

基于改革后的托福口语考试研究

——2026年1月版本综述与备考指南

仲慕秋

新东方国际教育培训事业部上海分中心，上海

收稿日期：2025年9月17日；录用日期：2025年10月10日；发布日期：2025年10月24日

摘 要

自2026年1月21日起，ETS将对托福iBT (TOEFL iBT)考试进行重大改革：考试采用更贴近现代学术场景的题材、在阅听两部分引入分段自适应设计，并对评分尺度与题型进行了整体调整。其中口语部分由原有的“一项独立口语题 + 多项综合口语题”整体重构为以“Listen and Repeat (听后复述)”与“Take an Interview (虚拟面试)”为主的新题型，评分尺度改为1~6 (0.5为步长，且在两年过渡期内，ETS依然将提供对应的120分换算分值)，并结合ETS的自动评分引擎与人工复核流程。本文先回顾托福口语的发展脉络，再详述2026年1月21日后口语题型与评分机制，对中国考生可能带来的影响与机遇进行分析，并给出针对性的备考建议。

关键词

托福，口语，备考，教学

A Study of the New TOEFL Speaking

—Overview and Study Guide of TOEFL Speaking for the January 2026 Edition

Muqiu Zhong

Shanghai Center, Global Education and Training Office, New Oriental Education & Technology Group, Shanghai

Received: September 17, 2025; accepted: October 10, 2025; published: October 24, 2025

Abstract

Starting from January 21, 2026, ETS will implement major reforms to the TOEFL iBT. The exam will adopt content that more closely reflects modern academic contexts, introduce a multistage adaptive design in both the reading and listening sections, and make overall adjustments to the scoring scale

文章引用：仲慕秋. 基于改革后的托福口语考试研究[J]. 国外英语考试教学与研究, 2025, 7(4): 153-158.
DOI: 10.12677/oetpr.2025.74018

and task types. In the speaking section, the original structure of “one independent speaking task + several integrated speaking tasks” will be fully restructured into new formats centered on “Listen and Repeat” and “Take an Interview.” The scoring scale will change to 1~6 (with 0.5 increments, and during a two-year transition period, ETS will continue to provide a conversion to the familiar 120-point scale). Scoring will combine ETS’s automated scoring engine with human review. This article first reviews the development of TOEFL Speaking, then explains in detail the new task types and scoring mechanisms after January 21, 2026, analyzes the potential impacts and opportunities for Chinese test-takers, and finally offers targeted preparation strategies.

Keywords

TOEFL, Speaking, Test Preparation, Instruction

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 托福口语的发展回顾

托福考试自 20 世纪后半叶成为衡量非英语母语学习者学术英语能力的重要标准，早期的 ETS 口语考试(Test of Spoken English)使用录音机和纸质题本作为载体，包含看图说话、回答问题、角色扮演等题型，并且依赖人工打分[1]。随着 2005 年以后计算机化、整合任务(integrated tasks)以及口语录音采集的应用，转向“独立 + 整合”并强调学术场景下的综合语言运用能力(例如：学术阅读 - 学术听力 - 口语整合)[2]。近年，ETS 又推出面向更短、更通用场景的 TOEFL Essentials，并研究“诱发模仿(elicited imitation)”任务在口语测评中的可靠性，为后续 iBT 改革提供研究依据。2025~2026 年的改革是在这些技术、测评研究与教育需求变化的背景下推进的。

2. 2026 年 1 月 21 日后托福口语考试的总体框架

根据 ETS 官方说明，2026 年 1 月 21 日及以后版本的托福 iBT 口语部分呈现以下总体特征：全新的任务类型、显著缩短的口语考试时长(11 题，约 8 分钟)，题材更贴近日常校园与现代学术场景；口语与写作的评分由 ETS 自动评分引擎主导，同时保留人工复核以确保公平性与一致性；考试总分体系也改为 1~6 量表(0.5 为步长)，并在两年过渡期内保留与 0~120 体系的对照分数以便机构适应。具体任务类型如下：7 题 Listen and Repeat (听后复述/Elicited Imitation)与 4 题 Take an Interview (虚拟面试/virtual interview)[3]。

3. 2026 年 1 月 21 日后口语题型详解

3.1. 听后复述(Listen and Repeat)

形式与目的：在每一个听后复述任务中，考生需要听一个句子，随后需尽可能保留原意与语音形式进行复述[3]。

该任务在 TOEFL Essentials 与 ETS 相关研究中被称为诱发模仿(EIT)，被证明能在较短时间内有效测量发音、语流与语法处理能力的综合表现[4]。

考查点：发音清晰度(intelligibility)、语调与节奏、短时记忆与语法重组能力、句子成分复述的准确性。因任务要求“尽量重现”，对语音与句子结构的敏感度将成为得分关键。

3.2. 虚拟面试(Take an Interview)

形式与目的：模拟与预录“考官”或面试官的对话，考生回答与个人经历、观点、学术场景或校园生活相关的开放性问题(通常为 5 个左右短问)。该任务意在评估考生以自然、连贯方式陈述经历与论点的能力，类似 TOEFL 以往的“独立口语”但更加面向真实对话语境[2] [3]。

考查点：组织表达(coherence)、信息量与细节支撑、语法与词汇多样性、口语流畅度与互动适应性(如回应问题焦点的准确性)。

3.3. 题量与时长

官方说明列出口语题型合计约 11 题、估计时长约 8 分钟。与旧版相比，答题时间更短、节奏更紧且任务数更多，但每题单次输出可能更短、更侧重即时表达与语音清晰度[3]。

4. 评分标准与评分机制解析

4.1. 分数设定与解释

自 2026 年 1 月 21 日起，托福采用 1~6 分尺度(步长 0.5)对每一科进行评分，四项平均取整至最接近的半分作为总分；在为期两年的过渡期内，成绩单仍同时显示与旧的 0~120 可比分值以便高校过渡。ETS 同时给出 CEFR 对照表，便于院校与学生理解分数含义(例如 5 分 $\approx$ C1) [5]。

4.2. 评分主体与流程

ETS 将以 AI 自动评分引擎(基于 NLP 与深度学习技术)为主对 Speaking 与 Writing 进行首轮评分，评分模型训练基于大量人工评分样本；人工评分保留为质量控制手段(用于模型校正、抽查与争议复核)。官方教师 FAQ 与评分指南显示，评分依据仍与以往相似：包括内容完整性、语用/连贯性、语言形式(发音、语法、词汇)与传达清晰度，但评分细则已被重新表述以适应短时高频的“复述”与“面试”任务类型[6]。

4.3. 对“听后复述”与“虚拟面试”的具体评分要点(摘要)

4.3.1. “听后复述”的评分要点

重视语音可懂度、句子信息保留与结构再现；过度删减原信息或改变核心意义会明显扣分；发音错误若影响理解则降分；自我纠正也会导致扣分[7]。

4.3.2. “虚拟面试”的评分要点

重视答案的相关性、拓展信息(举例、细节)、语篇组织与流畅度；语言形式错误若不影响交流则影响较小，但重复或停顿过多会影响评分[7]。

5. 对中国考生的影响与机遇

5.1. 影响

记忆与短时加工能力被放大：听后复述要求短时记忆与语音处理能力较强，部分弱项为“听写/复述”的考生可能需要额外训练，鉴于中国基础教育阶段大部分英文教学更注重书面英文使用，部分省份高考不包含听力项目，部分考生会在这一项目中遇到进一步的挑战[8]。

语音与可懂度更为重要：由于自动评分引擎同样敏感于发音清晰度(尤其是能否被自动识别为与原句高度一致)，相比于综合考量听力阅读等内容的托福综合题，新版本的托福对于发音清晰的考量比重相对

会更高。

时间管理与即时反应压力增加：口语总时长缩短、任务节奏加快，临场反应速度与组织能力成为新的瓶颈，对于平时习惯于“思考并翻译成英文再说”的考生而言，这一挑战会显著增加备考压力。

5.2. 机遇

任务更接近真实交流场景：虚拟面试模拟真实对话，对长期注重“交际能力”的考生(而非仅备考技巧)有利；具备日常英语使用经验者可能反而更占优势。考虑到大量的中国考生平时在学习和工作环境中可以使用英文，或可以通过视频或播客等接触到真实的英文使用对话，这些考生会较少感受到新题型的考察压力。

自动评分减少主观误差：AI 评分在稳定性与一致性上有潜力降低个别人工评分的主观波动(但仍需人工复核)。考生可能更少需要担心自己的回答是否在背景知识上、内容上或文化思想上可被考官个人理解。这对于平日并不完全处在美国式思想环境中的中国考生而言尤为难能可贵。考生可能不再需要认真准备可以被美国考官快速理解接受的西方例子，而可以转为使用自己生活中常见的经历来表达；这既可以降低思考与使用例子的难度、提升表达速度，也可以减少备考过程中花在积累理由和例子上的时间；增加备考效率。

短任务、多次表达：题目分散，考生有更多独立输出机会：即使个别题失误，整体仍可通过其他题型弥补表现。相比现行托福的 1 个独立题，未来新托福多个题目也会减少考生受到不熟悉话题影响从而整体无法作答的可能；同时新托福多个虚拟面试题相互之间话题相关，也会减少后续话题的思考难度。

6. 针对中国考生的备考建议

以下建议结合 ETS 官方说明与题型特性，供考生与教师参考。

6.1. 针对“听后复述”的训练

诱发模仿练习(EIT drills): 每天练习逐句听后复述，初期可选慢速句子，逐步提高句子长度与语速；重点练习保留句子核心信息与句子结构的完整再现。常见的托福备考网站，如新东方在线托福练习网站，均支持“精听”功能，考生可以利用网站的逐句播放功能，进行跟读。同时备考网站也支持变速播放听力，考生可以选择自己可以接受的速度开始练习，并逐步提升播放速度，提升练习效果[9]。

记忆与笔记法: 训练在听句子的 0~2 秒内抓关键词(主语、谓语、时间/地点短语)，通过短符号记下核心信息再复述。这一练习可以帮助考生提升短时记忆能力。

发音与声韵训练: 重点练习连读、弱读、重音位置，确保答题时语音清晰易懂。参考 ETS 的发音评分指南进行自检。考生可以使用类似于“新东方口语小助手”的软件来帮助自己提升连读、弱读、重读等语音现象的使用能力[10]。

积累场景短语词汇: 重点关注考试涉及的场景中常见的词汇和搭配，减少听力理解时处理信息、记忆信息的难度，也进一步降低口语复述组织时的难度[11]。

6.2. 针对“虚拟面试”的训练

话题框架模板: 准备常见口语话题(学习经历、团队合作、挑战与解决、短期/长期目标等)，并用“观点 - 解释 - 例子”结构训练回答，确保信息丰富且有条理。

提高信息密度: 练习在短时间内提供 1~2 个具体细节或例子支撑观点，避免泛泛而谈。一个常见的方式是在描述过程中刻意提及概念的下义词，比如考生在提及“在市区，公共交通很好用”的时候，可

以具体描述公共交通的下义词，比如地铁/轻轨等的好处。

口语流利度练习：通过限时答题、连讲练习减少填充词与长时间停顿。可用录音回放检查并改进。

6.3. 针对技术与考试策略的训练

熟悉界面与流程：在 ETS 提供的模拟练习中熟悉录音与时间提示等界面，减少考场因操作引起的紧张。

模拟真题与自动评分对齐：尽量使用 ETS 官方或获得 ETS 授权的在线练习资源进行模拟，以便理解自动评分对发音与信息保留的敏感度。

6.4. 教师与课程设计建议

把握技能而非只教“题型技巧”：训练重点应落在听力记忆、语音清晰度与短篇论述组织上；当前课堂上常见的“答题模板”或“背诵技巧”等内容，对于未来的托福口语帮助有限；在无法预判听力复述内容的前提下，背诵模板意义有限，同时在人工智能普及的考试里，背诵模板也比较容易被识别重复性高。所以，课堂应加入大量“听后复述”与“即时面试”实战练习，并且鼓励学生进行总结反思，积累自己无法顺利复述或顺利进行口头表达的语言并针对性复习，逐步提升表达能力。

数据驱动诊断训练：利用学生录音与自动评分反馈结合教师评分，针对共性弱点(如连读、停顿、信息省略)设计集中训练模块。

管理学生预期：从目前来看，新版托福考试更加关注语言使用基本功，应试技巧的作用下降。如果学生依然抱持当前版本考试下的“考前突击努力，用好方法技巧，高分唾手可得”的心态，会在学习和考试中面对困难。教师应说明新版考试特点，帮助学生树立合理的目标，制定长期的备考计划。

给予学生心理支持：在备考中，学生可能会面对一段时间内自身考试表现停滞不前的可能。教师应在需要的时候给予学生支持，帮助学生持续努力进步。

7. 结论

2026 年 1 月 21 日的托福 iBT 改革将口语测试向更“真实交流”与更“技术驱动评分”方向转变。对于中国考生而言，这既提出了对短时信息处理、发音可懂度与交际组织能力的新要求，也带来了更公平、可预测的评分机制与贴近日常学术情境的考查内容。备考策略应从单纯的模板训练转向技能训练：系统练习听后复述、强化发音与短篇表达、熟悉自动评分导向的反馈。教师与培训机构应及时调整教学设计，以技能为核心、以 ETS 官方样题与评分指南为基准开展训练；同时应当针对考生的心态和特点安排合适的目标、学习计划于考察节点。总体而言，适应这次改革、并以更贴近真实交流的训练方式准备托福的考生，将在未来更容易把语言能力转化为可衡量的成绩。

参考文献

- [1] archive.org.
<https://web.archive.org/web/20081204165457/http://www.ets.org/portal/site/ets/menu-item.1488512ecfd5b8849a77b13bc3921509/?vgnex-toid=e7242d3631df4010VgnVCM10000022f95190RCRD&vgnnextchan-nel=a4887f95494f4010VgnVCM10000022f95190RCRD>
- [2] ETS (2024) Official Guide to the TOEFL iBT Test. 7th Edition, Educational Testing Service.
- [3] TOEFL iBT Test Content (n.d.) www.ets.org. <https://www.ets.org/toefl/test-takers/ibt/about/content.html>
- [4] Deygers, B. (2020) Elicited Imitation: A Test for all Learners? Examining the EI Performance of Learners with Diverging Educational Backgrounds. *Studies in Second Language Acquisition*, 42, 933-957.
<https://doi.org/10.1017/S027226312000008X>

- [5] (2023) TOEFL Score Scale Update. <https://www.ets.org/toefl/china/toefl/score-scale-update.html>
- [6] ETS (2025) Teacher Frequently Asked Questions for TOEFL iBT® Enhancements. <https://www.ets.org/pdfs/toefl/teacher-faq.pdf>
- [7] ETS (2025) Speaking Rubrics. <https://www.ets.org/content/dam/ets-india/pdfs/toefl/speaking-rubrics.pdf>
- [8] 平顶山市人民政府. 明年河南新高考, 分数怎么算? 河南省教育考试院发布《等级赋分办法》及相关政策解读[EB/OL]. <https://www.pds.gov.cn/contents/22085/631319.html>, 2025-09-09.
- [9] 新东方在线. 托福听力-新托福听力真题[EB/OL]. <https://liuxue.koolearn.com/toefl/listen/>, 2025-09-09.
- [10] 新东方国际教育. 口语小助手[EB/OL]. <https://ielts.xdf.cn/202203/11745421.html>, 2025-09-09.
- [11] Yazdandoost, Z., Saleh, E.A. and Kafipour, R. (2014) The Relationship among Collocation Knowledge and Listening, Speaking, Reading and Writing Proficiency of Iranian EFL Learners. *Journal of International Scientific Publications: Language, Individual & Society/Language, Individual and Society*, **8**, 408-419.