

CAMBRIDGE

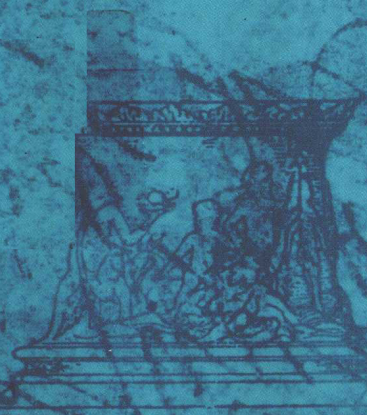
E n g l i s h T e a c h e r

Assessing Speaking.....

口语评价

Sari Luoma (芬) 著

ASSESSING
SPEAKING



CAMBRIDGE

E n g l i s h T e a c h e r

..... 剑 · 桥 · 英 · 语 · 教 · 师 · 丛 · 书

Assessing Speaking.....

口语评价

Sari Luoma (芬) 著
徐海铭 导读

外语教学与研究出版社 北京
FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS BEIJING
剑桥大学出版社
CAMBRIDGE UNIVERSITY PRESS

京权图字: 01 - 2010 - 6796

This is a reprint edition of the following title published by Cambridge University Press:

Assessing Speaking (ISBN: 0521804876)

© Cambridge University Press 2004

This reprint edition for the People's Republic of China (excluding Hong Kong SAR, Macao SAR and Taiwan Province) is published by arrangement with the Press Syndicate of the University of Cambridge, Cambridge, United Kingdom.

© Foreign Language Teaching and Research Press and Cambridge University Press 2010

This reprint edition is authorized for sale in the People's Republic of China (excluding Hong Kong SAR, Macao SAR and Taiwan Province) only. Unauthorized export of this reprint edition is a violation of the Copyright Act. No part of this publication may be reproduced or distributed by any means, or stored in a database or retrieval system, without the prior written permission of both Cambridge University Press and Foreign Language Teaching and Research Press.

本书版权由剑桥大学出版社和外语教学与研究出版社共同所有。本书任何部分之文字及图片,如未获得两社书面同意,不得用任何方式抄袭、节录或翻印。
只限中华人民共和国境内销售,不包括香港特别行政区、澳门特别行政区及台湾省。不得出口。

图书在版编目(CIP)数据

口语评价 = Assessing Speaking: 英文 / (芬) 卢奥玛 (Luoma, S.) 著;
徐海铭导读. — 北京: 外语教学与研究出版社, 2010.11
(剑桥英语教师丛书)
ISBN 978 - 7 - 5135 - 0338 - 9

I. ①口… II. ①卢… ②徐… III. ①英语—口语——语言教学—
教育评估—英文 IV. ①H319.9

中国版本图书馆 CIP 数据核字 (2010) 第 228593 号



悠游网—外语学习 一网打尽

www.2u4u.com.cn

阅读、视听、测试、交流、共享

封底刮刮卡, 获积分! 在线阅读、学习、交流、换礼包!

出版人: 于春迟
项目负责: 范海祥
责任编辑: 王海龙
封面设计: 袁璐
出版发行: 外语教学与研究出版社
社址: 北京市西三环北路 19 号 (100089)
网 址: <http://www.fltrp.com>
印 刷: 三河市北燕印装有限公司
开 本: 650×980 1/16
印 张: 17
版 次: 2010 年 12 月第 1 版 2010 年 12 月第 1 次印刷
书 号: ISBN 978 - 7 - 5135 - 0338 - 9
定 价: 34.00 元

* * *
购书咨询: (010)88819929 电子邮箱: club@fltrp.com
如有印刷、装订质量问题, 请与出版社联系
联系电话: (010)61207896 电子邮箱: zhijian@fltrp.com
制售盗版必究 举报查实奖励
版权保护办公室举报电话: (010)88817519
物料号: 203380001

出版说明

随着国家《英语课程标准》的实施以及《国家中长期教育改革和发
展规划纲要》(2010—2020年)的颁布,英语教育领域正在进行着全面
而深入的改革。新课程改革要求教师提高素质、更新观念、转变角色,
同时要求教师的教学行为发生相应的变化。研究学生如何学、教师如何
教、考试如何考和教师如何在教学实践中提高自我素质已经成为目前英
语语言教学研究的焦点。因此,教师培训工作的重要性和关键作用变得
越来越突出。为了更好地配合新课程改革,提高教师的自我素质,外语
教学与研究出版社精选并引进了剑桥大学出版社设计的这套“剑桥英语
教师丛书”。

这套丛书由当今国际语言教育界的权威人士 Gary Buck, John Read,
J. Charles Alderson 等编著,由国内英语教育界专家、学者撰写导读。这
些专家、学者在认真研读原著的基础上,结合中国外语教师的需求和现
状,有针对性地用深入浅出的语言对原书作了概括、总结与评价,有助
于英语教师及英语专业的研究生对原著的学习与理解。

本丛书的选题都是目前英语教师关注的教学实践中的热点专题。有
几本是关于英语教学中的测试方面的,如 *Assessing Vocabulary*, *Assessing
Listening*, *Assessing Speaking*, *Assessing Reading*, *Assessing Writing*; 有
关于语言教师开展自我研究的,如 *Psychology for Language Teachers*; 有
关于将信息技术用于语言教学之中的,如 *Language Learning in Distance
Education*, *Assessing Language through Computer Technology*。

本丛书从英语语言教学的理论和实践两个层面为英语教师的教学实践提供理论指导，并为英语教师的在职教育和终生职业发展提供丰富的资源，可以作为英语教师专业化教育与发展 and 英语专业研究生教育的教材，供在职英语教师和英语语言教学研究者，尤其是英语专业的研究生学习使用。

外语教学与研究出版社



语评价

导 读

如果您曾经观察过国内或者国际英语演讲比赛的场面，或者观摩过剑桥商务英语考试(BEC)考官培训的现场测试，可能会注意到这样一种现象：有些语音语调听起来不错的演讲选手或考生，最后的口语得分等级未必高。而有些语音语调听起来未必入耳的选手（尤其在许多中国教师听来），说话流利，虽有停顿和犹豫，但是使用了过渡性的小词如 *you see, kind of, that sort of thing*，且言之有物，口语表达给人自然、连贯的印象。其口语水平往往得到考官的高度认可。看来，口语优劣评判的尺度颇有讲究。

很自然地，你不禁要问：口语本身到底具有什么样的特征？我们需要设计出哪种类型的口试任务，才能诱导考生生产出话语，从而真实地反映出他们的口语水平呢？在我们采集了大量的口语样本之后，又需要设计出什么类型的打分量表来描述不同的口语等级？在设计好的打分量表中，应该包含哪些具体的衡量口语等级的指标？又该用什么样的描述词（*descriptors*）来简明、清晰而又确切地界定和描述这些具体指标？一旦确定好了打分量表和具体指标以后，又该如何实施口语评估以确保整体评估的效度和信度？

《口语评价》一书针对上述问题给出了具体、明确的回答。

1

全书结构和总体评价



全书结构

本书共八章。第一章导论部分扼要介绍口语评估领域包含的诸多环节,让读者从整体上把握口语评估涉及到哪些流程和具体要素。第二章从应用语言学视角讨论口语特征和口语技能的本质。其目的是廓清人们对口语和口语评估的一些误识,颇具新意。第三章详细讨论口语任务的设计、与口语任务相关的研究和实际操作。这是口语评估实际操作的开始。第四章详细介绍和论述口语打分量表的类型、开发、设计和操作时涉及的诸多问题,以及改进和研究打分量表的一些做法。本章的特点是以样本为例展开论述,有理论,有例证,具体翔实,毫不枯燥空洞。第五章有一定的理论性,作者从口语评估背后的几个语言学理论模型入手,讨论了当今国际上典型的把理论模型变成具有可操作性的口语评估设计概念框架的思路和实际操作方法,帮助读者清晰认识口语评估的不同理论构念。如果说,第二、第三、第四这三章的论述注重实践层面的操作,那么,第五章就是引领读者从实践层面进入理论层面,从认识上让

读者不仅知道口语评估该如何做，而且清楚为何如此做的理论依据。从第六章一直到第八章，作者提供了很多国际大型口语评估样本来支撑其关于如何开发口语评估的论述。但是，这三章中每一章的聚焦点都有所不同。第六章详细说明如何撰写测试规定。第七章集中讨论不同类型口语任务的特点和难度，并讨论如何设计开发这些任务。这一章与第三章呼应，但是比第三章讲述得更具体透彻。第八章从测试学角度阐述如何确保评估的信度和效度，在结束部分还论及未来提高口语评估信度和效度的可能手段以及研究方向。



总体评价

本书的适用范围较广。就英语学习者而言，他们迫切需要知道具体的口语测试标准，因为有了明晰、科学的标准作为参照，他们的学习才会有有的放矢，才能够针对明确的学习目标努力，并且时时用这些标准来作自我检查：自己的口语达到了哪个等级或者哪种等级？是否达到了某些职业对外语能力的要求？如果没有达到，还有多少个具体语言目标需要实现？

就语言教师而言，口语测试既是教学的一个组成环节，又是检测教学的一个有效工具或者手段。如果教师熟稔具体的口语测试标准，并且以此来检测课堂教学的效果，其教学努力就会更具针对性，其课堂语言也会更加真实地道，具体可体现在口语语法的使用、小词使用、停顿、犹豫、重复、修正、话轮和话语的衔接、话题的连贯一致等方面。尤为重要的是，他们会更加了解如何有效、科学地检测学生的口语水平，从而打出真实反映学生等级的口语分数。同时，教师还能明确今后该使用哪些具体教学手段来提高课堂口语教学效率。比如：不是一味地让学生练习说，而是要经常把学生的话语录下来，让学生转写自己的录音内容，然后反思，发现问题（包括语音语调问题、语法措辞问题、口

语性强弱问题等等),然后再操练提高。这种方法就是提高口语技能的“让学生自我反思”的口语教学法(self-reflecting of speaking skills)。其效果如何,有兴趣的教师不妨一试。

就语言测试开发者而言,他们需要知道口语测试的理论指导模式、具体的测试任务开发和设计的内在原理、具体过程,以及测试实施的程序等诸多要素。

就口语测试机构而言,它们需要对试题进行合理科学的设计和开发,保证测试的信度和效度,并对测试成绩作出合理权威的说明。本书对这些问题和技术细节都有翔实论述。

就口语培训机构而言,它们同样需要了解口语的本质和口语等级测量的依据。唯其如此,才能更好地把培训学员分成不同的班级,采取针对性措施,设计不同类型的口语任务,达到既定的教学目标。

尤其值得一提的是,就所有正规的专业口语考试而言(如英国剑桥大学的BEC考试、英国权威的雅思口试、中国的PET<公共英语等级考试>、上海的口译考试<中级、高级>以及其他各类入学考试中的口试),谙熟并深入了解口试的基本概念、设计过程、评分标准和根本性原理都是获取考官资格的必备要素。否则,考官的评估行为可能会流于经验和形式,缺乏科学测试的专业知识指导,很难具有权威性和说服力。

就测试研究人员而言,本书对口语测试的每个环节都有深入分析。更值得称道的是,本书把测试领域的最新动向和棘手难题呈现出来。研究测试的有心人肯定能从中发现许多可以深入探究的课题。举个简单的例子:从交际角度看,让考生与考生进行交流有哪些优点?又有哪些不足?这样的任务是否能够体现交际语言能力构念?如果不能,其不足体现在哪些方面?这样的任务有改进的可能吗?如果改进,应从哪些方面入手?考生对考生的影响体现在哪里?考生适应考生的能力具体体现在语言和交际能力的哪些方面?一个个看似微不足道的细节,其实对于测试研究和开发人员来说,都需要再三思量,反复检验,才能付诸应用。

此外,《口语评价》一书的表述深入浅出,举例引证丰

富权威（援引的例证都源自国际权威测试机构的口语测试项目），可读性和学术性兼具，概念阐释明晰，说理清楚翔实，以解决实际测试中遇到的问题为指归。综合起来看，本书对于从事口语测试设计和评估工作的读者是难得一见的好帮手。

读完本书，笔者印象最深的是：西方学者对于语言教育中的任何一个活动都从不马虎，从不含糊不清地一带而过，不会简单地、不加探究地重复同一活动若干年或若干代，而是耐心细致、锲而不舍地对每一个细节问题作深入分析，不断地进行实证性的尝试和探索，直至达到臻于完美的境地。可以说，正是这样的态度让他们深谙每种语言教育活动的长短优劣，让他们懂得如何综合各家之言，如何取长补短，从而把最完美的成果奉献给社会。也许，这就是西方人探究性、分析性和实证性思维之精髓所在。也许，这就是他们知识积累和增长速度如此之快的深层原因。本书虽论测试，但实乃西方人分析、探究和实证精神之典范之一。相信读者读后会自有体悟¹。

¹ 本文引用人名和出处来自原著的，不再附带参考文献。

2

各章节主要内容与导读

●●第一章 导论

本章扼要介绍口语评估的诸多环节。口语能力既是语言教学的目的，又是课程设计的内容。因此，如何评估口语能力也就很自然地构成了语言教学的重要组成部分。然而，口语评估非常棘手，原因之一是影响口语的因素很多；原因之二是我们期望口语分数具有很高的代表性，期望评估准确、公正，达到评估者预期的目的。本章列举了四种测试情形：一是交际式语言测试套卷的口试部分。这类测试很受年轻人欢迎，因为这类测试证书能够认证学生的语言等级。二是语言实验室测试。这是针对移民到英语国家的职业人士的测试。三是针对职校生的测试。测试题目基本上模拟的是工厂的实际工作任务和情景。四是针对成人的语言等级测试套卷的口试部分。这类测试与考生所学的特定课程无关，而是为了考查某一种语言能力。这类测试证书对于就业机构了解考生的口语能力有参考价值。

由此可以看出，口语评估过程包含了不少阶段。在评估的各个阶段，考官和考生互动，为下一个阶段做好准备。

因此，考生、交谈者、评分者和分数使用者都扮演着不同的角色。用文字描述的口语评估循环过程如下：

分数需要 ——► 系统开发（质量保证工作与规划和开发）——► 实施/表现（涉及任务、考生和考官）——► 评分/评估（涉及评分标准、口试表现和评分者）——► 分数使用

这个循环包含两个口语评估需要的互动过程。第一个过程是测试的实施或者测试的表现过程。在此过程中，考生彼此互动，或与考官互动，表现出口语能力的样本。第二个过程就是打分或者评估。在此过程中，评分者应用评分标准，评估考生的口语表现，最后得到口语分数，或关于口语能力的文字描述，或文字和分数两者兼而有之。值得一提的是，质量保证工作贯穿整个测试环节。质量工作主要包括构念的效度和信度。构念是描述测试目标的专门术语。口语构念就是指测试过程中所评估的特定的口语类型。评估质量主要体现在构念效度上。让测试具有效度涉及到口语测试循环每个阶段的每个过程和每个结果。构念信度就是要确保测试要给出一致、可靠的结果。关于这两个核心概念，第八章有专门分析。

●● 第二章 口语本质

本章内容涉及口语的语言学描述、口语的互动性及口语的社会性和基于情景的特点。所有这些视角都把口语看成是日常生活中不可或缺的部分。从这些视角出发，口语评估开发人员就能清晰理解口语能力到底意味着什么，然后再把这一理解落实到测试任务的设计和评分标准的制定过程中。测试中涵盖的语言特征愈能与口语本身的特征接近，测试结果就愈能有效地说明口语能力。这是对口语测试的根本认识。那么，口语有哪些具体特

征呢?

作者首先用一小节从发音(发音准确与发音的可理解性的关系,后者其实比前者更重要)、口语语法特征²(如思想单位 <idea units> 短小、说话人根据谈论主题采用连接词和重复等手段把思想内容组织起来、句子长度均值很低)、口语词汇、口误和错误³、加工和互动性⁴等几个方面描述了口语特征。其描述目的就是说明口语有其内在的语音、词汇、句法和语篇特征(让人觉得自然、有人际性)。与书面语不同,这些特征教师一般在课堂上不会传授。因此,在评估口语能力时,不能机械地根据书面语的语法标准来衡量。一般情况下,许多人评估考生的词汇能力就是看其对词汇的使用是否准确和具有表现力,是否多样化。但是,简单常见的词汇同样可以反映一个人的口语能力是否高级(Read, 2000)。此外,一些词组和短语为典型的口语词汇,使用这些词汇可以使口语显得更流利,因此,考查词汇的使用能力不能忽视这一特征。

此外,使用类指词(generic words)和特指词(specific words),到底哪种更能反映出口语能力呢?作者认为,类指词虽然不像特指词那么精确,但是它们在口语情境中完全能够为听者所理解,还可以使口头交际显得快捷、轻松。在外语学习语境中,学习者除了在课堂上使用外语,平时很少有机会使用所学语言,因此,使用特指词的可能性较小。鉴于此,不妨把“有效使用类指词”这一项纳入到打分量表中。

口语的另一个特征就是使用固定短语、填充词(fillers)和犹豫标志词(hesitation markers)。作者认为,在打分量

² 关于口语语法,参考Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (2000). *Longman Grammar of Spoken and Written English*. Beijing: Foreign Language Teaching and Research Press. 此书通过对语料库的分析得出关于口语语法的具体特征,可能是描述口语特征最为详尽的一本书。关于口语语法在口语教材中缺失的论文有Cullen, R. and Kuo, I-C. (2007). Spoken grammar and ELT course materials: A missing link? *TESOL Quarterly* 41 (2), 361-386。该论文对口语语法特征亦做了简明描述。

³ 关于口误的研究,参阅Garman, M. (1990). *Psycholinguistics*. Cambridge: CUP.

⁴ 关于会话特征,参阅Tsui, A. B. M. (2001). *English Conversation*. Oxford: OUP.

表中, 口试开发人员应该考虑考生是否在测试中策略性地、成功地使用了这样的小词⁵ (smallwords)。如果使用了, 应该给予褒奖, 因为小词的使用能够增加流利感 (Hasselgren, 1998)。有学者 (Nikula, 1996) 对外语学习者口语中小词的使用进行研究之后发现, 即使是高级的学习者, 使用“像口语般”的表达和语篇标示词的范围较之于本族人还是很窄, 而且, 这种差异不是因为个人或植根于文化的交际风格所致, 而是与语言能力相关。这充分表明, 使用“像口语般”的词在口语表现方面相当重要。鉴于此, 口语测试者应该意识到小词使用其实是说话者口语能力的标志之一。

口误 (slips) 和错误 (errors) 正常言语中包含了不少口误和错误, 比如: 注意力不集中导致发音错误、词汇使用错误等等。在母语交谈中, 出现这样的口误和错误, 倾听者一般都能原谅。但是, 二语和外语学习者的口误和错误可能说明他们知识不足。

加工和互动性 (reciprocity) 口语与书面语的不同体现在加工和互动性这两种情况上。口语的听说过程涉及到时间的短暂性; 解决这个短暂性的办法就是互动性, 即说话人和听话人彼此应答, 接替话轮, 共同完成言语交际任务。这其实也就是口语交际的社会性。

那么, 口语交际的社会性又是如何具体体现的呢? 作者用一小节介绍了“作为有意义言谈的口语”。言语交际的目的很多, 可以是消遣打发时光, 可以是交流信息……关键是言谈的参与者共同合作, 以言行事。言谈者和倾听者不断互换角色, 共同构建言谈事件, 分享影响言谈结果的权利。很自然地, 这就导致了言谈中语义的开放性 (openness of meanings in interaction)。在交谈和倾听的过程中, 语义并不总是清楚明确的。另外, 一句话除了其本身的含义外, 还能说明说话人对言谈话题的态度、他/她本人拥有的关于言谈话题历史的知识、他/她对即将发生的情况的看法等等。

⁵ Hasselgren认为, “口语中出现频率高的小词或者短语有助于保持口语流畅, 但是对于信息本身没有实质性的作用。”参见Hasselgren, A. (1998). *Smallwords and Valid Testing*. PhD Thesis. Department of English, University of Bergen, Bergen, Norway.

正因为言谈语义具有开放性,才使得言谈变得“险象环生”(被误解)。言谈者必须策略性地使用语言,或刻意追求表达模糊,或使用间接手段表明意图,或让语义留有解释余地。作者认为,能够保持言谈中的语义开放性说明说话者具有高超的交际能力,这也是言语交流自然性(naturalness)的一部分。口语评分者切不可忽视这一常常被遮蔽的口语特征。要达到这一评估目的,评估设计者要对评分者进行培训、指导,在评估量表中也通过描述语句来体现这一点。

在接下来的“口语使用变化”这一小节中,作者专门就口语用途做了分门别类的介绍。口语目的不同,评估方式和侧重点也就不同。如作为交往闲谈的口语和用于传递信息的口语。前者考查口语的社会性,即人际沟通,因此,测试应当让考生知道交谈的目的。后者着重信息传递,应该让听话人理解说话人传递的信息,因此,考查的是说话人在确立双方交谈的共同点、将信息分解为易于理解的语块、有逻辑地组织信息和展开信息、询问、重复、核查对方是否理解等诸多方面的表现。这些概念都应该体现在打分量表中,评分者凭借量表区分不同等级的考生。这一小节还介绍了 Hymes (1972) 经典的“言谈框架”(speaking framework) 的组成构件,即 S (situation, 情境)、P (participants, 言谈参与者)、E (ends, 言谈目的)、A (act sequence, 言谈行为次第)、K (key, 言谈口气)、I (instrumentalities, 言谈方式或渠道)、N (norms, 言谈规范)、G (genre, 言谈样式或者风格)。评估开发者在做初步计划时可以参照这一框架来描述测试构念。在随后的开发过程中,可以参照这一框架对不同的测试进行对比,这对于保证测试的公正性非常重要。比如:可以把测试中设计的交谈与考生在工作中的交谈情境进行比较,这样,测试开发者可以根据考生的实际考分来预测考生应对非测试情境的实际语言能力。

角色、角色关系和礼貌 语言使用过程中,言谈方式的一个重要方面就是礼貌原则 (Leech, 1985)。这一原则制约着言谈者的遣词造句。也正是这一原则,导致了人们常常在言谈中出现话语冗长、表达不直接、重复、甚至答非

所问的现象。按照 Grice (1975) 的分析, 日常言谈多数在表面上违背了会话原则 (即言谈量应充分而不冗余, 言谈的内容应只说你认为是真的东西, 言谈应与说话主题相关, 言谈的方式应简明、清晰、有条不紊), 但是遵循礼貌原则常常意味着违背会话原则。礼貌是一个很难评估的概念。一是因为评估过程中遵循的是模糊的原则而非明确的规则。二是因为它具有人际性和社会性, 而测试参与者之间的社会关系却是虚构的。因此, 在打分量表中常常用粗略的评语来评价考生言语中所体现的“礼貌”。

概括起来, 这一章是关于口语特征的理论铺陈, 作者一方面介绍口语的真实面貌, 另一方面颇有“纠偏”的意图。因为作者意识到, 如今许多语言教育者对口语能力持有许多错误的看法, 原因在于他们自己只是部分了解、甚至根本不了解口语本身的特征。作者如此大篇幅铺陈的要旨很直白: 测试者以其昏昏, 如何能给考生口语作出正确的评估? 关于这一点, 笔者认为, 对口语特征的误解和偏见, 甚至无视, 其实不仅仅体现在口语测试中, 而且还体现在口语教材的编写上。Cullen 和 Kuo (2007: 362) 认为, 口语教材编写中存在忽视口语语法特征的现象。可见, 如何描写和分析口语本身的特征, 其特征如何体现在口语测试和教材编写中, 以及如何体现在课堂教学的语言中, 是目前值得我国应用语言学界探讨的课题。

●● 第三章 口语任务

考生的口语能力最终落实在口试任务的完成上。因此, 恰当的口试任务的设计是口语评估的核心之一。口试任务, 说到底, 就是让考生置身于特定的语境中, 根据语境的要求来完成某一交际任务, 实现交际目的。那么, 什么是语境? 什么是任务呢?

语境 语境是语言使用的核心概念之一, 涵盖范围很宽, 包括语言使用情境的语言、物理、心理和社会等层面。

在实际交际活动中，语境指除了特定时刻语言产出之外的言谈情境中的一切因素，还包括言谈环境、语言使用经验、交际意图等等。语境还引导言谈内容。评估设计者通过任务⁶特性来控制语境特征，指导测试中的口语交流，因为明显的语境特征会影响考生的言语产出。语境不同，言谈就有所变化。因此，在口语测试过程中，语境的设计直接影响到语言使用的数量和质量。

任务 在语言学习情境下，任务的定义通常是基于语言的使用。Nunan 把交际任务定义为：

“课堂中的一项工作；这项工作涉及到学习者用所学语言来理解、控制、说话或交流，同时，学习者把注意力集中在语言意义上，不再是语言形式上。一项任务至少由一些语码输入、一个或多个关联的活动和步骤构成。语码输入指学习者加工的素材，可以是语言的（如广播新闻），也可以是非语言的（如一组照片），或者是“杂糅的”（如道路图）。此外，或明或隐的（更多时候是隐含的）任务包含了目标、教师角色、学生角色和情景。”（1993: 59）

这一定义有助于测试开发，因为测试开发者在设计口语任务时要包含语码输入、目标、角色和情景几大要素。

综合以上语境和任务定义，我们可以这样来定义口语任务：涉及说话人在特定场合使用语言以达到特定目标或者目的的活动。重点在目标或者目的之上。

初始任务设计 任务设计有几个环节。初始任务设计指关于测试该是什么形式、包含什么内容的决策。这个决策包括给考生哪些测试任务、如何具体安排测试实施。换言之，口语测试的理论构念在设计中有了具体操作化的体现。口试任务是大大而泛的概念，一定要具体化、可操作化。根据 Bygate（1987）的分类，口语任务可以细分为以下几种类型：

⁶ 关于任务型教学和任务设计的理论和实证研究可以参考Nunan, D. (2004). *Task-based Language Teaching*. Cambridge: CUP.和Plough, I. and Gass, S. (1993). Interlocutor and task familiarity: Effect on instructional structure. In G. Crookes and S. Gass (eds), *Tasks in a pedagogical context: integrating theory and practice*. Clevedon: Multilingual matters.