对比情况、优势和不足之处等,帮助考生了解自己的英语水平和存在的问题,以便有针对性地进行学习和提高。

11.2.2 班级/群体报告

面向一个班级、学校或其他特定群体的综合报告。除了展示群体的整体英语听说读写测评水平统计数据,如平均分、最高分、最低分、各分数段人数分布等,还可以分析群体在不同知识点、题型上的表现,以及与其他类似群体的对比情况,为教学管理者和教师提供宏观的教学参考,以便制定更适合的教学计划和教学方法。

11.2.3 机构评估报告

主要为教育机构、考试主办方等提供的报告。内容侧重于对整个测评过程的评估,包括试题的难易程度、区分度,测评系统的稳定性、准确性,以及不同地区、不同年龄段考生的表现差异等。帮助机构了解测评的有效性和可靠性,为后续的测评改进、课程设计和市场推广等提供数据支持。

11.3 报告模块

11.3.1 基本信息模块

包含考生的姓名、性别、年龄、考试时间、考试类型等基本信息,以便明确报告的对象和背景。

11.3.2 整体评价模块

给出考生在听说读写各项能力及综合能力的总体得分和等级,用简洁明了的语言对考生的英语综合能力进行概括性评价,如“英语综合能力优秀,听力理解准确、阅读理解深入、书面表达流畅、口语发音标准且交流自如;或英语综合能力有待提高,在听力捕捉信息、阅读理解分析、书面语法词汇运用和口语表达方面存在一些问题,需要加强练习”。

11.3.3 维度分析模块

将学生的英语听说读写能力拆解为听力理解、口语表达、阅读理解、书面表达四大核心维度展开深度剖析。每个维度均提供具体评分及针对性描述,具体如下:

听力理解:从信息抓取的准确性、细节辨析能力、逻辑推断能力等层面评估。明确标注

学生能精准捕捉关键信息（如主旨大意、数字、人物关系等）的表现，同时指出存在的问题（如漏听细节、混淆近音词、难以理解长难句逻辑等），并提供分层训练建议（如精听与泛听结合、针对连读弱读现象专项练习等）。

口语表达：融合发音、流利度两大子维度综合评估。发音维度聚焦音标准确性、重音语调合理性，指出清晰标准的部分及存在的问题（如特定音标混淆、语调平淡无起伏等）；流利度维度关注语速稳定性、停顿逻辑性，分析表达流畅的表现（如自然断句、语速适中）及影响因素（如频繁卡顿、重复用词等），并配套针对性改进方案（如跟读模仿、制定话题限时表达训练等）。

阅读理解：考查文本解读的深度与广度，包括主旨概括能力、细节定位能力、词义推测能力、篇章结构分析能力等。肯定学生在快速抓取主旨、精准定位细节等方面的亮点，同时指出存在的不足（如难以理解抽象词汇、忽略上下文逻辑关系、对复杂句式理解偏差等），并提供阅读策略指导（如学习略读跳读技巧、建立长难句分析笔记等）。

书面表达：整合语法、词汇两大子维度进行评估。语法层面重点关注语法规则运用的正确性（如时态一致性、句型结构完整性等），梳理常见错误（如主谓不一致、从句连接词误用等）并附具体修正案例；词汇层面评估词汇运用的丰富度与适配性，肯定高级词汇、地道表达的使用亮点，指出词汇量不足、搭配不当等问题，同时提供提升建议（如积累主题词汇库、学习同义替换技巧等）。

11.3.4 对比分析模块

将考生在听说读写各项能力及综合能力的成绩与同批次考生的平均成绩、优秀水平进行对比，以图表或数据的形式直观地展示考生在群体中的位置。还可以与考生过去的测评成绩进行对比，观察其英语综合能力的发展趋势，让考生和教师能够清晰地看到在不同能力维度上的进步或退步情况。

11.3.5 典型问题示例模块

从考生的口语回答中选取具有代表性的问题片段，如发音错误、语法错误、表达不流畅等，进行详细分析和标注，让考生更直观地了解自己的问题所在，同时也为教师提供具体的教学案例。

从考生的听力答题记录中选取理解错误的典型片段，分析原因并标注；

从阅读答题中选取对文本理解偏差的案例进行详细解读；

从写作答案中选取语法错误、逻辑混乱的段落进行分析和标注，让考生更直观地了解自己在英语各方面的问题所在。

11.3.6 学习建议模块

根据听说读写各项能力的测评结果,提供更全面的个性化学习建议。如“如果考生在听力方面存在问题,建议可以包括多听标准的英语发音材料、进行精听和泛听练习、尝试听写训练等;如果阅读理解能力不足,建议增加阅读量,学习阅读技巧,分析不同类型文章的结构和特点;对于写作薄弱的考生,建议多进行写作练习,学习优秀范文的结构和表达方式,注重语法和词汇的积累与运用;口语方面存在问题的考生,建议多模仿原声进行跟读练习、参加英语交流活动等。

11.3.7 附录模块

可能会包含一些补充信息,如测评所依据的评分标准、参考的语料库信息、相关的英语学习资源推荐等,以便考生和教师更好地理解测评的依据和方法,同时也为考生提供进一步学习的参考资料。

参考文献

- [1] 中华人民共和国教育部国家语言文字工作委员会 [2018]《中国英语能力等级量表》
- [2] 中华人民共和国教育部 [2022]《义务教育英语课程标准(2022版)》
- [3] 中华人民共和国教育部 [2017]《普通高中英语课程标准(2017版)》
- [4] 中华人民共和国教育部 [2000]《高职高专教育英语课程教学基本要求(试行)》
- [5] 中华人民共和国教育部 [2020]《大学英语教学指南》(2020版)
- [6] 中华人民共和国教育部 [2020]《普通高等学校本科外国语言文学类专业教学指南》

中国智慧工程研究会

英语能力人工智能分级测评实施规范

版权专有 侵权必究