

中文二语学习者与母语者的线上与线下口语表现差异
之语言学分析

高雨茹

教育博士学位论文

香港教育大学

2024年5月



The Education University
of Hong Kong Library

For private study or research only.
Not for publication or further reproduction.

**Face-to-face Versus Virtual Communication: A
Linguistic Analysis of Chinese Language Usage Differences
Among Non-native and Native Chinese Speakers**

By

Gao, Yuru

A Thesis Submitted to
The Education University of Hong Kong
in Partial Fulfillment of the Requirement for
the Degree of Doctor of Education
May 2024



The Education University
of Hong Kong Library

For private study or research only.
Not for publication or further reproduction.

原创声明

本人高雨茹在此声明，本人是论文唯一作者，除致谢中注明的内容外，本论文中的材料均为本人原创作品。本人谨再声明，本人在撰写论文时遵守了大学关于学术诚信、版权和剽窃的政策和规定，并且没有提交本论文中的任何材料以获得该大学或其他大学的学位。

Statement of Originality

I, Gao, Yuru, hereby declare that I am the sole author of the thesis and the material presented in this thesis is my original work except those indicated in the acknowledgement. I further declare that I have followed the University's policies and regulations on Academic Honesty, Copyright and Plagiarism in writing the thesis and no material in this thesis has been submitted for a degree in this or other universities.

摘要

本文采用看图说话和多种访谈问题的实验任务，借鉴叙述分析中的一些语言参数，对 40 位中文二语学习者在线上与线下的口头表达能力进行量化分析。作为参照组，40 位母语为汉语者也参与了实验。实验不仅分析线上与线下的口语表现，还分析不同的产出顺序、性别和语言背景对语言参数的影响，分析多种口语产出任务对于语言参数的影响，如单一图片描述和长故事图片描述、涉及情绪表达的问题等。结果可见不同的沟通方式、口语产出任务等因素对各项语言参数有或多或少的影响，且呈现趋势有差异。实验得出的主要结论有：

(1) 线下产出任务时的口语表现整体优于线上产出任务，体现在重复次数和犹豫次数更少，总字数和句子总数更少但是信息传输效率更高；(2) 在对各语言参数进行主体间效应检验时，口语产出环境出现显著关系次数较多，说明环境对于说话者的口语表现影响较大；(3) 单一图片任务比复杂长故事任务的信息传输效率更高，名词率更高，但同时错误率也较高；(4) 涉及开心情绪问题的总字数、句子总数和错误次数都少于不开心情绪问题等。本文的实证研究对于国际中文教学中的听说教学以及口语评核都有实际的参考意义。

关键词：线上，线下，中文二语学习者，看图说话，访谈问题

Abstract

This paper employs experimental tasks of speaking through pictures and several interview questions, drawing on some language parameters from narrative analysis, to quantitatively examine the oral expression abilities of 40 Chinese second language learners both online and in person. A reference group of 40 native Chinese speakers also participated in the experiment. The experiment compares online and face-to-face oral performance and investigates how different orders, genders, and language backgrounds affect language parameters. It analyzes the effects of multiple interview tasks on linguistic parameters, such as single and extend story picture descriptions, as well as interview questions about emotional expression. The findings reveal that different communication styles, assignments, and other factors have an impact on various language parameters, and their trends differ. The main conclusions drawn from the experiment are: (1) The face-to-face assignments provide superior oral performance than online assignments, as seen by fewer repetitions and hesitations, fewer overall words and sentences, and greater communication efficiency; (2) When conducting the test of between-subjects effects on several language parameters, the interview mode produced more significant relationships, demonstrating that the interview mode has a greater impact on the speaker's oral performance; (3) A single image task has higher communication efficiency and noun rate than complex long tale task, but there is also a larger error rate; (4) Happy-mood interview questions have fewer words, phrases, and errors than non-happy-mood interview questions, etc. The

empirical research presented in this paper has practical implications for listening and speaking education, as well as oral evaluation in international Chinese language teaching.

Keywords: Online, Face-to-face, Non-native Chinese speakers, Picture description, Interview questions

致谢声明

今天，在我就要毕业，行将离开学校的时候，感慨万千。虽然很不舍，但是也要和母校说再见了。多年前，当我来到这里就读博士的时候，我由衷地感到找到了一个温馨的家，这就是香港教育大学。教大为我提供了良好的学习和研究环境，让我能够在这里追求自己的学术梦想。在撰写毕业论文的过程中，我经历了无数次的挑战和困惑，也收获了满满的成长和感悟。此刻，我怀着感激和崇敬的心情，写下这篇致谢，向在我求学道路上给予我帮助和支持的人们表达最诚挚的谢意。

首先我要感谢我的导师张凌老师，如果没有您，我的博士道路绝对不会有现在这么顺利。从入学之初，张老师就以渊博的学识、严谨的治学态度和高尚的师德深深地影响着我。在论文的选题、研究设计、数据分析以及撰写的每一个环节，张老师都给予了我悉心的指导和耐心的解答。张老师总是能够敏锐地指出我的问题所在，并为我提供富有启发性的思路和建议，让我在学术的道路上不断前行。正是因为张老师的言传身教，我才得以在学术研究中逐渐成长，学会了如何独立思考、如何解决问题。老师不仅是我的学术导师，更是我人生道路上的引路人，我们一起去吃烧烤，逛大美督，这些都是美好的回忆，我将永远铭记张老师的教诲和关怀。

同时，我也要感谢我的副导师延晶老师，延老师在我的研究过程中给予了

很多宝贵的意见和支持。延老师的细心和认真，让我在论文的细节处理上更加严谨和完善。感谢李贵生老师和研究生院的老师帮忙组织答辩；感谢中山大学的洪炜老师和美国乔治华盛顿大学的张航老师，作为外审专家在百忙之中抽出

时间审阅我的毕业论文，让论文更加完善和成熟。老师们的专业素养和严谨的治学精神让我深受启发，也让我对学术研究有了更高的追求和标准。感谢张海静姐姐在我的论文写作过程中给予的建议；王国玉学长、石柳学姐、张筱涵、徐净径和张晓瑞在我准备答辩 PPT 时提出的宝贵意见；褚燕南同学帮助我调整论文格式的细节。

最后，我要感谢我的家人和朋友们，尤其是我的母亲梁影春、爱人曹盐、姥爷梁学忠和姥姥曾毓学。在我攻读博士学位的过程中，大家一直陪伴在我身边，给予我关心和支持。是爱和鼓励让我能够坚持到底，实现自己的人生目标。博士学位的获得是人生中的一个重要里程碑，但我深知这只是一个新的起点。学校永远是我的家，张老师永远是我终生的老师。谨以此致谢学校、领导、张老师以及所有给过我教诲的老师和学友们。

高雨茹

目录

第一章 绪论	1
1.1 引言	1
1.2 研究目的	4
1.3 文献综述	8
1.3.1 口语产出环境探讨	9
1.3.2 影响因素	11
1.3.3 叙述分析	13
1.3.4 测量指标	17
第二章 研究方法	20
2.1 被试	20
2.2 实验任务与材料	20
2.3 具体测量指标	24
第三章 口语产出环境（线上线下）的差别探讨	31
3.1 图片描述任务	31
3.1.1 单一图片任务	31
3.1.2 长故事图片任务	32
3.2 问答任务	34
3.3 任务分类对比	50
3.3.1 长短图片描述任务之间的差异	50

3.3.2 正式问答和非正式问答之间的差异	53
3.3.3 回答不同情绪的问题之间的差异	57
第四章 产出顺序（第一次和第二次）的差别探讨	60
4.1 图片描述任务	60
4.1.1 单一图片任务	60
4.1.2 长故事任务	61
4.2 问答任务	62
4.3 任务分类对比	76
4.3.1 长短图片描述任务整合	76
4.3.2 正式问答和非正式问答整合	78
4.3.3 情绪相关问题整合	78
第五章 性别和语言背景的差别探讨及整体交互作用	79
5.1 图片描述任务	79
5.1.1 单一图片任务	79
5.1.2 长故事任务	81
5.2 问答任务	84
5.3 任务分类对比	100
5.3.1 长短图片描述任务整合	100
5.3.2 正式问答和非正式问答整合	103
5.3.3 情绪相关问题整合	105
5.4 交互作用	106

5.4.1 单一图片任务的主体间效应检验	106
5.4.2 长故事图片任务的主体间效应检验	116
5.4.3 问答任务的主体间效应检验	125
5.4.4 简单主效应分析	159
第六章 总论	161
6.1 数据结果分析	161
6.1.1 图片描述任务数据结果分析	163
6.1.2 图片任务主体间效应检验数据结果分析	167
6.1.3 问答数据结果分析	169
6.1.4 相似任务的发现	177
6.2 教学应用与启示	179
6.2.2 整体趋势	179
6.2.3 进一步研究	183
参考文献	187
附录	199
附录一 问卷信息表	199
附录二 问卷 A	200
附录三 问卷 B	201

缩写表

AI=Artificial Intelligence, 人工智能

L1=First Language, 第一语言

L2=Second Language, 第二语言

NC=Number of Characters, 总字数

NS=Number of Sentences, 句子字数

SFPs=Sentence Final Particles, 句末语气词

IW=Number of Informative Word, 信息词数

FW=Number of Functional Word, 虚词数

CE=Communication Efficiency, 信息输出效率

HSK=汉语水平考试



列表清单

表 1 单一图片线上线下任务数据的成对样本统计与检验	32
表 2 长故事线上线下任务数据的成对样本统计与检验	34
表 3 十四个口语产出问答列表	34
表 4 问题 (1) 线上线下数据的成对样本统计与检验	37
表 5 问题 (2) 线上线下数据的成对样本统计与检验	38
表 6 问题 (3) 线上线下数据的成对样本统计与检验	38
表 7 问题 (4) 线上线下数据的成对样本统计与检验	39
表 8 问题 (5) 线上线下数据的成对样本统计与检验	40
表 9 问题 (6) 线上线下数据的成对样本统计与检验	41
表 10 问题 (7) 线上线下数据的成对样本统计与检验	42
表 11 问题 (8) 线上线下数据的成对样本统计与检验	43
表 12 问题 (9) 线上线下数据的成对样本统计与检验	44
表 13 问题 (10) 线上线下数据的成对样本统计与检验	45
表 14 问题 (11) 线上线下数据的成对样本统计与检验	46
表 15 问题 (12) 线上线下数据的成对样本统计与检验	47
表 16 问题 (HAPPY) 线上线下数据的成对样本统计与检验	48
表 17 问题 (SAD) 线上线下数据的成对样本统计与检验	49
表 18 长短图片描述任务的成对样本统计与检验	50
表 19 图片描述任务线上线下的成对样本统计与检验	53
表 20 正式问答和非正式问答的成对样本统计与检验	55

表 21 问答任务线上线下数据的成对样本检验	56
表 22 问题 H 和问题 S 的成对样本统计与检验	57
表 23 情绪问题线上线下的成对样本统计与检验	59
表 24 单一图片产出顺序数据的成对样本检验	61
表 25 长故事产出顺序数据的成对样本检验	62
表 26 问题 (1) 产出顺序的成对样本统计与检验	63
表 27 问题 (2) 产出顺序的成对样本统计与检验	64
表 28 问题 (3) 产出顺序的成对样本统计与检验	65
表 29 问题 (4) 产出顺序的成对样本统计与检验	66
表 30 问题 (5) 产出顺序的成对样本统计与检验	67
表 31 问题 (6) 产出顺序的成对样本统计与检验	68
表 32 问题 (7) 产出顺序的成对样本统计与检验	68
表 33 问题 (8) 产出顺序的成对样本统计与检验	69
表 34 问题 (9) 产出顺序的成对样本统计与检验	70
表 35 问题 (10) 产出顺序的成对样本统计与检验	71
表 36 问题 (11) 产出顺序的成对样本统计与检验	72
表 37 问题 (12) 产出顺序的成对样本统计与检验	73
表 38 问题 (HAPPY) 产出顺序的成对样本统计与检验	74
表 39 问题 (SAD) 产出顺序的成对样本统计与检验	75
表 40 图片任务第一次和第二次的成对样本统计与检验	77
表 41 问题中不同产出顺序的成对样本检验	78

表 42 情绪问题不同产出顺序的成对样本检验	79
表 43 单一图片不同语言背景的样本统计和 ANOVA	80
表 44 长故事不同性别的样本统计和 ANOVA	81
表 45 长故事不同语言背景的样本统计和 ANOVA	82
表 46 问题 (1) 中语言背景的单因素方差分析及相关参数	84
表 47 问题 (2) 中语言背景的单因素方差分析及相关参数	86
表 48 问题 (3) 中语言背景的单因素方差分析及相关参数	87
表 49 问题 (3) 中性别的单因素方差分析及相关参数	87
表 50 问题 (4) 中语言背景的单因素方差分析及相关参数	88
表 51 问题 (4) 中性别的单因素方差分析及相关参数	89
表 52 问题 (5) 中语言背景的单因素方差分析及相关参数	89
表 53 问题 (6) 中语言背景的单因素方差分析及相关参数	90
表 54 问题 (6) 中性别的单因素方差分析及相关参数	91
表 55 问题 (7) 中语言背景的单因素方差分析及相关参数	91
表 56 问题 (7) 中性别的单因素方差分析及相关参数	92
表 57 问题 (8) 中语言背景的单因素方差分析及相关参数	93
表 58 问题 (9) 中语言背景的单因素方差分析及相关参数	94
表 59 问题 (9) 中性别的单因素方差分析及相关参数	94
表 60 问题 (10) 中语言背景的单因素方差分析及相关参数	95
表 61 问题 (11) 中语言背景的单因素方差分析及相关参数	96
表 62 问题 (12) 中语言背景的单因素方差分析及相关参数	97

表 63 问题 (HAPPY) 中语言背景的单因素方差分析及相关参数	98
表 64 问题 (HAPPY) 中性别的单因素方差分析及相关参数	98
表 65 问题 (SAD) 中语言背景的单因素方差分析及相关参数	99
表 66 图片描述任务中不同性别的报告和 ANOVA	101
表 67 图片描述任务中不同语言背景的报告和 ANOVA	102
表 68 问题中不同性别的报告和 ANOVA	103
表 69 问题中不同语言背景的报告和 ANOVA	104
表 70 两个情绪问题中不同语言背景的报告和 ANOVA	106
表 71 单一图片部分语言参数的主体间效应检验	107
表 72 长故事部分语言参数的主体间效应检验	116
表 73 问题 (1) 部分语言参数的主体间效应检验	125
表 74 问题 (2) 部分语言参数的主体间效应检验	131
表 75 问题 (3) 部分语言参数的主体间效应检验	132
表 76 问题 (4) 信息输出效率的主体间效应检验	134
表 77 问题 (5) 部分语言参数的主体间效应检验	136
表 78 问题 (6) 部分语言参数的主体间效应检验	138
表 79 问题 (7) 部分语言参数的主体间效应检验	139
表 80 问题 (8) 部分语言参数的主体间效应检验	142
表 81 问题 (9) 部分语言参数的主体间效应检验	144
表 82 问题 (10) 部分语言参数的主体间效应检验	145
表 83 问题 (11) 部分语言参数的主体间效应检验	147

表 84 问题 (12) 部分语言参数的主体间效应检验	151
表 85 问题 (HAPPY) 部分语言参数的主体间效应检验	155
表 86 问题 (SAD) 部分语言参数的主体间效应检验	158
表 87 单一图片任务的简单主效应分析	159
表 88 长故事任务的简单主效应分析	159
表 89 问答任务的简单主效应分析	160

图形清单

图 1 单一图片产出环境和语言背景在总字数上的交互作用	108
图 2 单一图片产出环境和语言背景在句子总数上的交互作用	109
图 3 单一图片产出环境和语言背景在句末语气词上的交互作用	110
图 4 单一图片产出环境和语言背景在重复次数上的交互作用	111
图 5 单一图片产出环境和性别在信息词上的交互作用	112
图 6 单一图片产出环境和语言背景在虚词上的交互作用	113
图 7 单一图片产出环境和性别在错误次数上的交互作用	113
图 8 单一图片产出环境和语言背景在名词上的交互作用	114
图 9 单一图片产出环境和语言背景在动词上的交互作用	115
图 10 单一图片产出环境和性别在犹豫次数上的交互作用	115
图 11 长故事产出环境和语言背景在总字数上的交互作用	118
图 12 长故事产出环境和语言背景在句子总数上的交互作用	118
图 13 长故事产出环境和语言背景在句末语气词上的交互作用	120
图 14 长故事产出环境和语言背景在重复次数上的交互作用	120
图 15 长故事产出环境和语言背景在信息词上的交互作用	121
图 16 长故事产出环境和语言背景在虚词上的交互作用	122
图 17 长故事产出环境和性别在错误次数上的交互作用	122
图 18 长故事产出环境和语言背景在名词上的交互作用	123
图 19 长故事产出环境和语言背景在动词上的交互作用	123
图 20 长故事产出环境和性别在犹豫次数上的交互作用	124

图 21 长故事产出环境和语言背景在句子长度上的交互作用	124
图 22 长故事产出环境和语言背景在信息输出效率上的交互作用	125
图 23 问题 (1) 产出环境和语言背景在总字数上的交互作用	127
图 24 问题 (1) 产出环境和语言背景在句子总数上的交互作用	127
图 25 问题 (1) 产出顺序和性别在错误次数上的交互作用	128
图 26 问题 (1) 产出顺序和语言背景在信息传输效率上的交互作用	129
图 27 问题 (2) 产出环境和产出顺序在总字数上的交互作用	129
图 28 问题 (2) 语言背景和性别在句子总数上的交互作用	130
图 29 问题 (2) 语言背景和产出环境在句末语气词上的交互作用	130
图 30 问题 (5) 语言背景和产出环境在句末语气词上的交互作用	135
图 31 问题 (5) 产出顺序和性别在重复次数上的交互作用	135
图 32 问题 (6) 产出环境和语言背景在重复次数上的交互作用	137
图 33 问题 (6) 产出环境和产出顺序在虚词上的交互作用	138
图 34 问题 (7) 产出顺序和性别在动词上的交互作用	140
图 35 问题 (7) 产出环境和产出顺序在犹豫次数上的交互作用	141
图 36 问题 (7) 语言背景和产出顺序在信息传输效率上的交互作用	142
图 37 问题 (9) 语言背景和产出顺序在错误次数上的交互作用	143
图 38 问题 (9) 产出环境和性别在名词上的交互作用	143
图 39 问题 (10) 产出环境和产出顺序在虚词上的交互作用	145
图 40 问题 (10) 语言背景和性别在信息传输效率上的交互作用	145
图 41 问题 (11) 语言背景和产出顺序在动词上的交互作用	147

图 42 问题 (12) 产出环境和性别在句子总数上的交互作用	148
图 43 问题 (12) 产出环境和产出顺序在重复次数上的交互作用	149
图 44 问题 (12) 产出环境和性别在信息词上的交互作用	149
图 45 问题 (12) 产出环境和性别在名词上的交互作用	150
图 46 问题 (12) 产出环境和性别在动词上的交互作用	150
图 47 问题 (HAPPY) 产出环境和性别在总字数上的交互作用	152
图 48 问题 (HAPPY) 产出环境和语言背景在句子总数上的交互作用	153
图 49 问题 (HAPPY) 产出环境和语言背景在信息词上的交互作用	153
图 50 问题 (HAPPY) 产出环境和语言背景在名词上的交互作用	154
图 51 问题 (HAPPY) 产出环境和语言背景在动词上的交互作用	154
图 52 问题 (HAPPY) 性别和语言背景在犹豫次数上的交互作用	155
图 53 问题 (SAD) 性别和语言背景在句末语气词上的交互作用	157
图 54 问题 (SAD) 产出环境和产出顺序在错误次数上的交互作用	157
图 55 问题 (SAD) 产出环境和产出顺序在信息传输效率上的交互作用	158

第一章 绪论

1.1 引言

新冠肺炎疫情对国际学生的流动和学习体验产生了持续而深刻的影响，教育也做出了相应的调整，伴随着 AI 的飞速发展，网络教学成为了教育改革的发展趋势。疫情影响下的新全球化观念，是一种螺旋上升的全球化趋势，而不是逆全球化（Cui, 2022）。在这种新趋势下，人们对全球社区理念的理解更加深入，互联网成为了教育发展的新动力。我们生活在一个后数码时代，在线上 and 面对面领域进行教学和学习已经成为常态（Wang & Canagarajah, 2024）。因此，研究线上教学和线下教学在国际汉语教育中的实际应用是非常必要的，从而将在线教育的发展作为一项长期战略。虽然疫情对于教育的影响在逐渐减弱，但是疫情让人们更加熟悉了线上教育的教学产出环境。传统的面对面教学已不再是唯一的学习产出环境，线上教育正以其独特的优势逐渐受到广大学者和学生的青睐。过去的经验告诉教育者们当不能只以一贯的方式教学时，就需要找到与学生互动和支持群体互动的方法，重新思考更加高效的教学方法。通过对于线上和线下不同特点的研究，我们可以取长补短进行混合式教学，打造高效课堂。线上教育有其独特的优点，如方便快捷的即时沟通，融入不同的科技教学手段等，是线下教学所不具备的。语言不仅仅是沟通的工具，更是文化和认知的载体。随着全球化的深入，越来越多的非母语者希望通过学习中文来深入了解中国的历史、文化和社会。Yu 和 Xu（2022）的研究探讨了在中国的一群南亚学生的在线学习经历。该研究通过问卷调查的方法，揭示了线上学习状态对于个

体对学习的认知与态度、线上学习评价有显著的中介效应。因此教师不仅要转变对外汉语教学观念，也要促进科技与中文教育资源的融合，有效推动混合产出环境教学的发展，帮助学生适应不断变化的学习环境。本研究收集了四十名中文二语学习者和四十名中文母语者的语言数据，并对收集到的数据进行量化分析，根据分析结果提出能够帮助语言教师教学之建议及理论。整体来看，疫情后的世界跟疫情前还是发生了变化，线上学习资源比疫情前多了许多

(Scheg, 2014)。尤其是在语言学习中，网络教学能够为二语学习者创造良好的语言学习环境。中文作为全球重要的第二语言之一，越来越多的学习者希望通过在线学习来提升口语能力。留学生们多了一个选择，他们可以考虑在自己的国家里面，通过线上学习的方式来提前学习中文，从而节约课程成本和往返交通费。本研究从语言学的角度来分析线上跟线下的区别，从而提出线上教学的改进建议，更好的引导学生进行线上学习以达到最优的学习效果。了解在线学习环境下口语表现差异的特点和挑战，同时也可以帮助促进跨文化交流和交际的发展，为学习者提供更好的语言学习体验。本论文涵盖了中文国际教育、线上教学、语言学等主题，为探索线上线下语言教学领域提供了一个平台，为今后推进线上语言教学提供了参考。

中国正在大力发展以资讯交流和科技为基础 (ICT-based) 的线上中文教育。1995 年出现最早的中文学习网站，主要包括日常用语和相关汉字的学习，学习者可以点击音频学习拼音和句子读音。紧接着，博客、维基百科和广播的学习方式出现，打开了线上中文互动学习的视窗。线上中文教育主要有两大优势：线上教学成果丰富和线上教育平台多样。线上学习环境的普及和发展使得

远程语言学习更加便捷，但其对中文二语学习者语言习得的影响，尤其是口语能力的影响尚不明确。许多线上平台发展飞速，如慕课（Massive Open Online Course）、中文帮（Chinese Bon）、Italki 等中文学习网站，传统的“学校”与“教室”的概念被重新界定，教与学随时随地都可以发生（康叶钦，2014）。陆俭明、李宇明和贾益民等（2020）预测未来线上教学将发展成为一种独立的教学样态，即独立的线下教学、线下线上结合的混合式教学、独立的线上教学并存。然而，发展线上中文教育仍有不少阻碍，首先，对于教师而言就是一个不小的冲击力。线上教育面对的最大问题不是科技而是方法论。经验丰富的老教师教学效果却不尽人意，因为传统的线下教学方法无法完全照搬到线上，学生之间互动和交流难度增大，授课效果难以保证。教学评估更是难上加难，尤其是人数较多的班级，课堂活动受限，老师无法及时对每一位学生进行评估。本论文的研究结果对于制定在线中文教育政策和教育实践具有参考意义。教育决策者和教育机构可以根据研究的结论来制定更加科学和有效的在线中文教育策略，提供更高质量的在线中文教育，以满足不同学习者的需求。教材方面，国际汉语相关优秀教材一直处于供不应求的情况，权威的电子教材更是难以统一。语言点如何设置以符合习得顺序、练习的呈现方式以及教材脚本的编写等问题都需要全面的评估甚至学术研讨才能确定（宋晖、谭紫格，2018）。本论文从语言学的角度分析中文二语学习者线上与线下的口语表现差异，对于国际中文教学，尤其是听说教学和口语教学方面都有实际的参考价值，为语言教育提供有效的研究支持，从而推动普通话教学方法的发展和教材的编写进程。

1.2 研究目的

随着线上教学的普及，中文二语学习者更多地选择网络平台进行学习，学者们也越来越关注线上教学，如网络学习推荐系统领域的年度发文量总体呈现增长趋势，特别是自 2020 年以来呈现出快速上升态势（谢浩然、陈协玲、郑国城和王富利，2022）。然而，线上环境与线下环境在许多方面存在显著差异，如交际方式、反馈机制、学习策略等，这些差异可能会对学习者的口语表现产生影响。对外汉语相关研究文献的主题大多集中在教学产出环境、设计和资源等领域，且以宏观分析为主，目前对国际中文线上口语学习的实证性研究成果较少（刘士娟，2018）。在现有文献中，也有较少研究说明线上与线下的中文口语表现的差异和语言特点。以往的研究大多比较了短期在线教学和线下传统教学效果的差异，但对于已掌握中文的学习者（HSK 5 及以上）线上线下中文使用之间的比较研究不足。国内针对汉语二语口语习得的研究目前处于徘徊上升态势，然而就汉语二语习得整体而言，有关口语方面的研究同样不足（郭江华，2023）。丁安琪等（2021）学者提出，2018 年汉语水平口语考试（HSKK）考生主要集中在中国境外，并且考生整体口语能力处于中低级水平。口语的实时性和互动性强，对于中文二语学习者和母语者来说，口语差异在实时交流环境中更为显著和复杂。与听力、阅读和写作相比，口语在不同场景中的变化和适应性更能揭示语言学习的特点，也能够直接反映学习者在真实语境中的语言运用能力和思维反应速度。在数字化时代，线上交流愈发频繁，线上线下口语环境和条件的差异逐渐凸显。这种差异对于中文二语学习者的语言习得路径和效果有着重要的影响，深入研究有助于我们更好地理解和改进教

学方法，提升学习者的口语能力。陈默（2012）指出，从研究方法的角度看，国内外对于语言流利度的测量较少把二语学习者和母语者进行比较，但是母语者的基线数据是非常重要的。通过对比母语组和二语组的口语表现，能够清晰地揭示二语学习者与母语者直接存在的差距和相似之处。这有助于精准定位二语学习者的不足之处，为教学和学习策略的制定提供明确方向。其次，通过对比了解母语者语言能力的形成机制和特点，从而为促进二语学习者向母语者水平靠近提供有益的参考和借鉴，有助于探究语言学习过程中的普遍规律和特殊现象。发现母语者和二语学习者在口语处理、语言运用习惯等方面的差异，能够丰富语言学理论，推动语言教学方法的创新和改进。本研究计划的目的是填补现有研究的空白，为教育者提供更多线上线下口语表现不同特点的参考，提出关于如何优化线上和线下教学方法的见解。研究结合了产出顺序、性别、情绪、任务复杂度等其他因素进行分析。产出顺序（第一次和第二次参与访谈：先后）能够反映出参与者在不同阶段的心理状态和准备程度，可能导致口语表现的变化，从而揭示出练习效应用于口语输出的作用。江新（2006）提出语言输入的次数是影响汉语习得的重要因素之一，频率效应受到学习者特征、语言形式特征的影响。性别因素在语言表现方面可能存在差异，不同性别的思维方式、表达习惯以及社会期望等都可能对口语表达产生潜在影响，进而在口语表现上呈现出不同特点。一项关于性别和线上教育的研究发现（Guiller & Durndell, 2007），男性和女性在使用个别语言变量方面相似，但是女性使用强调类词语比男性多。男性比女性在语言互动中做出更多较负面回答，女性则更多的表示同意和支持他人，并传递出更多的个人情感语言。研究结果表明，性

别差异能够从面对面授课延续到网络教育环境中。关于情绪因素的研究有 Pekrun (2002) 等人在研究中使用学术情绪问卷 (Academic Emotions Questionnaire: AEQ) 对情绪认知效应模型进行了测试。结果表明, 学生的情绪与动机、学习策略、认知资源和考试成绩显著相关。王希竹和金晓艳 (2021) 以留学生中文口语课堂和考试为特定情境, 得出口语表达流利度与焦虑情绪具有相关性的结论。黄韧等人 (2017) 探讨了不同情绪和年龄之间的交互作用对口语产生年老化的影响, 要求青年组和老年组被试在不同的情绪启动条件下对图片进行命名, 结果表示随着年龄的增长, 消极情绪对口语产生的抑制作用更强。任务类型的多样性和特定要求会促使学习者运用不同的语言策略和技能, 进而影响口语的流利度、准确度和复杂度。在任务型语言教学中, 任务的设计特征对二语口语产出的影响是重要研究方向之一 (陈默、陈青, 2023)。以三十名初中级汉语二语学习者为研究对象的一项实验表明, 高复杂度任务的句法复杂度、词汇多样性和词汇难度明显更高, 但是低复杂度任务中的语言流利度、词汇和语法准确性更高 (吴继峰、高敏和赵晓娜, 2022)。综合考虑, 这几个因素对于理解和改善中文二语学习者的口语表现具有核心作用, 相比其他因素, 它们的影响更为直接、显著且具有较强的可观测性和可分析性。

在传统的线下教学环境中, 教师可以直接与学生面对面交流, 及时感知学生的学习需求和困惑, 并给予适当的指导和反馈。而在线上教学环境中, 由于缺乏面对面的实时交互, 教师与学习者之间的沟通会受到一定的限制。语言学习者线上学习的需求日渐增长, 但教师若完全将线上教学转变成线下教学的

替代产出环境，则是一种误区（岳海燕，2021）。线上教育的科技运用，其很大一部分取决于教师的态度和能力。疫情期间，教师的语言教学计划必须满足远程教育的需求，教师们将现有的教学材料直接转移到在线平台上，然而这些材料在设计时并没有考虑到远程学习的效果（文秋芳、杨佳，2020）。线上线下因教学产出环境的不同，授课方式也应该有所改变。新冠肺炎疫情引发的教育紧急情况表明，世界需要一个具有灵活性和适应力的教育系统，使教育者们够应对一些可能出现的紧急情况。在科技迅速发展的电子时代中，e-learning 已经成为语言教学的重要教学途径。有学者使用 Web of science 和 Scopus 作为数据库，对于高等教育中教师和学生的科技运用能力进行了研究（Zhao, Llorente & Gómez, 2021）。该研究提供了 2015 年至 2021 年高等教育中科技运用能力分析报告，包括评估科技运用能力的维度、研究方法、结果和局限性，主要研究结果表明大多数学生和教师都具有基本的科技运用能力。针对线上中文教学目前发展情况，本论文提出了以下研究问题：

1. 从语言学角度分析，中文二语学习者和母语者，线上与线下的口语表现差异体现在哪些方面？
2. 产出顺序、语言背景和性别因素是否影响中文二语学习者和母语者的口语表现？以上三个因素是否与口语产出环境有交互作用？
3. 不同形式的口语产出任务是否影响中文二语学习者和母语者的口语表现？如图片描述任务的长短、正式和非正式问题和表现不同情绪的问题。

本论文的研究假设是：中文二语学习者和母语者，在线上与线下的口语表

现之间有差异 ($p \leq 0.05$), 如线上组错误更多或信息词输出效率更低; 产出顺序、语言背景和性别因素同样也影响被试的口语表现, 并且与口语产出环境有交互作用; 不同形式的口语产出任务会影响中文二语学习者和母语者的口语表现。通过深入分析说话者线上和线下的口语表现差异, 这有助于了解非母语者学习中文口语的过程以及他们在不同环境下的表现。分析可能影响口语表现的各种因素, 包括语言能力水平、环境等, 通过对这些因素的研究, 可以更好地理解为什么会出现差异。此外, 本研究计划还希望为未来的教学实践提供实证依据。通过对线上和线下教学的综合研究, 我们可以为语言教育者提供关于如何结合两种产出环境的实践建议, 可以为在线学习平台提供有针对性的建议, 帮助提升中文二语学习者在线上学习环境中的口语表现, 提高他们的交际能力。通过实现以上研究目标, 将增进对中文二语学习者语言习得过程的理解, 为改进在线语言教育提供实践性的建议和策略, 进一步促进中文二语学习者口语能力的发展。

1.3 文献综述

重要文献被分为四个部分进行分类阐述, 分别是线上线口语产出环境探讨、其他影响因素、叙述分析理论和测量指标的选取。根据本文提出的研究问题, 主要目的之一是对比线上与线下不同语言环境中汉语口语表达能力的差异。尽管过往的文献中不乏对线上与线下教学模式的广泛探讨, 然而, 研究结论呈现出分歧与多变性, 缺乏一个综合性的视角来整合这些发现, 本研究能够为理解线上与线下汉语口语教学的效果提供更为全面的认识。为了达成这一目标, 研究选取了一系列变量作为分析的核心要素, 这些变量是基于先前文献中

已确立的重要影响因子。具体而言，产出顺序（即二语者在特定时间段内的第一次和第二次口语表达）、语言背景（母语者与二语者之间的差异）、性别特征（男性组与女性组之间的差异），以及任务多样性（长短图片描述任务、正式和非正式问答和不同情绪问答）构成了实验的四大核心维度。通过对这些变量的细致考察，旨在揭示它们如何单独及相互作用于线上与线下两种环境下的口语能力发展，从而为优化汉语口语教学策略提供实证依据。本研究主要采取叙述分析（*narrative analysis*）中的部分语言学测量指标（*linguistic measures*）来测量母语者和二语学习者线上与线下的中文口语表现差异，再收集语音数据后对其进行量化分析。通过实验设计与数据分析来填补现有知识体系中的空白，为汉语口语教学领域的实践创新做出贡献。此外，本研究重视理论与实践的结合，力求将所得结论转化为具有实际操作意义的教学建议，以促进汉语作为第二语言的教育实践，特别是在快速发展的数字化教育时代背景下，为线上汉语教学的设计与实施提供更具针对性的指导。

1.3.1 口语产出环境探讨

虽然在线教学的飞速发展带来了更多的学习机会，但同时也给学生和教育工作者带来了巨大的挑战（Giannikas, 2022）。创新性多层次、多元化的线上文化交流模式目前较少，需要积极探索其更多的发展空间（顾明远、滕璐，2020）。从张国雪（2023）的实验统计数据来看，线下汉语课堂中师生的身体动作行为时长远高于线上课堂，表明在虚拟空间中身体动作的展示受限。教师们需要充分了解线上教学的特点才能够更好的进行线上教育，而不是将传统线下

教学方法原封不动移往线上。Stickler 和 Shi (2013) 发现线上学习期间教师意图和学生感知之间的不匹配不仅会导致焦虑, 甚至会导致学生完全放弃在线语言学习体验。姜永超 (2016) 提出, 深度融合的线上汉语国际教育在宏观上应该做到融合教学理论、信息技术与汉语国际教育三者。Chen 和 Chew (2021) 通过访谈和焦虑程度调查问卷的方式, 研究了 40 个中国学生在四周内二语学习时的线上和线下口语表现, 并提出外语学习焦虑 (Foreign Language Anxiety: FLA) 是影响学生口语表达能力的重要因素。调查发现学生在面对面交流时的焦虑程度比线上高, 但是更多学生选择进行线下交流。赵雷 (2021) 提出, 如果成人学习者能够理性认识到如何进行高效学习, 将会取得事半功倍的效果。线上或线下的二语学习各有其优缺点, Firat 等 (2019) 在研究中发现, 当大学生更多地使用电子学习材料时, 学习成绩会显著提高。吴爱梅 (2023) 提出线上口语练习由于缺乏面对面社交属性, 教师常依靠音视频等媒体作为互动的媒介, 而忽视了由于环境变化带来的话题多元性的开发。因此在二语学习中的线上口语交流要以学生为中心, 提高互动性, 学生需要做到多说多练。郑艳群 (2001) 认为, 现代教学技术若与现代媒体技术结合, 建立新的师生和谐关系, 就能够创造出比传统教学方法更好的教学效果。陆俭明等 (2020) 发现对于一些日本的中文学习者来说, 网课的互动效果超出了线下课堂, 同学们更愿意在线上互动。其原因是线上交流具备一些线下交流没有的优势, 如可以使用文字、表情和语音等形式发言, 这让部分内向的学生减弱了展示自我的犹豫, 增强了表达的信心。Zhao 等学者 (2020) 同样认为线上交流有更轻松的氛围, 允许害羞的学生在聊天框中写答案而不是在全班面前发言能够增加他们的安全

感。但是 Na (2022) 的观点是线上教学交互性差，即使进行更具针对性的一对一线上中文学习，但是相比线下中文学习来说没有真实的交际空间会导致学习氛围不佳，教师在教学中也会变得更加被动。谢婷和李京南 (2021) 结合语言输入假说和输出假说的理论，采用“线上输入，线下输出”的精确混合教学进行了英语口语训练实践。研究结果表明，这种新的教学产出环境不仅可以提高学生的热情，而且可以通过多产出环境的在线语言输入和面对面课堂的口语训练，来提高他们的口语输出能力。中文二语口语学习也可以借鉴此英语口语训练产出环境，结合线上线下口语表达特点，通过多产出环境的混合学习来提高学习效果。

1.3.2 影响因素

Boyle (1987) 指出性别对于语言的产出有或多或少的的影响，女性的整体语言能力优于男性，但在听力词汇的理解方面男性优于女性。同样的，一项关于留学生中文口语产出的流利性实验表明，停顿时长受到留学生语言水平和性别因素交互作用的影响，研究使用的具体测量参数有：语速、平均语流长度、停顿、重复和自我纠正 (陈默, 2012)。该实验结论之一是汉语中级水平的男性组停顿时长比女性组长，汉语高级水平的男性组停顿时长比女性组短，母语组中男性和女性的停顿时长没有差异。Luo (2014) 调查了 257 名在美国学习汉语大学生的汉语口语焦虑程度及其相关因素，并表明相当多的学习者在说汉语时焦虑程度较高：方差分析 (ANOVA) 的结果表示性别对口语焦虑程度有显著影响，女性组在说汉语时比男性组更加焦虑。但是该研究中有 110 名女性被

试，147名男性被试，没有控制性别分布的数量，并且年龄跨度大：15岁至59岁。为了帮助减少汉语学习者的焦虑，建议对高度焦虑的学习者进行访谈研究，探索焦虑的来源。

频次是语言学习的根本性因素 (Ellis, 2002)。本研究中的每位被试会参与任务两次，口语产出顺序也是实验关注的因素之一。文秋芳 (2003) 表示学习语言的过程依靠的是输入类型频次和练习频次。王植洵和张锦坤 (2017) 提出，学习条件 (重复学习、提取练习) 与最终测验的时间间隔 (即时测验、延迟测验) 之间存在交互作用。杜娟 (2019) 研究了二语学习者的六次口语输出任务，结果表示随着练习次数的增加，二语学习者的口语流利性、准确性和词汇复杂性都有或多或少的提升。尹振宇 (2016) 提出，频次的增多可以增强学习者对于语言准确性的关注度，促使学习者将注意力由内容转向形式，从而加深记忆巩固学习内容。但是，也有学者提出练习效应有时对于二语习得的帮助并不明显：例如有部分在青春期过后开始学习二语的人，即便接触二语的时间长于母语，可是讲话仍然带有严重的母语口音，句法里依旧存在一些难以更改的差错 (Hulstijn, 2002)。在学习者不断摄入语言知识的过程中，输出次数与输入次数发挥着不同的作用，但是次数并非越多越好，而是要探寻到发挥练习效应的最优作用点，从而获取最佳的学习成果。在二语教学中借助练习效应，能让课堂学习更为高效，助力语言教学改革推进。

在任务多样性方面，Derwing, Rossiter, Munro 和 Thomson 的研究 (2004) 表示二语者在口语表达任务：图片描述中的流畅性明显低于独白或对话任务，

说明二语者在口语任务中的流利度因任务而异。这一发现与 Foster 和 Skehan (1996) 的发现相似，他们在比较三个不同的口语表达任务时发现，图片叙事任务比信息交换或协作决策任务包含更多的犹豫和停顿。徐琴芳 (2005) 同样分析对比了不同类型的三种口语任务：看图说文、话题讨论和个案研究，发现二语者在完成看图说文任务时准确率最低。但是，在探讨多样任务中的二语学习者口语表现方面，以英语为目标语的研究很多，但是以中文为第二语言的研究很少。并且，许多中文口语测试只涉及问答面试，这并不能全面的展示学习者的整体口语水平。个体差异也应更仔细地研究，考虑可能影响口语水平发展的变量，如学习者的母语、语言输入程度、互动机会和沟通意愿等。

1.3.3 叙述分析

叙事研究可以收集和分析人们讲述的描述经历并为其提供对应的解释。通过量化的数据，我们可以了解到母语者和二语学习者在线上 and 线下的中文表达能力是否有不同。生活是由叙述片段组成的，并且叙述片段具备不连续性

(Clandinin & Connelly, 2000)。定量叙述分析基于文本的语言特性进行研究，语言特性来源于长期努力发现这些特性的文学和语言学传统 (Franzosi, 2010)。

Andrade 和 Andersen (2020) 提出数码故事语法 (Digital Story Grammar: DSG) 是一种将叙述理论与数码文本分析相结合的方法论，通过一种算法来识别由主语、谓语和宾语组成的叙事单元。该方法提供了新的见解来识别叙事中的模型，并且研究这些模型如何与性别或社会阶层等社会结构相关。Güngör

(2021) 通过对考试设计者的叙述性询问，探讨了新冠肺炎期间在线语言能力

测试的实施。该研究以土耳其高等教育为背景，考察了考试设计者对测试适应性的发展、技术使用以及未来语言测试发展的描述。数据基于三维空间叙事结构：互动性、连续性和形式进行分析，具体分别是分析个人和社会的互动以更好地了解测试设计者的观点；连续性涉及他们的测试经验、源源不断的需求以及在线测试的发展方向；了解在线测试实施规则和条例的新形式。Poole

(2021)以叙事探究和关系伦理学为基础，通过和电子邮件通信，探讨了上海国际学校教师的生活经历和身份认同感；通过数据收集、转录和分析，说明了叙事探究如何推动整个研究过程。Boles (1997)在其研究中提出了对话语篇分析法(CDA: Conversational Discourse Analysis)，通过测量口语输出质量和说话者展示的交流特征，如输出词汇量总数和自我纠正次数，去测量四位失语症患者的交流过程。Saffran等学者(1989)的研究通过对叙事言语进行定量分析来分析失语症患者的口语输出。该研究设计了三组实验参与者，分别是失语症患者组、语言不流畅被试组和正常对照组，研究发现虽然在大多数测量参数上失语症组可以被轻易区分，但失语症组和语言不流畅组中的语言结构复杂性都很低。Menn、Ramsberger和Helm Estabrooks(1994)提出了另外一种分析语言叙事的快速方法，称之为语言沟通测量法(LCM: Linguistic Communication Measure)。语言沟通测量法设计的目的是辅助，而不是替代标准化考试。语言沟通测量法在之后被发展为粤语沟通测量法(CLCM: Cantonese Linguistic Communication Measure)(Kong & Law, 2005)。粤语沟通测量法使用图片叙述作为媒介，反映随着时间的推移语言叙事的变化过程。该方法使用了八个参数，反映了信息传递的数量、速度和效率，句子的语法和清晰度，以及语言输

出中的偏误和词汇多样性。CLCM 旨在记录或评估实验者语言能力，而不是交际或社交能力（Kong & Law, 2009）。为了进一步验证 CLCM，研究者使用判别函数分析（DFA: discriminant function analysis）重新分析了语言数据，以了解 CLCM 中选取的参数是否能够对这三组实验参与者进行分类，并预测实验参与者的组别。此方面的研究也会为中文教育中的认知心理学、语言学等跨学科研究人员提供参考（Law et al., 2008）。

Berman 和 Slobin 的研究（1994）提供了一个跨语言的视角，特别关注不同母语的结构属性和修辞偏好如何影响叙事能力的发展。该研究分析了超过 250 篇来自不同年龄段（包括学龄前儿童、学龄儿童和成人）和说五种不同语言（英语、德语、西班牙语、希伯来语和土耳其语）的人的文本，研究了不同年龄段的人们如何叙述事件。其中分析的语言虽然没有中文，但是也给我们进行叙述分析时的语言参数选择提供了参考意义。其中所有的叙事数据都是由同一本绘本《青蛙，你在哪里？》作为基础语料得到，我们也同样选择了这本青蛙的故事中文版本作为基础语料。使用单一图画故事书的原则是提供一个共同的内容，并且跨越年龄和语言，能够代表一个典型的儿童故事：有一个英雄（男孩和他的狗），一个问题（男孩有一只宠物青蛙跑掉了），一系列由这个问题引发的动作（男孩和狗搜寻失踪的青蛙），以及一个“快乐的结局”（男孩找到了他的青蛙，或者用另一只青蛙交换）。这本特定的图画书被证明适合进行叙述分析，因为它描述了一系列相当长和复杂的事件，并允许叙述者涉及多种话题。这些包括：（1）表达时间和相继发生的事件之间的关系，按顺序（例如，青蛙从罐子里出来然后跑掉了），与先前情况相关的事件（例如，青蛙曾被关在的罐

子是空的), 以及与同时发生或同时发生的事件有关的事件 (例如, 当男孩在树上的洞里看时, 狗被蜜蜂追赶); (2) 位置轨迹的描述, 例如鹿将男孩从悬崖上推下到下面的沼泽里; (3) 对事件采取不同视角的可能性 (例如, 狗被蜜蜂追赶, 或者蜜蜂在追狗); 以及 (4) 故事有两个主要的主角, 男孩和狗, 为表达谈论同时发生事件的时间关系提供了机会, 以及保持和转移参考, 从一个角色到另一个角色。青蛙的故事另一个特点是, 它允许在事件之间进行不同层次的认知推理。例如, 从图画中很容易认出青蛙已经从罐子里出来, 因为罐子是空的; 但是不容易理解爬上岩石的男孩抓住看似树枝的东西是什么, 通过观察之后的图片才会发现, 原来他不小心抓住的是站在岩石后面的鹿的鹿角。这些不同层次的解释能够引发不同类型和层次的语言表达。该研究为理解叙述的结构和语言特点提供了深入的见解, 此外, 研究还详细分析了语言形式如何用于叙事功能, 例如表达时间关系、事件观点和文本连接性。该研究选择测量的语言参数在叙事构建中起到了核心作用。它们是语言表达的系统性部分, 能够在时间和空间中定位叙事事件, 并与彼此相关联。语言参数涵盖了一系列语法形态词素和构造类型: (1) 连接的派生词素, 如动词的各种形态, 如过去式、现在时和进行时; (2) 自由的语法词素, 如介词、副词、介词和连词; (3) 句法构造, 如关系从句和补语从句; 以及 (4) 词序的系统性变化, 如左置位 (**left-dislocation**)。研究还包括了一些在短句中出现并以特殊方式与叙事话语形成相互作用的词汇或语句, 如时间副词和短语。从五个角度来分析文本: 时态 (从动词的时态区分)、叙事融合 (介词副词的使用等)、叙述视角 (男孩子视角或者青蛙的视角等)、连贯性 (句法连词和从句等) 和叙事风格 (语体色彩等)。

1.3.4 测量指标

二语口语产出中复杂性、准确性和流畅性（CAF: complexity, accuracy and fluency）的研究一直是学者们关注的焦点，近年来研究人员已经逐步开始调查 CAF 本身的结构，以提出更精确的定义和更先进的衡量标准（Michel, 2015）。基于 CAF，赵雷（2008）提出跨文化口语交际能力的具体指标应该包括准确性、流利性、得体性和多样性四个方面（四性），本实验测量指标的选择也是从 CAF 和以上四性发展而来的。Shewan（1988）提出了一个全面的语言量化分析系统（SSLA: Shewan Spontaneous Language Analysis），在该分析中，语言数据通过图片描述任务得出。研究者使用了 12 个 SSLA 参数，主要从语义、语音和句法角度对数据进行分析。图片描述任务需要一系列的词汇和语句输出，能够很完整的体现说话者的语言水平和语音特征。Smith 和 Wheeldon（1999）同样也是采用了线上图片描述任务，设计五个实验来研究口语句子生成过程中的概念和语法编码。因此，本研究选取了两个图片描述任务，包括一张家庭场景图片和一个《青蛙的故事》的二十九页连续长图（Mayer, 1969）。许多语言学研究都采用了《青蛙的故事》做为叙述故事的材料，来调查叙述者的语言能力和语音面貌。基于图片展示进行叙事描述是一项复杂的任务。实验参与者需从图片中获取信息，并将视觉信息转换为语言叙事结构输出（Holmqvist et al., 2005）。例如，静态图片需要被解释为发生的事件，图像元素需要被转换成语言代码。这就要求参与者结合词汇、语法和叙事结构各个方面组织语言并输出。除家庭场景的图片描述外，讲故事任务中能够提取到更多叙述信息，并且能够更好的考察说话者连接词的运用和语言逻辑的连贯性。叙事实践与语言的发

展、人类社会形态以及历史观念更迭密切相关。早期的叙事实践无论是民间故事、寓言或游记都记录了人类历史的发展，并被认为有助于创建社区教育。叙事描述是由相关的多个事件联系而成的。为了解读一个完整的故事，叙述者不仅需要理解单个事件，还需要理解事件之间的相互联系。叙述者需要能够用语言表达单个事件内容，并且能够推断出前后图片内容的关联性；同时也要关注故事情节的整体层次结构，以便将有关人物、事件和活动的信息组织成一个连贯的整体。图片选取确认后，数据分析参数的选取也应易于计算，并且能够反映出语言叙事特征。本研究的分析方向是语篇性的，关注句子之间的线性序列，即通过叙述形成连贯事件的方式。

锜宝香（2001）以看图、听故事、看图复述故事的方式收集了 131 位不同阅读能力学生的口语述说语言样本，并分析其语意、语法应用能力。研究使用统计分析软件 SPSS 中的双因子变异数分析（two-factor ANOVA）等方法来处理数据。得出结论是在语意的测量方面，一般阅读能力学生的相异词汇数

（number of different words）、校正后相异词汇数、成语数显著多于低阅读能力的学生，但是两组学生在平均句子长度方面并无显著差异。相异词汇数通常被当作测量语意变化（semantic diversity）的指标来诊断语言问题。在口语述说错误分析方面，一般阅读能力学生的词汇错用、语序颠倒的次数显著低于低阅读能力的学生。相似的发现还有 Stahl（1990）的研究，12 岁的学习障碍学生在复述故事时无法适当使用人称代名词。Justice 等人（2010）提出了一个细致的微观语言评估结构，包括从句结构、短语结构、修饰语、名词和动词的 18 个英语测量参数。但是由于英语和中文普通话之间的差异性，以上测量参数无法完

全借用于普通话语言研究中。汉藏语系和印欧语系是不同类型的语言，一个显著的区别在于形态语法方面。印欧语言通常有更丰富的形态句法转换，而中文则缺乏这些变化。我们以中文和英语之间的比较为例，英语使用“s”来表示复数，但中文没有类似的复数形式；英语通过动词变形来表现过去时、现在时、一般将来时的不同时间点，中文用助词或副词来体现。因此，在本研究中也会针对一些普通话学习重难点来进行调查，例如让被试描述带有复数物品的图片、描述过去发生的事情等。一项研究调查了不同母语（中文和英语、瑞典语、希腊语）、但同患有语言障碍儿童们的语言情况（Hao et al., 2018）。该研究不仅比较了研究对象的语言、动作、刺激事件和结果，还研究了其语言的信息传达、词汇多样性、句法复杂性和语法。此外，该研究使用了6个重点测量参数来评估特定的中文语言特征，包括量词、助词、副词、主动“把”字句和被动“被”字句等。在普通话中，有时删去一句话中的助词或副词并不影响句意。与普通话类似，粤语中的助词和副词有时也是可以选择使用的。但是，在大多情况下的普通话和粤语中，助词和副词的使用能够让句子更加完整，表达语义更加清晰。因此，汉语水平较高的二语学习者往往会使用助词和副词来帮助描述时态。与助词和副词不同的是，部分语言特征测量参数是必须要使用的：在“把”字句和“被”字句结构中，“把”字或者“被”字是必须包含的；在描述对象的数量时，量词必须出现在数字和名词之间。句子的语法和单词和语素的数量，以及从句的复杂性和类型都相关。

第二章 研究方法

2.1 被试

参与本研究的被试共 80 人，年龄区间都在 20-30 岁。其中 40 人是宁夏本地人，为对照组，即中文母语组，其中 20 名女生 20 名男生。另外 40 人是留学生组，为实验组，他们的中文水平都在 HSK5 及以上，同样也是相同性别数量 20 名女生 20 名男生。40 名留学生中，20 名是宁夏大学的学生（非洲留学生居多），另外 20 名是中山大学珠海分校的学生（东南亚留学生居多）。参与实验的留学生国籍分别是：尼日利亚、布隆迪共和国、老挝 4 人、喀麦隆 3 人、塞内加尔、乌兹别克斯坦 2 人、几内亚比绍、土库曼斯坦、刚果（布）、越南、吉尔吉斯斯坦、印度尼西亚 4 人、津巴布韦、牙买加、韩国、委内瑞拉、坦桑尼亚 2 人、苏丹、马来西亚 4 人、柬埔寨 4 人、尼泊尔、赞比亚、缅甸和俄罗斯。因不同水平二语学习者汉语水平差距过大，汉语水平等级低者在词汇量、句法的运用方面能力方面不足，可能会对数据的准确性有影响，所以在 40 位二语参与者中，会将所有参与者的汉语水平统一为 HSK5 级和 HSK6 级。通过分析四十名来自不同国家的留学生的音频，也能够丰富二语学习者的语言资料库。

2.2 实验任务与材料

本研究首先使用简单问卷来收集被试的基础背景信息，如姓名、年龄、语言背景和学习中文的年限等，之后以线上和线下口语产出任务的方式来收集被试的音频，从中文二语学习组和母语组、线上音频和线下音频数据的对比中做话语分析。每位实验参与者都需要做两次，线上一次线下一次，两次口语产出

任务之间至少间隔 48 小时，以减少熟悉程度对任务的影响。80 位参与者都需要参加两次任务，线上线下各一次，分别是 15 分钟左右。本研究将会控制性别和线上线下产出顺序这两个变量，两个组分别是男生 20 名，女生 20 名；先做线上任务的 20 名，另外 20 名先做线下任务。本研究主要会从音调、实词、虚词、句末语气词（SFP）等角度，来分析参与者们线上与线下交流时中文使用是否存在较大差距。录音数据收集完成后，会通过机器识别和人耳二次检查的方式转写成文字数据。所有的文字数据将从不同方面被重新处理分析，包括句子的数量，音调，带有 SFP 的句子等。在实验开始前，所有实验参与者都需要完成一份基础信息问卷调查，包括性别、年龄、国籍、语言背景、汉语水平、开始学习中文的年份和居住在中国的时间。完成问卷调查后，开始进行第一次口语产出任务，包括描述图片，回答日常问题，对假设情景的回应等。选取图片描述任务的原因包括：不需要对象进行冗长困难的词汇阅读，对于母语者和二语者难度均等；半正式形式能够提取出更加自然的叙述，减少被试压力；完全相同的图片描述，能够对不同对象的叙述进行对比；信息以视觉形式呈现更容易引发被试说话，对于二语者来说难度较低。看图说话是一种适用于中级以下水平学生的有效口语测量方式，能够控制被试说话的范围，不同学生之间的表现也有很强的可比性，因此在这里很适合用于衡量中文二语者的口语能力。图片描述任务分为两个，其一是关于一个家庭的图片，里面有五个人分别在进行不同的活动。研究者会给出相同的“请说说图片里有什么人？有什么东西？在发生什么事？”的回答引导。图片描述任务包含人物情绪的视觉呈现，以及对多种不同动作的描述，如：“图片里有一张长桌子，桌子上旁边有四个人。

其中爷爷在喝水，哥哥在吃汉堡，弟弟在哭，可能是因为哥哥不给弟弟吃汉堡，小女孩在读书。还有一个小男孩指着电视在看电视，电视里有一个女孩在唱歌。桌子上还有一个苹果，上面有一只苍蝇，可能苹果坏了。右边桌子上有一只猫，猫把花瓶打碎了。”

语言叙述评估通常包括宏观和微观分析 (Hao et al., 2018)。宏观分析重点考察故事描述的全局性，包括对人物、背景、事件、人物行动和后果的描述；微观分析重点考察语言本身，例如不同句法结构和特定类型词语的使用。微观分析还包括更多的测量参数，例如本研究选取的测量参数：总字数能够反映故事的长度，不同词性的数量表示词汇的多样性等。图片描述任务二是一个 29 页的长图，研究者会给出“请根据系列图片，讲述《青蛙的故事》”的回答引导。为避免研究者对于实验过程的影响，所有的任务都是同一位研究者进行实施，也就是说故事总是对一个已经了解青蛙的故事的听众讲述，并且研究者也能够看到图片。这种情况增强了对被试进行指示的倾向，尤其是语言运用还不完善的二语学习者，无论是在手势上还是在指示性表达的使用上。本研究中被试都是 HSK 水平 5 级及以上的中文水平，所以他们都有语言能力独立对故事进行完整的叙事描述。但是若出现单个单词被试想不起来的情况，研究者会给予提示，如“马蜂窝”等不常见的词汇。两个图片描述任务的不同之处有：长短不同；难易程度不同；故事发生的时间顺序不同。实验参与者在开始描述图片之前，可以提前通读所有图片。准备时间有助于参与者进行思考和组织语言，并且减轻压力和紧张情绪，从而更自然的进行语言输出。线上组在进行任务时会通过 APP 腾讯会议中的“共享屏幕”功能进行图片阅读。该图片描述任务要

求被试将各种语言成分融入到一个完整的故事中，因此，能够为语言研究提供丰富的数据样本。同时，图片与图片之间的逻辑性与关联性，更有助于理解语言能力较差的二语学习者想要表达的语意。

描述图片任务结束后，研究者会提问十个家常聊天问题，如“你喜欢吃什么食物？请描述一下？”。之后是三个较为正式的社会议题问题，如“你认为很好的老师应该是什么样子的？”。最后有两个带有表演性质的问题，需要被试在语言中表现出开心或不开心的情绪，如“假如你中了奖获得了一万人民币的奖金，你想告诉你的朋友，你会怎么说？”。不同情绪问答的音频之间会进行横向对比，是否有句末语气词使用次数差异等。在 48 小时后的第二次任务中，研究者会使用 B 卷进行采访。除了描述图片和阅读句子任务完全保留 A 卷内容，中间问答任务在问题的设计上有一些微小的改动。如 A 卷的问题是“请介绍你的一位朋友”，那么相对应的 B 卷问题则是“请介绍你的一位家人”；A 卷的问题是“你认为空气污染的来源是什么？”，那么相对应 B 卷的问题则是“你认为水污染的来源是什么？”两份问卷在难易程度、顺序上尽量保持一致，以避免其他外在因素对数据的影响。

实验过程中，在参与者同意的情况下会全程进行录音，之后将 M4A 格式的录音文件转换成 WAV 格式，并使用网页语音识别功能将 WAV 音频文件转换成文字。因说话者语音、语法错误或音量等问题，网页语音识别会有一定误差，所以会通过人耳识别对文字进行二次检查更正。文字信息将以数字形式整理制成 Excel 表格，通过 SPSS 中的配对样本 t 检验 (paired-samples t test) 可以分析出不同的语言产出环境 (F2F 与线上) 和产出顺序对语言的影响。其他因

素如性别和不同语言背景将使用单因素方差分析 (One-Way ANOVA) 进行数据分析, 因素 (factor) 分别为语言背景 (L1 与 L2) 和性别 (男与女), 因变量 (dependent variable) 为下述测量指标。此外, 为分析不同的产出环境和其他影响因素之间是否有交互作用, 我们还会运用一般线性模型 (general linear model) 下的单变量分析 (univariate analysis), 考察不同影响因素之间是否存在显著的主效应。在收集完所有的语言数据后, 音频文件将会被转换成文字, 数据由 IBM SPSS Statistics (version 28.1.1.0) 软件处理分析。

2.3 具体测量指标

本论文会采用定量和定性分析相结合的方法, 将拟定的测量指标以数字形式呈现。例如, 在名词一列中, 数字 10 表示本段对话出现 10 次名词, 数字 0 表示说话者在本段对话中没有使用名词。每个说话者会有单独的列表, 包括总字数、句子总数、时长、句末语气词的出现次数等。除了关注线上线下顺序因素, 其他因素也会被控制, 如考察性别和语言背景因素对于数据是否有影响。文献综述部分提到的研究是测量指标选取的理论基础, 如 Hao 等学者 (2018) 关于患有语言障碍儿童们语言情况的研究中, 使用了 6 个重点测量参数来评估特定的中文语言特征, 本文研究参数借鉴了该 6 个参数中的三个: 量词、助词和副词。在关于 CLCM (Kong & Law, 2005) 的研究中, 借鉴了其关于总词数、句末语气词、信息词和错误的测量参数。在得出基础数据后, 会进行多层次分析, 比如名词和动词之间的比例, 比如词汇输出效率 (时长和信息词的比例)。CLCM 的研究中另外采用的参数还有沟通效率 (Index of Communication Efficiency), 是通过信息词和时长的比例运算得出结果。语言的流利性、复杂

性和准确性是衡量口语表达能力的重要因素，以下是本论文进行语言分析时可能会选取的具体测量参数举例，这些语言参数将被运用于图片描述部分和问答部分：

复杂性测量参数：

(1) 名词 (**nouns**): 名词在图画描述中任务中能够表达主要信息，信息量的传递与名词使用的数量多少密切相关。名词通常是表达具体事物、概念和人物的词语，它们在口语中起着重要的作用，其使用次数可以反映被试在描述、解释和阐述观点时的详细程度。名词可以表示人、事、物、地点或抽象概念的统称，分为专有名词和普通名词。如“一个虫子在苹果上”中有两个名词。

(2) 动词 (**verbs**): 表示人或事物的动作或状态的词语，如“小女孩正在读书”中有一个动词。错误使用的动词也算在总数内，如“弟弟要喝哥哥的汉堡”中，虽然说话者把“吃”说成了“喝”，但是该句中也算有一个动词。丰富的动词运用有助于详细地叙述事情的发展、变化和情节，反映出个体在构建故事、描述场景和经历方面的能力水平。动词的使用次数较多，可能表明说话者在描述事件、动作和过程方面的能力较强，能够生动、动态地呈现信息，使表达更具活力和感染力。

(3) 句末语气词 (**SFPs: sentence final particles**): 句末语气词不仅与语体有关，一定程度上也能够体现出说话人的情绪，如惊讶、肯定、疑问、感叹等。较多地使用句末语气词可能意味着说话者更善于表达情感，使口语表达更富有感染力。虽然句末语气词本身并不直接代表口语流利度，但它们的使用可以增加口语表达的自然度和流畅性，也可以反映出说话者的个人语言风格特点。适

当运用句末语气词可以使话语更加连贯，减少停顿和犹豫。一项关于粤语 SFPs 的相关研究发现，在线上线下不同产出环境的任务中，线下产出环境中 SFP 出现的频率要比线上产出环境高得多，并且女性使用 SFP 的频率比男性稍微多一点 (Zhang & Law, 2022)。与线上产出环境相比，虽然音调 1 的 SFP 在线下产出环境中的出现频率只高出一一点 ($p=0.289$)，但是音调 3 的 SFP 在线下产出环境中比线上产出环境高出很多 ($p=0.001$)。本论文借鉴了该粤语 SFPs 的研究，同样测量普通话线上线下不同产出环境任务中的 SFP 出现频率并进行分析。

(4) 信息词 (IW: number of informative word)：信息词是承载主要内容和关键信息的词汇，其使用次数较多可能表明说话者能够提供丰富的内容和细节，使表达更加充实。语言学中，信息词指的是具有明确的词义和独立表达意义的词语，也被称为实词。与虚词相对，信息词能够独立构成一个词组或句子，并且具有具体的概念或意义。一个信息词可以被看作是一个小的文本单位，在数据统计时，只有使用正确的信息词才会被算入总数内。信息词被认为是图片描述时的关键词汇，类似于 Yorkston 和 Beukelman (1980) 在描述图片任务中提到的“内容词”。正确使用的人称代词不算作信息词，单独人称代词不算入信息词总数内，但是可以与其他字组合成为一个信息词。如“一个小孩他在哭”中有“一个小孩”和“哭”这两个信息词，“他”不算入信息词总数内。

(5) 虚词 (FW: number of functional words)：使用虚词的次数可以反映出二语者对汉语语法规则的掌握程度，使用次数恰当且正确说明对语法结构的理解和运用能力较强。在汉语中，虚词是指一类在句子中起辅助性作用的词语，与

实词相对，虚词并没有实际意义，其中包括介词、连词、助词、语气词、量词等。与英语和其他欧洲语系不同，量词是汉语中一项重要的句法元素，但是在英语中没有对应词。因量词数量众多并且使用情况复杂，汉语学习者往往在中后期学习阶段时才能掌握量词的使用（Erbaugh, 2006）。汉语的量词大概有 902 个，汉语中不同类型的名词与特定的量词相对应，如“一个苹果”中的“个”算一个量词，“一只猫”中的“只”也算一个量词。属于虚词的介词有“与、在”等；连词有“但是、然后”等；助词有“的，了”等；语气词有“吗，呀”等；代词有“他、这”等；副词有“非常、很”等。

准确性测量参数：

(6) 错误 (errors)：错误次数较多表明二语者在中文基本语言知识方面的掌握还不够扎实。词汇使用错误多，意味着词汇量有限或者对词汇的意义和用法理解不准确；语法错误多，则说明对汉语语法规则的学习和运用存在问题。统计了错误次数之后，就可以用总字数除以错误次数计算出错误率，从而得出线上线下错误率是否有区别的结论。错误的种类有很多：“把花的盆子打碎了”算一次错误，说话者想要表达的是“花瓶”，但是因词汇量的缺乏使用了错误词语。“桌子上还有苹果，一只苍蝇”算一次错误，说话者没有使用连接词，正确语句应该是“桌子上还有苹果和一只苍蝇”。“这个图片里有一家庭”也算一次错误，说话者未使用量词，句子成分缺失。在数据统计中会把错误种类大致分类为：语法错误、词语误用、遗漏偏误三个方面。

流利性测量参数：

(7) 犹豫 (hesitations)：犹豫次数较多表明说话者在语言产出过程中不够流

畅，可能在词汇检索、语法组织、语义构建等方面存在困难，导致表达过程中出现停顿、重复、延长音等犹豫现象。频繁的犹豫可能暗示二语者对自己的中文表达能力缺乏自信，担心出错或者不确定自己的表达方式是否正确，因此在表达时会表现出多次犹豫。在说话过程中，说话者沉默停顿三秒及以上算作一次犹豫，正常的语句停顿不算犹豫。使用“那个”、“嗯”等词语来填充或暂停描述，或使用“然后”“应该是”等模糊词语来表示不确定，在数据统计时都不算做犹豫。语句中的、语言时长、信息词数量和犹豫次数，都和说话者叙述信息的效率息息相关。

(8) 时长 (duration): 说话者从开始说出第一个有意义的词到最后一个词的时间长短，不包括研究者的提问，也不包括研究者提问和说话者回答中间的停顿时长。了解实验对象在回答单独问题、或者图片描述任务中的时长后，通过结合其中的信息词个数，能够计算出实验对象在传达信息时的效率百分比。话语量较多表明二语者具有较强的表达主动性和交流欲望，愿意积极参与对话并分享自己的想法和观点，在一定程度上也反映了二语者具备一定的语言输出能力，能够运用所学的中文知识组织较长和较复杂的语句来表达思想。

(9) 重复 (repetitions): 一定程度的重复可能是二语者为了确保表达的准确性，自我检查和修正语言输出的表现；较多的重复现象表明说话者在语言产出过程中不够流畅，在词汇提取、语句组织或表意传达上存在困难，导致需要重复表述来弥补。“有一个孙女正在看，正在看电视”算一次重复。当语言不流畅导致语句重复时，会影响语意表达清晰度，降低交流效率和语言质量。句意断开的两个句子不算重复，如“这是他们的爷爷，爷爷在喝水”。

(10) 总字数 (NC: number of characters): 总字数较多表明说话者具有较强的语言输出能力, 能够运用丰富的词汇和多样的表达方式来阐述观点和内容。这反映出他们在语言知识储备和运用方面有一定的水平, 能够较为充分地表达自己的想法。语言样本中的总字数包括不正确的字、重复、自我纠正、无关的陈述、离题的字; 不包括说话者的犹豫词、噪音、以及研究者的提问。因为汉语切词有模糊地带, 所以测量字数比测量词数数据会更加清晰。

(11) 句子总数 (NS: number of sentences): 句子总数较多且分布较为均匀, 每个句子之间的衔接自然流畅, 说明说话者的语言流畅度较高; 如果句子总数少, 且句子之间的停顿、犹豫较多, 可能反映出语言产出的流畅度欠佳。吕叔湘 (1979) 指出如果要探讨汉语句子的构造机制, 要根据句法形式在语言片段中将句子的边界划分清楚。徐晶凝 (2023) 在“话头自足结构”理论的框架下, 提出如何在连续的语言片段中划分句子的边界。因二语学习者有时不能表达出句意完整的句子, 所以缺失句子成分的短语, 如果前后有长时间停顿, 在数据分析时也算做一个句子。如“花的罐子掉下来, 然后碎”如果中间有超过三秒的停顿, 在本研究中算两个句子而不是一个句子。

(12) 句子长度 (SL: sentence length): 用总字数除以句子总数, 可以得出平均句子长度, 较长的句子通常意味着使用者能够运用更复杂的语法结构和更丰富的词汇。句子长度能够反映出被试说话的节奏, 也能体现其汉语水平和语言习惯。能够构造长句子并且语法正确, 说明说话者对中文的语法和句法规则有较好的掌握和运用能力; 而平均句子长度较短且存在较多语法错误的情况, 反映出语法和句法知识的欠缺。家庭活动图片描述任务因单一图片可描述内容较

少，未使用该测量参数；青蛙的故事描述任务中使用了该测量参数。

(13) 信息输出效率 (CE: **communication efficiency**): 用总信息词数除以时长的秒数, 可以得出信息输出效率。信息输出效率反映了学习者在特定时间内传递信息的能力, 输出效率高的说话者能够流畅地说出连贯的语句, 减少停顿、重复和犹豫。高效的信息输出表明学习者能够更流畅、更精准地表达想法。这项参数能够帮助区分不同等级汉语水平者, 了解被试的语言流利度。家庭活动图片描述任务中, 因被试对于单一图片的描述时长较短, 未使用该测量参数; 青蛙的故事描述任务中使用了该测量参数。

以上指标能够全面、多维度地评估被试的不同口语能力和表现: 错误次数可以衡量说话者对语言规则的掌握程度以及语言运用的准确性。犹豫次数反映出语言产出过程中的流畅性和语言思维转换的速度、语言组织规划能力。回答问题的时长能够展现语言反应速度、语言知识储备与运用能力等方面的水平, 话语量能体现语言表达的主动性、自信心、话题熟悉程度。重复次数可以揭示语言流畅性、语言监控能力。总字数反映出语言输出量和丰富度、语言思维活跃度, 句子总数可以了解语言组织能力、表意清晰度、思维活跃度与逻辑性。平均句子长度有助于判断语言复杂性、语法和句法掌握能力。信息输出效率综合了语言组织规划、词汇检索运用、语法掌握运用、思维敏捷度与逻辑能力等多方面的能力。并且这些测量指标能够相互补充和印证: 总字数、句子总数、平均句子长度等指标的发展变化, 反映出说话者在不同学习阶段或语言发展阶段的能力提升与变化, 并且与语言知识的积累、思维能力的发展相互影响。错误次数和犹豫次数都可以反映出说话者对语言知识的掌握和运用能力, 当错误

次数较多时，犹豫次数可能也会相应增加。语言表达流畅度与错误次数、犹豫次数、重复次数等指标都有关联，流畅度差导致错误、犹豫和重复现象增多，反之，错误、犹豫和重复多也会降低流畅度。信息输出效率与语言组织规划能力、词汇检索与运用能力、语法掌握与运用能力、思维敏捷度与逻辑能力多个指标密切相关。通过综合分析以上指标，可以更全面、深入地了解被试的口语表现及其背后的语言能力发展状况，为教学和学习提供有针对性的指导和反馈。

第三章 口语产出环境（线上线下）的差别探讨

3.1 图片描述任务

3.1.1 单一图片任务

首先，我们对看图说话中的单一家庭图片描述任务的语言数据进行配对样本 t 检验。下表是线上线下产出环境中不同测量语言参数的数据，线下为组 1，线上为组 2。从表 1 可以看出，除了重复次数外，其他语言参数的线上和线下的差别没有达到统计学意义的显著差异，也就是在家庭活动描述图片任务中，线上和线下的产出环境对各大多数语言参数没有显著的影响。重复次数 $t(df=79) = -2.05, p < 0.05$ 的数据表明，线上任务的重复次数要比线下任务的重复次数略多。在该短图片描述任务中其他参数无显著差异的情况下，线上重复次数较多表示实验参与者们在线上的口语表现较弱于线下。原因可能是多样的，比如在线下任务时图片观察更加直观等。具体结论会结合所有数据统一分析。

表 1 单一图片线上线任务数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差平均值	2:线上组 (N=80)	均值	标准差	标准误差平均值
总字数(线下)	89.04	45.092	5.041	总字数(线上)	95.51	54.515	6.095
句子总数(线下)	9.01	3.857	0.431	句子总数(线上)	9.57	4.854	0.543
SFPs(线下)	1.05	1.179	0.132	SFPs(线上)	0.85	1.244	0.139
重复次数(线下)	0.63	1.14	0.128	重复次数(线上)	0.94	1.521	0.17
信息词(线下)	20.4	8.641	0.966	信息词(线上)	20.98	11.004	1.23
虚词(线下)	10.79	6.644	0.743	虚词(线上)	11.26	8.417	0.941
错误(线下)	1.66	1.575	0.176	错误(线上)	1.74	2.157	0.241
名词(线下)	14.49	5.516	0.617	名词(线上)	14.9	7.505	0.839
动词(线下)	6.51	2.387	0.267	动词(线上)	6.76	2.63	0.294
犹豫次数(线下)	0.84	1.096	0.123	犹豫次数(线上)	0.91	1.275	0.143

	双侧 <i>p</i>
总字数(线下) - 总字数(线上)	0.086
句子总数(线下) - 句子总数(线上)	0.144
SFPs(线下) - SFPs(线上)	0.175
重复次数(线下) - 重复次数(线上)	0.044
信息词(线下) - 信息词(线上)	0.456
虚词(线下) - 虚词(线上)	0.416
错误(线下) - 错误(线上)	0.700
名词(线下) - 名词(线上)	0.441
动词(线下) - 动词(线上)	0.263
犹豫次数(线下) - 犹豫次数(线上)	0.574

3.1.2 长故事图片任务

郑咏滢和刘飞凤(2020)对不同难易程度的口语语料进行了计量语言学分析,结果表明二语学习者口语词汇复杂度各指标随着任务复杂度升高而显著变化。本研究看图说话中的两个任务难易程度不同,其一是单一家庭活动任务,其二是青蛙的故事描述任务。长故事描述任务更加能够体现被试的语言能力和

语言习惯，我们同样对长故事描述任务的数据进行分析，并在分析中加入了对于平均句子长度和信息输出效率参数的测量。在语言习得的不同阶段，平均句子长度可以反映学习者掌握的语言水平。随着语言能力的提高，学习者能够构造更复杂、信息含量更丰富的句子。信息输出效率关乎学习者在实际交流中的表现，在日常交流或专业领域中，能够高效传递信息的能力对于二语学习者来说尤其重要。这两项参数有助于了解学习者如何适应目标语言环境的要求，包括语速、准确性和流畅度等方面，进而指导他们如何在不同语境中有效使用语言。数据分析首先从两个角度对样本进行了配对样本 t 检验，分别是线上线下环境和参与口语产出任务产出顺序的角度。表 2 是线上线下不同环境时，青蛙的故事描述任务各项语言测量参数的数据。其中有两项测量参数呈现差异：犹豫次数 $t(df=79) = -4.38, p < 0.001$ 和信息输出效率 $t(df=79) = 2.9, p < 0.05$ 。数据表明，线下口语产出任务的犹豫次数比线上任务的犹豫次数少，线下任务的信息输出效率比线上任务的信息输出效率要高。在线上任务中，参与者需要同时处理多种任务，如操作电子设备、关注屏幕上的提示和反馈，这些额外的认知负荷可能导致更多的犹豫和降低沟通效率。与此相反，线下任务允许参与者更专注于故事内容和口头表达，减少了与技术相关的干扰，有利于提高沟通的流畅性和效率。线下任务中的社交互动可能提供了更多的情感支持和鼓励，这可以增强叙述者的自信心，减少犹豫，并促进信息的高效输出。线上任务可能缺乏足够的社交和情感支持，使得参与者在叙述长故事时感到更孤立、更不确定，从而影响了沟通的流畅性和效率。

表 2 长故事线上线下任务数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差 平均值	2:线上组 (N=80)	均值	标准差	标准误差 平均值
总字数(线下)	434.55	160.669	17.963	总字数(线上)	439.23	145.525	16.27
句子总数(线下)	40.16	14.447	1.615	句子总数(线上)	40.85	11.581	1.295
SFPs(线下)	7.51	4.003	0.448	SFPs(线上)	7	4.311	0.482
重复次数(线下)	3.41	3.61	0.404	重复次数(线上)	3.43	3.662	0.409
信息词(线下)	93.55	30.098	3.365	信息词(线上)	95.93	30.365	3.395
虚词(线下)	76.01	29.994	3.353	虚词(线上)	76.71	28.825	3.223
错误(线下)	6.18	7.341	0.821	错误(线上)	6.4	7.228	0.808
名词(线下)	55.95	23.181	2.592	名词(线上)	55.49	19.483	2.178
动词(线下)	36.46	12.065	1.349	动词(线上)	37.36	10.772	1.204
犹豫次数(线下)	4.31	4.108	0.459	犹豫次数(线上)	6.39	5.015	0.561
句子长度(线下)	10.847	1.396	0.156	句子长度(线上)	10.692	1.287	0.144
CE(线下)	0.6444	0.208	0.023	CE(线上)	0.584	0.228	0.026

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(线下) - 总字数(线上)	-0.317	79	0.752
句子总数(线下) - 句子总数(线上)	-0.509	79	0.612
SFPs(线下) - SFPs(线上)	1.014	79	0.314
重复次数(线下) - 重复次数(线上)	-0.031	79	0.976
信息词(线下) - 信息词(线上)	-0.786	79	0.434
虚词(线下) - 虚词(线上)	-0.252	79	0.802
错误(线下) - 错误(线上)	-0.487	79	0.628
名词(线下) - 名词(线上)	0.271	79	0.787
动词(线下) - 动词(线上)	-0.726	79	0.470
犹豫次数(线下) - 犹豫次数(线上)	-4.378	79	<0.001
句子长度(线下) - 句子长度(线上)	1.147	79	0.255
CE(线下) - CE(线上)	2.899	79	0.005

3.2 问答任务

表 3 十四个口语产出问答列表

	A 问卷 Questionnaire A	B 问卷 Questionnaire B
(1)	你可以自我介绍一下吗? Introduce yourself please.	你可以自我介绍一下吗? Introduce yourself please.
(2)	你喜欢在大城市工作还是小城市?	你喜欢在大城市生活还是小城市?

	Do you prefer to work in a big city or small city?	Do you prefer to live in a big city or small city?
(3)	请介绍你的一位朋友。比如说他的外表、性格和喜好? Please introduce a friend of yours, such as his/her appearance, personality, and hobbies?	请介绍你的一位家人。比如说外表、性格和喜好? Please introduce a family member of yours. Like his/her appearance, personality, and hobbies?
(4)	你平时的生活是怎样的?请描述你的一天。 What is your daily life like? Please describe your day.	你在学校的生活是怎样的?请描述你在学校的一天。 What is your school life like? Please describe your day in school.
(5)	周末一般会怎么度过?和谁一起度过? How do you usually spend your weekends? Who do you spend with?	假期一般会怎么度过?和谁一起度过? How do you usually spend your vacations? Who do you spend with?
(6)	请介绍你的国家。比如风俗文化。 Please introduce your country, such as customs and culture.	请介绍你的城市。比如风俗文化。 Please introduce your city, such as customs and culture.
(7)	你喜欢吃什么食物?请描述一下? Could you describe the food you like?	你不喜欢吃什么食物?请描述一下? Could you describe the food you dislike?
(8)	你最喜欢看什么书?为什么? What is your favorite book? Why?	你最喜欢什么电影?为什么? What is your favorite movie? Why?
(9)	你最难忘的一件事是什么? What is the most memorable thing in your life?	你最开心的一件事是什么? What is the happiest thing in your life?
(10)	你最喜欢的城市是哪里?为什么? Where is your favorite city? Why?	你最喜欢的国家是哪里?为什么? Where is your favorite country? Why?
(11)	你认为空气污染的来源是什么?你有什么建议? What are the sources of air pollution? Do you have suggestions on air pollution?	你认为水污染的来源是什么?你有什么建议? What are the sources of water pollution? Do you have suggestions on water pollution?

(12)	你认为很好的家长应该是什么样子的? What do you think good parents should be like?	你认为很好的老师应该是什么样子的? What do you think a good teacher should be like?
(happy)	假如你中了奖获得了一万人民币的奖金, 你想告诉你的朋友, 你会怎么说? If you won a prize of 10,000 RMB and you want to tell your friend, what would you say?	如果老师说你的论文很好, 你想告诉你的朋友, 你会怎么说? If the teacher said your paper is excellent, you are excited you want to tell your friend, what would you say?
(sad)	如果老师说你的论文很差, 你想告诉你的朋友, 你会怎么说? If the teacher said your paper is terrible, you are upset you want to tell your friend, what would you say?	假如你被人骗了一万人民币, 你很不开心, 你想告诉你的朋友, 你会怎么说? If you are cheated of 10,000 RMB, you are very upset, and you want to tell your friend, what would you say?

表 3 是 14 个口语产出问答的序号列表, 问卷 A 和问卷 B 相同序号的问题则是相似的问题。我们首先对第一个问题 (1) 的语言数据进行配对样本 t 检验, 表 4 是问题 (1) 线上线下产出环境中不同测量语言参数的数据, 线下为组 1, 线上为组 2。从中可以得出数据: 总字数 $t(df=79) = 3.223, p < 0.05$, 句子总数 $t(df=79) = 2.837, p < 0.05$, 犹豫次数 $t(df=79) = -3.266, p < 0.05$, 重复次数 $t(df=79) = 6.268, p < 0.001$, 信息词 $t(df=79) = -4.765, p < 0.001$, 虚词 $t(df=79) = -5.076, p < 0.001$, 错误 $t(df=79) = 6.008, p < 0.001$, 动词 $t(df=79) = -6.076, p < 0.001$, CE $t(df=79) = -5.576, p < 0.001$ 。数据表明, 在问题 (1) 中, 线下问答的总字数、句子总数、重复次数、错误数要比线上问答的多; 线下问答的信息词数、虚词数、动词数、犹豫次数和信息输出效率要比线上问答的少。

表 4 问题 (1) 线上线下数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差 平均值	2:线上组 (N=80)	均值	标准差	标准误差 平均值
总字数(线下)	34.88	17.839	1.994	总字数(线上)	26.51	19.452	2.175
句子总数(线下)	3.74	1.49	0.167	句子总数(线上)	3.16	1.618	0.181
SFPs(线下)	0.2	0.433	0.048	SFPs(线上)	0.18	0.382	0.043
重复次数(线下)	3.59	4.735	0.529	重复次数(线上)	0.25	0.666	0.074
信息词(线下)	3.95	3.044	0.34	信息词(线上)	6.18	3.795	0.424
虚词(线下)	0.75	1.336	0.149	虚词(线上)	2.41	2.823	0.316
错误(线下)	1.49	2.111	0.236	错误(线上)	0.09	0.284	0.032
名词(线下)	2.25	1.852	0.207	名词(线上)	2.73	2.037	0.228
动词(线下)	0.81	1.181	0.132	动词(线上)	1.95	1.395	0.156
犹豫次数(线下)	5.18	7.747	0.866	犹豫次数(线上)	8.23	5.983	0.669
CE(线下)	0.893	0.333	0.037	CE(线上)	3.015	3.252	0.364

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(线下) - 总字数(线上)	3.223	79	0.002
句子总数(线下) - 句子总数(线上)	2.837	79	0.006
SFPs(线下) - SFPs(线上)	0.424	79	0.673
重复次数(线下) - 重复次数(线上)	6.268	79	<0.001
信息词(线下) - 信息词(线上)	-4.765	79	<0.001
虚词(线下) - 虚词(线上)	-5.076	79	<0.001
错误(线下) - 错误(线上)	6.008	79	<0.001
名词(线下) - 名词(线上)	-1.953	79	0.054
动词(线下) - 动词(线上)	-6.076	79	<0.001
犹豫次数(线下) - 犹豫次数(线上)	-3.266	79	0.002
CE(线下) - CE(线上)	-5.576	79	<0.001

同样的，我们对接下来所有问题的语言数据都进行了配对样本 *t* 检验，表 5 是问题（2）线上线下产出环境中不同测量语言参数的数据，线下为组 1，线上为组 2。从数据分析得出的结论是问题（2）的各语言参数都没有显著差异，但是在表 6 的问题（3）数据中，虚词 $t(df=79) = -2.068, p < 0.05$ ，说明线上问答的虚词数比线下问答多。

表 5 问题（2）线上线下数据的成对样本统计与检验

1:线下组 (N=80)				2:线上组 (N=80)			
	均值	标准差	标准误差 平均值		均值	标准差	标准误差 平均值
总字数(线下)	30.1	13.309	1.488	总字数(线上)	32.21	21.745	2.431
句子总数(线下)	2.81	1.08	0.121	句子总数(线上)	3.06	1.656	0.185
SFPs(线下)	0.24	0.509	0.057	SFPs(线上)	0.29	0.697	0.078
重复次数(线下)	0.23	0.573	0.064	重复次数(线上)	0.2	0.433	0.048
信息词(线下)	6.58	2.56	0.286	信息词(线上)	6.86	3.631	0.406
虚词(线下)	5.15	2.994	0.335	虚词(线上)	5.46	4.248	0.475
错误(线下)	0.15	0.393	0.044	错误(线上)	0.18	0.444	0.05
名词(线下)	3.29	1.285	0.144	名词(线上)	3.5	2.019	0.226
动词(线下)	1.84	1.267	0.142	动词(线上)	1.86	1.209	0.135
犹豫次数(线下)	0.09	0.326	0.036	犹豫次数(线上)	0.06	0.244	0.027
CE(线下)	0.869	0.329	0.039	CE(线上)	0.79	0.258	0.029

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(线下) - 总字数(线上)	-0.860	79	0.392
句子总数(线下) - 句子总数(线上)	-1.245	79	0.217
SFPs(线下) - SFPs(线上)	-0.542	79	0.589
重复次数(线下) - 重复次数(线上)	0.352	79	0.726
信息词(线下) - 信息词(线上)	-0.652	79	0.516
虚词(线下) - 虚词(线上)	-0.627	79	0.533
错误(线下) - 错误(线上)	-0.406	79	0.686
名词(线下) - 名词(线上)	-0.856	79	0.395
动词(线下) - 动词(线上)	-0.161	79	0.872
犹豫次数(线下) - 犹豫次数(线上)	0.630	79	0.530
CE(线下) - CE(线上)	1.909	79	0.060

表 6 问题 (3) 线上线下数据的成对样本统计与检验

1:线下组 (N=80)				2:线上组 (N=80)			
	均值	标准差	标准误差 平均值		均值	标准差	标准误差 平均值
总字数(线下)	47.23	23.743	2.655	总字数(线上)	51.93	30.604	3.422
句子总数(线下)	4.9	2.168	0.242	句子总数(线上)	5.34	2.595	0.29
SFPs(线下)	0.6	0.836	0.093	SFPs(线上)	0.75	0.935	0.104
重复次数(线下)	0.61	0.921	0.103	重复次数(线上)	0.51	0.827	0.092
信息词(线下)	9.81	4.501	0.503	信息词(线上)	10.54	4.685	0.524
虚词(线下)	8.48	5.229	0.585	虚词(线上)	9.88	7.222	0.807
错误(线下)	0.46	0.728	0.081	错误(线上)	0.4	0.773	0.086
名词(线下)	4.46	2.407	0.269	名词(线上)	4.34	2.105	0.235

动词(线下)	2.78	2.044	0.228	动词(线上)	3.21	1.94	0.217
犹豫次数(线下)	0.28	0.616	0.069	犹豫次数(线上)	0.21	0.589	0.066
CE(线下)	0.701	0.247	0.028	CE(线上)	0.689	0.248	0.028

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(线下) - 总字数(线上)	-1.728	79	0.088
句子总数(线下) - 句子总数(线上)	-1.452	79	0.150
SFPs(线下) - SFPs(线上)	-1.136	79	0.259
重复次数(线下) - 重复次数(线上)	0.882	79	0.380
信息词(线下) - 信息词(线上)	-1.365	79	0.176
虚词(线下) - 虚词(线上)	-2.068	79	0.042
错误(线下) - 错误(线上)	0.672	79	0.504
名词(线下) - 名词(线上)	0.451	79	0.653
动词(线下) - 动词(线上)	-1.500	79	0.138
犹豫次数(线下) - 犹豫次数(线上)	0.779	79	0.438
CE(线下) - CE(线上)	0.397	79	0.693

表 7 是问题 (4) 线上线下产出环境中不同测量语言参数的数据, 线下为组 1, 线上为组 2。从数据分析得出的结论是总字数 $t(df=79) = -2.717, p < 0.05$, 句子总数 $t(df=79) = -2.831, p < 0.05$, SFPs $t(df=79) = -2.148, p < 0.05$, 虚词 $t(df=79) = -2.784, p < 0.05$, 名词 $t(df=79) = -2.002, p < 0.05$, 说明线上问答的总字数、句子总数、句末语气词数、虚词数和名词数都比线下问答的多。从 CE $t(df=79) = 2.016, p < 0.05$ 可以看出, 线上问答的信息输出效率比线下任务的低。

表 7 问题 (4) 线上线下数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差平均值	2:线上组 (N=80)	均值	标准差	标准误差平均值
总字数(线下)	47.75	26.11	2.919	总字数(线上)	57.11	36.4	4.07
句子总数(线下)	4.49	2.306	0.258	句子总数(线上)	5.4	3.108	0.348
SFPs(线下)	0.43	0.792	0.089	SFPs(线上)	0.63	0.832	0.093
重复次数(线下)	0.34	0.927	0.104	重复次数(线上)	0.53	0.9	0.101

信息词(线下)	11.38	6.425	0.718	信息词(线上)	13.03	7.826	0.875
虚词(线下)	7.4	4.621	0.517	虚词(线上)	9.18	6.587	0.736
错误(线下)	0.25	0.516	0.058	错误(线上)	0.33	0.689	0.077
名词(线下)	4.51	3.006	0.336	名词(线上)	5.41	4.034	0.451
动词(线下)	5.71	3.541	0.396	动词(线上)	6.18	3.89	0.435
犹豫次数(线下)	0.11	0.356	0.04	犹豫次数(线上)	0.1	0.302	0.034
CE(线下)	0.885	0.308	0.034	CE(线上)	0.81	0.235	0.026

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(线下) - 总字数(线上)	-2.717	79	0.008
句子总数(线下) - 句子总数(线上)	-2.831	79	0.006
SFPs(线下) - SFPs(线上)	-2.148	79	0.035
重复次数(线下) - 重复次数(线上)	-1.708	79	0.092
信息词(线下) - 信息词(线上)	-1.960	79	0.054
虚词(线下) - 虚词(线上)	-2.784	79	0.007
错误(线下) - 错误(线上)	-1.062	79	0.292
名词(线下) - 名词(线上)	-2.002	79	0.049
动词(线下) - 动词(线上)	-1.070	79	0.288
犹豫次数(线下) - 犹豫次数(线上)	0.276	79	0.783
CE(线下) - CE(线上)	2.016	79	0.047

表 8 是问题 (5) 线上线下产出环境中不同测量语言参数的数据, 线下为组 1, 线上为组 2。从犹豫次数 $t(df=79) = 2.188, p < 0.05$, CE $t(df=79) = 2.074, p < 0.05$ 可以看出, 线上问答的信息输出效率比线下问答的低, 但同时线上问答的犹豫次数也比线下问答少。

表 8 问题 (5) 线上线下数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差平均值	2:线上组 (N=80)	均值	标准差	标准误差平均值
总字数(线下)	29.88	21.3	2.381	总字数(线上)	33.38	20.167	2.255
句子总数(线下)	2.64	1.931	0.216	句子总数(线上)	2.81	1.722	0.192
SFPs(线下)	0.38	0.663	0.074	SFPs(线上)	0.45	0.884	0.099
重复次数(线下)	0.28	0.711	0.08	重复次数(线上)	0.25	0.606	0.068
信息词(线下)	6.71	4.249	0.475	信息词(线上)	7.14	3.655	0.409
虚词(线下)	5	4.091	0.457	虚词(线上)	5.59	3.831	0.428

错误(线下)	0.15	0.453	0.051	错误(线上)	0.24	0.534	0.06
名词(线下)	3.31	2.029	0.227	名词(线上)	3.68	2.042	0.228
动词(线下)	3.14	2.266	0.253	动词(线上)	3.25	1.761	0.197
犹豫次数(线下)	0.15	0.393	0.044	犹豫次数(线上)	0.05	0.219	0.025
CE(线下)	0.935	0.411	0.046	CE(线上)	0.843	0.333	0.037

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(线下) - 总字数(线上)	-1.289	79	0.201
句子总数(线下) - 句子总数(线上)	-0.684	79	0.496
SFPs(线下) - SFPs(线上)	-0.748	79	0.457
重复次数(线下) - 重复次数(线上)	0.271	79	0.787
信息词(线下) - 信息词(线上)	-0.791	79	0.431
虚词(线下) - 虚词(线上)	-1.145	79	0.256
错误(线下) - 错误(线上)	-1.222	79	0.225
名词(线下) - 名词(线上)	-1.242	79	0.218
动词(线下) - 动词(线上)	-0.441	79	0.661
犹豫次数(线下) - 犹豫次数(线上)	2.188	79	0.032
总字数(线下) - 总字数(线上)	2.074	79	0.041

表 9 是问题 (6) 线上线下产出环境中不同测量语言参数的数据, 线下为组 1, 线上为组 2。从虚词 $t(df=79) = -2.075, p < 0.05$ 得出结论, 线上问答的虚词数多于线下问答。

表 9 问题 (6) 线上线下数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差 平均值	2:线上组 (N=80)	均值	标准差	标准误差 平均值
总字数(线下)	49.83	32.941	3.683	总字数(线上)	58.96	42.645	4.768
句子总数(线下)	4.44	2.755	0.308	句子总数(线上)	5.14	3.408	0.381
SFPs(线下)	0.57	0.897	0.1	SFPs(线上)	0.81	1.08	0.121
重复次数(线下)	0.56	1.178	0.132	重复次数(线上)	0.55	0.926	0.104
信息词(线下)	10.05	6.677	0.746	信息词(线上)	11.45	7.289	0.815
虚词(线下)	8.63	5.823	0.651	虚词(线上)	10.43	7.69	0.86
错误(线下)	0.51	0.941	0.105	错误(线上)	0.54	1.252	0.14
名词(线下)	6.74	4.58	0.512	名词(线上)	7.58	4.991	0.558
动词(线下)	1.93	1.999	0.223	动词(线上)	1.8	1.724	0.193
犹豫次数(线下)	0.44	0.869	0.097	犹豫次数(线上)	0.33	0.632	0.071

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(线下) - 总字数(线上)	-2.374	79	0.020
句子总数(线下) - 句子总数(线上)	-1.314	79	0.193
SFPs(线下) - SFPs(线上)	-0.163	79	0.871
重复次数(线下) - 重复次数(线上)	0.266	79	0.791
信息词(线下) - 信息词(线上)	-1.696	79	0.094
虚词(线下) - 虚词(线上)	-2.441	79	0.017
错误(线下) - 错误(线上)	-2.109	79	0.038
名词(线下) - 名词(线上)	-0.424	79	0.673
动词(线下) - 动词(线上)	-1.698	79	0.093
犹豫次数(线下) - 犹豫次数(线上)	-1.754	79	0.083
CE(线下) - CE(线上)	1.687	79	0.096

表 11 是问题 (8) 线上线下产出环境中不同测量语言参数的数据, 线下为组 1, 线上为组 2。从总字数 $t(df=79) = -2.632, p < 0.05$, 句子总数 $t(df=79) = -2.137, p < 0.05$, 信息词 $t(df=79) = -3.015, p < 0.05$, 虚词 $t(df=79) = -2.260, p < 0.05$, 名词 $t(df=79) = -2.177, p < 0.05$, 动词 $t(df=79) = -2.340, p < 0.05$, CE $t(df=79) = 2.409, p < 0.05$ 得出结论, 线上的总字数、句子总数、信息词数、虚词数、名词数和动词数多于线下, 线上的信息传输效率低于线下。

表 11 问题 (8) 线上线下数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差 平均值	2:线上组 (N=80)	均值	标准差	标准误差 平均值
总字数(线下)	32.58	23.236	2.598	总字数(线上)	41.79	37.032	4.14
句子总数(线下)	2.75	1.539	0.172	句子总数(线上)	3.33	2.623	0.293
SFPs(线下)	0.36	0.661	0.074	SFPs(线上)	0.46	0.856	0.096
重复次数(线下)	0.33	0.632	0.071	重复次数(线上)	0.39	0.665	0.074
信息词(线下)	6.51	4.075	0.456	信息词(线上)	8.32	5.97	0.667
虚词(线下)	5.54	4.712	0.527	虚词(线上)	6.86	6.336	0.708
错误(线下)	0.21	0.495	0.055	错误(线上)	0.26	0.47	0.053
名词(线下)	2.86	2.608	0.292	名词(线上)	3.75	3.703	0.414
动词(线下)	2.73	1.713	0.192	动词(线上)	3.34	2.17	0.243

犹豫次数(线下)	0.15	0.359	0.04	犹豫次数(线上)	0.24	0.457	0.051
CE(线下)	0.872	0.697	0.078	CE(线上)	0.684	0.254	0.028

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(线下) - 总字数(线上)	-2.632	79	0.010
句子总数(线下) - 句子总数(线上)	-2.137	79	0.036
SFPs(线下) - SFPs(线上)	-0.905	79	0.368
重复次数(线下) - 重复次数(线上)	-0.685	79	0.496
信息词(线下) - 信息词(线上)	-3.015	79	0.003
虚词(线下) - 虚词(线上)	-2.260	79	0.027
错误(线下) - 错误(线上)	-0.783	79	0.436
名词(线下) - 名词(线上)	-2.177	79	0.032
动词(线下) - 动词(线上)	-2.340	79	0.022
犹豫次数(线下) - 犹豫次数(线上)	-1.409	79	0.163
CE(线下) - CE(线上)	2.409	79	0.018

表 12 和表 13 是问题 (9) 和问题 (10) 线上线下产出环境中不同测量语言参数的数据, 线下为组 1, 线上为组 2。从数据可以看出, 问题 (9) 中各语言参数都没有显著差异; 问题 (10) 中, 从 $CE\ t(df=79) = 2.144, p < 0.05$ 得出结论, 线上的信息传输效率低于线下。

表 12 问题 (9) 线上线下数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差 平均值	2:线上组 (N=80)	均值	标准差	标准误差 平均值
总字数(线下)	32.06	40.069	5.667	总字数(线上)	28.3	20.301	2.871
句子总数(线下)	2.78	3.489	0.493	句子总数(线上)	2.56	1.786	0.253
SFPs(线下)	0.54	0.93	0.132	SFPs(线上)	0.36	0.693	0.098
重复次数(线下)	0.2	0.495	0.07	重复次数(线上)	0.28	0.573	0.081
信息词(线下)	6.72	7.754	1.097	信息词(线上)	5.98	3.455	0.489
虚词(线下)	6.26	6.966	0.985	虚词(线上)	5.56	4.625	0.654
错误(线下)	0.12	0.594	0.084	错误(线上)	0.38	0.967	0.137
名词(线下)	2.64	3.269	0.462	名词(线上)	2.52	1.374	0.194
动词(线下)	2.26	2.898	0.41	动词(线上)	1.92	1.55	0.219
犹豫次数(线下)	0.28	0.536	0.076	犹豫次数(线上)	0.22	0.465	0.066
CE(线下)	0.915	0.515	0.073	CE(线上)	0.81	0.486	0.069

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(线下) - 总字数(线上)	0.545	79	0.588
句子总数(线下) - 句子总数(线上)	0.361	79	0.719
SFPs(线下) - SFPs(线上)	1.013	79	0.316
重复次数(线下) - 重复次数(线上)	-0.727	79	0.471
信息词(线下) - 信息词(线上)	0.586	79	0.561
虚词(线下) - 虚词(线上)	0.535	79	0.595
错误(线下) - 错误(线上)	-1.565	79	0.124
名词(线下) - 名词(线上)	0.237	79	0.814
动词(线下) - 动词(线上)	0.690	79	0.494
犹豫次数(线下) - 犹豫次数(线上)	0.573	79	0.569
CE(线下) - CE(线上)	1.002	79	0.321

表 13 问题 (10) 线上线下数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差 平均值	2:线上组 (N=80)	均值	标准差	标准误差 平均值
总字数(线下)	33.4	24.175	2.703	总字数(线上)	38.99	31.384	3.509
句子总数(线下)	3.68	4.968	0.555	句子总数(线上)	3.58	2.38	0.266
SFPs(线下)	0.29	0.556	0.062	SFPs(线上)	0.49	0.9	0.101
重复次数(线下)	0.28	0.636	0.071	重复次数(线上)	0.16	0.434	0.049
信息词(线下)	6.76	4.249	0.475	信息词(线上)	7.66	5.528	0.618
虚词(线下)	7.21	5.341	0.597	虚词(线上)	8.43	6.793	0.759
错误(线下)	0.23	0.527	0.059	错误(线上)	0.23	0.595	0.067
名词(线下)	3.94	2.528	0.283	名词(线上)	4.2	2.812	0.314
动词(线下)	1.56	1.112	0.124	动词(线上)	1.95	1.915	0.214
犹豫次数(线下)	0.11	0.318	0.036	犹豫次数(线上)	0.16	0.489	0.055
CE(线下)	0.89	0.418	0.047	CE(线上)	0.773	0.357	0.04

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(线下) - 总字数(线上)	-1.408	79	0.163
句子总数(线下) - 句子总数(线上)	0.159	79	0.874
SFPs(线下) - SFPs(线上)	-1.791	79	0.077
重复次数(线下) - 重复次数(线上)	1.265	79	0.210
信息词(线下) - 信息词(线上)	-1.217	79	0.227
虚词(线下) - 虚词(线上)	-1.400	79	0.165
错误(线下) - 错误(线上)	0.000	79	1.000
名词(线下) - 名词(线上)	-0.666	79	0.507
动词(线下) - 动词(线上)	-1.595	79	0.115

犹豫次数(线下) - 犹豫次数(线上)	-0.728	79	0.469
CE(线下) - CE(线上)	2.144	79	0.035

表 14 是问题 (11) 线上线下产出环境中不同测量语言参数的数据, 线下为组 1, 线上为组 2。从问题 (11) 中的句子总数 $t(df=79) = -2.117, p < 0.05$ 得出结论, 线上的句子总数比线下多。

表 14 问题 (11) 线上线下数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差平均值	2:线上组 (N=80)	均值	标准差	标准误差平均值
总字数(线下)	57.2	42.367	4.737	总字数(线上)	61.58	37.48	4.19
句子总数(线下)	4.31	2.949	0.33	句子总数(线上)	4.93	2.782	0.311
SFPs(线下)	0.69	0.908	0.102	SFPs(线上)	0.74	0.91	0.102
重复次数(线下)	0.66	0.954	0.107	重复次数(线上)	0.59	0.867	0.097
信息词(线下)	11.13	7.113	0.795	信息词(线上)	11.99	6.552	0.732
虚词(线下)	8.9	7.411	0.829	虚词(线上)	9.78	7.089	0.793
错误(线下)	0.63	0.877	0.098	错误(线上)	0.75	0.849	0.095
名词(线下)	6.59	4.212	0.471	名词(线上)	7.01	4.527	0.506
动词(线下)	3.65	2.668	0.298	动词(线上)	3.79	2.186	0.244
犹豫次数(线下)	0.36	0.68	0.076	犹豫次数(线上)	0.41	0.896	0.1
CE(线下)	0.694	0.319	0.035	CE(线上)	0.638	0.355	0.04

	t	自由度	双侧 p
总字数(线下) - 总字数(线上)	-1.223	79	0.225
句子总数(线下) - 句子总数(线上)	-2.117	79	0.037
SFPs(线下) - SFPs(线上)	-0.429	79	0.669
重复次数(线下) - 重复次数(线上)	0.555	79	0.581
信息词(线下) - 信息词(线上)	-1.291	79	0.200
虚词(线下) - 虚词(线上)	-1.338	79	0.185
错误(线下) - 错误(线上)	-1.275	79	0.206
名词(线下) - 名词(线上)	-0.995	79	0.323
动词(线下) - 动词(线上)	-0.415	79	0.679
犹豫次数(线下) - 犹豫次数(线上)	-0.601	79	0.550
CE(线下) - CE(线上)	1.300	79	0.197

表 15 是问题 (12) 线上线下产出环境中不同测量语言参数的数据, 线下为组 1, 线上为组 2。从总字数 $t(df=79) = -2.229, p < 0.05$, 句子总数 $t(df=79) = -2.229, p < 0.05$, 虚词 $t(df=79) = -2.113, p < 0.05$, CE $t(df=79) = 2.128, p < 0.05$ 得出结论, 线上的总字数、句子总数、虚词数多于线下, 线上的信息传输效率低于线下。

表 15 问题 (12) 线上线下数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差 平均值	2:线上组 (N=80)	均值	标准差	标准误差 平均值
总字数(线下)	35.29	24.982	2.793	总字数(线上)	44.6	42.217	4.72
句子总数(线下)	2.81	1.772	0.198	句子总数(线上)	3.5	2.925	0.327
SFPs(线下)	0.48	0.693	0.077	SFPs(线上)	0.53	0.9	0.101
重复次数(线下)	0.23	0.656	0.073	重复次数(线上)	0.45	0.926	0.104
信息词(线下)	7.38	4.734	0.529	信息词(线上)	8.59	6.806	0.761
虚词(线下)	6.33	4.547	0.508	虚词(线上)	8.01	8.196	0.916
错误(线下)	0.43	0.689	0.077	错误(线上)	0.49	0.795	0.089
名词(线下)	3.55	2.76	0.309	名词(线上)	4.1	3.455	0.386
动词(线下)	2.5	2.006	0.224	动词(线上)	2.96	3.042	0.34
犹豫次数(线下)	0.11	0.358	0.04	犹豫次数(线上)	0.15	0.426	0.048
CE(线下)	0.821	0.35	0.039	CE(线上)	0.727	0.316	0.035

	t	自由度	双侧 p
总字数(线下) - 总字数(线上)	-2.229	79	0.029
句子总数(线下) - 句子总数(线上)	-2.299	79	0.024
SFPs(线下) - SFPs(线上)	-0.415	79	0.679
重复次数(线下) - 重复次数(线上)	-1.738	79	0.086
信息词(线下) - 信息词(线上)	-1.882	79	0.064
虚词(线下) - 虚词(线上)	-2.113	79	0.038
错误(线下) - 错误(线上)	-0.649	79	0.518
名词(线下) - 名词(线上)	-1.653	79	0.102
动词(线下) - 动词(线上)	-1.416	79	0.161
犹豫次数(线下) - 犹豫次数(线上)	-0.652	78	0.516
CE(线下) - CE(线上)	2.128	79	0.036

表 16 是问题 (happy) 线上线下产出环境中不同测量语言参数的数据, 线下为组 1, 线上为组 2。从名词 $t(df=79) = -2.393, p < 0.05$ 得出结论, 线上的名词数多于线下。

表 16 问题 (happy) 线上线下数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差 平均值	2:线上组 (N=80)	均值	标准差	标准误差 平均值
总字数(线下)	32.58	25.095	2.806	总字数(线上)	34.7	21.794	2.437
句子总数(线下)	3.25	2.336	0.261	句子总数(线上)	3.48	2.239	0.25
SFPs(线下)	0.99	1.196	0.134	SFPs(线上)	0.8	0.863	0.096
重复次数(线下)	0.29	0.62	0.069	重复次数(线上)	0.34	0.693	0.077
信息词(线下)	6.79	4.602	0.515	信息词(线上)	7.79	4.304	0.481
虚词(线下)	5.2	4.58	0.512	虚词(线上)	5.31	3.873	0.433
错误(线下)	0.23	0.573	0.064	错误(线上)	0.16	0.462	0.052
名词(线下)	2.14	1.704	0.191	名词(线上)	2.78	1.889	0.211
动词(线下)	3.35	2.506	0.28	动词(线上)	3.6	2.254	0.252
犹豫次数(线下)	0.09	0.326	0.036	犹豫次数(线上)	0.11	0.39	0.044
CE(线下)	0.955	0.492	0.055	CE(线上)	0.889	0.374	0.042

	t	自由度	双侧 p
总字数(线下) - 总字数(线上)	-0.651	79	0.517
句子总数(线下) - 句子总数(线上)	-0.636	79	0.527
SFPs(线下) - SFPs(线上)	1.277	79	0.205
重复次数(线下) - 重复次数(线上)	-0.542	79	0.589
信息词(线下) - 信息词(线上)	-1.589	79	0.116
虚词(线下) - 虚词(线上)	-0.197	79	0.845
错误(线下) - 错误(线上)	0.844	79	0.401
名词(线下) - 名词(线上)	-2.393	79	0.019
动词(线下) - 动词(线上)	-0.721	79	0.473
犹豫次数(线下) - 犹豫次数(线上)	-0.445	79	0.658
CE(线下) - CE(线上)	1.247	79	0.216

表 17 是问题 (sad) 线上线下产出环境中不同测量语言参数的数据, 线下为组 1, 线上为组 2。从总字数 $t(df=79) = -2.732, p < 0.05$, 句子总数 $t(df=$

79) = -3.110, $p < 0.05$, SFPs $t(df=79) = -2.449, p < 0.05$, 虚词 $t(df=79) = -2.514, p < 0.05$, CE $t(df=79) = 1.296, p < 0.05$ 得出结论, 线上的总字数、句子总数、句末语气词数和虚词数多于线下, 线上的信息传输效率低于线下。

表 17 问题 (sad) 线上线下载数据的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差平均值	2:线上组 (N=80)	均值	标准差	标准误差平均值
总字数(线下)	33.2	21.816	2.439	总字数(线上)	40.89	26.919	3.01
句子总数(线下)	3.39	1.754	0.196	句子总数(线上)	4.13	2.218	0.248
SFPs(线下)	0.74	0.868	0.097	SFPs(线上)	1.1	1.165	0.13
重复次数(线下)	0.41	0.867	0.097	重复次数(线上)	0.45	0.967	0.108
信息词(线下)	6.94	4.193	0.469	信息词(线上)	7.65	4.361	0.488
虚词(线下)	5.04	3.723	0.416	虚词(线上)	6.53	5.236	0.585
错误(线下)	0.31	0.667	0.075	错误(线上)	0.38	0.7	0.078
名词(线下)	2.36	1.693	0.189	名词(线上)	2.54	1.981	0.221
动词(线下)	3.21	2.406	0.269	动词(线上)	3.55	2.116	0.237
犹豫次数(线下)	0.11	0.39	0.044	犹豫次数(线上)	0.05	0.271	0.03
CE(线下)	0.913	0.478	0.053	CE(线上)	0.812	0.344	0.038

	t	自由度	双侧 p
总字数(线下) - 总字数(线上)	-2.732	79	0.008
句子总数(线下) - 句子总数(线上)	-3.110	79	0.003
SFPs(线下) - SFPs(线上)	-2.449	79	0.017
重复次数(线下) - 重复次数(线上)	-0.340	79	0.735
信息词(线下) - 信息词(线上)	-1.428	79	0.157
虚词(线下) - 虚词(线上)	-2.514	79	0.014
错误(线下) - 错误(线上)	-0.712	79	0.479
名词(线下) - 名词(线上)	-0.679	79	0.499
动词(线下) - 动词(线上)	-1.150	79	0.253
犹豫次数(线下) - 犹豫次数(线上)	1.296	79	0.199
CE(线下) - CE(线上)	2.165	79	0.033

3.3 任务分类对比

我们接下来将问卷分为三个部分再次进行对比分析，第一个部分是图片描述：长短图片描述任务之间语言参数的区别。第二个部分是正式问答和非正式问答之间语言参数的区别。第三个部分是回答不同情绪的问题（开心和不开心）时语言参数的区别。

3.3.1 长短图片描述任务之间的差异

因为长短图片描述任务在总字数、句子总数等方面差异很大，我们将具体的数字数据转换成比例数据来进行配对检验。比如名词数转换为名词率，即名词数除以信息词数；动词数转换为动词率，即动词数除以信息词数；虚词数转换为虚词率，即虚词数除以总字数；句子长度，即总字数除以句子总数；句末语气词率，即句末语气词数除以总字数；重复次数率，即重复次数除以总字数；错误次数率，即错误次数除以总字数；犹豫次数率，即犹豫次数除以总字数。短图片家庭描述任务为组 1，长图片青蛙的故事描述任务为组 2。

表 18 长短图片描述任务的成对样本统计与检验

1:短任务 (N=160)	均值	标准差	标准误差平均值	2:长任务 (N=160)	均值	标准差	标准误差平均值
CE(短)	0.822	0.284	0.022	CE(长)	0.614	0.22	0.017
名词率(短)	0.724	0.118	0.009	名词率(长)	0.583	0.101	0.008
动词率(短)	0.335	0.082	0.006	动词率(长)	0.399	0.081	0.006
虚词率(短)	0.118	0.033	0.003	虚词率(长)	0.175	0.029	0.002
句子长度(短)	9.918	1.708	0.135	句子长度(长)	10.769	1.341	0.106
SFPs(短)	0.01	0.013	0.001	SFPs(长)	0.017	0.009	0.001
重复率(短)	0.008	0.013	0.001	重复率(长)	0.008	0.007	0.001
错误率(短)	0.018	0.017	0.001	错误率(长)	0.015	0.016	0.001
犹豫率(短)	0.01	0.014	0.001	犹豫率(长)	0.014	0.014	0.001

	<i>t</i>	自由度	双侧 <i>p</i>
CE(短) - CE(长)	13.030	159	<0.001
名词率(短) - 名词率(长)	13.167	159	<0.001
动词率(短) - 动词率(长)	-7.099	159	<0.001
虚词率(短) - 虚词率(长)	-19.313	159	<0.001
句子长度(短) - 句子长度(长)	-6.076	159	<0.001
SFPs(短) - SFPs(长)	-6.449	159	<0.001
重复率(短) - 重复率(长)	-0.169	159	0.866
错误率(短) - 错误率(长)	2.244	159	0.026
犹豫率(短) - 犹豫率(长)	-3.673	159	<0.001

从表 18 可以看出, 除重复率外, 所有的测量参数都呈现出差异。 $p < 0.001$ 的有 CE $t(df=159) = 13.030$ 、名词率 $t(df=159) = 13.167$ 、动词率 $t(df=159) = -7.099$ 、虚词率 $t(df=159) = -19.313$ 、句子长度 $t(df=159) = -6.076$ 、SFPs $t(df=159) = -6.449$ 和犹豫率 $t(df=159) = -3.673$ 。 $p < 0.05$ 的有错误率 $t(df=159) = 2.244$ 。数据说明短图片任务比长故事任务的信息传输效率更高, 名词率更高, 但同时错误率也较高。长故事任务比短图片任务的动词率、虚词率、句子长度、句末语气词率和犹豫率更高。短图片任务要求被试基于单张图片快速表达核心信息, 由于视觉刺激直接明了, 被试能够迅速聚焦于关键点, 用较少的词汇和句子传达较为明确的内容。然而, 由于时间限制和即兴反应的压力, 可能会导致更高的错误率。在描述图片时, 被试主要关注的是图片中的对象或事物, 因此名词使用频率更高, 因为它们是直观描绘场景的重要元素。长故事任务需要构建情节、描绘角色动态以及展现因果关系, 这必然涉及更多的动作描述和连接性词语的使用, 以保证故事连贯性和逻辑性。同时, 讲述一个完整的故事需要更复杂的句子结构和更多的情感表达, 因此句子长度会增加, 并且为了增强叙述效果和情感色彩, 被试使用了更多句末语气词来调整

语调和传递情绪。面对需要深度思考和组织语言的长故事任务，被试可能需要更多的时间来构思内容、选择恰当的词汇和表述方式，因此犹豫现象更为常见。总而言之，长短不同的图片描述任务会影响被试的语言产出特点，短图片任务更注重即时反应和精确描述，而长故事任务则侧重于叙事连贯性和情感丰富度。两者各有其语言学特征和挑战，在教师组织语言能力测试时可以参考以上发现。

为了对比线上线下口语产出环境对于整体图片描述任务的影响，我们整合了短图片任务角度和长故事任务，从线上线下口语产出环境角度对比数据，线下为组 1，线上为组 2。根据表 19 得出结论有两个语言参数呈现差异：犹豫次数 $t(df=79) = -4.086, p < 0.001$ 和信息传输效率 $t(df=79) = 3.765, p < 0.001$ 。数据说明在整体图片描述任务中，线下的犹豫次数比线上的多，但是信息传输效率大于线上。结合之前单独分析的两个图片描述任务的数据：在家庭活动短图片描述任务中，线上的重复次数要比线下的重复次数略多；在长故事描述任务中，线下的犹豫次数比线上的犹豫次数少，线下的信息输出效率比线上的信息输出效率要高。对于长故事这类要求叙述连贯性和深度的任务，线下的犹豫次数反而较少，说明面对面交流时，被试的叙述流畅性得到了提升，或是被试能够更快地适应和调整讲述节奏。同样，在长故事描述任务中，线下的信息输出效率更高，这进一步印证了面对面交流对于复杂内容叙述的优势，包括更有效的信息构建、更好的情感共鸣以及更强的叙事动力。综上所述，线上可能更易于产生重复现象，而线下虽然犹豫次数可能增多，但在信息传输效率方面有明显优势，特别是在处理复杂情境和深度信息时。

表 19 图片描述任务线上线下的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差 平均值	2:线上组 (N=80)	均值	标准差	标准误差 平均值
总字数(线下)	523.59	186.458	20.847	总字数(线上)	534.74	182.897	20.449
句子总数(线下)	49.18	16.451	1.839	句子总数(线上)	50.43	14.825	1.658
SFPs(线下)	8.56	4.455	0.498	SFPs(线上)	7.85	4.699	0.525
重复次数(线下)	4.04	4.095	0.458	重复次数(线上)	4.36	4.705	0.526
信息词(线下)	113.95	34.001	3.801	信息词(线上)	116.9	37.664	4.211
虚词(线下)	86.8	33.202	3.712	虚词(线上)	87.98	33.752	3.774
错误(线下)	7.84	8.174	0.914	错误(线上)	8.14	8.552	0.956
名词(线下)	70.44	26.168	2.926	名词(线上)	70.39	24.499	2.739
动词(线下)	42.98	13.407	1.499	动词(线上)	44.13	12.112	1.354
犹豫次数(线下)	5.15	4.623	0.517	犹豫次数(线上)	7.3	5.669	0.634
CE(线下)	1.507	0.476	0.053	CE(线上)	1.365	0.446	0.05

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(线下) - 总字数(线上)	-0.702	79	0.485
句子总数(线下) - 句子总数(线上)	-0.847	79	0.399
SFPs(线下) - SFPs(线上)	1.329	79	0.188
重复次数(线下) - 重复次数(线上)	-0.737	79	0.463
信息词(线下) - 信息词(线上)	-0.904	79	0.369
虚词(线下) - 虚词(线上)	-0.409	79	0.684
错误(线下) - 错误(线上)	-0.595	79	0.554
名词(线下) - 名词(线上)	0.029	79	0.977
动词(线下) - 动词(线上)	-0.888	79	0.377
犹豫次数(线下) - 犹豫次数(线上)	-4.086	79	<0.001
CE(线下) - CE(线上)	3.765	79	<0.001

3.3.2 正式问答和非正式问答之间的差异

在十四个问题中，我们选择问题 11 和 12 作为正式问答的代表，在问卷 A 和问卷 B 中分别是：“你认为空气（水）污染的来源是什么？你有什么建议？”和“你认为很好的家长（老师）应该是什么样子的？”。非正式问答我们选择了

问题 5 和问题 7, 在问卷 A 和问卷 B 中分别是: “周末 (假期) 一般会怎么度过? 和谁一起度过? 你 (不) 喜欢吃什么食物? 请描述一下?” 表 20 为正式问答 (formal) 和非正式问答 (casual) 的数据对比, 从数据可以看出, 除动词外, 所有的测量参数都呈现出很大差异 ($p < 0.001$)。其中总字数 $t(df=159) = -9.975$ 、句子总数 $t(df=159) = -8.144$ 、SFPs $t(df=159) = -4.583$ 、重复次数 $t(df=159) = -3.718$ 、信息词 $t(df=159) = -8.653$ 、虚词 $t(df=159) = -9.630$ 、错误次数 $t(df=159) = -7.972$ 、名词 $t(df=159) = -10.273$ 、动词 $t(df=159) = -1.133$ 、犹豫次数 $t(df=159) = -4.711$ 和 CE $t(df=159) = 8.323$ 。数据说明正式问答的总字数、句子总数、句末语气词、重复次数、信息词数、虚词数、错误次数、名词数、犹豫次数都比非正式问答多, 但是非正式问答的信息传输效率高于正式问答。

在正式问答环境中, 母语者和中文二语学习者都倾向于使用更多的字数、句子、和其他语言元素。这可能是因为正式问答通常对语言的准确性和规范性有更高的要求, 这可能导致被访者更频繁地进行自我修正 (增加错误次数) 和使用更正式的语言 (增加虚词和名词数)。被访者可能会更加谨慎地选择词汇和语句结构以避免误解或不准确的信息, 从而导致更多的犹豫和重复。正式问答往往需要更严谨的表述、完整的逻辑结构以及更多的解释性内容来确保信息的准确无误传递, 这导致其包含更多的信息词、名词、虚词等, 并由于思考时间增加而出现犹豫次数增多的现象。非正式交流倾向于更简洁和直接的表达, 减少了不必要的修饰和冗余, 从而提高了信息的传递速度。非正式沟通中, 人们倾向于使用更为口语化、简洁的语言表达, 这也是二语学习者在语言学习过

程中最先掌握的部分。在不拘泥于复杂的句式或严格的语法结构时，核心思想就能够被快速传达，提高信息传输效率。中文二语学习者在正式场合下表现出更多的犹豫和错误，因为他们在使用非母语进行复杂的语言表达时面临更大的挑战。以上数据揭示了正式和非正式问答环境对于语言使用和沟通效率的影响。在正式问答中，母语者和二语学习者都更注重语言的规范性和准确性，但这可能以沟通效率为代价。而在非正式问答中，更简洁、直接和自然的交流方式有助于提高信息传输效率。了解这些差异对于语言教学、跨文化交流和沟通策略的设计都是很有帮助的。

表 20 正式问答和非正式问答的成对样本统计与检验

c:非正式组 (N=160)	均值	标准差	标准误差 平均值	f:正式组 (N=160)	均值	标准差	标准误差 平均值
总字数(非)	54.86	30.608	2.42	总字数(正式)	99.33	66.027	5.22
句子总数(非)	5.12	2.549	0.201	句子总数(正式)	7.78	4.534	0.358
SFPs(非)	0.68	1.036	0.082	SFPs(正式)	1.21	1.38	0.109
重复次数(非)	0.56	1.032	0.082	重复次数(正式)	0.96	1.303	0.103
信息词(非)	12.96	5.957	0.471	信息词(正式)	19.54	10.886	0.861
虚词(非)	8.91	6.239	0.493	虚词(正式)	16.51	12.156	0.961
错误(非)	0.35	0.646	0.051	错误(正式)	1.14	1.307	0.103
名词(非)	5.81	2.96	0.234	名词(正式)	10.63	6.507	0.514
动词(非)	6.1	3.173	0.251	动词(正式)	6.45	3.872	0.306
犹豫次数(非)	0.2	0.486	0.038	犹豫次数(正式)	0.53	0.911	0.072
CE(非)	1.921	0.73	0.058	CE(正式)	1.44	0.542	0.043

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(非) - 总字数(正式)	-9.975	159	<0.001
句子总数(非) - 句子总数(正式)	-8.144	159	<0.001
SFPs(非) - SFPs(正式)	-4.583	159	<0.001
重复次数(非) - 重复次数(正式)	-3.718	159	<0.001
信息词(非) - 信息词(正式)	-8.653	159	<0.001
虚词(非) - 虚词(正式)	-9.630	159	<0.001

错误(非) - 错误(正式)	-7.972	159	<0.001
名词(非) - 名词(正式)	-10.273	159	<0.001
动词(非) - 动词(正式)	-1.133	159	0.259
犹豫次数(非) - 犹豫次数(正式)	-4.711	159	<0.001
CE(非) - CE(正式)	8.323	159	<0.001

我们接下来将非正式问答（问题 5 和 7）和正式问答（问题 11 和 12）的数据整合，从线上线下的口语产出环境角度对比整合后的问答数据。线下为组 1，线上为组 2。从表 21 可以看出， $p < 0.05$ 的有句子总数 $t(df=79) = -3.075$ 、信息词 $t(df=79) = -2.906$ 、虚词 $t(df=79) = -3.180$ 、错误次数 $t(df=79) = -2.692$ 、名词数 $t(df=79) = -2.290$ 、动词数 $t(df=79) = -2.049$ 和信息传输效率 $t(df=79) = 3.255$ ； $p < 0.001$ 的有总字数 $t(df=79) = -3.463$ 。数据说明线下的总字数、句子总数、信息词数、虚词数、错误次数、名词数和动词数都比线上的少，但是线下的信息传输效率比线上的高。

表 21 问答任务线上线下数据的成对样本检验

	t	自由度	双侧 p
总字数(线下) - 总字数(线上)	-3.463	79	<0.001
句子总数(线下) - 句子总数(线上)	-3.075	79	0.003
SFPs(线下) - SFPs(线上)	-0.790	79	0.432
重复次数(线下) - 重复次数(线上)	-0.402	79	0.689
信息词(线下) - 信息词(线上)	-2.906	79	0.005
虚词(线下) - 虚词(线上)	-3.180	79	0.002
错误(线下) - 错误(线上)	-2.692	79	0.009
名词(线下) - 名词(线上)	-2.290	79	0.025
动词(线下) - 动词(线上)	-2.049	79	0.044
犹豫次数(线下) - 犹豫次数(线上)	-0.394	79	0.695
CE(线下) - CE(线上)	3.255	79	0.002

3.3.3 回答不同情绪的问题之间的差异

在语言学领域，认知语言学的兴起为情感研究开辟了新的道路，因为它被视为理解认知加工和概念的来源（萧季桦，2007）。本论文也选取了不同情绪的问题进行探究，在十四个问题中，问题 13 和 14 是带有明显的情绪色彩（开心和不开心）的问题。开心的问题（问题 happy）在问卷 A 和问卷 B 中分别是：“假如你中了奖获得了一万人民币的奖金，你想告诉你的朋友，你会怎么说？如果老师说你的论文很好，你想告诉你的朋友，你会怎么说？”我们以“问题 h”表示。不开心的问题（问题 sad）在问卷 A 和问卷 B 中分别是：假如你被人骗了一万人民币的奖金，你很不开心，你想告诉你的朋友，你会怎么说？如果老师说你的论文很差，你想告诉你的朋友，你会怎么说？我们以“问题 s”表示。

表 22 问题h和问题s的成对样本统计与检验

h:开心组 (N=160)	均值	标准差	标准误差 平均值	s:不开心组 (N=160)	均值	标准差	标准误差 平均值
总字数(h)	33.64	23.453	1.854	总字数(s)	37.04	24.726	1.955
句子总数(h)	3.36	2.283	0.181	句子总数(s)	3.76	2.027	0.16
SFPs(h)	0.89	1.044	0.083	SFPs(s)	0.92	1.04	0.082
重复次数(h)	0.31	0.656	0.052	重复次数(s)	0.43	0.915	0.072
信息词(h)	7.29	4.47	0.353	信息词(s)	7.29	4.279	0.338
虚词(h)	5.26	4.228	0.334	虚词(s)	5.78	4.59	0.363
错误(h)	0.19	0.52	0.041	错误(s)	0.34	0.683	0.054
名词(h)	2.46	1.822	0.144	名词(s)	2.45	1.839	0.145
动词(h)	3.48	2.379	0.188	动词(s)	3.38	2.265	0.179
犹豫次数(h)	0.1	0.358	0.028	犹豫次数(s)	0.08	0.336	0.027
CE(h)	0.922	0.437	0.035	CE(s)	0.862	0.418	0.033

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(h) - 总字数(s)	-2.069	159	0.040

句子总数(h) - 句子总数(s)	-2.269	159	0.025
SFPs(h) - SFPs(s)	-0.239	159	0.811
重复次数(h) - 重复次数(s)	-1.644	159	0.102
信息词(h) - 信息词(s)	-0.021	159	0.984
虚词(h) - 虚词(s)	-1.418	159	0.158
错误(h) - 错误(s)	-2.636	159	0.009
名词(h) - 名词(s)	0.040	159	0.968
动词(h) - 动词(s)	0.560	159	0.576
犹豫次数(h) - 犹豫次数(s)	0.506	159	0.614
CE(h) - CE(s)	1.628	159	0.105

从表 22 数据可以得出，总字数 $t(df=159) = -2.069$ 、句子总数 $t(df=159) = -2.269$ 和错误次数 $t(df=159) = -2.636$ 呈现差异 ($p < 0.05$)，说明开心情绪问题的总字数、句子总数和错误次数都少于不开心情绪的问题，得出结论是情绪也会影响人们的口语表达。当被试讨论开心内容时，他们的回答通常更为精炼。当人们谈论积极或令人愉快的话题时，他们可能感到更加轻松自在，情感上的舒适可能导致语言表达更为流畅和自信，因此能用更少的词语和句子结构来精确传达信息，从而减少冗余和错误。积极情绪促进语言的流畅性和连贯性，而消极情绪可能导致言语中断和重述，增加了错误次数和句子的重构。心理动力学机制提出 (Fowler & Christakis, 2008)，讨论积极情绪的内容会激发人们的内在动机，产生分享欲望，这时的回答会更加直接、中心突出，不需要过多铺垫或赘述。积极情绪可以促进大脑的认知灵活性和工作效率，使得思考和表达过程更为迅速，因此可能在较短的时间内组织好答案，字数自然会相对较少。同时谈论开心的事情时，个体可能感觉社会评价的压力较小，无需过多修饰或谨慎措辞以避免负面反应，因此减少了由于过度思考而导致的错误。开心的情绪容易引发听者共鸣，为了保持这种积极互动氛围，被试可能会选择

更为简洁明快的方式来讲述故事，避免复杂化带来的消极影响。由于心理、社交和认知等多个因素相互作用的结果，导致了被试在面对不同情绪的问题时出现总字数、句子总数和错误次数上的差异。我们接下来将开心的问题和不开心问题的数据进行整合，从线上线下口语产出环境角度对比了整合后的数据，根据表 23 得出结论：句子总数 $t(df=79) = -1.998$ 、名词 $t(df=79) = -2.034$ 和 CE $t(df=79) = 2.439$ 呈现差异 ($p < 0.05$)，说明线下的句子总数和名词数少于线上，但是信息传输效率大于线上。

表 23 情绪问题线上线下的成对样本统计与检验

1:线下组 (N=80)	均值	标准差	标准误差平均值	2:线上组 (N=80)	均值	标准差	标准误差平均值
总字数(线下)	65.78	41.886	4.683	总字数(线上)	75.59	44.711	4.999
句子总数(线下)	6.64	3.537	0.395	句子总数(线上)	7.6	3.854	0.431
SFPs(线下)	1.73	1.492	0.167	SFPs(线上)	1.9	1.725	0.193
重复次数(线下)	0.7	1.141	0.128	重复次数(线上)	0.79	1.456	0.163
信息词(线下)	13.73	7.848	0.877	信息词(线上)	15.44	7.833	0.876
虚词(线下)	10.24	6.89	0.77	虚词(线上)	11.84	7.991	0.893
错误(线下)	0.54	1.006	0.112	错误(线上)	0.54	0.954	0.107
名词(线下)	4.5	2.801	0.313	名词(线上)	5.31	3.282	0.367
动词(线下)	6.56	4.36	0.487	动词(线上)	7.15	3.901	0.436
犹豫次数(线下)	0.2	0.537	0.06	犹豫次数(线上)	0.16	0.489	0.055
句子长度(线下)	1.868	0.798	0.089	句子长度(线上)	1.7	0.623	0.07

	t	自由度	双侧 p
总字数(线下) - 总字数(线上)	-1.863	79	0.066
句子总数(线下) - 句子总数(线上)	-1.998	79	0.049
SFPs(线下) - SFPs(线上)	-0.771	79	0.443
重复次数(线下) - 重复次数(线上)	-0.662	79	0.510
信息词(线下) - 信息词(线上)	-1.801	79	0.075
虚词(线下) - 虚词(线上)	-1.728	79	0.088
错误(线下) - 错误(线上)	0.000	79	1.000
名词(线下) - 名词(线上)	-2.034	79	0.045
动词(线下) - 动词(线上)	-1.092	79	0.278

犹豫次数(线下) - 犹豫次数(线上)	0.536	79	0.593
句子长度(线下) - 句子长度(线上)	2.439	79	0.017

第四章 产出顺序（第一次和第二次）的差别探讨

4.1 图片描述任务

4.1.1 单一图片任务

周丹丹（2010）提出在二语习得的过程中，同时注意语言内容和形式非常困难。其研究以 20 名大学生为对象，让他们对故事复述任务进行重复练习，得出结论是二语学习者在语言练习的过程中，注意力逐渐从内容转移到语言形式上，实现了口语表达内容和形式的平衡发展；同时语言形式的各个方面如准确性和复杂性也都得到了相应的进步。表 24 是单一图片任务中产出顺序中不同测量语言参数的数据，第一次参加数据为组 1（先），第二次参加为组 2（后）。数据得出，总字数 $t(df=79) = 4.47$ 、句子总数 $t(df=79) = 5.12$ 和犹豫次数 $t(df=79) = 4.14$ 呈现差异 ($p < 0.001$)，说明第一次的家庭图片描述任务的总字数、句子总数和犹豫次数都多于第二次的的数据。下表中呈现出差异 ($p < 0.05$) 的参数还有信息词数 $t(df=79) = 2.1$ 、虚词数 $t(df=79) = 2.83$ 和名词数 $t(df=79) = 2.2$ ，说明第一次的家庭图片描述任务的信息词数、虚词数和名词数也较多于第二次的的数据。在第一次单一图片描述任务中，图片的新颖性导致参与者更加详细地描述，以确保覆盖所有重要的细节。由于不熟悉任务要求，第一次尝试时会有更多的犹豫和探索性的语言使用。在第一次描述中，参与者在思考和表达上投入更多的努力，为第二次描述提供了练习和准备。而在

第二次描述任务中，参访者对图片变得更熟悉，因此使用更少的字数和句子来传达相同的信息。随着任务的重复，参与者练习了如何更有效地组织和传达信息，从而减少了犹豫和不必要的信息。

表 24 单一图片产出顺序数据的成对样本检验

	双侧 <i>p</i>
总字数(先) - 总字数(后)	<0.001
句子总数(先) - 句子总数(后)	<0.001
SFPs(先) - SFPs(后)	0.126
重复次数(先) - 重复次数(后)	0.691
信息词(先) - 信息词(后)	0.039
虚词(先) - 虚词(后)	0.006
错误(先) - 错误(后)	0.051
名词(先) - 名词(后)	0.031
动词(先) - 动词(后)	0.145
犹豫次数(先) - 犹豫次数(后)	<0.001

4.1.2 长故事任务

表 25 是从先后产出顺序角度分析时，青蛙的故事任务各项语言测量参数的数据。从下图数据可以看出，句末语气词 $t(df=79) = 2.7$ 、犹豫次数 $t(df=79) = 3.28$ 和信息输出效率 $t(df=79) = -3.62$ 呈现统计学差异 ($p < 0.05$)，说明第一次的中的青蛙的故事描述任务的句末语气词数和犹豫次数多于第二次的的数据，第二次的信息输出效率比第一次的高。在第一次讲述故事时，被访者在尝试组织故事细节的同时进行描述，这导致更多的犹豫和使用句末语气词来填充间断。随着故事内容的熟悉，第二次讲述时被访者更自信，语言表达更流畅，导致犹豫次数减少。在第一次讲述后，被访者对故事内容进行了反思和内化，使得第二次讲述时能够更高效地输出信息。对于母语者和中文二语学习

者，母语者在两次尝试之间的改进在于他们能够更快地适应任务并优化语言输出。二语学习者的提升更侧重于减少语言错误和增加准确性，他们在语言产出上更需注意和练习。

表 25 长故事产出顺序数据的成对样本检验

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(先) - 总字数(后)	1.135	79	0.260
句子总数(先) - 句子总数(后)	1.317	79	0.192
SFPs(先) - SFPs(后)	2.695	79	0.009
重复次数(先) - 重复次数(后)	0.956	79	0.342
信息词(先) - 信息词(后)	-0.886	79	0.378
虚词(先) - 虚词(后)	1.242	79	0.218
错误(先) - 错误(后)	0.814	79	0.418
名词(先) - 名词(后)	-0.654	79	0.515
动词(先) - 动词(后)	0.362	79	0.718
犹豫次数(先) - 犹豫次数(后)	3.278	79	0.002
句子长度(先) - 句子长度(后)	-0.395	79	0.694
CE(先) - CE(后)	-3.615	79	<0.001

4.2 问答任务

表 26 是问题 (1) 不同产出顺序中语言测量参数的数据，第一次参加数据为组 1 (先)，第二次参加为组 2 (后)。从数据可以看出，重复次数 $t(df=79) = 2.565, p < 0.05$ ，犹豫次数 $t(df=79) = 2.830, p < 0.05$ ，CE $t(df=79) = -3.369, p < 0.05$ ，得出结论是第一次的重复次数和犹豫次数多于第二次，第一次的信息传输效率低于第二次。

表 26 问题 (1) 产出顺序的成对样本统计与检验

1:第一次组 (N=80)	均值	标准差	标准误差 平均值	2:第二次组 (N=80)	均值	标准差	标准误差 平均值
总字数(先)	32.56	21.805	2.438	总字数(后)	28.83	15.795	1.766

句子总数(先)	3.6	1.776	0.199	句子总数(后)	3.3	1.344	0.15
SFPs(先)	0.24	0.457	0.051	SFPs(后)	0.14	0.347	0.039
重复次数(先)	0.34	0.728	0.081	重复次数(后)	0.13	0.369	0.041
信息词(先)	6.23	4.551	0.509	信息词(后)	5.8	2.702	0.302
虚词(先)	2.68	3.507	0.392	虚词(后)	2.08	2.103	0.235
错误(先)	0.06	0.244	0.027	错误(后)	0.09	0.284	0.032
名词(先)	2.68	2.282	0.255	名词(后)	2.49	1.543	0.172
动词(先)	2	1.821	0.204	动词(后)	2.03	1.458	0.163
犹豫次数(先)	0.13	0.333	0.037	犹豫次数(后)	0.01	0.112	0.013
CE(先)	0.809	0.312	0.035	CE(后)	0.978	0.471	0.053

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(先) - 总字数(后)	1.370	79	0.174
句子总数(先) - 句子总数(后)	1.428	79	0.157
SFPs(先) - SFPs(后)	1.727	79	0.088
重复次数(先) - 重复次数(后)	2.565	79	0.012
信息词(先) - 信息词(后)	0.885	79	0.379
虚词(先) - 虚词(后)	1.566	79	0.121
错误(先) - 错误(后)	-0.575	79	0.567
名词(先) - 名词(后)	0.732	79	0.466
动词(先) - 动词(后)	-0.127	79	0.899
犹豫次数(先) - 犹豫次数(后)	2.830	79	0.006
CE(先) - CE(后)	-3.369	79	0.001

表 27 是问题 (2) 不同产出顺序中语言测量参数的数据, 第一次参加数据为组 1 (先), 第二次参加为组 2 (后)。从数据可以看出, 重复次数 $t(df=79) = -2.218, p < 0.05$, CE $t(df=79) = -3.369, p < 0.001$, 得出结论是第一次的重复次数少于第二次, 第一次的信息传输效率高于第二次。

表 27 问题 (2) 产出顺序的成对样本统计与检验

1: 第一次组 (N=80)	均值	标准差	标准误差平均值	2: 第二次组 (N=80)	均值	标准差	标准误差平均值
总字数(先)	33.4	25.816	4.082	总字数(后)	34.23	17.809	2.816
句子总数(先)	3.05	1.825	0.289	句子总数(后)	3.15	1.477	0.234
SFPs(先)	0.2	0.516	0.082	SFPs(后)	0.38	0.774	0.122
重复次数(先)	0.1	0.379	0.06	重复次数(后)	0.38	0.628	0.099

信息词(先)	6.95	4.449	0.703	信息词(后)	7.05	2.81	0.444
虚词(先)	5.5	4.846	0.766	虚词(后)	5.93	4.053	0.641
错误(先)	0.13	0.404	0.064	错误(后)	0.25	0.494	0.078
名词(先)	3.68	2.433	0.385	名词(后)	3.63	1.213	0.192
动词(先)	1.73	1.396	0.221	动词(后)	2.03	1.441	0.228
犹豫次数(先)	0.05	0.221	0.035	犹豫次数(后)	0.13	0.404	0.064
CE(先)	0.923	0.294	0.046	CE(后)	0.65	0.177	0.028

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(先) - 总字数(后)	-0.191	79	0.849
句子总数(先) - 句子总数(后)	-0.296	79	0.769
SFPs(先) - SFPs(后)	-1.096	79	0.280
重复次数(先) - 重复次数(后)	-2.218	79	0.032
信息词(先) - 信息词(后)	-0.136	79	0.893
虚词(先) - 虚词(后)	-0.499	79	0.621
错误(先) - 错误(后)	-1.152	79	0.256
名词(先) - 名词(后)	0.126	79	0.900
动词(先) - 动词(后)	-1.149	79	0.258
犹豫次数(先) - 犹豫次数(后)	-1.000	79	0.323
CE(先) - CE(后)	4.830	79	<0.001

表 28 是问题 (3) 不同产出顺序中语言测量参数的数据, 第一次参加数据为组 1 (先), 第二次参加为组 2 (后)。从数据可以看出, 总字数 $t(df=79) = 4.115, p < 0.001$, 句子总数 $t(df=79) = 3.112, p < 0.05$, SFPs $t(df=79) = 2.532, p < 0.05$, 信息词 $t(df=79) = 3.744, p < 0.001$, 虚词 $t(df=79) = 3.359, p < 0.05$, 动词 $t(df=79) = 3.568, p < 0.001$, 犹豫次数 $t(df=79) = 2.072, p < 0.05$, 得出结论是第一次的总字数、句子总数、句末语气词数、信息词数、虚词数、动词数和犹豫次数都高于第二次。

表 28 问题 (3) 产出顺序的成对样本统计与检验

1: 第一次组 (N=80)	均值	标准差	标准误差平均值	2: 第二次组 (N=80)	均值	标准差	标准误差平均值
总字数(先)	54.96	31.214	3.512	总字数(后)	44.33	22.328	2.512

句子总数(先)	5.58	2.894	0.326	句子总数(后)	4.67	1.693	0.19
SFPs(先)	0.82	0.888	0.1	SFPs(后)	0.52	0.86	0.097
重复次数(先)	0.62	1.017	0.114	重复次数(后)	0.49	0.714	0.08
信息词(先)	11.14	5.111	0.575	信息词(后)	9.24	3.874	0.436
虚词(先)	10.34	7.457	0.839	虚词(后)	8.06	4.81	0.541
错误(先)	0.44	0.813	0.091	错误(后)	0.43	0.692	0.078
名词(先)	4.66	2.412	0.271	名词(后)	4.13	2.09	0.235
动词(先)	3.51	2.309	0.26	动词(后)	2.51	1.51	0.17
犹豫次数(先)	0.33	0.729	0.082	犹豫次数(后)	0.16	0.436	0.049
CE(先)	0.68	0.241	0.027	CE(后)	0.712	0.255	0.029

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(先) - 总字数(后)	4.115	79	<0.001
句子总数(先) - 句子总数(后)	3.112	79	0.003
SFPs(先) - SFPs(后)	2.532	79	0.013
重复次数(先) - 重复次数(后)	1.134	79	0.260
信息词(先) - 信息词(后)	3.744	79	<0.001
虚词(先) - 虚词(后)	3.359	79	0.001
错误(先) - 错误(后)	0.134	79	0.894
名词(先) - 名词(后)	1.959	79	0.054
动词(先) - 动词(后)	3.568	79	<0.001
犹豫次数(先) - 犹豫次数(后)	2.072	79	0.042
CE(先) - CE(后)	-0.982	79	0.329

表 29 是问题 (4) 不同产出顺序中语言测量参数的数据, 第一次参加数据为组 1 (先), 第二次参加为组 2 (后)。从数据可以看出, $CE\ t(df=79) = -2.546, p < 0.05$, 得出结论是第一次的信息传输效率低于第二次。

表 29 问题 (4) 产出顺序的成对样本统计与检验

1: 第一次组 (N=80)	均值	标准差	标准误差平均值	2: 第二次组 (N=80)	均值	标准差	标准误差平均值
总字数(先)	53.74	31.323	3.502	总字数(后)	51.13	32.657	3.651
句子总数(先)	5.09	2.825	0.316	句子总数(后)	4.8	2.716	0.304
SFPs(先)	0.57	0.823	0.092	SFPs(后)	0.48	0.811	0.091
重复次数(先)	0.5	1.019	0.114	重复次数(后)	0.36	0.799	0.089
信息词(先)	12.31	7.244	0.81	信息词(后)	12.09	7.169	0.802

虚词(先)	8.36	5.648	0.632	虚词(后)	8.21	5.867	0.656
错误(先)	0.33	0.689	0.077	错误(后)	0.25	0.516	0.058
名词(先)	4.58	3.252	0.364	名词(后)	5.35	3.852	0.431
动词(先)	6.28	3.789	0.424	动词(后)	5.61	3.634	0.406
犹豫次数(先)	0.14	0.381	0.043	犹豫次数(后)	0.08	0.265	0.03
CE(先)	0.801	0.272	0.03	CE(后)	0.895	0.273	0.031

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(先) - 总字数(后)	0.727	79	0.469
句子总数(先) - 句子总数(后)	0.854	79	0.396
SFPs(先) - SFPs(后)	1.051	79	0.296
重复次数(先) - 重复次数(后)	1.242	79	0.218
信息词(先) - 信息词(后)	0.261	79	0.795
虚词(先) - 虚词(后)	0.225	79	0.823
错误(先) - 错误(后)	1.062	79	0.292
名词(先) - 名词(后)	-1.713	79	0.091
动词(先) - 动词(后)	1.544	79	0.127
犹豫次数(先) - 犹豫次数(后)	1.395	79	0.167
CE(先) - CE(后)	-2.546	79	0.013

表 30 和表 31 是问题 (5) 和问题 (6) 不同产出顺序中语言测量参数的数据, 第一次参加数据为组 1 (先), 第二次参加为组 2 (后)。从数据可以看出问题 (5) 中的各语言测量参数都没有体现显著差异。在问题 (6) 中, 总字数 $t(df=79) = 2.035, p < 0.05$, 句子总数 $t(df=79) = 2.426, p < 0.05$, 信息词 $t(df=79) = 2.441, p < 0.05$, 动词 $t(df=79) = 2.378, p < 0.05$, 得出结论是问题 (6) 第一次的总字数、句子总数、信息词数和动词数都多于第二次。

表 30 问题 (5) 产出顺序的成对样本统计与检验

1: 第一次组 (N=80)	均值	标准差	标准误差 平均值	2: 第二次组 (N=80)	均值	标准差	标准误差 平均值
总字数(先)	32.48	23.347	2.61	总字数(后)	30.78	17.891	2
句子总数(先)	2.83	2.097	0.235	句子总数(后)	2.63	1.513	0.169
SFPs(先)	0.45	0.84	0.094	SFPs(后)	0.38	0.718	0.08
重复次数(先)	0.29	0.766	0.086	重复次数(后)	0.24	0.534	0.06

信息词(先)	7.24	4.358	0.487	信息词(后)	6.61	3.509	0.392
虚词(先)	5.46	4.401	0.492	虚词(后)	5.13	3.487	0.39
错误(先)	0.23	0.551	0.062	错误(后)	0.16	0.434	0.049
名词(先)	3.68	2.151	0.24	名词(后)	3.31	1.913	0.214
动词(先)	3.35	2.234	0.25	动词(后)	3.04	1.789	0.2
犹豫次数(先)	0.13	0.369	0.041	犹豫次数(后)	0.08	0.265	0.03
CE(先)	0.907	0.355	0.04	CE(后)	0.871	0.397	0.044

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(先) - 总字数(后)	0.621	79	0.536
句子总数(先) - 句子总数(后)	0.783	79	0.436
SFPs(先) - SFPs(后)	0.748	79	0.457
重复次数(先) - 重复次数(后)	0.542	79	0.589
信息词(先) - 信息词(后)	1.168	79	0.246
虚词(先) - 虚词(后)	0.654	79	0.515
错误(先) - 错误(后)	0.869	79	0.387
名词(先) - 名词(后)	1.242	79	0.218
动词(先) - 动词(后)	1.234	79	0.221
犹豫次数(先) - 犹豫次数(后)	1.070	79	0.288
CE(先) - CE(后)	0.794	79	0.429

表 31 问题 (6) 产出顺序的成对样本统计与检验

1:第一次组 (N=80)	均值	标准差	标准误差 平均值	2:第二次组 (N=80)	均值	标准差	标准误差 平均值
总字数(先)	59.4	46.583	5.208	总字数(后)	49.39	26.93	3.011
句子总数(先)	5.3	3.843	0.43	句子总数(后)	4.28	2.037	0.228
SFPs(先)	0.76	1.034	0.116	SFPs(后)	0.63	0.96	0.107
重复次数(先)	0.57	1.251	0.14	重复次数(后)	0.54	0.826	0.092
信息词(先)	11.88	8.409	0.94	信息词(后)	9.63	5.043	0.564
虚词(先)	10.24	8.142	0.91	虚词(后)	8.81	5.232	0.585
错误(先)	0.63	1.372	0.153	错误(后)	0.43	0.742	0.083
名词(先)	7.78	5.77	0.645	名词(后)	6.54	3.489	0.39
动词(先)	2.15	2.129	0.238	动词(后)	1.58	1.508	0.169
犹豫次数(先)	0.39	0.849	0.095	犹豫次数(后)	0.38	0.663	0.074
CE(先)	0.689	0.317	0.035	CE(后)	0.664	0.274	0.031

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(先) - 总字数(后)	2.035	79	0.045

句子总数(先) - 句子总数(后)	2.426	79	0.018
SFPs(先) - SFPs(后)	0.992	79	0.324
重复次数(先) - 重复次数(后)	0.228	79	0.820
信息词(先) - 信息词(后)	2.441	79	0.017
虚词(先) - 虚词(后)	1.626	79	0.108
错误(先) - 错误(后)	1.380	79	0.172
名词(先) - 名词(后)	1.957	79	0.054
动词(先) - 动词(后)	2.378	79	0.020
犹豫次数(先) - 犹豫次数(后)	0.191	79	0.849
CE(先) - CE(后)	0.871	79	0.387

表 32 是问题 (7) 不同产出顺序中语言测量参数的数据, 第一次参加数据为组 1 (先), 第二次参加为组 2 (后)。从数据可以看出名词 $t(df=79) = 2.897, p < 0.05$, 得出结论是第一次的名词数多于第二次。

表 32 问题 (7) 产出顺序的成对样本统计与检验

1: 第一次组 (N=80)				2: 第二次组 (N=80)			
	均值	标准差	标准误差 平均值		均值	标准差	标准误差 平均值
总字数(先)	24.35	16.906	1.89	总字数(后)	22.11	12.597	1.408
句子总数(先)	2.53	1.509	0.169	句子总数(后)	2.26	1.099	0.123
SFPs(先)	0.29	0.578	0.065	SFPs(后)	0.25	0.436	0.049
重复次数(先)	0.3	0.701	0.078	重复次数(后)	0.3	0.624	0.07
信息词(先)	6.29	3.712	0.415	信息词(后)	5.79	2.67	0.299
虚词(先)	3.89	3.253	0.364	虚词(后)	3.34	3.828	0.428
错误(先)	0.19	0.393	0.044	错误(后)	0.13	0.333	0.037
名词(先)	2.64	1.788	0.2	名词(后)	1.99	1.471	0.164
动词(先)	2.9	2.114	0.236	动词(后)	2.91	1.593	0.178
犹豫次数(先)	0.06	0.244	0.027	犹豫次数(后)	0.14	0.347	0.039
CE(先)	0.992	0.464	0.052	CE(后)	1.072	0.549	0.061

	t	自由度	双侧 p
总字数(先) - 总字数(后)	1.245	79	0.217
句子总数(先) - 句子总数(后)	1.456	79	0.149
SFPs(先) - SFPs(后)	0.491	79	0.625
重复次数(先) - 重复次数(后)	0.000	79	1.000
信息词(先) - 信息词(后)	1.201	79	0.233
虚词(先) - 虚词(后)	1.135	79	0.260

错误(先) - 错误(后)	1.149	79	0.254
名词(先) - 名词(后)	2.897	79	0.005
动词(先) - 动词(后)	-0.048	79	0.962
犹豫次数(先) - 犹豫次数(后)	-1.754	79	0.083
CE(先) - CE(后)	-1.049	79	0.298

表 33 是问题 (8) 不同产出顺序中语言测量参数的数据。从数据可以看出 SFPs $t(df=79) = 2.080, p < 0.05$, 动词 $t(df=79) = 3.528, p < 0.001$, CE $t(df=79) = 2.120, p < 0.05$, 得出结论是第一次的句末语气词数、动词数和信息传输效率都高于第二次。

表 33 问题 (8) 产出顺序的成对样本统计与检验

1:第一次组 (N=80)	均值	标准差	标准误差 平均值	2:第二次组 (N=80)	均值	标准差	标准误差 平均值
总字数(先)	38.66	31.033	3.47	总字数(后)	35.7	31.412	3.512
句子总数(先)	3.11	2.068	0.231	句子总数(后)	2.96	2.264	0.253
SFPs(先)	0.53	0.886	0.099	SFPs(后)	0.3	0.604	0.068
重复次数(先)	0.39	0.646	0.072	重复次数(后)	0.33	0.652	0.073
信息词(先)	7.9	5.279	0.59	信息词(后)	6.94	5.058	0.565
虚词(先)	6.41	5.774	0.646	虚词(后)	5.99	5.46	0.61
错误(先)	0.28	0.551	0.062	错误(后)	0.2	0.403	0.045
名词(先)	3.59	3.113	0.348	名词(后)	3.03	3.326	0.372
动词(先)	3.48	2.074	0.232	动词(后)	2.59	1.77	0.198
犹豫次数(先)	0.21	0.412	0.046	犹豫次数(后)	0.18	0.414	0.046
CE(先)	0.861	0.694	0.078	CE(后)	0.694	0.269	0.030

	t	自由度	双侧 p
总字数(先) - 总字数(后)	0.815	79	0.418
句子总数(先) - 句子总数(后)	0.543	79	0.589
SFPs(先) - SFPs(后)	2.080	79	0.041
重复次数(先) - 重复次数(后)	0.685	79	0.496
信息词(先) - 信息词(后)	1.539	79	0.128
虚词(先) - 虚词(后)	0.705	79	0.483
错误(先) - 错误(后)	1.180	79	0.242
名词(先) - 名词(后)	1.356	79	0.179
动词(先) - 动词(后)	3.528	79	<0.001

犹豫次数(先) - 犹豫次数(后)	0.598	79	0.552
CE(先) - CE(后)	2.120	79	0.037

表 34 是问题 (9) 不同产出顺序中语言测量参数的数据。从数据可以看出总字数 $t(df=79) = 4.497, p < 0.001$, 句子总数 $t(df=79) = 5.023, p < 0.001$, 信息词 $t(df=79) = 4.195, p < 0.001$, 虚词 $t(df=79) = 4.736, p < 0.001$, 名词 $t(df=79) = 3.723, p < 0.001$, 动词 $t(df=79) = 3.808, p < 0.001$, 犹豫次数 $t(df=79) = 2.170, p < 0.05$, CE $t(df=79) = -2.762, p < 0.05$, 得出结论是第一次的总字数、句子总数、信息词数、虚词数、名词数、动词数和犹豫次数都高于第二次, 但是第一次的信息传输效率低于第二次。

表 34 问题 (9) 产出顺序的成对样本统计与检验

1:第一次组 (N=80)	均值	标准差	标准误差 平均值	2:第二次组 (N=80)	均值	标准差	标准误差 平均值
总字数(先)	39.66	40.486	4.526	总字数(后)	23.33	15.015	1.679
句子总数(先)	3.54	3.401	0.38	句子总数(后)	1.95	1.179	0.132
SFPs(先)	0.61	0.893	0.1	SFPs(后)	0.39	1.153	0.129
重复次数(先)	0.25	0.54	0.06	重复次数(后)	0.21	0.441	0.049
信息词(先)	8.19	7.741	0.865	信息词(后)	5.13	2.655	0.297
虚词(先)	7.74	7.452	0.833	虚词(后)	4.75	3.75	0.419
错误(先)	0.34	0.941	0.105	错误(后)	0.15	0.393	0.044
名词(先)	3.46	3.39	0.379	名词(后)	2.13	1.162	0.13
动词(先)	2.55	2.709	0.303	动词(后)	1.56	1.281	0.143
犹豫次数(先)	0.29	0.556	0.062	犹豫次数(后)	0.14	0.347	0.039
CE(先)	0.780	0.372	0.042	CE(后)	0.941	0.533	0.060

	t	自由度	双侧 p
总字数(先) - 总字数(后)	4.497	79	<0.001
句子总数(先) - 句子总数(后)	5.023	79	<0.001
SFPs(先) - SFPs(后)	1.481	79	0.143
重复次数(先) - 重复次数(后)	0.555	79	0.581
信息词(先) - 信息词(后)	4.195	79	<0.001

虚词(先) - 虚词(后)	4.736	79	<0.001
错误(先) - 错误(后)	1.832	79	0.071
名词(先) - 名词(后)	3.723	79	<0.001
动词(先) - 动词(后)	3.808	79	<0.001
犹豫次数(先) - 犹豫次数(后)	2.170	79	0.033
CE(先) - CE(后)	-2.762	79	0.007

表 35 是问题 (10) 产出顺序中不同语言测量参数的数据, 第一次参加数据为组 1 (先), 第二次参加为组 2 (后)。从数据可以看出总字数 $t(df=79) = 2.046, p < 0.05$, 信息词 $t(df=79) = 2.173, p < 0.05$, 名词 $t(df=79) = 2.182, p < 0.05$, 得出结论是第一次的总字数、信息词数、虚词数和名词数高于第二次。

表 35 问题 (10) 产出顺序的成对样本统计与检验

1: 第一次组 (N=80)	均值	标准差	标准误差 平均值	2: 第二次组 (N=80)	均值	标准差	标准误差 平均值
总字数(先)	40.2	33.979	3.799	总字数(后)	32.19	19.953	2.231
句子总数(先)	3.68	2.56	0.286	句子总数(后)	3.58	4.878	0.545
SFPs(先)	0.48	0.914	0.102	SFPs(后)	0.3	0.537	0.06
重复次数(先)	0.23	0.573	0.064	重复次数(后)	0.21	0.52	0.058
信息词(先)	8	6.053	0.677	信息词(后)	6.43	3.337	0.373
虚词(先)	8.64	7.392	0.826	虚词(后)	7	4.407	0.493
错误(先)	0.24	0.621	0.069	错误(后)	0.21	0.495	0.055
名词(先)	4.49	3.218	0.36	名词(后)	3.65	1.903	0.213
动词(先)	1.86	1.888	0.211	动词(后)	1.65	1.181	0.132
犹豫次数(先)	0.13	0.369	0.041	犹豫次数(后)	0.15	0.453	0.051
CE(先)	0.794	0.341	0.038	CE(后)	0.869	0.436	0.049

	t	自由度	双侧 p
总字数(先) - 总字数(后)	2.046	79	0.044
句子总数(先) - 句子总数(后)	0.159	79	0.874
SFPs(先) - SFPs(后)	1.560	79	0.123
重复次数(先) - 重复次数(后)	0.139	79	0.890
信息词(先) - 信息词(后)	2.173	79	0.033
虚词(先) - 虚词(后)	1.911	79	0.060

错误(先) - 错误(后)	0.307	79	0.760
名词(先) - 名词(后)	2.182	79	0.032
动词(先) - 动词(后)	0.865	79	0.390
犹豫次数(先) - 犹豫次数(后)	-0.363	79	0.717
CE(先) - CE(后)	-1.349	79	0.181

表 36 和表 37 是问题 (11) 和问题 (12) 不同产出顺序中语言测量参数的数据。从数据可以看出, 问题 (11) 的各语言测量参数无显著差异, 问题 (12) 中的总字数 $t(df=79) = 2.071, p < 0.05$, 重复次数 $t(df=79) = 2.352, p < 0.05$, 信息词 $t(df=79) = 2.045, p < 0.05$, 名词 $t(df=79) = 2.618, p < 0.05$ 。得出结论是问题 (12) 第一次的总字数、重复次数、信息词数和名词数都高于第二次。

表 36 问题 (11) 产出顺序的成对样本统计与检验

1:第一次组 (N=80)	均值	标准差	标准误差 平均值	2:第二次组 (N=80)	均值	标准差	标准误差 平均值
总字数(先)	60.51	36.634	4.096	总字数(后)	58.26	43.184	4.828
句子总数(先)	4.78	2.877	0.322	句子总数(后)	4.46	2.881	0.322
SFPs(先)	0.81	0.915	0.102	SFPs(后)	0.61	0.893	0.1
重复次数(先)	0.59	0.822	0.092	重复次数(后)	0.66	0.993	0.111
信息词(先)	11.34	6.639	0.742	信息词(后)	11.78	7.051	0.788
虚词(先)	9.38	6.817	0.762	虚词(后)	9.3	7.688	0.859
错误(先)	0.73	0.811	0.091	错误(后)	0.65	0.915	0.102
名词(先)	6.63	4.428	0.495	名词(后)	6.98	4.319	0.483
动词(先)	3.55	2.266	0.253	动词(后)	3.89	2.59	0.29
犹豫次数(先)	0.36	0.75	0.084	犹豫次数(后)	0.41	0.837	0.094
CE(先)	0.661	0.368	0.041	CE(后)	0.671	0.299	0.033

	t	自由度	双侧 p
总字数(先) - 总字数(后)	0.625	79	0.534
句子总数(先) - 句子总数(后)	1.058	79	0.293
SFPs(先) - SFPs(后)	1.747	79	0.084
重复次数(先) - 重复次数(后)	-0.555	79	0.581

信息词(先) - 信息词(后)	-0.650	79	0.518
虚词(先) - 虚词(后)	0.113	79	0.910
错误(先) - 错误(后)	0.760	79	0.450
名词(先) - 名词(后)	-0.818	79	0.416
动词(先) - 动词(后)	-1.024	79	0.309
犹豫次数(先) - 犹豫次数(后)	-0.601	79	0.550
CE(先) - CE(后)	-0.240	79	0.811

表 37 问题 (12) 产出顺序的成对样本统计与检验

1:第一次组 (N=80)	均值	标准差	标准误差 平均值	2:第二次组 (N=80)	均值	标准差	标准误差 平均值
总字数(先)	44.29	42.709	4.775	总字数(后)	35.6	24.249	2.711
句子总数(先)	3.45	2.877	0.322	句子总数(后)	2.86	1.867	0.209
SFPs(先)	0.59	0.852	0.095	SFPs(后)	0.41	0.741	0.083
重复次数(先)	0.49	1.006	0.112	重复次数(后)	0.19	0.506	0.057
信息词(先)	8.64	6.927	0.774	信息词(后)	7.33	4.542	0.508
虚词(先)	7.94	8.309	0.929	虚词(后)	6.4	4.365	0.488
错误(先)	0.46	0.711	0.079	错误(后)	0.45	0.778	0.087
名词(先)	4.25	3.627	0.405	名词(后)	3.4	2.488	0.278
动词(先)	3.03	3.093	0.346	动词(后)	2.44	1.908	0.213
犹豫次数(先)	0.1	0.343	0.039	犹豫次数(后)	0.16	0.436	0.049
CE(先)	0.782	0.357	0.040	CE(后)	0.766	0.315	0.035

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(先) - 总字数(后)	2.071	79	0.042
句子总数(先) - 句子总数(后)	1.947	79	0.055
SFPs(先) - SFPs(后)	1.470	79	0.145
重复次数(先) - 重复次数(后)	2.352	79	0.021
信息词(先) - 信息词(后)	2.045	79	0.044
虚词(先) - 虚词(后)	1.916	79	0.059
错误(先) - 错误(后)	0.129	79	0.897
名词(先) - 名词(后)	2.618	79	0.011
动词(先) - 动词(后)	1.813	79	0.074
犹豫次数(先) - 犹豫次数(后)	-1.092	79	0.278
CE(先) - CE(后)	0.350	79	0.727

表 38 是问题 (happy) 不同产出顺序中语言测量参数的数据, 第一次参加数据为组 1 (先), 第二次参加为组 2 (后)。从数据可以看出, 句子总数 $t(df=79)=2.023, p<0.05$, 重复次数 $t(df=79)=2.839, p<0.05$, 动词 $t(df=79)=4.121, p<0.001$, CE $t(df=79)=-2.619, p<0.05$ 。得出结论是第一次的句子总数、重复次数和动词数高于第二次, 信息传输效率低于第二次。

表 38 问题 (happy) 产出顺序的成对样本统计与检验

1:第一次组 (N=80)	均值	标准差	标准误差 平均值	2:第二次组 (N=80)	均值	标准差	标准误差 平均值
总字数(先)	36.35	28.288	3.163	总字数(后)	30.93	17.085	1.91
句子总数(先)	3.71	2.887	0.323	句子总数(后)	3.01	1.382	0.155
SFPs(先)	0.9	1.197	0.134	SFPs(后)	0.89	0.871	0.097
重复次数(先)	0.44	0.777	0.087	重复次数(后)	0.19	0.48	0.054
信息词(先)	7.68	5.199	0.581	信息词(后)	6.9	3.588	0.401
虚词(先)	5.01	4.944	0.553	虚词(后)	5.5	3.379	0.378
错误(先)	0.19	0.506	0.057	错误(后)	0.2	0.537	0.06
名词(先)	2.28	1.889	0.211	名词(后)	2.64	1.745	0.195
动词(先)	4.13	2.73	0.305	动词(后)	2.83	1.756	0.196
犹豫次数(先)	0.14	0.443	0.05	犹豫次数(后)	0.06	0.244	0.027
CE(先)	0.855	0.427	0.048	CE(后)	0.989	0.439	0.049

	t	自由度	双侧 p
总字数(先) - 总字数(后)	1.688	79	0.095
句子总数(先) - 句子总数(后)	2.023	79	0.046
SFPs(先) - SFPs(后)	0.084	79	0.933
重复次数(先) - 重复次数(后)	2.839	79	0.006
信息词(先) - 信息词(后)	1.224	79	0.225
虚词(先) - 虚词(后)	-0.856	79	0.394
错误(先) - 错误(后)	-0.168	79	0.867
名词(先) - 名词(后)	-1.329	79	0.188
动词(先) - 动词(后)	4.121	79	<0.001
犹豫次数(先) - 犹豫次数(后)	1.348	79	0.181
CE(先) - CE(后)	-2.619	79	0.011

表 39 是问题 (sad) 不同产出顺序中语言测量参数的数据。从数据可以看出, 信息词 $t(df=79)=3.060, p<0.05$, 虚词 $t(df=79)=3.125, p<0.05$, 名词 $t(df=79)=4.167, p<0.001$ 。得出结论是第一次的信息词数、虚词数和名词数高于第二次。

表 39 问题 (sad) 产出顺序的成对样本统计与检验

1:第一次组 (N=80)	均值	标准差	标准误差 平均值	2:第二次组 (N=80)	均值	标准差	标准误差 平均值
总字数(先)	39.84	25.095	2.806	总字数(后)	34.25	24.185	2.704
句子总数(先)	3.8	1.958	0.219	句子总数(后)	3.71	2.106	0.235
SFPs(先)	0.8	1.036	0.116	SFPs(后)	1.04	1.037	0.116
重复次数(先)	0.51	1.031	0.115	重复次数(后)	0.35	0.781	0.087
信息词(先)	8.03	4.438	0.496	信息词(后)	6.56	4.009	0.448
虚词(先)	6.69	4.924	0.55	虚词(后)	4.88	4.061	0.454
错误(先)	0.35	0.658	0.074	错误(后)	0.34	0.711	0.079
名词(先)	2.94	1.898	0.212	名词(后)	1.96	1.649	0.184
动词(先)	3.59	2.443	0.273	动词(后)	3.18	2.067	0.231
犹豫次数(先)	0.11	0.421	0.047	犹豫次数(后)	0.05	0.219	0.025
CE(先)	0.901	0.482	0.054	CE(后)	0.823	0.341	0.038

	t	自由度	双侧 p
总字数(先) - 总字数(后)	1.943	79	0.056
句子总数(先) - 句子总数(后)	0.349	79	0.728
SFPs(先) - SFPs(后)	-1.571	79	0.120
重复次数(先) - 重复次数(后)	1.493	79	0.139
信息词(先) - 信息词(后)	3.060	79	0.003
虚词(先) - 虚词(后)	3.125	79	0.002
错误(先) - 错误(后)	0.142	79	0.887
名词(先) - 名词(后)	4.167	79	<0.001
动词(先) - 动词(后)	1.412	79	0.162
犹豫次数(先) - 犹豫次数(后)	1.296	79	0.199
CE(先) - CE(后)	1.656	79	0.102

4.3 任务分类对比

4.3.1 长短图片描述任务整合

从第一次和第二次的产出顺序角度，对比整合后的短图片任务角度和长故事任务数据，第一次为组 1（先），第二次为组 2（后）。从表 40 可以看出， $p < 0.05$ 的有总字数 $t(df=79) = 2.047$ 、句子总数 $t(df=79) = 2.431$ 和句末语气词 $t(df=79) = 2.993$ ； $p < 0.001$ 的有犹豫次数 $t(df=79) = 4.028$ 和信息传输效率 $t(df=79) = -4.901$ 。数据说明第一次的总字数、句子总数、句末语气词数和犹豫次数比第二次的多，但是第二次的信息传输效率大于第一次。结合之前单独分析的两个图片描述任务的数据：在家庭活动短图片描述任务中，第一次的家庭图片描述任务的总字数、句子总数、犹豫次数、信息词数、虚词数和名词数都多于第二次的的数据；在长故事描述任务中，第一次的句末语气词数和犹豫次数多于第二次的的数据，第二次的信息输出效率比第一次的高。无论是第一次还是第二次，在短图片任务中，名词使用量都占有较大比例，因为描述图片通常涉及对场景中人物、物品及动作主体的识别和表述。在长故事任务方面，第一次的句末语气词数和犹豫次数多于第二次，说明首次尝试讲述复杂故事时，被试的叙述节奏会更慢，情绪表达也更为丰富。第二次的长故事描述任务显示出更高的信息输出效率，即在同等时间内传达的有效信息更多，是因为被试经过第一次的实践后，对任务有了更深的理解，能够更好地规划和组织自己的讲述内容，从而减少无效或冗余的信息传递，提高表达效率。练习效应在图片相互之间有联系的长故事任务中体现更明显，在单一图片

描述任务中体现不明显。长故事任务需要更高层次的语言组织能力和故事叙述能力，这使得被访者更加依赖于他们的语言处理和练习提升的效果。当描述长故事时，无论是中文二语学习组还是母语组，他们的表现都因为练习而有所提高。相比之下单一图片描述任务相对简单，不需要复杂的语言组织和叙述技巧，因此练习带来的效果不如长故事描述任务那么显著。因此为了提高语言学习者的语言能力，教师们在设计练习时应该考虑任务的复杂性和连贯性。更具挑战性的任务如连贯的长故事描述，可能会更有效地促进语言能力的提升。

表 40 图片任务第一次和第二次的成对样本统计与检验

1:第一次组 (N=80)	均值	标准差	标准误差 平均值	2:第二次组 (N=80)	均值	标准差	标准误差 平均值
总字数(先)	545.06	191.15	21.371	总字数(后)	513.26	176.722	19.758
句子总数(先)	51.54	15.965	1.785	句子总数(后)	48.06	15.173	1.696
SFPs(先)	8.98	4.889	0.547	SFPs(后)	7.44	4.133	0.462
重复次数(先)	4.43	4.839	0.541	重复次数(后)	3.98	3.93	0.439
信息词(先)	114.88	32.945	3.683	信息词(后)	115.98	38.64	4.32
虚词(先)	89.89	35.346	3.952	虚词(后)	84.89	31.309	3.5
错误(先)	8.36	8.521	0.953	错误(后)	7.61	8.192	0.916
名词(先)	70.43	24.953	2.79	名词(后)	70.4	25.736	2.877
动词(先)	43.94	13.152	1.47	动词(后)	43.16	12.403	1.387
犹豫次数(先)	7.29	5.494	0.614	犹豫次数(后)	5.16	4.835	0.541
CE(先)	1.348	0.451	0.050	CE(后)	1.524	0.466	0.052

	<i>t</i>	自由度	双侧 <i>p</i>
总字数(先) - 总字数(后)	2.047	79	0.044
句子总数(先) - 句子总数(后)	2.431	79	0.017
SFPs(先) - SFPs(后)	2.993	79	0.004
重复次数(先) - 重复次数(后)	1.024	79	0.309
信息词(先) - 信息词(后)	-0.336	79	0.738
虚词(先) - 虚词(后)	1.774	79	0.080
错误(先) - 错误(后)	1.504	79	0.136
名词(先) - 名词(后)	0.015	79	0.988

动词(先) - 动词(后)	0.597	79	0.552
犹豫次数(先) - 犹豫次数(后)	4.028	79	<0.001
CE(先) - CE(后)	-4.901	79	<0.001

4.3.2 正式问答和非正式问答整合

表 41 是从第一次和第二次的产出顺序角度对比整合后的问答数据。数据表明, $p < 0.05$ 的有总字数 $t(df=79) = 2.324$ 、句子总数 $t(df=79) = 2.394$ 、句末语气词 $t(df=79) = 2.102$ 和名词 $t(df=79) = 2.418$ 。数据说明第一次的总字数、句子总数、句末语气词数和名词数都比第二次多。

表 41 问题中不同产出顺序的成对样本检验

	t	自由度	双侧 p
总字数(先) - 总字数(后)	2.324	79	0.023
句子总数(先) - 句子总数(后)	2.394	79	0.019
SFPs(先) - SFPs(后)	2.102	79	0.039
重复次数(先) - 重复次数(后)	1.113	79	0.269
信息词(先) - 信息词(后)	1.760	79	0.082
虚词(先) - 虚词(后)	1.775	79	0.080
错误(先) - 错误(后)	1.431	79	0.156
名词(先) - 名词(后)	2.418	79	0.018
动词(先) - 动词(后)	0.961	79	0.340
犹豫次数(先) - 犹豫次数(后)	-0.990	79	0.325
CE(先) - CE(后)	-0.325	79	0.746

4.3.3 情绪相关问题整合

从第一次和第二次的产出顺序角度对比整合后的情绪问题数据, 第一次为组 1 (先), 第二次为组 2 (后)。表 42 得出结论: 总字数 $t(df=79) =$

2.103、重复次数 $t(df=79) = 3.323$ 、信息词 $t(df=79) = 2.388$ 、动词 $t(df=79) = 3.380$ 和犹豫次数 $t(df=79) = 2.013$ 呈现差异 ($p < 0.05$)，说明第一次的总字数、重复次数、信息词数、动词数和犹豫次数多于第二次。

表 42 情绪问题不同产出顺序的成对样本检验

	t	自由度	双侧 p
总字数(先) - 总字数(后)	2.103	79	0.039
句子总数(先) - 句子总数(后)	1.621	79	0.109
SFPs(先) - SFPs(后)	-0.994	79	0.323
重复次数(先) - 重复次数(后)	3.323	79	0.001
信息词(先) - 信息词(后)	2.388	79	0.019
虚词(先) - 虚词(后)	1.422	79	0.159
错误(先) - 错误(后)	0.000	79	1.000
名词(先) - 名词(后)	1.516	79	0.133
动词(先) - 动词(后)	3.380	79	0.001
犹豫次数(先) - 犹豫次数(后)	2.013	79	0.048
CE(先) - CE(后)	-0.786	79	0.434

第五章 性别和语言背景的差别探讨及整体交互作用

5.1 图片描述任务

5.1.1 单一图片任务

接下来，我们从另外两个角度，性别角度和语言背景角度，对看图说话中的家庭图片描述任务的语言数据进行单因素方差分析（One-Way ANOVA）。在本研究关于性别角度的数据分析中，各语言参数都没有显著差异，说明性别因素对于家庭活动描述图片任务没有显著影响。Hellman（2011）的研究调查了二语学习者在长时间的语言学习后是否达到了母语水平的词汇，结论是高水平

二语学习者能够达到母语水平的词汇量和知识的深度，并且词汇可能是成人二语学习中最成功的领域。表 43 是不同语言背景中（L1 和 L2）测量参数的数据，可以发现四个语言参数在 L1 和 L2 参与者中呈现差异：其中 $p < 0.001$ 的有重复次数 $F(1,158) = 11.93$ 、错误数 $F(1,158) = 31.23$ 和犹豫次数 $F(1,158) = 21.92$ ， $p < 0.05$ 的有名词数 $F(1,158) = 5.91$ 。数据说明在家庭图片描述任务中，母语者比二语者的重复次数、错误数、犹豫次数更少，名词数也整体较少。母语者在说话过程中的重复次数较少表明他们在表达思想时更加流畅和自信。相比之下，二语学习者会因为寻找正确的词汇或是结构上的不确定而更频繁地进行重复，也因为对语言规则的不熟悉而更容易犯错。母语者在表达过程中犹豫的次数更少，这反映了他们在语言产出上的自然和流畅性。二语学习者需要更多时间来思考适当的词汇、语法结构或是发音，从而导致更多的犹豫。母语者使用的名词总数也整体较少，这可能是因为他们能够更有效地使用语言，通过更丰富的词汇和语法结构来传达信息，而不是仅仅依赖于名词。二语学习者会更更多地依赖名词来描述事物或情况，因为相比动词变化和复杂的语法结构，名词通常更容易学习和使用。

表 43 单一图片不同语言背景的样本统计和ANOVA

语言背景		重复次数	错误	名词	犹豫次数
L1	平均值	0.43	0.94	15.94	0.46
	个案数	80	80	80	80
	标准偏差	0.897	1.306	7.472	0.871
L2	平均值	1.14	2.46	13.45	1.29
	个案数	80	80	80	80
	标准偏差	1.613	2.062	5.279	1.314
总计	平均值	0.78	1.70	14.69	0.88
	个案数	160	160	160	160
	标准偏差	1.349	1.883	6.568	1.186

		自由度	均方	<i>F</i>	显著性
重复次数	组间	1	20.306	11.925	<0.001
	组内	158	1.703		
	总计	159			
错误	组间	1	93.025	31.234	<0.001
	组内	158	2.978		
	总计	159			
名词	组间	1	247.506	5.914	0.016
	组内	158	41.851		
	总计	159			
犹豫次数	组间	1	27.225	21.916	<0.001
	组内	158	1.242		
	总计	159			

5.1.2 长故事任务

接下来我们从性别角度和语言背景角度，对看图说话中青蛙的故事语言数据进行单因素方差分析。表 44 是从不同性别角度分析的数据，从数据中可以看出，有两个测量参数：重复次数 $F(1,158) = 6.28$, $p < 0.05$ 和错误次数 $F(1,158) = 5.69$, $p < 0.05$ 呈现显著区别。说明在青蛙的故事任务中，男性组比女性组的重复次数更多，错误次数更多。导致性别差异的因素十分复杂，本数据呈现的性别差异也可能受到个别较大差异的影响。因为研究的样本量不大，无法充分代表广泛的人群，但是该发现能够为后续性别与语言相关的研究提供参考。

表 44 长故事不同性别的样本统计和ANOVA

性别		重复次数	错误
男	平均值	4.13	7.64
	个案数	80	80
	标准偏差	4.073	8.533
女	平均值	2.71	4.94
	个案数	80	80

	标准偏差	2.973	5.446
总计	平均值	3.42	6.29
	个案数	160	160
	标准偏差	3.625	7.263

		自由度	均方	<i>F</i>	显著性
重复次数	组间	1	79.806	6.276	0.013
	组内	158	12.716		
	总计	159			
错误	组间	1	291.600	5.691	0.018
	组内	158	51.235		
	总计	159			

表 45 是从语言背景角度分析的数据，从数据中可以看出，有九个测量参数都呈现显著区别，分别是总字数 $F(1,158) = 8.8, p < 0.05$ 、信息词总数 $F(1,158) = 26.26, p < 0.001$ 、虚词数 $F(1,158) = 9.47, p < 0.05$ 、错误次数 $F(1,158) = 119.41, p < 0.001$ 、名词数 $F(1,158) = 19.06, p < 0.001$ 、动词数 $F(1,158) = 24.1, p < 0.001$ 、犹豫次数 $F(1,158) = 17.17, p < 0.001$ 、句子长度 $F(1,158) = 40.21, p < 0.001$ 和信息输出效率 $F(1,158) = 150.71, p < 0.001$ 。将这九个测量参数的平均值进行比较后得出结论：母语组的总字数、信息词总数、虚词数、名词数、动词数、句子长度和信息输出效率数据比二语组的数据大，母语组的错误数、犹豫次数数据比二语组的数据小。

表 45 长故事不同语言背景的样本统计和ANOVA

语言背景		总字数	信息词	虚词	错误	名词	动词	犹豫次数	句子长度	CE
L1	平均值	471.88	106.09	83.31	1.54	62.70	41.05	3.89	11.371	0.767
	个案数	80	80	80	80	80	80	80	80	80
	标准偏差	169.420	30.808	30.179	1.622	21.899	11.706	3.611	1.173	0.178
L2	平均值	401.90	83.39	69.41	11.04	48.74	32.78	6.81	10.167	0.461
	个案数	80	80	80	80	80	80	80	80	80
	标准偏差	125.780	24.914	26.873	7.605	18.406	9.503	5.181	1.228	0.134
总计	平均值	436.89	94.74	76.36	6.29	55.72	36.91	5.35	10.769	0.614

个案数	160	160	160	160	160	160	160	160	160
标准偏差	152.819	30.160	29.325	7.263	21.346	11.410	4.687	1.341	0.220

		自由度	均方	<i>F</i>	显著性
总字数	组间	1	195860.025	8.798	0.003
	组内	158	22261.936		
	总计	159			
信息词	组间	1	20611.600	26.259	<0.001
	组内	158	784.920		
	总计	159			
虚词	组间	1	7728.400	9.465	0.002
	组内	158	816.485		
	总计	159			
错误	组间	1	3610.000	119.407	<0.001
	组内	158	30.233		
	总计	159			
名词	组间	1	7798.056	19.058	<0.001
	组内	158	409.166		
	总计	159			
动词	组间	1	2739.025	24.096	<0.001
	组内	158	113.669		
	总计	159			
犹豫次数	组间	1	342.225	17.165	<0.001
	组内	158	19.938		
	总计	159			
句子长度	组间	1	57.990	40.209	<0.001
	组内	158	1.442		
	总计	159			
CE	组间	1	3.749	150.710	<0.001
	组内	158	0.025		
	总计	159			

5.2 问答任务

从另外两个角度：性别角度和语言背景角度，对 14 个问题的语言数据进行单因素方差分析（One-Way ANOVA）后发现，在问题（1）性别角度中各语言参数都没有显著差异。根据表 46 问题（1）中不同语言背景中（L1 和 L2）测量参数的数据，可以发现多个语言参数在 L1 和 L2 参与者中呈现差异：其中 $p < 0.001$ 的有总字数 $F(1,158) = 47.379$ 、句子总数 $F(1,158) = 42.657$ 和句末语气词 $F(1,158) = 13.140$ ， $p < 0.05$ 的有重复次数 $F(1,158) = 6.831$ 、信息词 $F(1,158) = 9.228$ 、虚词 $F(1,158) = 8.424$ 、信息输出效率 $F(1,158) = 7.665$ 、错误 $F(1,158) = 9.427$ 和动词 $F(1,158) = 7.546$ 。数据说明在问题（1）中，母语组的重复次数和信息输出效率比二语组大，母语组的总字数、句子总数、句末语气词、错误、动词、虚词、信息词比二语组少。

表 46 问题（1）中语言背景的单因素方差分析及相关参数

		自由度	均方	F	显著性
总字数	组间	1	13340.756	47.379	<0.001
	组内	158	281.577		
	总计	159			
句子总数	组间	1	84.100	42.657	<0.001
	组内	158	1.972		
	总计	159			
SFPs	组间	1	2.025	13.140	<0.001
	组内	158	0.154		
	总计	159			
重复次数	组间	1	2.256	6.831	0.010
	组内	158	0.330		
	总计	159			
信息词	组间	1	122.500	9.228	0.003
	组内	158	13.275		
	总计	159			
虚词	组间	1	67.600	8.424	0.004
	组内	158	8.025		

	总计	159			
CE	组间	1	1.218	7.665	0.006
	组内	158	0.159		
	总计	159			
错误	组间	1	0.625	9.427	0.003
	组内	158	0.066		
	总计	159			
动词	组间	1	19.600	7.546	0.007
	组内	158	2.597		
	总计	159			

语言背景	总字数	句子总数	CE	SFPs	动词	错误	虚词	信息词	重复次数
L1 平均值	21.56	2.73	0.981	0.08	1.66	0.01	1.72	5.14	0.11
个案数	80	80	80	80	80	80	80	80	80
标准偏差	18.832	1.559	0.475	0.265	1.466	0.112	2.480	3.568	0.356
L2 平均值	39.83	4.17	0.807	0.30	2.36	0.14	3.03	6.89	0.35
个案数	80	80	80	80	80	80	80	80	80
标准偏差	14.440	1.230	0.304	0.488	1.745	0.347	3.146	3.718	0.731
总 平均值	30.69	3.45	0.894	0.19	2.01	0.08	2.38	6.01	0.23
计 个案数	160	160	160	160	160	160	160	160	160
标准偏差	19.071	1.577	0.407	0.407	1.644	0.264	2.898	3.737	0.585

在问题（2）中，从性别角度来看各语言参数都没有显著差异。表 47 是问题（2）中不同语言背景中（L1 和 L2）测量参数的数据，可以发现多个语言参数在 L1 和 L2 参与者中呈现差异：其中 $p < 0.001$ 的有重复次数 $F(1,158) = 15.301$ 和信息输出效率 $F(1,158) = 34.583$ ， $p < 0.05$ 的有错误 $F(1,158) = 5.291$ 和动词 $F(1,158) = 4.285$ 。数据说明在问题（2）中，母语组的信息输出效率比二语组大，母语组的重复次数、错误数和动词数比二语组少。

表 47 问题（2）中语言背景的单因素方差分析及相关参数

		自由度	均方	F	显著性
重复次数	组间	1	3.600	15.301	<0.001
	组内	158	0.235		
	总计	159			
错误	组间	1	0.900	5.291	0.023

	组内	158	0.170		
	总计	159			
动词	组间	1	6.400	4.285	0.040
	组内	158	1.494		
	总计	159			
CE	组间	1	2.520	34.583	<0.001
	组内	158	0.073		
	总计	159			

语言背景		重复次数	错误	动词	CE
L1	平均值	0.06	0.09	1.65	0.9552
	个案数	80	80	80	80
	标准偏差	0.291	0.326	1.213	0.328
L2	平均值	0.36	0.24	2.05	0.704
	个案数	80	80	80	80
	标准偏差	0.621	0.484	1.231	0.195
总计	平均值	0.21	0.16	1.85	0.830
	个案数	160	160	160	160
	标准偏差	0.506	0.418	1.235	0.297

在问题（3）中，从不同语言背景角度进行单因素方差分析（表 48）得出结论：重复次数 $F(1,158) = 8.795$, $p < 0.05$ 、错误 $F(1,158) = 16.755$, $p < 0.001$ 和信息输出效率 $F(1,158) = 32.167$, $p < 0.001$ ；从性别角度进行单因素方差分析（表 49）得出结论：错误 $F(1,158) = 0.045$, $p < 0.05$ 。数据说明在问题（3）中，男性组比女性组错误次数多，母语组的信息输出效率比二语组大，母语组的重复次数和错误数比二语组少。

表 48 问题（3）中语言背景的单因素方差分析及相关参数

		自由度	均方	F	显著性
重复次数	组间	1	6.400	8.795	0.003
	组内	158	0.728		
	总计	159			
错误	组间	1	8.556	16.755	<0.001
	组内	158	0.511		
	总计	159			

CE	组间	1	1.638	32.167	<0.001
	组内	158	0.051		
	总计	159			

语言背景		重复次数	错误	CE
L1	平均值	0.36	0.20	0.796
	个案数	80	80	80
	标准偏差	0.680	0.488	0.255
L2	平均值	0.76	0.66	0.594
	个案数	80	80	80
	标准偏差	0.997	0.885	0.191
总计	平均值	0.56	0.43	0.695
	个案数	160	160	160
	标准偏差	0.874	0.749	0.247

表 49 问题 (3) 中性别的单因素方差分析及相关参数

		自由度	均方	<i>F</i>	显著性
错误	组间	1	2.256	4.098	0.045
	组内	158	0.551		
	总计	159			

性别		错误
男	平均值	0.55
	个案数	80
	标准偏差	0.870
女	平均值	0.31
	个案数	80
	标准偏差	0.587
总计	平均值	0.43
	个案数	160
	标准偏差	0.749

在问题 (4) 中, 从不同语言背景角度进行单因素方差分析 (表 50) 得出结论: 重复次数 $F(1,158) = 5.594$, $p < 0.05$ 、错误 $F(1,158) = 22.145$, $p < 0.001$ 和信息输出效率 $F(1,158) = 27.435$, $p < 0.001$; 从性别角度进行单因素方差分

析（表 51）得出结论：重复次数 $F(1,158) = 4.020$, $p < 0.05$ 。数据说明在问题（4）中，男性组比女性组重复次数多，母语组的信息输出效率比二语组大，母语组的重复次数和错误数比二语组少。

表 50 问题（4）中语言背景的单因素方差分析及相关参数

		自由度	均方	F	显著性
重复次数	组间	1	4.556	5.594	0.019
	组内	158	0.814		
	总计	159			
错误	组间	1	7.225	22.145	<0.001
	组内	158	0.326		
	总计	159			
CE	组间	1	1.793	27.435	<0.001
	组内	158	0.065		
	总计	159			

语言背景		重复次数	错误	CE
L1	平均值	0.26	0.08	0.954
	个案数	80	80	80
	标准偏差	0.568	0.265	0.254
L2	平均值	0.60	0.50	0.742
	个案数	80	80	80
	标准偏差	1.143	0.763	0.259
总计	平均值	0.43	0.29	0.848
	个案数	160	160	160
	标准偏差	0.915	0.608	0.247

表 51 问题（4）中性别的单因素方差分析及相关参数

		自由度	均方	F	显著性
重复次数	组间	1	3.306	4.020	0.047
	组内	158	0.822		
	总计	159			

性别		重复次数
男	平均值	0.57
	个案数	80
	标准偏差	1.123
女	平均值	0.29

	个案数	80
	标准偏差	0.620
总计	平均值	0.43
	个案数	160
	标准偏差	0.915

在问题（5）中，从不同语言背景角度进行单因素方差分析（表 52）得出结论：重复次数 $F(1,158) = 7.247$, $p < 0.05$ 、错误 $F(1,158) = 12.018$, $p < 0.001$ 和信息输出效率 $F(1,158) = 30.167$, $p < 0.001$ ；从性别角度进行单因素方差分析，各语言测量参数之间无显著差异。数据说明在问题（5）中，母语组的信息输出效率比二语组大，母语组的重复次数和错误数比二语组少。

表 52 问题（5）中语言背景的单因素方差分析及相关参数

		自由度	均方	F	显著性
重复次数	组间	1	3.025	7.247	0.008
	组内	158	0.417		
	总计	159			
错误	组间	1	2.756	12.018	<0.001
	组内	158	0.229		
	总计	159			
CE	组间	1	3.597	30.167	<0.001
	组内	158	0.119		
	总计	159			

语言背景		重复次数	错误	CE
L1	平均值	0.13	0.06	1.04
	个案数	80	80	80
	标准偏差	0.402	0.332	0.400
L2	平均值	0.40	0.33	0.739
	个案数	80	80	80
	标准偏差	0.821	0.591	0.280
总计	平均值	0.26	0.19	0.889
	个案数	160	160	160
	标准偏差	0.659	0.495	0.376

在问题(6)中,从不同语言背景角度进行单因素方差分析(表53)得出结论:错误 $F(1,158)=7.713$, $p < 0.05$ 和信息输出效率 $F(1,158)=24.869$, $p < 0.001$; 从性别角度进行单因素方差分析(表54)得出结论,动词 $F(1,158)=4.240$, $p < 0.05$ 。数据说明在问题(6)中,男性组使用的动词数比女性组多,母语组的信息输出效率比二语组大,母语组的错误数比二语组少。

表 53 问题(6)中语言背景的单因素方差分析及相关参数

		自由度	均方	F	显著性
错误	组间	1	9.025	7.713	0.006
	组内	158	1.170		
	总计	159			
CE	组间	1	1.889	24.869	<0.001
	组内	158	0.076		
	总计	159			

语言背景		错误	CE
L1	平均值	0.29	0.785
	个案数	80	80
	标准偏差	1.058	0.326
L2	平均值	0.76	0.568
	个案数	80	80
	标准偏差	1.105	0.213
总计	平均值	0.53	0.676
	个案数	160	160
	标准偏差	1.104	0.296

表 54 问题(6)中性别的单因素方差分析及相关参数

		自由度	均方	F	显著性
动词	组间	1	14.400	4.240	0.041
	组内	158	3.396		
	总计	159			

性别	动词	
男	平均值	2.04
	个案数	80

	标准偏差	2.003
女	平均值	1.69
	个案数	80
	标准偏差	1.703
总计	平均值	1.86
	个案数	160
	标准偏差	1.862

在问题（7）中，从不同语言背景角度进行单因素方差分析（表 55）得出结论：错误 $F(1,158) = 11.286$, $p < 0.001$ 和信息输出效率 $F(1,158) = 26.762$, $p < 0.001$ ；从性别角度进行单因素方差分析， $p < 0.05$ 的有总字数 $F(1,158) = 9.532$ 、句子总数 $F(1,158) = 6.848$ 、信息词 $F(1,158) = 8.120$ 和动词 $F(1,158) = 6.463$ 。数据说明在问题（7）中，男性组的总字数、句子总数、信息词数和动词数都比女性组少，母语组的信息输出效率比二语组大，母语组的错误数比二语组少。

表 55 问题（7）中语言背景的单因素方差分析及相关参数

		自由度	均方	F	显著性
错误	组间	1	1.406	11.286	<0.001
	组内	158	0.125		
	总计	159			
CE	组间	1	5.958	26.762	<0.001
	组内	158	0.223		
	总计	159			

语言背景		错误	CE
L1	平均值	0.06	1.225
	个案数	80	80
	标准偏差	0.244	0.530
L2	平均值	0.25	0.839
	个案数	80	80
	标准偏差	0.436	0.406
总计	平均值	0.16	1.032
	个案数	160	160

标准偏差	0.364	0.509
------	-------	-------

表 56 问题 (7) 中性别的单因素方差分析及相关参数

		自由度	均方	<i>F</i>	显著性
总字数	组间	1	2009.306	9.532	0.002
	组内	158	210.792		
	总计	159			
句子总数	组间	1	11.556	6.848	0.010
	组内	158	1.688		
	总计	159			
信息词	组间	1	81.225	8.120	0.005
	组内	158	10.003		
	总计	159			
动词	组间	1	21.756	6.463	0.012
	组内	158	3.366		
	总计	159			

性别		总字数	句子总数	信息词	动词
男	平均值	19.69	2.13	5.33	2.54
	个案数	80	80	80	80
	标准偏差	11.917	1.151	2.613	1.614
女	平均值	26.77	2.66	6.75	3.28
	个案数	80	80	80	80
	标准偏差	16.720	1.432	3.630	2.031
总计	平均值	23.23	2.39	6.04	2.91
	个案数	160	160	160	160
	标准偏差	14.903	1.323	3.233	1.866

在问题 (8) 中, 从不同语言背景角度进行单因素方差分析 (表 57) 得出结论: 错误 $F(1,158) = 14.078$, $p < 0.001$ 、重复次数 $F(1,158) = 13.518$, $p < 0.001$ 和信息输出效率 $F(1,158) = 7.665$, $p < 0.05$; 从性别角度进行单因素方差分析, 各语言测量参数无显著差异。数据说明在问题 (8) 中, 母语组的信息输出效率比二语组大, 母语组的错误数和重复次数比二语组少。

表 57 问题 (8) 中语言背景的单因素方差分析及相关参数

		自由度	均方	<i>F</i>	显著性
重复次数	组间	1	5.256	13.518	<0.001
	组内	158	0.389		
	总计	159			
错误	组间	1	3.025	14.078	<0.001
	组内	158	0.215		
	总计	159			
CE	组间	1	2.076	7.665	0.006
	组内	158	0.271		
	总计	159			

语言背景		错误	重复次数	CE
L1	平均值	0.10	0.17	0.892
	个案数	80	80	80
	标准偏差	0.302	0.414	0.695
L2	平均值	0.37	0.54	0.664
	个案数	80	80	80
	标准偏差	0.582	0.779	0.241
总计	平均值	0.24	0.36	0.778
	个案数	160	160	160
	标准偏差	0.482	0.648	0.531

在问题 (9) 中, 从不同语言背景角度进行单因素方差分析 (表 58) 得出结论: 错误 $F(1,158) = 14.023$, $p < 0.001$ 、重复次数 $F(1,158) = 5.999$, $p < 0.05$; 从性别角度进行单因素方差分析, 犹豫次数 $F(1,158) = 5.769$, $p < 0.05$ 。数据说明在问题 (9) 中, 男性组比女性组的犹豫次数少, 母语组的错误数和重复次数比二语组少。

表 58 问题 (9) 中语言背景的单因素方差分析及相关参数

		自由度	均方	<i>F</i>	显著性
重复次数	组间	1	1.406	5.999	0.015
	组内	158	0.234		
	总计	159			
错误	组间	1	6.806	14.023	<0.001
	组内	158	0.485		

总计		159		
语言背景		错误	重复次数	
L1	平均值	0.04	0.14	
	个案数	80	80	
	标准偏差	0.191	0.381	
L2	平均值	0.45	0.33	
	个案数	80	80	
	标准偏差	0.967	0.569	
总计	平均值	0.24	0.23	
	个案数	160	160	
	标准偏差	0.725	0.492	

表 59 问题 (9) 中性别的单因素方差分析及相关参数

		自由度	均方	<i>F</i>	显著性
犹豫次数	组间	1	1.225	5.769	0.017
	组内	158	0.212		
	总计	159			

性别		犹豫次数
男	平均值	0.13
	个案数	80
	标准偏差	0.333
女	平均值	0.30
	个案数	80
	标准偏差	0.560
总计	平均值	0.21
	个案数	160
	标准偏差	0.468

在问题 (10) 中, 从不同语言背景角度进行单因素方差分析 (表 60) 得出结论: $CE F(1,158) = 17.080, p < 0.001$; 从性别角度进行单因素方差分析, 各语言测量参数无显著差异。数据说明在问题 (10) 中, 母语组的信息输出效率比二语组高。

表 60 问题 (10) 中语言背景的单因素方差分析及相关参数

		自由度	均方	<i>F</i>	显著性
CE	组间	1	2.382	17.080	<0.001
	组内	158	0.139		
	总计	159			

语言背景		CE
L1	平均值	0.953
	个案数	80
	标准偏差	0.448
L2	平均值	0.709
	个案数	80
	标准偏差	0.279
总计	平均值	0.831
	个案数	160
	标准偏差	0.392

在问题 (11) 中, 从不同语言背景角度进行单因素方差分析 (表 61) 得出结论: $p < 0.001$ 的有重复次数 $F(1,158) = 13.004$ 、信息词 $F(1,158) = 11.261$ 、错误 $F(1,158) = 43.598$ 、名词 $F(1,158) = 13.609$ 、CE $F(1,158) = 22.819$, $p < 0.05$ 的有总字数 $F(1,158) = 4.466$ 、句子总数 $F(1,158) = 4.110$ 、动词 $F(1,158) = 7.204$ 、犹豫次数 $F(1,158) = 9.420$; 从性别角度进行单因素方差分析, 各语言测量参数无显著差异。数据说明在问题 (11) 中, 母语组的信息词数、名词数、总字数、句子总数、动词数和信息输出效率比二语组高, 母语组的重复次数、错误数和犹豫次数比二语组低。

表 61 问题 (11) 中语言背景的单因素方差分析及相关参数

		自由度	均方	<i>F</i>	显著性
总字数	组间	1	6969.600	4.466	0.036
	组内	158	1560.597		
	总计	159			
句子总数	组间	1	33.306	4.110	0.044
	组内	158	8.104		

	总计	159			
重复次数	组间	1	10.000	13.004	<0.001
	组内	158	0.769		
	总计	159			
信息词	组间	1	493.506	11.261	<0.001
	组内	158	43.823		
	总计	159			
错误	组间	1	25.600	43.598	<0.001
	组内	158	0.587		
	总计	159			
名词	组间	1	240.100	13.609	<0.001
	组内	158	17.642		
	总计	159			
动词	组间	1	41.006	7.204	0.008
	组内	158	5.692		
	总计	159			
犹豫次数	组间	1	5.625	9.420	0.003
	组内	158	0.597		
	总计	159			
CE	组间	1	2.245	22.819	<0.001
	组内	158	0.098		
	总计	159			

语言背景	总字数	句子		重复次		信息		犹豫次		CE
		总数	总数	数	数	词	错误	名词	动词	
L1	平均值	34.91	3.80	0.18	7.05	0.16	3.97	1.85	0.14	0.953
	个案数	80	80	80	80	80	80	80	80	80
	标准偏差	32.063	5.137	0.497	5.518	0.561	2.756	1.969	0.470	0.448
L2	平均值	37.47	3.45	0.26	7.37	0.29	4.16	1.66	0.14	0.709
	个案数	80	80	80	80	80	80	80	80	80
	标准偏差	23.533	1.974	0.590	4.303	0.556	2.592	1.043	0.347	0.279
总计	平均值	36.19	3.63	0.22	7.21	0.23	4.07	1.76	0.14	0.831
	个案数	160	160	160	160	160	160	160	160	160
	标准偏差	28.064	3.883	0.546	4.935	0.560	2.669	1.573	0.412	0.392

在问题（12）中，从不同语言背景角度进行单因素方差分析（表 62）得出结论： $p < 0.001$ 的有错误 $F(1,158) = 26.649$ 、CE $F(1,158) = 16.870$ ；从性别

角度进行单因素方差分析，各语言测量参数无显著差异。数据说明在问题

(12) 中，母语组的信息输出效率比二语组高，母语组的错误数比二语组低。

表 62 问题 (12) 中语言背景的单因素方差分析及相关参数

		平方和	自由度	均方	<i>F</i>	显著性
错误	组间	12.656	1	12.656	26.649	<0.001
	组内	75.038	158	0.475		
	总计	87.694	159			
CE	组间	1.726	1	1.726	16.870	<0.001
	组内	16.168	158	0.102		
	总计	17.894	159			

语言背景		错误	CE
L1	平均值	0.18	0.878
	个案数	80	80
	标准偏差	0.497	0.343
L2	平均值	0.74	0.670
	个案数	80	80
	标准偏差	0.838	0.295
总计	平均值	0.46	0.774
	个案数	160	160
	标准偏差	0.743	0.335

在问题 (happy) 中，从不同语言背景角度进行单因素方差分析 (表 63) 得出结论：SFPs $F(1,158) = 9.176$, $p < 0.05$ 、重复次数 $F(1,158) = 12.197$, $p < 0.001$ 、错误 $F(1,158) = 15.790$, $p < 0.001$ 、CE $F(1,158) = 30.910$, $p < 0.001$ ；从性别角度进行单因素方差分析 (表 64)，重复次数 $F(1,158) = 4.822$, $p < 0.05$ 、错误 $F(1,158) = 6.929$, $p < 0.05$ 。数据说明在问题

(happy) 中，男性组的重复次数和错误数比女性组多，母语组的句末语气词和信息输出效率比二语组高，母语组的重复次数和错误数比二语组低。

表 63 问题 (happy) 中语言背景的单因素方差分析及相关参数

	自由度	均方	<i>F</i>	显著性
--	-----	----	----------	-----

SFPs	组间	1	9.506	9.176	0.003
	组内	158	1.036		
	总计	159			
重复次数	组间	1	4.900	12.197	<0.001
	组内	158	0.402		
	总计	159			
错误	组间	1	3.906	15.790	<0.001
	组内	158	0.247		
	总计	159			
CE	组间	1	4.966	30.910	<0.001
	组内	158	0.161		
	总计	159			

语言背景		重复次数	SFPs	错误	CE
L1	平均值	0.14	1.14	0.04	1.098
	个案数	80	80	80	80
	标准偏差	0.443	1.166	0.191	0.417
L2	平均值	0.49	0.65	0.35	0.746
	个案数	80	80	80	80
	标准偏差	0.779	0.843	0.677	0.384
总计	平均值	0.31	0.89	0.19	0.922
	个案数	160	160	160	160
	标准偏差	0.656	1.044	0.520	0.437

表 64 问题 (happy) 中性别的单因素方差分析及相关参数

		自由度	均方	F	显著性
重复次数	组间	1	2.025	4.822	0.030
	组内	158	0.420		
	总计	159			
错误	组间	1	1.806	6.929	0.009
	组内	158	0.261		
	总计	159			

性别		重复次数	错误
男	平均值	0.43	0.30
	个案数	80	80
	标准偏差	0.759	0.644
女	平均值	0.20	0.09
	个案数	80	80
	标准偏差	0.513	0.326

总计	平均值	0.31	0.19
	个案数	160	160
	标准偏差	0.656	0.520

在问题 (sad) 中, 从不同语言背景角度进行单因素方差分析 (表 65) 得出结论: 重复次数 $F(1,158) = 8.505, p < 0.05$ 、犹豫次数 $F(1,158) = 4.587, p < 0.05$ 、错误 $F(1,158) = 32.548, p < 0.001$ 、CE $F(1,158) = 37.194, p < 0.001$; 从性别角度进行单因素方差分析, 各语言测量参数无显著差异。数据说明在问题 (sad) 中, 母语组的信息输出效率比二语组高, 母语组的重复次数、犹豫次数和错误数比二语组低。

表 65 问题 (sad) 中语言背景的单因素方差分析及相关参数

		自由度	均方	F	显著性
重复次数	组间	1	6.806	8.505	0.004
	组内	158	0.800		
	总计	159			
错误	组间	1	12.656	32.548	<0.001
	组内	158	0.389		
	总计	159			
犹豫次数	组间	1	0.506	4.587	0.034
	组内	158	0.110		
	总计	159			
CE	组间	1	5.296	37.194	<0.001
	组内	158	0.142		
	总计	159			

语言背景		重复次数	犹豫次数	错误	CE
L1	平均值	0.23	0.03	0.06	1.044
	个案数	80	80	80	80
	标准偏差	0.527	0.157	0.244	0.446
L2	平均值	0.64	0.14	0.62	0.680
	个案数	80	80	80	80
	标准偏差	1.150	0.443	0.848	0.293
总计	平均值	0.43	0.08	0.34	0.862
	个案数	160	160	160	160

标准偏差	0.915	0.336	0.683	0.418
------	-------	-------	-------	-------

5.3 任务分类对比

5.3.1 长短图片描述任务整合

从性别角度对图片描述任务进行单因素方差分析。从表 66 可以看出，其中重复次数 $F(1,158) = 5.438$, $p < 0.05$ 和错误数 $F(1,158) = 5.408$, $p < 0.05$ 。数据说明在图片描述任务中，男性组的重复次数和错误次数都比女性组多。结合之前单独分析的两个图片描述任务的数据：在家庭活动短图片描述任务中，性别因素对于各语言参数没有显著影响；在长故事描述任务中，男性组比女性组的重复次数更多，错误次数更多。这一现象可能反映出在构建复杂情节、叙述连续性故事时，男性可能存在更强的修正倾向以及更高的犯错概率。这或许与男女在叙事策略、情感表达方式以及处理抽象思维方面的差异有关。但是实际个体差异会受到多种因素的影响，包括但不限于文化背景、教育经历、个人性格特质等。

表 66 图片描述任务中不同性别的报告和ANOVA

性别		重复次数	错误
男	平均值	5.00	9.50
	个案数	80	80
	标准偏差	4.739	9.610
女	平均值	3.40	6.48
	个案数	80	80
	标准偏差	3.899	6.558
总计	平均值	4.20	7.99
	个案数	160	160
	标准偏差	4.400	8.340

		自由度	均方	<i>F</i>	显著性
重复次数	组间	1	102.400	5.438	0.021
	组内	158	18.830		
	总计	159			
错误	组间	1	366.025	5.408	0.021
	组内	158	67.683		
	总计	159			

从语言背景角度对图片描述任务进行单因素方差分析，从表 67 可以看出，其中总字数 $F(1,158) = 6.640$, $p < 0.05$ 、重复次数 $F(1,158) = 5.796$, $p < 0.05$ 和虚词数 $F(1,158) = 6.834$, $p < 0.05$ 。其中 $p < 0.001$ 的有信息词数 $F(1,158) = 23.018$ 、错误数 $F(1,158) = 123.944$ 、名词数 $F(1,158) = 18.858$ 、动词数 $F(1,158) = 22.794$ 、犹豫次数 $F(1,158) = 23.088$ 和信息传输效率 $F(1,158) = 182.842$ 。数据说明在图片描述任务中，母语组的总字数、虚词数、信息词数、名词数、动词数和信息传输效率比二语组多，母语组的重复次数、错误数和犹豫次数比二语组少。结合之前单独分析的两个图片描述任务的数据：在家庭活动短图片描述任务中，母语者比二语者的重复次数、错误数、犹豫次数和名词数更少；在长故事描述任务中，母语组的总字数、信息词总数、虚词数、名词数、动词数、句子长度和信息输出效率比二语组的大，母语组的错误数、犹豫次数比二语组小。毋庸置疑，母语者在各种类型的图片描述任务中表现出更强的语言能力和流畅性，而二语者面临一定程度的语言挑战和表达困扰。

表 67 图片描述任务中不同语言背景的报告和ANOVA

语言背景	总字数	SFPs	重复次数	信息词	虚词	错误	名词	动词	犹豫次数	CE
L1 平均值	566.04	7.90	3.38	128.15	94.16	2.48	78.64	48.06	4.35	1.776

	个案数	80	80	80	80	80	80	80	80	80	80
	标准偏差	212.787	4.793	4.017	38.395	36.178	2.062	27.282	13.395	3.914	0.329
L2	平均值	492.29	8.51	5.03	102.70	80.61	13.50	62.19	39.04	8.10	1.096
	个案数	80	80	80	80	80	80	80	80	80	80
	标准偏差	142.298	4.360	4.631	27.872	28.991	8.614	20.092	10.317	5.780	0.306
总	平均值	529.16	8.21	4.20	115.42	87.39	7.99	70.41	43.55	6.22	1.436
计	个案数	160	160	160	160	160	160	160	160	160	160
	标准偏差	184.189	4.578	4.400	35.797	33.378	8.340	25.267	12.749	5.268	0.466

		自由度	均方	<i>F</i>	显著性
总字数	组间	1	217562.500	6.640	0.011
	组内	158	32763.502		
	总计	159			
重复次数	组间	1	108.900	5.796	0.017
	组内	158	18.789		
	总计	159			
信息词	组间	1	25908.100	23.018	<0.001
	组内	158	1125.538		
	总计	159			
虚词	组间	1	7344.100	6.834	0.010
	组内	158	1074.657		
	总计	159			
错误	组间	1	4862.025	123.944	<0.001
	组内	158	39.228		
	总计	159			
名词	组间	1	10824.100	18.858	<0.001
	组内	158	573.979		
	总计	159			
动词	组间	1	3258.025	22.794	<0.001
	组内	158	142.934		
	总计	159			
犹豫次数	组间	1	562.500	23.088	<0.001
	组内	158	24.363		
	总计	159			
CE	组间	1	18.485	182.842	<0.001
	组内	158	0.101		
	总计	159			

5.3.2 正式问答和非正式问答整合

以整合后的四个问题作为数据库, 从不同性别角度对其进行数据分析。表 68 数据表明, 总字数 $F(1,158) = 4.359$, $p < 0.05$ 、信息词 $F(1,158) = 4.699$, $p < 0.05$ 和动词 $F(1,158) = 6.134$, $p < 0.05$, 男性组的总字数、信息词数和动词数少于女性组。

表 68 问题中不同性别的报告和ANOVA

性别		总字数	信息词	动词
男	平均值	140.13	30.01	11.41
	个案数	80	80	80
	标准偏差	64.680	11.470	4.851
女	平均值	168.25	34.99	13.69
	个案数	80	80	80
	标准偏差	101.658	17.025	6.631
总计	平均值	154.19	32.50	12.55
	个案数	160	160	160
	标准偏差	86.094	14.683	5.903

		自由度	均方	F	显著性
总字数	组间	1	31640.625	4.359	0.038
	组内	158	7258.897		
	总计	159			
信息词	组间	1	990.025	4.699	0.032
	组内	158	210.696		
	总计	159			
动词	组间	1	207.025	6.134	0.014
	组内	158	33.750		
	总计	159			

以整合后的四个问题作为数据库, 从不同语言背景角度对其进行数据分析。表 69 数据表明, 重复次数 $F(1,158) = 12.908$, $p < 0.001$ 、信息词 $F(1,158) = 4.748$, $p < 0.05$ 、错误次数 $F(1,158) = 70.961$, $p < 0.001$ 、名词数 $F(1,158) = 4.748$, $p < 0.05$ 、

1,158) = 8.871, $p < 0.05$ 、犹豫次数 $F(1,158) = 10.209$, $p < 0.05$ 和信息传输效率 $F(1,158) = 63.686$, $p < 0.001$ 。母语组的信息词数、名词数和比信息传输效率比二语组的高; 母语组的重复次数、错误次数和犹豫次数比二语组的低。

表 69 问题中不同语言背景的报告和ANOVA

语言背景		重复次数	信息词	错误	名词	犹豫次数	CE
L1	平均值	1.00	35.00	0.59	18.31	0.44	3.926
	个案数	80	80	80	80	80	80
	标准偏差	1.646	16.271	1.040	9.124	0.840	0.981
L2	平均值	2.05	30.00	2.40	14.55	1.01	2.796
	个案数	80	80	80	80	80	80
	标准偏差	2.031	12.510	1.620	6.664	1.373	0.801
总计	平均值	1.52	32.50	1.49	16.43	0.72	3.361
	个案数	160	160	160	160	160	160
	标准偏差	1.916	14.683	1.633	8.185	1.171	1.0579

		自由度	均方	F	显著性
重复次数	组间	1	44.100	12.908	<0.001
	组内	158	3.416		
	总计	159			
信息词	组间	1	1000.000	4.748	0.031
	组内	158	210.633		
	总计	159			
错误	组间	1	131.406	70.961	<0.001
	组内	158	1.852		
	总计	159			
名词	组间	1	566.256	8.871	0.003
	组内	158	63.829		
	总计	159			
犹豫次数	组间	1	13.225	10.209	0.002
	组内	158	1.295		
	总计	159			
CE	组间	1	51.119	63.686	<0.001
	组内	158	0.803		
	总计	159			

5.3.3 情绪相关问题整合

二语习得的影响因素有社会文化因素和情感因素，社会文化因素中的性别因素是影响二语习得的因素之一（冯华，2012）。从不同性别角度对比本研究中情绪问题的数据，各语言参数都没有显著影响，说明性别因素对于情绪问题的数据没有显著影响。从不同语言背景角度对比情绪问题的数据（表 70）得出结论：其中句末语气词 $F(1,158) = 4.776, p < 0.05$ 、重复次数 $F(1,158) = 14.862, p < 0.001$ 、错误次数 $F(1,158) = 39.940, p < 0.001$ 、犹豫次数 $F(1,158) = 4.102, p < 0.05$ 和信息传输效率 $F(1,158) = 52.735, p < 0.001$ 。说明母语组的句末语气词和信息传输效率高于二语组，母语组的重复次数、错误次数和犹豫次数低于二语组。母语者更擅长通过句尾语气来准确地表达情绪，语言准确度更高，在情感交流时更为熟练与自然。

表 70 两个情绪问题中不同语言背景的报告和ANOVA

语言背景		SFPs	重复次数	错误	犹豫次数	CE
L1	平均值	2.09	0.36	0.10	0.10	2.143
	个案数	80	80	80	80	80
	标准偏差	1.722	0.767	0.302	0.377	0.703
L2	平均值	1.54	1.12	0.98	0.26	1.426
	个案数	80	80	80	80	80
	标准偏差	1.449	1.594	1.201	0.611	0.533
总计	平均值	1.81	0.74	0.54	0.18	1.784
	个案数	160	160	160	160	160
	标准偏差	1.610	1.304	0.977	0.512	0.718

		自由度	均方	F	显著性
SFPs	组间	1	12.100	4.776	0.030
	组内	158	2.533		
	总计	159			

重复次数	组间	1	23.256	14.862	<0.001
	组内	158	1.565		
	总计	159			
错误	组间	1	30.625	39.940	<0.001
	组内	158	0.767		
	总计	159			
犹豫次数	组间	1	1.056	4.102	0.045
	组内	158	0.258		
	总计	159			
CE	组间	1	20.520	52.735	<0.001
	组内	158	0.389		
	总计	159			

5.4 交互作用

5.4.1 单一图片任务的主体间效应检验

学习环境和学生的自主性学习存在复杂的交互影响，而这在中文教学尤其是线上教学中更加凸显（王改花、张李飞和傅钢善，2021）。除了分析线上线下面口语产出环境中口语表现的差异，我们还关注不同口语产出环境和其它影响因素（语言背景、产出顺序和性别）之间是否有显著的交互作用，是否存在显著的主效应，因此运行了一般线性模型（general linear model）下的单变量分析（univariate analysis），进行主体间效应检验（test of between-subjects effects）。表 71 是呈现出统计学意义上显著差异的部分数据检验结果，没有呈现交互作用的数据结果未列出。表中可以看出，语言背景和性别的交互作用对于句末语气词有影响，但是语言背景和性别分别对于句末语气词无显著影响；语言背景对于重复次数和名词有影响，口语产出环境、顺序、语言背景和性别的交互作用对于重复次数和虚词有影响；口语产出环境、顺序和性别的交互作用对于名词

有影响。

表 71 单一图片部分语言参数的主体间效应检验

句末语气词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.153	0.315
截距	1	98.263	<0.001
口语产出环境	1	1.075	0.302
语言背景*性别	1	4.956	0.028

重复次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.900	0.028
截距	1	56.222	<0.001
口语产出环境	1	2.819	0.095
语言背景	1	11.141	0.001
口语产出环境*产出 顺序*语言背景*性别	1	7.065	0.009

虚词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	0.783	0.694
截距	1	329.992	<0.001
口语产出环境	1	0.126	0.723
产出顺序*口语产 出环境*语言背景*性别	1	5.274	0.023

名词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.091	0.370
截距	1	796.828	<0.001
口语产出环境	1	0.160	0.690
语言背景	1	5.730	0.018
产出顺序*口语 产出环境*性别	1	4.433	0.037

因总字数和各参数都未达到统计学意义上的显著差异，在此未列出总字数检验结果。但是从图 1 我们可以看出，母语者在线上任务时的总字数多于线下，二语者的总字数线上线下差距不大。图 2 数据表明，母语组和二语组都是

线上任务时的句子总数较多于线下。在线上环境中，尤其是在本研究中不需要视频的情况下，被试可能感到更少的社交压力。这可能使母语者更放松，从而使用更多的词汇来表达自己的。而非母语者由于语言水平的限制，在线上和线下环境中的表达差异不大，非母语者主要集中于准确和恰当地使用语言。

图 1 单一图片产出环境和语言背景在总字数上的交互作用

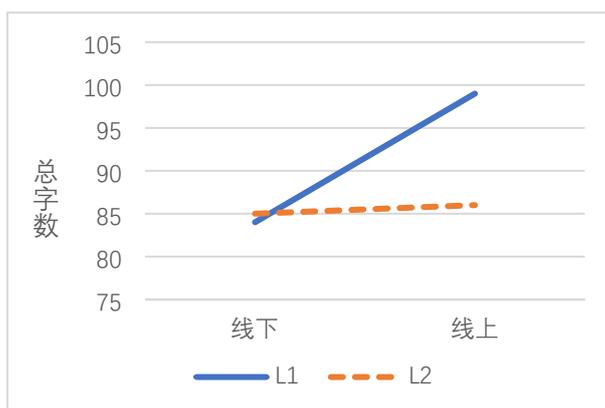
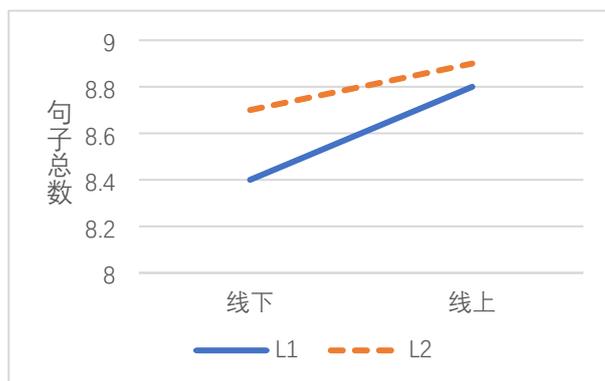


图 2 单一图片产出环境和语言背景在句子总数上的交互作用

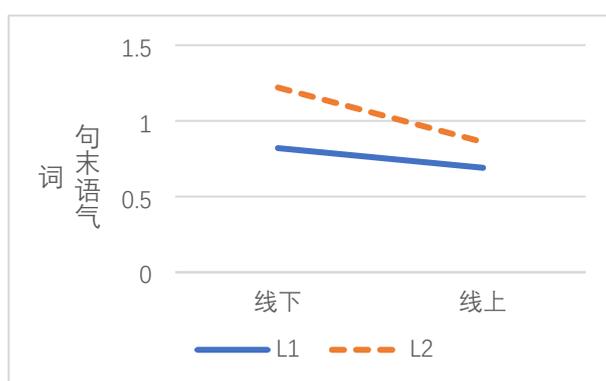


句子总数和产出顺序之间的关系达到显著 $F(df=1) = 6.089, p < 0.05$ ，第一次的句子总数多于第二次。第一次任务时，参与者可能会感到新鲜和好奇，因此更加投入和积极地参与对话。到了第二次，这种新鲜感会降低，参与者可能变得更加节约用词或直接。在第一次任务时，参与者可能没有完全明确的预期，因此他们可能会用更多的句子来表达自己的想法和观点。随着对过程的熟悉，第二次他们可能更加直接和高效地表达。

男性组线上线下的句末语气词大致持平，女性组线下的句末语气词明显多于线上。句末语气词和语言背景和性别的交互作用之间的关系达到显著 $F(df=1) = 4.956, p < 0.05$ 。图 3 数据表示，母语组和二语组都是线下的句末语气词多于线上，且二语组差距较大。线下环境提供了更丰富的非语言沟通机会，如面部表情和手势等，这可能鼓励参与者使用更多的语气词来配合这些非语言信号，增强表达的情感和强度。在线上，这种非语言沟通的缺失可能使得语气词的使用减少。对于非母语者，线下环境可能提供了更真实的语言使用场景，他们可能更多地模仿母语者的语气词使用习惯。在线上环境，非母语者可能更侧重于语法结构和词汇的准确性，从而减少语气词的使用。研究表明，女性在语言使用中往往更注重情感表达和社交互动 (Abdul-Ghani et al., 2022)。因此，在线下环境中，女性可能更倾向于使用语气词来表达情感和建立关系。相比之下，男性可能在语言使用上更为直接，无论是在线上还是线下环境。线下更侧重于建立关系和社交互动，这可能鼓励所有参与者，尤其是女性，使用更多的语气词来增强沟通的亲密度和表达的情感色彩。而在线上，沟通可能更任务驱动和目标导向，从而减少了语气词的使用。在一项关于线上和线下口语表现的研究中 (Warschauer, 1996)，一组英语为第二语言的学生以两种不同的产出环境（线上和线下）进行了课堂讨论。该研究的测量参数有国籍、语言能力、学习英语的时间以及学生态度等。研究结果显示，在线上讨论中，学生们分别参与讨论的时间比在线下更平等，也就是说学生们的线上课堂参与度更趋于平均。研究还发现，与面对面讨论相比，学生在线上讨论中使用的语言在词汇和句法上更加正式和复杂。这也解释了为什么在本研究的中被试线下的句末语气

词的使用比线上的多。在某些文化背景下，句末语气词可能被视为一种礼貌或温和的交流方式。在线下交流中，由于人与人之间的直接互动，被试可能更倾向于使用这些语气词来建立良好的社交关系。而在线上，由于缺乏直接的身体语言和面部表情，这种倾向可能会减弱。互动在线下汉语教学中比较容易实现，教师可以通过非言语交际手段和言语交际手段实现师生互动，如眼神、点头都可以理解为一个有意义的互动行为。但在线汉语教学的互动更常采取言语交际手段；而非言语交际手段的使用，则受到其本身的不明确性和网络技术 etc 客观因素的限制和影响（宋晖、谭紫格，2018）。

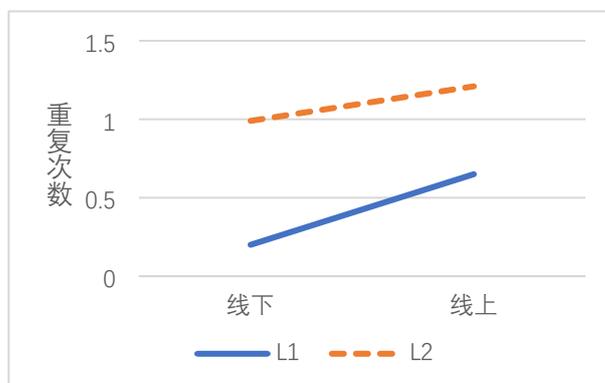
图 3 单一图片产出环境和语言背景在句末语气词上的交互作用



重复次数和语言背景之间的关系达到显著 $F(df=1) = 11.141, p < 0.05$ ，和产出顺序、口语产出环境和性别的交互作用之间的关系达到显著 $F(df=1) = 11.715, p < 0.05$ 。图 4 表示母语组和二语组都是线上的重复次数多于线下，且母语者差距更大。由此可以推测出在单一图片任务中，越接近母语的二语学习者，线上线下的不同口语产出环境对口语中的重复次数影响更大。男性组和女性组的对比未列出图表，同样两组都是线上的重复次数多于线下。线上环境缺乏即时的非语言反馈可能使得参与者重复自己的话，以确保信息的清晰度和

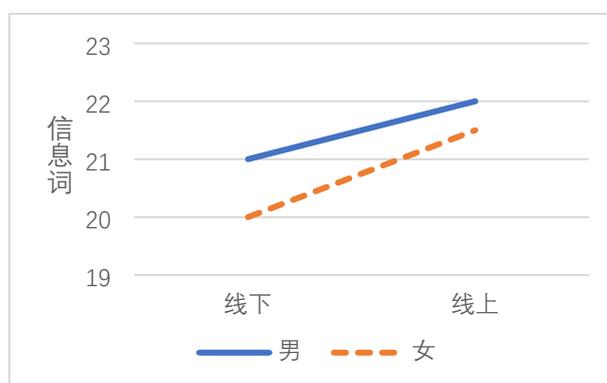
理解。对于非母语者来说，他们需要在思考和表达时更多地使用重复词语以加强语言的准确性和流利性。

图 4 单一图片产出环境和语言背景在重复次数上的交互作用



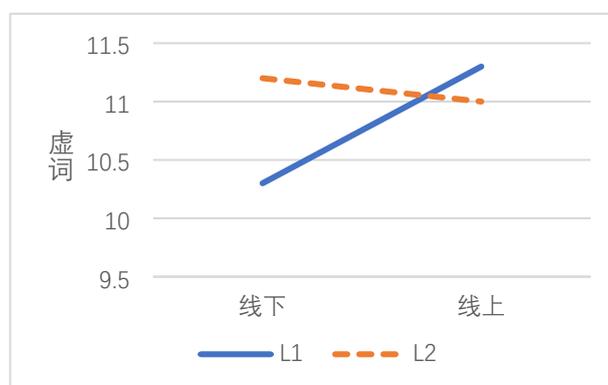
因信息词和各语言参数都未达到统计学意义上的显著差异，在此未列出信息词检验结果。虽然母语组和二语组都是线上的信息词较多于线下，但是信息词数多并不代表口语表达的信息丰富，因为重复次数和信息词数也有相关性，仅仅依靠信息词数多并不能说明口语表现能力好。因此我们结合后续的错误次数、信息传递效率等参数共同分析进行总结。从图 5 我们可以看出男性组和女性组都是线上任务时的信息词数量较多于线下，女性组线上线下信息词数差距更大。

图 5 单一图片产出环境和性别在信息词上的交互作用



虚词和产出顺序、口语产出环境、语言背景和性别的交互作用关系达到显著 $F(df=1) = 5.274, p < 0.05$, 男性组线上任务时的虚词数量多于线下, 女性组则是线下任务时的虚词数量较多于线上。图 6 表示母语组线上任务时的虚词数量多于线下, 二语者则是线下任务时的虚词数量较多于线上。虚词的数量与口语能力并不是正相关关系。虚词在语言中起到连接词与词组、调节句子结构、表达情感和语气等功能, 使用适当的虚词可以使语言更加流畅和自然。然而, 过多或过少的虚词使用都会影响语言的准确性和表达能力。在母语组中, 过多的虚词使用可能反映出口语习惯或个人语言风格, 但并不一定意味着口语能力好; 在二语组中虚词使用较多可能是因为尚未掌握虚词的正确用法或缺乏与母语者相同的语言习惯。

图 6 单一图片产出环境和语言背景在虚词上的交互作用

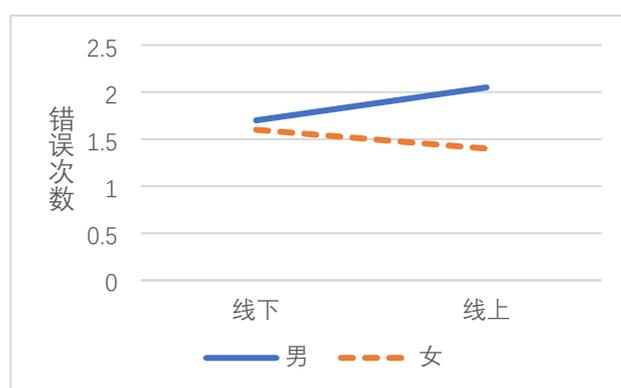


错误次数和语言背景之间的关系达到显著 $F(df=1) = 29.939, p < 0.001$ 。母语组线上的错误次数较多于线下, 二语组的错误次数远多于母语组, 线上线下的错误次数大致持平。对于母语组来说, 尽管他们在语言掌握上十分自如, 但是口语产出环境等外部因素依旧影响到他们的表现。非母语者的错误次数因

为语言能力的限制而整体多于母语者，且这种差异在线上 and 线下中保持一致。

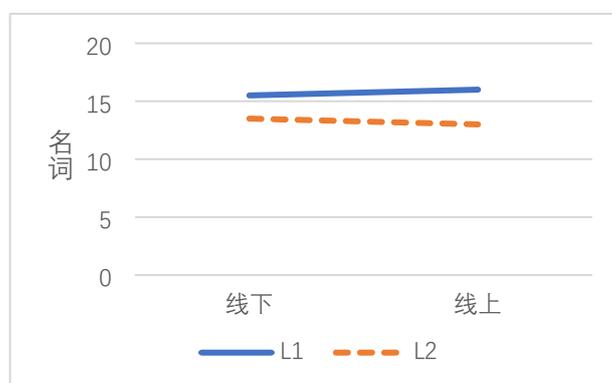
图 7 表示男性组线上的错误次数多于线下，但是女性组线上的错误次数少于线下。

图 7 单一图片产出环境和性别在错误次数上的交互作用



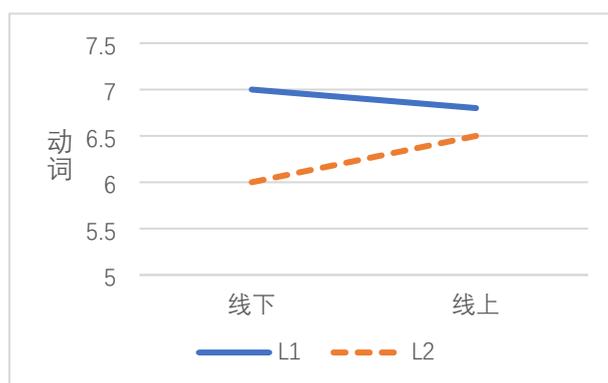
名词数量和语言背景之间的关系达到显著 $F(df=1) = 5.73, p < 0.05$ ，和产出顺序、口语产出环境和性别的交互作用之间的关系达到显著 $F(df=1) = 4.433, p < 0.05$ 。母语组的名词数量整体多于二语组，母语者拥有更丰富的词汇量，因此在描述和表达时使用更多的名词。图 8 表示母语组和二语组线上线下的名词量差别不大，母语组线上使用的名词数略多于线下，二语组相反。男生组和女生组在线上线下名词的使用中差异不大。

图 8 单一图片产出环境和语言背景在名词上的交互作用



因动词和各语言参数都未达到统计学意义上的显著差异，在此未列出动词次数检验结果，但是从图 9 我们可以看出，母语者线上任务时的动词数量多于线下，二语组相反且差距更大。女性者线上线下的动词数量大致持平，男性组线上任务时的动词数量多于线下。结合名词和动词的数据来看，母语组和二语组在线下任务时名词数差距和动词数差距几乎一致，但是在线上任务时名词数差距大于动词数差距。母语者和二语学习者在语言习得和使用中采用的策略不同。母语者可能更倾向于使用名词来构建细节丰富的叙述，而二语学习者可能在词汇选择上更为保守，特别是在对线上沟通环境感到不那么自信时。

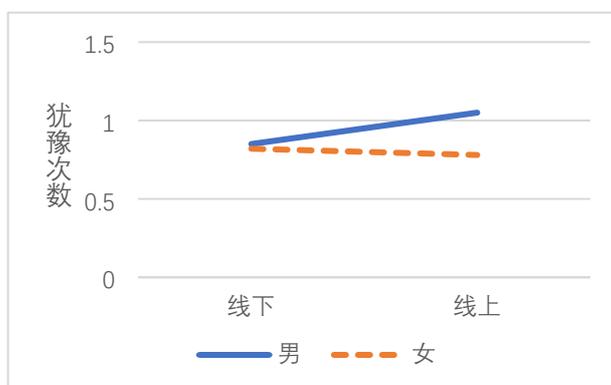
图 9 单一图片产出环境和语言背景在动词上的交互作用



犹豫次数和产出顺序之间的关系达到显著 $F(df=1) = 7.976, p < 0.05$ ，和语言背景之间的关系达到显著 $F(df=1) = 21.858, p < 0.001$ 。母语组线上线下的犹豫次数较为接近，二语组线上的犹豫次数多于线下。非母语者可能在线上环境中感到更多的不确定性，因此犹豫次数更多。二语者在使用非母语沟通时已经面临较高的认知负荷，线上环境进一步增加这种负荷。因为他们需要同时处理语言表达和技术操作，这可能导致他们在表达时更加犹豫，线下环境提供

了更多即时的反馈，帮助他们更快地做出反应。图 10 得出女性组线上线下的犹豫次数较为接近，男性组线上的犹豫次数略多于线下。

图 10 单一图片产出环境和性别在犹豫次数上的交互作用



5.4.2 长故事图片任务的主体间效应检验

为了解在长故事描述任务中，线下和线上的口语产出环境和其它影响因素（语言背景、产出顺序和性别）之间是否有显著的交互作用，我们同样对该任务的各项数据进行了主体间效应检验。表 72 是长故事任务中呈现出统计学意义上显著差异的部分数据检验结果。

表 72 长故事部分语言参数的主体间效应检验

总字数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.278	0.223
截距	1	1324.833	<0.001
口语产出环境	1	0.043	0.835
语言背景	1	8.583	0.004
语言背景*产出顺序*口语产出环境	1	4.068	0.046

重复次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.665	0.064
截距	1	147.944	<0.001
口语产出环境	1	0.008	0.931
性别	1	6.711	0.011
口语产出环境*产出 顺序	1	4.219	0.042
口语产出环境*产出 顺序*语言背景*性别	1	5.817	0.017

信息词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	3.534	<0.001
截距	1	1935.617	<0.001
口语产出环境	1	0.322	0.571
语言背景	1	27.969	<0.001
语言背景*产出 顺序*口语产出环境	1	15.443	<0.001

虚词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.490	0.116
截距	1	1117.777	<0.001
口语产出环境	1	0.039	0.844
语言背景	1	9.550	0.002
语言背景*性别*产出 顺序*口语产出环境	1	3.936	0.049

错误次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	10.848	<0.001
截距	1	227.184	<0.001
口语产出环境	1	0.118	0.732
语言背景	1	129.261	<0.001
性别	1	11.047	0.001
语言背景*性别	1	12.070	<0.001

名词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	2.649	0.001
截距	1	1249.663	<0.001
口语产出环境	1	0.021	0.885
语言背景	1	19.636	<0.001
语言背景*产出 顺序*口语产出环境	1	10.456	0.002

信息输出效率	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	13.371	<0.001
截距	1	2674.561	<0.001
口语产出环境	1	6.287	0.013
语言背景	1	168.063	<0.001
产出顺序	1	10.069	0.002
语言背景*口语产出 环境*性别	1	4.388	0.038

数据表明，总字数和语言背景之间的关系达到显著 $F(df=1) = 8.583$, $p < 0.05$ ；和语言背景、产出顺序和口语产出环境之间的关系达到显著 $F(df=1) = 4.068$, $p < 0.05$ 。因句子总数同各项测量参数都未达到统计学意义上的显著，在此未展示其数据描述。如图 11，母语者在线上任务时的总字数较多于线下，二语者的总字数线上线下差距不大，该发现与家庭图片描述任务的语言数据结论一致。图 12 表明母语者在线上任务时的句子总数较多于线下，二语者的句子总数线上线下一致。由于线上沟通缺少非言语交流的元素，母语者可能通过增加语言输出（总字数和句子总数）来补偿这一点，以确保他们的信息被准确理解。二语者的语言能力有限，语言产出相对稳定，不受线上或线下环境影响太大，语言能力限制了他们在语言产出上的灵活性。二语者在进行语言产出时面临的认知负荷较高，不管是线上还是线下，他们可能都已经达到了自己在语言表达上的最大能力，因此在两种环境中的语言表现差异不大。女性组线上任务

时的总字数和句子总数都较多于线下，但是男性组相反。

图 11 长故事产出环境和语言背景在总字数上的交互作用

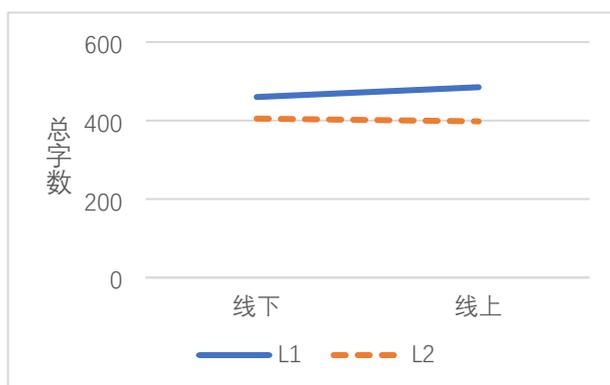
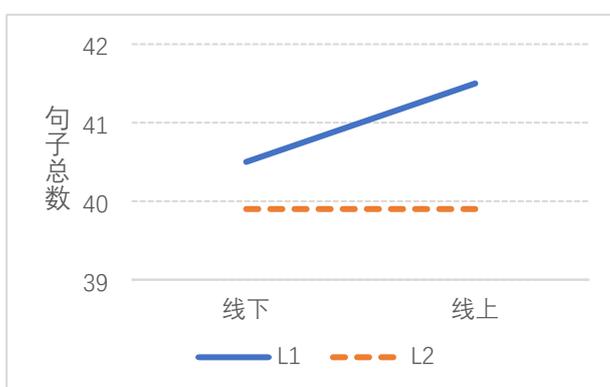


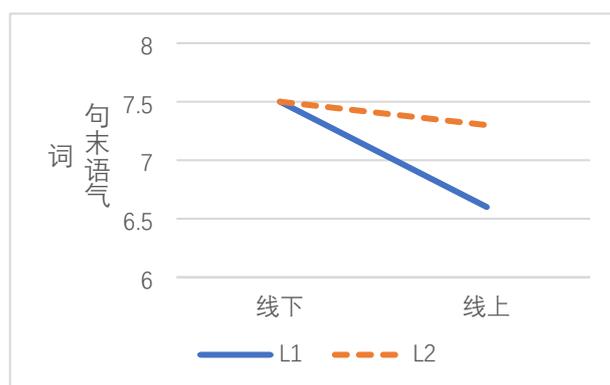
图 12 长故事产出环境和语言背景在句子总数上的交互作用



句末语气词同各项测量参数也未达到统计学意义上的显著，在此不展示其数据描述。男性组线下的句末语气词多于线上，女性组数据持平。图 13 所示，母语者和二语者都是线下的句末语气词多于线上，且母语者线上线下的差距更大。在面对面的交流中，语气词可以与非言语信号（如肢体语言、面部表情和语调）结合，以增强说话者的意图和情感表达。线下交流中更频繁使用语气词可能是为了加强这种即时互动的效果。在线交流使得说话者更依赖于文字或者简洁的语言表达，尤其是在没有视频只有语音传递的交流中，语气词可能被视为不必要的附加元素，从而减少其使用。其次，由于网络延迟或者交流工具的限制，线上交流中互动不如线下即时，从而减少了语气词的必要性或者使用频

率。线上交流可能促使参与者采用更直接或更正式的沟通风格，相比之下，线下交流更为随意，被试更倾向于使用语气词。母语者更有能力运用语言的各种细微之处包括语气词的使用，而二语者可能更专注于基本的语法和词汇，以确保信息的清晰传达。母语者在语言习得方面拥有更深层的直觉和习得，他们能够更敏感地感知到线上和线下环境中沟通方式的差异，并据此调整自己的语言使用，这导致了母语者线上线下语气词使用差距更大。这些解释都指出，线下环境倾向于鼓励更自然、丰富的语言表达，包括频繁使用语气词，而线上环境可能促成更简洁、直接的沟通方式。对于语言教学和交流策略的设计而言，理解这些差异非常重要，可以帮助教师们更好地适应和优化不同环境下的教学。

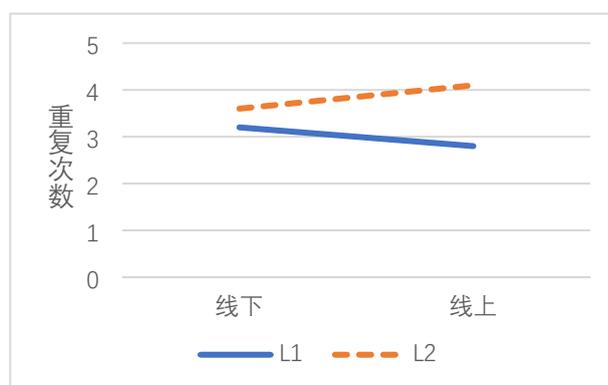
图 13 长故事产出环境和语言背景在句末语气词上的交互作用



重复次数和性别之间的关系达到显著 $F(df=1) = 6.711, p < 0.05$ ；和产出顺序、口语产出环境交互作用之间的关系达到显著 $F(df=1) = 4.219, p < 0.05$ ；和语言背景、产出顺序、口语产出环境和性别的交互作用之间的关系达到显著 $F(df=1) = 5.817, p < 0.05$ 。如图 14，二语组线上的重复次数多于线下，母语组相反。对于母语者，线下交流是最轻松、最自然的沟通方式；在线上交流时，他们更为谨慎和有目的性，语言输出更加正式。对于非母语者，线下交流

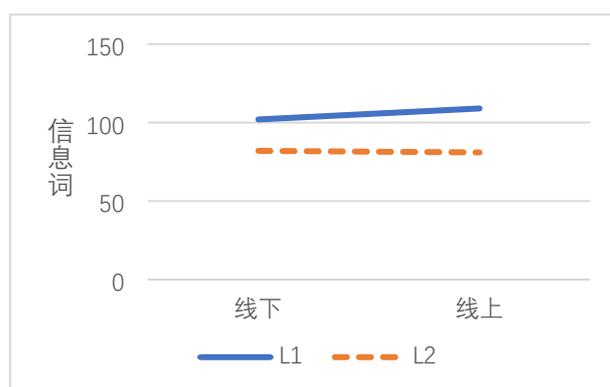
可能会带来更大的压力。他们可能担心自己的发音不准确或词汇使用不当，更少地重复语句以避免暴露自己的不足。而在线上，由于缺乏即时的面对面反馈，他们可能不确定自己是否被理解，从而更多地重复语句来确保信息传达的准确性。从性别角度来看，男性组线下的重复次数比线上多，然而女性组正好相反，线上的重复次数比线下多。影响因素有他们的社会化经验、沟通习惯、感知的压力和焦虑以及认知策略的影响。

图 14 长故事产出环境和语言背景在重复次数上的交互作用



信息词和语言背景之间的关系达到显著 $F(df=1) = 27.969, p < 0.001$ ；和语言背景、产出顺序和口语产出环境的交互作用之间的关系达到显著 $F(df=1) = 15.443, p < 0.001$ 。男性组线下的信息词更多，而女性组线上的信息词更多。图 15 表明母语者的线上的信息词数比线下略多，二语者线上线下的信息词数大概持平。

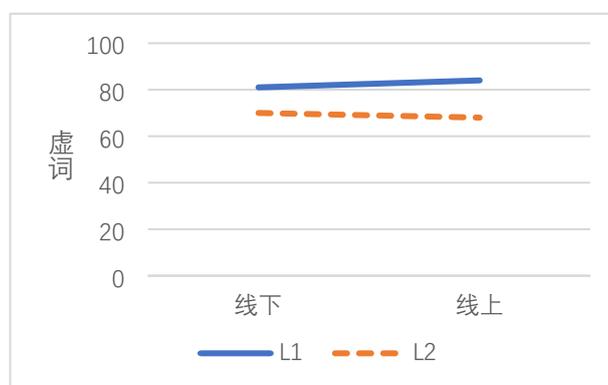
图 15 长故事产出环境和语言背景在信息词上的交互作用



虚词和语言背景之间的关系达到显著 $F(df=1) = 9.55, p < 0.05$, 和语言背景、性别、产出顺序和口语产出环境之间的关系达到显著 $F(df=1) = 3.936, p < 0.05$ 。男性组线下的虚词数较多于线上, 女性组线上的虚词数较多于线下。

图 16 表示母语组线上的虚词数较多于线下, 二语组线下的虚词数较多于线上。

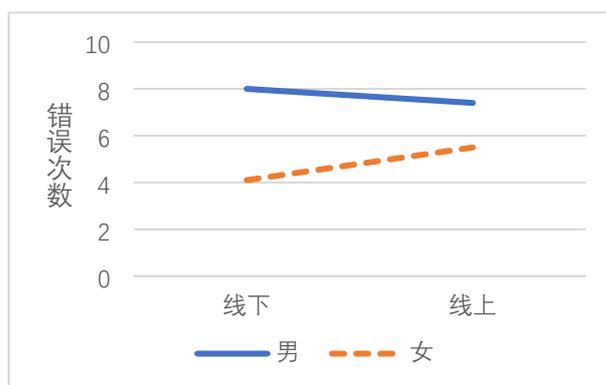
图 16 长故事产出环境和语言背景在虚词上的交互作用



错误次数和语言背景之间的关系达到显著 $F(df=1) = 129.261, p < 0.001$; 和性别之间的关系达到显著 $F(df=1) = 11.047, p < 0.05$; 和语言背景和性别之间的交互作用达到显著 $F(df=1) = 12.07, p < 0.001$ 。母语组和二语组都是线上的错误次数略多于线下的错误次数。图 17 显示男性组线下的错误次数比线上的

多，女性组相反。

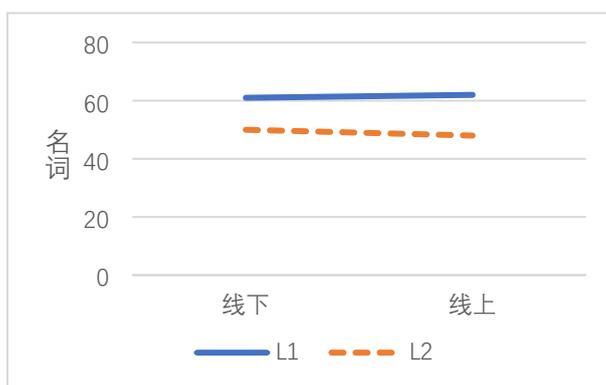
图 17 长故事产出环境和性别在错误次数上的交互作用



名词数量和语言背景之间的关系达到显著 $F(df=1) = 19.636, p < 0.001$ ；和语言背景、产出顺序、口语产出环境之间交互作用的关系达到显著 $F(df=1) = 10.456, p < 0.05$ 。女性组线上的名词数较多于线下，男性组相反。在线下任务时，男性组的名词数多于女性组，但是在线上任务时名词数量大致持平。

图 18 表示母语组线上的名词数较多于线下，二语组相反。

图 18 长故事产出环境和语言背景在名词上的交互作用

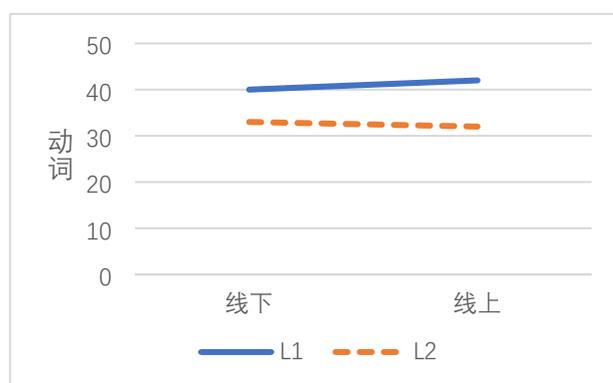


动词数量和语言背景之间的关系达到显著 $F(df=1) = 23.767, p < 0.001$ 。

女性组线上的动词数量多于线下，男性组相反。如图 19 所示，母语者线上的动

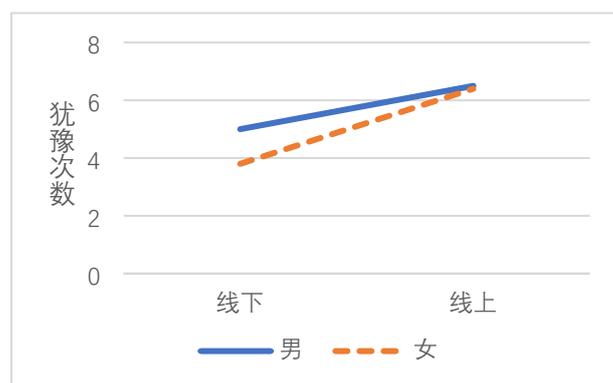
词数量较多于线下，二语组大致持平。

图 19 长故事产出环境和语言背景在动词上的交互作用



犹豫次数和语言背景之间的关系达到显著 $F(df=1) = 18.335, p < 0.001$ ；和产出顺序之间的关系达到显著 $F(df=1) = 5.757, p < 0.05$ ；和口语产出环境之间的关系达到显著 $F(df=1) = 8.753, p < 0.05$ 。母语组和二语组线上的犹豫都多于线下。如图 20，男性组和女性组线上的犹豫次数都多于线下，且女性组线上线下的犹豫次数数量的差异更大。

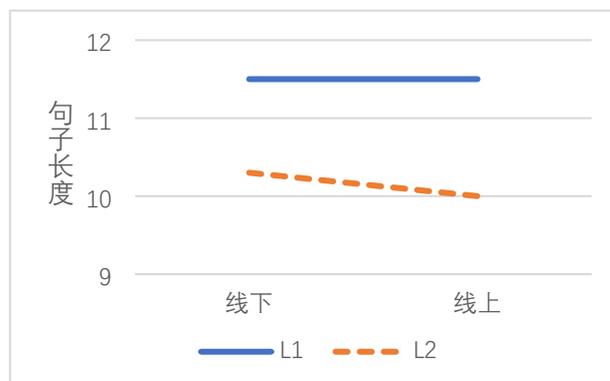
图 20 长故事产出环境和性别在犹豫次数上的交互作用



句子长度和语言背景之间的关系达到显著 $F(df=1) = 38.606, p < 0.001$ 。女性组线上和线下的句子长度持平，男性组线下的句子长度较多于线

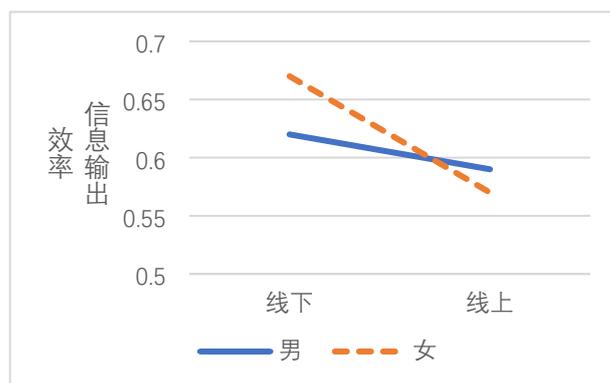
下。图 21 表示母语组线上和线下的句子长度持平，二语组线下的句子长度较多于线上。

图 21 长故事产出环境和语言背景在句子长度上的交互作用



信息输出效率和语言背景之间的关系达到显著 $F(df=1) = 168.063, p < 0.001$ ；和产出顺序之间的关系达到显著 $F(df=1) = 10.069, p < 0.05$ ；和口语产出环境之间的关系达到显著 $F(df=1) = 6.287, p < 0.05$ ；和语言背景、性别和口语产出环境之间交互作用的关系达到显著 $F(df=1) = 4.388, p < 0.05$ 。母语组和二语组都是线下的信息输出效率高于线上。图 22 表示男性组和女性组都是线下的信息输出效率高于线上。在线上任务中男性组的信息输出效率高于女性组；在线下任务中男性组的信息输出效率低于女性组。

图 22 长故事产出环境和语言背景在信息输出效率上的交互作用



5.4.3 问答任务的主体间效应检验

(1) 问题 1

和看图说话内容相同的，为研究线下和线上的口语产出环境和其它影响因素（语言背景、产出顺序和性别）之间是否有显著的交互作用，是否存在显著的主效应，我们同样运行了一般线性模型下的单变量分析，进行主体间效应检验。表 73 列出了达到统计学意义上显著差异的交互作用数据 ($p < 0.05$)，无显著差异的数据未列出检验结果。

表 73 问题 (1) 部分语言参数的主体间效应检验

总字数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	7.968	<0.001
截距	1	684.457	<0.001
产出环境	1	12.702	<0.001
语言背景	1	60.577	<0.001
产出环境*语言背景	1	35.536	<0.001

句子总数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	6.526	<0.001
截距	1	1158.745	<0.001
产出环境	1	8.047	0.005
语言背景	1	51.171	<0.001
产出环境*语言背景	1	28.126	<0.001
产出环境*顺序	1	6.708	0.011

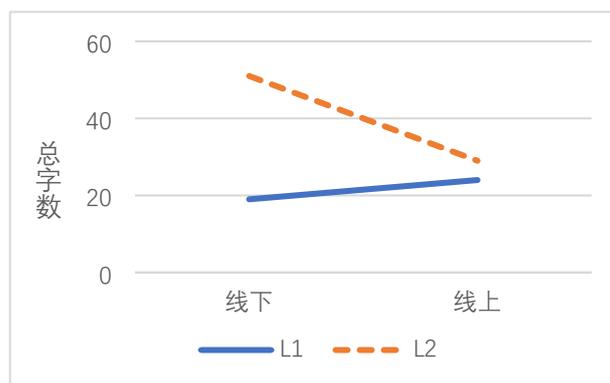
错误次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.772	0.044
截距	1	14.172	<0.001
语言背景	1	9.914	0.002
顺序*性别	1	6.414	0.012

犹豫次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.434	0.139
截距	1	12.067	<0.001
顺序	1	8.078	0.005
顺序*性别*产出环境	1	4.887	0.029

信息传输效率	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.965	0.022
截距	1	843.851	<0.001
语言背景	1	7.461	0.007
顺序	1	6.979	0.009
顺序*语言背景	1	4.074	0.045
产出环境*顺序	1	4.931	0.028

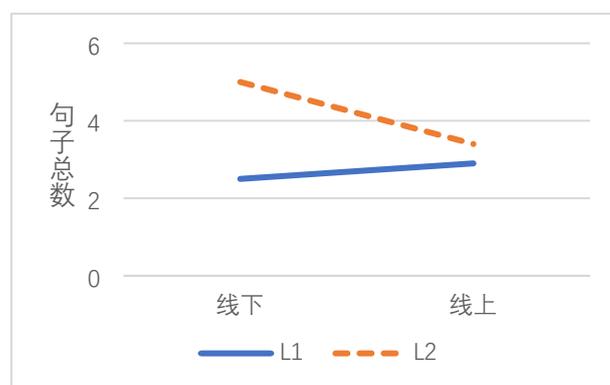
数据可以得出结论，总字数和口语产出环境之间的关系达到显著 $F(df=1) = 12.702, p < 0.001$ ；和语言背景之间的关系达到显著 $F(df=1) = 60.577, p < 0.001$ ；和语言背景和口语产出环境的交互作用之间的关系达到显著 $F(df=1) = 35.536, p < 0.001$ 。图 23 表示母语组线上总字数多于线下，二语组相反。

图 23 问题 (1) 产出环境和语言背景在总字数上的交互作用



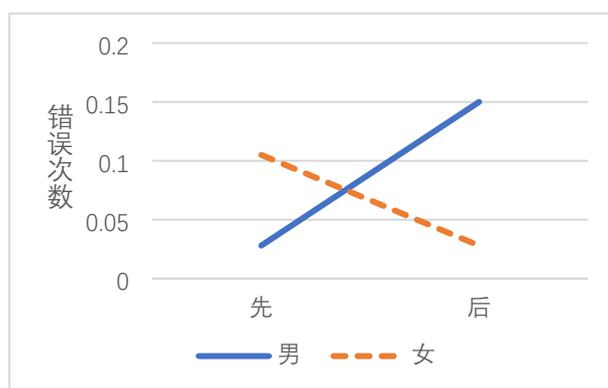
句子总数和口语产出环境之间的关系达到显著 $F(df=1) = 8.047, p < 0.05$; 和语言背景之间的关系达到显著 $F(df=1) = 51.171, p < 0.001$; 和语言背景和口语产出环境的交互作用之间的关系达到显著 $F(df=1) = 28.126, p < 0.001$; 和产出顺序和口语产出环境的交互作用之间的关系达到显著 $F(df=1) = 6.708, p < 0.05$ 。母语组和二语组都是第一次的句子总数多于第二次。图 24 说明母语组线上句子总数多于线下，二语组相反。

图 24 问题 (1) 产出环境和语言背景在句子总数上的交互作用



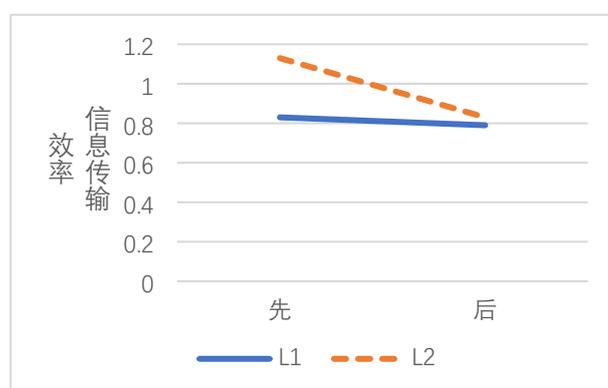
错误次数和语言背景之间的关系达到显著 $F(df=1) = 9.914, p < 0.05$; 和产出顺序和性别交互作用之间的关系达到显著 $F(df=1) = 6.414, p < 0.05$ 。图 25 说明男性组第二次的错误次数高于第一次，女性组相反。

图 25 问题 (1) 产出顺序和性别在错误次数上的交互作用



犹豫次数和产出顺序之间的关系达到显著 $F(df=1) = 8.078, p < 0.05$; 和产出顺序、口语产出环境和性别交互作用之间的关系达到显著 $F(df=1) = 4.887, p < 0.05$ 。信息传输效率和语言背景之间的关系达到显著 $F(df=1) = 7.461, p < 0.05$; 和产出顺序之间的关系达到显著 $F(df=1) = 6.979, p < 0.05$; 和语言背景和产出顺序的交互作用之间的关系达到显著 $F(df=1) = 4.074, p < 0.05$; 和产出顺序和口语产出环境的交互作用之间的关系达到显著 $F(df=1) = 4.931, p < 0.05$ 。母语组和二语组都是线上的信息传输效率大于线下，二语组之间差异较大。图 26 说明母语组和二语组都是第一次的信息传输效率大于第二次，同样二语组之间差异较大。

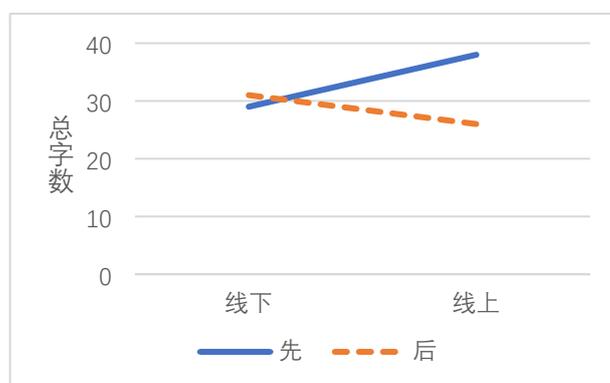
图 26 问题 (1) 产出顺序和语言背景在信息传输效率上的交互作用



(2) 问题 2

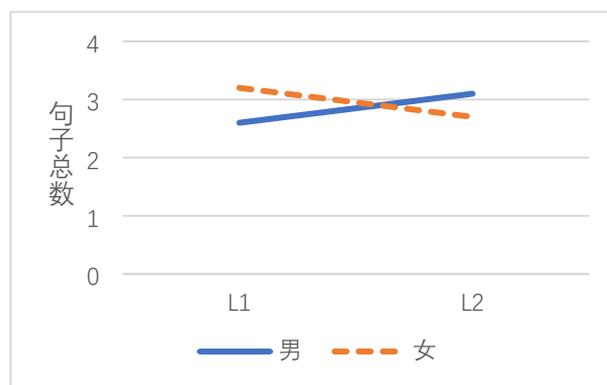
表 74 数据表明, 总字数和产出顺序之间的关系达到显著 $F(df=1)$
 $=4.050$, $p < 0.05$; 和语言背景和性别交互作用之间的关系达到显著 $F(df=1)$
 $=5.040$, $p < 0.05$; 和产出顺序和口语产出环境的交互作用之间的关系达到显著
 $F(df=1) = 6.452$, $p < 0.05$ 。母语组中, 女性比男性总字数多, 二语组相反。
 图 27 说明第一次口语产出任务中, 线上的总字数比线下多且差异较大, 第二次
 任务相反。

图 27 问题 (2) 产出环境和产出顺序在总字数上的交互作用



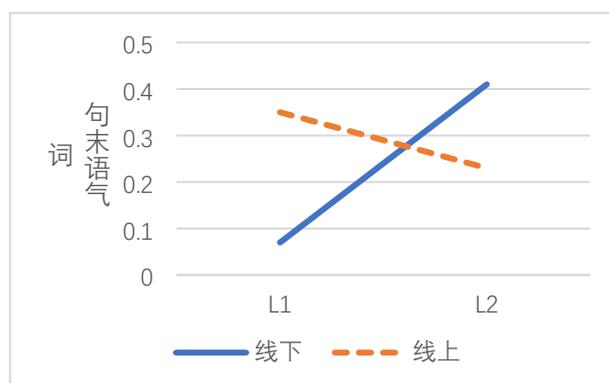
句子总数和语言背景和性别交互作用之间的关系达到显著 $F(df=1)$
 $=4.730$, $p < 0.05$; 和产出顺序和口语产出环境的交互作用之间的关系达到显著
 $F(df=1) = 11.920$, $p < 0.001$ 。图 28 说明母语组中, 女性比男性句子总数多,
 二语组相反。

图 28 问题 (2) 语言背景和性别在句子总数上的交互作用



句末语气词和语言背景和口语产出环境交互作用之间的关系达到显著 $F(df=1) = 5.569, p < 0.05$ 。图 29 表示母语组线上的句末语气词多于线下，二语组相反。

图 29 问题 (2) 语言背景和产出环境在句末语气词上的交互作用



信息词和口语产出环境与产出顺序交互作用之间的关系达到显著 $F(df=1) = 8.986, p < 0.05$ ；虚词和口语产出环境与产出顺序交互作用之间的关系达到显著 $F(df=1) = 4.080, p < 0.05$ ；名词和口语产出环境与产出顺序交互作用之间的关系达到显著 $F(df=1) = 8.047, p < 0.05$ 。

表 74 问题 (2) 部分语言参数的主体间效应检验

总字数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.518	0.106
截距	1	495.206	<0.001
顺序	1	4.050	0.046
性别*语言背景	1	5.040	0.026
顺序*产出环境	1	6.452	0.012

句子总数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.707	0.056
截距	1	740.585	<0.001
性别*语言背景	1	4.730	0.031
顺序*产出环境	1	11.920	<0.001

句末语气词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	0.823	0.650
截距	1	29.490	<0.001
产出环境*语言背景	1	5.569	0.020

信息词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.512	0.108
截距	1	758.935	<0.001
产出环境*顺序	1	8.986	0.003

虚词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	0.981	0.478
截距	1	329.441	<0.001
产出环境*顺序	1	4.080	0.045

名词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.403	0.153
截距	1	659.391	<0.001
产出环境*顺序	1	8.047	0.005

(3) 问题 3

根据表 75 的数据结果来看, 总字数和产出顺序之间的关系达到显著 F (df =1) =6.272, $p < 0.05$; 和语言背景、口语产出环境、产出顺序和性别的交互作

用之间的关系达到显著 $F(df=1) = 7.147, p < 0.05$ 。句子总数和产出顺序之间的关系达到显著 $F(df=1) = 6.104, p < 0.05$ ；和语言背景、口语产出环境、产出顺序和性别的交互作用之间的关系达到显著 $F(df=1) = 6.442, p < 0.05$ 。句末语气词和产出顺序之间的关系达到显著 $F(df=1) = 4.138, p < 0.05$ ；和语言背景、产出顺序和性别的交互作用之间的关系达到显著 $F(df=1) = 4.471, p < 0.05$ 。重复次数和语言背景之间的关系达到显著 $F(df=1) = 8.432, p < 0.05$ ；和口语产出环境、产出顺序和性别的交互作用之间的关系达到显著 $F(df=1) = 5.921, p < 0.05$ 。信息词和产出顺序之间的关系达到显著 $F(df=1) = 7.367, p < 0.05$ ；和语言背景、口语产出环境、产出顺序和性别的交互作用之间的关系达到显著 $F(df=1) = 10.290, p < 0.05$ 。虚词和产出顺序之间的关系达到显著 $F(df=1) = 5.598, p < 0.05$ ；和语言背景、口语产出环境、产出顺序和性别的交互作用之间的关系达到显著 $F(df=1) = 4.147, p < 0.05$ 。名词和语言背景、口语产出环境、产出顺序和性别的交互作用之间的关系达到显著 $F(df=1) = 10.870, p < 0.05$ 。

表 75 问题 (3) 部分语言参数的主体间效应检验

总字数	自由度	F 值	p 值
修正模型	15	1.365	0.172
截距	1	537.322	<0.001
顺序	1	6.272	0.013
产出环境*语言背景 *性别*顺序	1	7.147	0.008

句子总数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.587	0.084
截距	1	767.766	<0.001
顺序	1	6.104	0.015
产出环境*语言背景 *性别*顺序	1	6.442	0.012

句末语气词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.356	0.177
截距	1	95.852	<0.001
顺序	1	4.138	0.044
语言背景*性别*顺序	1	4.471	0.036

重复次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.488	0.117
截距	1	67.911	<0.001
语言背景	1	8.432	0.004
产出环境*性别*顺序	1	5.921	0.016

信息词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.661	0.065
截距	1	825.363	<0.001
顺序	1	7.367	0.007
语言背景*产出环境 *性别*顺序	1	10.290	0.002

虚词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.451	0.132
截距	1	348.814	<0.001
顺序	1	5.598	0.019
语言背景*产出环境 *性别*顺序	1	4.147	0.044

名词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.319	0.198
截距	1	619.860	<0.001
语言背景*产出环境 *性别*顺序	1	10.870	0.001

(4) 问题 4

表 76 表示, 信息输出效率和语言背景之间的关系达到显著 $F(df=1) = 29.928, p < 0.001$; 和产出顺序之间的关系达到显著 $F(df=1) = 6.152, p < 0.05$; 和语言背景、产出顺序和性别的交互作用之间的关系达到显著 $F(df=1) = 4.285, p < 0.05$ 。

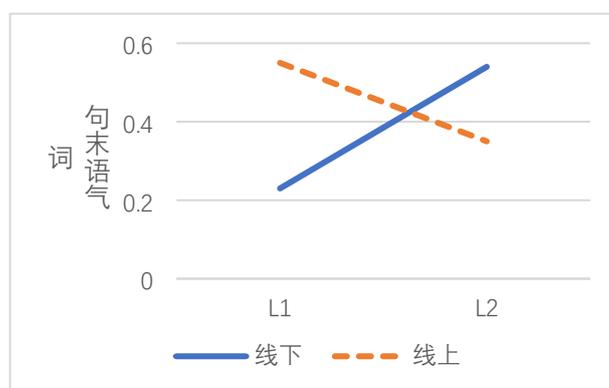
表 76 问题 (4) 信息输出效率的主体间效应检验

信息输出效率	自由度	F 值	p 值
修正模型	15	3.600	<0.001
截距	1	1852.526	<0.001
语言背景	1	29.928	<0.001
顺序	1	6.152	0.014
语言背景*性别*顺序	1	4.285	0.040

(5) 问题 5

总字数和语言背景、口语产出环境、产出顺序和性别的交互作用之间的关系达到显著 $F(df=1) = 4.285, p < 0.05$ 。句末语气词和语言背景和口语产出环境交互作用之间的关系达到显著 $F(df=1) = 4.288, p < 0.05$ 。图 30 说明在母语组中, 线上的句末语气词比线下多且差异更大, 二语组相反。

图 30 问题 (5) 语言背景和产出环境在句末语气词上的交互作用



重复次数和语言背景的关系达到显著 $F(df=1) = 7.191, p < 0.05$ ；和产出顺序和性别交互作用之间的关系达到显著 $F(df=1) = 4.602, p < 0.05$ 。图 31 表示男性组第一次的重复次数多于第二次且差距较大，女性组相反。

图 31 问题 (5) 产出顺序和性别在重复次数上的交互作用

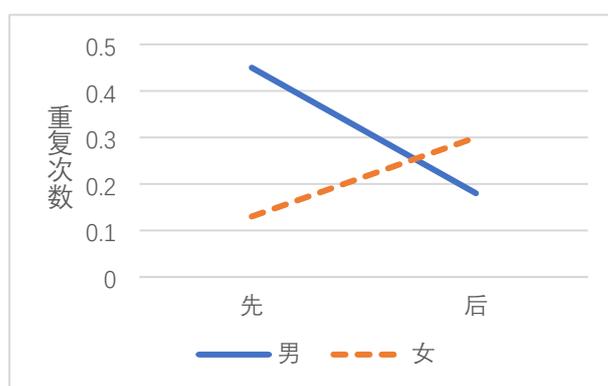


表 77 表示，虚词和口语产出环境、语言背景、产出顺序和性别交互作用之间的关系达到显著 $F(df=1) = 4.433, p < 0.05$ 。犹豫次数和口语产出环境的关系达到显著 $F(df=1) = 4.235, p < 0.05$ ；和口语产出环境和产出顺序交互作用之间的关系达到显著 $F(df=1) = 4.235, p < 0.05$ ；和口语产出环境、产出顺序和性别交互作用之间的关系达到显著 $F(df=1) = 6.295, p < 0.05$ 。信息输出效率和语言背景的关系达到显著 $F(df=1) = 32.270, p < 0.001$ ；和口语产出环境、产出顺序和性别交互作用之间的关系达到显著 $F(df=1) = 5.205, p < 0.05$ 。

表 77 问题 (5) 部分语言参数的主体间效应检验

总字数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	0.932	0.530
截距	1	371.233	<0.001
语言背景*产出环境 *性别*顺序	1	4.285	0.028

句末语气词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	0.885	0.582
截距	1	44.625	<0.001
语言背景*产出环境	1	4.288	0.040

重复次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.243	0.247
截距	1	25.958	<0.001
语言背景	1	7.191	0.008
顺序*性别	1	4.602	0.034

虚词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	0.896	0.570
截距	1	284.821	<0.001
语言背景*产出环境 *顺序*性别	1	4.433	0.037

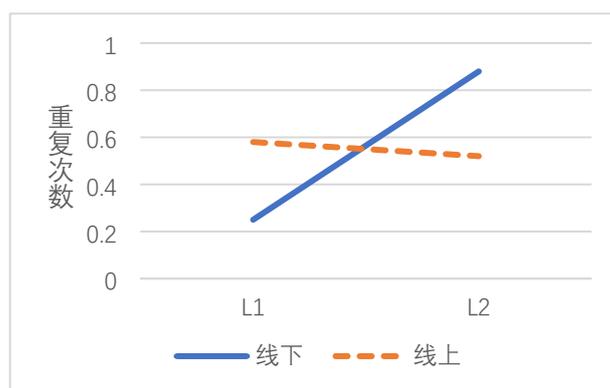
犹豫次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.798	0.040
截距	1	16.708	<0.001
产出环境	1	4.235	0.041
产出环境*顺序	1	4.235	0.041
产出环境*顺序*性别	1	6.295	0.013

信息输出效率	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	3.136	<0.001
截距	1	1073.085	<0.001
语言背景	1	32.270	<0.001
产出环境*顺序*性别	1	5.205	0.024

(6) 问题 6

表 78 说明重复次数和语言背景和口语产出环境的关系达到显著 $F(df=1) = 4.306, p < 0.05$; 和口语产出环境、语言背景和性别交互作用之间的关系达到显著 $F(df=1) = 4.099, p < 0.05$; 和口语产出环境、产出顺序和性别交互作用之间的关系达到显著 $F(df=1) = 5.420, p < 0.05$ 。图 32 说明母语组中，线上任务的重复次数比线下多，二语组相反。

图 32 问题 (6) 产出环境和语言背景在重复次数上的交互作用



虚词和口语产出环境和产出顺序交互作用之间的关系达到显著 $F(df=1) = 32.270, p < 0.05$ 。从图 33 可以看出第一次任务之间虚词的差距比第二次更大。

图 33 问题 (6) 产出环境和产出顺序在虚词上的交互作用

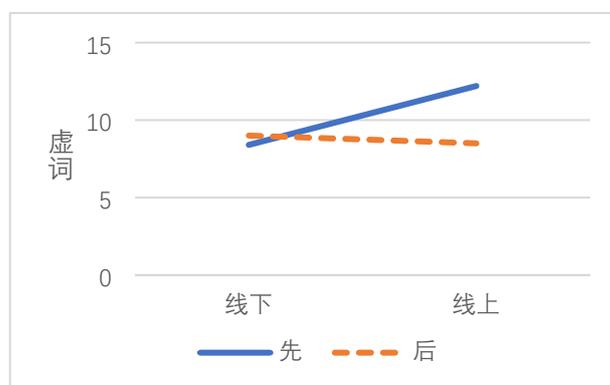


表 78 问题 (6) 部分语言参数的主体间效应检验

重复次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.435	0.139
截距	1	45.996	<0.001
语言背景*产出环境	1	4.306	0.040
语言背景*产出环境* 性别	1	4.099	0.045
顺序*产出环境*性别	1	5.420	0.021

虚词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.439	0.137
截距	1	323.150	<0.001
产出环境*顺序	1	4.151	0.043

(7) 问题 7

表 79 说明句末语气词和口语产出环境和产出顺序交互作用之间的关系达到显著 $F(df=1) = 4.922, p < 0.05$ 。信息词和口语产出环境、语言背景、性别和产出顺序交互作用之间的关系达到显著 $F(df=1) = 5.210, p < 0.05$ 。虚词和口语产出环境、语言背景、性别和产出顺序交互作用之间的关系达到显著 $F(df=1) = 5.109, p < 0.05$ 。错误次数和语言背景的关系达到显著 $F(df=1) = 11.416, p < 0.001$ ；和口语产出环境的关系达到显著 $F(df=1) = 4.996, p < 0.05$ ；和口语产出环境、语言背景和性别交互作用之间的关系达到显著 $F(df=1) = 5.969, p < 0.05$ 。名词次数和产出顺序的关系达到显著 $F(df=1) = 5.748, p < 0.05$ ；和口语产出环境、产出顺序、语言背景和性别交互作用之间的关系达到显著 $F(df=1) = 5.292, p < 0.05$ 。

表 79 问题 (7) 部分语言参数的主体间效应检验

句末语气词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	0.832	0.641
截距	1	42.502	<0.001
顺序*产出环境	1	4.922	0.028

信息词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.781	0.043
截距	1	593.597	<0.001
性别	1	8.338	0.004
语言背景*产出环境 *顺序*性别	1	5.210	0.024

虚词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.547	0.096
截距	1	172.823	<0.001
产出环境	1	4.197	0.042
语言背景*产出环境 *顺序*性别	1	5.109	0.025

错误次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	2.442	0.003
截距	1	32.693	<0.001
语言背景	1	11.416	<0.001
产出环境	1	4.996	0.027
语言背景*顺序*性别	1	5.969	0.016

名词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.291	0.215
截距	1	318.694	<0.001
顺序	1	5.748	0.018
语言背景*顺序*性 别*产出环境	1	5.292	0.023

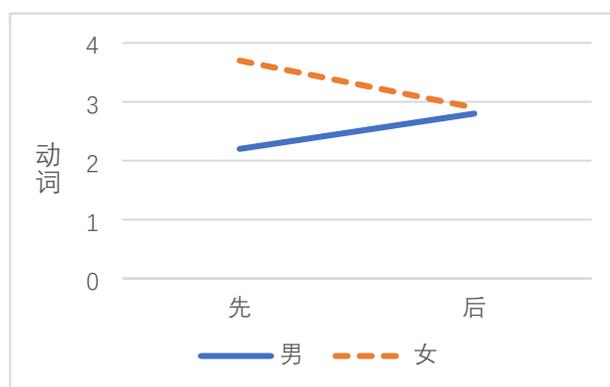
动词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.384	0.163
截距	1	401.221	<0.001
性别	1	6.867	0.010
顺序*性别	1	5.585	0.019

犹豫次数	自由度	F 值	p 值
修正模型	15	1.899	0.028
截距	1	18.935	<0.001
产出环境*顺序	1	4.734	0.031
产出环境*顺序*性别	1	7.397	0.007

信息传输效率	自由度	F 值	p 值
修正模型	15	3.033	<0.001
截距	1	776.519	<0.001
语言背景	1	27.030	<0.001
语言背景*顺序	1	4.436	0.037
语言背景*性别	1	7.101	0.045

动词次数和性别的关系达到显著 $F(df=1) = 6.867, p < 0.05$; 和产出顺序和性别交互作用之间的关系达到显著 $F(df=1) = 5.585, p < 0.05$ 。图 34 表示第一次任务中, 女生组比男生组使用了更多的动词, 第二次任务虽然也是女生组动词次数更多但是差异不大; 男生组中, 第二次比第一次动词次数更多, 女生组相反。

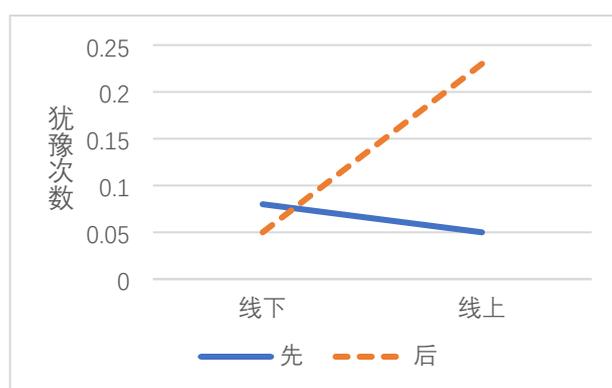
图 34 问题 (7) 产出顺序和性别在动词上的交互作用



犹豫次数和口语产出环境和产出顺序的关系达到显著 $F(df=1) = 4.734, p < 0.05$; 和口语产出环境、产出顺序和性别交互作用之间的关系达到显著 $F(df$

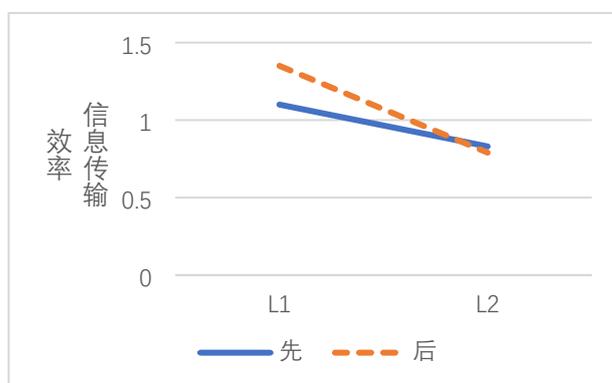
$=1) = 7.397, p < 0.05$ 。图 35 说明第一次任务中，线上线下犹豫次数差异不大，但是第二次任务中差异较大；线下任务中犹豫次数差异不大，线上任务中差异较大。

图 35 问题 (7) 产出环境和产出顺序在犹豫次数上的交互作用



信息传输效率和语言背景的关系达到显著 $F(df=1) = 27.030, p < 0.001$ ；和语言背景和产出顺序交互作用之间的关系达到显著 $F(df=1) = 4.436, p < 0.05$ ；和语言背景和性别交互作用之间的关系达到显著 $F(df=1) = 7.101, p < 0.05$ 。母语组中男生比女生信息传输效率高，二语组中相反。图 36 说明母语组第二次任务比第一次信息传输效率高，二语组相反。

图 36 问题 (7) 语言背景和产出顺序在信息传输效率上的交互作用



(8) 问题 8

表 80 说明总字数和口语产出环境、语言背景、性别和产出顺序交互作用之间的关系达到显著 $F(df=1) = 4.474, p < 0.05$ 。虚词和口语产出环境、语言背景、性别和产出顺序交互作用之间的关系达到显著 $F(df=1) = 4.216, p < 0.05$ 。

表 80 问题 (8) 部分语言参数的主体间效应检验

总字数	自由度	F 值	p 值
修正模型	15	0.946	0.515
截距	1	225.115	<0.001
语言背景*顺序*产出环境*性别	1	4.474	0.036

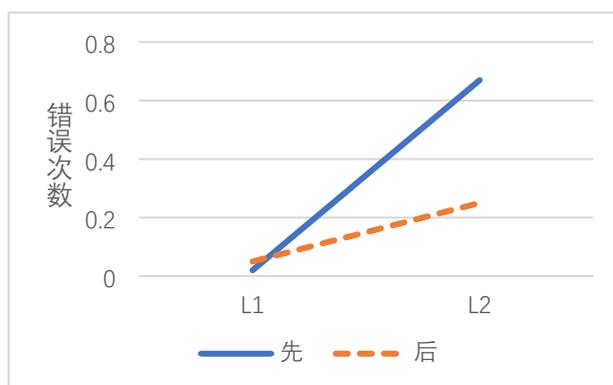
虚词	自由度	F 值	p 值
修正模型	15	0.680	0.801
截距	1	189.478	<0.001
语言背景*顺序*产出环境*性别	1	4.216	0.042

(9) 问题 9

表 81 表示错误次数和语言背景的关系达到显著 $F(df=1) = 16.035, p < 0.001$ ；和语言背景和产出顺序交互作用之间的关系达到显著 $F(df=1) = 4.442, p < 0.05$ ；和口语产出环境和性别交互作用之间的关系达到显著 $F(df=1) = 4.442, p < 0.05$ ；和口语产出环境、语言背景和性别交互作用之间的关系达到显著 $F(df=1) = 5.497, p < 0.05$ ；和口语产出环境、产出顺序和性别交互作用之间的关系达到显著 $F(df=1) = 4.442, p < 0.05$ 。男性组整体比女性组错误次数多，男性组线上的错误次数多于线下，女性组相反。图 37 说明在母语组

中，第一次和第二次的错误次数相差不大，甚至第二次的错误次数略多；但是在二语组中第一次的错误次数远高于第二次。

图 37 问题 (9) 语言背景和产出顺序在错误次数上的交互作用



名词和产出顺序的关系达到显著 $F(df=1) = 12.151, p < 0.001$ ；和口语产出环境和性别交互作用之间的关系达到显著 $F(df=1) = 4.724, p < 0.05$ 。图 38 表示男性组线上的名词数多于线下，女性组相反。

图 38 问题 (9) 产出环境和性别在名词上的交互作用

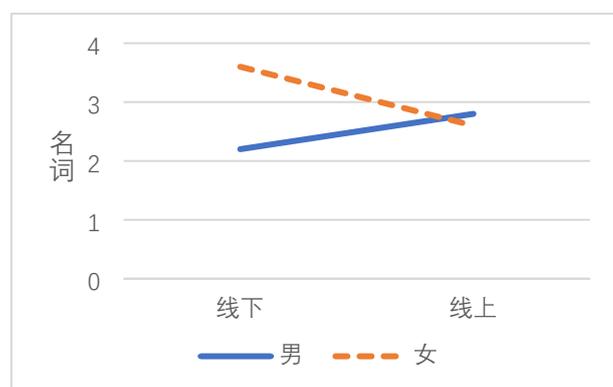


表 81 问题 (9) 部分语言参数的主体间效应检验

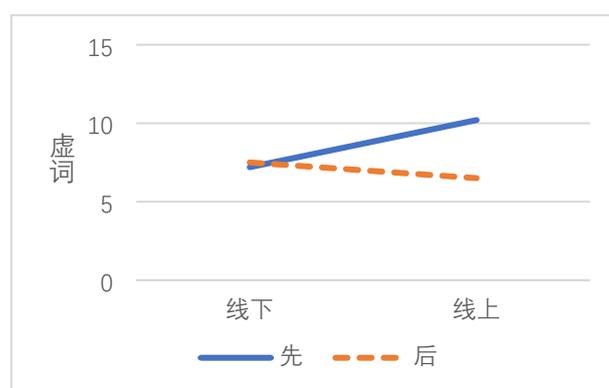
错误次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	3.061	<0.001
截距	1	22.237	<0.001
语言背景	1	16.035	<0.001
语言背景*顺序	1	4.442	0.037
产出环境*性别	1	4.442	0.037
语言背景*产出环境*性别	1	5.497	0.020
产出环境*顺序*性别	1	4.442	0.037

名词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.824	0.037
截距	1	199.094	<0.001
顺序	1	12.151	<0.001
产出环境*性别	1	4.724	0.031

(10) 问题 10

表 82 表示总字数和语言背景、口语产出环境和性别交互作用之间的关系达到显著 $F(df=1) = 4.378, p < 0.05$ 。句子总数和语言背景、口语产出环境和性别交互作用之间的关系达到显著 $F(df=1) = 4.624, p < 0.05$ 。虚词和口语产出环境和产出顺序交互作用之间的关系达到显著 $F(df=1) = 4.765, p < 0.05$ 。从图 39 可以看出线上任务时，两次之间虚词数差距较大，线下任务时差距较小。

图 39 问题 (10) 产出环境和产出顺序在虚词上的交互作用



错误次数和口语产出环境、性别、语言背景和产出顺序交互作用之间的关系达到显著 $F(df=1) = 5.266, p < 0.05$ 。名词数和产出顺序的关系达到显著 $F(df=1) = 4.314, p < 0.05$ ；和口语产出环境、性别和语言背景交互作用之间的关系达到显著 $F(df=1) = 4.314, p < 0.05$ 。信息传输效率和语言背景的关系达到显著 $F(df=1) = 18.052, p < 0.001$ ；和性别和语言背景交互作用之间的关系达到显著 $F(df=1) = 5.747, p < 0.05$ 。图 40 说明母语组中，男性比女性信息传输效率高，二语组相反。

图 40 问题 (10) 语言背景和性别在信息传输效率上的交互作用

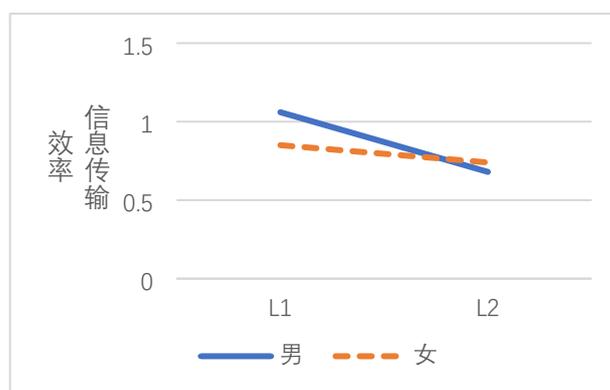


表 82 问题 (10) 部分语言参数的主体间效应检验

总字数	自由度	F 值	p 值
修正模型	15	1.262	0.234
截距	1	271.969	<0.001
语言背景*产出环境*	1	4.378	0.038
性别			

句子总数	自由度	F 值	p 值
修正模型	15	0.822	0.652
截距	1	135.869	<0.001
语言背景*产出环境*	1	4.624	0.033
性别			

虚词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.475	0.122
截距	1	272.815	<0.001
产出环境*顺序	1	4.765	0.031

错误次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.181	0.293
截距	1	26.255	<0.001
语言背景*产出环境* 顺序*性别	1	5.266	0.023

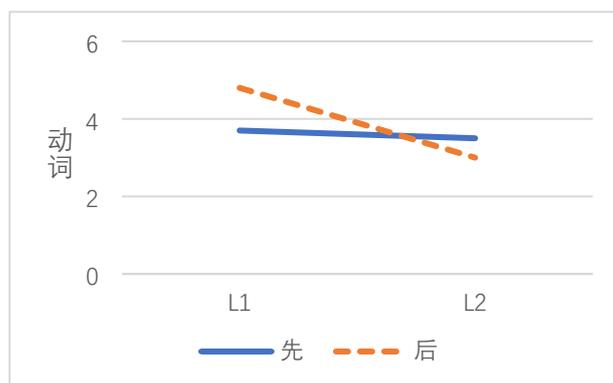
名词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.196	0.281
截距	1	377.622	<0.001
顺序	1	4.314	0.040
语言背景*产出环境* 性别	1	4.314	0.040

信息传输效率	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	2.280	0.006
截距	1	796.197	<0.001
语言背景	1	18.052	<0.001
语言背景*性别	1	5.747	0.018

(11) 问题 11

表 83 表示句子总数和语言背景的关系达到显著 $F(df=1) = 4.018, p < 0.05$; 句子总数和口语产出环境、语言背景、产出顺序和性别交互作用之间的关系达到显著 $F(df=1) = 4.713, p < 0.05$ 。动词数和语言背景的关系达到显著 $F(df=1) = 6.934, p < 0.05$; 和语言背景和产出顺序交互作用之间的关系达到显著 $F(df=1) = 4.523, p < 0.05$ 。图 41 表示母语组中, 第二次的动词数多于第一次, 二语组相反且动词数差距缩小。

图 41 问题 (11) 语言背景和产出顺序在动词上的交互作用



犹豫次数和语言背景的关系达到显著 $F(df=1) = 9.546, p < 0.05$; 和口语产出环境和产出顺序交互作用之间的关系达到显著 $F(df=1) = 5.049, p < 0.05$; 和口语产出环境、语言背景、性别和产出顺序交互作用之间的关系达到显著 $F(df=1) = 10.434, p < 0.05$ 。

表 83 问题 (11) 部分语言参数的主体间效应检验

句子总数	自由度	F 值	p 值
修正模型	15	1.301	0.209
截距	1	420.182	<0.001
语言背景	1	4.018	0.047
产出环境*语言背景* 顺序*性别	1	4.713	0.032

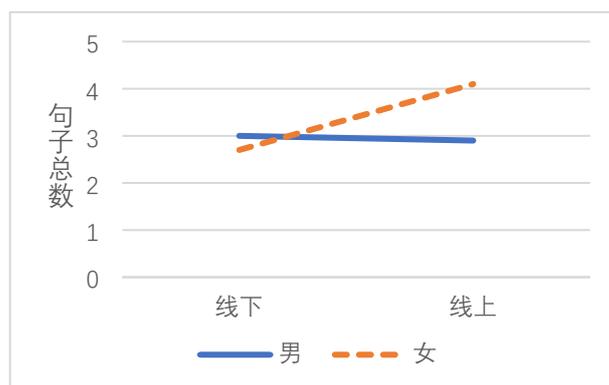
动词	自由度	F 值	p 值
修正模型	15	1.726	0.052
截距	1	401.071	<0.001
语言背景	1	6.934	0.009
语言背景*顺序	1	4.523	0.035

犹豫次数	自由度	F 值	p 值
修正模型	15	2.335	0.005
截距	1	41.736	<0.001
语言背景	1	9.546	0.002
产出环境*顺序	1	5.049	0.026
语言背景*产出环境*	1	10.434	0.002
顺序*性别			

(12) 问题 12

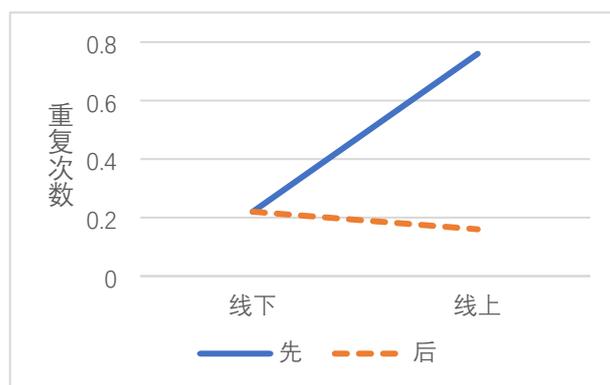
表 84 表示句子总数和口语产出环境和性别交互作用之间的关系达到显著 $F(df=1) = 4.083, p < 0.05$ 。图 42 说明男性组线上线下的句子总数大致持平，线下较多；女性组线上的句子总数高于线下且差距较大。

图 42 问题 (12) 产出环境和性别在句子总数上的交互作用



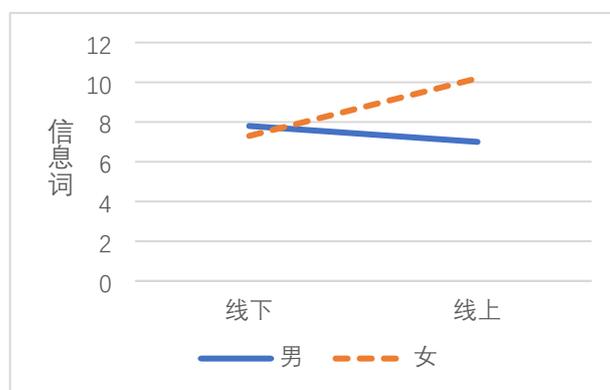
重复次数和产出顺序的关系达到显著 $F(df=1) = 5.798, p < 0.05$ ；和口语产出环境和顺序交互作用之间的关系达到显著 $F(df=1) = 5.692, p < 0.05$ ；和口语产出环境和性别交互作用之间的关系达到显著 $F(df=1) = 3.945, p < 0.05$ 。男性组线上线下的重复次数大致持平，线下较多；女性组线上的重复次数高于线下且差距较大。从图 43 可以看出，第一次任务时线上线下的重复次数差异较大，第二次任务重复次数大致持平。

图 43 问题 (12) 产出环境和产出顺序在重复次数上的交互作用



信息词和口语产出环境和性别交互作用之间的关系达到显著 $F(df=1) = 4.500, p < 0.05$; 和语言背景、口语产出环境、产出顺序和性别交互作用之间的关系达到显著 $F(df=1) = 4.764, p < 0.05$ 。图 44 表示, 男性组线上的信息词略少于线下, 女性组相反且差距较大。

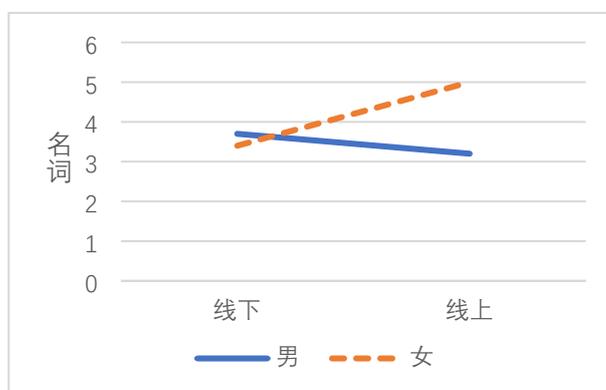
图 44 问题 (12) 产出环境和性别在信息词上的交互作用



错误次数和语言背景的关系达到显著 $F(df=1) = 428.416, p < 0.001$; 和语言背景、产出顺序和口语产出环境交互作用之间的关系达到显著 $F(df=1) = 8.182, p < 0.05$ 。名词数和口语产出环境和性别的关系达到显著 $F(df=1) = 4.674, p < 0.05$; 和语言背景、产出顺序、性别和口语产出环境交互作用之间

的关系达到显著 $F(df=1) = 4.000, p < 0.05$ 。图 45 表示男性组线下名词数多于线上，女性组相反且差距较大。

图 45 问题 (12) 产出环境和性别在名词上的交互作用



动词数和口语产出环境和性别的关系达到显著 $F(df=1) = 6.191, p < 0.05$ ；和语言背景、产出顺序、性别和口语产出环境交互作用之间的关系达到显著 $F(df=1) = 7.421, p < 0.05$ 。图 46 说明男性组线下动词数多于线上，女性组相反且差距较大。

图 46 问题 (12) 产出环境和性别在动词上的交互作用

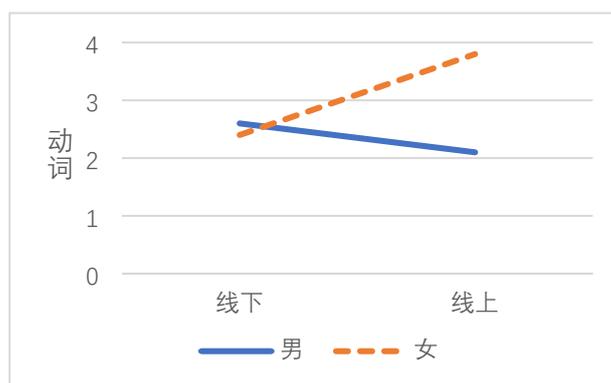


表 84 问题 (12) 部分语言参数的主体间效应检验

句子总数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.526	0.103
截距	1	282.054	<0.001
产出环境*性别	1	4.083	0.045

重复次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.572	0.089
截距	1	29.204	<0.001
顺序	1	5.798	0.017
产出环境*顺序	1	5.692	0.018
产出环境*性别	1	3.945	0.049

信息词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.667	0.064
截距	1	311.882	<0.001
产出环境*性别	1	4.500	0.036
语言背景*产出环境* 顺序*性别	1	4.764	0.031

错误次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	3.333	<0.001
截距	1	74.083	<0.001
语言背景	1	28.416	<0.001
语言背景*产出环境* 顺序	1	8.182	0.005

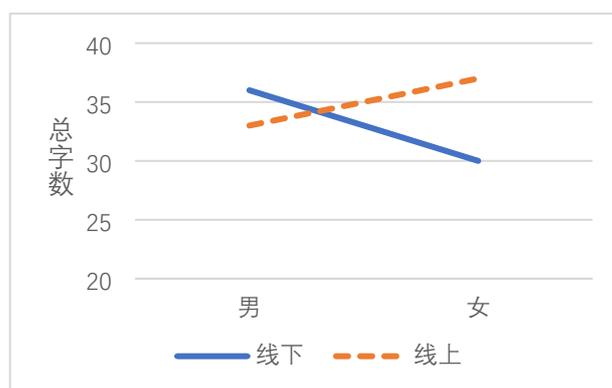
名词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.779	0.043
截距	1	254.181	<0.001
产出环境*性别	1	4.674	0.032
语言背景*产出环境* 顺序*性别	1	4.000	0.047

动词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	2.191	0.009
截距	1	199.203	<0.001
产出环境*性别	1	6.191	0.014
语言背景*产出环境* 顺序*性别	1	7.421	0.007

(13) 问题 happy

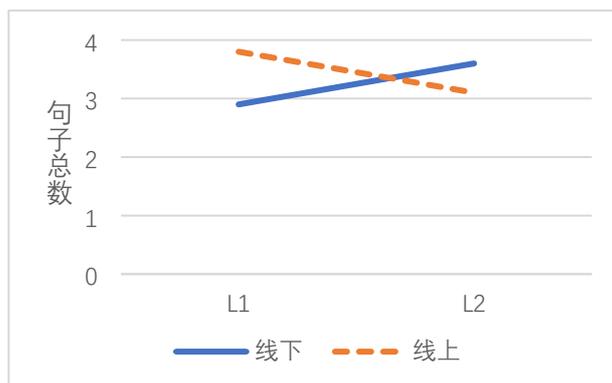
表 85 表示总字数和口语产出环境和性别交互作用之间的关系达到显著 $F(df=1) = 5.628, p < 0.05$ 。图 47 说明男性组线上的总字数少于线下，女性组相反。

图 47 问题 (happy) 产出环境和性别在总字数上的交互作用



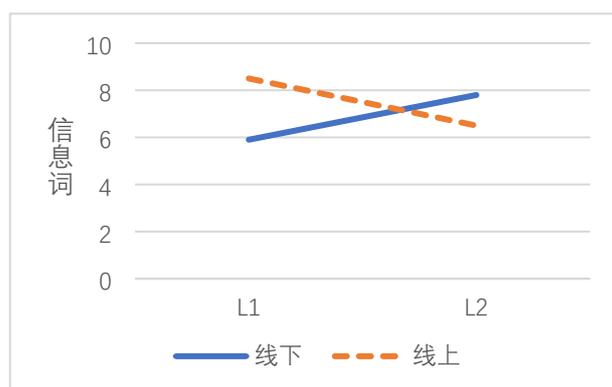
句子总数和产出顺序的关系达到显著 $F(df=1) = 4.210, p < 0.05$ ；和口语产出环境和语言背景交互作用之间的关系达到显著 $F(df=1) = 4.502, p < 0.05$ 。男性组线上的句子总数多于线下，女性组相反。图 48 表示母语组线上任务时的句子总数大于线下，但二语组结论相反。

图 48 问题 (happy) 产出环境和语言背景在句子总数上的交互作用



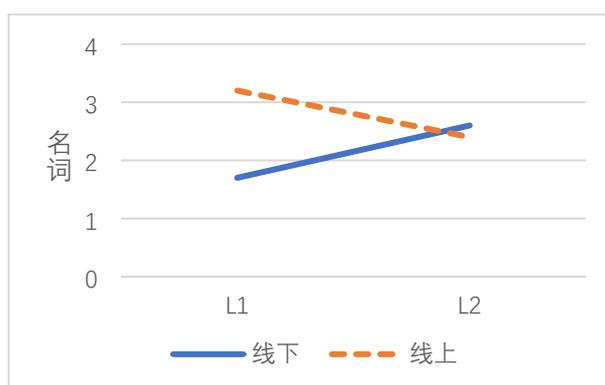
信息词和口语产出环境和语言背景交互作用之间的关系达到显著 $F(df=1) = 6.206, p < 0.05$ 。男性组线上的信息词数多于线下，女性组相反。图 49 表示母语组线上任务时的信息词数大于线下，二语组结论相反。

图 49 问题 (happy) 产出环境和语言背景在信息词上的交互作用



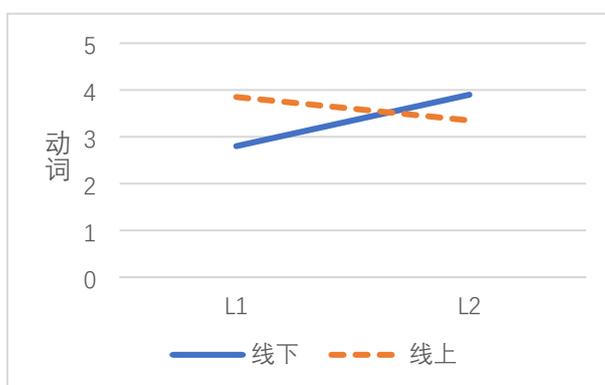
名词和口语产出环境和语言背景交互作用之间的关系达到显著 $F(df=1) = 6.206, p < 0.05$ 。男性组线上的名词数多于线下，女性组相反。图 50 表示母语组线上线上任务时的名词数相差较大（线上多），二语组两次任务中的名词数相差不大。这是因为二语组的中文名词词汇量有限，所以较少受到不同口语产出环境的影响。

图 50 问题 (happy) 产出环境和语言背景在名词上的交互作用



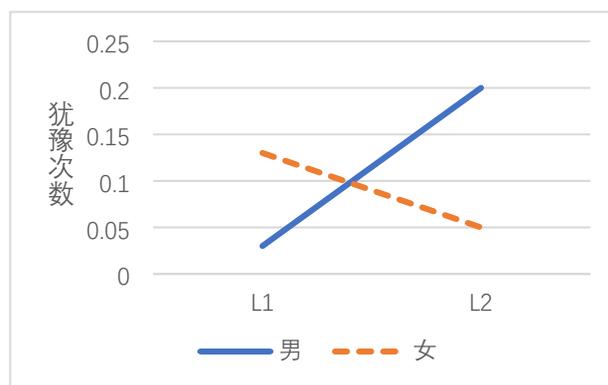
动词和产出顺序的关系达到显著 $F(df=1) = 13.646, p < 0.001$; 和口语产出环境和语言背景交互作用之间的关系达到显著 $F(df=1) = 4.731, p < 0.05$ 。男性组线上的动词数多于线下, 女性组相反。图 51 表示母语组线上任务时的动词数大于线下, 二语组相反。

图 51 问题 (happy) 产出环境和语言背景在动词上的交互作用



犹豫次数和性别和语言背景交互作用之间的关系达到显著 $F(df=1) = 4.721, p < 0.05$ 。图 52 表示男性组线上的犹豫次数多于线下且差距较大, 女性组相反。

图 52 问题 (happy) 性别和语言背景在犹豫次数上的交互作用



信息传输效率和语言背景的关系达到显著 $F(df=1) = 31.539, p < 0.001$;
和产出顺序的关系达到显著 $F(df=1) = 4.510, p < 0.05$; 和口语产出环境、性别和产出顺序交互作用之间的关系达到显著 $F(df=1) = 5.095, p < 0.05$ 。

表 85 问题 (happy) 部分语言参数的主体间效应检验

总字数	自由度	F 值	p 值
修正模型	15	1.234	0.254
截距	1	336.510	<0.001
产出环境*性别	1	5.628	0.019

句子总数	自由度	F 值	p 值
修正模型	15	1.284	0.219
截距	1	357.579	<0.001
顺序	1	4.210	0.042
语言背景*产出环境	1	4.502	0.036

信息词	自由度	F 值	p 值
修正模型	15	1.346	0.183
截距	1	441.038	<0.001
语言背景*产出环境	1	6.206	0.014

名词	自由度	F 值	p 值
修正模型	15	1.721	0.053
截距	1	309.324	<0.001
产出环境	1	5.007	0.027
语言背景*产出环境	1	8.655	0.004

动词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.908	0.027
截距	1	372.437	<0.001
顺序	1	13.646	<0.001
语言背景*产出环境	1	4.731	0.031

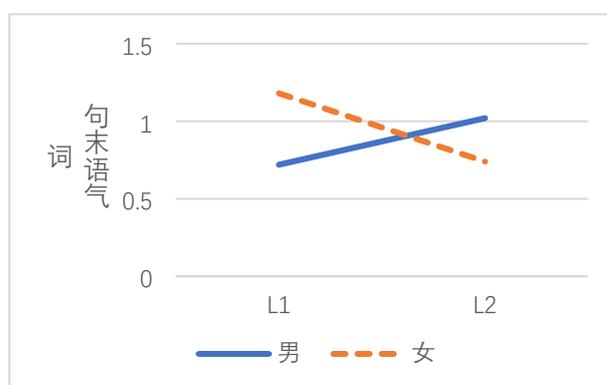
犹豫次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	0.806	0.669
截距	1	12.086	<0.001
语言背景*性别	1	4.721	0.031

信息传输效率	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	3.426	<0.001
截距	1	867.774	<0.001
语言背景	1	31.539	<0.001
顺序	1	4.510	0.035
产出环境*顺序*性别	1	5.095	0.026

(14) 问题 sad

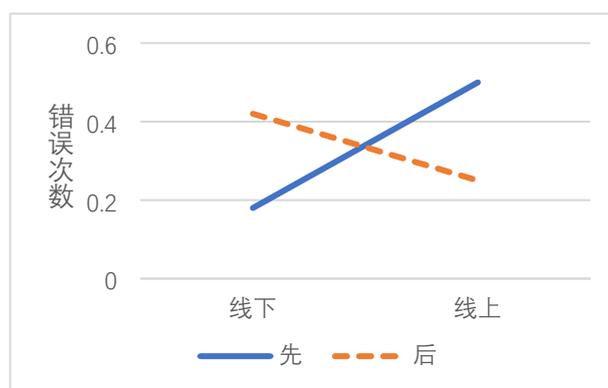
表 86 表示句末语气词和口语产出环境的关系达到显著 $F(df=1) = 4.979$, $p < 0.05$; 和语言背景和性别交互作用之间的关系达到显著 $F(df=1) = 4.979$, $p < 0.05$ 。图 53 表示母语组中女性比男性使用更多的句末语气词, 二语组相反。

图 53 问题 (sad) 性别和语言背景在句末语气词上的交互作用



错误次数和语言背景的关系达到显著 $F(df=1) = 31.522, p < 0.001$ ；和口语产出环境和产出顺序交互作用之间的关系达到显著 $F(df=1) = 6.408, p < 0.05$ ；和语言背景、产出顺序和口语产出环境交互作用之间的关系达到显著 $F(df=1) = 7.762, p < 0.05$ 。图 54 表示第一次任务时线上线下的错误次数差距较大于第二次。

图 54 问题 (sad) 产出环境和产出顺序在错误次数上的交互作用



信息传输效率和语言背景的关系达到显著 $F(df=1) = 43.809, p < 0.001$ ；和口语产出环境和产出顺序交互作用之间的关系达到显著 $F(df=1) = 16.345, p < 0.001$ ；和性别和口语产出环境交互作用之间的关系达到显著 $F(df=1) = 4.182, p < 0.05$ 。男性组中线上线下的信息传输效率差距不大，线上比线下略高；女性组线上的信息传输效率低于线下。图 55 表示第一次任务时线上线下的信息传输效率差距较大于第二次。

图 55 问题 (sad) 产出环境和产出顺序在信息传输效率上的交互作用

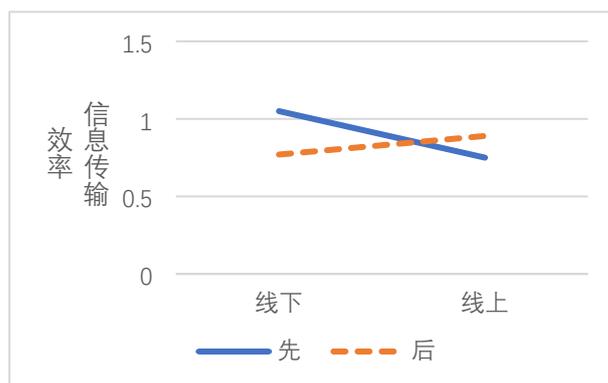


表 86 问题 (sad) 部分语言参数的主体间效应检验

句末语气词	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	1.152	0.316
截距	1	124.465	<0.001
产出环境	1	4.979	0.027
语言背景*性别	1	4.979	0.027

错误次数	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	3.284	<0.001
截距	1	47.434	<0.001
语言背景	1	31.522	<0.001
产出环境*顺序	1	6.408	0.012
语言背景*产出环境*顺序	1	7.762	0.006

信息传输效率	自由度	<i>F</i> 值	<i>p</i> 值
修正模型	15	5.437	<0.001
截距	1	953.239	<0.001
语言背景	1	43.809	<0.001
产出环境*顺序	1	16.345	<0.001
产出环境*性别	1	4.182	0.043

5.4.4 简单主效应分析

本论文发现了大量的交互效应，因此进行简单主效应（simple main effect）分析以进一步厘清变量之间的关系。

(1) 单一图片任务

表 87 单一图片任务的简单主效应分析

单一图片任务		自由度	<i>F</i>	显著性
语言背景	重复次数	1	11.105	0.001
	错误	1	30.267	<0.001
	名词	1	5.797	0.017
产出顺序	犹豫次数	1	22.137	<0.001
	句子总数	1	6.366	0.013
	犹豫次数	1	8.036	0.005

表 87 展示了单一图片任务数据的简单主效应分析结果。语言背景对各因变量的影响不尽相同：语言背景对重复次数 $F=11.105$, $p<0.01$ 、错误次数 $F=30.267$, $p<0.001$ 、名词 $F=5.797$, $p<0.05$ 和犹豫次数 $F=22.137$, $p<0.001$ 有显著影响。产出顺序对句子总数 $F=6.366$, $p<0.05$ 和犹豫次数 $F=8.036$, $p<0.01$ 有显著影响。产出环境和性别对单一图片任务采用的测量指标都没有显著影响。

(2) 长故事任务

表 88 长故事任务的简单主效应分析

长故事任务		自由度	<i>F</i>	显著性
语言背景	总字数	1	8.472	0.004
	信息词	1	25.654	0.000
	虚词	1	9.382	0.003
	错误	1	123.288	0.000
	名词	1	18.495	0.000
	动词	1	23.653	0.000
	犹豫次数	1	18.978	0.000
产出环境	句子长度	1	39.454	0.000
	CE	1	163.775	0.000
	犹豫次数	1	8.921	0.003
	CE	1	6.039	0.015

	SFPs	1	4.064	0.046
产出顺序	犹豫次数	1	5.989	0.016
	CE	1	9.882	0.002
性别	重复次数	1	6.4	0.012
	错误	1	9.988	0.002

表 88 展示了长故事任务数据的简单主效应分析结果。语言背景对总字数 $F=8.472$, $p < 0.01$ 、信息词 $F=25.654$, $p < 0.001$ 、虚词 $F=9.382$, $p < 0.01$ 、错误次数 $F=123.288$, $p < 0.001$ 、名词 $F=18.495$, $p < 0.001$ 、动词 $F=23.653$, $p < 0.001$ 、犹豫次数 $F=18.978$, $p < 0.001$ 、句子长度 $F=39.454$, $p < 0.001$ 和 CE $F=163.775$, $p < 0.001$ 有显著影响。产出环境对犹豫次数 $F=8.921$, $p < 0.01$ 和 CE $F=6.039$, $p < 0.05$ 有显著影响。产出顺序对句末语气词 $F=4.064$, $p < 0.05$ 、犹豫次数 $F=5.989$, $p < 0.05$ 和 CE $F=9.882$, $p < 0.01$ 有显著影响。性别对重复次数 $F=6.4$, $p < 0.05$ 和错误 $F=9.988$, $p < 0.01$ 有显著影响。

(3) 问答任务整合

表 89 问答任务的简单主效应分析

问答任务		自由度	F	显著性
语言背景	重复次数	1	12.096	0.001
	信息词	1	4.595	0.034
	犹豫次数	1	9.896	0.002
	错误	1	70.034	0.000
	名词	1	8.937	0.003
	CE	1	65.684	0.000
	总字数	1	4.657	0.032
性别	动词	1	6.702	0.011
	信息词	1	5.040	0.026
产出环境	CE	1	6.692	0.011

表 89 展示了整合后问答任务数据的简单主效应分析结果。语言背景对重复次数 $F=12.096$, $p < 0.01$ 、信息词 $F=4.595$, $p < 0.05$ 、犹豫次数 $F=9.896$, $p <$

0.01、错误次数 $F=70.034$, $p<0.001$ 、名词 $F=8.937$, $p<0.01$ 和 CE $F=65.684$, $p<0.001$ 有显著影响。产出顺序对本研究采用的测量指标都没有显著影响。性别对动词 $F=6.702$, $p<0.05$ 和信息词 $F=5.04$, $p<0.05$ 有显著影响。产出环境对 CE $F=6.692$, $p<0.05$ 有显著影响。综上所述, 在单一图片任务中语言背景对重复次数、错误次数和名词数有显著影响; 产出顺序对句子总数和犹豫次数有显著影响。在长故事任务中语言背景对总字数、信息词数、虚词数、错误次数、名词数、动词数、犹豫次数、句子长度和信息传输效率有显著影响; 产出环境对犹豫次数和信息传输效率有显著影响; 产出顺序对句末语气词数、犹豫次数和信息传输效率有显著影响; 性别对重复次数和错误次数有显著影响。整合后问答任务中语言背景对重复次数、信息词数、犹豫次数、错误次数、名词数和信息传输效率有显著影响; 性别对动词数和信息词数有显著影响; 产出环境对信息传输效率有显著影响。

第六章 总论

6.1 数据结果分析

本研究的数据结果回答了提出的研究问题并印证了研究假设: 从语言学角度分析, 口语产出环境、产出顺序、语言背景和性别因素影响参与者的口语表现, 有交互作用且在某些语言参数上呈现显著差异。同时, 不同形式的问题(图片描述任务的长短、正式和非正式问题和表现不同情绪的问题)也对于中文二语学习者和母语者的口语表现有影响。在本文文献综述部分提及的锜宝香(2001)关于一般阅读水平学生和低阅读水平学生的研究, 结论之一是一般阅

读水平学生相异词汇数比低阅读水平学生多，但是平均句子长度两组相差不大。该结论与本研究的发现有相同之处：在描述短图片任务时母语组使用较少名词，整体来看二语组和母语组时的平均句子长度相差不大。这说明在考察学生的语言水平时，平均句子长度并不是一个很重要的评判指标。在 Kong (2009) 关于失语症人群中文语言表达的研究中，区分他们语言表达流利度最有用的两个参数是信息传输效率 (Index of Communication Efficiency: ICE) 和语法支持参数 (Index of Grammatical Support: IGS)，同时错误次数 (Index of Error: IEr) 可以用来区分失语症组和普通人群组。在本研究的多个语言测量参数中 (总字数、句子总数、时长、句末语气词、重复次数、信息词数、虚词数、错误次数、名词数、动词数、犹豫次数、句子长度、和信息输出效率)，数据呈现显著差异次数较多的，较为重要的语言参数有信息词数、信息输出效率和错误次数。信息词数与被访者的词汇量密切相关，信息词数与信息输出效率在线上 and 线下交流中的对比中差距最大。错误次数在母语组和二语组之间差距最大。本实验回答了第一章研究目的部分提出的第一个研究问题：“从语言学角度分析，中文二语学习者和母语者，线上与线下的口语表现差异体现在哪些方面？”，实验得出结论是差异主要体现在总字数、句子总数、重复次数、错误次数、犹豫次数和信息输出效率上。在不同任务中，造成被试口语表现差异的具体分析和原因如下。

6.1.1 图片描述任务数据结果分析

从家庭图片描述任务的语言数据来看，线上的重复次数要比线下的重复次

数略多。对于一些人来说，线上沟通可能更加令人不安，这种紧张可能导致言语表达上的重复和不流畅。从青蛙的故事描述任务的语言数据来看，线上的犹豫次数比线下的犹豫次数多，并且在故事描述任务中线下的信息输出效率比线上的信息输出效率要高。面对面的交流往往伴随着更高的社交期望和规范。在面对面交流中，被试可能会觉得需要更快速地响应，以避免在沟通中出现长时间的尴尬和沉默。与线上相比，线下往往具有更自然的交流节奏，双方可以根据对方的反应和语言提示来调整自己的说话速度和内容，从而实现更流畅、更高效的信息输出。在线下交流中，实验者与被试之间的面对面互动更有利于捕捉和解读非言语的信息，如肢体语言、面部表情和眼神交流。这些非言语线索能够为被试提供即时的反馈，使其更容易、更迅速地理解问题，从而减少犹豫次数。综合性的交流渠道使得信息传达更为丰富和准确，从而提高信息输出效率。采访者可以利用身体语言，如点头、微笑或摇头，来给予反馈，而这些在线上中可能不太明显或不易察觉。因此，线下的被试可能会更少地使用言语来填补这些沟通空白。与真人面对面交流，相对于屏幕上的虚拟形象，会让人有更强烈的连接和沟通感，有助于建立和维护人际关系中的信任和连接。真实的人际互动促使被试更加直接、简洁地表达自己的想法，减少犹豫。这种情感上的连接会促使被试更加开放，更愿意分享深层次的思考和感受，从而增加信息的输出效率。其次，线上交流尤其是视频，常常受到技术问题的困扰。

Budianto (2023) 的研究表示在印度尼西亚某些地区的网络设施与环境较为落后，尚不能很好地满足教学的需要，也不易构建理想的线上汉语教学产出环境。网络延迟、声音中断或图像模糊会导致被试需要额外的时间来处理和理解

问题，从而增加了停顿和犹豫的次数，受访者在尝试维持流畅的对话时也更容易重复自己的话。除了网络因素，环境也是另外一个的影响因素。本研究的线下任务通常在一个预先设定并为此目的设计的环境中进行，例如一个安静的小亭子或房间内。这种环境让被试感到更为放松和舒适，有助于其集中精力。而线上任务可能在被试的家中或其他非正式的地点进行，这些地方可能存在各种外界干扰，导致被试更容易分心。

从产出顺序角度来分析家庭图片描述任务的语言数据，第一次的总字数、句子总数和犹豫次数都多于第二次；第一次的信息词数、虚词数和名词数也较多于第二次。该发现说明练习效应对家庭图片描述任务的结果影响不大，第二遍任务并没有体现出整体更高的语言描述水平。同样从产出顺序角度来分析青蛙故事描述任务的语言数据，第一次的句末语气词数和犹豫次数多于第二次的的数据，第二次的信息输出效率比第一次的高。和家庭图片描述任务不同，青蛙的故事描述任务在第二遍任务中体现出整体更高的语言描述水平。这说明练习效应对单一图片，或者说对于内容很少的语言输出任务影响不大，但是对于内容较多的语言输出任务，如一整个绘本故事影响很大，被试对于该语言任务的熟悉度和信息输出效率呈正相关。与单一图片相比，长故事包含更多的信息、角色和情节。对于长故事，被试需要处理和记忆的信息量更大。在第一次任务中，为了确保自己没有遗漏任何细节或情节，被试可能会更加小心谨慎，这导致更多的犹豫和使用句末语气词。同样的，第一次任务时被试需要更多时间来整理思路并概括故事，这导致更多的犹豫和使用句末语气词。到了第二次任务，他们对故事的内容、主题和细节更为熟悉，能更加有条理地进行回答，

更自信地表达自己的观点，从而提高了信息输出效率。被试在第二次任务时对故事的内容有了更清晰的记忆，他们也许对第一次任务中提到的内容进行了自主学习，使得第二次更为简练和有针对性。

从性别因素分析语言数据，性别因素对于家庭图片任务没有显著影响，但是在青蛙的故事描述任务中，男性组比女性组的重复次数更多，错误次数更多。性别与语言使用之间的关系历来是社会语言学和性别研究的热门话题。尽管男性和女性在生物结构上有所不同，但语言能力与性别的关系往往受到多种社会和文化因素的影响。研究显示，某些情境下男性可能更容易感受到压力或焦虑 (Shaywitz et al., 1995)。例如，在需要进行公开演讲或长时间叙述的任务中，男性可能比女性更容易感到不自在或焦虑，这导致他们更多地重复或出错。研究指出，男性更多地依赖大脑的左半球来处理线性和逻辑任务，而女性可能更多地使用双半球来处理综合和关联任务。因此，男性和女性在处理语言任务时经常采用不同的策略。例如，男性更倾向于线性思考，即按照逻辑顺序描述故事，而女性更善于整体性思考，能够更好地综合和组织信息。这种差异可能导致男性在长故事描述任务中更容易出错或重复。另一个可能影响结果的因素是提问者的性别因素，相对于和异性交流，与同性交流会更为舒适和自然。因为提问者是女性，所以男性被试可能会在与女性者交谈时感到更大的压力，尤其是当涉及到展现自己的知识或能力时。

从语言背景角度分析家庭图片描述任务，母语者比二语者的重复次数、错误数、犹豫次数更少，名词数也整体较少。很明显的原因是，因为母语者比二语者更熟练地掌握语言，因此在语言描述时，他们往往能更流利、准确地表达

自己的想法。这也解释了为什么母语者的重复次数、错误数和犹豫次数都较少。母语者通常拥有更丰富的词汇，这使得他们能够更精确地描述图片或故事的细节，而不必过度依赖一般性的名词。这可能是为什么在家庭图片描述任务中，母语者的名词数整体较少的原因。在青蛙的故事任务中，母语组的总字数、信息词总数、虚词数、名词数、动词数、句子长度和信息输出效率数据比二语组的数据大，母语组的错误数、犹豫次数比二语组的数据小。母语者能够使用更复杂的句子结构，包括多种从句和修饰语。这解释了为什么在青蛙的故事任务中，母语组的句子长度数据会比二语组的大。由于母语者不必花费太多的认知资源去处理语言本身，他们可以将更多的注意力集中在如何组织和传达信息上。母语者通常更了解如何在他们的语言中有效地表达意思，这意味着他们可能会采用更简洁易懂的口语化语言来描述事物。而二语者可能会在尝试寻找合适的表达时使用更多的虚词或冗余的语言。这可能是为什么母语者在青蛙的故事任务中的信息输出效率更高的原因。对于母语者来说，语言处理很大程度上是自动化的，这使得他们能够更快速、更自然地进行描述。相比之下，二语者可能需要更多的认知努力来规划、组织和形成他们的描述，这会导致更多的犹豫和错误。总而言之，母语者与二语者在语言任务中的表现差异可以从多个角度来解释，其中认知处理、语言熟练度、词汇知识和句子结构都是关键因素。这种差异不仅在图片描述任务中存在，在其他语言任务中也同样明显。

6.1.2 图片任务主体间效应检验数据分析

本实验对交互作用的分析得出的结论回答了第一章提出的第二个研究问

题：“产出顺序、语言背景和性别因素是否影响中文二语学习者和母语者的口语表现？以上三个因素是否与口语产出环境有交互作用？”结论是肯定的。在此我们通过比较两个看图说话任务的主体间效应检验数据，找出其共同的结论并进行分析。家庭活动图片描述任务因只有单张图片，提供的信息量较少，因此结合青蛙的故事共同分析看图说话任务中被试的口语表现。以下将单一家庭活动图片描述任务称为任务一，将青蛙的故事描述任务称为任务二，对比分析四个组别：母语组、二语组、女性组和男性组的数据。在看图说话任务的主体间效应检验中，只选择任务一和任务二相同的结论，得出的结论如下：母语组和女性组在线上任务时的总字数、句子总数、名词数和信息词数多于线下。在线下交流时如果有过多过长时间的停顿会让人感觉尴尬，但是线上交流的环境创设了一些疏离感。在不需要即时反馈的线上环境中，这两个群体可能更愿意详细地描述图片，或者有更多的时间来构思和表达自己的想法。母语组和二语组线下的句末语气词多于线上。线下的互动更为即时和直接，句末语气词的使用可以帮助说话者更好地表达自己的情感态度，或者是为了寻求听者的反馈。二语组和女性组线上的重复次数多于线下。这可能是由于线上通讯的延迟或者不稳定性导致的需要重复来确保信息传达，也可能是因为线上环境下更高的压力或不确定性导致的言语重复。母语组线上的虚词数、错误次数较多于线下，二语组线下的虚词数较多于线上。母语组、二语组和男性组都是线上的犹豫多于线下。二语组和男性组线下的句子长度较多于线下。在面对面的交流中，这两个群体更倾向于构建更复杂或完整的句子。母语组、二语组、男性组和女性组都是线下的信息输出效率高于线上。

任务一和任务二主体间效应检验呈现显著差异的相同语言测量参数如下：

任务一重复次数和语言背景之间的关系达到显著，和产出顺序、口语产出环境和性别的交互作用之间的关系达到显著。任务二重复次数和性别之间的关系达到显著；和产出顺序、口语产出环境交互作用之间的关系达到显著；和语言背景、产出顺序、口语产出环境和性别的交互作用之间的关系达到显著。任务一虚词和产出顺序、口语产出环境、语言背景和性别的交互作用关系达到显著。任务二虚词和语言背景之间的关系达到显著，和语言背景、性别、产出顺序和口语产出环境之间的关系达到显著。任务一错误次数和语言背景之间的关系达到显著。任务二错误次数和语言背景之间的关系达到显著，和性别之间的关系达到显著；任务一名词数量和语言背景之间的关系达到显著，和产出顺序、口语产出环境和性别的交互作用之间的关系达到显著，母语组的名词数量多于二语组。任务二名词数量和语言背景之间的关系达到显著；和语言背景、产出顺序、口语产出环境之间交互作用的关系达到显著。任务一犹豫次数和产出顺序之间的关系达到显著，和语言背景之间的关系达到显著。任务二犹豫次数和语言背景之间的关系达到显著；和产出顺序之间的关系达到显著；和口语产出环境之间的关系达到显著。如果只选择语言测量参数之间交互关系达到显著、并且任务一和任务二体现同样趋势的数据，有重复次数、虚词和名词：任务一中被访者的重复次数和产出顺序、口语产出环境和性别的交互作用之间的关系达到显著。这意味着在进行短图片描述时，被访者的重复次数受到他们进行的顺序、口语产出环境以及不同性别的共同影响。任务二重复次数和产出顺序、口语产出环境交互作用之间的关系达到显著；和语言背景、产出顺序、口语产出

环境和性别的交互作用之间的关系达到显著。在讲述长故事时，被访者因为语言背景、产出顺序、口语产出环境和性别不同，在言语中有不同程度的重复。在任务一和任务二中，虚词和产出顺序、口语产出环境、语言背景和性别的交互作用关系达到显著。这表示虚词的使用受到四个变量综合作用的影响，而这种影响在短图片描述和长故事描述两种不同类型的任务中都是一致的。任务一名词数量和产出顺序、口语产出环境和性别的交互作用之间的关系达到显著。在描述短图片时，使用的名词数量不仅受到产出顺序和产出环境的影响，还与性别有关。任务二名词数量和语言背景、产出顺序、口语产出环境之间交互作用的关系达到显著。讲述长故事时使用的名词数量也会受到被访者的语言背景、产出顺序以及口语产出环境的复合影响。从主体间效应检验结果可以看出，语言表达中的重复次数、名词使用以及虚词的使用都不是孤立发生的，它们受到产出顺序、产出环境、性别和语言背景等多个因素的交互影响。这些发现对语言学习、教学以及跨文化交流等领域提供洞察，它们揭示了在不同环境中语言使用的复杂性。了解这些交互作用如何影响语言输出，可以帮助我们设计更有效的沟通策略和教学方法。

6.1.3 问答数据结果分析

从线上和线下角度对比 14 个问题的数据分析结果，在问题（1）“你可以自我介绍一下吗？”中，线下的总字数、句子总数、重复次数、错误数要比线上的多。线下的信息词数、虚词数、动词数、犹豫次数和信息输出效率要比线上的少。问题（2）“你喜欢在大城市工作（生活）还是小城市？”和问题（9）“你最

难忘（开心）的一件事是什么？”的各语言参数都没有显著差异。在问题（3）“请介绍你的一位朋友（家人）比如说他的外表、性格和喜好？”和问题（6）“请介绍你的国家（城市）比如风俗文化。”中，线上的虚词数比线下多。在问题（4）“你平时（在学校）的生活是怎样的？请描述你的一天。”中，线上的总字数、句子总数、句末语气词数、虚词数和名词数都比线下的多，线上的信息输出效率比线下的低。问题（5）“周末（假期）一般会怎么度过？和谁一起度过？”线上的信息输出效率比线下的低，但同时线上的犹豫次数也比线下少。问题（7）“你（不）喜欢吃什么食物？请描述一下？”线上的总字数、虚词数和错误次数多于线下。问题（8）“你最喜欢看什么书（电影）？为什么？”线上的总字数、句子总数、信息词数、虚词数、名词数和动词数多于线下，线上的信息传输效率低于线下。问题（10）“你最喜欢的城市（国家）在哪里？为什么？”线上的信息传输效率低于线下。问题（11）“你认为空气（水）污染的来源是什么？你有什么建议？”线上的句子总数比线下多。问题（12）“你认为很好的家长（老师）应该是什么样子的？”线上的总字数、句子总数、虚词数多于线下，线上的信息传输效率低于线下。问题（happy）“假如你中了奖获得了一万人民币的奖金，你想告诉你的朋友，你会怎么说？/如果老师说你的论文很好，你想告诉你的朋友，你会怎么说？”线上的名词数多于线下。问题（sad）“假如你被人骗了一万人民币，你很不开心，你想告诉你的朋友，你会怎么说？/如果老师说你的论文很差，你想告诉你的朋友，你会怎么说？”线上的总字数、句子总数、句末语气词数和虚词数多于线下，线上的信息传输效率低于线下。

在测量线上线下不同产出环境的 11 个语言测量参数中，总字数和句子总数

呈现差异次数中多次都是线上比线下多，只有一次相反（问题 1）。句末语气词数、动词数、信息词数、虚词数和名词数都是线上比线下多。错误次数呈现差异次数一次线下多（问题 1），一次线上多（问题 7）。犹豫次数呈现差异次数一次线下多（问题 5），一次线上多（问题 1）。信息传输效率呈现差异次数其中多次都是线上的信息传输效率低于线下，只有一次相反（问题 1）。以上数据可以看出问题 1 和整体趋势相反的次数较多，这是因为第一个问题是自我介绍，二语学习者在初学中文时就会学习自我介绍内容，所以对该问题的回答已经准备的很充分。如果将 14 个问题数据结论进行总结，得出被访者在线上任务时会进行更多的语言输出的结论，如信息词数、句子总数更多。但是这并不说明线上的口语表现优于线下，因为线下的信息传输效率更高。口语表现涉及到清晰度、准确性、流利度和效率等方面。尽管线上语言输出的总量更多，但线下信息的传达更为直接和精炼，被访者在相同的时间内能够传递更多的有效信息。在线上任务时，因为缺乏即时反馈性以及面对面的社交压迫，被访者会有更多时间来构思回答。此外，线上减少了非语言交流的元素，如肢体语言和面部表情，可能导致被访者通过增加语言输出来补偿。

从产出顺序对比 14 个问题的数据分析结果，在问题（1）中第一次的重复次数和犹豫次数多于第二次，第一次的信息传输效率低于第二次。问题（2）中第一次的重复次数少于第二次，第一次的信息传输效率高于第二次。问题（3）中第一次的总字数、句子总数、句末语气词数、信息词数、虚词数、动词数和犹豫次数都高于第二次。问题（4）中第一次的信息传输效率低于第二次。问题（5）和问题（11）的各语言测量参数都没有体现显著差异。问题（6）第一次

的总字数、句子总数、信息词数和动词数都多于第二次。问题（7）第一次的名词数多于第二次。问题（8）第一次的句末语气词数、动词数和信息传输效率都高于第二次。问题（9）第一次的总字数、句子总数、信息词数、虚词数、名词数、动词数和犹豫次数都高于第二次，但是第一次的信息传输效率低于第二次。问题（10）第一次的总字数、信息词数、虚词数和名词数高于第二次。问题（12）第一次的总字数、重复次数、信息词数和名词数都高于第二次。问题（happy）第一次的句子总数、重复次数和动词数高于第二次，信息传输效率低于第二次。问题（sad）第一次的信息词数、虚词数和名词数高于第二次。在测量产出顺序的 11 个语言测量参数中，呈现差异次数由大到小排列分别为信息词数、信息传输效率>总字数、名词数和动词数>句子总数、重复次数和虚词数>犹豫次数>句末语气词数，且都是第一次多于第二次。重复次数呈现差异多是第一次重复次数多于第二次。信息传输效率呈现差异次数多是第一次访信息传输效率小于第二次。受访者在第一次时语言输出的总量更多，如总字数、名词数和动词数更多，但同时重复数和犹豫次数也较多，这就导致了信息传输效率较低。被试在第一次任务时可能表现出更高的积极性，尝试提供更多的信息，这也反映在总字数、名词和动词数量的增加上。原因可能是初次任务时被试更加努力尝试给出完整的回答，或者是因为在初次任务中他们还在适应的过程和内容，从而导致更多的语言尝试。增加的语言输出伴随着更多的重复和犹豫，被试在尝试表达自己的思想时缺乏信心或语言能力欠佳。犹豫包括使用填充词（如“嗯”、“啊”等）或是在讲话时三秒以上的停顿。尽管输出了更多的词汇和句子，但其中有相当部分并未携带新的信息，而是重复或修正了之前的信

息导致信息传输效率的降低。问题任务在产出顺序方面并没有很明显的相同趋势，但是如果问卷 A 和问卷 B 中的问题完全一样的话，因为练习效应的影响，趋势可能会更加明显。从沟通效率的角度来看，理想的情况是被试能够以较少的重复和犹豫来清晰、准确地传达信息。今后在教学中为了提高学生的口语表达效率，可以要求学生事先准备说话的要点，练习清晰的表达，或者在表述中采取更有条理的沟通方式。

对 14 个问题进行两个角度的单因素方差分析，从性别角度来分析各语言参数，在问题（1）（2）（5）（8）（10）（11）（12）（sad）中各语言参数没有显著差异。在问题（3）中，男性组比女性组错误次数多。在问题（4）中，男性组比女性组重复次数多。在问题（6）中，男性组使用的动词数比女性组多。在问题（7）中，男性组的总字数、句子总数、信息词数和动词数都比女性组少。在问题（9）中，男性组比女性组的犹豫次数少。在问题（happy）中，男性组的重复次数和错误数比女性组多。在 14 个问题中，重复出现显著差异的语言测量参数较少，只有重复次数和错误数出现了两次，且都是男性组的重复次数和错误数比女性组多。总体的趋势表明男性组比女性组的口语表现要弱一些，体现在错误次数、重复次数更多，且总字数和信息词数较少，说明男性组给出较少的信息但犯了更多的语言错误。但是性别因素的数据受个体差异影响较大，且本文的样本量不够，所以该结论暂时并不能应用于教学中，只能对后续相关研究提供参考。通过基于性别的语言表现差异的发现，教师可以更好地理解学生在课堂上的行为，从而提供更有有效的支持和指导。这有助于优化学习环境，确保每个学生都能够在语言学习中取得进步。教师需要意识到不同性别

间可能存在的差异，但是也要避免因性别刻板印象而降低对某一性别群体的期望。教师应为所有学生提供平等的学习机会，并鼓励他们根据自身的潜力和能力进行发展。

从不同语言背景（L1 和 L2）角度来分析各语言参数，在问题（1）中，L1 的重复次数和信息输出效率比 L2 大，L1 的总字数、句子总数、句末语气词、错误、动词、虚词、信息词比 L2 少。在问题（2）中，L1 的信息输出效率比 L2 大，L1 的重复次数、错误数和动词数比 L2 少。在问题（3）中，L1 组的信息输出效率比 L2 组大，L1 组的重复次数和错误数比 L2 组少。在问题（4）中，L1 的信息输出效率比 L2 大，L1 的重复次数和错误数比 L2 少。在问题（5）中，L1 的信息输出效率比 L2 大，L1 的重复次数和错误数比 L2 少。在问题（6）中，L1 的信息输出效率比 L2 大，L1 的错误数比 L2 少。在问题（7）中，L1 的信息输出效率比 L2 大，L1 的错误数比 L2 少。在问题（8）中，L1 的信息输出效率比 L2 大，L1 的错误数和重复次数比 L2 少。在问题（9）中，L1 组的错误数和重复次数比 L2 组少。在问题（10）中，L1 组的信息输出效率比 L2 组高。在问题（11）中，L1 组的信息词数、名词数、总字数、句子总数、动词数和信息输出效率比 L2 组高，L1 组的重复次数、错误数和犹豫次数比 L2 组低。在问题（12）中，L1 组的信息输出效率比 L2 组高，L1 组的错误数比 L2 组低。在问题（happy）中，L1 组的句末语气词和信息输出效率比 L2 组高，L1 组的重复次数和错误数比 L2 组低。在问题（sad）中，L1 组的信息输出效率比 L2 组高，L1 组的重复次数、犹豫次数和错误数比 L2 组低。在不同语言背景的 14 个问题数据中，信息词数、虚词数和名词数呈现差异次数为一

次，L1 组的信息词数和名词数比 L2 组多，L1 组的虚词数比 L2 组少。总结总字数、句末语气词、动词数和句子总数的呈现差异次数，一次 L1 组总字数、句末语气词、动词数和句子总数比 L2 组少，另一次相反。犹豫次数呈现差异次数两次都是 L1 组的犹豫次数比 L2 组少。重复次数呈现差异次数为十次，其中多次 L1 组重复次数比 L2 组少，只有一次相反。错误数呈现差异次数都是 L1 组错误次数比 L2 组少。信息输出效率呈现差异次数都是 L1 组信息输出效率比 L2 组高。

从问题的实验整体来看，线上输出的字词数量一般会多一些，比如总字数、虚词数、句子总数等，但是线下的信息输出效率通常比线上的信息输出效率要高。第一次的多个语言参数如句子总数、动词数和犹豫次数等，也较多于第二次的参数，但是第二次的信息输出效率通常比第一次的信息输出效率高。性别因素对于的影响较小，但是综合的口语表现男性组要比女性组稍弱一些，如重复次数和错误次数更多。很明显的差异是 L1 组比 L2 组口语表现强，体现在信息词数和信息输出效率等参数高，错误数和犹豫次数少于 L2 组。为了解 L1 组和 L2 组哪个组别线上线下的差异更大，本研究对两个组的线上线下数据都进行了配对样本 t 检验。在单一图片任务中，L1 组的重复次数呈现差异 $t(df=39) = -2.814, p < 0.05$ ；L2 组无参数呈现差异。在长故事任务中，L1 组的犹豫次数呈现差异 $t(df=39) = -3.537, p < 0.05$ ；L2 组的犹豫次数呈现差异 $t(df=39) = -2.758, p < 0.05$ ，时长 $t(df=39) = -2.779, p < 0.05$ ，CE $t(df=39) = 4.479, p < 0.001$ 。在 14 个问题中选择了两个正式问题和两个非正式问题的整合，共四个问题，并探讨 L1 组和 L2 组在这四个问题中线上和线下的区别。L1 组的多项

参数呈现差异：总字数 $t(df=39) = -4.239, p < 0.001$ ，句子总数 $t(df=39) = -4.115, p < 0.001$ ，信息词数 $t(df=39) = -3.552, p < 0.05$ ，虚词数 $t(df=39) = -3.901, p < 0.001$ ，错误次数 $t(df=39) = -2.554, p < 0.05$ ，名词数 $t(df=39) = -2.894, p < 0.05$ 和 CE $t(df=39) = 2.830, p < 0.05$ ；L2 组无参数呈现差异。根据以上数据得出结论是 L1 组线上线下的口语表现的差异更大，也就是语言水平越接近母语者越会受到不同口语产出环境的影响。母语者通常在日常生活中频繁使用母语进行面对面交流，因此他们可能对线下的环境更为适应。母语者对语言的掌握程度更高，因此他们更能够感知和利用非语言沟通的细微差别，比如面部表情、肢体语言和语调变化，这些在线上交流中不如线下那么明显或容易传达。在线下环境中更容易建立社交联系和亲密感，而这些因素在线上交流中不那么强烈，从而影响被访者的语言表现。所以在考核初级和中级水平的中文二语学习者的口语表现时，线上交流可以是一个有效的替代方式。在高级中文水平的二语学习者进行口语练习时，最好使用线下的方式。整体分析四个因素（口语产出环境、顺序、语言背景和性别）及其交互作用对于多个语言测量参数的影响，得出以下结论。口语产出环境与其他因素的交互作用对多个语言测量参数的影响最大，如总字数、句子总数、信息词、名词和重复次数。虽然二语组有几次相反趋势，但是整体趋势是线上的总字数、句子总数、信息词和名词多于线下。产出顺序与其他因素的交互作用对虚词影响最大，虽然母语组有几次相反趋势，但是整体趋势是第二次的虚词数要比第一次少。语言背景与其他因素的交互作用对错误次数和信息传输效率影响最大，母语组信息传输效率大于二语组，母语组的错误次数少于二语组。性别与其他因素的交互作用影响

较小，但是在名词、动词、重复次数中对数据也有影响，整体趋势是女性组使用的名词数、动词数多于男性组，男性组的重复次数多于女性组。

6.1.4 相似任务的发现

本文将相似任务的对比分析分为了三个部分：长短图片描述任务的数据分析、正式问答和非正式问答的数据分析、回答不同情绪的问题（开心和不开心）的数据分析，得出的结论回答了本文在第一章研究目的部分提出的第三个研究问题：“不同形式的问题是否影响中文二语学习者和母语者的口语表现？如图片描述任务的长短、正式和非正式问题和表现不同情绪的问题。”数据说明被试在完成短图片任务时比长故事任务的信息传输效率更高，名词率更高，但同时错误率也较高。长故事任务比短图片任务的动词率、虚词率、句子长度、句末语气词率和犹豫率更高。短图片由于其直观、明了的特点，汉语学习者在描述时能更快更直接地传达核心信息，尤其体现在对具体对象和概念名称的表达上。教师可以利用这一特点设计针对词汇识别、实物命名等初级语言技能的教学活动，帮助学生快速积累汉语名词词汇。长故事任务在测试语言复杂性上有其独特的优势，被试的动词使用更加频繁，虚词运用也更多样，这有助于他们练习复杂的句式构造和语境适应能力。通过讲述和复述长故事，学生能够逐渐掌握汉语中的各种语气表达和句子连接方式，提高口语和书面表达的连贯性和丰富度。长故事任务促使学生尝试构建较长且完整的句子，并学会正确使用句末语气词来表达情感色彩和态度。以上发现为汉语作为第二语言的教学提供了实证依据，可以根据学生在不同任务类型中的表现调整教学策略。例

如，在初期阶段更多借助短图片任务训练基本词汇和句型，随着学生水平提升，逐渐引入长故事任务以促进更高阶的语言综合运用能力发展。同时，针对不同类型任务反映出的学生弱点，开展个性化的语言修正和拓展训练。在编写汉语作为第二语言的教材或测试试卷时，上述数据也可以作为参考。可以设计更多基于图片的题目来检验留学生对名词的理解和快速表达能力，如看图说话、根据图片写关键词等。考虑到错误率较高，试题应确保图片内容清晰易懂，避免歧义，并且设置适当的梯度难度，允许学生从简单词汇描述逐渐过渡到较复杂的句式。在阅读理解部分，可以加入较长的故事文本，考查留学生的动词应用、虚词使用以及复杂句子结构的理解与把握。通过让学生复述故事、改写结局或者讨论故事中的人物动机等活动，考察他们在实际语境中的语法运用和情感表达。

在正式问题和非正式问题部分，数据说明正式问答的总字数、句子总数、句末语气词、重复次数、信息词数、虚词数、错误次数、名词数、犹豫次数都比非正式问答多，但是非正式问答的信息传输效率高于正式问答。教师可以结合两种问答产出环境的特点，灵活调整课堂氛围。在需要深入探讨、严谨思考的问题上创造正式环境，在需要激发创新思维、锻炼即时反应能力的主题上创造宽松、活跃的非正式环境。在正式问答环节的教学中，教师可以注重对学生语法、词汇运用及逻辑表达能力的培养，通过讲解和练习复杂句型、虚词用法以及如何准确使用信息词来提高书面或口头表达的准确性。对于非正式问答，可采用更生活化、自由开放的教学方式，鼓励学生在轻松环境中快速反应并清晰传达信息，如进行角色扮演、小组讨论、实时问答等互动性强的教学活动。

在考核评价时，可以根据不同类型的问答特点设置相应题目，既测试学生的严谨表述能力（例如写论文、报告），也考察他们简洁高效沟通的能力（如简报、口头汇报）。在回答不同情绪问题（开心和不开心）的中，数据表明涉及开心情绪问题的总字数、句子总数和错误次数都少于不开心情绪问题。在日常的教学和人际沟通中，了解并尊重情绪对言语表达的影响至关重要。教师可以利用这一特点，在传授知识时尽量营造积极、轻松的学习氛围，以提高学生的信息接收和反馈效率；同时，在处理学生情绪问题时，要允许他们有足够的表达时间和机会，耐心倾听，帮助他们理清思路，准确表达内心感受。由于不开心情绪的问题往往包含更多的字词数和错误数，教师可以有针对性地进行辅导和纠正，帮助学生更准确、流畅地表达负面情绪，提高语言表达能力。在教学设计方面，教师可以根据不同情绪下的语言特点，设计更具针对性的教学活动和练习。例如，针对不开心情绪的表达找到常常出现的错误并更正，减少错误次数；针对开心情绪的表达鼓励学生多说多练，以强化学生在情感表达方面的语言技能。

6.2 教学应用与启示

6.2.2 整体趋势

首先，线上输出的字词数量一般会多一些，比如总字数和句子总数等，但是线下的信息输出效率比线上的信息输出效率要高。Chen 和 Chew（2021）的研究结论是学生在线下任务时的焦虑程度比线上高，但是更多学生选择进行线下。结合该研究结论分析本文数据结果，参与者们可能在线上任务时更加轻松

一些，紧张程度更低，所以口语输出量更大。但是从信息输出效率来评判口语表现，线下的口语表现优于线上，在线下环境中交流更加直接和实时。面对面社交环境的构成要求被试及时回应提问，这种直接性会提高线下的信息输出效率，因为交流更快速，实验者和被试之间可以更容易地建立联系，避免误解或信息传递的延迟。线上教学虽然可能产生较多的文字内容记录，但线下直接互动的教学方式在信息传输效率上具有优势。在这两种不同的产出环境下，被试的口头中文表达并没有巨大差异，尤其是语言水平较低的二语者。因此在线下教学不方便实施的时候，线上教学可以成为有效的替代方法，并不会对二语者的中文口语表达造成太大的障碍。同时，教师在设计课程时，也可以结合线上线下教学的特点，合理安排不同形式的教学活动，进行高效混合式教学产出环境的探索。例如，可以利用线上平台进行基础词汇、语法知识点的讲解和练习，确保学生有充足的时间消化吸收；而复杂对话、文化理解等需要深度交流的内容，则可更多地采用面对面的线下方式进行；运用现代信息技术提供个性化学习资源，同时保持线下课堂的高效互动，从而实现对外汉语教学的有效性和趣味性的双重提升。在线下教学中注重提高学生信息输出的效率，培养他们精准、简洁表达的能力；而在线上教学时，则引导学生充分发挥语言输出多、内容丰富的优势，使表达更完整、深入，彰显线上教育的灵活性、便捷性和高效性（邵宏第、任雪梅，2023）。教师可以利用这一发现设计更具针对性的练习和作业，比如布置线下的限时表达任务以强化效率，安排线上的详细阐述作业以拓展思维和丰富表达。线下交流中实时反馈的优势也提示了教师们在实际教学过程中要重视即时性沟通，鼓励学生之间的面对面讨论和角色扮演等活动，

以便及时纠正发音、语法错误，加深对语言结构和语境使用的理解。教研人员也可以研究如何提升线上教学的信息传输效率，比如改进在线教学工具的功能以支持实时互动、模拟真实对话情境等。同时，在评估教学质量时，既要关注学生的知识掌握程度，也要考虑到信息传递效率这一重要指标，构建更加全面的教学效果评价体系。

第一次任务中的多个语言参数如句子总数、动词数和犹豫次数等，也较多于第二次的的数据，但是第二次的信息输出效率比第一次的信息输出效率要高。第一次任务对被试来说是一个较为新颖的体验，因此他们更倾向于表达更多的观点和想法，导致了更多的句子总数、动词数等语言参数。犹豫次数在第一次中更高，可能因为被试可能在第一次任务时感到更紧张或不确定，导致他们在表达时更频繁地停顿或犹豫。第二次会因为被试已经熟悉了的过程和环境，所以他们在表达自己的观点时可能更加流畅和自信，从而提高了信息输出效率。被试在第二次任务中更加熟悉了实验者的提问方式和的目的，因此能够更快地理解问题并给出相应的回答，进一步提高了信息输出效率。从第一次到第二次的进步表明，随着时间推移和练习，学生的学习效果得到了明显改善。因此，教师应引导学生开展自主学习，通过反复实践、自我反思来逐步减少语言输出中的犹豫现象，提高表达流畅度。同时也可以通过有效的反馈机制加速学生从初次接触汉语到熟练运用的过程，在首次沟通后给予详细的个性化反馈，帮助学生识别并改正错误，同时强化他们已经掌握的部分，从而缩短学习曲线。关于产出顺序的另一个发现是在单一家庭图片中，练习效应的体现没有长故事描述任务明显。长故事任务提供了更多的情境线索和上下文信息，使得被试在多

次讲述时能够逐步丰富和完善自己的叙述内容，从而展现出更显著的练习效应。通过复述或重构故事，学习者有机会反复操练词汇、语法结构以及叙事逻辑，这有助于提高他们的语言表达能力和思维组织能力。教师可以根据这一发现调整教学方法，对于需要强化语言应用和表达流利度的教学目标，可以适当增加长故事叙述类的任务设计，鼓励学生从多角度、多层次进行思考和表述，以充分发挥练习效应的作用。也可以进一步优化基于图片的任务设计，使其既能保持一定的挑战性，又能在重复练习过程中体现出明显的进步。比如可以通过提供一系列相关联的图片，引导学生构建一个完整的故事线来讲述，从而更好地发挥练习效应。

一项研究调查了 474 名信息技术专业学生在没有实践环节和有实践环节在线课程的课堂表现焦虑程度，并探讨了他们的课堂表现焦虑程度是否因性别和在线学习的平均时间而异。该研究结果表明，学生的两种课堂表现焦虑程度（没有实践环节和有实践环节的在线课程）因性别和不同的在线学习时间组而异，并且女性的两种课堂表现焦虑程度均显著高于男性（Zhang, Zhou & Xu, 2022）。另一项研究调查了 1377 名大学生在上网课时的学习状态（Saadé, Kira, Mak & Nebebe, 2017），30% 的学生都表示在上网课时会有一些焦虑心态，且女性女生比男性学生们焦虑情况更严重。但是在本论文研究结果中，综合的口语表现男性组要比女性组稍弱一些，如重复次数和错误次数更多，且性别因素对于的影响较小。因此性别在本研究的语言学习者中不是一个显著影响口语表现的因素。尽管男性组的口语表现略弱于女性组，但差异不是非常显著，这可能是因为个体差异在性别之间更大，而性别本身对于口语表现的影响较小。在参

与本次研究的 40 名二语学习者中，男性组 HSK5 级者 17 人，HSK6 级者 3 人；女性组 HSK5 级者 13 人，HSK6 级者 7 人，这也是男性组整体口语水平弱于女性组的影响因素。

实验结论中十分显著且清晰可见的差异在于，L1 组在口语表现方面明显强于 L2 组。这一优势具体体现在信息词数以及信息输出效率参数上，L1 组在这些方面的数值较高。同时，L1 组的错误数显著少于 L2 组，犹豫次数也明显低于 L2 组。这一发现与事先所设定的研究假设高度一致。对于二语学习者而言，除了语言本身所带来的挑战外，他们还必须去努力适应汉语文化的深厚背景和独特语境，而这一因素不可避免地会对他们的口语表现产生一定程度的影响。针对 L2 学习者所开展的教学工作应当更加侧重于实践性以及情境化。可以通过精心设计模拟真实场景的活动、为学生提供大量互动练习的宝贵机会等方式，助力学生在实际操作中切实提高信息处理的速度，有效降低错误率，并逐步减少在表达过程中出现的犹豫现象。在对学生的学习进展进行评价时，除了密切关注他们在语法上的正确性和词汇量的多少外，还应当对信息传输效率、错误纠正能力以及表达流畅度等关键维度给予充分且足够的重视，并及时为学生提供具有高度针对性、极具建设性的反馈意见，以此推动学生语言综合技能的稳步提升。在教学规划的具体操作方面，教师需要着重加强词汇积累、语法训练等基础性的内容，从而有效提高信息词数和语言输出效率，显著减少错误次数。在教学重点的精准把握上，教师还应当注重培育二语学习者的语言自信心，积极鼓励他们大胆地进行表达，最大程度地减少犹豫次数。与此同时，针对那些常见的错误进行重点的讲解和有针对性的纠正，协助他们循序渐

进地提高口语表达的准确性。教师还能够引导母语学生与二语学生展开交流互助，促使二语学习者在真实的交流过程中借鉴母语组的丰富经验，进而显著提高口语水平。特别是对于汉语水平相对较高的二语学习者来说，他们在已经成功掌握足量的词汇和语法规则之后，需要用心学习并模仿母语者地道的语言运用方式，例如句末语气词的巧妙运用、语音语调的合理改变以及文化语言的准确使用等，以便能够更为出色地提高语言交流能力和文化理解能力。

6.2.3 进一步研究

之后的研究可以分析线上线下的课堂组织形式和互动方式如何影响口语发展，比如同伴互评、教师反馈的有效性、小组讨论的活跃程度，以及线上异步交流与面对面即时反馈在提高口语能力上的差异。Kong (2009) 的研究发现被访者在四个图片中都有相似的语言产出，但是更年轻的被访者倾向于产出更多的信息词并且词汇多样性 (Lexical Diversity) 更丰富，王勃然等人 (2019) 认为暂无有力证据证实或否定二语习得语法关键期的存在。之后的研究也可以加入年龄等因素，研究在线上 and 线下环境中语言表达和交流效率的具体差异，比较不同性别、年龄和文化背景的参与者在这两种环境下的表现。探讨技术介入 (比如视频质量、互动工具的使用) 如何影响语言产出和沟通质量，运用 AI 技术赋能对外汉语教学。教师可以设计并实施混合式教学产出环境，结合线上线下的优势，探究这种产出环境下二语学习者的口语习得效果，并评估其对提高整体语言能力的贡献。探讨二语学习者在不同环境下如何利用线上线下资源进行口语练习，比如在线平台的自主学习资源使用情况，以及线下活动参与度对

口语进步的作用。在产出顺序方面，可以进行多次口语产出任务，探讨顺序和次数如何影响语言表现，包括语言输出的数量和质量。考察随着次数的增加参与者语言表现的改善程度，以及这种改善是否因人而异。其他因素方面也可以研究线上与线下环境对学习者心理状态的不同影响，例如：焦虑水平、自我效能感、注意力集中度，以及虚拟环境中身份匿名性和距离感等因素对口语表达的潜在作用。对于本文数据中体现的性别差异现象，今后的研究可以缩小语言数据之间的个性化差异，如控制参与者的汉语水平、年龄、语言背景等。教研人员可以进一步探讨其背后的原因，包括生物学、社会学、心理学等多个层面，以科学的态度探究性别对语言习得和使用是否存在影响，深入研究性别如何影响语言错误的类型和频率，并讨论出特定教学策略来帮助男性和女性改善他们的语言技能。后续研究在测量参数中可以加入不同国籍的信息，扩大样本量，缩小被试个体间差异，结合母语迁移理论研究不同国家的中文二语学习者学习过程中的异同。将母语按照语系（如印欧语系、汉藏语系等）或具体语言（如英语、日语、韩语、法语等）分类，对比分析各母语背景对中文二语学习者线上线下的口语输出有何特殊影响。深入研究不同母语背景的学生在语音、词汇、语法结构、语用习惯等方面的迁移现象，以及这些迁移如何影响他们的中文口语表现。根据不同国籍中文二语学习者的语言习得特点和困难点，探索并设计更具针对性的教学策略和方法，以促进学习者在口语能力方面的有效提升。

参考文献

- Abdul-Ghani, E., Kim, J., Kwon, J., Hyde, K.F., & Cui, Y.G. (2022). Love or like: gender effects in emotional expression in online reviews. *European Journal of Marketing*, 56(12), 3592-3616. <https://doi.org/10.1108/EJM-01-2021-0064>
- Andrade, S. B., & Andersen, D. (2020). Digital story grammar: a quantitative methodology for narrative analysis, *International Journal of Social Research Methodology*, 23(4), 405-421. DOI: 10.1080/13645579.2020.1723205
- Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: a crosslinguistic developmental study*. L. Erlbaum Associates.
- Boles, L. (1997). Conversational analysis as a dependent measure in communication therapy with four individuals with aphasia. *Asia Pacific Journal of Speech, Language and Hearing*, 2(1), 43-61.
- Boyle, J. P. (1987). Sex Differences in Listening Vocabulary. *Language Learning*, 37(2), 273–284. <https://doi.org/10.1111/j.1467-1770.1987.tb00568.x>
- Budianto, P. (2023). Survey and Research on Online Chinese Teaching in Indonesia. *Jurnal Cakrawala Mandarin*, 7(2), 1-10.
- Clandinin, D. J., & Connelly, F. M. (2000). *Narrative inquiry: experience and story in qualitative research*. Jossey-Bass Inc.
- Chen, Y. Q., & Chew, S. Y. (2021). Speaking Performance and Anxiety Levels of Chinese EFL Learners in Face-to-Face and Synchronous Voice-Based Chat. *Journal of Language and Education*, 7(3), 43-57.

<https://doi.org/10.17323/jle.2021.11878>

- Cui, G. (2022). The SWOT analysis and research on the development strategy of online Chinese teaching international Chinese education under the new situation. *International Chinese Language Teaching*, 1(1), 58–71. Retrieved from <https://so04.tci-thaijo.org/index.php/iclt/article/view/252125>
- Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second Language Fluency: Judgments on Different Tasks. *Language Learning*, 54(4), 655–679. <https://doi.org/10.1111/j.1467-9922.2004.00282.x>
- Erbaugh, M. (2006). *Chinese classifiers: Their use and acquisition*. Cambridge, United Kingdom: Cambridge University Press.
- Ellis, N. C. (2002). FREQUENCY EFFECTS IN LANGUAGE PROCESSING. *Studies in Second Language Acquisition*, 24(2), 143–188. <https://doi.org/10.1017/S0272263102002024>
- Foster, R., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18(3), 299–323.
- Fowler, J. H., & Christakis, N. A. (2008). Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *BMJ*, 337(7685), 23–26. <https://doi.org/10.1136/bmj.a2338>
- Firat, M., Öztürk, A., Günes, I., Çolak, E., Beyaz, M., & Büyük, K. (2019). How e-learning engagement time affects academic achievement in e-learning environments. *Open Praxis*, 11(2), 129–141.

- Franzosi, R. (2010). *Quantitative narrative analysis*. SAGE.
- Giannikas, C. N. (2022). *Transferring language learning and teaching from face-to-face to online settings*. IGI Global.
- Guiller, J., & Durndell, A. (2007). Students' linguistic behaviour in online discussion groups: Does gender matter? *Computers in Human Behavior*, 23(5), 2240-2255.
<https://doi.org/10.1016/j.chb.2006.03.004>
- Güngör, M. N., & Güngör, M. A. (2021). Reconceptualizing Testing in Times of COVID-19 from an Activity Theory Lens: A Narrative Inquiry of a Test-designer's Experiences. *The Asia-Pacific Education Researcher*, 30(6), 531-539.
- Hao, Y., Sheng, L., Zhang, Y., Jiang, F., Villiers, J. de, Lee, W., & Liu, X. L. (2018). A Narrative Evaluation of Mandarin-Speaking Children with Language Impairment. *Journal of Speech, Language, and Hearing Research*, 61(2), 345–359. https://doi.org/10.1044/2017_JSLHR-L-16-0367
- Hellman, A. B. (2011). Vocabulary size and depth of word knowledge in adult-onset second language acquisition. *International Journal of Applied Linguistics*, 21(2), 162-182.
- Holmqvist, K., Holsanova, J., Johansson, V., & Strömquist, S. (2005). Perceiving and Producing the Frog Story. In *Perspectives on Language and Language Development* (pp. 289–292). Springer, US. https://doi.org/10.1007/1-4020-7911-7_21
- Hulstijn, J. H. (2002). What does the impact of frequency tell us about the language

acquisition device? *Studies in Second Language Acquisition*, 24(2), 269–273.

<https://doi.org/10.1017/S0272263102002115>

Justice, L. M., Bowles, R., Pence, K., & Gosse, C. (2010). A scalable tool for assessing children's language abilities within a narrative context: The NAP (narrative assessment protocol). *Early Childhood Research Quarterly*, 25(2), 218–234.

Kong, A.P.H., & Law, S.P. (2005). *Validation of the Cantonese linguistic communication measure (CLCM)* [Paper presentation], The 43rd Annual Meeting of Academy of Aphasia, Amsterdam, Netherlands.

Kong, A. P., & Law, S. P. (2009). A linguistic communication measure for monitoring changes in Chinese aphasic narrative production. *Clinical linguistics & phonetics*, 23(4), 255–269. <https://doi.org/10.1080/02699200802673234>

Law, S. P., Wong, W., & Wong, E. (2008). A semantic treatment for Cantonese anomic patients: implications for the relationship between impairment and therapy. In S. P. Law, B. Weekes, & A. Wong (Eds.), *Language disorders in speakers of Chinese* (pp. 273–292). Clevedon, UK: Multilingual Matters.

Luo, H. (2014). Foreign Language Speaking Anxiety: A Study of Chinese Language Learners. *Journal of the National Council of Less Commonly Taught Languages*, 15(1), 99-117.

Mayer, M. (1969). *Frog, where are you?* New York: Dial Press.

Menn, L., Ramsberger, G., & Helm-estabrooks, N. (1994). A linguistic

- communication measures for aphasic narratives. *Aphasiology*, 8(4), 343 – 359.
- Meng, N. (2021). “Repeat After Me”: Is There a Better Way to Correct Tone Errors in Teaching Mandarin Chinese as a Second Language? In: Yang, C. (Eds.), *The Acquisition of Chinese as a Second Language Pronunciation* (pp. 163-173). Springer, Singapore. https://doi.org/10.1007/978-981-15-3809-4_7
- Michel, M. C. (2015). Book review: A Housen, F Kuiken and I Vedder, *Dimensions of L2 performance: Complexity, accuracy and fluency in SLA* [Review of *Book review: A Housen, F Kuiken and I Vedder, Dimensions of L2 performance: Complexity, accuracy and fluency in SLA*]. *Language Teaching Research*, 19(1), 123–124. SAGE Publications. <https://doi.org/10.1177/1362168814542719>
- Na, Z. (2022). The Research of Online Chinese Teaching to Ethnic Chinese Children. *MANDARINABLE: Journal of Chinese Studies*, 1(2), 83-90.
- Poole, A. (2021). Narrative inquiry and relational ethics: Negotiating the lived experiences of international school teachers in China. *International Journal of Research & Method in Education*, 44(2), 113-124.
- Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: procedure and data. *Brain and Language*, 37(3), 440-479.
- Scheg, A. G. (2014). *Identifying the Differences Between Online and F2F Teaching*. In *Reforming Teacher Education for Online Pedagogy Development*. IGI Global.
- Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P.,

- Fulbright, R. K., ... & Fletcher, J. M. (1995). Sex differences in the functional organization of the brain for language. *Nature*, 373(6515), 607-609.
- Shewan, C.M. (1988). The Shewan Spontaneous Language Analysis (SSLA) system for aphasic adults: Description, reliability and validity. *Journal of Communication Disorder*, 21(2), 103-138.
- Saadé, R. G., Kira, D., Mak, T., & Nebebe, F. (2017). *Anxiety and Performance in Online Learning*. Proceedings of the Informing Science and Information Technology Education Conference, Vietnam.
<http://www.informingscience.org/Publications/3736>
- Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition*, 73(3), 205–246. [https://doi.org/10.1016/S0010-0277\(99\)00053-0](https://doi.org/10.1016/S0010-0277(99)00053-0)
- Stickler, U., & Shi, L. (2013). Supporting Chinese speaking skills online. *System*, 41(1), 50-69.
- Stahl, V. (1990). *Reference cohesion in LLD preadolescents' procedural and narrative discourses* [Paper presentation]. the American Speech-Language-Hearing Convention, Seattle, WA.
- Wang, C., & Canagarajah, S. (2024). Postdigital ethnography in applied linguistics: Beyond the online and offline in language learning. *Research Methods in Applied Linguistics*, 3(2), 100111.
- Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the

- second language classroom. *CALICO Journal*, 13(2), 7-26.
- Yorkston, K. M., & Beukelman, D. R. (1980). An analysis of connected speech samples of aphasic and normal speakers. *Journal of Speech and Hearing Disorders*, 45(1), 27–36.
- Yu, L., & Xu, L. (2022). Research on the Online Learning Experience of South Asian Students during COVID-19: The Mediating Effect of Online Learning Status on Autonomous Learning and Learning Effect Evaluation, *Journal of International Students*, 12(SI), 45-60.
- Zhang, L., & Law, C. C. (2022). *Face-to-face versus virtual communication: A comparison of occurrences of Cantonese SFPs* [Paper presentation]. The 6th Workshop on Innovations in Cantonese Linguistics (WICL-6), The Ohio State University, USA.
- Zhao, L. X., Blankinship, B., Duan, Z., Huang, H., Sun, J., & Bak, T. H. (2020). Comparing face-to-face and online teaching of written and spoken Chinese to adult learners: An Edinburgh-Sheffield case study. *International Journal of Chinese Language Teaching*, 1(1), 83-98.
- Zhang, X., Zhou, F., & Xu, J. (2022). Technical College Students' Practical Performance Anxiety during Online Learning: Difference in Gender and Average Time of Online Learning. *Sustainability (Basel, Switzerland)*, 14(13), 8218. <https://doi.org/10.3390/su14138218>
- Zhao, Y., Llorente, A. M. P., & Gómez, M. C. S. (2021). Digital competence in

higher education research: A systematic literature review. *Computers and Education*, 168(1), 104212–104212.

<https://doi.org/10.1016/j.compedu.2021.104212>

- 陈默（2012）：美国留学生汉语口语产出的流利性研究，《语言教学与研究》，25（2），17-24。
- 陈默、陈青（2023）：任务因素对中文二语口语产出复杂度、准确度和流利度的影响，《陕西师范大学学报（哲学社会科学版）》，52（6），129-143。
- 丁安琪、程海婷、于艳、丁涵和陈文景（2021）：全球汉语学习者口语能力表现的现状与问题——基于2018年HSKK大数据的分析，《语言教学与研究》，25（6），13-23。
- 杜娟（2019）：频次效应下二语学习者口语产出的变化机制，《开封教育学院学报》，39（3），53-54。<https://doi.org/10.3969/j.issn.1008-9640.2019.03.022>
- 冯华（2012）：性别因素对留学生汉语口语习得的影响，《现代语文》，22（8），86-88。
- 郭江华（2023）：近十年国内汉语二语口语习得研究述评——基于四种汉语类专业期刊，《现代语言学》，11（6），2390-2397。
- 顾明远、滕珺（2020）：后疫情时代教育国际交流与合作的新挑战和新机遇，《比较教育研究》，42（9），3-7。
- 黄韧、张清芳和李丛（2017）：消极情绪抑制了老年人的口语产生过程，《心

理与行为研究》，15（3），372-378。

姜永超（2016）：信息技术驱动的汉语国际教育现状与反思，《教学研究》，39（2），77-81。

江新（2006）：汉字频率和构词数对非汉字圈学生汉字学习的影响，《心理学报》，38（04），489-496。

康叶钦（2014）：在线教育的“后 MOOC 时代”，《清华大学教育研究》，35（1），85-93。

陆俭明、崔希亮、张旺熹、张鹏、梁霞、郑艳群、冯丽萍、孙瑞、郝美玲和古川裕（2020）：“新冠疫情下的汉语国际教育：挑战与对策”大家谈（下），《语言教学与研究》，46（5），1-16。

陆俭明、李宇明和贾益民等（2020）：“新冠疫情对国际中文教育影响形势研判会”观点辑录，《世界汉语教学》，34（04），435-450。

刘士娟（2018）：在线中文教学：什么和为什么？，《国际汉语教育》，3（02），11-25。doi:CNKI:SUN:GZZY.0.2018-02-005

吕叔湘（1979）：《汉语语法分析问题》，北京：商务印书馆。

锒宝香（2001）：国小低阅读成就学生的口语述说能力：语言层面的分析，《特殊教育学报》，90（15），129-175。

<https://doi.org/10.6768/JSE.200109.0129>

邵宏第、任雪梅（2023）：新时代外语院校对外文化交流方式创新研究——以大连外国语大学“汉语桥”线上团组交流项目为例，《Евразия. Диалог культур》，6（1），683-687。DOI: 10.24412/cl-37114-2023-1-683-687

- 宋晖、谭紫格（2018）：对外汉语在线教学的“三教”问题，《国际汉语教育》，3（02），4-10。doi:CNKI:SUN:GZZY.0.2018-02-003
- 吴爱梅（2023）：学生引发的线上汉语口语互动教学研究，《汉学与国际中文教育》，4（1），84-92。
- 吴继峰、高敏和赵晓娜（2022）：汉语二语学习者口语产出的任务复杂度效应研究，《四川师范大学学报（社会科学版）》，49（6），169-176。
- 王勃然、黄桂玲和代佳馨（2019）：基于年龄效应的二语/外语语法习得研究述评，《唐山师范学院学报》，41（4），1-8。
- 王改花、张李飞和傅钢善（2021）：学习者特征对混合学习效果影响研究，《开放教育研究》，27（1），71-83。doi:10.13966/j.cnki.kfjyyj.2021.01.008
- 文秋芳（2003）：频率作用与二语习得，《外语教学与研究》，35（6），151-154。
- 文秋芳、杨佳（2020）：从新冠疫情下的语言国际教育比较看国际中文在线教育的战略价值，《语言教学与研究》，206（06），5-12。
- 王希竹、金晓艳（2021）：二语学习者口语表达流利度的影响因素研究——基于特定情境焦虑视角，《汉语学习》，10（6），85-93。
- 王植洵、张锦坤（2017）：提取练习效应的产生机制：编码、提取的整合演变，《心理科学进展》，25（9），1512-1520。
- 谢浩然、陈协玲、郑国城和王富利（2022）：人工智能赋能个性化学习：E-Learning推荐系统研究热点与展望，《现代远程教育研究》，34（03），15-23。

- 萧季桦（2007）：《汉语口语之情绪语言》，台湾，台湾大学博士论文。
- 徐晶凝（2023）：如何在流水句中划定句子的边界——兼谈单句、复句、句群的划分问题，《语言教学与研究》，46（4），58-69。
- 徐琴芳（2005）：不同任务下的口语准确性研究，《山东外语教学》，45（6），65-68。<https://doi.org/10.3969/j.issn.1002-2643.2005.06.017>
- 谢婷、李京南（2021）：大学英语口语实验教学产出环境实践，《实验科学与技术》，1（5），108-114。
- 岳海燕（2021）：线上线下融合教学，《江西教育》，8（21），1。
<https://doi.org/10.3969/j.issn.1006-270X.2021.21.003>
- 尹振宇（2016）：频次效应及其对外语教学的启示，《教育文化论坛》，8（5），77-80。
- 周丹丹（2010）：口语练习频次效应与注意力资源的利用，《外语研究》，41（6），50-55。
- 张国雪（2023）：多模态话语在线上、线下语法课中的对比研究，《Advances in Social Sciences》，12（6），2903-2911。
- 赵雷（2008）：对外汉语口语教学目标的实现，《汉语学习》，10（6），85-91。
- 赵雷（2021）：线上任务型汉语口语教学中的学习者互动策略，《国际汉语教学研究》，11（3），82-89。
- 郑艳群（2001）：课堂上的网络和网络上的课堂——从现代教育技术看对外汉语教学的发展，《世界汉语教学》，38（4），98-104。

<https://doi.org/10.3969/j.issn.1002-5804.2001.04.012>

郑咏滢、刘飞凤（2020）：复杂理论视角下任务复杂度对二语口语表现的影响，《现代外语》，43（3），365-376。

附录

附录一 问卷信息表

香港教育大学

The Education University of Hong Kong

研究项目：中文非母语者线上与线下中文使用的对比

Face-to-face versus virtual communication: A comparison of Chinese language usage among non-native Chinese speakers in China

实验问卷（实验人员版本）

请写出以下个人信息 please answer the questions below:

姓名 name:

性别 gender:

生日 birth date:

国籍 nationality:

教育程度 degree of education:

职业 occupation:

第一语言 first language:

第二语言 second language:

第三语言及以上（若有）third language or above (if have):

汉语水平 Chinese level:

开始学习中文的年份 the date of starting learning Chinese:

居住中国的时间 the length of staying in China:

其它相关信息 other information:



附录二 问卷 A

A 问卷 Questionnaire A

1. 看图说话 *read the pictures* :

- (1) 请说说图片里有什么人? 有什么东西? 在发生什么事?

Please tell me who is in the picture? What is there? What's going on?

- (2) 请根据系列图片, 讲述《青蛙的故事》。

Please tell 《The Frog Story》 according to the series of pictures.

2. 家常聊天 *small conversations* :

- (1) 你可以自我介绍一下吗? Introduce yourself please.

- (2) 你喜欢在大城市工作还是小城市?

Do you prefer to work in a big city or small city?

- (3) 请介绍你的一位朋友。比如说他的外表、性格和喜好?

Please introduce a friend of yours, such as his/her appearance, personality and hobbies?

- (4) 你平时的生活是怎样的? 请描述你的一天。

What is your daily life like? Please describe your day.

- (5) 周末一般会怎么度过? 和谁一起度过?

How do you usually spend your weekends? Who do you spend with?

- (6) 请介绍你的国家。比如风俗文化。

Please introduce your country, such as customs and culture.

- (7) 你喜欢吃什么食物? 请描述一下?

Could you describe the food you like?

- (8) 你最喜欢看什么书? 为什么?

What is your favorite book? Why?

- (9) 你最难忘的一件事是什么?

What is the most memorable thing in your life?

- (10) 你最喜欢的城市是哪里? 为什么?

Where is your favorite city? Why?

3. 社会议题 *formal conversations* :

- (1) 你来到中国后, 认为中国和自己的国家不一样的地方在哪里?

After you came to China, what are the differences between China and your country?

- (2) 你认为空气污染的来源是什么? 你有什么建议?

What are the sources of air pollution? Do you have suggestions on air pollution?

- (3) 你认为很好的家长应该是什么样子的?

What do you think good parents should be like?

4. 请根据以下假设的场景, 做出相应的表演 *Please act accordingly according to the following hypothetical scenarios* :

- (1) 假如你中了奖获得了一万人民币的奖金, 你想告诉你的朋友, 你会怎么说?

If you won a prize of 10,000 RMB and you want to tell your friend, what would you say?

- (2) 如果老师说你的论文很差, 你想告诉你的朋友, 你会怎么说?

If the teacher said your paper is terrible, you are upset you want to tell your friend, what would you say?

附录三 问卷 B

B 问卷 Questionnaire B

1. 看图说话 *read the pictures* :

- (1) 请说说图片里有什么人? 有什么东西? 在发生什么事?

Please tell me who is in the picture? What is there? What's going on?

- (2) 请根据系列图片, 讲述《青蛙的故事》。

Please tell 《The Frog Story》 according to the series of pictures.

2. 家常聊天 *small conversations* :

- (1) 你可以自我介绍一下吗? Introduce yourself please.

- (2) 你喜欢在大城市生活还是小城市?

Do you prefer to live in a big city or small city?

- (3) 请介绍你的一位家人。比如说外表、性格和喜好?

Please introduce a family member of yours. Like his/her appearance, personality and hobbies?

- (4) 你在学校的生活是怎样的? 请描述你在学校的一天。

What is your school life like? Please describe your day in school.

- (5) 假期一般会怎么度过? 和谁一起度过?

How do you usually spend your vacations? Who do you spend with?

- (6) 请介绍你的城市。比如风俗文化。

Please introduce your city, such as customs and culture.

- (7) 你不喜欢吃什么食物? 请描述一下?

Could you describe the food you dislike?

- (8) 你最喜欢什么电影? 为什么?

What is your favorite movie? Why?

- (9) 你最开心的一件事是什么?

What is the most happy thing in your life?

- (10) 你最喜欢的国家是哪里? 为什么?

Where is your favorite country? Why?

3. 社会议题 *formal conversations* :

- (1) 你来到中国后, 认为中国和自己的国家相似的地方在哪里?

After you came to China, what are the similarities between China and your country?

- (2) 你认为水污染的来源是什么? 你有什么建议?

What are the sources of water pollution? Do you have suggestions on water pollution?

- (3) 你认为很好的老师应该是什么样子的?

What do you think a good teacher should be like?

4. 请根据以下假设的场景, 做出相应的表演 *Please act accordingly according to the following hypothetical scenarios* :

- (1) 假如你被人骗了一万人民币, 你很不开心, 你想告诉你的朋友, 你会怎么说?

If you are cheated of 10,000 RMB, you are very upset, and you want to tell your friend, what would you say?

- (2) 如果老师说你的论文很好, 你想告诉你的朋友, 你会怎么说?

If the teacher said your paper is excellent, you are excited you want to tell your friend, what would you say?