

汉语二语口语流利性的量化测评： 表达与认知维度的实证研究

陈家茜^{1,2} 高清辉¹

1. 厦门大学国际中文教育学院，厦门；

2. 上海青浦区协和双语学校，上海

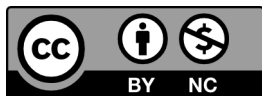
摘要 | 汉语二语学习者的口语流利性测评是二语教学与测试领域的重要研究议题。本研究构建了一种基于客观量化指标的汉语口语流利性测评方法，并对其信度与效度进行实证检验。研究采用计算机化测试（机考实验），从“表达流利性”和“认知流利性”两个维度对30名中高级水平来华留学生的汉语口语流利性进行测评。受试完成三项任务：看图说话（口语输出）、组句（词汇提取）和范畴判断（注意力控制）。本研究提取六项关键流利性指标，包括语速、平均语流长度、发音速度（表达流利性指标）以及组句反应时、反应时变异系数、注意力转移消耗（认知流利性指标），并系统分析其信度和效度。研究结果表明，表达流利性指标具有较高的内部一致性信度，并与受试者的口语成绩显著相关；认知流利性指标的整体信度较低，但组句任务的反应时与口语能力表现出显著相关性。回归分析进一步显示，表达流利性指标能够有效预测口语成绩，而认知流利性指标的预测作用相对有限。本研究验证了基于客观量化指标测评汉语口语流利性的可行性，同时揭示了其局限性，研究结果为汉语口语测试的客观化评价提供了实证支持，并对二语口语教学与评测体系的优化具有一定的启示意义。

关键词 | 口语流利性；认知流利性，国际中文教育；量化测评

Copyright © 2025 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

<https://creativecommons.org/licenses/by-nc/4.0/>



1 引言

在第二语言教学与评估领域，口语流利性被普遍视为衡量口语能力的重要指标，对有效交流至关重要。流利性不仅影响信息传递的效率，还直接关系到听者对讲话内容的理解。在日常交际中，人们常以“语言是否流利”作为评价二语学习者口语水平的依据，而流畅的口

语能显著增强交际效果，反之，不流畅的表达则可能引发沟通障碍。因此，如何客观、公正地评估汉语作为第二语言学习者的口语流利性，是当前语言测试研究的重要课题之一。

传统口语测试评分主要依赖人工主观判断，往往受到考官经验、情绪及其他非语言因素的影响，从而导致评分的不一致性和一定程度的主观偏差。近年来，为提

基金项目：本文为国家社科基金一般项目（项目编号：20BYY117）阶段性成果。

通讯作者：高清辉，博士，厦门大学副教授，研究方向：国际中文教育、教育测评。

文章引用：陈家茜，高清辉. 汉语二语口语流利性的量化测评：表达与认知维度的实证研究 [J]. 教育研讨, 2025, 7 (5): 561-569.

<https://doi.org/10.35534/es.0705109>

高口语测评的客观性和可重复性，国内外学者纷纷探索基于计算机技术的口语考试形式及自动评分系统。客观量化的评测方法不仅能有效减少人为误差，还能满足大规模测试的需求。在这一背景下，本研究以来华留学的汉语二语学习者对象，旨在建立一套基于客观指标的汉语口语流利性测评方法，为汉语口语考试的客观化评分提供实证依据。

本研究的独特之处在于将口语流利性细分为“表达流利性”和“认知流利性”两个维度，并分别进行测量和分析。Segalowitz (2000, 2010) 最早提出，口语流利性可分为言语输出层面的表达流利性（关注语速、停顿等口语产出特征）和认知过程层面的认知流利性（反映支撑口语产出的心理加工效率）。借鉴这一理论框架，本文不仅考察学习者实际讲话的流畅程度，同时评估其语言加工过程的效率。通过这种多维度的测评，本文期望深入揭示留学生汉语口语流利性的外在表现与内在机制，从而为二语口语能力的培养和测评提供新的理论视角和实践思路。

综上所述，本研究主要探讨以下问题：

(1) 如何界定并量化“口语流利性”，涉及哪些具体方面和测量指标？

(2) 表达流利性与认知流利性之间存在怎样的关系，它们各自与学习者口语能力水平的相关性如何？

(3) 能否通过客观量化指标有效测评二语者的口语流利性，并在多大程度上预测其整体口语能力？

2 文献综述及理论背景

2.1 口语流利性的概念与分类

口语流利性 (fluency) 在第二语言习得研究中占有重要地位，但其概念内涵一直存在争议。Starweather (1987) 将流利性细分为语音流利性、句法流利性、语义流利性和语用流利性四个方面。Lennon (1990) 则区分了广义和狭义的流利性：广义流利性指整体口语表达能力，包括发音、词汇、语法等方面的熟练程度；狭义流利性仅指说话的流畅程度，即话语的顺畅性和停顿的减少。在语言能力评估的实践中，一些语言标准或量表也对流利性进行了定义，美国外语教学委员会 (ACTFL, 2012) 明确提出流利性是指口语表达的流畅性、思想连贯性及语境适应性。

Segalowitz (2010) 进一步提出应将口语流利性分为表层的表达流利性 (utterance fluency) 和深层的认知流利性 (cognitive fluency) 两个维度。表达流利性主要指说话人在产出话语时的顺畅程度，包括语速、停顿频率与长度、重复与自我修正等话语特征；认知流利性则指支撑口语产出的认知加工过程的效率，涉及词汇提取、句法编码、语音产出以及注意力分配等心理加工过程。孔文和方洁 (2019) 进一步指出，认知流利性是二语口语流利性的核心要素，因为它反映了说话者将语言知识转换为实时言语输出的能力，包括加工速度、加工稳定性

和注意控制的灵活性，直接影响学习者语言知识到言语输出的效率与稳定性。认知流利性是指在言语规划、编码和执行过程中说话人提取词汇和使用语言资源表达意义、实现交际功能和追求心理目标的加工效率，包括加工速度、加工稳定性和注意力灵活性。

2.2 口语流利性的影响因素

二语口语流利性受到多种因素的影响，包括语言知识水平、心理认知能力以及交际情境等。一方面，学习者的词汇和语法等语言知识储备是流利表达的基础；另一方面，认知因素如工作记忆容量、注意力控制能力也与流利性密切相关。例如，金霞 (2012) 的研究表明，学习者的工作记忆容量越大，其汉语口语输出的流利度和准确度越高；同时，随着学习者汉语水平的提升，口语流利性对工作记忆资源的依赖逐渐降低。此外，赵宁 (2010) 的研究发现，学习者的人格特质（如内向或外向）会显著影响口语流利程度；焦虑水平较高的学习者往往在口语表达中表现出较低的流利性。

此外，母语迁移效应也是影响流利性发展的重要因素。Handley等人 (2024) 的研究表明，母语口语模式对初期二语口语流利性存在影响，但这种影响随着二语水平的提高而逐渐减弱。孔文和方洁 (2019) 进一步归纳出影响二语口语流利性的四个主要因素：认知流利性、交际动机、语言使用环境及认知体验，这些因素之间形成动态互动关系。其中认知流利性对第二语言话语流利性的影响最大，是整个动态系统中的核心。

任务类型对口语流利性也有显著影响。曹贤文、王文雅 (2021) 的研究发现，不同难度的任务会引发学习者不同程度的认知加工，进而影响其口语流利性表现。在低难度、熟悉内容的任务中，学习者能够更轻松地组织语言，表现出较高的语速、较少的停顿和修正；而在高难度、抽象复杂的任务下，其语言输出的流畅性往往会受到抑制，出现更多停顿、重复和自我修正等非流利现象。这表明任务的难度与认知要求会直接影响学习者的注意力分配和语言加工过程，最终体现在口语流利性的外在表现上。

2.3 口语流利性的提高方法

提升口语流利性的教学策略研究表明，重复输出练习与语块使用是提高二语流利度的有效手段。Maurice (1983) 提出的“4-3-2”训练法让学习者反复缩短时间讲述同一内容，通过逐次减少叙述时间，有效提升了学习者的语速并减少了停顿。邹晶 (2009) 也证实了通过多次复述练习可以在短期内明显改善口语流利度，但需要注意避免学生产生倦怠情绪。此外，加强语块（即常用搭配和句型）的习得和运用也有助于提升流利性。原萍和郭粉绒 (2010) 的研究发现，口语语料中使用更多惯用语块的学生，其语速和流利度指标更高。而随着技术的发展，语音助手与自学App也能提升流利性，Rahman & Tomy (2022) 研究了语音助手对英语流利性学习的影响，发现学习者在

使用语音助手进行练习后流利性有所提升。可见，教学中增加学生开口练习的机会，鼓励合理的重复输出和语块运用，能促进语言知识向流利表达的转化。

2.4 口语流利性的量化测评

近年来，随着口语流利性研究的深入，量化评估方法逐步受到重视，研究者们开始尝试将口语流利度转化为可测量的量化指标，并探讨这些指标与口语能力之间的关系。大量研究表明，时间维度的指标是衡量口语流利性的有效手段。常用的时间性指标包括：语速（每分钟输出的音节或词数）、发音/讲话时间比（讲话时间占总时间比例）、平均语流长度（两次停顿之间平均输出的音节数）、平均停顿时长等。郭修敏（2005）提出了包括语速、发音速度、平均语流长度等多个时间维度指标，这些指标在汉语二语研究中表现出了良好的信度与效度。相较之下，语法复杂度或词汇多样性等指标与流利性的关系没有那么直接。近年来，还有学者引入错误和修正维度的指标，如每百词（或音节）的修正频次、重复率等，以综合评估流利性。例如，柴省三和鲍杰（2022）运用人工神经网络分析韩国学习者口语，发现从14项流利度指标中筛选出的5项指标（包括百音节更改次数、T单位正确率、修正比例、语速等）对预测学习者口语水平具有较高的效力。

在认知流利性的测量方面，常用方法是通过特定实验任务记录反应时间。Segalowitz等人（2004）采用词汇判断、注意力转换等任务来测试学习者的认知处理速度，发现这些认知任务的表现与口语流利程度显著相关。胡伟杰和王建勤（2017）针对50名韩语背景汉语学习者的研究也表明，认知流利性指标相较于表达流利性指标，对口语能力具有更高的预测效力。然而，部分研究也表明认知指标如工作记忆容量与流利性的关系并不一致，提示认知流利性测评仍需进一步探索。认知流利性的量化研究起步较晚，目前成果有限，但已引起越来越多关注，被视为流利性研究的重要发展方向。

近年来，自动化口语流利性测评技术取得了快速发展。Liu等人（2023）结合语音声学特征与自动语音识别（ASR）技术构建了多模态评分模型，有效降低了主观评分带来的偏差。在汉语二语口语领域，自动评分与机器学习算法的应用也逐步兴起，体现了口语流利性测评向客观化、智能化方向发展的趋势。

综上，本研究在综合考虑表达与认知维度的基础上，选取关键流利性指标开展汉语二语口语流利性的量化测评分析，旨在填补现有研究中的不足，为汉语口语能力测评提供实证依据。

3 研究方法

3.1 研究对象

本研究选取某大学30名具有中高级汉语水平（HSK

5级及以上）的来华留学生作为研究对象。为确保样本的代表性及研究结果的有效性，在招募阶段充分考虑性别、年龄和母语背景的平衡性。最终有效参与的30名受试者中，男性18人、女性12人，年龄在18~28岁之间（平均年龄21.4岁），分别来自印度尼西亚、韩国、日本、越南、美国等13个国家。

3.2 测试设计

本研究采用了全程计算机化测试（机考）的形式进行数据采集。实验过程中利用E-Prime实验软件呈现实验刺激并记录被试的反应时间，采用Praat软件录音记录被试的口语输出。测试包括三个任务：（1）看图说话任务，用于测量表达流利性；（2）组句任务，用于测量词汇提取速度，即认知流利性的一部分；（3）范畴判断任务，用于测量注意力转换能力，也是认知流利性的另一部分。所有受试者在两天内完成所有测试。为保护参与者隐私，测试中以代号记录数据（如Sub1，Sub2，...）。

在正式测试前，研究者进行了小规模预试（5名留学生参与），根据预试结果对测试材料和操作流程进行了调整，以确保正式被试能够充分理解任务要求并顺利完成实验。

3.3 测试任务与材料

（1）看图说话任务

看图说话任务选用了德国经典漫画《父与子》中的六幅连环漫画作为材料（如图1所示）。被试需观察漫画情节后，用汉语叙述完整的故事，每位被试有1分钟准备时间用于浏览图片、构思故事，随后口头讲述并录音。选用连环漫画材料的优点在于其直观的视觉呈现，避免文字输入造成的理解差异。由于限定了主题和内容，每位被试面对的交际情境基本一致，确保了所有受试面对一致的语言输出情境，从而提高了测试结果的可比性。

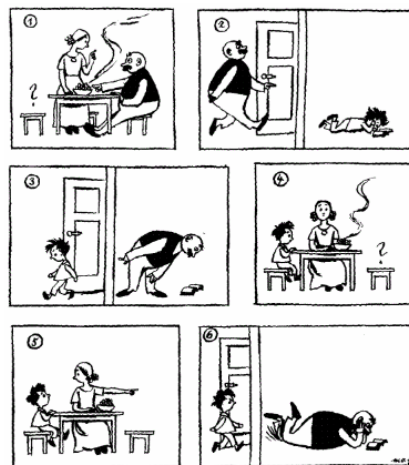


图1 看图说话使用图片

Figure 1 The pictures used for "Picture Description"

(2) 组句任务

组句任务借鉴了Segalowitz等人(2004)以及胡伟杰和王建勤(2017)的实验范式。材料选自汉语中级口语教材《发展汉语(第二版)中级口语I》中的句子。本研究挑选10个句子,将每个句子拆分为3个片段,并通过E-Prime将片段以随机顺序呈现在计算机屏幕上。被试的任务是尽快将这三个部分组合成一个符合汉语语法的完整句子,并大声朗读出来。E-Prime记录从句子呈现到被试开始朗读的时长作为反应时间。每个句子的词汇和语法均已通过预试确认为适合受试者语言水平,以排除知识因素的干扰。

(3) 范畴判断任务

范畴判断任务用于测试注意力转换与分配能力。本研究采用经典的“交替分类”(Alternating Categorization)实验范式。材料由12个汉语词汇组成,其中6个为时间类词汇,6个为频率类词汇(如表1所示)。时间类词汇再分为描述“过去”类的“以前”“昨天”“刚才”和描述“将来”类的“以后”“明天”“马上”两类;频率词分为描述“频率低”的如“有时”“几次”“很少”和描述“频率高”的如“经常”“多次”“一直”两类。E-Prime按照“T-T-F-F...”的顺序交替呈现时间词和频率词,即连续两个时间词后切换呈现连续两个频率词,以此类推。被试需要对出现的每个词语进行类别判断:看到时间词就按键判断该词属于“过去”范畴还是“将来”范畴,看到频率词则判断其属于“频率低”还是“频率高”范畴。E-Prime记录每次判断的反应时。由于词语呈现类别交替变化,被试在部分判断时无需转换注意焦点(如连续判断两个时间词,属于“重复”情境),而在部分判断时需在不同范畴之间切换(如从判断时间词切换到判断频率词,属于“转换”情境)。一般而言,被试在重复判断时的反应会比转换判断更快,这种差异即为“注意力转移消耗”(也称“转换代价”)。本研究以每位被试“转换”试次平均反应时减去“重复”试次平均反应时的差值来量化其注意力转移消耗——差值越大,说明被试在转换注意焦点时花费的额外时间越多,注意控制的灵活性越差;差值越小,则注意力转换效率越高。

表1 范畴判断测试材料

Table 1 Category judgment test materials

时间词		频率词	
描述“过去”	描述“将来”	描述“频率低”	描述“频率高”
以前、昨天、 刚才	以后、明天、 马上	很少、几次、 有时	经常、多次、 一直

3.4 测量指标

基于文献综述与理论分析,本研究选取以下六项客观指标进行口语流利性的量化评估:对于表达流利性,

从被试看图说话任务的录音中提取(1)语速(总音节数/总语料时长,反映平均说话速度)、(2)平均语流长度(总音节数/总语流段数,即两次停顿之间平均输出的音节数)以及(3)发音速度(总发声时长/总音节数,即剔除静默停顿后每秒发出的音节数)。需要说明的是,本研究将0.3秒及以上的静默视为停顿,用以划分语流段落。认知流利性方面,本研究记录(4)组句任务的平均反应时(10个句子完成的平均用时)、(5)组句反应时的变异系数(反应时标准差除以平均值,以衡量反应速度的稳定性)以及(6)注意力转移消耗(范畴判断任务中“转换”试次反应时减去“重复”试次反应时的差值,衡量注意力灵活控制能力)。上述六项指标中,(1)–(3)属于口语表达流利性指标,(4)–(6)属于口语认知流利性指标。

3.5 数据收集与处理

实验过程完整记录受试者在所有任务中的表现数据。看图说话任务的录音经转写后,由Praat软件提取总音节数、总语流段数、总静默时长等信息,再计算表达流利性的各项指标值;组句任务和范畴判断任务由E-Prime直接输出每次试次的反应时,并据此计算认知流利性的指标值。数据整理后使用SPSS 24.0软件进行统计分析,分析步骤包括:(1)指标的内部一致性信度(Cronbach's α);(2)各指标的描述性统计分析(均值和标准差);(3)指标间及指标与流利性总分的Pearson相关分析;(4)效标关联效度检验,采用受试者课程口语成绩作为效标;(5)回归分析,评估表达与认知流利性指标对口语成绩的预测能力,包括单变量和逐步多元线性回归分析。

4 研究数据与分析

4.1 信度检验

本研究对表达流利性和认知流利性两个维度的指标进行了内部一致性信度分析(Cronbach's α 系数)。分析结果如表2所示,表达流利性指标的内部一致性信度较高($\alpha=0.900$),而认知流利性指标的内部一致性信度较低($\alpha=0.269$),整体量表的信度系数为0.748,达到口语测试一般信度要求($\alpha>0.70$)。认知流利性指标信度偏低的原因可能在于各认知任务性质差异较大或被试对任务熟悉度不足,此外样本容量有限也可能影响结果稳定性。

表2 信度分析结果(Cronbach's α 系数)

Table 2 Reliability analysis results (Cronbach's α coefficient)

指标类别	表达流利性指标	认知流利性指标	全卷(全部指标)
α 系数	0.900	0.269	0.748

4.2 描述统计与指标对比

本研究统计了30名被试在表达流利性和认知流利性测试中的表现，记录了被试表达流利性指标（语速、发音速度、平均语流长度）和认知流利性指标（组句任务反应时、反应时变异系数及注意力转移消耗）的平均表现（原始数据略）。由于被试在平均语流长度上得分是正向的，

即该项得分越高，则被试口语流利性越好；被试在组句反应时上得分是负向的，即得分越高，则被试口语流利性越差。因此，在导出分数时将负向得分指标记为负数，即在该项上原始得分越高，转换后得分越低。在六项指标中，组句反应时、组句反应时变异系数和注意力转移消耗均为负向指标。转换后30位被试的平均表现如表3所示。

表 3 表达流利性指标与认知流利性指标描述统计

Table 3 Descriptive statistics of expressive fluency indicators and cognitive fluency indicators

编号	语速 (音节 / 秒)	发音速度 (音节 / 秒)	平均语流长度 (音节 / 个)	组句反应时 (秒)	组句反应时变异系数	注意力转移消耗 (秒)	总分
Sub1	2.7914	3.6909	8.1429	-11.5830	-0.2310	-0.2735	2.5377
Sub2	2.6193	4.4902	6.7778	-14.4875	-0.4539	-6.0517	-7.1059
Sub3	2.9071	3.8473	7.8824	-13.9451	-0.6956	-4.3094	-4.3133
Sub4	2.1854	4.1346	6.6875	-13.0661	-0.3915	-8.2354	-8.6854
Sub5	1.6464	3.4644	5.4231	-13.1101	-0.2849	-6.7995	-9.6607
Sub6	2.2488	3.7574	6.8000	-9.6142	-0.5510	-0.4095	2.2315
Sub7	4.7818	5.4868	14.0000	-4.6969	-0.4140	-7.1630	11.9947
Sub8	1.7637	3.0339	5.6364	-8.1312	-0.1731	-2.2037	-0.0741
Sub9	1.8016	2.7213	4.2000	-10.1314	-0.2903	-2.1238	-3.8224
Sub10	2.4067	3.7401	6.1111	-11.0953	-0.4325	-0.4660	0.2641
Sub11	4.3558	5.1757	13.6667	-8.8301	-0.2678	-9.4946	4.6056
Sub12	1.7504	4.3661	4.4483	-10.1963	-0.5972	-5.5224	-5.7511
Sub13	5.8892	7.2706	11.5714	-4.5634	-0.1893	-0.4656	19.5130
Sub14	3.7076	4.8670	8.9000	-5.4036	-0.4572	-3.9310	7.6827
Sub15	1.8987	3.2161	4.6429	-13.2331	-0.6608	-6.7736	-10.9099
Sub16	3.1890	4.5240	8.9000	-6.6444	-0.3994	-2.4777	7.0915
Sub17	1.6905	3.0759	4.3125	-15.0794	-0.4988	-3.7672	-10.2665
Sub18	0.9838	3.3411	4.1905	-16.8233	-0.5749	-1.7595	-10.6423
Sub19	1.0724	2.9497	3.0556	-20.3270	-0.5516	-6.4734	-20.2744
Sub20	3.0073	4.1911	9.7917	-11.9412	-0.6604	-3.1077	1.2809
Sub21	2.2512	3.2604	6.5000	-11.1401	-0.4104	-6.0626	-5.6015
Sub22	1.5720	2.9610	5.5455	-16.0042	-0.4412	-5.0445	-11.4114
Sub23	1.1254	2.1923	5.0833	-10.5287	-0.5257	-3.1233	-5.7766
Sub24	2.4495	3.7204	5.0000	-14.3003	-0.6122	-0.8653	-4.6079
Sub25	1.3989	2.3125	8.3750	-19.4671	-0.4592	-0.6738	-8.5136
Sub26	2.3204	2.6636	13.5000	-7.2549	-0.5133	-4.7562	5.9596
Sub27	2.6066	4.1735	6.0000	-6.4259	-0.3845	-1.9153	4.0544
Sub28	1.8084	2.9465	6.0435	-8.5719	-0.5361	-6.4187	-4.7283
Sub29	1.3595	3.1088	4.4231	-12.6021	-0.8892	-0.4318	-5.0317
Sub30	2.7116	4.0833	7.0833	-14.1123	-0.5970	-2.7627	-3.5937

数据显示，被试个体间表达和认知流利性存在明显差异，且两类流利性指标并非完全同步，有些被试口语输出流畅，但认知任务表现一般，反之亦然。这进一步验证了

口语流利性应从表达和认知两个维度分别考察的必要性。为便于对比分析，本研究将原始分数转换为标准化的T分数，各项指标的导出分数统计如表4所示。

表 4 各指标导出T分数统计

Table 4 Statistical T-scores of each indicator

编号	语速	发音速度	平均语流长度	组句反应时	组句反应时变异系数	注意力转移消耗	总分
Sub1	53.478	49.338	53.649	49.650	65.259	63.510	334.884
Sub2	51.908	57.119	48.919	42.359	51.114	41.345	292.765
Sub3	54.533	50.860	52.746	43.721	35.782	48.028	285.671

续表

编号	语速	发音速度	平均语流长度	组句反应时	组句反应时变异系数	注意力转移消耗	总分
Sub4	47.952	53.658	48.606	45.927	55.076	32.969	284.188
Sub5	43.038	47.133	44.225	45.817	61.840	38.477	280.528
Sub6	48.530	49.986	48.996	54.593	44.956	62.988	310.048
Sub7	71.625	66.821	73.945	66.936	53.649	37.082	370.059
Sub8	44.107	42.942	44.964	58.315	68.935	56.106	315.368
Sub9	44.453	39.899	39.986	53.294	61.498	56.412	295.544
Sub10	49.970	49.817	46.609	50.874	52.472	62.771	312.514
Sub11	67.740	63.793	72.790	56.561	62.925	28.138	351.947
Sub12	43.986	55.911	40.847	53.131	42.025	43.375	279.276
Sub13	81.721	84.187	65.529	67.271	67.905	62.773	429.387
Sub14	61.830	60.787	56.273	65.162	50.904	49.480	344.437
Sub15	45.338	44.715	41.521	45.508	37.989	38.576	253.648
Sub16	57.103	57.448	56.273	62.047	54.576	55.055	342.502
Sub17	43.440	43.351	40.376	40.874	48.268	50.108	266.417
Sub18	36.997	45.932	39.953	36.496	43.442	57.809	260.630
Sub19	37.805	42.122	36.021	27.701	44.920	39.727	228.296
Sub20	55.446	54.208	59.362	48.751	38.016	52.638	308.422
Sub21	48.552	45.147	47.956	50.762	53.877	41.303	287.597
Sub22	42.360	42.232	44.649	38.552	51.924	45.208	264.926
Sub23	38.288	34.749	43.047	52.297	46.564	52.578	267.523
Sub24	50.360	49.625	42.759	42.829	41.070	61.240	287.883
Sub25	40.781	35.919	54.453	29.859	50.783	61.974	273.771
Sub26	49.183	39.337	72.212	60.515	47.348	46.314	314.910
Sub27	51.793	54.036	46.224	62.596	55.520	57.212	327.380
Sub28	44.515	42.091	46.374	57.209	45.901	39.937	276.028
Sub29	40.422	43.671	40.759	47.092	23.494	62.903	258.342
Sub30	52.750	53.158	49.978	43.301	42.034	53.962	295.183

4.3 指标与口语成绩的相关分析

为检验指标与总体口语能力的关系，本研究进行了各指标与被试汉语口语成绩（标准化后）之间的相关性分析。结果如表5所示，三项表达流利性指标均与口语成绩呈显著正相关，其中语速与口语成绩的相关系数最高（约 $r=0.47$ ， $p<0.01$ ），发音速度次之（约 $r=0.44$ ， $p<0.05$ ），平均语流长度也达到显著相关水平（约 $r=0.40$ ， $p<0.05$ ）。这意味着说话速度越

快、连续语流越长的学习者，其口语成绩往往越高。相比之下，认知流利性指标中只有组句反应时与口语成绩显著相关（约 $r=0.48$ ， $p<0.01$ ），而反应时变异系数和注意力转移消耗与口语成绩的相关均不显著（ $p>0.05$ ）。可见，在本研究选取的认知指标中，词汇提取速度对整体口语水平具有一定的指示作用，而认知加工稳定性和注意控制灵活性尚未体现出明显的作用。

表5 各项流利性指标与口语成绩之间的相关分析结果（Pearson相关系数）

Table 5 Results of the correlation analysis between the fluency indicators and speaking scores (Pearson's correlation coefficient)

指标	语速	发音速度	平均语流长度	组句反应时	反应时变异系数	注意力转移消耗	口语成绩
语速	1	0.90**	0.79**	0.65**	0.37*	-0.06	0.47**
发音速度	0.90**	1	0.56**	0.54**	0.30	-0.08	0.44*
平均语流长度	0.79**	0.56**	1	0.55**	0.30	-0.18	0.40*
组句反应时	0.65**	0.54**	0.55**	1	0.38*	0.01	0.48**
反应时变异系数	0.37*	0.30	0.30	0.38*	1	-0.06	0.03
注意力转移消耗	-0.06	-0.08	-0.18	0.01	-0.06	1	-0.03
口语成绩	0.47**	0.44*	0.40*	0.48**	0.03	-0.03	1

注：** $p<0.01$ （双尾），显著相关；* $p<0.05$ （双尾），显著相关。

4.4 回归分析

为进一步探究两类流利性指标对口语能力的预测作用，本研究进行了多层次的回归分析。首先，以表达流利性和认知流利性总分为自变量分别进行单因素回归分析。表达流利性对口语成绩有显著预测作用（ $R^2=0.229$, $p<0.01$ ），而认知流利性指标的预测力未达显著水平（ $R^2=0.063$, $p>0.05$ ）。具体结果如表6和表7所示。

表6 表达流利性对口语成绩的单因素回归分析结果^b

Table 6 Univariate regression analysis results of expressive fluency on speaking scores

模型	R	R 方	调整后 R 方	标准估算的误差
1	0.478 ^a	0.229	0.201	7.534

注：a. 预测变量：（常量），表达流利性；b. 因变量：口语成绩。

表7 认知流利性对口语成绩的单因素回归分析结果^b

Table 7 Results of univariate regression analysis of cognitive fluency on speaking scores

模型	R	R 方	调整后 R 方	标准估算的误差
1	0.251 ^a	0.063	0.029	8.306

注：a. 预测变量：（常量），表达流利性；b. 因变量：口语成绩。

其次，将六个具体指标分别进行单独预测分析。结果表明语速、发音速度、平均语流长度、组句任务反应时均对口语成绩有显著预测作用，认知流利性的其他两个指标预测效果不明显。

最后，通过逐步多元回归分析进一步验证预测作用的稳定性。逐步分析模型的具体回归分析结果如表8所示。总体而言，以语速为主要预测变量时模型的预测效果最好，加入过多变量后模型预测效能并未显著增强，这可能受到多重共线性和样本容量限制的影响。

表8 逐步多元回归分析结果汇总^g

Table 8 Summary of stepwise multiple regression analysis results

模型	R	R 方	调整后 R 方	标准估算的误差
1	0.471 ^a	0.222	0.194	7.570
2	0.472 ^b	0.223	0.165	7.702
3	0.479 ^c	0.230	0.141	7.816
4	0.536 ^d	0.287	0.173	7.668
5	0.571 ^e	0.326	0.186	7.607
6	0.571 ^f	0.326	0.151	7.770

注：a. 预测变量：（常量），语速；b. 预测变量：（常量），发音速度，语速；c. 预测变量：（常量），平均语流长度，发音速度，语速；d. 预测变量：（常量），组句反应时，发音速度，平均语流长度，语速；e. 预测变量：（常量），组句反应时变异系数，平均语流长度，发音速度，组句反应时，语速；f. 预测变量：（常量），注意力转移消耗，组句反应时，组句反应时变异系数，发音速度，平均语流长度，语速；g. 因变量：口语成绩。

5 结论

本研究针对汉语二语学习者群体，构建并验证了一种结合表达流利性和认知流利性的汉语口语流利性量化测评体系。通过实证分析，获得了以下主要结论：

首先，在表达流利性方面，语速、平均语流长度和发音速度等客观指标能够稳定、有效地反映汉语口语表达的流畅性。这些指标之间具有较高的内部一致性，并与传统口语考试成绩存在显著相关关系，表明此类指标能够较好地作为汉语二语口语水平客观评估的有效手段。通过量化分析学习者口语输出的时间性特征，可以更为客观和精准地评估口语表达的实际熟练程度。

其次，在认知流利性方面，各项指标的表现存在明显差异。组句任务的反应时指标与学习者实际口语成绩表现出显著相关，证实了词汇提取和句子生成效率对口语流畅度的重要作用。然而，组句反应时变异系数和注意力转移消耗指标在本研究中未表现出显著的效度，说明当前认知流利性的测量方法仍存在一定局限性。这可能与测量任务设计的差异、被试对任务的熟悉程度以及样本容量限制有关。因此，未来研究中可通过优化任务设计、扩大样本容量，或结合更多类型的认知测量任务，以提高认知流利性指标的信度与效度。

最后，表达流利性和认知流利性作为口语流利性的两个独立维度，各自揭示了口语表现不同层次的特征。本研究发现，两维度之间的相关性并不显著，这意味着表达输出的流畅性和认知加工效率并非同步发展。因此，在口语能力评估过程中，综合考虑表达流利性与认知流利性两个维度，将能够更全面、更深入地诊断和评估二语学习者的口语能力。一方面，表达流利性指标可直观量化学习者的语言输出表现；另一方面，认知流利性指标则能揭示语言产出的内在加工效率和潜在制约因素。二者结合，更能体现口语能力测评的客观性与全面性。

综上所述，本研究证实了基于计算机化客观指标的汉语口语流利性量化测评方法的可行性与有效性，为汉语口语测评的客观化提供了实践参考和有益启示。表达流利性指标在口语测评领域具有较高的实用价值，认知流利性指标则为口语能力评估提供了新的研究视角与方向。未来研究应进一步扩大样本规模，对认知流利性测量方法开展更为深入和细致的探索，以不断完善汉语作为第二语言口语流利性测评体系，推动汉语口语测评理论与实践的科学发展。

参考文献

[1] 孔文，方洁. 第二语言口语流利性的研究现状与展望[J]. 中国考试，2019（8）：8-14.

- [2] 金霞. 工作记忆容量限制对二语学习者口语产出的影响[J]. 外语教学与研究, 2012(4).
- [3] 赵宁. 中国高中生内外向性格与英语口语流利性与准确性的相关研究[D]. 上海: 华东师范大学, 2010.
- [4] 邹晶. 4/3/2重复练习法与外语口语流利性的提高[J]. 天津市经理学院学报, 2009(6): 69-70.
- [5] 原萍, 郭粉绒. 语块与二语口语流利性的相关性研究[J]. 外语界, 2010(1): 54-62.
- [6] 郭修敏. 汉语作为第二语言的口语流利性的量化测评[D]. 北京: 北京语言大学, 2005.
- [7] 柴省三, 鲍杰. 基于人工神经网络的韩国汉语学习者口语流利性评价研究[J]. 汉语学习, 2022(5): 72-81.
- [8] 胡伟杰, 王建勤. 第二语言口语认知流利性对口语能力的预测作用[J]. 世界汉语教学, 2017, 31(1): 105-115.
- [9] Segalowitz N. Automaticity and attentional skill in fluent performance [M] //H. Riggenbach (ed.), Perspectives on fluency. Michigan: University of Michigan Press, 2000.
- [10] Segalowitz N. Cognitive bases of second language fluency [M]. New York, NY: Routledge, 2010.
- [11] Starweather C W. Fluency and Stuttering [M]. Englewood Cliffs, NJ: Prentice Hall, 1987: 11.
- [12] Lennon P. Investigating fluency in EFL: A quantitative approach [J]. Language Learning, 1990, 40(3): 387-417.
- [13] ACTFL. ACTFL Proficiency Guidelines 2012 [EB/OL]. (2019-03-22) [2025-04-08]. <https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012>.
- [14] Handley Z L, Andringa S J, Wang H. Exploring the Influence of Individual Speaking Style, Cross-Linguistic Transfer of Language-Specific Fluency Patterns, and Proficiency on Second Language Oral Fluency [J]. Journal of the European Second Language Association, 2024, 8(1): 34-47.
- [15] Maurice K. The fluency workshop [M]. Washington DC: TESOL Newsletter, 1983: 4, 29.
- [16] Segalowitz N, Freed B F. Context, Contact and Cognition in Oral Fluency Acquisition: Learning Spanish in At Home and Study Abroad Contexts [J]. Studies in Second Language Acquisition, 2004(26).
- [17] Liu J, Wumaier A, Fan C, et al. Automatic Fluency Assessment Method for Spontaneous Speech without Reference Text [J]. Electronics, 2023, 12(8): 1775.

Quantitative Assessment of Oral Fluency in Chinese as a Second Language: An Empirical Study on Expressive and Cognitive Dimensions

Chen Jiayi^{1,2} Gao Qinghui¹

1. School of International Chinese Language Education, Xiamen University, Xiamen;

2. Shanghai Qingpu District Xiehe Bilingual School, Shanghai

Abstract: The assessment of oral fluency in learners of Chinese as a second language (L2) is a key topic in second language teaching and testing. This study developed a measurement framework for Chinese oral fluency based on objective and quantifiable indicators and empirically examined its reliability and validity. Thirty intermediate-to-advanced international students learning Chinese participated in a computer-based test (CBT), which assessed oral fluency from two dimensions: expressive fluency and cognitive fluency. Participants completed three tasks: picture-based storytelling (oral output), sentence construction (lexical retrieval), and category judgment (attentional control). Six core fluency indicators were extracted, including speech rate, mean length of runs, and articulation rate (expressive fluency), as well as sentence construction response time, coefficient of variation in response time, and attention-shifting cost (cognitive fluency). Results showed that the expressive fluency indicators demonstrated high internal consistency and were significantly correlated with participants' oral proficiency scores. Although the overall reliability of cognitive fluency indicators was relatively low, the response time in the sentence construction task showed a significant correlation with oral proficiency. Regression analyses further indicated that expressive fluency indicators significantly predicted oral scores, whereas the predictive power of cognitive fluency indicators was limited. These findings support the feasibility of using objective quantitative indicators to assess L2 Chinese oral fluency and reveal some of the limitations of such approaches. The study provides empirical evidence for the objectification of Chinese oral testing and offers pedagogical insights for enhancing L2 oral fluency instruction and assessment systems.

Key words: Oral fluency; Cognitive fluency; International Chinese Education; Quantitative assessment