

語言間相互理解度測試的理論與實踐
Theory and Practice of Mutual Intelligibility Measurement between Languages

by

溫涵
WEN Han

哲學博士（語言學—漢語）
Doctor of Philosophy in Linguistics (Chinese)



UNIVERSIDADE DE MACAU

UNIVERSITY OF MACAU

人文學院
Faculty of Arts and Humanities

澳門大學
University of Macau



澳門大學

UNIVERSIDADE DE MACAU

UNIVERSITY OF MACAU

語言間相互理解度測試的理論與實踐
Theory and Practice of Mutual Intelligibility Measurement between Languages

By

溫涵
WEN Han

導師：徐杰教授
SUPERVISOR: Professor Xu Jie

學系：中國語言文學系
DEPARTMENT: Department of Chinese Language and Literature

哲學博士（語言學—漢語）
Doctor of Philosophy in Linguistics (Chinese)

2019

UNIVERSITY OF MACAU

人文學院
Faculty of Arts and Humanities

澳門大學
University of Macau



澳門大學

UNIVERSIDADE DE MACAU

UNIVERSITY OF MACAU

Author's right 2019 by
WEN, Han



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU



澳門大學

UNIVERSIDADE DE MACAU

UNIVERSITY OF MACAU

致謝

暮然回首，發現已到南國求學六載。碩士畢業後，承蒙老師抬愛，讓我可以有機會拜入徐門，從此踏入學術的大門。老師在博士入學之初就告訴我，在博士階段，要從一個學生轉變為學者，要做知識的主人。導師是一名嚮導，給學生指一條路讓學生自己去探險。每每想到此，都能感受到老師給予的期待和自身的使命感。感謝老師高瞻遠矚地為我指的這條路，雖然荊棘密布，鮮有人至，然而越走越能夠發現，這是一條正確的路，一條可以通往學術中心的路，一條四通八達的路。有時候自認為已經走了很遠，其實才剛剛邁出了一小步。

博士論文的寫作並不只是一個人努力就可以完成的，在此首先要感恩導師徐傑教授，老師把這個題目交給我就是給予我充分的信任，要求我既要腳踏實地田野調查，又要有天馬行空的想象力與悟性。老師對於這篇論文的指導讓這篇論文更加厚重，更具說服力。感恩老師的批評，老師的每一個教誨學生都銘記於心，害怕辜負了老師的期望，惟以刻苦學習，努力工作來踐行老師的指導。

在論文的實驗部分，我還要感謝天津師範大學董悅凱老師、天津外國語大學易正中教授、湖南師範大學徐超老師、天津中學王鑫副校長、張錯灃律師、廈門一中張蘭玉老師以及北京大學、廈門大學、中南林業科技大學、廈門一中、天津中學的諸位老師同學的協助與支持。沒有以上領導老師同學的協助，我很難在較短的時間內找到近500名被試來完成所需的實驗。很多人都是萍水相逢，僅通過基本的介紹就提供給我無私熱情的幫助，讓我倍受感動。

在讀博士期間，我也受到了澳門大學各位老師，師兄師姐師弟師妹們的照

顧。在我徬徨、猶豫和困惑的時候，都有他們一直鼓勵、安慰和勸導。還要感謝我的女朋友對我的支持與信任。最後要感激我的父母，感謝他們對我的教育，讓我能夠沒有後顧之憂地在知識的海洋中暢遊。

一篇博士論文背後凝結了老師、同學、家人的關懷和期許，溫涵沒齒難忘，唯有繼續努力，將所有支持我的人的期許轉化為動力，不斷鞭策自己進步，方才對得起所有真正關心批評我的人。



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

摘要

理解度理論在醫學、信號學等眾多學科內均已有相關理論與實踐標準。語言學視野下的理解度概念也已被學界所熟知，但多停留在簡單應用層面，鮮有系統的理论研究，致使實踐研究成果往往因缺乏理論依託而大大降低可用性。

本文在充分掌握前人文獻的基礎上，詳細構建起在語言學視野下的理解度理論，對理解度的概念、性質、影響因素、測試方法、參照標準以及理論意義等內容進行了系統闡述，並討論了理解度理論的局限以及未來發展方向。本文論證了在理解度測試中加入時間參數的必要性與可行性，同時完善有關時間參數的理論和計算方法，並以粵方言、湘方言及閩南方言的理解度實驗為例，驗證有關時間參數的理論和可操作性。經實驗得到粵方言與湘方言之間的相互可理解度為0.533，粵方言與閩南方言之間的相互可理解度為0.506，湘方言與閩南方言之間的相互可理解度為0.476。

此外本文利用理解度測試作為標準判定粵式普通話的性質，將其與普通話和非粵方言區被試對粵方言的理解度相對比。經實驗得到普通話的理解度為0.952，非粵方言區被試對粵方言的理解度為0.641。整體對粵式普通話的理解度為0.858，與前一組實驗結果對比後發現粵式普通話理解度更接近普通話的理解度，且與普通話相同具有理解度上的共享性，故認為粵式普通話應屬於寬式普通話的範疇。

本文第一次直接且完整地在語言學視野下研究理解度理論，自上而下地建構理解度理論框架，為其從理論層面上奠定堅實的基礎，明確了理解度意義與價值，即將其作為衡量語言間距離的標準。首次在理解度測試中加入反應時

間，能夠獲得更符合實際的理解度數值。最後，利用前文所述的理論方法指導個案研究，為方言普通話研究提供定量研究的案例，為方言普通話的性質研究提供新的研究方法。

理解度理論綜合多門學科，觸及語言的本質與核心，同時應用範圍廣泛。本文構建起理解度理論也是搭建起與其他學科交互的平台，多學科交叉合作共同促進理論與實踐的發展。



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

Abstract

There have been some theories and practical standards of in many disciplines such as medical and signal science. The concept of Intelligibility in the field of linguistics has also been well known, but it stays at the simple application level, but there are few systematic theoretical studies on it, which makes the practical research lack of theoretical support and applicability.

On the basis of fully grasping the former literature, this dissertation constructs the theory of intelligibility in the perspective of linguistics, and systematically expounds the concept, nature, influencing factors, methodology, reference standards and theoretical significance. It also discusses the limitations of the theory and the direction of future development. This dissertation demonstrates the necessity and feasibility of adding time parameters in the intelligibility test, and at the same time perfects the theory and calculation method of time parameters.

This dissertation takes the intelligibility experiment of Cantonese, Xiang dialect and Min dialect as an example to verify the theory and operability of time parameters. The intelligibility between Cantonese and Xiang dialect is 0.533; the intelligibility between Cantonese and Min dialect is 0.506; the intelligibility between Xiang dialect and Min dialect is 0.476.

In addition, this dissertation take the result of the intelligibility as a standard to analyze the category of Cantonese-Mandarin, comparing with the intelligibilities of Cantonese and Mandarin. The intelligibility of Mandarin is 0.952, the intelligibility of non-cantonese speaker is 0.641, and the intelligibility of Cantonese-Mandarin is 0.797. Compared with the results of the first experiments, it is found that the intelligibility of Cantonese-Mandarin is closer to the intelligibility of Putonghua, and it has the same pattern of intelligibility as Mandarin. Therefore, it is believed that Cantonese-Mandarin should widely belong to the category of Putonghua.

For the first time, this dissertation directly and completely studies the theory of intelligibility, constructing a theoretical framework of intelligibility, founding a solid foundation on the theoretical level, and clarifying the meaning and value of intelligibility, which is the standard for measuring the distance between languages. For the first time, adding reaction time to the intelligibility test provides a more realistic value of intelligibility. Finally, using the theoretical methods described above to guide case studies, provide quantitative research for dialectical Mandarin research, and

demonstrate a new research method for the study of the nature of dialect Mandarin. Intelligibility theory integrates multiple disciplines, touches the essence and core of language, and has a wide range of applications. This dissertation builds a theory of intelligibility which is a platform for interaction with other disciplines. Multidisciplinary cooperation promotes the development of theory and practice.



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

聲明

本人所提交的論文，除了經清楚列明來源出處的資料外，其他內容均為原創；

本論文的全部或部分未曾在同一學位或其他學位中提交過。

本人聲明已知悉及明白《澳門大學學生學術誠信處理規條》及《澳門大學學生紀律規章》。



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

目錄

致謝.....	i
摘要.....	iii
聲明.....	vii
圖表目錄.....	xiv
第一章 緒論.....	1
1.1 研究內容與研究目標.....	1
1.1.1 理解度.....	1
1.1.2 理解度測試.....	2
1.1.3 相互理解度測試.....	4
1.1.4 普通話研究.....	5
1.2 研究現狀.....	7
1.2.1 國際研究現狀.....	7
1.2.2 國內研究現狀.....	8
1.3 研究方法.....	9
1.4 研究問題.....	10
1.5 潛在研究意義.....	11
1.5.1 理解度理論的研究意義.....	11
1.5.2 理解度測試的研究意義.....	13
1.5.3 普通話理解度的研究意義.....	16
1.6 原創聲明.....	17
第二章 可理解度概念及相關理論.....	18
2.1 理解度的源流.....	18
2.2 理解度的概念.....	22
2.2.1 通信模式.....	23
2.2.2 理解度的概念.....	28
2.3 理解度的性質.....	35
2.4 理解度的影响因素.....	39
2.5 理解度的測試方法.....	50
2.5.1 實驗材料.....	51
2.5.2 實驗方法.....	52
2.5.3 實驗對象.....	56
2.5.4 測試工具.....	57

2.6 理解度常模參照標準.....	59
2.6.1 常模參照與標準參照	59
2.6.2 理解度常模參照標準	62
2.7 理解度的應用.....	64
2.8 理解度的意義.....	71
2.8.1 建立標準	71
2.8.2 驗證結論	76
2.8.3 探索未來	77
2.9 理解度的局限.....	79
2.9.1 言外之意的理解度	79
2.9.2 相互理解度的實時監控	80
2.9.3 樣本數量	82
2.10 理解度理論的未來發展方向.....	83
2.10.1 人工智能（Artificial Intelligence）領域	83
2.10.2 大數據化、常模建立、量表化	85
2.11 小結.....	86
第三章 理解度測試與時間參數.....	92
3.1 時間參數的概念.....	92
3.2 添加時間參數的必要性.....	93
3.3 意義和作用.....	95
3.4 改進公式.....	97
3.5 實驗目的.....	99
3.6 實驗驗證.....	100
3.6.1 實驗材料	100
3.6.2 實驗對象	102
3.6.3 實驗流程	104
3.6.4 實驗地點	105
3.6.5 實驗結果	106
3.6.6 時間參數計算	114
3.7 理解度分析與結論.....	121
3.7.1 理解度分析	121
3.7.2 相互理解度及相互理解率	124
3.8 小結.....	129
第四章 粵式普通話理解度研究.....	134

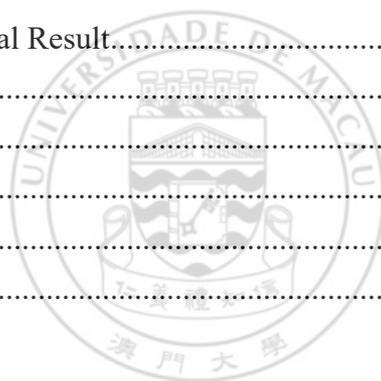
4.1 普通話與普通話水平測試.....	134
4.2 方言普通話.....	138
4.2.1 方言普通話.....	138
4.2.2 粵式普通話.....	146
4.3 實驗目的.....	151
4.4 粵方言與普通話理解度實驗.....	152
4.4.1 實驗材料.....	152
4.4.2 實驗對象.....	154
4.4.3 實驗流程.....	159
4.4.4 實驗地點.....	160
4.4.5 實驗結果.....	160
4.4.6 理解度分析.....	169
4.4.7 相互理解度與相互理解率.....	179
4.5 粵式普通話理解度實驗.....	180
4.5.1 實驗材料.....	180
4.5.2 實驗流程.....	184
4.5.3 實驗對象.....	185
4.5.4 實驗地點.....	189
4.5.5 實驗結果.....	190
4.5.6 分析與討論.....	204
4.6 小結.....	208
第五章 結論.....	211
附錄.....	225
參考文獻.....	237

Table of Contents

Acknowledgement	i
Abstract	v
Declaration	vii
List of Tables and Figures.....	xi
Chapter 1 Introduction	1
1.1 Research Contents And Objectives	1
1.1.1Intelligibility	1
1.1.2Intelligibility Test	3
1.1.3Mutual Intelligibility.....	4
1.1.4Putonghua Study	5
1.2 Research Status.....	7
1.2.1 Overseas Research Status	7
1.2.2 Domestic Research Status.....	8
1.3 Research Methodology	10
1.4 Research Problems	10
1.5 Potential Contributions.....	11
1.5.1 Potential Contributions of Intelligibility.....	11
1.5.2 Potential Contributions of Intelligibility Test.....	13
1.5.3 Potential Contributions of Intelligibility of Putonghua	17
1.6 Statement of Originality	17
Chapter 2 The Concept and Theory of Intelligibility.....	18
2.1The Origin of Intelligibility	18
2.2The Concept of Intelligibility	23
2.2.1 Communication Mode	23
2.2.2The Concept of Intelligibility	28
2.3The Character of Intelligibility.....	35
2.4 Influence Factors of Intelligibility.....	40
2.5The Method of Intelligibility Test	51
2.5.1 Experimental Material	52
2.5.2 Experimental Methodology	53
2.5.3 Experimental Subject.....	57
2.5.4 Test Tool.....	58
2.6Norm-reference of Intelligibility	61
2.6.1Norm-reference and Criterion-Referenced	61
2.6.2Norm-reference of Intelligibility	63

2.7 Application of Intelligibility	65
2.8 The Significant of Intelligibility	73
2.8.1 Establish Standard.....	73
2.8.2 Verify Conclusions	78
2.8.3 Discover future.....	79
2.9 Limitation of Intelligibility	81
2.9.1 Intelligibility of Implication.....	82
2.9.2 Real Time Monitoring of Mutual Intelligibility.....	83
2.9.3 Sample Quantity.....	84
2.10 Future Development	85
2.10.1 Artificial Intelligence	85
2.10.2 Big Data	87
2.11 Brief Summary	88
Chapter 3 The Intelligibility Test and Time Parameter.....	94
3.1 The Concept of Time Parameter.....	94
3.2 Necessity.....	96
3.3 Significant.....	97
3.4 Improvement of Equation.....	99
3.5 Experimental Purpose.....	102
3.6 Experimental Verification	102
3.6.1 Experimental Material	102
3.6.2 Experimental Object.....	104
3.6.3 Experimental Procedure.....	106
3.6.4 Experimental Location.....	108
3.6.5 Experimental Result.....	109
3.6.6 Time Parameter.....	117
3.7 Conclusion of Intelligibility.....	124
3.7.1 Analyses of Intelligibility	124
3.7.2 Mutual Intelligibility	127
3.8 Brief Summary	132
Chapter 4 The Intelligibility of Putonghua with Cantonese Characters.....	138
4.1 Putonghua and PSC	138
4.2 Putonghua with Dialect Characters	142
4.2.1 Putonghua with Dialect Characters.....	143
4.2.2 Putonghua with Cantonese Characters.....	150
4.3 Experimental Purpose.....	156
4.4 The Intelligibility of Putonghua and Cantonese	156

4.4.1 Experimental Material	156
4.4.2 Experimental Subject.....	158
4.4.3 Experimental Procedure.....	164
4.4.4 Experimental Location.....	165
4.4.5 Experimental Result.....	165
4.4.6 Analyses of Intelligibility	175
4.4.7 Mutual Intelligibility.....	185
4.5 The Intelligibility of Putonghua with Cantonese Character	186
4.5.1 Experimental Material	186
4.5.2 Experimental Procedure.....	190
4.5.3 Experimental Subject.....	191
4.5.4 Experimental Location.....	195
4.5.5 Experimental Result.....	196
4.5.6 Discussion.....	211
4.6 Brief Summary	215
Chapter 5 Conclusion.....	218
Appendix.....	232
References.....	245



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

圖表目錄

表格 1 理解度影響因素等級和處理方法.....	49
表格 2 實驗材料類型.....	100
表格 3 被試年齡描述性統計.....	103
表格 4 實驗地點.....	106
表格 5 時間參數計算過程 1.....	118
表格 6 時間參數計算過程 2.....	118
表格 7 時間參數計算過程 3.....	119
表格 8 時間參數計算過程 4.....	120
表格 9 粵方言、湘方言和閩南方言的單向理解度.....	123
表格 10 粵方言、湘方言和閩南方言的相互理解度.....	124
表格 11 粵方言、湘方言和閩南方言的相互理解率.....	127
表格 12 粵方言與普通話實驗材料類型.....	152
表格 13 粵方言與普通話測試被試籍貫統計.....	155
表格 14 粵方言與普通話測試被試第一語言統計.....	156
表格 15 粵方言與普通話測試被試第一語言方言區統計.....	156
表格 16 粵方言與普通話測試被試教育背景統計.....	157
表格 17 語言能力自評.....	158
表格 18 粵方言與普通話測試地點.....	160
表格 19 被試第一語言因素下的正確值和反應時間.....	163
表格 20 反應時長影響因素及影響量.....	165
表格 21 不同組別普通話和粵方言的正確值與反應時長.....	167
表格 22 句子與詞彙的反應時長與正確值.....	168
表格 23 粵式普通話實驗例句信息.....	182
表格 24 粵式普通話實驗被試籍貫信息統計.....	187
表格 25 粵式普通話測試被試第一語言信息統計.....	188
表格 26 粵式普通話官話方言區被試頻率統計.....	188
表格 27 粵式普通話被試教育信息統計.....	189
表格 28 粵式普通話實驗地點.....	189

圖表 1 通信模式.....	23
圖表 2 測試流程示意圖.....	104
圖表 3 粵方言者對粵方言、湘方言與閩南方言三者的正確值與反應時間.....	108
圖表 4 閩方言者對粵方言、湘方言與閩南方言三者的正確值與反應時間.....	110
圖表 5 閩南方言者對粵方言、湘方言與閩南方言三者的正確值與反應時間..	113
圖表 6 教育背景因素下的正確值與反應時間.....	162
圖表 7 被試第一語言因素下的正確值和反應時間.....	163
圖表 8 被試不同組別的正確值和反應時間.....	164
圖表 9 第一語言背景下被試對普通話與粵方言的理解度.....	172
圖表 10 教育背景因素下被試對普通話與粵方言的理解度格局.....	173
圖表 11 不同組別的普通話與粵方言理解度格局.....	173
圖表 12 不同組別的普通話與粵方言詞彙理解度格局.....	175
圖表 13 不同組別的普通話與粵方言句子理解度格局.....	175
圖表 14 澳門非粵方言被試年級分組粵方言綜合理解度.....	177
圖表 15 澳門非粵方言被試年級分組粵方言句子與詞彙理解度格局.....	178
圖表 16 不同教育背景下粵式普通話理解度.....	191
圖表 17 本族語和外族聽話人對粵式普通話的理解度.....	193
圖表 18 口音感評價與理解度.....	195
圖表 19 不同性別的口音感評價.....	195
圖表 20 本族和外族聽話者的口音感評價.....	196
圖表 21 理解度主觀評價.....	197
圖表 22 不同性別的理解度主觀評價.....	198
圖表 23 本族與外族聽話者的主觀理解度評價.....	199
圖表 24 粵式普通話性質主觀評價.....	200
圖表 25 不同性別的普通話性質主觀評價.....	201
圖表 26 本族與外族聽話者粵式普通話性質主觀評價.....	202
圖表 27 理解障礙調查.....	203
圖表 28 不同性別的理解困難點調查.....	203
圖表 29 本族與外族聽話者理解障礙調查.....	204

圖表 30 本族人三種語言變體理解度散點圖.....206

圖表 31 論文章節內在關聯.....211



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

List of Tables and Figures

Table 1	Level of Influencing Factors and Processing Method	49
Table 2	Types of experimental materials	100
Table 3	Age Descriptive Statistics of Subjects	103
Table 4	Laboratory Site.....	106
Table 5	Time parameter calculation process 1.....	118
Table 6	Time parameter calculation process 2.....	118
Table 7	Time parameter calculation process 3.....	119
Table 8	Time parameter calculation process 4.....	120
Table 9	Intelligibility of Cantonese, Xiang and Minnan Dialects	123
Table 10	Mutual Intelligibility of Cantonese, Xiang and Minnan Dialects.....	124
Table 11	Mutual Intelligibility Rate of Cantonese, Xiang and Minnan Dialects ...	127
Table 12	Cantonese Dialect and Putonghua Experimental Materials Types.....	152
Table 13	Statistics of Cantonese Dialect and Putonghua Test Subjects'Origins.....	155
Table 14	First Language Statistics of Cantonese and Putonghua Test Subjects.....	156
Table 15	Cantonese and Putonghua Subjects First Language Area Statistics	156
Table 16	Education background statistics of test subjects.....	157
Table 17	Language Competence Self-assessment	158
Table 18	Laboratory Site.....	160
Table 19	ACC and RT under First Language Factor	163
Table 20	Influencing factors and influencing factors of response time.....	165
Table 21	ACC and RT of Putonghua and Cantonese Dialects in Different Groups	167
Table 22	RT and ACC of Sentences and Vocabulary.....	168
Table 23	Cantonese Putonghua Experimental Example Sentences Information.....	182
Table 24	Statistics of Origin Information of Cantonese Putonghua Test Subjects..	187
Table 25	First Language Information of Cantonese Putonghua Test Subjects.....	188
Table 26	Frequency Statistics of Cantonese Mandarin Dialect Areas.....	188
Table 27	Education Information Statistics of Cantonese Putonghua Subjects	189
Table 28	Laboratory Site.....	189
Figure 1	Communication Mode	23
Figure 2	Test Process Diagram	104
Figure 3	ACC, RT of Cantonese speakers to Cantonese, Xiang & Minnan dialect	108

Figure 4 ACC, RT of Xiang speakers to Cantonese, Xiang & Minnan dialect	110
Figure 5 ACC, RT of Minnan speakers to Cantonese, Xiang & Minnan dialect....	113
Figure 6 ACC and RT under Educational Background Factors.....	162
Figure 7 ACC and RT under First Language Factor.....	163
Figure 8 ACC and RT of subjects in different groups	164
Figure 9 The Intelligibility of Putonghua and Cantonese in the first language	172
Figure 10 The Intelligibility of Putonghua and Cantonese under the Educational Background Factor	173
Figure 11 The Intelligibility of Putonghua and Cantonese in Different Groups.....	173
Figure 12 Vocabulary Intelligibility of Putonghua and Cantonese.....	175
Figure 13 Sentence Intelligibility of Putonghua and Cantonese	175
Figure 14 Intelligibility of Cantonese of Macao non-Cantonese subjects	177
Figure 15 Intelligibility of Cantonese of Macao non-Cantonese subjects	178
Figure 16 Education background of Cantonese and Putonghua test subjects.....	191
Figure 17 Intelligibility of Cantonese Putonghua by native and foreign listeners...	193
Figure 18 Acoustic evaluation and Intelligibility.....	195
Figure 19 Acoustic evaluation of different genders	195
Figure 20 Acoustic evaluation of native and foreign listeners.....	196
Figure 21 Subjective evaluation of Intelligibility	197
Figure 22 Subjective evaluation of Intelligibility in different gender.....	198
Figure 23 Evaluation of subjective Intelligibility of native and foreign listeners...	199
Figure 24 Subjective evaluation of the nature of Cantonese Putonghua.....	200
Figure 25 Subjective evaluation of the nature of Putonghua in different genders..	201
Figure 26 Subjective evaluation of the nature of Cantonese Putonghua.....	202
Figure 27 Intelligibility Obstacles Survey	203
Figure 28 Survey of Intelligibility Difficulties of Different Genders	203
Figure 29 Survey of Intelligibility Barriers of native and foreign listeners	204
Figure 30 Scattered Diagram of Intelligibility of Three Variations.....	206
Figure 31 Intrinsic Relevance of Chapters in Papers	211

第一章 緒論

語言是人與人之間溝通的工具。隨著人類文明的進步和發展，人類對於世界的認識不斷加深，改造世界能力不斷增強，語言也在不斷地變化。人口的遷移將語言又遷移到世界各個角落，不同地區的語言變化速度並不相同，並且變化的層級也並不相同。有些地區的變化大，有些地區的變化小，有些地區依舊保留了很多歷史語言的特徵，有些地區吸納了別的語言的特徵。久而久之，語言和語言之間便不能夠溝通，一種語言的不同方言之間也不能夠溝通，一種方言的南北兩端也不太能夠聽懂對方說的究竟是什麼。使用不同地域方音、不同派系方言甚至不同源流的語言的人們發現，並不能立刻聽懂對方說的是什麼，有一些是立刻完全能夠聽懂的，有一些是可能需要反應一下就大概能夠理解的，更多的是完全聽不懂的「外語」。語言學家也注意到這個問題，完全理解和完全不理解這兩個極端是很容易描述清楚。然而還有大量的事實是介乎於這兩端的，並且很難單純通過文字來描述清楚的，這需要借助定量的方法來進行精細化的描寫，借以此來準確地描述操兩種不同方音、方言或語言的人究竟能夠理解對方多少。本文將會著眼於理解度的理論與實踐，對理解度的理論進行深入建構，對理解度的測試進行深度優化，最後綜合利用理解度理論進行個案分析。

1.1 研究內容與研究目標

1.1.1 理解度

理解度（Intelligibility），又稱可理解度、可懂度，被認為是交際的行為標準（the behavioral standard of communication）（Kent, R.D. 1992）。理解度可以被定義為是說話者的意圖通過傳輸系統與聽話者的反饋之間的匹配程度。

當聽話者所聽到的詞與說話者說出的詞一致，那麼理解度是滿分；如果聽話者所聽到的每一個詞都與說話者說出的不一致，那麼理解度就是零分。在這兩個極端之間，理解度是一個聽話者聽懂多少的程度上的連續統（Schiavetti, N. 1992）。理解度的基礎是建立在語言交流上的，本文所研究的理解度是純粹對語言的理解。誠然在真實的交流過程中，外部的手勢、體態、表情等因素也會給語言增加更豐富的內容，有時候可以增加對語言的理解。但是在本文中，外部的因素會被排除，從而可以更加專注於語言本身所帶來的理解度的變化。

理解度作為一個早已經被學界接受並且廣泛使用的概念，並沒有得到預期使用程度相應的研究。絕大多數的前人研究是直接使用理解度這個概念，而並沒有對其進行深入的討論，也沒有相應的理論建構，這種情況會讓人對理解度這個概念本身產生懷疑。學界更多的是對理解度的應用產出大量的成果，而忽視了理解度理論本身視其為理所當然的存在。但是，往往被忽視的那些理所應當的甚至是看不見的東西卻是最重要的，比如空氣在人的生活中似乎完全不起到任何作用，以至於我們經常忽略了空氣的存在，然而空氣對於每個人來說又是如此的至關重要，沒有了空氣，整個人類社會就是不存在的。所以對理解度本身的理論架構是非常基礎性的和必要的。本文也將會全面深入地研究並完善理解度的理論框架，總結前人對理解度的認識，釐清理解度的相關概念，分析討論理解度概念的外延和內涵，剖析剝離影響理解度的因素，構建理解度使用平台，最終建立語言學視野下的理解度的理論結構框架。

1.1.2 理解度測試

理解度測試顧名思義是對理解度進行的測試是一個人或一類人對一門語言或方言的理解程度進行的測試。理解度的測試的前提是理解度是一個客觀的

值，且這個值可以通過特定手段被測量。目前能夠見到大量的關於理解度測試的文獻成果，前人通過各種手段，控制各種各樣的變量對理解度進行試驗和測試。實驗方法各式各樣，實驗結果更是各不相同，理解度測試似乎也變成了任人打扮的小姑娘，學者為了自己的研究目的，增加或刪減試驗步驟和方法。這樣帶來的結果是每一篇文章的實驗結果都是孤立的，因為研究者所採用的方法不盡相同，這就導致無法就其實驗結果進行數量上的比對，而只能夠在結論層面進行比較。這顯然是缺乏足夠的說服力的，同時也造成了大量學術資源的浪費，孤立的結論並不能放到整個學術網絡中進行比較和聯繫，也是一件非常可惜的事情。誠然研究方法是為研究的目的服務的，但是研究也應該有固定的研究範式和研究標準。學者可以根據自己的需要進行微調，而不是進行結構化的改變。

本文所研究的語言間相互理解度測試，其測試對象並不拘泥於兩種語言之間，或是兩種方言之間，抑或是兩種語言變體之間。以上這三者在語言學視角下，僅存在量的差異，並沒有本質性的差異，測試不同的語言並不會改變相互理解度測試的測試範式，換句話說相互理解度測試的測試範式可以包容各種類型的語言或語言變體。其中的個性化差異體現在具體的實驗材料的選擇上。實驗範式的確定就能夠建立起一個科學化的研究標準和一個研究平台，任何兩種語言都可以進行理解度的測試，並且因其使用同一範式，即可以與其他已有研究結果進行比較，從而保證了學科發展的延續性，同時也能夠提供更多科學化的數據。本文認為應該制定出一套理解度測試的標準，使其結果能夠更加標準化和科學化。標準的確立不僅有利於得到更準確科學的結果，還有助於將理解度測試進一步應用於更廣闊的範圍，取得更多更有效的成果。這就類似於一個

電腦的運算能力是有限的，但是把眾多電腦的運算能力通過互聯網的作用整合起來，那麼其運算能力並不是簡單的相加的結果，則是指數型增長的結果。與此同時，對理解度測試標準的制定也是理解度理論本身的重要一環，標準的制定是對理解度理論的豐富和完善，將標準的制定納入到理論研究的範疇中，可以使該標準得到充分的重視、討論和研究。因此本文將會在原有可理解測試的基礎上對可理解測試進行優化改進，使其能夠更好地為語言學服務，並且最終制定出一套標準化的測試範式。

1.1.3 相互理解度測試

因為理解度測試是單向的，實驗者是作為聽話人的角色出現在實驗中。真實的對話場景並不是一個人長時間持續充當聽話人，而是兩人或多人互相輪替承擔說話人和聽話人的角色。所以單向的理解度測試並不能夠很好地呈現出交互的情景。而相互理解度測試能夠更好地彌補這樣的一個遺憾，相互理解度的測試可以是雙向的甚至可以是多向的。但是鑒於目前技術條件的限制，並不能非常有效地實時測試出兩個人或多個人交談中的每個人的理解度，因此相互理解度的測試是在理解度單向測試的基礎上將雙向或多向拆分成兩個單項或多個單項進行測試。利用技術手段將同一時空下的對話拆分成不同時空下的測試題目，分別進行測量，最後再對兩個實驗結果進行整合。

相互理解度測試是建立在單向的理解度測試的基礎上，是通過理解度測試的累積和疊加得出的結果。本文會在充分論證理解度測試的基礎上，對相互理解度測試進行實踐測試，利用兩次理解度測試來達到相互理解度測試的效果。並且通過實踐也可以反觀理解度理論的合理性和理解度測試的可操作性，從而證實理論假設同時豐富理論建構。

1.1.4 普通話研究

普通話作為中華人民共和國通用語，是以北方話為基礎方言，以北京音為標準音，以典範的現代白話文作為語法規範。自從1955年「全國文字改革會議」和「現代漢語規範研討會」確定普通話的明確定義以來，普通話已經隨著大眾媒體和政府的大力推廣之下進入了人們生活的方方面面。普通話已經成為中國人語言生活中必不可少的一部分。普通話推廣的進行達到了一個前所未有的深度和廣度。從深度上看，很多城市中的夫妻已經不再用方言作為家庭語言，因而孩子的第一語言便直接是普通話。越來越多的孩子失去了方言背景，在教育語言中也是直接以普通話作為教學語言，在學校範圍內杜絕方言的存在。從廣度上看，國家語言文字事業「十三五」發展規劃（教育部、國家語委2016：3）「十二五」期間，特別是黨的十八大以來，國家通用語言文字普及程度進一步提高，普通話普及率達到70%以上。也就是說，目前中國已經有將近10億人能夠使用普通話。對於普通話的讀音標準問題，普通話的標準音應以《現代漢語詞典》中的注音作為標準讀音。普通話等級考試的最終參考答案也是依照《現代漢語詞典》中的注音進行裁定。然而真正能夠達到標準的人寥寥無幾。也就是說絕大多數人是不能夠達到所謂的標準的，也就是不能夠標準地講出他們的國家通用語。這將極大地傷害全國人民對於普通話的情感與認同。

方言在我國的語言生活中同樣扮演著不可或缺的作用。每個方言都承載著當地獨特的歷史人文信息，是語言研究的重要資源。方言的影響力雖然在普通話的推廣過程中被削弱了，但是在一些強勢方言區的方言，其使用頻率還相當高。例如港澳地區很多家庭往往會傾向於選擇用方言來作為家庭語言，讓孩子先學習方言。在這樣語言背景成長起來的孩子，在之後接受普通話教育時，會

多多少少帶有一些方言的特徵，於是就形成了所謂的地方普通話。

地方普通話的形成與普通話的推廣有著密切的關係。隨著普通話的推廣，各地的地方普通話也隨之產生，進入到語言學研究的視野中，學術界也有越來越多的學者關注到這個問題，並對其展開研究與討論。

本文關於普通話的觀點是採納自徐傑、董思聰（2013）中的觀點，將普通話的標準放開，不以北京音為標準音，而是以北京音為基礎音的普通話，即所謂的「寬式普通話」。在這一標準下，將會吸納更多的語言材料進入研究視野，並且能夠將「體系性」、「穩定性」、「可理解性」作為寬式普通話的一項重要標準。在此定義下的普通話更有生命力，更鮮活也更加自然真實。追究普通話的命名之初就在於「普遍通行」，因此本文也會將研究材料選擇的重點放在普遍通行的普通話上。從概念的外延上看，它包括涵蓋了前人研究中的地方普通話和動態普通話等概念範疇。因為從這種寬式普通話是有方言背景的人在向普通話靠攏的過程中產生的，所以從主觀意願上來看，他們是在盡可能講普通話，而不是故意要講有地方方言特色的普通話。雖然其客觀結果來看，講話人並不能夠一下子說到播音員的水平，做到完全沒有錯誤，但從其心理感知上可知，他們已經盡可能放棄方言的特徵，而轉向使用普通話的特徵。所以我們有理由把這種帶有地方特色的普通話納入到普通話的範疇中，對其進行測試和研究。本文將會選取帶有粵方言特色的普通話進行理解度測試，帶有粵方言特色的普通話也就是通常所謂的「廣普」或「粵普」，當然其中還包括很有代表性的「港普」和「澳普」。這些普通話在細節處存在或多或少的差異，例如「港普」中通常會夾雜大量的英文或借詞，但基本上他們帶有同樣的特徵。通過對其進行測試來瞭解這種普通話在其他地區的理解度，並再與標準的普通話

和標準的粵方言的理解度進行比較，研究三者之間的異同和內在規律。

1.2 研究現狀

1.2.1 國際研究現狀

國際對理解度的研究最早可以追溯到1920年代，出現在通信領域和病理學領域。隨著電子錄音技術等系統的進步，研究者可以對音節、詞、句子等內容進行分析，從而能夠得到真正的聽懂度的結果（Weismer, G. 2008）。時至今日這個問題依舊在病理學領域有新的成果出現（Fuse, A., Navichkova, Y., & Alloggio, K. 2017）。二十世紀五十年代之後，理解度研究慢慢從醫學轉入人文學科，較早的關於理解度的論文是Bansal（1969）及Tiffen（1974）分別研究印度英語口音和尼日利亞英語口音的理解度問題，隨後Smith, L. E., & Rafiqzad, K.（1979）就談到跨文化交流中理解度是一個非常重要的不可回避的問題。之後理解度研究才逐步擴大到語言學研究領域，其研究主要集中於對帶有口音的英語進行理解度的測試，例如：Brennan, E. M., & Brennan, J. S.（1981），Wardrip-Fruin, C., & Constantinou, A.（1993），Rogers, C.L.（1997），Yang, H. T.（2005），Pickering, L.（2006），Hardman, J.（2010），Hayes-Harb, R., & Watzinger-Tharp, J.（2012），Nejjari, W., Gerritsen, M., Van der Haagen, M., & Korzilius, H.（2012），Huong, T. T.（2017）。他們的實驗對象多是不同地區的帶有口音的說話人，把他們的語音錄下來，讓英語母語者進行判斷。其中的被試被分成兩類，一類是以英語為母語，但他們生活的地方並不是在英國或美國本土；另外還有一類是以英語為第二語言，有的研究是越南的英語、中國的英語以及帶有西班牙、墨西哥等地口音的英語等等，例如：Munro, M. J., & Derwing, T. M.（1995），Munro, M. J., Derwing, T. M., & Morton, S. L.

(2006) 等。Nelson, C. L. (2012) 對世界眾多口音英語的理解度進行了較為詳細的理論和實踐研究，是目前較為系統的關於英語世界的理解度研究。另外，在國際研究中，也存在著如何區分語言和方言的問題，較為常見且普遍接受的做法就是利用理解度進行測試，其中有Voegelin, C. F., & Harris, Z. S. (1951), Comrie, B. (2009), Melinger, A. (2018)。

除此之外，還有將理解度用於手語理解的問題中，如：Sáfar, A., Meurant, L., Haesenne, T., Nauta, E., De Weerd, D., & Ormel, E. (2015)。這也是為理解度的應用做出非常重要的嘗試，大膽地擴寬了理解度的應用範圍。將對有聲語言的理解擴展到手勢語的理解。

1.2.2 國內研究現狀

國內對於語言理解度的研究起步比較晚。但學界早已注意到方言之間的溝通在程度上是有差異的，於是在二十世紀八十年代末到九十年代初期，學者採用了計算語言距離的方式來解釋方言之間溝通差異的問題，最有代表性的是陸致極 (1987)，鄭錦全 (1988)，鄭錦全 (1994)，陳海倫 (1996)，游汝傑、楊蓓 (1998)。語言距離的研究固然能夠解釋一定的差異現象，但是理解度是動態的，而非靜態的音系差異就可以充分說明的。進入21世紀後，更多的學者注意到語言理解度的研究模式，也逐漸使用到鑒定方言間親疏關係以及語言距離的研究領域中。其中有林素娥、蘇憲萍 (2007)，Tang, C., & van Heuven, V. J. J. P. (2008)，Tang, C., & Van Heuven, V. J. (2009)，石林 (2010)，黃行 (2011)，王璐、張吉生 (2014)，王璐 (2014) 等。其中林素娥、蘇憲萍 (2007) 對方言理解度測試的測試單位進行了較為深入的探討，結合了認知規律闡述了方言可懂度所應該注意的問題，對如何開展計量研究進

行了初步構思，用以指導方言可懂度計量研究的具體調查和統計分析。

除此之外，在其他研究領域也逐漸採用了理解度的研究方法。將英語設為研究對象，以英語作為第二外語研究中國大學生帶有漢語口音的英語的理解度。其中有何家勇、李珊、何雲亭（2010），莊小榮（2011），張伶俐（2012），高超、文秋芳（2012），李景娜、王遙（2015）等。將漢語作為研究對象，以漢語作為第二外語研究外國留學生或外族學生學習漢語時的理解度問題：高潔、姚勇（2013），張林軍（2015a）等。還有魏梅（2017）對中國英文報刊中本土化英語的可理解度研究，不過解決的是閱讀文本過程中的理解度問題，與本文所論述的理解度不同。

與此同時關於理解度的應用研究還涉及到醫學層面，主要針對聽障人群的診斷與康復方面，其中包括袁海軍（1994），陳雪清、程佳佳、劉博等（2012），韓軍寧、陳雪清、郭倩倩等（2017）。與理解度相關的研究還存在於信號學領域，研究信號傳遞時的信息衰減、噪音干預等問題，其研究成果豐富，並且已經形成相對完整的理論與實踐體系。這些研究成果也可以作為研究理解度的重要參考資料，但與本文所涉及領域差別略大，暫不做深入研究。

1.3 研究方法

本文採用理論推導和實踐驗證兩種方法對理解度理論進行分析和討論。在理論方面，首先對理解度理論進行深入細緻的探討。通過先行文獻和前人研究來找到理解度理論的學術位置，明確理論的邊界，確定理論的職能。之後再對理論的內部進行細緻論述。在實踐方面，根據理論的指導和前人實驗模式制定出改進版的科學實驗流程，結合本文所要論述的內容定制相關的實驗流程和

材料，分步驟分批次進行實驗，並收集實驗數據。最後使用統計軟件SPSS將實驗數據進行統計學運算，從而得出在統計學意義下有意義的結論。最後再反觀實驗中的問題，對理論進行適當的修正和補充。

1.4 研究問題

本文主要研究三個核心問題，分別是建立理解度理論，對現有理解度理論進行改進並進行實驗驗證，最後是對粵式普通話進行個案分析。每個核心問題裡面又會有若干個小問題需要論述：

1、建立理解度理論

- (1) 理解度的源流
- (2) 理解度的概念
- (3) 理解度的影響因素
- (4) 理解度的應用和意義
- (5) 理解度理論局限與發展
- (6) 理解度理論及其相關問題

2、對現有理解度測試進行改進

- (1) 為何引入時間參數
- (2) 如何引入時間參數
- (3) 引入時間參數的意義

3、個案分析1：粵方言、湘方言及閩南方言之間的相互理解度研究

- (1) 三者之間的相互理解度與相互理解率



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

(2) 與前人研究之間的關聯和比較

(3) 理解度數值的內涵與意義

4、個案分析2:粵式普通話的理解度研究

(1) 普通話、方言與方言普通話的關係

(2) 普通話、粵方言和粵式普通話相應的理解度

(3) 引入理解度測試對地方普通話研究的意義

1.5 潛在研究意義

1.5.1 理解度理論的研究意義

理解度概念的應用已經非常普及，在醫學、信號學等眾多範疇內都有涉及，且已經形成一套理論和一定的行業規範與固定範式。語言學視野下的理解度概念也已經被學術界熟知且廣泛應用了。是因為理解度對於語言來說，是一個基礎性的屬性，語言的使用的目的就是要讓聽話者聽懂、理解，而為了滿足聽懂、理解的結果，其中涉及到的因素卻非常多，既涉及到語言本體的語音、詞彙、語法的方面，又涉及到語言外部的社會屬性。可見理解度對於語言學來說是一個非常重要的一環。然而，對理解度理論的研究尚不足以匹配得上其在語言學中的重要意義，可以說對理解度理論的深度挖掘還是一片處女地，尚沒有得到很好的探討，也還沒有建立起足夠堅實的理論框架。絕大多數情況下，研究者忽略了對其理論的探討而更多關注於實際應用與操作。這樣的結果往往是陷入了一種瓶頸，不能夠繼續深入發現問題，而是重復在一個層次上做量的累積而不是質的突破。並且現有的理解度的理論最早都源於西方，國內對於理解度的關注起步晚、關注低、範圍窄。雖然理解度的提法常常可以聽到，現實是漢語語言學界還沒有正式的聚焦於理解度的研究，可以說幾乎還是一片空

白。然而漢語擁有眾多方言資源、少數民族語言資源，這就亟需一個完善的理論體系將其整合起來並建立漢語在理解度領域的話語權。

本文將會是第一次在漢語語言學視野內對理解度理論進行深入系統地探討，用以彌補漢語語言學界對這一問題的理論空白。本文會在前半部分系統論述理解度的理論，包括概念的內涵和外延，理論背景，影響因素以及相關理論問題，試圖建立起一個較為完善的理論框架和學術體系，有利於日後在該理論的基礎上充分討論升級改進。建立起系統完善的理論架構就是建立起一個理論的平台，可以更好的與其他理論在同一水平上對話。創造出形式化體系化的理論接口，更好地為理解度理論在知識網絡中找到屬於自己的節點和定位。本文還會充分考慮到漢語的特點對理解度理論進行修改，建立起既符合語言一般性的理論框架，又適用於個別語言的理論細節，可以讓理解度理論容納更多的語言事實，充實豐滿自身的理論，又能夠使其最大程度地與應用相結合。以材料支撐理論，以理論指導實踐，以實踐補充材料。

語言本體的研究似乎與語言應用研究之間有一條不可跨越的鴻溝，本體追求理論的完美和精簡，而應用更多的著重於語言的細節的特殊性，理解度理論就是介於語言本體與應用之間的一道重要橋梁。影響理解度的原因是多元的，是複雜的，是綜合的，因此它將本體中語音、詞彙、語法的各個層面的特徵與語言應用聯繫起來，同時還綜合考慮到了社會、心理等多方面因素。理論的建設是為應用理論打下堅實的基礎，用理論去指導實踐能夠在實踐的過程中發現更多有意義、有價值的問題。反過來，通過實踐的實驗結果反觀理論的猜想也可以促使理論的更新與進步。它為本體研究提供了一塊試驗田，讓本體研究的種子可以得到生根發芽的沃土；同時又為語言應用找到了根基，讓語言應用

可以汲取更多的養分。

1.5.2 理解度測試的研究意義

理解度的測試是一個綜合利用了語言學、心理學、語言測試理論、統計學等幾大學科作為理論背景的測試，是一個有著交叉學科背景的產物。這就要求實驗設計者在充分掌握語言本體事實的基礎上，瞭解其他語言學科的理論基礎和框架，熟悉一般實驗的相關流程和對數據的統計，才能夠得到滿意的結果。其結果也可以為以上幾個學科領域甚至更廣泛的領域做出貢獻。以下列舉幾個理解度測試在語言學理論和實踐層面上的意義：

1、為判定方音、方言、語言提供了客觀可信的數據基礎。使用聽懂度作為判斷兩種語言變體是否是一種語言的兩種方言還是兩種語言，這最早由 Comrie, B. (1987) 提出。以往對兩種語音系統之間關係的界定究竟是方音、方言還是語言經常會有爭議。然而理解度的測試數據能夠給出可驗證的數據。前人研究從語言距離的角度入手，通過分析音系材料或共有詞彙從而發現兩個語言系統之間的差異，最終得到二者之間的距離。然而這種距離是從一份材料到另一份材料的距離，是一套系統與另一套系統的差異，它是影響理解度的一個因素，但絕對不是理解度的實際數值，因為它並不是一個講話人到另一個聽話人之間的距離。文獻中總結的語言系統是規律的，也是靜態的；但實際中使用的言語卻是動態的、複雜的。動態的、複雜的言語研究相較於規律的靜態的語言研究多了很多不可控的個體差異甚至未知的因素，可它並不是一個不能夠被研究的區域。理解度測試不僅剛好為研究這個問題帶來瞭解決的方法，還將這個問題引入了一個更深入的層次：第一，語言的理解是一個複雜的過程，只考慮材料是不夠的，更要考慮到處理材料的處理器——大腦。大腦對語言的處

理功能往往超出了我們目前的認知水平，對語言系統的聯想、整合甚至再學習的能力十分驚人，因此研究者可以借由理解度測試窺探一些大腦對語言認知理解的過程。第二，學界通常認為方音、方言、語言這三者可能會構成一個由特徵構成的連續統，從方音到方言再到語言，特徵的增減導致理解度的降低或升高，量的累積導致了質的變化，當特徵差異達到一定標準後，兩個系統就被區分為兩種方音、方言或語言。這裡涉及到兩個問題，其一是這個衡量的尺子是什麼，其二是尺子上的刻度應該如何得到。如果僅僅使用材料來作為尺子，那麼會忽略掉大量的外部因素，包括語用因素、社會屬性等等。若從影響溝通度的因素角度出發，那麼可能將是一個巨大的挑戰，因為既想要同時控制語言內部的語音、詞彙、語法變量，又要控制語言外部的應用和社會屬性，這幾乎是一個不可能完成的任務。從這個思維中跳出來去思考，為了測量理解度，就直接對理解度進行測量，暫時不管理解度是如何被影響的。理解度就比如是一棵樹上的果子，以往的研究希望通過對樹的粗細、大小、枝葉等進行測量從而得出果子的大小，而忽略了光照、肥料、通風等外部因素。這些因素固然或直接或間接地影響了果子的大小，但是與其顧及到所有因素，倒不如直接把果子摘下來，用尺子量一下就知道了實際果子的大小。果子的大與小究竟是受到了樹本身因素的影響還是外部因素的影響，則是下一步需要研究的內容。直接測量果子大小的尺子就是理解度測試。尺子定下來之後，還需要尋找特定的方法來測量進而確定尺子的具體刻度。對於國際公制的標準制定是採取一個自然的物理常數然後再對其進行等分或翻倍。例如：一攝氏度指在1標準大氣壓下，純淨的冰水混合物的溫度為0度，水的沸點為100度，其間平均分為100份，每一等份為1度，記作 1°C 。一米是被定義為光在真空中 $1/299792458$ s（秒）的時間內

所通過的距離。一千克是1立方分米的純水在4°C時的質量。同理，理解度就可以通過理解的兩端進行分析，確定完全理解和完全不理解的參數之後，再就其內部進行深入劃分。雖然前人在研究中已經試圖給出一些標準，本文認為以目前的技術手段，暫時不能夠一步到位，又找到尺子，又划出刻度。不過，理解度測試給這個問題提出了一個正確的方向和可行的解決途徑。

2、理解度測試能夠通過實驗手段測量出理解度的具體數值，為理解度理論找到事實的落腳點。最早的理解度測試並不是服務於語言學研究，因此有一些實驗環節、實驗方法、數據統計需要加以完善和改進，使其更符合語言學精細化的研究需要。尤其是當語言和語言本身差異並不大的時候，對實驗精度的要求就會很高，故在以往的實驗中忽略的技術參數和細節需要在如今的實驗中體現出來，並且計入到最後的統計計算中。這些都為標準化、模式化的測驗模式提供基礎條件。理解度測試的適用面極大，因此設計一個基礎的、通用的理解度測試模型是必需的，也是必要的。這樣做一方面能夠保證單個測試結果的有效性，同時模型的可複製性就保證了同樣使用這個測試模型的理解度試驗都可以科學有效。如此，可理解度測試就構成了一個比較的平台，在同一平台上可以進行數字與數字之間的比較。兩顆果子並不能放到同一時空下進行比較，那麼用同一把尺子對兩顆果子進行分別測量，就可以利用兩次測量的數字進行比較，結果一目瞭然。另一方面更重要的是，從共識的角度去看，隨著測試結果越來越多，測試和測試之間的比較越來越多，就可以發現其中的規律和法則。這有利於研究者更好地認識語言之間的親疏關係及其譜系，甚至有助於描繪語言的演變和發展軌跡。從歷時的角度再去思考，隨著時間的推移，可能語言和語言之間的理解度越來越大，或者越來越小，這也反映了語言接觸、語言

變異等理論現象。

時間參數的加入首先是為了語言學研究的需要，同時也是將大腦心理處理過程納入到理解度的理論中。因為本文認為，理解度應該考慮到真實對話場景的要素，如果真實對話時，對語言材料的處理時間變長就會增加其理解的時間，理解時間的延長勢必會導致理解度的下降，因為對方講話者的語速如果超過了聽話者的反應速度，那麼別人講到下面的句子的時候，上面的句子還在處理中，這樣只會越來越跟不上從而無法理解。所以理解度測試中也一定要考慮到時間參數。看似僅僅只是增加了一個時間參數，但其意義卻是十分深遠。這讓理解度測試進入了一個更深入的層次，可以探尋到大腦處理語言的過程。這一個層次的深入，可以進一步擴展理解度測試與其他重要領域的接口，可以與心理語言學和神經語言學相聯繫，共同探討在反應時間內大腦的神經運轉機制等內容，將理解度測試推向語言學研究的前沿。

1.5.3 普通話理解度的研究意義

關於普通話的研究也已經由來已久，爭論的焦點一直集中在如何處理普通話與方言，普通話與地方普通話之間的關係上。尤其是後者，普通話與大量廣泛存在的地方普通話的界限應該定在什麼位置，一直是學界爭論的焦點問題。理解度測試給普通話的重新劃界提供了新的途徑。普通話的概念本源就是「普遍通行的話」。若換一種說法來解釋「普遍通行」就是「普遍理解」或者是「普遍可懂」。對於普遍理解和普遍可懂的普遍性最終還是要落腳到「度」的問題上。究竟聽懂多少就可以算作是普遍理解、普遍可懂；究竟理解度達到一個什麼樣的程度後，方言口音的特徵就不會對理解構成影響，這最終還是要依賴於理解度測試提供有效的證據。本文會利用理解度測試的方法首次嘗試用定

量的方式對該問題進行解答，讓普通話的討論不只停留在理論層面，也能直接和實驗數據相結合，為理論提供更好的技術和數據支持。

1.6 原創聲明

本論文，除了經清楚列明來源出處的資料外，其他內容均為本人原創；本論文的全部數據並未曾在同一學位或其他學位中提交過。論文中之觀點除非明確註明引用外，均為本人原創。



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

第二章 可理解度概念及相關理論

2.1 理解度的源流

幾十年來，理解度測試被應用於各個領域中。國外對理解度的研究最早可以追溯到1920年代。隨著電子錄音技術等系統的進步，研究者可以對音節、詞、句子等內容進行分析，從而能夠得到真正的理解度的結果（Weismer, G. 2008）。Miller, G. A., Heise, G. A., & Lichten, W. (1951) 的《可理解度測試中測試材料的語境功能》（The Intelligibility of Speech as a Function of the Context of the Test Materials）。文章開篇第一句就說明，在很長的時間中，通信工程師們有一個心理學的測試方法叫做清晰度測試（Articulation Test），其文中所介紹和使用的測試方法與基本結構都與現在的方法大致相同。請一個發音人向一組被試朗讀音節、詞或句子的列表，被試聽過之後來報告他們所聽到的內容，清晰度的數據就是根據他們所報告的內容的正確率的比值進行計算。文章中區分了聽到（Audible）和理解（Intelligible）之間的區別，強調對理解度的測量中要清楚明白測試材料的意思。文章還強調背景對於理解的重要作用，通過改變背景信息的性質和數量，從而研究其對理解度正確性的影響。

清晰度測試作為理解度測試的前身，早在上個世紀上半葉就已經對清晰度測試有過相關論述Fletcher, H., and Steinberg, J.C. (1929)，Egan, J.P (1942, 1948)，清晰度測試專門應用於通信領域中，研究在通信領域中不同的通信條件、節制等諸多因素對信息傳播效果的影響大小。從方法論的角度上看，理解度測試的方法基本繼承自清晰度的測試。但是其實驗目的卻和理解度有著很大的差異，理解度更強調感知，而清晰度更注重於信號和噪聲之間的關係。兩者互有關係，卻處於不同的層面。理解度的測試早期確實受到了清晰度測試的影

響。

由此可見，清晰度測試確是理解度測試（Intelligibility Test）的前身。而之後清晰度測試的術語和理解度測試的術語之間有著重合現象，隨著時間的演進，理解度測試的術語慢慢佔據了主流甚至替代了清晰度測試成為主要學術術語。在概念外延和學術範疇上，二者逐漸各自分化成兩個不同領域的支線學科，其研究方法和研究程式也有較大的差異。二者之間的術語外延部分略有混亂，後文中將會就該問題進行詳細討論並加以區分。

理解度測試的其中一個應用方向集中在病理學領域，主要處理言語混亂（Speech Disorder）的問題。Kent, Raymond D. (1992) 認為言語理解度是言語病理學中的首要問題。理解度測試的興起得益於傳信工具的升級，尤其是電話的出現（Fletcher, 1953）。在病理學領域，理解度是作為一條判斷言語混亂的標準（Metz, Schiavetti, and Sittler, 1980）。理解度測試可以很好地解釋言語混亂的原因，並且對病人的發音缺陷進行分析。同時，理解度測試也應用於聾兒的語言康復方面，主要包括檢測言語治療的學習過程，比較言語訓練的方法，以及評價主流的被試（Kent, Raymond D., 1992）。

在這一階段，理解度測試主要停留在應用的層面，通過測試來找到理解度的使用意義和應用價值。最初的理解度主要針對於怎樣診斷、評價和治療語言障礙的方面，雖然處理的是語言材料，但是並沒有對語言本身進行分析和研究，而是通過對說話人進行評估和分析，以便更好地診斷病人的病情，檢測病人的恢復過程等。並且，理解度測試是一種非常有用的測試，是因為測試的結果很容易可以與講話人（病人）的家屬和其他同行交流（Beukelman，

1981)。對語言材料本身的關注也是更多地集中在語音層面，具體的音素對於發音的困難以及對具體音素的矯正和治療方面。

目前所見的最早直接論述理解度的文獻是Catford, J. C. (1950)的《理解度》(Intelligibility)這篇文章。文章開篇指出，發音教學的主要目的是滿足理解度的需求，文章還指出講話人和聽話人在理解度中各自的關係。之後Bansal (1969)的博士論文《印度英語口音的可理解性》(The Intelligibility of Indian English: Measurements of the Intelligibility of Connected Speech, and Sentence and Word Material, Presented to Listeners of Different Nationalities)問世後，開啓了將理解度引入對英語口音的研究中，本文認為從此理解度進入語言學的研究範圍內，理解度理論和研究範式開始服務於語言學研究。主要原因有：

(1) 理解度的研究主要目的不再是為了其他學科，而是直接針對語言學的研究，語言的理解度作為研究的終極目標而不是研究的手段和方法，研究的出發點從生理現象轉移到了心理現象；

(2) 實驗材料是基於語言材料的特點和特性進行篩選，以特定語言特點或語言現象為中心選擇變量，採用不同語言背景的健康實驗對象，而非以語言障礙人士作為主要實驗對象；

(3) 研究的成果是為了語言學理論服務，新的研究成果主要解決語言學關心的問題，雖然其結果也可應用於其他學科，其第一目的已經轉入語言學研究框架內；

(4) 研究者更多地來自語言學背景。

時至今日，對英語口音的理解度的研究仍在進行中，而且佔據了理解度研究的重要陣地。隨著研究範圍的擴大，其他語言口音的理解度研究也陸續也得到陸續開展。隨著研究內容的深入，語音系統中各個層次的因素對語言理解度的影響認識不斷加深。除了以口音作為理解度的主要變量加以研究之外，使用理解度對語言和方言進行區分也是另一塊研究重鎮。

Voegelin, C. F., & Harris, Z. S. (1951) 對如何確定語言間的理解度進行了深入地討論。文章認為在當時那個時期，將會有越來越多的機會和條件去進行實驗和田野調查。為此作者認為有必要來總結一下對此類問題的研究方法，作者總結了四種研究方法，分別是（1）詢問調查法（the 「ask the information」 method）；（2）計算相同法（the 「count sameness」 method）；（3）解析結構法（the 「structure status」 method）；（4）測試信息法（the 「test the information」 method）。作者認為可以通過這四種方法來得到語言之間的理解度，再根據理解度去確定語言與語言之間的關係。這四類方法各自擁有其適用方面，也有其局限性。仍需肯定的是，這四類方法為之後的研究奠定了非常重要的方法論基礎。漢語語言學界曾經將方法（2）應用於計算漢語中各個方言之間距離，例如：陸致極（1987），鄭錦全（1988、1994），陳海倫（1996）等。該方法對理解度理論在語言學的研究有著重要的意義，因為世界上究竟有多少種語言一直都是語言學的基本問題之一，也是眾說紛紜、尚未能達成共識的問題。其原因可能是多樣的，可能受條件、人員、材料等多方面的制約，本文認為之所以沒有能夠達成共識，給出一份滿意的答案，是因為從根本上說，還缺乏一個完整的理論體系，缺少一個通用的方法，所以導致了學者們陷入各自為戰的境地。

理解度理論一直被提及，卻很少有學者對其進行詳細專門的論述。理解度測試的前身是清晰度測試。清晰度測試是作為通信領域長期關注、使用且研究的課題之一，後又轉入到病理學範疇。直至20世紀50年代，才回歸語言學領域內。時至今日，對於理解度的研究仍再繼續，然而依舊有一些基礎性問題沒有得到很好的解答，需要分析和論證。例如：

(1) 對於理解度理論本身的研究尚且不夠，研究者在做理解度測試的過程中對理解度理論只是簡單帶過，因此理解度理論本身還有很多結構性內容亟需補充。

(2) 理解度理論與理解度測試的概念是不同的，首先要明確理解度的理論研究和理解度測試的應用的界限，二者之間是有區別的。理解度理論是基礎，理解度測試是方法，理解度測試也只是探尋語言間理解度的其中一種方式。理解度測試從哪些方面可以補充理解度理論，理解度理論又可以從哪些方面指導理解度測試。

(3) 理解度測試與清晰度測試之間的異同，清晰度測試已經形成了非常完善的測試程式，並且已經制定出相應的國家標準。理解度測試是否能夠在比對二者異同後，找出清晰度測試可供學習，改進的地方，最終也形成自己的標準。

本文會在前人研究的基礎上，在後文中對以上這些問題進行詳盡論述，試圖推進理解度理論研究和理解度測試研究，在理論上有所突破和發展。

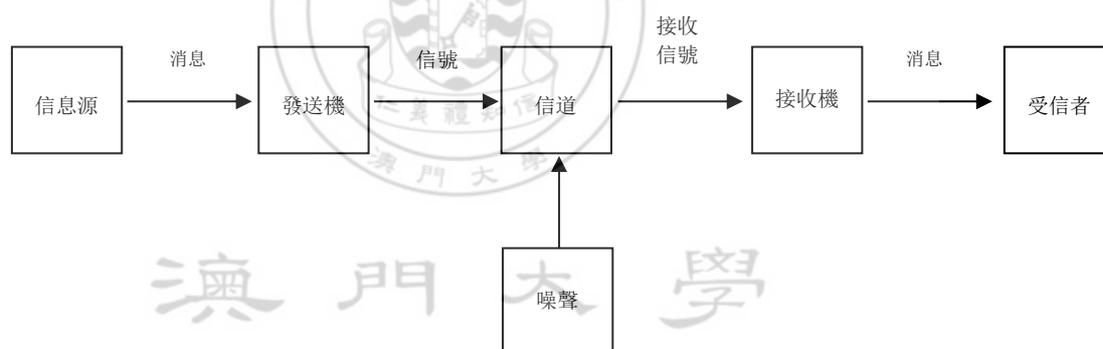
2.2 理解度的概念

本節將會在語言學範疇內為理解度進行定義。在對理解度概念進行界定之

前，首先需要明確理解度指的是語言的理解。Beukelman & Yorkston (1979) 注意到理解度的得分與信息傳遞之間有非常強的關係，語言的理解度測試可以提供一個有效的關於交際表現的指數。從信息學的角度上看，語言也是信息的一種，因此想要對理解度有一個清楚的理解之前，可以先回顧信息交換的基本結構，從中找到理解度的定位，從而能夠更好地認識和理解理解度的定義，也可以更好地對其他概念進行區分。

2.2.1 通信模式

理解度理論的研究中，Kent, Raymond D. (1992) 認為語言理解度的測量就是對說話者、傳信系統、聽話者三者之間交互結果的測量。無論理解度測試



的目的是什麼，講話者、傳信系統、聽話者三者之間的作用都要被考慮進去。三分法的信息傳遞結構是一種較為直觀樸素的結構，其實在信息科學領域，克勞德·申農 (Claude Elwood Shannon, 1916—2001) 早在《通信的數學理論 (A mathematical theory of communication)》(1948) 就提出了五分的通信系統的基本模式 (見圖1)。

圖表 1 通信模式

該信息結構模型能夠更清晰地說明信息傳輸結構內部的過程，而且使用這種通信結構也可以更好地與通信學領域相結合，找到理解度理論與通信學領域之間的聯繫。

1、信息源 (Source) —— 它產生將要傳輸給接收端的消息 (message) 或消息序列，也是信息 (information) 的來源，可以是人、動物或其他機器。信息可以有各種類型，可以是電報系統中的字母，可以是電磁波，也可以是語言。在語言學的角度中，大腦就是交際系統中的信息源，此時的消息還是抽象的，還沒有物質形式。

2、發送機 (Transmitter) —— 採用某種方法把消息轉換為可以傳送的信號 (signal)。信號的種類也是多樣的，可以是光信號、電信號或聲信號。

3、信道 (Channel) —— 它是發送機到接收機之間傳遞信號的通道。信道可以是導線、電纜、空氣、文字等等。在傳輸的過程中可能會被噪聲干擾，因此信道和信道之間是交叉的。

4、噪聲 (Noise) —— 它是對信息的干擾，發生在信息傳輸的階段。噪聲的出現勢必會影響信號的傳輸，接收機接受到的信號其實是原始信號加上噪音後的信號。噪音也可以被看成是一個信息源通過發送機發出的信號，只不過這個信號並不是原本的信息源希望發出的。噪聲也分為系統內部的噪聲和系統外部的噪聲。系統內部的噪聲，比如在無線電傳輸的過程中傳輸信號的衰減，或者是在電線中，由於電線的電阻導致信號的衰減。系統外部的噪聲則是在傳輸過程中受到了如電磁信號或聲光信號等信號，這種信號對原始信號具有很強的干擾作用，信號與噪聲之間的比例稱之為信噪比。

5、接收機 (Receiver) —— 承擔與發送機相反的工作，把得到的信號解碼成受信者可以理解的消息。

6、受信者 (Destination) —— 信息的接受者，可以是人，或者是動物，也

可以是機器。受信者既可以和信息源是同一類的事物，也可以是不同類的。

從宏觀的角度來看理解度理論所要處理的是講話人、傳信系統和聽話人之間的關係，但是這三者並不是在同一個層次上的；換句話說，這三者和理解度理論中的地位並不相同。講話人對應的是講話人所發出的信號，即講話人所說的言語產品，講話人包含了信息源和發送機兩部分；聽話人包含了接收機和受信者兩部分。理解度最終的目的是理解，理解的主體是聽話人，所以研究的核心在於通過聽話人的一系列反饋來得到理解的最終結果，言語產品和傳信系統則是影響理解度的主要要素。因此可以把理解度看作是因變量，言語產品和傳信系統是自變量。再看兩個自變量之間，二者也並不是處於一個絕對的平等地位。對於理解度而言，理解的對象是言語產品而不是噪聲，是言語產品的變化導致了語言的可理解度降低，最極端的情況是完全理解和完全不理解。在完全不理解這種語言的情況下，即便是完全沒有噪聲，還是不可以理解的。在完全能夠理解這種語言的情況下，須得有非常強大的噪音干擾，例如機器的轟鳴、飛機起飛、信號微弱等情況才會使得理解度接近零。在理解度理論的範圍內，言語產品應該是理解度的首要自變量，更何況理解度是語言學範疇下的研究，語言的作用首當其衝，言語產品背後所代表的正是語言的屬性，人的屬性在這裡就變得不那麼重要了。噪聲這個變量並不是不考慮進來，而是作為一個可控的變量，可以在實驗室環境下，最大可能地降低噪聲的影響，在本文中並不作為最重要的決定因素之一。

在信息結構中，參與理解的重要一個環節就是接收機，接收機接收到信號並對其進行解碼，解碼成受信者能夠理解的信息。需要探尋的是在何種言語信號的刺激下，接收機能夠成功轉化多少正確信息；在何種干擾的情況下，接收

機就不能轉化出正確的信息等等。明確理解度中什麼是自變量，什麼是因變量是非常重要的，這為理論的驗證和應用都打下基礎。因為明確了變量，就可以對變量進行控制，從而得到自變量和因變量之間的關係。

如果使用這套模型再來審視病理學範疇中的理解度的研究，也不難發現病理學研究與語言學研究之間的差別。病理學首先面對的是言語混亂的患者，患者本身作為信號的信息源和發送機出現了障礙，並不能很好地將信息傳達出來。在這種情況下人的健康屬性則成為一個重要的變量，也是首要值得關注的對象。相對的，這時候患者具體產出的言語產品的具體內容就沒有那麼重要了。可以控制言語產品的屬性來檢測患者的患病程度，從而可以更好地對症治療。

這套通信結構同樣適用於清晰度測試，恰好通信結構可以很好地展示出清晰度測試與理解度測試之間的相同點和不同點。清晰度與理解度最大的不同在於，理解度追求的是信號理解的程度，而清晰度追求的是信號接收的程度。也就是說，清晰度最關心的問題，是信號從發出機到接收機的這個過程中間，信號的損失和噪聲的干擾究竟對最後信號的接收起到了什麼樣的影響。以語言為例，清晰度關注的是聽到，而理解度關注的是聽懂。所以對於清晰度來說，噪聲這個變量就變得尤為關鍵，因為如果沒有噪聲，清晰度就是百分之百。不同類型的噪聲會對信號產生何種的干擾才是清晰度的基礎問題。相比較之下，用什麼類型的語言材料便不是最重要的考慮因素。清晰度不需要去關注聽沒聽懂，只需要關注是否聽到；而理解度是要保證在聽到的情況下，還要聽懂。

通信系統的結構對理解度理論非常重要：

首先，它幫助理解度理論認清了自己；找到了理解度中的關鍵因素，並且能夠更直觀地認識到不同因素之間的關係。因素和因素之間並不是並列的關係，而是因變量和自變量之間的關係，自變量之間也有主次之分，在數學公式的表達中可以通過系數的大小來進行加權。找到了理解度中的關鍵點也就是找到了理解度問題中的主要矛盾和著力點。這樣可以更好地排除其他的干擾因素，分配更多的研究精力和研究資源在主要矛盾上。

其次，它幫助理解度理論划定了邊界；通信系統就像是如來佛的五指山，它涵蓋了並且限制了理解度理論的運行模式，系統外的問題不是理解度理論需要關注的問題，也不是理解度理論可以解答的問題。為理解度理論划定了邊界的同時也為理解度理論畫出了自己的疆界，在這個疆界內，理解度理論可以充分地被構建，被研究，被討論，被應用充分地發揮其自身的優越性。

最後，它幫助理解度理論找到了兄弟；找到了兄弟的前提是兄弟一直都存在，兄弟不僅被找到了，並且之間還能夠互相分辨。理解度理論和清晰度理論之間的關係，以及病理學中的理解度和語言學中的理解度之間的關係，放在通信系統的結構中一目瞭然。不同理論使用的通信系統是一致的，但是其研究側重的不同引發出不同學科在不同層面上的不同的研究。在「術業有專攻」的分工研究下，不能夠陷入盲人摸象的狹隘的視野中，而要在適當的階段跳出學科的藩籬，更全面更加系統地對通信系統加以研究和思考。兄弟雖然分開過，總歸還是一家人。

這三點意義是有機統一的，且層層遞進的。只有先認清了理論本身，才能夠找到與相關理論之間的界限。理解度理論不需要也不可能包打天下，知道哪

裡是局限性，哪裡是特長就可以揚長避短。

2.2.2 理解度的概念

理解度的概念在早期的論述中有過論述，但是大多較為簡單。僅僅靠一兩句話就帶過，似乎認為理解度的概念太過於簡單而沒有得到其應有的重視。隨著研究的深入，對於理解度概念的討論又顯得有些陳舊和過時。本文認為對理解度的概念探討並不是老生常談，恰恰是因為這個問題尚沒有得到充分地討論，在當下的語言學研究中，理解度這一概念仍有待釐清、更新和拓展。

2.2.2.1 內涵

Nicholas Schiavetti (1992) 專門在文中列出一節來給可理解度下定義，他認為：「語言可理解度可以被定義為：說話者的意圖與通過傳輸系統傳遞給聽話者的響應之間的匹配程度。當聽話者的回答列表中的所有單詞與講話者產生的所有單詞全部匹配時，語言理解度是完美的。當聽話者的回答列表中的單詞與講話者產生的單詞都不匹配時，語言理解度為零。在完美和零理解度這兩個極端之間存在一個連續統，在這個連續統上，我們可以量化聽話者的答復列表與說話人的預期說出內容的匹配程度。」中華人民共和國國家標準聲學語言清晰度測試方法 (GB/T 15508-1995) 中對語言可懂度下的定義是：「由發音人發出的語言單位 (句、詞或音節)，經語言傳遞系統，為聽音人正確識別的比例。」《社會語言學詞典 (A Dictionary of Sociolinguistics)》(2004) 在相互可理解度 (Mutual Intelligibility) 這一詞條下沒有直接給出定義，而是說如果說話者能夠理解對方，那麼兩種語言變體就可以說是相互理解的。相互可理解度可以作為標準用來判斷方言和語言。由此我們知道，理解是理解度的首要要

素。《語言學和語音學詞典（Dictionary of Linguistics and Phonetics）》

（2011）將相互可理解度定義為人們相互理解的能力，同樣也提到了可以用理解度來區分方言和語言。

前文中已經提到，理解度與相互理解度事實上並不相同。本文認為有必要在此區分下二者的概念。

（1）理解度是從說話者到聽話者的過程，是一次信號傳遞與理解的過程。相互理解度是雙程的甚至是多程的，一程中的信號源與發送機，可以作為下一程的接收機和受信者，角色存在轉換。

（2）因為相互理解度是雙程或多程的，所以就要考慮到兩次或多次的噪聲干擾。

（3）理解度適用於單向話語模式，例如聽一場演講，聽話者不能通過語言給說話者予以反饋，其理解度相對穩定；真實的相互理解度更適用於對話交流模式，二者交流時會因為交流過程中的即時反饋而增強二者之間的相互理解度，因此真實相互理解度會處於一個動態的過程中。

上文提到真實的相互理解度是設想一個真實的對話場景，兩個人或多個人互相說話的場景，此時的相互理解度是真實的相互理解度。與之相對的是實驗室相互理解度，將相互理解度的對話場景折分成實驗室中的兩次單向的理解度測試。雖然這樣的實驗得到的結果，可能與真實的相互理解度之間存在誤差，但從另一方面看，它可以幫助有效規避許多干擾因素，最終可以通過不斷改進實驗方法，逐漸縮小與真實理解度之間的誤差值。

對比上述幾個定義不難發現一個有意思的現象，那就是每個學科都對理解度下定義，似乎理解度是從屬於病理學，有從屬於社會語言學，還屬於聲學的一部分。顯然這種現象是混亂的，是有問題的。論證理解度理論的學科從屬也是對該理論研究的前提。本文認為應該從心理語言學的範疇來為理解度理論下定義，理由如下：

(1) 理解度的核心在於理解，對自然語言的理解是一個心理過程，是大腦對語言信號的認知和處理。這是一個高級的複雜的心理現象其中涉及聽覺信息處理，言語信息的記憶提取，語法分析，語義理解等階段，這些全部屬於心理語言學的研究範疇。而病理學的核心問題在於如何診斷和治療病人，理解度測試只是作為醫生診斷的一個方法，其主要目的並不是為了更好地發展理解度理論。

(2) 心理語言學的研究方法能夠更適用於理解度的研究方法，心理語言學研究方法的基本原則、實驗方法、建立理論的方法還有統計手段很大程度上符合理解度的研究方法。社會語言學的調查方法偏重於社會調查，例如問卷法，訪談法，問路法或非介入式調查法等等，不適用於對理解度的測量，但是理解度的結果可以應用到社會語言學中，後文會對這一問題進行詳述。聲學的研究方法確實與理解度的研究方法有大量重合之處，甚至絕大多數方法和實驗材料可以直接借鑒用於理解度的研究，然而二者的研究目的迥然有別，使用相同的方法，但終點卻不相同。

本文將理解度理論納入到心理語言學研究範疇下，鑒於當前心理語言學雖然對詞語或句子的理解度進行了深入的研究，但尚沒有形成系統的理解度理

論。本文主張將理解度理論納入心理語言學研究範疇，在汲取心理語言學已有的理論成果的基礎上，再綜合其他學科對理解度的應用，進而確立深厚的理解度根基和明確的發展方向。

由於以往的文獻中對於理解度的定義論述並不夠充分，更多的是著眼於理解度的用法。本文希望能夠建構起理解度的理論，那麼理解度的概念則是不能繞過的基石。從理解度的概念出發可以對條件的約束和限制再衍生推導出其他相關下位概念。

語言信號由說話者說出經傳信系統到達聽話者的這一過程中，聽話者對原信號即時的理解程度，為單向理解度。在對話過程中，聽話者與講話者的身份交替轉變，雙方或多方輪流成為聽話者，對話參與者相互之間的理解度即為相互理解度。

這其中包含四個重要的關鍵要素，分別是信號、過程、主體和即時理解，他們內部各自都有需要界定的內容，通過這幾方面對理解度進行定義，既可以符合前人的研究成果，也可以支撐起後續的深入研究。建立任何一個理論，除了討論理論本身的架構外，都還要建立理論的基本假設作為理論的根基。在討論概念的基本要素之前，有必要對理解度中包含的基本假設做出說明。對基本假設的說明，有助於更好的應用理解度概念，能夠更明確理解度的使用範圍以及各個要素之間的職能。

基本假設1：理解是可以被量化的。可以通過科學的實驗和統計方法對理解度進行定量研究。

基本假設2：理解度的理論應反映真實人際對話中的交流模式。即便是在

實驗室環境內，使用人機測試的方法，也是要考慮到實際對話中的因素。

基本假設3：原信號是可以被人類所理解的有意義的信號，若信號本身沒有有意義，是不能夠被理解的信號，那麼就無法對其測量理解度，理解度也就無從談起了。

基本假設4：理解度僅限於對信號的字面意思的理解，暫不考慮對字面意思意外內容的理解。言外之意的理解是更高層次的語言理解，其理解機制與普通對話的理解並不完全相似，這個理解過程會涉及到更多的文化百科知識和個人經驗。例如：「今天星期天。」字面意思就是今天是星期天。如果一個孩子對父母說今天星期天，可能就是想說今天想出去玩；一個老人對孩子說今天星期天，可能是希望他們今天能夠來家裡看望自己。對於這種情況可以在實驗材料的選擇上加以規避，消除言外之意對理解度的干擾。

1、信號：信號可以是任何信號，但在本文中如不做特殊說明則指的是語音信號，是人類語言的聲音信號。本文將刺激的性質定義在聲音信號中，排除了文字、符號、標識等非聽覺性刺激。這並不代表文字、符號、標識不能夠有理解度的問題，他們的理解度問題不在本文的考察研究範圍內，但是本文所給理解度下的定義仍然適用。即便是無線電信號，依舊可以納入到本定義中，無非是接受者完全無法直接接受解碼無線電信號，因此理解度是零。

在語言學研究範疇下，人類語言作為理解度的信號是毋庸置疑的，但是人類語言具有層級性，從最小的單個音素，再到音義結合單位，短至語素，長至語篇都是不同層級單位的人類語言。並不是任何層級的語言材料都適合作為理解度信號的。首先純粹的語音信號，例如音素或無意義的語音，是不可以作為

理解度的信號的。因為其信號中沒有可理解的內容，這樣測出的並不是理解度而是清晰度而已。在各個音義結合的層級當中段落和語篇雖然內容豐富，但是內容過長，大大超越了人類短時記憶（short-term memory 簡稱 STM）「7±2」的極限，並不適合作為單次實驗的信號。如果將語篇拆分會更有利於被試的理解，也能夠利於理解度的測量，排除了短時記憶的影響。剩下的詞素、詞、短語和句子中，詞素因為過於簡短，且通常有歧義成分存在，所以一般不適用做理解度測試的材料。後三者也是眾多已有研究使用過的材料，林素娥、蘇憲萍（2007）認為應選取詞為主要的測試單位，同時結合少量的體現方言語法特徵的語句進行理解度測試。游汝傑、楊蓓（1998）認為詞彙是影響方言之間可懂度的最重要因素。詞彙相同，語音不同，有時還能聽懂或勉強聽懂；詞彙不同，則可懂度為零。由於詞彙必須借助語音的以表達，所以口語詞彙可懂度實際上已經包含語音因素。

本文認為句子是最適合做理解度測試的語言材料。句子優於短語和詞的原因有以下幾點：

（1）句子相比於短語和詞能夠提供一個完整的意義和語境，並且完整的句子也包含相應的語氣。從信息結構上看，句子更完整。完整的語境能夠有效地消除歧義。並且在不同語言中詞彙的概念和外延往往不同，其理解程度和正確性都會隨之受到影響，而一個完整的句子能夠盡可能的將這類問題降到最低。

（2）句子相比於短語和詞更貼近實際交際情況，無論是單向還是雙向的交際模式，句子都是扮演著基本的單位。測量理解度也是要以真實交際情況為準，對句子的理解相比於對短語和詞的理解更具有現實意義，更具有說服力。

(3) 句子的音節數量相比於短語和詞更符合人類短時記憶的區間，信息量適中，所用時間也適中，既沒有詞或短語的短暫感，也沒有語篇信息過多的冗長感。符合語言交際的實際情況也便於測量。

2、過程：從信號發出者經由傳信系統到達聽話人的過程，這過程中信號有可能會被噪音干擾，從而影響理解度的數值。聽話人包含了信息系統中的接收機和受信者兩個角色。

3、主體：信號發出者可以是人或者是機器，但是受信者一定是人。理解度的主體是聽話人，聽話人接收信號，理解信號，最後對信號做出反應，一系列的理解過程都是在聽話人這一環節進行的，因此聽話人的主體地位是應該確定的。因此在設計相關實驗時，對聽話人的把控則是保證實驗進行的重要環節。理解度測量最大的問題在於流程複雜導致變量太多，突出聽話人的主體地位，就是將非聽話人的變量控制住，找到聽話人有限的變量再對其進行控制，達到最終的控制變量。以聽話人作為主體，聽話人的身份背景和語言背景就作為理解度研究的重點。之所以沒有選擇說話者作為主體，是因為在理解的過程中，信號發出者並不對理解這個過程產生影響，而是發出的信號對聽話者產生影響；實驗中，信號發出者並不一定是人，可以是機器，可見說話者是可以被替代的，不可被替代的是說話者的信號；說話人也並不是完全不重要，從相互可理解度的角度來看，說話人和聽話人的身份是交替互相轉換的，所以在實際情況下二者是相互交替成為理解度的主體的。

4、即時理解：本定義對聽話人的反應進行了具體的界定，需要其對信號理解後做出反應。首先要對聽話人是否理解了信號內容進行考量，需要設置相

應機制對是否理解進行監督篩查，理解度的前提條件就是要理解，如果沒有理解就沒有度。其次是及時地做出反應，及時性是前人定義中從未出現過的內容。之所以會將即時性添加到定義中，第一，理解度是處理自然語言語流中的理解，在這種情況下，反應的時效性尤為重要，如果未能及時作出反應，則勢必會影響到語言的理解；第二，即時性也是對理解性的篩查機制，其目的在於可以篩查出非理解因素導致的正確反應，即通過大腦的聯想，猜想得到的正確答案。

2.3 理解度的性質

理解度的性質也是理解度的特點。本文從即時性、聽覺性、單向性、客觀性、動態性、相對穩定性這六個角度去對其進行描述。對理解度性質的描述一方面可以更好地理解認識理解度本身，更重要的是利用理解度的性質去更好地解決理解度的相關問題。之後的研究都會利用這六個屬性指導和約束材料的篩選、實驗的設計、數據的分析等內容。

理解度是即時性的。《現代漢語詞典》（2012：795）理解的詞條下面解釋為：「懂，瞭解」。這是通常意義上的「理解」，理解度中的理解與通常意義中的「理解」有些許差異。其差異在於通常意義上的「理解」可以使用在例如「我當時沒聽懂你說了什麼，後來才理解了你的意思。」等情況中是經過一段時間後，才理解當時說的話。這種「懂」無論是對字面意義上的理解，還是超越字面意義外的意義上的理解，都是花費了一段時間的。理解度中的「懂」應是有即時性的，有時間約束的，在很短的時間內能夠做到的「秒懂」。這種理解在理解度上才是有意義的「懂」，因為設想在真實的語言環境中，對於一種熟悉的語言，即便一個人語速多快，你都可以清楚地聽懂每個字每句話，如

果是一個不熟悉的語言，一個人的語速稍微快一些，理解上就會吃力，因為大腦在處理這個詞的時候，下一個詞已經出現了，不斷出現處理時間的延遲導致最後已經不記得上一個詞是什麼的時候，下面的詞也已經說完了，最終導致無法完全理解對方在說什麼，這時的理解度就大打折扣了。也就是反應時間對理解度產生了重要的影響，可見瞬間的理解和經過處理後的理解二者之間是有差異的，反應在理解度上就是通過反應的時間來對理解度進行量化。這個度既可以看成是單位時間內理解的程度，也可以看成理解相同內容所花費的時間，二者在量上是等價的。

理解度是聽覺性的。理解度中聽話人是對語音信號進行理解並做出反應。理解度中的信號刺激是之所以採用聽覺信號，而不採用視覺信號刺激，原因有以下幾點。首先，聽覺信號相較於視覺信號更單一，通過技術手段將聲音信號中的雜音濾去減少噪音的干擾。而視覺信號較為複雜，不能很好控制噪音對信號的干擾。使用聽覺信號刺激更具有可操作性。其次，相較於文字而言，語音是第一性的。聽力水平對於語言能力來講是處於第一位的，也是最基礎的語言能力。有些語言或方言只有語音形式沒有文字形式，而且漢語方言中，更多的現象是文字形式相同而讀音不同。還存在著相當一部分的方言詞並不能用漢字表示出來。所以語音形式更能夠具有普遍性。再次，從日常交際的實際情況出發，使用聽覺刺激也是符合日常生活的，最終所反映出的理解度也更能夠符合常識。最後，大腦對於語音的處理時間是快於對文字的處理時間的。從理解效率的角度上看，聽覺信號也是優於文字信號的，更有效率性。本文認為理解度並不是絕對不能夠使用視覺信號作為實驗的刺激材料的，目前見到的使用視覺信號作為材料的實驗多是用於對標識的理解，這已經超出語言學的範疇。在

語言學研究範疇內，對於特定的問題可以採用專門的方法來研究。

理解度是單向的。是從說話者經由傳信系統抵達聽話者的過程，此一個過程發生一次理解，此一程就擁有一個理解度。如果是兩個人交替對話，一個對話中就有兩次理解的過程，那就有兩個理解度，信號每被理解一次就產生一次理解度。每有一次理解度，則必定會有一個理解主體的出現和一個傳信過程，然而發信者既有可能是一個人也有可能是多個。通過實驗得到的理解度是將多個單次理解度進行統計計算得到的。確定理解度是單向性的有助於將複雜的交際模式簡單化，找到其內部最簡單最基本的單位，對其進行測量和計算。一個人對一種語言的理解度應處於一個恆定的值，而一個人並不是對該語言的每句話的理解度都處於那個值，所以最終的理解度就是基於多個單次理解度的平均值。通過多個單次理解度的數值有效地抵消了一部分隨機誤差，從而獲得一個更符合實際的數值。在這個過程中，理解度的性質沒有發生變化，僅僅是數量上的差異，依舊保持理解度的單向性。另一方面，理解度的單向性隱含著一個非常重要的問題，那就是相互理解度中的不平衡性，兩種語言之間的相互理解度是可以拆分成兩組單向的理解度，那麼二者之間就必然存在著差異，這種差異就是相互理解度中的不平衡性。不平衡性很好地解釋了為什麼兩種語言系統之間的特徵差異數量是固定的，也就是說二者之間的距離是一樣的，但是相互理解度中一方高於另一方的這個矛盾。

理解度是客觀性的。一個人對一種語言的理解度是客觀的，與他的個人身高，體重一樣都是一個人的客觀屬性。不像是忠誠，熱心，漂亮這種屬性是隨其他人的態度而轉移的。一個人的理解度之所以是客觀的，是因為理解主體的語言和聽到的語言之間的距離是客觀的；並且他對語言信號的理解這一過程是

客觀的；除非聽話人不配理解之外，影響理解度的因素也是客觀的。理解度客觀性決定了理解度是可以通過科學的方法進行測量的。交流過程中個體的主觀感受也能夠從側面反映出理解度的客觀水平，這其中的差值會是明顯的且必然的。不能依靠主觀感受來對理解度進行估算，但是主觀感受也能夠說明一些問題，例如當主觀感受與客觀水平差距太大的時候，有可能說明實驗出現了一些較大的問題。若將溝通的雙方從個體分別擴大到兩個同質的群體，那麼兩個群體間的理解度也是客觀的，可以通過每對個體間的理解度計算得來。一個人對一種語言的理解度可以與身高體重一樣成為一個常量，一個群體中相同的人的理解度也可以構成一個集合，從而形成這個群體的常量。這個常量可以作為一個常模，運用這個常模來和其他群體的理解度進行對比。這就好像是黃種人的身高是一個常量，白種人和黑種人的身高也各是一個常量，三者可以相互對比得到誰高誰矮的結論。同樣的方法依舊適用於理解度。群體內部的語言能力的同一性也決定了是否能夠得到客觀一致的語言間的理解度。

理解度是動態性的。理解度雖然是一個常量，但並不是一成不變的。一個人通過對一門非母語的語言的學習可以不斷增強其對該門語言的理解度，伴隨著他學習的進步，理解度有可能達到百分之百，有可能止步於一個固定的水平；也有可能長時間的不使用這門語言而逐漸忘記了，理解度逐漸下降。而上升到群體層面，一個人的理解度的變化不足以撼動整個群體，但是隨著量變的增加，群體的理解度也會隨之改變。放在歷時的角度去觀察，那麼群體與群體之間，即語言和語言之間的理解度就是處在一個動態的狀態，取決於兩種語言之間的人口流動，互動等等因素。兩個村落本來屬於一種語言，但是由於長期缺乏互動，最終兩個村落之間相互理解度就可能降到非常低的水平，甚至完全

不能溝通。通過長期的頻繁互動，兩個單語群體也可能形成了一個雙語群體，或者是兩個單語群體合併成為一個單語群體。因此從總體來看，語言間的理解度是一個動態的常量。可以通過觀測這個常量的變化趨勢，瞭解到一個群體對一種語言的熟悉程度，也可以認識到一種語言對一個群體的影響。從中可以為語言接觸和語言演化提供實際的數據和信息。認識到理解度的動態性就應該將研究視野從共時層面拓展到歷時層面，對不同時間下的理解度進行分析，找到引起質變的時間節點下的理解度，從理解度角度去審視語言變化的規律。

理解度是相對穩定的。理解度的值雖然是一個動態的過程，但是在外界變化較小的時候，理解度還是保持在一個相對穩定的狀態。首先成人的語言學習是一個累積的過程，這個過程中理解度有可能是緩步提升的。在一個特定的時間點內，或較短的時間內，理解度是保持相對穩定的。相對穩定也給理解度的研究提供了可能性。在實際交流中，兩個人的相互理解度可能會因為交流時間的增加而增加，一旦交流結束，理解度又會回到與原來差不多的水平，維持在一個相對穩定的區間。在群體角度，短時間內，個人的語言能力提升，對群體的改變是微乎其微的，處在一個穩定的階段。所以在共時的層面上，群體和群體之間的相互理解度是一個穩定的常量。通過這個穩定的常量可以衡量出群體和群體之間的親疏遠近關係。因此可以利用理解度作為標準之一判定語言和方言之間的關係。

2.4 理解度的影响因素

本節將會討論理解度的影響因素，理解度之所以在漢語學界沒有得到非常廣泛的應用，其中一個原因就是因為影響理解度的因素太過於複雜。不能夠控制好各個變量之間的關係。所以導致不能非常好的對理解度進行分析。Varonis

E.M. & Gass S. (1982:132) 通過實證研究發現言語理解度涉及多個因素，可表示如下：

$$C = P\alpha + G\beta + F_1\gamma + F_2\delta + F_3\varepsilon + \dots + FL\zeta + S\eta \dots$$

其中，C代表聽話人對說話人言語的理解度，P代表說話人的語音，G代表說話人的語法，F1代表聽話人對話題的熟悉度，F2代表聽話人對說話人的熟悉度，F3代表聽話人對說話人母語的熟悉度，FL代表說話人的語音流利度，S代表社會因素，公式中的希臘字母代表各個因素的影響權重。

陳海倫（1996:366）對於在方言間的溝通度的影響因素方面給出了八個主要因素：

- 一、話語中提供的信息量；
- 二、交際者對對方方言知識的累積程度；
- 三、系統中對應語音的相似程度；
- 四、各語音對應類型在總體中所佔的比重；
- 五、各語音成分對應規律的強弱；
- 六、詞彙共有量；
- 七、詞的現次率；
- 八、語法一致性的程度。

在這八個因素中，第一、第二和第七個因素可以在語言材料中排除，此外

的五個因素是必須要考慮的。其中因素三、四、五決定了方言語音的溝通度，因素六、七決定了方言詞彙的可溝通度。

可見影響溝通度的因素確實很複雜，但是並不能因為複雜就放棄研究，首先需要認識到影響理解度因素的特點，然後對其進行歸類，最後根據其特點不同類別不同對待。主要的目的是化繁至簡，能夠通過理性科學的分析，正確的認識不同的影響因素，從而更好的指導實踐。

本文經過研究發現，影響理解度的影響因素有以下幾個特點：

(1) 類型多：Varonis E.M. & Gass S. (1982:132) 總結有至少有八種因素影響理解度，但是在他的公式中，也使用了省略號來表示可能會有更多的因素影響到理解度。陳海倫 (1996:366) 也總結出八個影響方言間溝通度的主要因素。可能對於自然科學來說八個因素並不能算作非常多的變量，但是對於傳統的語言學研究來說，八個變量本身已經是一個非常複雜的問題了。然而這八個變量並不僅僅是八個，更準確的說是八類。每一個變量其實是一類變量的集合，每一個下面又會包含著若干的細節，如果最終落實到具體特徵上，估計會是一個非常龐大的數字。

(2) 層級多：理解度的影響因素中，最明顯的就是語音的差異。語音差異可以說是首要影響因素，然而語音差異中包含了不同層級的因素，其中包括音段音位的差異、超音段音位的差異；音段音位裡面包含元音、輔音的差異，超音段音位中就更加複雜包含了聲調、重音、停頓、句調。其中元音還包含著共振峰F1、F2的格局差異，輔音的VOT等等細節特徵的差異。然而這僅僅是眾多影響因素中的一個類別。若考慮到語法或詞彙因素的影響，那麼需要考慮的層

級將會更多。

(3) 範圍廣：除去語言系統內部的眾多特徵的差異，非語言的因素也會對理解度產生非常大的影響，例如談論某個問題時候，聽話人對於說話人所談論事物的大背景是否有所瞭解。背景知識能夠在某個階段顯著提升理解度。除此之外。還有噪聲的影響也會影響理解度，周圍其他聲音或語音的混入，會對理解度產生一定的影響。甚至聽話人的性別，年齡都會對理解度產生一定的影響。這些都不是語言內部的差異，而是包含了社會、心理等其他相關方面的因素。

(4) 容忍度不一：大腦對與變量之間的容忍度是有差異的，語言系統中，特徵的核心是清晰的，但是很多特徵的邊界是模糊的。在部分人聽來一個音可以是[b]，但同時另一部分可能聽起來就像是[d]，這是不同的大腦對於這個音的分類是有出入的，這也會導致特徵的不清晰和混亂。

以上是對影響理解度的因素的特點的描述，可見影響理解度的因素是一個非常龐雜的系統，認識到這個系統的特點僅僅是第一步。如此紛繁複雜的系統，一定有其內部的邏輯和規律，如果要深入研究，首先需要對影響因素進行分類分層處理。通過分層和分類的方法找到其中核心的因素，與邊緣的因素區別對待，從而實現對理解度影響因素的簡化。

首先根據影響因素的產生位置進行分類，分為語言系統內部影響因素和非語言系統影響因素兩大類。

一、語言系統內部影響因素

語言系統內部對理解度產生影響是在於兩種語言之間的差異和共性。語言本身就是一個複雜的系統，主要包括語音、詞彙和語法三大方面。從這三大方面分別來看其對理解度的影響因素。

1、語音：前人研究主要集中於研究語音對理解度影響。在一個語音系統中，根據的音位屬性可以分為音段音位和超音段音位。音段音位中包含元音、輔音等音位系統；超音段音位中包含聲調、重音、停頓、句調等音位系統。在西方早期理解度的研究中，由於涉及到語言治療的部分，所以更加關注音段音位對理解度的影響，他們利用語言中的最小對立對進行測試，從而發現不同音位對理解度的影響。傳統音系學對漢語方言的分析主要集中在聲韻調的類型上，將音系差異與理解度相聯繫的研究出現在20世紀90年代，以語言距離作為研究對象，把兩個方言的所有音系特徵列出，計算他們之間的差異值後比上全部特徵數量得到二者之間的差異值。通過差異值的大小來推論兩種方言之間的理解度。本文認為語言距離確實能夠說明二者之間存在著的差異，並且這些差異確實對理解度產生影響；語言距離不是對理解度的唯一影響因素，只是影響因素之一。語言距離理論並不能很好地解釋理解度的不對稱性，比如粵方言母語者對於普通話的理解度要高於普通話母語者對粵方言的理解度。而二者之間的語言距離是一個固定的值，並不會有差異。本文也贊同陳海倫（1996）的觀點，認為溝通度是動態的度量，語言距離是一種靜態的度量。因為是一種動態的度量方式，就會存在很多不確定的變量在影響著溝通度的大小，於是就要求新的研究中加入實驗的方法去控制各種變量更好的能夠測量實際的溝通度。語言距離並不因為是靜態的而沒有價值，是能夠從客觀實際材料對實驗測量出的理解度進行參考。

超音段音位中，除去聲調外，其他的因素暫時還鮮有研究，其中一部分原因在於早期多是選擇使用單詞作為實驗材料的基本單位。因此沒有對其他超音段音位進行研究。在漢語中通常情況下沒有詞重音，在句子層面上重音作為一種表示強調的手段，重音落在哪個詞上，句子的焦點就落在哪裡。如果沒有明顯的重音，那麼自然焦點就是落在句子的結尾。句調的升降必然會引起語義的變化，在漢語中句調的作用尤其重要，因為漢語疑問句的疑問詞並不需要在語音層面發聲位移，從句式角度來看，一個句子擁有疑問詞也不一定就是一個疑問句。例如：「我也不是你的誰。」因此句調的差異對句子的理解產生較為重要的作用。疑問句和感嘆句與陳述句之間都有非常明顯的句調差異。

漢語句子中的停頓主要出現於主謂之間，對句子意思的影響並不是很大；但是句子中各個詞語之間的停延率確實會形成不同的語言風格，這種差異是否會影響句子意思的理解，目前還沒有見到相關論證，暫時不把這類特徵納入影響理解度的因素中，不過會在具體的實驗材料中注意到詞語間停延率的相關問題。

2、詞彙：前人研究中也注意到詞彙對理解度的影響，盡量選用的是基本詞彙其中包括人體相關部位，數字，日期等核心詞彙。避免各個方言區的方言詞對理解度的影響。這樣做的好處是可以排除一部分詞彙上的干擾因素，讓聽話人的注意力集中於語音部分。詞彙要素不同於語音要素之處在於詞彙對於理解度的障礙是極性的，懂就是懂，不懂就是不懂，不是一個漸變的過程，粵方言的勞駕是「唔該」，這兩個詞哪怕用普通話讀出來，非粵方言區的人照樣是不明白的。甚至粵方言的外來詞「士多」，來自於英語的store，是小賣部的意思，對於非粵方言區的人用任何方言語音都沒辦法理解到這類詞彙。所以在很

多港式普通話中就會出現大量的粵方言特有詞。這些特有詞彙的運用導致了港式普通話即便使用普通話的音系系統，仍舊有一部分信息不能夠被理解。另外，詞彙的同源性也能夠從側面體現出方言間的理解度。與語言距離類似，詞彙的同源性也希望通過詞彙系統的差異來計算理解度。本文認為語音和詞彙都是理解度的重要影響因素，但是單純通過計算一個因素的差異度來得到理解度是片面的，不能否認的是，這二者的意義是很重要的，可以是理解度一個側面的評價指數，如果把理解度、語言距離和詞彙源流綜合來看，可能可以得到一個回歸曲線，瞭解到他們三者之間的關係，也是十分有意義的一種探索。

3、句法：句法方面目前還沒有將其作為影響理解度的主要因素而加以研究。原因大致有以下幾個：（1）句法內部系統結構複雜，很難剝離其中對應的細節特徵進行控制變量；（2）句法結構對句子的影響多是判斷句子的合法與不合法，通常合法的句子在理解上就不存在理解的困難；（3）在一門語言中，句法的可變空間並不大，所以通常不能夠產生差異足夠大的變項。句法因素在漢語方言間卻可能作為一個影響理解度的因素，當然影響程度可能不及語音和詞彙。因為漢語方言不同程度地保留了一定的古代漢語的句法要素，其中包括一些古代漢語的句式和語序，這些特徵提供了句法充當理解度影響因素的條件。比如在粵方言中，會出現一些特殊的句式；例如：直接賓語和間接賓語位置的互換「聽日畀錢我（明天給我錢）」；比較句「佢高過你。（他比你高。）」等等，這些可能會為理解造成一定的障礙。

二、非語言系統影響因素

1、信噪比：信噪比（SNR）是所需的信號的強度和背景噪聲的強度之間

的比值。噪聲的增大勢必會降低原有信號的清晰度，而清晰度的降低也會影響信號的理解度。噪音的存在是客觀的，也是難以避免的，尤其是在複雜的工作環境比如工廠，鬧市，水下，高電磁干擾環境等。人耳接受聲音是不具有選擇性的，沒有濾波器的，所有的聲音統統都要轉化成信號傳入大腦。當然大腦存在著一定的對聲音的選擇性，對自己可以理解的聲音更具有敏感性。目前，有相當一部分的研究是針對如何降低噪音對信號的影響。通常會採用不同算法以實現語音的增強效果，其中包括譜減法、統計模型算法、維納濾波法等。這些研究對具體的實際生活都具有重要意義，但是並不適用於語言學研究中的實驗。其原因有以下幾點：

(1) 對於語言清晰度的研究方向與理解度的研究方向是不同的。語言清晰度中的語言材料並不是最重要的實驗變量，而對於語言學研究來說，語言材料對理解度產生的影響才是首要研究的目的。

(2) 由於二者之間的研究目的不同，所以變量的層級就是不同的，在語言學研究中，噪聲是可以通過技術手段進行規避的，讓噪聲降低到一個不會產生影響的水平上。另外，也不會人工增加噪聲來測試其對理解度的影響，只是將噪聲作為一個變量進行控制，弱化其在理解度中的干擾作用。

雖然通過人為手段將噪聲的影響控制到最低，並不是要忽略其影響力。在正視噪聲對理解度的影響情況下，對其進行有效地規避，目的是為了降低實驗成本減少變量數量。噪聲對理解度的影響出現在信號傳輸的過程中，理解度實驗其實把這一過程拆分開來，因此噪聲就會對語音信號產生兩段干擾，一段是在錄音的階段，另外一段是在給聽話者聽音的階段。所以若想對噪聲進行有效

控制，就需要在這兩段中進行降噪處理。

2、背景信息：對句子的理解除了需要語言學的知識外，還需要豐富的百科知識對理解進行輔助。Bransford, J.D., & Johnson, M.K. (1972) 的經典實驗中給被試提供一段文字，一組是完全不給題目，一組是先呈現題目然後呈現材料，最後一組是先呈現材料後出現題目。最後先呈現題目組的回憶成績明顯好於其他兩組，其他兩組的表現相似。題目在此項實驗中充當了背景信息的作用，提示了理解文章所需的基礎信息，此實驗得到的重要結論就是背景知識促進了理解度的增長。為理解話語，需要的是更多有結構的知識，這種知識是由較高層次的結構來解釋的（王亞同、李文崗2011）。背景信息並不是在任何理解度階段都可以起作用的，例如一個不會外語的人無論他擁有多麼豐富的百科知識，多麼瞭解對方所講的內容的背景，都沒辦法知道一個講外語的人在向他具體講些什麼。背景信息對於那種似懂非懂的情況下，能夠對理解度有顯著的提升。因為背景信息提供了一個基本的語義場，在這個語義場中，大致會包含若干個相關的詞彙。聽話者在聽到這些詞彙的時候可以更方便更容易更快速地將語音與大腦中存儲的相關內容進行聯繫，並在大腦中提取出相應的意思。情景模型對於理解非常重要。在書面話語與口頭話語如對話中都有情景模型發揮的作用（王亞同、李文崗2011）。事件索引（event-indexing）模型已經區分了情景模型的五個成分：空間、時間、實體、動機以及因果關係。聽者將聽到的內容與記憶每個成分的當前心理表徵結合起來，利用情景來激活句子的理解。最後背景信息可以讓聽者快速地消除歧義，在正確的情境下獲得正確的意思。

但是究竟背景信息對理解度產生怎麼樣的影響，影響的程度是什麼樣的，

在理解度達到什麼樣的水平的時候，背景信息會提供顯著的幫助，也是是很多尚未解答的問題，需要不斷的探索和實驗才能夠得到有效的結論。因此目前的實驗暫時先不考慮背景信息對理解度的影響，盡量選擇沒有專門背景信息就可以理解的句子，使用常用的語義場內的詞彙便可以降低直至消除其影響。

3、社會因素：理解度曾經在社會語言學領域進行過較為詳細的論述，可見語言的社會屬性對理解度的影響也是十分明顯的。從一個言語社區的內部來看，不同性別，年齡，職業，教育程度都會表現出對語言的敏感性的差異化分布。在多個言語社區之間來看，通常高地位言語社區的人是可以理解低地位言語社區人的語言，相反卻不行。就如古羅馬的貴族使用的規範的拉丁語，下層民眾使用的是簡化的不規範的拉丁語，這樣貴族就可以理解下層民眾的語言，而下層民眾就不能完全理解貴族的語言。除此之外，Wolff（1959）還認為語言態度也可以影響一種語言的理解。相關更多論述更多屬於社會語言的研究範疇，社會語言學對於社會角色屬性的要求更加細緻，也體現出理解度的社會屬性。

4、認知因素：不同語言之間會存在著認知方面的差異，例如漢語與英語之間存在著對空間認知方面很大的不同。在漢語中往前翻書對應著英語中的backward，而向後翻書對應著英語中的forward。再例如英語中通常會將郵寄地址的最小的單位放在前面，而漢語則是完全相反，將最大的單位放在前面。這種認知差異必然會對理解度產生影響。不過這類認知方面的差異，多體現在語言距離較大的兩種語言之間，方言與方言之間通常並不會出現明顯的認知模式的差異。具體認知因素會對理解度產生怎樣的影響，目前還不能得到明確的答案，首先認知因素內部還會存在諸多子因素，若想具體知道究竟認知因素在某

對語言的理解度上有多少影響，還需要個別問題個別分析。

上文對影響理解度的幾大方面的影響因素進行梳理，分別從語言因素和非語言因素兩大部分進行分析。其目的是為了在眾多影響因素中確定哪些是重要影響因素，哪些是不重要的因素，哪些因素是符合實驗目的需要控制變量的，哪些因素是不符合實驗目的需要合理規避的。於是可以將各個影響因素分層分級處理，不同級別的影響因素對其有不同的處理方式。

表格 1 理解度影響因素等級和處理方法

	影響因素	處理方法
第一級	語音	重點分析
第二級	詞彙、句法	綜合考量
第三級	背景信息、社會因素、認知因素	加以注意，暫不深入探討
第四級	噪音	合理規避

在語言學研究的框架內，首先要著重處理語言本身對理解度的影響，所以將語音、詞彙和句法三者放在第一級和第二級中。在這三類裡面，語音對理解度的影響更加明顯也十分基礎。因此需要對其進行重點分析，詞彙和句法的部分需要和語音綜合考量，如果單獨對詞彙和句法進行變量控制並不是不可以，可能會服務於其他的研究目的，比如對新詞彙的理解或特殊句法結構的理解等方面。第三級和第四級中主要是非語言因素：第三級的背景信息和社會因素需要在實驗中加以注意，不能忽略其影響力，在實驗材料的選擇和實驗被試的選擇中充分考慮到這二者對理解度的影響，力圖求簡，只是不將其作為討論和分析的重點。最後一級是噪音這個影響因素，排在最後一級不是說噪音對理解度的影響不重要，相反，噪音對理解度的影響有可能超過以上所有。但是在語言學的研究中，不把噪音作為一個主要考量的因素，而把噪音只作為一個標準，

在實驗室的噪音環境下對理解度進行測量，從而合理規避了噪音對理解的影響。總之通過對理解度影響因素的深入討論，本文總結出四個級別，每個級別中還可以細分為若干小級別。總體來看，通過將因素分級考量，可以有效的優先分析高級因素。可以將眾多複雜的因素化繁為簡，大大提高研究效率。

2.5 理解度的測試方法

前文中提到Voegelin, C. F., & Harris, Z. S. (1951) 對如何確定語言間的理解度的方法進行了深入地討論。作者總結了四種研究方法，分別是（1）詢問調查法（the 「ask the information」 method）；（2）計算相同法（the 「count sameness」 method）；（3）解析結構法（the 「structure status」 method）；（4）測試信息法（the 「test the information」 method）。Kate Elizabeth Bunton (1999: 2) 寫到有兩種測量理解度的方法，分別是評價法（Scaling Procedures），另外一種是認詞法（Word Identification Tasks）。認詞法中包括轉寫和選擇這兩種模式用於不同的研究目的。Casad, E. H. (1974) 使用一種叫做記錄文本測試（Recorded Text Testing, RTT）。這種測試的標準流程是提前準備一段話，讓被試者聽一遍之後，在聽第二遍的過程中向他詢問十個問題，所有的內容都是被試者的母語（Blair 1990:73）。Kluge, A. (2006) 針對RTT總結的優勢在於可以對全文進行測試而不是選擇性的片段，還有讓被試復述故事要比直接答題更加合適，並且測試過程不需要預先設計問題，也不需要向被試翻譯。同時也存在著一些不足，主要是體現在測試的打分和評價過程中會費時費力，容易造成不客觀的現象。

Weismer (2008) 介紹了兩種，一種是特徵分析法（Feature-analytic Measures）；另一種是轉寫法（Transcription Measures）。前者是對不同聲學特

徵的詞語進行選擇，後者是計算轉寫的詞與音頻刺激中的詞語的比例。在前文提到的文獻中，對方法論的使用也是各不相同。經過半個世紀的發展，目前通行的理解度的測試方法是運用實驗的方法來對理解度進行測試，其實目前的方法其核心概念是繼承自「測試信息法」。通過規範的實驗的方法來對理解度進行實驗測試。

Gooskens, C., & Schneider, C. (2016) 在他的研究中改進了之前的研究方法，採用兩種測試理解度的方法，一個是選擇圖案的任務，另外一個是單詞翻譯任務。在選圖任務中，被試要求在聽完單詞之後在四個圖片中選擇一個正確的。這種實驗方法也被應用於 (Schüppert, A., & Gooskens, C. 2011, 2012) 中。這種方法可以讓更多的人在短的時間內接受測試，並且不受閱讀和識字的限制。

以上的眾多研究中，雖然都採用的是實驗的方法，但是實驗中的實驗要素的不同導致實驗的過程相差非常大。實驗過程中也涉及到眾多參數及變項，研究者可以根據實驗目的的不同進行相應的調整。本文將實驗中會涉及到的各個變項列出，並且根據理解度的性質對其進行甄別，留下符合要求的選項，進而優化實驗設計。

2.5.1 實驗材料

實驗材料可以根據語言材料的屬性來進行選擇，理解度的測試要求實驗材料首選一定是語音材料，如果希望被試能夠將注意力全部集中到語音材料上，可以選擇不使用相應的文字信息，只呈現語音刺激。不建議只使用文字材料，因為文字材料相較於語音材料是第二性的，並且文字的理解度是和教育水平相

關的，並不是一個全民性的語言能力，對文字的理解則又進入到閱讀理解的範疇中，因此本文認為不適用文字材料。語音材料的長短可以根據內容的長短分成字、詞、句、段等不同的類型，之前在理解度的性質一節討論過對於理解度測試最好的單位應該是句子這一層級，既可以完整地表達意思又可以盡量保持一個較短的時長，不至於被被試在實驗過程中快速遺忘。除了句子之外，詞也是一個相對較好的單位，詞彙也是語言中基本的構成單位，對詞彙的理解也是構成對更複雜結構理解的基礎，因此詞彙單位也可以作為理解度測試中可以選擇的材料單位。

其他單位並不是完全不可以，但是各自有各自的優缺點，除非單獨研究某個問題，一般不建議使用其他的單位。在漢語中，對一個單字的理解通常有太多意義，另外同音字太多導致很難辨別具體是哪個字。語篇的內容過多，篇幅過長，通常能夠理解其大致意思，並不能做到每個細節都理解到，這樣對理解度的計算就產生了很大程度上的灰色地帶，從而造成不準確性。

2.5.2 實驗方法

實驗根據評價機制應該分為主觀評價和客觀評價兩種方式。主觀評價是被試者接受完實驗的刺激後，根據自己主觀的感受，回答實驗者設計的問題，通常多是通過萊克特等級量表格式。例如Murray J.Munro, Tracey M.Derwing & Susan L.Morton (2006) 讓被試在聽完材料後，根據剛才聽到的錄音就口音感 (Accentedness) 理解度 (Intelligibility) 和掌握度 (Comprehensibility) 幾個方面進行打分。用主觀的打分的方式得到的結果，與客觀實驗的結果相結合相互支撐和印證。

客觀的評價方法是通過一套成型的測試對被試進行實驗，通過對被試的實驗數據的收集得到最終的結果。實驗的方法可以有很多種類型，有轉寫法，完形填空，單項選擇，判斷正誤，選擇圖形等等方法。根據被試反饋的類型可以分為兩類，一類是轉寫類，一類是選擇類。轉寫類是要求被試將聽到的內容在轉寫成文字寫下來，轉寫的內容可以是關鍵詞，或者是整句話，還可以是提煉觀點。本文認為轉寫法並不能實際反映理解度，表面上看，轉寫法是將所有的要素都表現了出來，但是實際上因為在轉寫的過程中牽扯到把聽覺材料轉化為文字，是一個重新編碼輸出的過程，該過程需要一個較長的反應時間和處理流程。因此在轉寫的過程中，因為轉寫的速度趕不上語音的速度，所以勢必會造成聽過後面的內容忘記了前面的，有可能需要播放兩至三遍才能夠讓被試將文字轉寫完畢，這樣實際上已經不能夠完全反應出實際的理解度了，因為在實際的交談中，如果對方有耐心用同一句話講三遍，基本上已經可以得出結論是對方不能夠聽懂說話者講的話。另外，文字的轉寫並不能夠說明就一定是理解的產物，有可能只是將聽到的內容轉化下來的一個過程。這個過程可以很好地反映出被試的「聽到」，但是並不能很好地反映出被試的「聽懂」。這個過程類似於手機的聽寫功能，手機可以幾乎完美地將語音信號轉化為文字信號，但是在此過程中，手機並不能夠完全理解這些話的內容。尤其是對於使用拼音文字的語言，將聲音信號用字母記錄下來並不是非常困難的事情，對於一個外族人，即便完全聽不懂內容，只要熟悉拼寫規則和讀音規則也是能夠將聲音信號轉寫下來。除此之外，轉寫法還有一個可能存在的問題就是忽略了一個句子中信息是有層級性的，有些信息是關鍵的，而有些信息是次要的。理解有些關鍵信息就可以理解大部分內容，而那些次要信息起著一些結構性的作用，並不能

夠對理解核心內容產生有效幫助。將全部內容轉寫出來實際上就是忽略了信息的層級性。假如聽到的句子是：「今天星期三。」然而轉寫下來的句子是「今天星期天」。這樣表面上看只有一個字出現了錯誤，那麼正確率是80%。然而重點的新信息是星期三，而不是星期天。這樣僅僅是20%的錯誤就造成了句義100%的偏差。因此本文認為轉寫類的測試方法並不適理解度測試。

另外一類是選擇類的任務類型，其中包括單項選擇、判斷正誤和選擇圖片等方法。這幾種方法總本質上都是讓被試在理解的前提下做出正確的選擇和判斷。之所以把判斷正誤這類的任務也納入到選擇類當中，是因為判斷正誤從本質上也是在正確和錯誤的兩個選項中選擇出一個正確的選項。如果需要測量被試的反應時間，那麼建議不要設置過多的干擾選項，否則被試會花費更多的時間在理解選項上，而並不能反映出真實的理解時間。選擇類的任務類型的優勢在於雖然不能夠將一個句子中的所有內容全部體現出來，但是可以有側重地體現句子的意思，找出句子中的關鍵信息點，對其進行測試從而測試句子的理解度。在這個過程中可能會遺漏一些信息點，往往這種信息點在不影響句子整體意思的情況下，可以暫時忽略不計。其次，選擇類任務還有著操作簡便、易與統計的優點。測試過程中，被試不需要經過太長時間的培訓就可以掌握測試的操作，更容易上手實驗，這樣得到的數據也更加真實有效。選擇類題目，相比於轉寫類題目，結果更加直觀，也更加容易統計。最後，選擇類的任務不需要被試對信號進行二次加工，也不需要將聲音信號轉化為另外一種形式的信號，而是直接在大腦中進行判斷，然後做出反應即可。從時間上看更能夠接近於真實的理解時間。

選擇類的任務類型也有一定的劣勢，需要在實驗設計中盡量避免，或者在

實際操作過程中注意觀察。由於是選擇題目，就會存在著並沒有完全理解，甚至是理解錯誤卻能夠選對的情況，這種情況有可能是被試隨機選擇的，往往在被試不太理解測試材料的情況下，為了作答強行給出一個答案，這樣可能會造成一些誤差在其中。另外被試也可能沒有完全聽懂，而是通過選項的提示信息經過大腦分析後給出一個答案，這時的正確率與實際的理解度也會出現一些誤差。這種誤差常常出現在兩種語言相近的情況下，例如北方官話和粵方言都屬於漢語的方言，二者之間有著非常大的差距，但是相比於其他印歐語系的語言又有著相當密切的聯繫。對於這類的情況就需要在實驗設計中增加更多的測量值，來消除上述情況所帶來的誤差。

本文認為引入反應時間參數可以很好的消除選擇類題型所帶來的誤差，反應時間可以很好的反映出被試對於實驗材料的熟悉度和理解度。如果能夠很順利的理解，那麼就可以很輕鬆快速的完成題目；如果是因為不熟悉隨機做出的選擇，或者是經過思考後得出的答案，其時間必然要慢與正常作答的時間不同。反應時間可以對測試的正確率起到補充修正的作用，以求得到更加接近於真實的理解度數值，具體關於加入反應時間的論述請詳見第二章。

以上是關於理解度測試的測試題型的討論，從題型的角度上看，理解度測試與普通的聽力測試十分相似，並且從形式上看，二者也是相似的。不過本文需要指出理解度測試與聽力測試有相似指出，但並不能簡單地等同於聽力測試。首先是因為聽力測試通常是對學生的第二語言進行的測試，是對學習成果的檢驗和測評。最終結果是學生的學習成績。而理解度測試並不是檢驗學生的學習成果，而是要通過被試的語言能力得到語言之間的相互理解度。聽力測試的成績對於個人來說是有一定的評價作用，可以評判一個學生的學習水平和個

人的語言能力。而理解度測試並不是針對於個人的語言水平和學習能力，與單獨個體的努力並沒有太大的關係。而是通過大量的樣本得到一個常量，這個常量代表的是整個群體所對應的語言對另外一種語言的理解度。從結果上看是一個宏觀的水平，沒有優劣之分，是一個客觀自然的常量。並不能利用理解度的分數去進行價值衡量。雖然二者之間擁有着類似的測試模式和題型，但是二者之間從理論背景到結果闡釋是有著本質性差異的，故而不能夠將二者簡單的等同起來。

2.5.3 實驗對象

實驗對象的選擇也是保證實驗數據真實有效的重要環節。挑選合適的測試對象需要遵守以下幾個原則：首先是廣泛性，廣泛性包含兩方面含義，一方面是範圍的廣泛，另外一方面是數量的廣泛。範圍的廣泛是要求實驗篩選被試的時候包含盡量多的類型，充分考慮到人群的各個社會屬性的群組。比如性別因素、年齡因素、國籍祖籍、受教育因素等背景信息，同時也包括被試的語言背景信息，例如第一語言，家庭語言，外語能力等信息。因此也需要對被試進行細緻的背景調查，然後對其進行分組。為了保證數據的代表性，所以在時間條件和物質條件允許的情況下，可以盡量保證每個分組都能夠有至少多於一個人。這樣可以有效提高數據的準確性，從而保證數量的廣泛，因為只有足夠多的數量才能夠更好的吸收單個個體的誤差對整體的影響，用隨機誤差來消除個別的誤差使得數據整體保持一個穩定的範圍。除此之外同質性也是需要注意的一個重要原則，在分組的過程中，要盡量保證每一個組別下的被試保證同質性，這樣可以更好地控制組與組之間的變量，從而有效的得到變量對最終結果的影響。分組並不是為了分組而分組，分組的數量也不是越多越好，而是要根

據實際需要和實驗目的進行設計。如果A組和B組之間的數據並沒有統計學上的差異，那麼可以合併兩個分組。從理論上看就是A組和B組的區別性特徵並沒有對最終結果產生足夠大的影響，所以就可以忽略二者之間的這個差異，從而可以更集中地表現出其他組間差異。最後要保證實驗對象的隨機性，被試的選擇不能夠依照人為因素進行篩選，而是要通過被試的個人屬性進行篩選，充分保證了樣本的隨機性能夠最大程度上避免人為因素對樣本本身的主觀控制，同時可以利用樣本的隨機性來抵消掉個別樣本偏離主體的現象。綜上在進行測試的過程中，根據以上的原則篩選出適合的被試有助於更準確地得到實驗數據，保證數據的真實性和有效性。

2.5.4 測試工具

傳統的測試多是使用調查問卷的方式，讓被試聽過錄音後在調查問卷上做出選擇、判斷或者轉寫等任務。這種測試的優勢在於（1）操作簡單，被試不需要特意的熟悉實驗方法就可以非常快速地進入到測試中去。不需要前期的練習，也不需要複雜的儀器和準備工作，實驗過程非常快速高效。由於測試操作簡單，對實驗的操作人員的要求並不高，通過簡單的訓練就可以培訓出很多實驗操作人員，從而可以多線進行測試，快速取得大量的數據。（2）對測試設備的要求並不高，傳統測試使用問卷進行測試，問卷的成本非常低可以近似忽略不計，而且問卷易於複製。問卷的設計相比較於測試程序的设计更加容易上手。（3）可多人同時進行，使用問卷進行測試可以將被試者集中在教室內或實驗室內同時進行實驗。這樣可以在較短的時間內同時獲得較多份問卷。然而這種傳統的測試方法也存在著一些劣勢。（1）不能夠對每個實驗個體的測試細節進行把控，以傳統的形式進行測試，通過問卷獲得被試的反饋是極其有限

的，僅僅局限於正確率，並不能反映出被試具體的反應時間和其他心理或生理活動。這樣得到的結果可能與真實的情況產生偏差，需要更多的參數來完善最終的結果。（2）增加干擾因素，由於對實驗環境的要求並不高，並且可以集中測試，因此在播放錄音的過程中，不能保證每個被試都能夠接收到同樣質量的語音信號。很難控制噪音在測試過程中對錄音的影響。（3）實驗數據的統計需要人力統計並且數字化，面對大量的數據和樣本的過程中，極有可能產生人為錯誤，影響了最終的結果。

經過改良後的測試就會採用電腦以及其他儀器來收集被試的相關信息，如果是一般的刺激反應實驗會用到心理學的實驗軟件進行編程，利用軟件來收集反應數據；高階的測試可以加入眼動儀，利用眼睛注視點的變化軌跡來反映出大腦的處理過程；還可以配合腦電圖掃描儀（EEG）等精密儀器實時監控被試在測試過程中的大腦電位變化，從而可以更加直觀地認識大腦的工作機制。這種測試方法的優勢在於（1）可以採集到更加豐富的實驗數據，更加全面的觀測到被試在測試過程中的各項心理和生理反應。這使得最終計算出的理解度數值更加全面並且更加具有可信度。（2）利用測試軟件編程生成特定的測試系統，可以保證測試中的流程不會受人為因素而干擾，能夠保證所有人的測試程序都是一致的。同時也可以利用測試程序中的隨機順序設置來呈現刺激，進一步避免固定的順序性呈現方式給被試帶來的學習效應或其他影響因素。（3）便於數據的回收與統計，測試結束後軟件會自動生成測試結果，這個過程並不需要人力判斷和統計，從過程中減少了人為因素對測試結果的影響。直接生成的電子表格也更方便進行後續的統計和分析。然而改良後的測試方法也存在著一些劣勢。（1）實驗成本增加，無論是時間成本還是物質成本都大大增加。

實驗所需的軟件、儀器都需要相當高的成本。在非常大的程度上限制了實驗的進行，為實驗者設立了相當高的門檻和障礙。並且由於軟件、儀器等設備的數量有限，並不能夠在同一時間里進行大量的測試，通常情況下需要一對一進行測試，並且測試時間要長於傳統測試方法。若利用腦電圖掃描儀，則還需要測試前為被試洗頭，帶電極帽，注射導電凝膠等複雜的前期準備工作。（2）培訓實驗人員難度提高，由於使用了高級電子設備，需要自行編程生成測試程序，還需要調試相關設備，對實驗操作人員的水平要求大大提高，簡單的培訓已經不能夠滿足實驗規程的需求。需要經過長時間培訓才能夠掌握實驗的相關操作流程，不利於廣泛開展大規模的測試。

傳統方法和改良方法各有利弊，從技術進步的角度上看，改良的方法代表著進步的方向，雖然還存在著諸多不便，其目的是為了得到更加豐富和準確的數據。所以從長遠來看，改良的方法會逐漸成為主流，最終取代傳統的方法。就目前階段，既不可以不顧技術發展學術需要依舊使用傳統方法進行測試，也不能夠陷入高科技的設備中不可自拔，需要在二者之間找到平衡點，利用最經濟最可操作的方式方法盡可能地獲取更多的有效數據。在實踐中不斷完善改進測試方法獲得更有意義的成果。

2.6 理解度常模參照標準

2.6.1 常模參照與標準參照

理解度的結果是通過一個個單一的測試中得來的，單一的結果並不能夠反映整體的狀況，因此勢必要通過擴大樣本數量來得到更加有代表性的結果，通過大樣本來抵消掉實驗中存在的誤差。樣本擴大後的數據需要通過統計學的手段對其進行再次處理，利用處理後的數據就可以建立有意義的理論模型，從而

形成標準。測驗的原始分數本身沒有多大的意義，必須另外設立一種明確規定的、統一的參照框架來解釋測驗分數（王佶旻 2011）。本文認為需要利用語言測試學的理論來建立科學化的標準模型。利用語言測試成熟的理論來建立和完善理解度的標準體系。語言測試是應用語言學的一門重要分支學科，現代測量統計意義上的語言測試是一門跨學科的綜合性科學（黃銳 2012：8）。語言測試學已經形成其完整的理論和實踐體系，理解度測試可以充分借鑒其研究成果，從而將理解度測試形成更具有操作性的測試體系，並且形成標準，這樣可以更好地應用於實踐當中。

目前語言測試的主要參照模型有兩種：常模參照（Norm-referenced Test）與標準參照（Criterion-referenced Test）。所謂常模（Norm）是指標準化樣本中測驗的分數分布（安妮·安娜斯塔西、蘇珊娜·厄比娜 1997：63），常模參照測試是通過測試將應試者個體的能力發展水平與某一個特定群體（測試對象總體）的能力發展水平進行比較。常模參照測試可以使應試者的能力發展水平得到最大程度的區分。（黃銳 2012：59）常模參照測驗一般都是大規模的標準化測驗，多用於醫學、學業成就考試及各種能力方面的水平性考試。常模指的是通過採集標準化樣本，並經過測試得到測試數據，根據這些數據的分布型態，進行一定的統計處理後，建立的一套關於應試者水平的評價參照標準。因此常模就是根據標準化樣本在測試上的測試結果，在經過一定的統計技術處理後，建立起來的具有參照點和單位的測試結果評價參照系統（黃銳 2012：59）。常模參照測試的分數本身沒有確切的含義，不能說明被試達到了什麼樣的標準或者水平。這個分數是要和其他所有全部已經參加測試的人得到的分數進行比較，從而能夠得到被試在整體中的地位。常模是一個相對的數量，隨著

時間和空間的不同而變化。任何常模，不管怎樣，都限於導出常模的特定的常模總體。心理與教育測驗的常模不是絕對的、普遍的和永久的。他們只是表示組成標準化樣本的個體的測驗成績（安妮·安娜斯塔西、蘇珊娜·厄比娜 1997）。

建立常模的過程大致有九個步驟：

- (1) 確定目標團體；(2) 確定這個常模的個關鍵的統計量；(3) 確定個統計量能容忍的樣本誤差；(4) 設計一個抽樣方案；(5) 估計出最小樣本數；(6) 抽取樣本實施測量；(7) 計算出各種數據及其標準誤差；(8) 確定常模分數的類型，並制定常模分數的轉換表；(9) 把建立常模的程序寫成材料（王佶旻 2011）。
- (2) 一般來說，常模可以分為發展常模和團體內常模。發展常模表示個體在正常發展線路上處於怎樣的水平，例如智力測試的比奈-西蒙智力量表，團體內常模是根據最可比較的標準化樣本的成績來評價當時成績的。目前絕大多數的標準化測驗都會提供某種形式的團體內常模（王佶旻 2011）。托福（TOEFL）考試就是典型的常模參照測試。

標準參照測試是根據特定的標準直接對測試的結果進行分級。參照絕對標準對考試分數加以解釋，標準參照測驗主要用於鑒定被試是否達到了某個設定的等級。因此廣泛引用於例如大學英語四六級考試或普通話測試中。常模參照與標準參照測驗的根本區別在於解釋分數的參照點不同，常模參照解釋分數的參照點會依賴於整體水平。標準參照測驗的參照點是預先確定的目標或任務，

不與其他測試者發生關係。這裡的「標準」是指編制項目和考試、解釋和報告分數時所依據的特定的知識領域或範疇。標準需要嚴格明確和界定，它是不依賴於樣本的。

Bechman (1990:75) 認為常模參照測驗和標準參照測驗的基本區別有兩個，一個是設計和開發的過程不同，二是分數的產生和解釋不同。彭康洲 (2013:7) 歸納出二者之間的區別，他認為常模參照語言測試的測量類型是一般語言能力，是個體相對於群體位置的相對解釋類型，分數會呈正態分布，並且主要用於測量的目的。標準參照語言測試的測量類型基於課程目標的具體語言點，是個體成績和預先的學習目標作比較得出的，分數會呈負偏態分布，主要用於教學目的。

2.6.2 理解度常模參照標準

本文認為理解度測驗應該建立在常模參照的理論框架下，並且使用常模參照標準。其原因主要有以下幾點：

- 1、從理解度的性質出發，某種特定的語言對於特定的人群在一段時間內的理解度應該是一個恆定的範圍，並且形成正態分布，這也符合常模的屬性。理解度若低於這個範圍，則說明被試對於該語言的熟練程度還沒能達到基本水平。若遠遠低於這個範圍，那麼就有可能說明被試完全不懂這一門語言。因此每一個參與理解度測試的被試的成績是和所有已經參加過測試的被試的成績進行比對，得到他自身的一個排名，由此可以知道他在整個群體中是處於一個什麼地位。事實上是每個人和其他的人進行的比較，而他的成績也會成為下一個人比較的對象。正如一個正常人血液內的白細胞數量是 $(3.5\sim9.5) \times 10^9$

9/L，如果一個人的白細胞數量低於 $3.5 \times 10^9/L$ ，那麼他可能患有白細胞減少症，對應的是貧血等疾病；如果一個人的白細胞數量高於 $9.5 \times 10^9/L$ ，那麼他可能患有一些炎症。

2、語言的理解度是一種語言能力。母語者對語母語的理解度是一種從屬於語言能力的一種必備的能力，並不是教育的結果。對語正常的母語者來說，只要他習得了這門語言，那麼他的理解度就一定是在正常的範圍內。不會因為教育程度，努力程度或其他因素影響，也不存在需要努力練習就能夠提高一個等級的可能性。所以這種語言能力是一個言語社區內部所有人共享的能力。這種能力的量化範圍也必定在一個固定的區間內。就如白細胞的數量是恆定的，不會因為受教育程度高，或者早餐吃得豐盛就會超過這個範圍。正常健康的個體都會保持在這個區間內，低於或者高於都是不正常的。

3、不選擇標準參照的理由：

標準參照測試是認為建立起一套標準，達到一定標準，那麼就拿到某個等級。事實上是被試者和標準之間的比較，是與其他人沒有關係的。也就是說，假設標準是60分，得到60分就是合格，而59分就是不合格，哪怕平均分只有30分。這並不符理解度測試的性質。理解度測試並不是普通話測試，聽力考試，更不是大學英語四六級考試。它是具有普遍意義的一種測試，可以利用這種測試去衡量不同語言之間的關係；而標準參照則是使用多把尺子去測量，測量出來的結果並不能相互比較，就例如，大學英語過了四級的評價體系和普通話的評價體系是兩個完全不同的體系。

標準的建立是理解度理論的必經之路，標準的建立就是建立理解度的標

尺。可以通過理解度去衡量語言間的距離，衡量語言間的親疏關係，亦或是衡量一個人的語言水平等等。選擇合適的標尺模式也是值得思考的問題，否則差之毫釐謬以千里。這把標尺的刻度並不能依靠人為划定，而是需要通過實驗得到的事實來進行划定。

2.7 理解度的應用

一種理論如果僅僅停留在理論，而不能有實際的應用的話，那麼這種理論的意義將大打折扣。一個優秀的理論不僅擁有深遠的理論意義，還需要對具體的實踐有指導意義。理解度理論不僅需要理論支撐，同時更加需要實踐來對理論加以補充和輔助，二者之間相輔相成並且缺一不可。其他理論往往是理論先行，理論指導實踐。對於理解度來說，恰恰是實踐先行，而理論略顯薄弱。因為理解度的概念容易理解，實踐操作又容易上手，所以導致在沒有理論指導下就進行了實踐操作。正是因為缺乏了直接的理論指導，所以導致理解度理論的實踐應用並不十分成體系，目前是呈現出一種多點開花的狀態，各個領域各個方面都涉及一些理解度的概念和實踐，卻沒能進入一個應有的框架和體系。不過也為理論的建設和今後的實踐奠定了基礎划定了範圍。本文總結了幾項比較有代表性的目前國內外的理解度實踐成果，通過對其進行綜述分析後，可以為日後實踐提出一些借鑒方案。

在國內的理解度研究中，有使用到「溝通度」，「可理解度」等提法，概念上大同小異，總體屬於理解度的理論範疇，主要處理方言之間的理解度問題，也有帶口音的第二語言的理解度等問題。

黃行（2011）提到語言溝通度的計算方法，他認為：「語言溝通度的測試

方法一是測算兩種或幾種方言之間一定數量常用詞的語音語義近似度,二是使用某一方言話語故事的錄音去測試另一方言聽話人的理解率來測算方言之間的可溝通程度,即用不同方言人主觀聽辨測試來確定方言的溝通程度。」作者認為方言之間的可溝通度主要是指先天性溝通度,並且十分重要的是作者給出了近似詞比例60%和先天溝通度80%的臨界值,近似詞比例低於60%則可以不視為同一種語言,也不再考慮先天溝通度的情況;近似詞比例高於60%的情況下再用先天溝通度的80%來區分,高於80%則說明先天溝通度高,幾種非常相似的言語變體可以視為相似方言。低於80%則說明先天性溝通度低,幾種非常相似的言語變體可以視為不相似方言或不同語言。該文章根據調查給出近似詞比例和先天溝通的標準來加以確定方言之間的關係非常具有啟發性,同時文章也提到我國民族語言學界目前基本還沒有採用溝通度的方法來測試語言或方言之間的可通話程度。僅有石林(2010)使用了溝通度的方法對侗語方言土語間的理解度進行了調查。作者對侗語的11個點進行了110人次的調查,通過溝通度的測試方法,作者發現南侗語講話人對北侗語的理解度為51.2%,北侗語講話人對南侗語的理解度為51.3%。文章中引述世界少數民族語文研究院的經驗,語言或方言間的理解度如果在70%以上,二者間的通話基本無問題,如果在70%以下,二者的通話將困難。這與我們此次對侗語調查得出的結果相吻合。總之,理解度越高,通話越容易,反之,理解度越低,通話越困難。在文章中作者通過自己的測試也基本符合這一標準。

何家勇、李珊、何雲亭(2010)文章對90名學生的英語錄音進行可理解度實驗,發現非英語母語者對背景有很強的依賴性,在零語境的條件下的語音可理解度不到0.6。文章還分析了造成錯誤的原因,認為語調、節奏、停頓等此

標準特徵有時容易對聽者產生誤導，在理解的時候需要付出更大的努力才能理解成功。文章實驗的方法是先找到學生進行錄音，然後請被試前來實驗，實驗的手段是聽寫，聽者可以最多重複聽三遍錄音。

高超、文秋芳（2012）以60名在英國學習的英語本族語者和非本族語者為研究對象，探討了中國語境中本土化英語的可理解和可接受度。研究發現中國本土化特徵的英語具有較高的可理解性和可接受性。作者認為可理解度不是一次性的、簡單的概念，而是多層次的、受多重因素影響的、個體差異很明顯的動態概念，之所以重視個體差異，是因為交際過程常常是一對一的交流，假如有一方沒有聽懂，就會導致交際行為失敗。文章選用的語料庫是CENAC（China 's English News Articles Corpus）從中挑選文本素材，然後邀請六十名在英國讀書的英國本地或留學的大學生共同填寫問卷，問卷使用萊克特等級量表格式。結果對於測試材料的可理解度和可接受度的均值是4.22和4.16。通過對數據的分析，英語本族語者和非英語本族者的評價有顯著差異。

方瑞、趙玉婷、陳小潔（2017）收集50名來自成都雙流機場主要航線乘務人員的英語播報錄音，並進行了可理解度實驗。實驗的被試是25名來自英語為母語的國家的留學生或外籍教師。請這25名被試根據錄音進行轉寫。50組錄音每位被試抽取兩個進行實驗，每個錄音聽兩遍。經測試得到總體的理解度為0.89，標準差為0.07，總體可理解度較高。文章還分析了造成錄音不被理解的原因，並分析了錯誤類型。該實驗中讓被試進行轉寫，通過轉寫的答案進行評分，得到有效數據。

Murray J.Munro, Tracey M.Derwing & Susan L.Morton（2006）指出在理解

或評估外來口音時，聽者不僅受到語言材料本身的內容的影響，還會受到本身的語言背景和他處理不同語言變體經驗的影響。鑒於後者的對理解的影響，該文希望得到不同背景的聽者是否會對第二語言的材料有著同樣的反應。該研究中作者請到母語為廣東話、日語、普通話和英語的聽者來給背景是廣東話、日語、波蘭語和西班牙語的發音人所講的英語打分。得到的結論是無論母語的背景差異，聽者對於理解度和理解度以及口音打分都表現出較高的相關性。雖然有一些組間差異，但是所有的聽者對於判斷所有48位發音人中哪一位是最難聽懂的，哪一位是最好聽的是相同的，且組間差異小。與前人研究相同的是，聽者對與自己相同背景的口音沒有顯示出特別的理解度優勢。最終文章支持了：儘管聽者是來自不同的語言背景，但是語言自身的性質才是決定第二語言如何被感知的一個重要的因素這一觀點。該實驗是通過轉寫的方式記錄聽者對於語音材料的理解度，此外還通過一個問卷來測試聽者主觀的對與聽到材料的理解度，理解度和口音，作為對實驗結果的補充。

Tang Chaoju & Vincent J. van Heuven (2009) 證明互通度是一個測量方言間距離的有效的方式。文章測試了15種漢語方言詞和句子的互通度。對單詞的測試是通過給詞彙分類，例如身體，植物，動物等等，句子的測試是讓受試者把一個目標詞彙翻譯成自己的方言。他們共獲得了47250個數據，並得到結論：互通度可以被意見收集調查和實驗結果決定，這兩種方式得到的結果都與客觀實驗的結果顯著性相關。在句子層面的互通度實驗反映出來的漢語方言的分類情況要優於意見收集調查的結果。

後逐步擴大其在語言學視野下的應用，其主要研究方面在於對帶有口音的英語進行理解度的測試，例如：Brennan, E. M., & Brennan, J. S. (1981),

Wardrip-Fruin, C., & Constantinou, A. (1993), Yang, H. T. (2005), Pickering, L. (2006), Hardman, J. (2010)。

Hayes-Harb, R., & Watzinger-Tharp, J. (2012) 這篇文章探討了口音與理解度之間的關係。文章認為母語者對本民族語言的外族習得者會形成一種語言認識，這種認識是如何與口音與可理解兩個維度上的判斷相互作用的。分析表明，持有不同語言認識的聽話人對非本族言語並不一定會產生不同的判斷。然而，數據指出德國母語者對第二語言習得的發音的評估有相當大的差異。這篇文章使用了兩種實驗任務方式，一種是口音判斷任務 (Accentedness Judgment Task)，另一種是聽覺詞彙識別任務 (Auditory Word Identification Task)。口音判斷任務是在口音判斷任務的每個試驗中，被試連續聽到由單個發音人產生的五個目標句子。每個聽眾以不同的隨機順序聽取了15組不同發音人的句子。在使用DMDX演示軟件 (Forster & Forster, 2003) 的帶聲音的展台內通過耳機提供聽覺刺激，聽眾在紙張評分表上標記他們的反應。根據每個說話者，被試通過六度評價法 (1代表發音人聽起來不像德語的母語者，6代表發音人聽起來像是德語的母語人者) 回答以下幾個問題：

你認為這位發音人的德語：

- (1) 在講話的速度方面處於什麼水平？
- (2) 在講話的流暢性方面處於什麼水平 (這位發音人說話是否順暢或停頓或暫停很多)？
- (3) 在發出德語的發音能力方面處於什麼水平？

(4) 在言語的節奏和韻律方面處於什麼水平？

(5) 總體處於什麼水平？

聽覺詞彙識別任務是用來評估發音人的語音的可理解度。使用對聽到的詞彙的識別正確率作為這項實驗的可理解度數值。被試者會聽到單獨的德語詞彙，並被要求在兩個備選答案中選擇一出一個。呈現的單詞是從前一個任務中的五個句子中拼接出來的，備選答案是之間構成最小對立對，所使用的都是德語中最常見的最小對立對。

文章最後認為儘管非母語者非常努力「克服」他們的口音，但其實口音本身並不一定構成可理解性的障礙。事實上，口音僅僅是說話者所表現出來的不同於母語者的特徵。無論母語者還是非母語者，雙方在交流的過程中都會承擔交際的負擔，為了達到交際目的雙方都是平等地適應對方的口音，意識到這一點而對於學習者來說也是有益的。

Nejjari, W., Gerritsen, M., Van der Haagen, M., & Korzilius, H. (2012),
Huong, T. T. (2017), 他們的實驗對象多是不同地區的帶有口音的說話人，把他們的語音錄下來，讓英語母語者進行判斷的實驗。其中的被試應該分成兩類，一類是以英語為母語的，只不過生活的地方並不是在英國或美國本土；另外還有一類是以英語為第二語言的，有的研究是越南的英語、中國的英語以及帶有西班牙、墨西哥等地口音的英語等等，例如：Munro, M. J., & Derwing, T. M. (1995), Munro, M. J., Derwing, T. M., & Morton, S. L. (2006) 等。

以上可見國內眾多研究還都處於嘗試和試探的起步階段，也涉及了較多方面。有許多值得借鑒的地方：首先在方言研究中，對於是否能夠共同的尺度，

石林（2010）給出了自己的標準，通過實驗和經驗都能給出一個合理的取值，是今後對理解度標準問題的探討有著極大的借鑒作用。其次，對於區分方言間的先天可理解度和後天的可理解度也是非常有啟發性的一種思路，這樣可以將音系和詞彙差異的研究與理解度研究有機的結合在一起，形成一個非常有創建性的體系。再次，高超、文秋芳（2012）使用問卷的方式來調查理解度和可接受度，文章採用了萊克特等級量表格式，提高了問卷的規範性。最後對於機場英語播報的理解度也是很好地將理解度的大致思路運用到了生活中去，極大地拓展了理解度的應用範圍，確實，一項理論從書本走向應用都要有一個適配的過程，這種擴展性地嘗試也是非常具有創建性的，由此可以引發一些列應用理解度解決實際問題的研究。在外國研究中，能夠見到對理解度的應用形成了一定的規模，會對集中一類的問題進行討論，不過是針對零散的問題進行討論，未能夠將每個研究成果綜合形成理論化體系化的成果。另外，其研究目前還是主要以英語為核心的應用偏多，以中文或其他語言為核心的理解度應用尚未能形成規模。

在眾多可借鑒的內容背後，依舊可以見到許多值得改進的地方，需要在今後的理論和實踐中加以改善。

（1）標準的制定要以定量的方法來制定，不能以定性的方法大約給出一個標準，這樣給出的標準並沒有實際的引用價值，並不能成為之後引用的一個真實數據，只能作為大致參考。

（2）僅靠問卷調查得到的結果並不能精確反應理解度中度的概念，只能給出一個大概的值。如果可以不僅依靠問卷再綜合理解度測試可到的結果會更加

精准，可以更好地反應真實的理解度數值。

(3) 實驗中重複播放實驗語音材料並不能反映出真實的對話狀態。並且播放兩遍得到的理解度應該是兩次播放的和，而並不是單次的理解度。另外，採用轉寫的方式來收集理解度的結果也值得考量，因為文字是第二性的，語音是第一性的，轉寫的過程是大腦將語音重新編碼的過程，可能會造成又一次信息的干擾或丟失。既有可能在轉寫時出現誤差從而影響實驗結果。

2.8 理解度的意義

理解度理論是一個綜合利用了實驗語言學、心理語言學、語言測試理論、二語習得理論、統計學等幾大學科的應用型理論，其結果也可以在各個學科領域內受益。理解度理論是綜合多個學科的理論，同時也是為多個學科搭建起了一個理論平台，在這個平台上，將各個學科的理論融會貫通，最大程度上發揮各個學科的優勢。本節從理解度理論的功能角度出發，探討理解度理論能夠為其他理論貢獻怎樣的功能，將理解度理論作為眾多理論的工具，實際解決理論上的相關問題，建立相應的標準，簡化學習過程。

2.8.1 建立標準

標準是任何一個行業學科都需要討論的一個問題，國家層面會建立相應的國家標準，行業內部會制定行業標準，學科內部也有成型的研究標準，寫作標準等等。秦始皇的功績之一就是將戰國末年各國混亂的標準統一建立新的標準，這個標準包括了文字的標準，度量衡的標準等。正式由於標準的建立，才有之後文化經濟的進步。秦代兵馬俑的燒制都是採用標準化的方式進行才會形成如此大的規模。現代科學技術的進步也得益於各項標準的確立，現代工業的

發展也是標準化生產、流水線生產的產物。不符合標準的產品都不能夠出廠，符合標準的產品才能夠流通到市場中去。標準就像是一把尺子，用這把尺子衡量出不同事物的屬性。在同一標準的情況下，才能夠將兩樣不同的事物進行比較，也就是用標準作為聯繫不同事物之間的共同的參照物。標準本身是虛的，而被測量的事物卻是實實在在地客觀存在。甲乙兩根「木頭」長短不一，利用「尺子」來進行測量，木頭甲長「兩米」，木頭乙「三米」，木頭乙比木頭甲「長」。這裡的「木頭」「尺子」「兩米、三米」分別象徵著不同類型的事物，木頭是被測量物，是一個實體，具有客觀性，可測量性。尺子是標準的客觀載體，是標準的物質呈現。具體是什麼樣的物質並不重要，尺子究竟是木頭做的，還是鐵的、還是銅的都不重要，唯一不變的是他們上面的刻度即標準。這把測量長度的「尺子」可以是通常意義上的尺子，也可以是一段繩子，也可以是通過計算光通過的時間等等，手段並不是唯一的，但是可以選擇最經濟的，最簡單的，最適合的方式。最後「兩米和三米」是測量的結果，是通過和標準比較後得到的相應的結果，這個結果對應的事物是客觀的，但是數值並不是一定的，換一種單位即換一種標準就會得到其他的數值。數值的改變並不會影響事物本身的屬性，雖然兩米可以換算成二十分米或者二百釐米，但是無論怎麼換算，這些數值對應的內容都是等價的，且不以人的意志為轉移，也不會因為測量者的變化而變化。最後二者比較的結果乙比甲「長」也是客觀的，肉眼都可以觀測到的，而具體長多少，那麼就只有通過標準，才能知道二者之間是相差一米。由此一個簡單的事實可以推廣到許多相關的問題上，不過需要注意到其中隱含的前提和預設，也就是首先這個事物的客觀屬性，高矮、長短、胖瘦都是可以通過實際測量，然而美醜、好惡、愛恨等屬性並不能通過這種模

式進行判斷，在一個人的眼睛里，可以區分張三美，李四醜，但是換到另外一個人的審美，有可能發生變化。當然有些選美冠軍也是公認的美，這種公認性似乎是賦予了美的客觀屬性，然而並不能說這個選美冠軍比普通的某個女孩子美多少。也就是說缺乏了通用的對美的客觀評判標準才會導致這樣的問題。

米是長度單位，一根木棍的長度是測量木棍的頂端到木棍的底端之間的距離，就是測量一個點到另一個點的距離。那麼一個語言和另一個語言之間的距離應該如何計算，本文認為理解度理論就是語言和語言之間的「尺子」。最早 Comrie, B. (1987) 提出使用理解度作為判斷兩種語言變體是否是一種語言的兩種方言還是兩種語言。語言系統是一個客觀實在的系統，語言的物質存在方式可以是文字，也可以是錄音、錄像等介質。兩個語言系統之間存在著或大或小的差異，單純從感知上可以判斷出A語言和B語言之間的差異較大，B語言和C語言之間的差異較小。就像三條長短不一的木棒，甲3米，乙2米，丙1米，可以通過直觀的視覺來測量甲比乙長，乙比丙長，如果沒有尺子，就沒辦法知道究竟甲比乙長多少，乙比丙長多少。往往事情沒有那麼直觀，同樣是三條木棒，甲0.999米，乙1米，丙1.001米。那麼僅僅靠直觀的視覺已經不容易分辨出究竟誰比較長，似乎也能有些感覺，但並不能十分有把握。語言也是相同的，同樣是三種不同的語言，英語、法語、漢語，似乎很好區分三者之間的差異，擁有一些語言學常識的人就可以知道英法的關係要更近，漢語與其他兩種語言的關係更遠。然而同是漢語中的閩方言、粵方言、贛方言三者之間究竟是什麼樣的關係，恐怕就不是那麼簡單就可以說清的。

理解度就是這樣的一把尺子，用來客觀地衡量語言和語言之間的距離，以理解度作為衡量語言間距離、差異以及親疏的標準，理解度測試就是測量的方

式。游汝傑、楊蓓（1998）認為口語的可懂度影視方言之間接近率的重要標誌。也就是說ABC三種語言，AB之間的相互理解度是90%，BC之間的相互理解度是20%，AC之間的相互理解度是15%，那麼顯而易見A語言和B語言之間的差異較小，距離相對較近，而B語言與C語言之間的差異很大，距離很遠，A語言和C語言之間差異最大，距離最遠。由此可以判斷AB語言之間擁有的親緣關係肯定近與他們與C語言之間的關係。另外一種情況是ABC三種語言，AB之間的相互理解度是89%，BC之間的相互理解度是82%，AC之間的相互理解度是86%。這種情況下很難通過直觀的感受來判斷三者之間的關係，使用了理解度這把尺子，就可以清晰地看出三者之間的親疏遠近關係，三者之間都擁有較高的理解度，那麼這三者之間的距離一定是很近的語言變體。既有可能是剛剛分化的三個變體，也有可能是即將融合的三個變體。誠然測算語言之間的距離的方法並不是只有理解度這一個途徑，已有的研究已經做出了相當多的嘗試，前文中已經對其的優劣進行了論述，利用詞彙差異計算語言親疏程度的研究是一種共時的靜態的比較研究。但它忽視語言或方言的系統性和社會性。僅通過計量不同方言聲韻調的差異來反映方言間差異程度，顯然將系統中處於不同位置的要素一視同仁地處理，並以要素之間的比較代替系統本身的比較，是不科學的；同時，也忽視了方言的特殊性，有別於普通話的是，方言一般沒有自己的書面形式，只存在於人們的口語交際中，以普通話為參照系的字音或詞音計量單位也極可能忽視方言自身的特點（林素娥、蘇憲萍 2007）。理解度理論相對於傳統方法的優勢就在於它具有更廣泛的適用性，更便捷的操作性，為學科帶來更大的視野和前瞻性。換句話說，理解度理論就是語言學的一把更好用，更科學的尺子。這把尺子理解度測試為判定方言、語言提供了客觀可信的數

據。對理解度測試模型的改進能夠大大增加測試的標準化、模式化的程度，可以作為典型的測試適用於世界上任何語言。這種模型的建立需要做到可複製、可驗證，重復後對結果的影響小，為之後的研究提供便利的實驗方案。同時，如果使用同樣的實驗流程和統計方法，不同語言之間的數據就可以放到統一平台上進行對比。這樣的好處在於不再是結論與結論之間的比較，而是真正落實到數字與數字的比較，將理解度變為一種定量研究。

除此之外，尺子還有一個更重要的屬性，那就是尺子上有刻度。這個刻度也是尺子本身最大的意義所在。通過理解度測量出語言之間的距離後，並不是研究的終點。也就是說測量出兩根木頭之間誰長誰短固然很重要，然而這只是第一步，下一步是要說明長於 a 米的木頭可以稱之為木棍，短於 b 米的木頭只能稱之為木塊，介於中間的需要稱之為木條。這樣做的目的是在於木條可以用來製作兵器，而木塊只能用來燒火，木條可以用來製作板凳等等。不同的類型決定了不同的屬性，也就是決定了不同的用途。是為了更好的物盡其用。同理，理解度可以為語言提供同樣的功能，以往對兩種語言變體之間關係的界定常常會有爭議，理解度測試的數據能夠給出有效的證據。如果相互理解度大於 a ，那麼二者之間是方言關係；如果相互理解度小於 a ，那麼二者之間是語言關係。這樣可以更深刻更精確地認識語言之間的關係，正確地認識其中的差別可以為語言規劃，語言調查提供重要的基礎信息，能夠更好地服務於語言的應用。這種從數據中得到的結論更具有說服力，由定量研究走向定性研究，真正將二者有機結合起來。

冉啓斌、索倫·維希曼（2018）認為利用理解度標準來區分語言和方言有兩個問題，第一個理解度往往是不對稱的，第二是量化的結果不可用。其實理

解度不對稱是正常現象，是客觀事實。從距離的角度上來看，確實兩個點之間的距離是確定的，不應該是不對稱的，然而語言之間理解度的不對稱恰恰反映了兩個語言之間的關係，這種理解度上的不對稱正是二者之間關係的體現。因為語言和語言之間不一定就是兄弟關係，也可以是父子關係，叔侄關係。兄弟關係之間當然可以是對稱的可以異位的，互相之間你來我往，稱兄道弟。可是父子關係之間天生就是不對稱的，是不可異位的。因此理解度上的不對稱並不能夠作為不能區分語言和方言的例證，而恰恰證明瞭利用理解度理論揭示出的不對稱性更能有助於揭示出語言間的關係。除此之外量化的結果不可用其原因是標準的不統一，標準的不統一原因是在方法的不統一，本文正是認識到這一問題並希望能夠對此構建起統一的標準。雖然然方言和語言分界的具體數值現在尚不得知，這需要進行大量的實驗建立常模後才能夠逐步實現。這並不能削弱理解度理論在語言學中的意義，因為一個優秀的理論往往能夠對事實進行合理地推測，並給出正確的探尋方向。理解度理論正是這樣的一種理論，它具有足夠的前瞻性能夠通過理論探索照亮前方的迷霧，指出前進的方向，並預測出前方的成果。

2.8.2 驗證結論

理解度測試可以為以往研究的結論進行實際的驗證，對之前的定性研究進行定量的驗證，以定量的方式補充和完善定性的結論。Tang Chaoju (2009) 在第五章中引用了文獻中方言之間的客觀距離值，也通過作者自己的計算方式計算出方言之間的距離值。並與傳統的方言譜系進行對比，經對比後發現二者之間並無太大差異，有明顯的相關性。之後就利用作者收集到的方言之間的客觀距離值和實際理解度測試的結果比較，用客觀距離值作理解度的預測自變

量，並與作者論文的第三章和第四章的理解度值進行相關性計算，通過回歸分析得到判斷相似值比理解度更難從客觀語言距離測量中預測出來。功能測試法生成的樹形結構圖能夠反映傳統的漢語方言系譜圖。從而得到方言之間的判斷理解度能和判斷語言相似性顯著相關的結論，進一步印證了傳統方言學分類的結論。同時論文還之處對漢語方言而言，音調信息對理解度只起次要作用，因為從韻律版和無韻律版測試材料獲得的結果基本上很相近。發現二者之間同樣存在一定的相關性。也就是在印證前人結論的同時，還能夠對深化補充以往的結論。這種對前人研究的驗證和補充是建立在真實實驗和數據的基礎上的，可以作為有力的證據支撐輔助前人的研究結論，同時又進行了補充和創新，極大的促使學術的進步，而且這種進步更加扎實。反過來理解度理論在實踐中發現的問題和觀測到的現象，又為傳統本體理論製造了更多的可供研究的課題。

2.8.3 探索未來

理解度理論受益於心理語言學，二語習得理論等學科，同時也為這些學科開闢新的研究思路和實驗方法。理解度理論可以在心理語言學的理论指導下，結合心理認知的理論進一步探索人類度語言的理解機制。探索人類對語言的理解究竟是以何種方式進行的，人的大腦對於語言信號究竟是如何解碼的，這個解碼的過程具體是分成幾個步驟的，影響理解度的因素之間究竟哪個因素是主要因素，各個影響因素之間是否有相互影響，如何能夠快速提高理解度等等問題。這都與心理認知理論息息相關，只有能夠很好的回答了這些問題，才能夠推動人工智能等高科技的實質性進展。若想試圖回答上述問題，只能通過理解度實驗的方式進行實證研究。大腦是一個黑盒，人類目前對大腦的實際工作原理並不完全瞭解，只能通過外部施加不同的刺激來監測大腦做出相應的反應，

從而推測大腦的運行機制。理解度測試也是秉承這樣的思路，通過對大腦釋放不同的刺激，來獲得並且記錄下大腦的反應。從中可以瞭解到大腦對語言的認識和理解。這與單純從材料中得到的結論相互對照，可以發現大腦在處理語言材料時會有很多彈性和灰色地帶。表面上差異很大的事物，有可能會被大腦認為是一致的，而表面上差異並不大的事物，大腦也是可以敏銳的發現其中的差異，因此究竟差異的標準是什麼也是需要通過制定一套標準來決定的。理解度測試恰好是制定這樣一套標準的很好的方法。傳統方言研究的利用方言材料計算出方言之間的距離，這個數據固然正確，卻又在某種程度上並不正確，因為語言雖然是死的，但是使用語言的主體人是活的。語言和語言之間的距離，並不完全是音系結構和音系結構，詞彙和詞彙之間，句法和句法之間的距離。兩種語言之間的距離，最有發言權的應該是說這兩種語言的人。大腦在很多情況下會忽略掉很多差異，又放大某種差異，使得真正的語言距離和考語言材料算出的相差甚遠。理解度理論就是可以搭建起人與人之間的橋梁，語言和語言之間的橋梁，向更多未知的領域發起探索的號角。

二語習得教學中，檢測學生的語言能力通常只能通過語言測試來進行評價，然而測試的評價僅僅是一方面，有時並不能對學生的實際交流理解能力進行一個準確的測算，因此就不能根據學生的實際情況給出一個確切的評價。可以通過理解度測試對學生在模擬真實的環境下進行測試，從而得到更符合真實環境下的理解度評分。因為理解度測試不僅僅是以正確率作為唯一標準，還同時綜合了其他參數，因此可以為學生提供更全面的參考意見，也可以給老師提供多方面的反饋。在二語教學中同樣存在著標準的問題，若以字典上標準的發音作為標準教授、考核學生，對學生來說無疑是一項不可能完成的任務。以漢

語為例，漢語母語者尚且不能自信地認為自己的普通話是標準普通話，漢語教師也並不能夠確保一字一詞都能夠發音準確，只要能達到相互可理解這個水平就可以視作是及格的普通話，從而可以節約更多的時間精力去學習其他語言知識。除此之外，可以將二語習得的學生的理解度與母語者的理解度常模進行比較，從而可以直觀地比較二語學習者和母語者之間的差異，並且可以知道二者之間的差異距離，這樣可以讓教師能夠更好的掌握學習者的學習進度，也可以更好地檢測教學的實際效果。

最后，一個語言變體究竟是稱其為一種語言還是一種方言，其背後通常蘊含著複雜的政治、經濟、歷史、文化等諸多方面的影響。語言本身的屬性往往是較少考慮的因素。因此才會出現相互理解度很高的兩個變體稱為兩種語言，而相互理解度很低的兩個變體稱為一種語言的兩種方言的現象。在政治世界中，語言與方言是明顯的政治符號和意識形態標籤，然而在語言學的視角下，可以忽略這些非語言本體的標籤來重新審視語言本身，通過理解度理論與測試去衡量不同變體之間的距離，依靠理解度作為標準，重新繪製在純粹語言學視角下的世界語言版圖。這種理論研究並不涉及政治及意識形態，純粹從科學的角度來對語言的親屬關係進行劃分。正如同DNA測試一般，使用DNA作為一把標尺，就可以知道不同物種之間的親疏遠近程度，從而為不同物種之間劃分不同的綱目。在語言學中，理解度測試從某種程度上講就是語言的DNA測試，其目的也會更好地反作用于語言學及相關學科的進一步研究，進一步幫助我們去認識語言的內部世界。

2.9 理解度的局限

2.9.1 言外之意的理解度

言外行為是語用學中最重要的理論之一，它與生活中的活語言息息相關（楊曉 2011）。英國哲學家 J. Austin（1962）認為一個人在說話的時候，在大多數情況下可能在同時實施三種行為：言內行為（Locutionary Act）、言外行為（Illocutionary Act）和言後行為（Perlocutionary Act）。目前的理解度模式僅僅適用於對語言材料字面意義的理解即言內行為的理解，尚沒有涉及到對言外行為的理解中。因為研究理解度的重心在於不同語言變體之間溝通的過程中，對理解度進行定量研究。若是使用同一語言，而該實驗材料字面意思非常簡單，卻通常包含著豐富的言外之意，這種情況下不同人接收到的信號的理解度又是一個新的問題。在目前的理解度理論框架內尚不能對該問題做出合理的解釋。因為如果聽話人僅僅聽出了字面意思，按常理說已經是做到了百分之百的理解度，可是在明顯的言外之意的的情況下，字面意思往往並不是說話人想要說的內容，恰好言外之意才是真正想要傳達的信號，這時聽話人沒能解讀出其中蘊含的意思，那麼這時的理解度又是百分之零。這個矛盾也是理論中實際存在的。通常情況言外之意的解讀與情景和人物背景等預設有關係。如果一句話在相同的預設內可能有多重言外之意的解讀方法，那麼對於這種形式的理解度該如何判定，暫時還不能給出一個非常圓滿的答案。就比如春節聯歡晚會上的相聲或小品，在北方人看來，可能十分好笑，但是在南方人看來，字面意思完全理解，就覺得完全不能捕捉到笑點。這顯然是沒能很好地理解演員想要傳達的意思。這種對於言外之意的理解度應該屬於更高級的語言理解度範疇，也是語言應用的更高級水平，還需要持續更深入地研究才能夠將其內部機制解釋清楚，暫時超出了本文研究的理解度的解釋範圍。

2.9.2 相互理解度的實時監控

在兩人和兩人以上的交流模式中，若二人使用同一種語言，那麼其中的相互理解度應該基本不會發生變化，維持在很高的水平。若二人使用完全不同的兩種語言，那麼其中的相互理解度應該也不會發生太大的變化，維持在一個非常低的水平。若二人使用不同的語言變體，例如使用相近的方言，那麼其中的相互理解度可能會隨著二人交流時間的增加而上升。這個上升的過程是怎樣的根據現有的研究方法還不能知道。從理論上推論，相互理解度會上升到一定的高度然後又保持穩定，不會繼續上升。由於這是一個變化的趨勢，對話開頭和對話結尾可以通過現有的手段測量溝通度，但是其中的變化軌跡和具體過程依照目前的手段暫時無法得知。目前僅僅能夠做到從理論的角度推論結果是這樣的趨勢，但是尚沒有非常好的實驗方法能夠實時測量出二者之間的溝通度。研究這一問題的意義在於能夠及時地發現提升理解度和降低理解度的相關參數，即兩個語言變體中哪些特徵是輕易被識別哪些特徵不容易被識別；以及如何快速提高理解度的方法，是否存在一些關鍵性的特徵，意識到這個特徵後就能夠讓理解度快速提高等等問題。

目前對該問題還不能有所進展的主要問題在於：

(1) 實際對話中，有涉及到眾多的非語言參數。二者對話的過程中，可能不僅僅會通過語言來傳達信息，還會夾雜更多的肢體語言、表情等多方面的信息。通過非語言性的手段傳遞的信號會幫助聽話者理解對方的語言性內容，從而增強了語言的理解度，這顯然是不符合實際交流的情況的。

(2) 二者對話的過程中，語義大致上是連續的，這也勢必造成背景信息的重復出現。背景信息會有助於提升聯想能力，從而加大理解度，這種理解度的

增大並不是雙方對語言的熟悉程度提高所造成的結果。但是目前的實驗手段很難將這一部分的誤差消除。

(3) 多人的對話過程相比於兩人的對話其難度不是簡單的數量的加和，而會是成倍遞增的。多重信道的疊加意味著多重噪音的疊加，並且變量的也會成幾何倍數增加。所以這樣複雜的交際模型目前是很難控制的。

上述原因導致現有的理解度測試方法尚不能夠對其進行很好的測量，有賴於更好的研究模式和改進的實驗方法，才能夠實現做到真正的相互理解度的實時測量。

2.9.3 樣本數量

樣本數量對於理解度的應用至關重要，兩種語言之間的相互理解度是一個客觀的常量，這個常量是一個固定的區間範圍。這個常量就如同一個人去醫院驗血的時候，每一個檢驗項目後面都有一個正常值範圍，如果檢測值超過了這個正常範圍，或者低於正常範圍那麼說明這個人得了相應的疾病。這個常量的範圍稱之為常模，常模是經過大規模的科學檢測得到的一個相對穩定的取值範圍。常模參照方法也是眾多學科都會應用的一種方法，理解度理論也是可以利用常模參照這個方法應用理解度的數據。因此如何得到這個常量則是應用理解度的必經過程。這個過程需要大量的樣本通過同一個實驗手段才能夠建立起有參考價值的常模數據。

經過大量的實驗，每兩種語言之間都可以建立起這樣的一個常模。每一個常模都需要足夠多的被試才能夠建立起來。在共時層面建立起來的常模可以放到歷時層面上去觀察。只要按照程序地進行理解度的測試，那麼就可以對不同

時間段的理解度常模進行總結。從而得到語言演變的規律和瞭解語言演變的階段，也同時為語言演變提供一個指數化的參數作為階段的標記。那麼這種實驗所需要的樣本數量也是空前巨大的。這一切都依賴於足夠大的樣本，在目前階段只能作為理論上的推演，尚不能夠實現，有賴於日後不斷完善的實驗工具和方法，能夠在節約實驗所需的人力成本和實踐成本，可以盡快建立真實有效的常模體系。

2.10 理解度理論的未來發展方向

理解度理論不僅能夠涵蓋大量的語言事實，同時還能夠對語言學的本質性問題進行深入探討，例如人類是如何理解語言的，這個理解的機制究竟是什麼樣的一個過程，大腦究竟是如何處理聽覺信號的，以及相關眾多基礎問題都與理解度理論相關，這類問題嘗嘗被人們忽略，因為看起來並不是問題的問題，往往因為太過於習以為常而忽略，恰恰是這些容易被忽略的問題才是問題的關鍵。就像空氣一樣，雖然看不到，摸不著，但是失去了空氣，人就是無法生存，所以看不見的空氣對人來說卻是至關重要的。因此理解度理論提升到非常高的理論層次。目前，人類理解語言的具體機制都還處於假想和推論階段，還沒有一個學科可以對人類語言理解機制做出革命性進展。本文認為是缺乏一個各個學科共同合作的平台，單兵作戰不如聯合作戰，理解度理論就提供了這樣的一個理論和實踐的平台，可以集各方所長。

2.10.1 人工智能（Artificial Intelligence）領域

人工智能屬於計算機領域，希望通過計算機來模擬人類的智能，甚至延伸、超越人類的智能，幫助人類實現很多重要的科研突破。人工智能所要處理的問題也有很多包括圖像識別，自然語言識別，語言理解等諸多方面。其核心

在於讓計算機擁有自我學習的能力，不斷自我更新。人工智能最重要的問題之一就是讓機器去理解自然語言。目前所見到的大多數的智能機器人能夠初步地實現與人類的對話。但是對話的內容和數量都是有限的，並且是設定好的。其機制是通過語音識別軟件將聲音信號轉化為可檢索的機器語言，隨後進入數據庫中，與數據庫中已有的句子進行檢索和比對，檢索到相同的或類似的結果後，提取相對應的答語隨之通過揚聲器播放出來。目前的技術不斷在更新，對語音的識別能力越來越強，準確率越來越高，並且數據庫的內容也越來越豐富，甚至採用雲端的數據庫，用大數據進行分析後檢索到適當的答語。能夠讓人感覺到幾乎是真實的對話模式。但是，機器的反應機制還不符合人類使用語言的模式，原因有以下幾點：

- (1) 機器還不能夠自我造句，還不能夠說出這世界上沒有過的話語。
- (2) 對同一問題的答案大致相同，如果持續向機器問一個問題，得到的只有三四個相同的答案，不會有其他的答案出現。
- (3) 並不能理解言外之意，只能夠對字面意思進行解讀和反應。

本文認為人工智能若想做到對語言理解的真正突破，一定是建立在對人類理解語言的模式有充分瞭解的基礎上的。只有這樣，才能夠稱得上是名副其實的「智能」。

在機器學習人類語言的過程中，理解度也可以作為檢驗人工智能語言的一個重要手段。也就是說理解度可以作為一把尺子來衡量人工智能的語言能力，只有當機器產生語言的理解度達到真實人類語言的理解度水平才能符合要求。利用理解度理論對人機對話進行檢測，一方面可以檢驗人類對機器說話的理解

度，對於真正屬於機器產出的語言進行理解度分析，另一方面可以檢驗機器對人類說話的理解度，也就是人機之間的相互理解度。這也為理解度理論和研究提高到一個新的高度和應用層面。

另外，人工智能如果能夠非常好的模擬了人類的語言理解模式，那麼人工智能可以通過兩台電腦相互交流模擬兩種語言之間的交際過程，這樣可以通過模擬的交際過程中記錄下實時的理解度，這樣可以省去真實的實驗操作，並且機器可以自我記錄，得到的結果也會更加準確有效，如此便可以大大節省人力成本和時間成本。推而廣之，利用人工智能的交際模式模擬出世界上任何兩種語言之間的相互理解度。可以為語言學研究帶來非常多重要的研究成果。

最後，理解度理論可以應用於機器翻譯領域。人說的話是否能夠被機器所理解，機器翻譯的內容是否能夠很好被對方理解，都是提升機器翻譯質量的重要環節，也可以更好的監測機器翻譯的結果是否真正有效。理解度理論也可以為機器翻譯改進理解和翻譯的算法，從而更快速準確地理解句子，並翻譯出來。

2.10.2 大數據化、常模建立、量表化

理解度常模的建立必須建立在大數據的基礎上，才能建立起猶如醫學測量標準一樣的常模數據。為了達到這樣的目的，需要做到以下幾點：

(1) 實驗過程程式化，統一所有的實驗過程，保證所有的實驗流程的同一性。從實驗方法和實驗材料上都盡可能一致，才能保證結果的可比性。

(2) 實驗過程簡易性，簡易型增強了實驗的可操作性，並且簡易的實驗也是將大量冗余的步驟刪除後得到的，只把必要的流程保留。

(3) 實驗過程電子化，利用現代科技的各種軟件或傳感器對人類大腦活動進行檢測，可以更加準確地記錄下真實的反應。

(4) 實驗材料量表化，將實驗的材料總結成科學的量表。通過量表的制定，可以更加規範實驗過程，並且將人為干擾因素降到最低。量表的使用，可以講過簡單的培訓就可以訓練出大量的合格的調查員，從而進行廣泛的大面積的調查，無疑可以大大提高效率。

依照目前的技術發展速度來看，大數據的時代已經來到，並且勢不可擋。大數據為研究者提供了一個更宏觀的視角，更廣闊的視野去認識和處理一個複雜的問題。尤其像人類對自然語言理解的模式，並不是通過幾個簡單的變量就可以說清的，也不是通過十幾個被試就能夠涵蓋所有可能性的，因此必須要擁有大數據的數據觀去分析和思考這個問題，才能夠得到一個全面真實的結論。理解度理論的進步需要充分趕上大數據的潮流，跟隨者技術進步的發展，不斷自我革新，與時俱進。

2.11 小結

本章內容是本篇論文的核心，從各個方面闡述了理解度理論，第一次將理解度理論完整地進行總結，歸納及論述，構建起理解度理論的理論框架。本節將各節中的重要內容和結論進行總結和梳理，將主要的觀點歸納精煉。

第一節中文章回顧了理解度的源流，從上個世紀二十年代開始逐漸興盛，在各個領域相繼對理解度這一問題進行研究。隨著學科的不斷分化，研究目的的不斷細化，信號學主要研究在信號傳遞過程中的信號衰弱與增強等問題，演化出清晰度測試與相應算法。理解度測試最先應用在病理學領域中，後逐漸引

入到語言學研究範疇中。在語言學研究領域中，心理語言學，社會語言學都對該問題有所涉及，但未見其形成一個完整獨立的理論，其理論多見於論文中零散的論述。

第二節是對理解度的概念進行界定。首先利用經典五分的通信系統模式來對信息傳播的過程進行描述。通過對不同模塊的性質特徵進行描述說明後，找到其在理解度理論中的對應項並分析二者之間的關係，發現通信系統模式可以很好地描述理解度發生的整個過程，並且更加明確理解度的研究重心是在於信號理解的程度，而信號學的清晰度是研究信號接收的程度。病理學是患者本人作為信號的信息源和發出機出現了障礙，理解度的研究方向則是偏重接收機到收信者的過程。接著本節對理解度的內涵進行梳理，最終提出理解度的概念是：語言信號由說話者說出經傳信系統到達聽話者的這一過程中，聽話者對原信號即時的理解程度，為單向理解度。在對話過程中，聽話者與講話者的身份交替轉變，雙方或多方輪流成為聽話者，對話參與者相互之間的理解度即為相互理解度。並對概念所蘊含的基本假設進行闡述以及概念中的關鍵要素進行詳細說明。

第三節列舉了理解度的性質，對理解度性質的總結和歸納一方面可以深入瞭解認識理解度本質，另外更重要的是要利用理解度的特性去合理地解決實際問題。本文從即時性、聽覺性、單向性、客觀性、動態性、相對穩定性這六個角度去對其進行描述。這六個基本屬性可以較為全面地描述理解度，從這六個角度去把握理解度理論可以更好地指導進一步的理論建設和實驗設計。

第四節詳細討論了理解度的影響因素，理解度研究的瓶頸問題就是在於影

響理解度的因素具有種類多，層級多，範圍廣等幾大問題。理解度的變化是受到了眾多影響因素的共同作用，過多的因素阻礙了傳統方法對理解度研究的深入。本節從語言系統內部和語言系統外部的影響因素進行探查，語言系統內部包含了語音、詞彙、句法三個主要方面因素以及內部若干子因素，語言系統外部因素歸納總結了信噪比，背景信息，社會因素等因素。充分闡述這些因素可能會對理解度產生怎樣的影響。僅僅是闡述和說明還並沒有給出更好的解決方法，本文認為影響理解度的眾多因素對理解度的影響是有層次的，並不是每個因素都對理解度產生相同量級的影響，因此對因素進行分級也是必須的。分級處理的方法一方面可以簡化操作，有側重地考慮每個因素的影響力；另一方面可以具體問題具體分析，不同層級不同對待，應對每個層級的影響因素的對策並不相同。試圖做到既不是一刀切，也不是一視同仁。

第五節討論理解度的測試方法，本節主要是從理論層面來對測試方法進行討論，具體的技術性的測試內容會在後文具體討論。這一節主要從實驗材料、實驗方法、實驗對象、和測試工具四個方面來對理解度的測試進行闡述。由於理解度測試適用性非常廣，因此會利用理解度測試研究不同具體的問題。所以實驗中的各個要素差別就會非常大。本節的思路是將實驗中可能會出現到的變項列出，通過理解度的性質對其進行甄別，留下符合要求的選項。實驗材料的部分本文認為首選應使用句子作為實驗材料，其次可以使用詞語。不推薦使用字，語篇等單位作為實驗材料。實驗方法部分，本文認為應該將主觀判斷和客觀實驗結合起來，客觀實驗為主，主觀判斷為輔，二者之間相互印證，相互補充；其次使用選擇類的任務類型的方法比使用轉寫類的任務類型更能夠得到真實的數據，且實驗材料錄音只播放一次；最後需要記錄被試的反應時間，用來

消除選擇類型任務的誤差。實驗對象的選擇方面盡量保證來源的廣泛性，杜絕利用主觀意識來選擇被試。測試工具的選擇也是本節的重點內容之一，本文分析了傳統測試和改良後的測試的利弊，雖然目前來看，傳統測試的經濟型，便捷性要大大勝過改良後的測試；改良後的測試在數據上的準確性和多樣性上是傳統測試不可比擬的。因此從長遠的角度來看，改良的方法代表著進步的方向，會逐漸成為主流，最終取代傳統的方法。而現階段，應該將二者有機結合起來，發揮二者的優勢，在實踐中不斷完善測試的方法。

第六節討論了理解度應該使用什麼樣的參照標準，本文認為根據理解度的性質和特點，應該使用常模參照標準。這一節首先對常模參照標準和標準參照標準進行了介紹，闡述了二者之間有什麼差異，各自適用於什麼樣的測試背景。接著列舉選擇常模參照標準的原因，先從理解度的性質出發，同時考慮到理解度是一種語言能力，最後列舉了為什麼不選標準參照的理由。因此理解度理論需要使用常模參照標準來設置理論的標準。

第七節列舉了國內外一些重要的相關應用理解度的文獻，通過已有文獻來總結理解度的應用。總結後發現，國內眾多研究還都處於嘗試和起步階段，涉及方面很廣，不過還需要借鑒國外類似文獻及其操作方法。國外的文獻主要注重二語習得方面，口音對英語理解度的影響，研究範圍多集中在印歐語系，對其他語言討論並不多。最後本文總結了三點需要在日後的研究中加以改善的環節。

第八節集中討論了理解度理論的意義，是本文的重要核心內容之一。本節從理解度理論的功能角度出發，探討理解度理論能夠為語言學貢獻怎樣的功

能。首先是尺子的功能，理解度是作為衡量語言與語言之間距離的尺子，首先明確以理解度作為衡量標準，接著是建立理解度的標準。標準對於一個學科的重要性不言而喻，理解度理論就是這樣的一個標準建立者，也是一個平台的搭建者，在這個平台上可以找到與其他理論的接口，共同協作解決一個問題。其次是理解度理論可以對已有研究的結論進行驗證，驗證的過程中又對原理論進行補充，可以極大促進學術的進步。並且理解度實踐中發現的問題和觀測到的現象，又為傳統理論創造更多可供研究的課題。最後是理解度理論可以提出更多更有意義的問題來引領學科的發展和學術的進步。在實踐中不斷創新，向未知的領域不斷探索。

第九節列舉了理解度理論的困境，一個理論一定不是十全十美包打天下的。理解度理論也有它本身的局限，至少是在目前階段不可回避的缺陷。首先是對言外之意的理解，目前的理解度理論並不能夠處理這種問題，目前的理解度僅僅是對字面意思的理解，言外之意的理解又屬於更高一個層次的理解範疇，暫時不在本文的研究範圍內。其次是對相互理解度的實時監控，依靠目前的技術手段還不能支撐這一概念的實施，只能夠將相互理解度拆分成兩個單向的理解度，之後通過統計手段計算成相互理解度。最後是樣本數量的不足，若想真正建立語言之間的理解度標準，那麼需要對世界上大多數語言進行研究測試，而每一門語言就需要大量的被試來形成基礎性的數據。所以對於構建語言之間的理解度常模，還需要大量的樣本。隨著樣本數量的擴大，理解度數值的標誌作用會逐漸加強。

第十節展望了理解度理論未來的發展方向。本文設想能夠進一步完善和改進人工智能技術，利用理解度理論得到的成果來幫助人工智能技術，測試人工

智能的語言能力，優化人工智能的算法。並且希望能夠讓理解度理論與人工智能領域相結合，利用人工智能技術的智能性來模擬人類的理解度模式，從而通過人工智能獲得大量的模擬數據，這樣可以替代傳統測試的方法，極大地提高效率。除此之外，可以應用到機器翻譯的領域，理解度理論可以為機器翻譯改進優化算法，檢測機器翻譯的結果是否能夠被理解，機器理解人類的語言是否是真正的意思等等。最後是理解度理論需要建立在大數據的視角下來考慮問題，宏觀全面地認識問題，要站在上帝視角去審視理解度的問題，就是要充分意識到問題的複雜性，簡單機械的求簡最終會導致失真，承認問題的複雜性，但是並不畏懼複雜，而是用相應的技術和手段去處理複雜，從浩如煙海的數據中找到簡單真實的規律。



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

第三章 理解度測試與時間參數

第二章第五節中已經從理論上討論了理解度的測試，本章將會從實踐的角度進行理解度測試的探討，並且試圖對其進行改進，即在理解度測試中增加對時間參數的測量和計算，讓理解度測試更加貼近實際情況，更好地測量出被試者的真實水平。

3.1 時間參數的概念

時間參數時間參數是利用被試的反應時間，經計算後得出的參數，用以參與理解度的計算。在實際測試中，實驗需要測量被試做出每一個反應的反應時長，用這個反應時長，經一些列計算，從而將被試的真實時間轉化為可供運算的時間參數，該參數最終參與理解度數值的運算。

在概念中有幾個關鍵要素需要注意：首先是反應時長，這是時間參數的核心，是被試在進行測試的過程中，完成相關指定任務所做出反應的時間，也就是從被試看到題目後，到被試做出判斷或選擇這段時間。其次是計算，時間參數並不是使用直接實際的時間數據的，而是將採集到的數據進行歸一化等一系列運算後形成的參數。最終是以參數形式進入理解度的計算公式並參與計算。

本文使用反應時間作為計算時間參數的初始數值，而這裡的時間並不是理解時間。根據前文所說的理論，被試者因為聽不懂或不熟悉測試內容，因此在選擇的過程中會產生比完全理解所用的更長的時間。此處延長的時間是大腦處理和運算的時間，屬於理解時間。本文並不使用這個理解時間，而是選擇反應時間，主要有以下幾點原因：

- 1、理解度實驗是屬於行為實驗的層級和範疇，還沒有上升到神經實驗的

層級。在行為實驗中，使用反應時間是通行的做法，而如果是在神經層級，那麼反應時間就是遠遠不夠的，是絕對不符合實驗規程的做法。反應時間通常都在一千毫秒左右的數量級，而神經實驗中的反應時間是幾十到幾百毫秒的數量級。通常是大腦還沒有意識到事情的發生，神經就已經完成了反應的工作，因此二者並不是在一個層級的，在行為實驗的層級中使用反應時間是沒有問題的。

2、反應時間雖然與真正的理解時間不同，但是二者之間具有等效性。可以將反應時間分成兩部分，一部分是大腦處理信息所用的理解時間，另一部分是得到答案後大腦指揮身體做出反應的行為時間，反應時間可以看成是這兩部分的加和。通常得到答案後身體做出反應的行為時間是大致相同的，因此造成反應時間差異的就是大腦處理信息所使用的理解時間。換句話說，行為時間並不會因為理解的難度而延長，反應時間的長短主要是隨著理解時間的變化而變化，理解時間是隨著理解難度的變化而變化，因此反應時間也是隨著理解難度的變化而變化。所以反應時間和理解時間二者之間是有等效性的。

3、如果要檢測真正的理解時間，是需要借助一系列高級儀器，例如腦電圖掃描儀，才能夠實時監控到大腦真正理解的時間點。這類實驗需要高級的實驗儀器和實驗場所，同時也需要較長的準備和實驗時間，以及足夠的資金。本實驗考慮到自身的條件限制，沒有選擇神經實驗，而是選擇了行為實驗，利用反應時間作為時間系數的基礎數據。

3.2 添加時間參數的必要性

在以往傳統的理解度測試中，僅僅是以被試完成指定任務的正確率作為最

終的理解度數值。然而在實際的情況中，正確率卻不一定能夠反映出理解度的真實水平。前文中提到理解度的測試方法主要分為兩種，一種是轉寫類任務，另一種是選擇類任務。目前通行的理解度測試也是使用選擇類任務，本文也推薦選擇類任務，因為在選擇性任務中，作答的過程簡單易學，有了答案之後就可以迅速做出反應，不需要再重新編碼轉寫下來，因此被試可以將注意力集中進行理解分析。這種方法的優勢就是簡單、便捷，而劣勢是被試者作出反饋過於簡單以至於會出現即使不會也做出一個選擇，亦或是即便沒有聽太懂，也是通過後期大腦的分析聯想得到的結果，也就是通常所說的「蒙」。通過這種方式答對題目就會造成最終結果的不準確，因為一種情況是完全不會做出的選擇，可能是蒙對的，這種情況還能夠通過一些技術方法杜絕和避免；另外一種情況更具有迷惑性，就是被試並不是完全不懂，而是通過題乾的提示，再通過聯想、回憶等方法將題目做對，這種情況下被試也覺得自己聽懂了理解了。然而這樣的理解方式並不是滿足理解度定義中即時性的要求；換句話說，這種理解並不符合語言實際情況的理解，而是一種後期語言信號加工處理的結果。人腦對於理解複雜的理解工作需要付出更多的時間和注意力。注意好比資源，它的容量是有限的（桂詩春 2011）。單位時間內不能夠解決複雜的問題，那麼只有付出更多的時間。若將被試者處於真實的語言環境中，在自然語流的速度下，被試者的大腦並沒有足夠的時間去處理、聯想、加工聽話人的語言，往往是還沒聽懂第一個詞的時候，第二或第三個詞就已經說完了。這樣的結果就是每一個詞都因為來不及處理所以導致聽不懂，於是就不能夠達成交際目的。

除此之外，在處理同一語言中的不同支系，或同一方言中不同片區的語言理解度時，雙方的語音系統差異並沒有非常大，如果用傳統的方式，僅僅依靠

正確率來評價，那麼可能會造成結果的整體偏高，並不利於得到真實的結果。

3.3 意義和作用

本文首先提出在理解度測試中應增加時間參數，以往的理解度研究和實踐中並沒有真正加入時間參數參與理解度運算。Yorkston & Beukelman (1981) 就首先提出了溝通效率 (communication efficiency ratio) 這個概念，其中就在理解度的問題中提及時間概念，作者認為不同的說話者即使說出同樣的可理解的詞語，也是有不同的，說話者能夠在「單位時間」內說出越多的可理解的詞語，那麼這個人就可以稱得上是高效的溝通者。由此可以引申出相似的結論，如果是聽話者在單位時間內聽懂的詞語越多，也說明這個聽話者對這個語言更熟悉。

理解度測試中之所以沒有人加入時間參數，其原因主要是為早期的理解度出現是處理病理學的問題，對於生病的患者，首先要考慮的事情是對與錯，不是快和慢。對於病患來說，能說正確，理解正確已經是不容易的事情了，至於反應速度的快慢就暫時不考慮了。另外，對於測驗不同語言之間的相互理解度時，如果語言差異過大，那麼就沒有考慮時間的必要性，因為聽得懂就是聽得懂，可以理解就是可以做出判斷，相反，如果聽不懂，不能理解的情況，那麼無論給被試多長時間都不能夠將題目做對，那麼這種情況下，引入時間參數也並沒有太大的必要。最後還有一種情況就是傳統的理解度測試方法無論是轉寫法還是選擇法，並沒有條件記錄下被試做出的每一個選擇的時間，也就是在技術層面上還達不到記錄時間的條件。

時間參數對於理解度測試的意義首先是在數據上，毋庸置疑，時間參數可

以大大精確理解度測試的數值。因為可以通過反應時間來測試出被試者究竟是真的理解了，還是經過一番思考後答對的題目。增加了反應時間就比原有的數據多增加了一個維度。以往的數據僅僅使用正確率來作為評價理解度的指標，那麼只會有兩種可能性，正確就表示理解，錯誤就表示不理解。

選擇	正確	錯誤
表示含義	理解	不理解

加入時間這個維度後，數據就可以反映出更多更豐富的内容。

	正確	錯誤
用時長	經過思考後答對了=打折扣的理解	經過思考後答錯了=不理解
用時短	理解	不理解

用時長的正確和用時短的正確之間是有區別的，用時長的理解並不能說是不理解，而是打了折扣的理解。因為雖然結果對了，但是卻不是像母語者一樣的「秒懂」，而是大腦處理後的結果。錯誤也分用時長和用時短，不過結果都是錯誤，都是沒有理解，就不需要再進一步分析了。

由此可見，以往的測試中並不統計時間，因此包含著相當大的一部分正確的結果並不是真正的理解，而是摻入了一部分其他成分。加入了反應時間後，能夠對數據做到更加精確，減少實驗過程中的誤差。

其次，時間參數對於理解度測試的意義是將理解度測試引入心理學測試的範疇。因為在心理學測試中，反應時間是一項必須涉及的重要參數，是能夠將

很多問題引入深層的探討中去。如果卻是了反應時間的測試就不能夠稱其為合格的心理學測試。增加時間參數也能夠讓理解度測試轉變為到更符合學術規程的測試，也就能進一步保證測得數據的有效性。

3.4 改進公式

鄭錦全（1994）提出利用音韻材料計算漢語方言之間的溝通度。他的計算方法是把聲母、介音、元音、韻尾和聲調的類型列出來，成素數目等於或者大於平均的類型就是多素類型，成素數目小於平均的類型就是少素類型，多素類型提供的是信息，而少素類型提供的是雜訊，所以根據平均決定該類型是信息或是雜訊，再以對當成素的異同決定每個成素的權重，因為多素類型提供信息所以賦予正值，少數類型提供雜訊，所以賦予負值，以字頻乘權重，得出數值，所有的數值加起來得總值。最後總值除以總字數，就是單向的溝通度。文章中舉例北京話對濟南話的溝通度是通過北京話中零聲母的字和濟南話中零聲母的字共有290個，根據所有配對的頻數的平均數來計算出零聲母的平均值是160，此項的權重是0.20，也就是屬於多素類型，可以提供信息，頻數乘以權重得到數值是58，然後再將所有的配對的數值加總，得到總值是2004.25，再除以總字數2763，所以等於0.725。也就是北京話對濟南話的溝通度是0.725。

可以將他的公式總結成：

$$D = \frac{F_1 \times W_1 + F_2 \times W_2 + \dots + F_n \times W_n}{F_1 + F_2 + \dots + F_n}$$

這裡的D代表的是語言距離，F是頻數，W是權重。鄭錦全僅僅是通過對音系材料的歸納整理得到的這樣的算法，並沒有對理解度做出實際的實驗測試。

陳海倫（1996：361）提出方言的語音單向溝通度可通過一下公式求得：

$$Y(I, J) = \Sigma X(i, j) \times P(i, j) \times R(i, j)$$

其中Y(I, J)為甲方言對乙方言的語音單向可溝通度，X(I, j)為甲方言對乙方言某一對應類型的語音相似度，P(I, j)為該類型在總體中的比重，R(I, j)為該類型對語音對應規律的影響。

方言詞彙單向可溝通度可以通過下式求得：

$$C(I, J) = \frac{\Sigma K_v}{\Sigma K_v + \Sigma j_v}$$

其中C(I, J)為方言詞彙單向可溝通度，K_v為某一雙方共有的詞的現次率， ΣK_v 為雙方共有的所有詞的現次率之和，j_v為某一方言獨有的某個詞的現次率， Σj_v 為該方言獨有的所有詞的現次率之和。式子表示的就是共有詞彙量的總現次率與共有詞彙量總現次率加上一方言獨有詞彙的總現次率的和的比值，也就是交際中在詞彙方面一方能聽懂另一方的話語的覆蓋率。

除此之外，陳海倫（1996:367）認為鄭錦全（1994）在計算方言間溝通度的過程中，針對各種情況所設定的計算權值為1、-1、0.5、-0.5、0.25、0.25，這樣的權重值太過於主觀，缺乏客觀依據的支持。

其餘文獻中對理解度做測試的相關文獻並沒有給出相關的計算方法，因為主要的計算方法就是統計被試的正確率後計算平均數。大概可以歸納成：

$$I = \frac{A_1 + A_2 + \dots + A_n}{n}$$

本文認為單純的平均數已經無法滿足現今理解度計算的需要，需要在以往的計算公式中加入時間參數T，從而改進原有公式。因為時間參數是被試完成每一道題目所用時間計算而來的，因此需要跟每一道題的正確選項相乘，最終求整體的平均值。可將公式歸納成：

$$I = \frac{\sum_{i=1}^n A_n \times T_i}{n} = \frac{A_n \times T_1 + A_n \times T_2 + \dots \dots A_n \times T_n}{n}$$

這裡的A是被試答對題目的得分，答對得一分，答錯得零分。T是時間參數，這個時間參數是因人而異，因題而異的。時間參數可以理解為是給答題得分加權，增加了一個權重。也就是將用時長的得分打一個折扣，並不能說被試者完全不理解，而又與完全理解有區別對待，所以要乘以一個系數，讓被試者的得分並不能夠得到完全理解的滿分。

需要說明的是這裡的I還並不是最終的兩種語言的理解度，而是一個人的理解度。最終兩種語言之間的理解度是要通過若干個I求得的。簡單的處理方式是給出一個平均數；然而單純給出一個平均是僅僅是簡化了問題，並不能夠真實的反應事實，真正的理解度應該是一個範圍，這也是要通過常模的建立才能夠得到的，絕大多數的人都落在這個範圍中，那麼就是真正的理解度取值。具體的時間參數的計算方法詳見後文3.6.5.2.2節。

3.5 實驗目的

此次實驗會利用前文所述的理解度理論作為支撐，指導實驗的設計與結果的分析。另一方面也是用一個實際的實驗來驗證理論的可行性，將抽象的理論轉化為實際的實驗結果，最後可以利用實驗結果來說明相關的理論和實際問題。其中此次實驗主要研究的問題有：

- (1) 怎樣計算時間參數，並如何將時間參數納入到理解度的計算中？
- (2) 粵方言、湘方言和閩南方言之間的相互理解度是多少？相互理解率是多少？
- (3) 粵方言、湘方言和閩南方言之間是否存在著理解度不對稱的現象。
- (4) 此次實驗所得出的理解度數值與前人相關研究有何異同。

3.6 實驗驗證

3.6.1 實驗材料

此次實驗選擇了 24 個詞和 24 個句子作為實驗材料，實驗共涉及到三個方言分別是：粵方言、閩南方言和湘方言。實驗材料的產生方式是在現代漢語基本常用詞彙表中簡單隨機抽樣得到的 48 個詞，然後將 24 個詞造句。唯一的篩選條件是這個詞不可以有同音異型詞，即不允許聽聲音產生歧義。每個句子的長度平均值 7.625 個字，句子字數在人腦±7 個單位的短時工作記憶區間範圍內。其中每個方言分別有 8 個不同的詞和句子座位實驗材料。粵方言發音人是以粵方言為第一語言，女性，25 歲，廣州人；閩南方言發音人是以廈門話為第一語言，女性，24 歲，廈門人；湘方言發音人是以長沙話為第一語言，女性，長沙人。音頻採集使用 Praat 軟件，單聲道，44100Hz，48 個音頻總時長 73104ms，平均時長 1523.01ms，最大值 2856ms，最小值 378ms，標準差 1456.52。在視覺呈現方面，全部材料採用的是圖片呈現。圖像資料來自百度圖片關鍵詞搜索，選擇較為靠前且能夠很好傳達意思的圖片。同時在每個類別中增添正向與反向的答案。具體材料類型分布請見表 2。

表格 2 實驗材料類型

	詞	句子

	正向	反向	正向	反向
粵方言	4	4	4	4
湘方言	4	4	4	4
閩南方言	4	4	4	4

之所以是要選擇粵方言、閩南方言和湘方言作為閱讀測試材料的語言是因為以下幾個考量因素：

(1) 首先這三者都屬於南方方言中具有代表性的方言，歷來作為方言研究的中心。粵方言的使用範圍主要分布在廣東省內，閩方言的適用範圍主要分布在福建省內，湘方言的適用範圍主要分布在湖南省內。粵閩和粵湘都相互接壤，但這三者之間語音差異很大，從直觀感受來看，三者相互之間的可理解度水平不高。

(2) 粵方言的影響範圍在所有漢語方言中影響最大；閩南方言作為閩方言代表方言，相較於其他子方言，閩南方言的影響範圍都遠遠超出其他子方言區，但總體上來看，閩南方言的影響略遜於粵方言；而湘方言的影響範圍較小。雖然三者影響力各有不同，然而三者之間的方言都保持著各自的獨立性，相互之間的影响較小。因此方言之間的界限分明，彼此之間並沒有相互影響。因此更加有利於對其進行相互理解度的測量。

(3) 最後，若增加更多方言作為實驗材料，必然會增加被試的測試時間。同時若出現大量被試不能夠識別的方言材料，被試會產生一定程度的疲勞和厭倦反應，甚至會做出隨意作答的情況。從而必然會大大影響試驗結果。因此也不宜添加更多方言作為實驗材料。

3.6.2 實驗對象

此項的被試共有 122 人，本實驗通過被試的第一語言進行分組，分別分成粵方言組、閩南話組和湘方言三個組別。其中粵方言組共有 32 人，閩南話組共有 45 人，湘方言組共有 45 人。所有被試均是有償招募，實驗者向所有實驗參與者支付相關報酬或等值的禮品以示回饋。被試絕大部分是在校大學生，其餘一部分是高中生。為保證實驗的有效性和數據的真實性，保證被試的方言背景的純潔性，此次實驗先後在廈門、長沙和澳門三地分別進行。在被試招募的過程中選擇本地人，且要求被試的第一語言一定是本地方言。以下是從性別、年齡等角度分別對實驗被試背景信息的描寫和描述性統計結果。

3.6.2.1 性別

此次試驗的被試中女性共有 66 人，佔全部人數的 54.1%；男性共有 58 人，佔全部人數的 45.9%。其中，第一語言為粵方言的女性被試共有 17 人，男性被試共有 15 人；第一語言為閩南語的女性被試共有 25 人，男性被試共有 20 人；第一語言為湘方言的女性被試共有 24 人，男性被試共有 21 人。雖然男女生比例並不是完美的 1:1 比例，至少保證了每一組中男性被試數超過 15 人，數據有效。

3.6.2.2 年齡

此次試驗的主要被試群體是在校學生，因此年齡主要集中在青少年年齡段。未選擇其他年齡段的主要原因是：考慮到不同年齡層的測試者反應時長可能有較明顯的差異；其次大學生的語言情況基本定型，並且可以通過對這一代學生的數據預測之後的語言狀況；最後，本次實驗旨在測試語言之間的相互可理

解度，因此將一些相關變量進行控制，因此就著重選擇同一個年齡層的被試進行測試。年齡的最小值是 15 歲，年齡最大是 40 歲，平均年齡在 19.39 歲，標準差 3.114，方差 9.698，總體水平趨於集中。本測試中所有年齡均為週歲。其中 15-17 歲共有 16 人，佔總數的 13.1%；18-22 歲共有 99 人，佔總數的 81.2%；22 歲以上共有 7 人，佔總人數的 5.74%。

表格 3 被試年齡描述性統計

	數字	範圍	最小值	最大值	平均值	標準差	方差
					統計		
年齡	5856	25	15	40	19.39	3.114	9.698

男性被試的平均年齡是 19.27 歲，標準差 4.117；女性被試的平均年齡是 19.48 歲，標準差 1.877。粵方言組中，總體平均年齡是 20.22 歲，標準差 2.736，其中男性被試的平均年齡是 20.13 歲，標準差 3.032；女性被試的平均年齡是 20.29 歲，標準差 2.445。閩南話組中，總體平均年齡是 18.16 歲，標準差 2.385，其中男性被試的平均年齡是 16.85 歲，標準差 2.266；女性被試的平均年齡是 19.20 歲，標準差 1.919。湘方言組中，總體平均年齡是 20.02 歲，標準差 3.588，其中男性被試的平均年齡是 20.95 歲，標準差 4.983；女性被試的平均年齡是 19.21 歲，標準差 1.000。

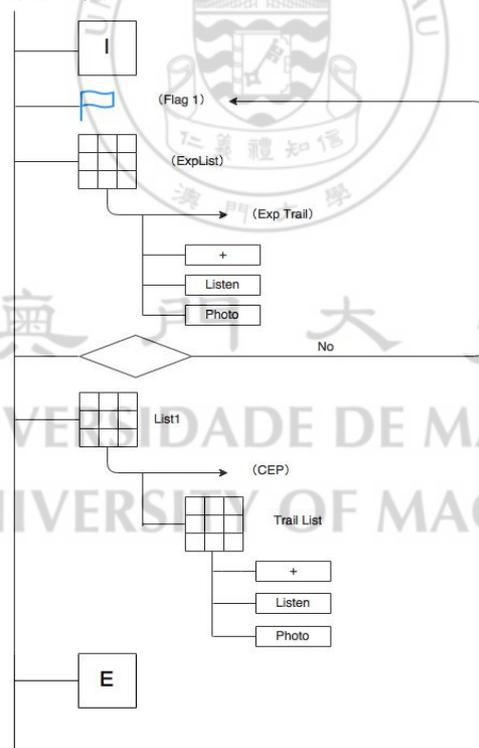
3.6.2.5 教育背景

此次實驗的被試主要組成成分是學生，其中高中生共有 15 人，佔總體的 12.30%，本科生或已擁有本科學歷，共有 105 人，佔總體的 86.06%，在讀研究生 2 人，佔總體的 1.64%。本科生中大一學生共有 32 人，大二學生 24 人，大三學生 37 人，大四學生 9 人。

3.6.3 實驗流程

實驗使用由卡內基梅隆大學、匹茲堡大學學習研究與發展中心和美國心理學軟件工具公司研發的心理學行為研究實驗軟件E-prime 2.0。實驗過程中沒有任何人為干擾因素，完全由軟件進行刺激呈現和數據記錄。實驗共分四大部分：

1、歡迎屏幕：停留時間無限制，在歡迎屏幕上介紹實驗的流程和如何操作按鍵做出反應。鍵盤上「F」鍵表示聽到的內容與看到的內容相匹配，「J」鍵表示聽到的內容與看到的內容不匹配。被試閱讀完畢後，表示已經熟悉實驗流程及操作後，可按任意鍵繼續。



圖表 2 測試流程示意圖

2、練習環節：此環節是為了讓被試更加直觀瞭解到實驗流程及操作而設計。此環節分三個單元。（1）凝視點：屏幕中央上出現凝視點「+」，時長200ms，不需做反應直接進入第二單元；（2）聲音刺激：屏幕上不出現任何信息，被試從耳機中聽到聲音刺激後不需做反應，語音播放完畢後直接進入第三

單元；（3）圖像展示：屏幕上顯示圖片或文字，總時長5000ms，被試需要根據判斷做出選擇，若被試沒有在規定時間內做出選擇，則自動跳轉至下一刺激。

實驗列表中共有4對音頻與圖像刺激組合，兩組匹配，兩組不匹配，其中兩組是圖片，兩組是文字。音頻內容是簡單的雙字詞。實驗列表循環一遍進入選擇界面，此時如果實驗者已經熟悉流程和操作，按「P」鍵進入正式實驗；如還希望繼續練習，按「Q」鍵回到練習模塊初始位置（Flag 1）。

3、實驗模塊：實驗與練習模塊設置相同，包含凝視點、聲音刺激和圖像顯示三個單元，三個單元依次出現，三個單元為一組。不同之處在於，實驗模塊的第三單元記錄被試的反應內容，正確率以及反應時長。實驗列表中共有48組刺激組合，會按照軟件的隨機種子隨機播放，並且每一組刺激只會出現一次。

4、結束屏幕：感謝被試的參與。並結束全部實驗。

實驗全程大約需要3-5分鐘，中間不設置休息時間。實驗結束後，會對被試進行一個簡單的訪談，首先瞭解實驗狀況，是否出現一些非正常的情况，以保證數據真實有效，不是隨意做出的反應。然後會對被試的背景做一個簡要調查，之後詢問被試的第一語言，籍貫、年齡、年級等問題。所有問題不涉及個人隱私，背景調查也不記錄被試姓名，並口頭得到被試的授權允許。調查結束後，會贈送給被試相應的禮品，並請其簽收。

3.6.4 實驗地點

本實驗先後在四所大學和一所高中進行實驗，這五所學校分別位於廈門、

長沙和澳門三座城市。為保證實驗的順利進行，防止噪音對實驗的干擾。實驗地點都選自安靜的教室、圖書館或語言實驗室。具體地點見下表：

表格 4 實驗地點

學校	實驗地點
廈門大學漳州校區	主二樓四樓
廈門一中集美分校	教師辦公室
湖南師範大學	圖書館
中南林業科技大學	學生活動中心辦公室
澳門大學	人文學院語言實驗室

3.6.5 實驗結果

全部實驗過程，主要收集被試的反應（Actual Response）和反應時間（Response Time，簡寫成RT）兩項數據，根據正確反應值（Correct Response）與被試的反應相比較得到正確值（Accuracy，簡寫成ACC）。若被試答對，那麼正確值為1，如果答錯，則正確值為0。反應時間的單位是毫秒（ms）；音頻結束後立即顯示文字或圖片，因此反應時間就從音頻結束後開始計算。本文數據分析所用到的基礎數據就是正確值和反應時間。本實驗共有122名被試，每名被試均有48道題目，每道題目記錄兩項數據，因此本實驗共有 $122 \times 48 \times 2 = 11712$ 個有效數據。本實驗所有的數據分析均使用IBM SPSS Statistics 22版，32位版。

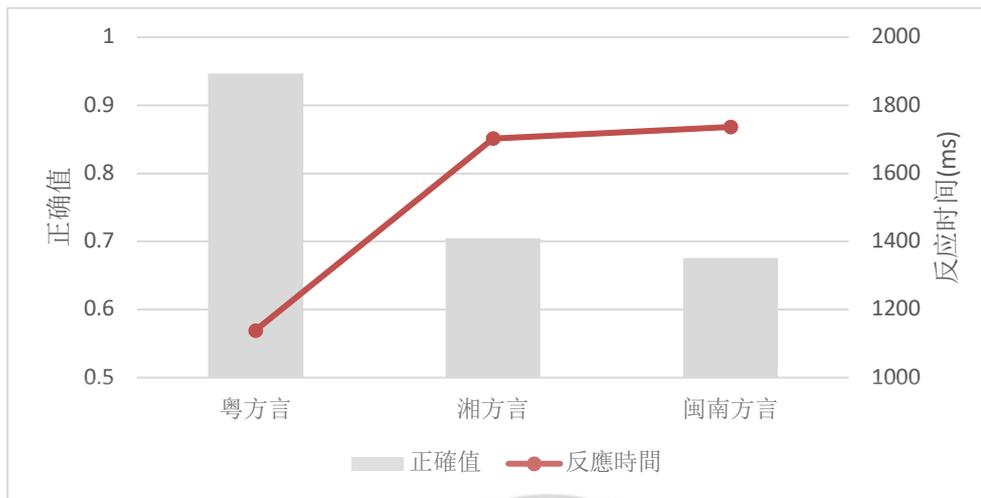
3.6.5.1 正確值與反應時間

本節分別考察不同方言者在聽不同的方言的情況下所表現出的具體數據。細節方面會從聽話者、材料語言、句子詞彙與正向反向四個方面來分別呈現。首先是粵方言者在聽粵方言句子的情況下，正向問題的正確率均值是0.977，標準偏差0.152，反應時間均值為1130.656毫秒，標準偏差667.250，反向問題

的正確值均值是0.977，標準偏差0.152，反應時間均值為1195.297毫秒，標準偏差664.737；在聽粵方言詞彙的時候，正向問題的正確率均值是0.977，標準偏差0.152，反應時間均值為918.633毫秒，標準偏差389.942，反向問題的正確值均值是0.859，標準偏差0.349，反應時間均值為1311.945毫秒，標準偏差986.489；對粵方言整體的正確率均值是0.947，標準偏差0.224，反應時間均值為1024.645毫秒，標準偏差555.653。

粵方言者在聽湘方言句子的情況下，正向問題的正確率均值是0.422，標準偏差0.496，反應時間均值為1911.00毫秒，標準偏差1048.738，反向問題的正確值均值是0.891，標準偏差0.313，反應時間均值為1861.891毫秒，標準偏差952.149；在聽湘方言詞彙的時候，正向問題的正確率均值是0.711，標準偏差0.455，反應時間均值為1434.703毫秒，標準偏差597.631，反向問題的正確值均值是0.797，標準偏差0.403，反應時間均值為1602.773毫秒，標準偏差783.037；對湘方言整體的正確率均值是0.705，標準偏差0.456，反應時間均值為1702.332毫秒，標準偏差881.774。

粵方言者在聽閩南方言句子的情況下，正向問題的正確率均值是0.602，標準偏差0.491，反應時間均值為1708.539毫秒，標準偏差862.088，反向問題的正確值均值是0.758，標準偏差0.430，反應時間均值為1920.258毫秒，標準偏差1164.237；在聽閩南方言詞彙的時候，正向問題的正確率均值是0.539，標準偏差0.500，反應時間均值為1664.281毫秒，標準偏差952.791，反向問題的正確值均值是0.804，標準偏差0.398，反應時間均值為1652.195毫秒，標準偏差745.729；對閩方言整體的正確率均值是0.676，標準偏差0.469，反應時間均值為1736.318毫秒，標準偏差947.182。



圖表 3 粵方言者對粵方言、湘方言與閩南方言三者的正確值與反應時間

以正確值為因變量，利用單因素方差分析發現方言差異與對正確值存在顯著性的影響， $[F(2)=78.370, sig=0.00<0.05]$ ，正向與反向題目也對正確值存在顯著性影響， $[F(1)=54.308, sig=0.00<0.05]$ ，詞彙和句子沒有對正確值呈現出顯著性影響， $[F(1)=0.278, sig=0.592>0.05]$ 。三個變量的交互影響呈現出顯著性影響， $[F(2)=13.389, sig=0.00<0.05]$ 。經過Tukey HSD事後檢驗，聽粵方言材料和聽湘方言材料之間呈現出顯著性差異， $sig=0.00<0.05$ ，聽粵方言材料和聽閩南方言材料之間呈現出顯著性差異， $sig=0.00<0.05$ ，聽閩南方言材料和聽湘方言材料之間沒有呈現出顯著性差異， $sig=0.435>0.05$ 。

以反應時間為因變量，利用單因素方差分析發現方言差異與對正確值存在顯著性的影響， $[F(2)=80.857, sig=0.00<0.05]$ ，正向與反向題目也對正確值存在顯著性影響， $[F(1)=9.026, sig=0.003<0.05]$ ，詞彙和句子對正確值呈現出顯著性影響， $[F(1)=19.559, sig=0.00<0.05]$ 。三個變量的交互呈現出顯著性影響， $[F(2)=3.833, sig=0.022<0.05]$ 。經過Tukey HSD事後檢驗，聽粵方言材料和聽湘

方言材料之間呈現出顯著性差異， $\text{sig}=0.00<0.05$ ，聽粵方言材料和聽閩南方言材料之間呈現出顯著性差異， $\text{sig}=0.00<0.05$ ，聽閩南方言材料和聽湘方言材料之間沒有呈現出顯著性差異， $\text{sig}=0.798>0.05$ 。

根據Pearson相關性檢驗，粵方言聽話者對三種方言的正確值與反應時間呈現出顯著的相關性，顯著性系數 $\text{sig}=0.00<0.05$ ，相關性系數是-0.294，因此二者之間呈現出強烈的負相關關係。

其次是湘方言者在聽粵方言句子的情況下，正向問題的正確率均值是0.700，標準偏差0.459，反應時間均值為1509.933毫秒，標準偏差723.188，反向問題的正確值均值是0.622，標準偏差0.486，反應時間均值為1995.489毫秒，標準偏差956.746；在聽粵方言詞彙的時候，正向問題的正確率均值是0.567，標準偏差0.497，反應時間均值為1736.683毫秒，標準偏差894.274，反向問題的正確值均值是0.867，標準偏差0.341，反應時間均值為1740.089毫秒，標準偏差695.483；對粵方言整體的正確率均值是0.633，標準偏差0.482，反應時間均值為1745.548毫秒，標準偏差799.954。

湘方言者在聽湘方言句子的情況下，正向問題的正確率均值是0.983，標準偏差0.128，反應時間均值為1180.650毫秒，標準偏差596.132，反向問題的正確值均值是0.939，標準偏差0.240，反應時間均值為1532.411毫秒，標準偏差852.310；在聽湘方言詞彙的時候，正向問題的正確率均值是0.972，標準偏差0.165，反應時間均值為1136.056毫秒，標準偏差549.704，反向問題的正確值均值是0.933，標準偏差0.250，反應時間均值為1496.367毫秒，標準偏差900.968；對湘方言整體的正確率均值是0.957，標準偏差0.212，反應時間均值

為1336.371毫秒，標準偏差760.780。

湘方言者在聽閩南方言句子的情況下，正向問題的正確率均值是0.494，標準偏差0.501，反應時間均值為1743.117毫秒，標準偏差762.356，反向問題的正確值均值是0.678，標準偏差0.468，反應時間均值為1960.533毫秒，標準偏差1071.041；在聽閩南方言詞彙的時候，正向問題的正確率均值是0.522，標準偏差0.508，反應時間均值為1690.394毫秒，標準偏差860.177，反向問題的正確值均值是0.683，標準偏差0.466，反應時間均值為1961.067毫秒，標準偏差874.865；對閩方言整體的正確率均值是0.594，標準偏差0.491，反應時間均為1838.778毫秒，標準偏差905.697。



圖表 4 閩方言者對粵方言、湘方言與閩南方言三者的正確值與反應時間

以正確值為因變量，利用單因素方差分析發現方言差異與對正確值存在顯著性的影響， $[F(2)=159.629, sig=0.00<0.05]$ ，正向與反向題目也對正確值存在顯著性影響， $[F(1)=3.504, sig=0.00<0.05]$ ，詞彙和句子沒有對正確值呈現出顯著性影響， $[F(1)=0.245, sig=0.215>0.05]$ 。三個變量的交互影響呈現出顯著性影響， $[F(2)=14.074, sig=0.00<0.05]$ 。經過Tukey HSD事後檢驗，聽粵方言材料和

聽湘方言材料之間呈現出顯著性差異， $\text{sig}=0.00<0.05$ ，聽粵方言材料和聽閩南方言材料之間呈現出顯著性差異， $\text{sig}=0.00<0.05$ ，聽閩南方言材料和聽湘方言材料之間呈現出顯著性差異， $\text{sig}=0.00<0.05$ 。

以反應時間為因變量，利用單因素方差分析發現方言差異與對正確值存在顯著性的影響， $[F(2)=75.705, \text{sig}=0.00<0.05]$ ，正向與反向題目也對正確值存在顯著性影響， $[F(1)=63.005, \text{sig}=0.003<0.05]$ ，詞彙和句子對正確值未呈現出顯著性影響， $[F(1)=0.576, \text{sig}=0.448>0.05]$ 。三個變量的交互呈現出顯著性影響， $[F(2)=5.846, \text{sig}=0.003<0.05]$ 。經過Tukey HSD事後檢驗，聽粵方言材料和聽湘方言材料之間呈現出顯著性差異， $\text{sig}=0.00<0.05$ ，聽粵方言材料和聽閩南方言材料之間沒有呈現出顯著性差異， $\text{sig}=0.081>0.05$ ，聽閩南方言材料和聽湘方言材料之間呈現出顯著性差異， $\text{sig}=0.00<0.05$ 。

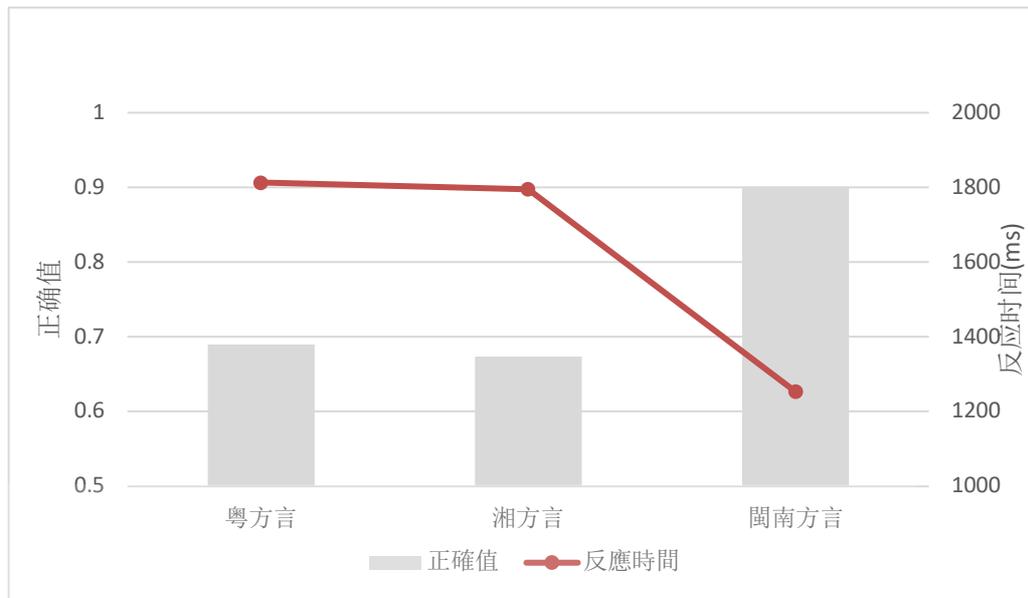
根據Pearson相關性檢驗，湘方言聽話者對三種方言的正確值與反應時間呈現出顯著的相關性，顯著性系數 $\text{sig}=0.00<0.05$ ，相關性系數是-0.250，因此二者之間呈現出強烈的負相關關係。

最後是閩南方言者在聽粵方言句子的情況下，正向問題的正確率均值是0.722，標準偏差0.449，反應時間均值為1449.949毫秒，標準偏差782.097，反向問題的正確值均值是0.700，標準偏差0.459，反應時間均值為1794.760毫秒，標準偏差978.524；在聽粵方言詞彙的時候，正向問題的正確率均值是0.522，標準偏差0.501，反應時間均值為1522.949毫秒，標準偏差847.487，反向問題的正確值均值是0.817，標準偏差0.388，反應時間均值為1678.168毫秒，標準偏差895.088；對粵方言整體的正確率均值是0.690，標準偏差0.463，

反應時間均值為1611.457毫秒，標準偏差888.237。

閩南方言者在聽湘方言句子的情況下，正向問題的正確率均值是0.483，標準偏差0.501，反應時間均值為1635.982毫秒，標準偏差961.348，反向問題的正確值均值是0.867，標準偏差0.340，反應時間均值為1766.564毫秒，標準偏差943.832；在聽湘方言詞彙的時候，正向問題的正確率均值是0.694，標準偏差0.462，反應時間均值為1372.268毫秒，標準偏差683.168，反向問題的正確值均值是0.650，標準偏差0.478，反應時間均值為1632.365毫秒，標準偏差874.910；對湘方言整體的正確率均值是0.674，標準偏差0.469，反應時間均值為1601.795毫秒，標準偏差883.811。

閩南方言者在聽閩南方言句子的情況下，正向問題的正確率均值是0.978，標準偏差0.148，反應時間均值為1446.086毫秒，標準偏差818.563，反向問題的正確值均值是0.917，標準偏差0.277，反應時間均值為1713.404毫秒，標準偏差1056.884；在聽閩南方言詞彙的時候，正向問題的正確率均值是0.867，標準偏差0.341，反應時間均值為1532.520毫秒，標準偏差878.386，反向問題的正確值均值是0.844，標準偏差0.363，反應時間均值為1692.076毫秒，標準偏差873.509；對閩方言整體的正確率均值是0.901，標準偏差0.298，反應時間均值為1596.022毫秒，標準偏差917.342。



圖表 5 閩南方言者對粵方言、湘方言與閩南方言三者的正確值與反應時間

以正確值為因變量，利用單因素方差分析發現方言差異與對正確值存在顯著性的影響， $[F(2)=70.677, sig=0.00<0.05]$ ，正向與反向題目也對正確值存在顯著性影響， $[F(1)=25.441, sig=0.00<0.05]$ ，詞彙和句子對正確值呈現出顯著性影響， $[F(1)=6.768, sig=0.009>0.05]$ 。三個變量的交互影響呈現出顯著性影響， $[F(2)=38.777, sig=0.00<0.05]$ 。經過Tukey HSD事後檢驗，聽粵方言材料和聽湘方言材料之間沒有呈現出顯著性差異， $sig=0.715>0.05$ ，聽粵方言材料和聽閩南方言材料之間呈現出顯著性差異， $sig=0.00<0.05$ ，聽閩南方言材料和聽湘方言材料之間呈現出顯著性差異， $sig=0.00<0.05$ 。

以反應時間為因變量，利用單因素方差分析發現方言差異與對正確值存在顯著性的影響， $[F(2)=95.587, sig=0.00<0.05]$ ，正向與反向題目也對正確值存在顯著性影響， $[F(1)=34.835, sig=0.00<0.05]$ ，詞彙和句子對正確值未呈現出顯著性影響， $[F(1)=0.042, sig=0.838>0.05]$ 。三個變量的交互呈現出顯著性影響， $[F(2)=3.465, sig=0.031<0.05]$ 。經過Tukey HSD事後檢驗，聽粵方言材料和聽湘

方言材料之間沒有呈現出顯著性差異， $\text{sig}=0.922>0.05$ ，聽粵方言材料和聽閩南方言材料之間呈現出顯著性差異， $\text{sig}=0.00<0.05$ ，聽閩南方言材料和聽湘方言材料之間呈現出顯著性差異， $\text{sig}=0.00<0.05$ 。

根據Pearson相關性檢驗，閩南方言聽話者對三種方言的正確值與反應時間呈現出顯著的相關性，顯著性系數 $\text{sig}=0.00<0.05$ ，相關性系數是-0.218，因此二者之間呈現出強烈的負相關關係。

3.6.6 時間參數計算

3.6.6.1 反應時間

此前的理解度測試中，均沒有對測試的反應時間進行記錄，因此也並未涉及關於反應時間的計算方式。本文首次將反應時間融入到理解度的運算中，首先要對反應時間進行系統的描述與論證之後，才能夠使其科學地加入到實際的運算中。

從總體上看，此次實驗共收集到反應時間數據是 $122 \times 48 = 5856$ 個數據，本文中所有的真實反應時間的單位均為毫秒（ms）。本測試中反應時間的極小值是51ms，極大值是5000ms，平均值是1603.09ms，標準差896.46。

3.6.6.1.1 背景差異

本節會從被試的性別、學歷和第一語言等幾個維度進行分析，分析不同條件下的反應時間是否有顯著性差異。首先從性別的方面來看，男性被試的平均反應時間是1724.78毫秒，標準差是942.12；女性被試的平均反應時間是1499.84毫秒，標準差是842.28。經獨立樣本T檢驗後，得到二者之間有顯著性差異， $[T=9.64, df=5854, \text{sig}=0.00<0.05]$ 。因為此次實驗的被試有部分為高中生，其

餘均是大學本科以上學歷，因碩士生被試僅有兩人，數量極低，故將其併入大學本科組進行統計。高中生被試的平均反應時間是1617.95毫秒，標準差是863.53；大學以上學歷的平均反應時間是1601.01毫秒，標準差是901.04。經獨立樣本T檢驗後，得到二者之間沒有顯著性差異， $[T=0.490, df=952.125, sig=0.624>0.05]$ 。最後從被試的第一語言方面來看，分別是粵方言組、湘方言組和閩方言組。粵方言組的被試平均反應時間是1526.01毫秒，標準差是897.71；湘方言組的被試平均反應時間是1640.23毫秒，標準差是856.49；粵方言組的被試平均反應時間是1620.76毫秒，標準差是922.73。經單因素方差檢驗，組間呈現顯著性差異， $[F(2, 5853)=7.97, sig=0.000<0.05]$ 。經Tukey HSD事後檢驗可知，粵方言組與其他兩組被試之間均存在顯著性差異，顯著性系數 $sig=0.00<0.05$ ；而湘方言組與閩方言組之間不存在顯著性差異，顯著性系數 $sig=0.755>0.05$ 。因為並不是所有組別均包含高中生和大學學生這兩個水平，因此不能夠對這三個變量進行交互效應的檢驗。

3.6.6.1.2 實驗材料差異

本節會根據實驗材料的性質不同，分別考察其是否對反應時間產生影響。實驗材料的性質包括語言差異，其中包含粵方言、湘方言、閩南方言三個水平；詞彙句子差異，其中包含詞彙和句子兩個水平；正向與反向題目差異，其中包含正反各兩個水平。首先，語言差異方面，聽粵方言材料的平均反應時間是1611.46毫秒，標準差是888.24；聽湘方言材料的平均反應時間是1601.79毫秒，標準差是883.81；聽閩南方言材料的平均反應時間是1596.02毫秒，標準差是917.34。經單因素方差檢驗，組間沒有呈現顯著性差異， $[F(2, 5853)=0.148, sig=0.863>0.05]$ 。經Tukey HSD事後檢驗可知三組兩兩之間也並未呈現出顯著性

差異。

詞彙和句子材料方面，聽詞彙材料的平均反應時間是1571毫秒，標準差851.67；聽句子材料的平均反應時間是1634.46，標準差是938.20。二者經過獨立樣本T檢驗後，得到二者之間呈現出顯著性差異， $[T=-2.679, df=5854, sig=0.07<0.05]$ 。因此聽句子材料的時間要顯著長於聽詞彙材料的時間，因為句子帶有更多的信息，大腦要處理更多的信息自然需要消耗更多的時間，另外句子材料本身錄音的時長也是長於詞彙材料錄音的時長。

正向與反向題目方面，正向題目的平均反應時間是1712.89毫秒，標準差是940.18；反向題目的平均反應時間是1493.29毫秒，標準差是836.35。二者通過獨立樣本T檢驗後，得到二者之間呈現出顯著性差異， $[T=-9.443, df=5854, sig=0.00<0.05]$ 。可見正向題目的反應時間是顯著高於反向題目所花的時間。通常來說，被試在聽不懂的情況下，更加傾向於選擇否定的選項。

最後經過單因素方差分析，將三者結合起來，得到三組之間共同作用呈現出對結果顯著的交互效應， $[F(2)=4.252, sig=0.014>0.05]$ 。可見在三者的共同作用下產生了對反應時間的交互效應。

本文在計算理解度的時間參數的過程中充分考慮到各個變量對反應時間的影響，根據變量的性質採取不同的手段。對於個體的差異，採用歸一化的方法對其進行統一處理，將其納入到同一範圍內進行考量。而對實驗材料的差異，則採取分類處理的辦法對數據進行加工。具體方法可參加3.6.6.2節的計算過程。

3.6.6.2計算過程

此次試驗對所有被試的所有反應時間進行記錄，因此共有有效反應時間數據 $277 \times 48 = 13296$ 個數據。根據上文所討論的內容將反應時間按照實驗材料的類型進行分開，分別得到句子的反應時間和單詞的反應時間。以下是計算時間參數的逐個步驟：

(一) 歸一化 (Normalization)

不同的人對於相同的事物所需要的反應時間也是不同的，有些人就是反應略快一些，有些人就是稍微慢一些，總體不會相差太多，不過為了提高實驗的科學性，並且考慮到被試都存在著個性化差異，因此首先將反應時間進行歸一化處理。對這個歸一化系數簡寫成N，該系數需要有以下幾個特點：(1) N值的取值範圍是零到一的閉區間([0,1])；(2) N值和反應時間呈負相關關係，反應時間越長，N值越小，相反反應時間越短，N值就越大。這樣下一步計算出的時間參數才能夠表現出反應時間越長，得分越低的效果；(3) N值是個性化的，是因人而異的，需要對每個被試單獨統計和計算。

本文採用線性的歸一方法對反應時間進行歸一化處理，最終的T值是用1減去歸一後的數值，就可以達到與反應時間負相關的效果。採用的公式是：

$$T = 1 - \frac{RT - \text{Min}_{RT}}{\text{Max}_{RT} - \text{Min}_{RT}}$$

在這個公式中，RT是某個被試某道題的反應時間，MaxRT和MinRT分別對應著整個測試中，該被試的最長反應時間和最短反應時間。歸一化的好處是將之間的絕對的反應時長轉化成相對的歸一化數據，使得數據和數據之間擁有可以比較的可能性，同時也是一種數據的簡化計算的方式，將有量綱的表達

式，轉化為無量綱的表達式，成為純量。將一個是在的表示時間長短的數值轉化為一個用於計算、比較、分析的參數，將原始數據映射到[0,1]的區間內部，可以讓數據處理變得更便捷快速。此處以被試5-1作為示例：

表格 5 時間參數計算過程 1

5-1	白菜	山羊	拖鞋	微博	夏天	眼鏡	水杯	……	水果
原始數據	1771	661	874	970	652	970	454	……	768
歸一化	0.132	0.872	0.752	0.574	0.6447	1	0.9937		0.7746

(二) 分類

根據上一節所述，因為實驗材料的性質不同，影響了反應時間的不同，歸一化只能夠將個人因素的影響消除，並不能夠消除實驗材料性質對反應時間所造成了影響，因此這一步需要根據實驗材料的性質來對歸一化之後的數據進行處理。

根據實驗材料性質分別分成 $3 \times 2 \times 2 = 12$ 組，根據上述表格中的分類方式，分別得到被試對不同類型的實驗材料的平均反應時間及相關數據。

表格 6 時間參數計算過程 2

方言	詞彙句子	正確錯誤	均值	標準 偏差
粵方言	句子	錯誤	1794.7602	978.52469
		正確	1449.9488	782.09750
	詞彙	錯誤	1678.1680	895.08825
		正確	1522.9488	847.48679
湘方言	句子	錯誤	1766.5635	943.83212
		正確	1635.9816	961.34785
	詞彙	錯誤	1632.3648	874.91057
		正確	1372.2684	683.16797
閩南方言	句子	錯誤	1713.4037	1056.88391
		正確	1446.0861	818.56306
	詞彙	錯誤	1692.0758	873.50859

		正確	1532.5205	878.38623
--	--	----	-----------	-----------

以上的数据均包括了不同语言背景的被试的反应时间，进一步通过分类，区分本地人的反应时间和其他方言区的反应时间。最终形成 $3 \times 3 \times 2 \times 2 = 36$ 类。

(三) 参照运算

依照上述分类，可以最终得到本地人对本地方言在各种情况下的反应时间，在排除其他非本地人的两个水平量之后，可以得到 $(3-2) \times 3 \times 2 \times 2 = 12$ 类。

表格 7 時間參數計算過程 3

方言	詞彙句子	正確錯誤	均值	標準 偏差	樣本數量
粵方言	詞彙	正確	0.825488	0.2034225	128
		錯誤	0.69976	0.3590547	128
	句子	正確	0.777859	0.2435658	128
		錯誤	0.816137	0.232365	128
湘方言	詞彙	正確	0.778884	0.2473991	180
		錯誤	0.67122	0.325562	180
	句子	正確	0.737815	0.262617	180
		錯誤	0.721466	0.3041391	180
閩南方言	詞彙	正確	0.690535	0.3302428	180
		錯誤	0.67743	0.3172963	180
	句子	正確	0.845703	0.2066286	180
		錯誤	0.729042	0.3157271	180

以上數值是本地人對本地方言的理解平均時間，那麼以本地人理解本地方言的時間作為標準，將該類下的歸一化結果除以本地人理解本地方言的平均時間得到最終的時間參數。其背後的理論依據是，相較於本地人的平均水平，非本地人在同樣題目上的相對位置。這種方法實際上是將每個題目單獨縱向來看，以本地人的平均水平為標準，計算每個數據與標準之間的相對關係。這樣做可以更好地從時間層面顯示出真正理解與後期加工之間的差異。此處以「白菜」這道題目作為處理過程示例：

表格 8 時間參數計算過程 4

白菜	5-1	5-2	5-3	5-4	5-5	5-6	……	5-122
歸一化	0.1326	0.934	0.9617	0.9655	0.9667	0.4865	……	0.6283
參照運算	0.1606	1.0257	1.1651	1.1412	1.1056	0.9027	……	0.843

此步驟結束後，就得到了最終的理解度參數，就將原先的反應時長轉化為可以計算的時間參數。安妮·安娜斯塔西、蘇珊娜·厄比娜（1997）認為設計導出分數有雙重目的。首先導出分數表明瞭個體在團體中的相對位置，因此參照其他個體，可以評價其得分。其次，導出分數提供了個比較的度量，對個體在不同測驗上的得分可以進行比較。

在此過程中，歸一化、分類、參照運算三個步驟不能調換順序，也不可以缺少某個環節。如果失去了歸一化的環節，就沒有辦法消除個體之間的差異，因為有些人對同樣事情的反應可能會比其他人稍慢一些，但是這對於他本身來講已經是在很短時間內做出的選擇，那麼因為他的絕對反應時間比別人稍長，所以就不能得到靠前的排序，以至於他不能得到應有的時間參數，那麼就是錯誤地使用了數據。

如果失去了分類這個步驟，將歸一化的數據直接進行參照運算成時間參數，這樣實際上是將所有的題目看成是一個常模，用一套標準來衡量。題目和題目之間因為性質差異，在反應時間上有顯著性差異，如果籠統來看，可能會使得結果不甚精確。

最後需要再次說明通過這樣計算出時間參數具有動態性。因為隨著樣本數量的擴大，有可能本地人的均值會有所波動，那麼就會似的整體的理解度參數向上或向下浮動。但是總體來看，本地人聽本地方言應該維持在一個基本的水

平上，最終會形成一個常模的範圍，那麼縱使日後增加更多的個案，也並不會造成整體性的波動。

3.7 理解度分析與結論

3.7.1 理解度分析

根據上文所述，將原始的反應時長通過一系列的計算使其轉化為計算理解度的時間參數。利用這個時間參數，與每一題的正確值相乘，就可以得到該題目的得分，最終根據不同的變量提取出相應的理解度數值。得到理解度數值後，就可以根據每名被試的每一道題目的理解度得分計算此次實驗的信度。根據被試第一語言的不同分組，利用得到的理解度數值分別計算克隆巴赫系數，分別是 $\alpha=0.995$ 、 0.994 、 0.995 ，說明此次測試具有極好的信度，經由測試得到的結論是可信的。

本次實驗中男性被試的理解度均值為 0.667 ，標準差 0.510 ，女性被試的理解度均值為 0.646 ，標準差 0.517 。經獨立樣本T檢驗可知，二者之間沒有顯著性差異， $[T=1.554, df=5722.508, sig=0.120>0.05]$ 。高中生被試的理解度均值為 0.6465 ，標準差 0.516 ，非高中生被試的理解度均值為 0.656 ，標準差 0.513 ，經獨立樣本T檢驗可知，二者之間沒有顯著性差異， $[T=-0.481, df=929.841, sig=0.631>0.05]$ 。

前文根據被試的第一語言情況分成了粵方言組，湘方言組和閩南方言三組。下面會先考察這三組內部的理解度結果，然後再考察三組之間的理解度結果。本文的最終實驗目的是考察三種方言兩兩之間的相互理解度。

首先是粵方言組，粵方言組對湘方言句子材料整體的理解度是 0.504 ，標

標準差0.489；其中對句子材料的正向題目的理解度是0.307，標準偏差0.439，反向題目的理解度是0.701，標準偏差0.457；對詞彙材料的整體理解度是0.611，標準偏差0.490，詞彙材料的正向題目的理解度是0.539，標準偏差0.467，反向題目的理解度是0.683，標準偏差0.503。整體上對湘方言的理解度是0.557，標準偏差0.492。

粵方言組對閩南方言句子材料整體的理解度是0.5187，標準差0.481；其中對句子材料的正向題目的理解度是0.419，標準偏差0.423，反向題目的理解度是0.618，標準偏差0.515；對詞彙材料的整體理解度是0.541，標準偏差0.499，詞彙材料的正向題目的理解度是0.442，標準偏差0.520，反向題目的理解度是0.640，標準偏差0.458。整體上對閩南方言的理解度是0.530，標準偏差0.490。

其次是湘方言組，湘方言組對粵方言句子材料整體的理解度是0.475，標準差0.463；其中對句子材料的正向題目的理解度是0.548，標準偏差0.483，反向題目的理解度是0.401，標準偏差0.430；對詞彙材料的整體理解度是0.544，標準偏差0.468，詞彙材料的正向題目的理解度是0.397，標準偏差0.434，反向題目的理解度是0.691，標準偏差0.456。整體上對粵方言的理解度是0.509，標準偏差0.466。

湘方言組對閩南方言句子材料整體的理解度是0.441，標準差0.491；其中對句子材料的正向題目的理解度是0.299，標準偏差0.396，反向題目的理解度是0.583，標準偏差0.507；對詞彙材料的整體理解度是0.469，標準偏差0.505，詞彙材料的正向題目的理解度是0.470，標準偏差0.522，反向題目的理解度是0.467，標準偏差0.455。整體上對閩南方言的理解度是0.452，標準偏差0.491。

最後是閩南方言組，閩南方言組對粵方言句子材料整體的理解度是0.489，標準差0.447；其中對句子材料的正向題目的理解度是0.552，標準偏差0.474，反向題目的理解度是0.426，標準偏差0.409；對詞彙材料的整體理解度是0.475，標準偏差0.454，詞彙材料的正向題目的理解度是0.337，標準偏差0.408，反向題目的理解度是0.613，標準偏差0.458。整體上對粵方言的理解度是0.482，標準偏差0.450。

閩南方言組對湘方言句子材料整體的理解度是0.477，標準差0.467；其中對句子材料的正向題目的理解度是0.327，標準偏差0.431，反向題目的理解度是0.626，標準偏差0.455；對詞彙材料的整體理解度是0.515，標準偏差0.496，詞彙材料的正向題目的理解度是0.551，標準偏差0.478，反向題目的理解度是0.479，標準偏差0.511。整體上對湘方言的理解度是0.496，標準偏差0.482。

綜上可以得到這三種方言兩兩之間的理解度。在前文中論述過，理解度具有不對稱性，因此通過此數據就可以明顯看出，這種方言間單向的不對稱性的差異。粵方言人聽湘方言的理解度是0.557，標準偏差0.492；粵方言人聽閩南方言的理解度是0.530，標準偏差0.490；湘方言人聽粵方言的理解度是0.509，標準偏差0.466；湘方言人聽閩南方言的理解度是0.455，標準偏差0.491；閩南方言人聽粵方言的理解度是0.482，標準偏差0.450；閩南方言人聽湘方言的理解度是0.496，標準偏差0.482。見下表：

表格 9 粵方言、湘方言和閩南方言的單向理解度

	粵方言	湘方言	閩南方言
粵方言人		0.557	0.530
湘方言人	0.509		0.455
閩南方言人	0.482	0.496	

3.7.2 相互理解度及相互理解率

上文中所計算出的理解度均是單向的理解度，本文的主題是計算語言間相互理解度。囿於技術上的限制，暫時不能夠實現實時的相互理解度的測算。暫時只能通過技術手段將相互對話的過程用電腦分割開，讓對話在不同的時空發生。於是目前可以得到六組相異的單向理解度數值，將其轉化為三對相互的理解度。與此之外，還可以通過實驗得到三組本地方言者對本地方言的理解度數值。在計算相互理解率的情況下，先將九組單項理解度數據進行歸一化處理，得到九組理解度的相對理解率，之後將六組相異的單項理解率組合，得到最終的三對相互理解率。

由表9可知，不同方言背景的被試對不同方言的材料呈現出不同的理解度。不同方言背景的被試，在實際上是背後代表著該被試所操方言與目標方言之間的可理解度。所以最終可以抽象出兩種方言之間的理解度，總結出兩種方言之間的相互理解度。最後通過二者之間的相互平均，得到最終的相互可理解度。粵方言與湘方言之間的相互可理解度為0.533，粵方言與閩南方言之間的相互可理解度為0.506，湘方言與閩南方言之間的相互可理解度為0.476。

表格 10 粵方言、湘方言和閩南方言的相互理解度

	單項可理解度		相互理解度
粵方言-湘方言	0.557	0.509	0.533
粵方言-閩南方言	0.530	0.482	0.506
湘方言-閩南方言	0.455	0.496	0.476

鄭錦全（1994：40）曾經對方言之間的距離進行過全面的考察，作者利用音系方面的差異來對中國17個方言點進行了語言距離的計算。最終得到了272個單向溝通度，136個相互溝通度。得出全部漢語方言的平均溝通度是0.518。

其中也包括了本文中所涉及到的粵方言、湘方言和閩南方言三者之間的溝通度。

湘方言者聽粵方言的溝通度是0.412，粵方言者聽湘方言的溝通度是0.454，粵方言與湘方言之間的相互平均溝通度是0.433。粵方言者聽閩南方言的溝通度是0.515，閩南方言者聽粵方言的溝通度是0.434，粵方言與閩南方言之間的相互平均溝通度是0.474。湘方言者聽閩南方言的溝通度是0.448，閩南方言者聽湘方言的溝通度是0.388，湘方言與閩南方言之間的相互平均溝通度是0.418。

將鄭錦全得到的相互溝通度與本文得到的數字相對比，可以說大致屬於同樣的數據分布，因為二者之間是完全不同的實驗方法，數據的處理也是完全不同的。但是所基於的方言是相同的，因此二者之間差異不大。僅從數字上說，本文得到的數字與鄭錦全所得到的數字相比，總體偏高。這一點是非常符合實際情況且正確的。因為鄭錦全的溝通度是基於材料的差異所得到的數字；而本文通過心理語言學測試的方法得到的二者溝通度，從心理層面上看，大腦對與具體語音的容忍度是非常高的，通常情況下可以把兩個完全不同的音處理為一個音位，這樣就會在一定程度上提高相互理解度。

另外Tang（2009：76-78）也對漢語方言進行過理解度測量。她也是分別對句子和詞彙進行理解度實驗，在句子實驗中，被試需要將聽到的句子的最後一個詞翻譯成自己本方言；在詞彙實驗中，被試需要將聽到的詞彙划入詞彙所屬的類別。這樣的出的相互理解度的結果，每個方言有100道題。作者也是利用正確率直接轉化為理解度。粵方言者聽湘方言的句子理解度是8%，詞彙理

解度是23%；粵方言聽閩南方言的句子理解度8%，詞彙理解度是31%；湘方言者聽粵方言的句子理解度是13%，詞彙理解度是19%；湘方言者聽閩南方言的句子理解度13%，詞彙理解度是12%；閩南方言聽粵方言的句子理解度10%，詞彙理解度是25%；閩南方言者聽湘方言的句子理解度是23%，詞彙理解度是24%。

雖然該文也是使用實驗的方式調查了各大方言區之間的理解度，然而她的實驗設計有著明顯的問題，她的實驗中不僅考察了理解度，還涉及到了方言之間的翻譯的問題，以及詞彙的歸類等問題。直觀的來看，廈門本地人做閩南方言詞彙任務的正確率也僅為39%，廣州本地人做粵方言詞彙任務的正確率為55%，長沙本地人做湘方言的詞彙任務的正確率為48%。這個數值基本不能夠反映出真實的情況。因此僅將其作為參考，不進行具體對比。

在實驗數據中還有三組關於本地聽本地方言材料的數據，分別是粵方言被試聽粵方言的句子材料整體的理解度是0.997，標準差0.306；其中對句子材料的正向題目的理解度是1，標準偏差0.313，反向題目的理解度是0.994，標準偏差0.300；對詞彙材料的整體理解度是0.976，標準偏差0.437，詞彙材料的正向題目的理解度是0.985，標準偏差0.284，反向題目的理解度是0.967，標準偏差0.549。整體上對粵方言的理解度是0.987，標準偏差0.377。

湘方言被試聽湘方言的句子材料整體的理解度是0.971，標準差0.422；其中對句子材料的正向題目的理解度是0.986，標準偏差0.370，反向題目的理解度是0.955，標準偏差0.468；對詞彙材料的整體理解度是0.956，標準偏差0.455，詞彙材料的正向題目的理解度是0.974，標準偏差0.357，反向題目的理

解度是0.938，標準偏差0.536。整體上對湘方言的理解度是0.963，標準偏差0.439。

閩南方言被試聽閩南方言的句子材料整體的理解度是0.971，標準差0.393；其中對句子材料的正向題目的理解度是0.994，標準偏差0.261，反向題目的理解度是0.947，標準偏差0.490；對詞彙材料的整體理解度是0.927，標準偏差0.545，詞彙材料的正向題目的理解度是0.953，標準偏差0.529，反向題目的理解度是0.900，標準偏差0.562。整體上對閩南方言的理解度是0.949，標準偏差0.476。

根據目前得到的九組單向可理解度數值，進行線性歸一化處理。將本地人聽本地方言的數值作為全部理解的標準，另一邊本文選擇的三對方言，就是要選擇互相之間不懂的情況，所以就以三對方言之間六對相互理解度作為不理解的標準。最終以這兩個端點作為理解率的格局，得到最終三對方言之間的相互理解率。粵方言與湘方言的單向可理解度為10.16%，湘方言與粵方言的單向可理解度為19.20%；粵方言與閩南方言的單向可理解度為5.03%，湘方言與粵方言的單向可理解度為14.06%；湘方言與閩南方言的單向可理解度為7.64%，湘方言與粵方言的單向可理解度為0.01%。粵方言和湘方言之間的相互可理解率是14.68%，粵方言與閩南方言之間的相互可理解度是9.54%，湘方言與閩南方言的相互理解度是3.82%。

表格 11 粵方言、湘方言和閩南方言的相互理解率

(A—B)	方向	單向可理解度	單向可理解率	相互可理解率
粵方言— 湘方言	A 方言人聽 B 方言	0.557	19.20%	14.68%
	B 方言人聽 A 方言	0.509	10.16%	
粵方言—	A 方言人聽 B 方言	0.530	14.06%	9.54%

閩南方言	B 方言人聽 A 方言	0.482	5.03%	
湘方言—	A 方言人聽 B 方言	0.455	0.01%	3.82%
閩南方言	B 方言人聽 A 方言	0.496	7.64%	

最後本文需要討論理解度數值的性質及意義。理解度本身具有動態性，這種動態性不僅是理論上的，實際操作中也驗證這理解度的動態性。動態性一方面是受試者本身對語言的掌握會逐漸加深，正如實驗中澳門非粵方言受試者處於粵方言母語者組和內地非粵方言受試者組之間，相對於內地非粵方言受試者處於一個上升的進程中。澳門非粵方言受試者內部也是隨著在澳時間的增加，理解度也在總體上升的趨勢。另一方面是理解度的計算機制賦予的，由於計算理解度時的時間參數是在歸一化後，對所有參與測試的受試者的歸一化結果進行排序。也就是說每一個新的受試者的加入都會使得排名順序發生變化，排名順序發生變化就會影響到最終的賦值結果，進而影響到理解度的得分。動態性並不是說理解度數值是無規律的動態性，而是處在一個動態穩定的狀態。最終的理解度數值會隨著大樣本的收集穩定在一個理解度區間內。

最後探討理解度數值所代表的含義。理解度數值是每一名受試者所做的每一道題目產生的，最終通過平均數的方式統計起來。正如前文所述，理解度是測算語言之間距離的尺子，理解度數值就是通過尺子測算出的數值。數值本身除了描述理解度之外，是沒有其他意義的，只有通過對比才能產生意義。就比如2米，1米，這兩個數字本身是無意義的，將這兩個數字賦予同一類事物，就能夠產生對比，比如一個人高2米，另一個人高1米，能夠得到的基本結論是2米的人比1米高，但是還有另一層含義在裡面，就是如果這兩個人都是成年人的話，那麼身高2米的人是高於人類平均身高的，身高1米的是低於人類平均身高

的。之所以能夠得出後面的深層含義，是因為在每個人的大腦中有一個大致人類身高的常模，之後才對新的數值與那個常模進行比較。所以若想賦予一個數字更多的意義，就需要常模的建立。遺憾的是在目前階段對於理解度的常模尚沒有建立，也就是說目前的理解度數值只能夠停留在第一層次的比較，僅能獲得第一層次的意義，尚不能夠達到第二層次的，獲得具有普適性的意義。不過沒有第一層次量的累積，是無法達到第二層次質的突破。因此從這一點來看，又賦予了第一層次更多一層的意義。理解率就給出一個相對直觀的數字，可以直觀的看出兩種方言之間的理解率，可以通俗的理解。

3.8 小結

這一章是利用一個具體真實的理解度測試，進一步從實踐角度深入闡釋理解度測試，並且對傳統的理解度測試進行改進，在理解度測試中增加時間參數。並且在本節中首先用理論的方式論證時間參數的必要性，另外通過實踐的方式計算出時間參數，最終計算出理解度數值並得出結論。

第一節中，本文探討了時間參數的概念，時間參數是利用被試的反應時間，經計算後得出的參數，用以參與理解度的計算。在理解度測試中，時間參數是由被試在測試中的反應時長經由一系列計算得來的。需要注意的是，此處原始數據中是反應時長，而並不是理解時間。之所以選擇反應時長而沒有選擇理解時間，本文給出了三條主要原因。第一是理解度實驗目前屬於行為實驗的層級和範疇，還沒有上升到神經實驗的層級。在行為是嚴重，提取反應時間是通行做法；第二是理解時間和反應時間雖然不同，但是具有等效性；最後是考慮到實驗條件的限制選擇了行為實驗而放棄了神經實驗，因此利用反應時間作為時間系數的基礎數據。

第二節討論了為什麼要添加時間參數。本文認為在選擇類任務中，由於被試做出反饋的過程過於簡單，因此有可能出現並不是完全理解，或者乾脆不理解的情況就做出回答，就是通常所說的「蒙」。另外還有一種更加具有迷惑性的現象就是被試不是完全不懂，而是似懂非懂，是通過調動自己的語言知識輔之以聯想及各種提示將題目做對，而這種做對的方式僅僅是題目正確，卻不是理解的結果，而是大腦一系列複雜分析運算的結果。換句話說，本文所謂的理解是「秒懂」，而經過聯想、回憶等一系列複雜運算之後的懂與真正的「秒懂」本質上是有區別的。因為在正常的語流中，只有「秒懂」才能夠更順暢地交流，否則雙方或者一方要需要對每句話每個詞都進行深入分析後才能夠交流，恐怕這樣的對話很難進行下去。所以為了區分上述情況，就有必要引入時間參數來進入運算。尤其是在處理兩種語言變體沒有語系上的差異，而是分述不同方言支系這類的情況的時候，引入時間參數的必要性愈加凸顯。並且恰恰是這種在一個語言內部，兩種相近的語言變體的理解度往往是爭論的關鍵點，因此更需要引入時間參數來對理解度測試做出精細化地區分和考察。

第三節探討了時間參數的意義，首先，之所以在之前的理解度實踐中沒有引用時間參數，是鑒於前人研究的目的和實驗精度並不需要多引入一個維度的數據就可以得到想要的結果。而在語言學的研究中，本文認為時間參數的引入可以對語言研究有更精細化的分析結果，因為多引入了一個維度的數據，就可以將以往簡單的對錯二元對立擴展成二維的矩陣，擴展了對理解度測試的解釋空間。其次是利用時間參數去修正在實際測試中可能會產生的誤差，從而得到更符合事實的結果。最後是為了將理解度測試納入到科學的測試中，充分借鑒心理學測試的通常範式即採集反應時間這一做法，使得理解度測試本身更具有

科學性，同時更好地建立模式化，規範化的理解度測試。

第四節本文利用前文所考察了三個方言之間的相互理解度與相互理解率。這三個方言分別是粵方言，湘方言與閩南方言，此實驗共調查被試122名，其中32名是以粵方言為第一語言，45名是以湘方言為第一語言，45名以閩南方言作為第一語言。女性被試共有66人，佔全部人數的54.1%，男性共有58人，佔全部人數的45.9%。被試的主要群體是在校學生，其中有15名為高中生，均是來自於閩南方言區。本實驗的被試在徵集過程中主要要求是：1、必須是本地人；2、被試的第一語言和家庭語言；3、被試不會其他方言，且沒有長期在外地生活學習。所有被試均是有償招募，實驗者向所有實驗參與者支付相關報酬或等值的禮品以示感謝。

實驗分別選擇了24個詞和24個句子作為實驗材料，實驗共涉及到三個方言分別是：粵方言、閩南方言和湘方言。實驗材料的產生方式是在現代漢語基本常用詞彙表中簡單隨機抽樣得到的48個詞，然後將24個詞造句。最終得到24個句子及24個詞。實驗材料共有三個變量，除了詞彙和句子之外，還分為三種方言和正向與反向題目。三個變量中，除方言變量有三個水平外，其餘兩個變量均有兩個水平。最終共有12種組合，每種組合都包含四個材料。本實驗分別在澳門、長沙和廈門三座城市進行實驗。

本實驗通過E-Prime軟件，主要收集了正確值和反應時間兩個量，因此共收集了 $122 \times 48 \times 2 = 11712$ 個有效數據，後經過IBM SPSS Statistics進行數據分析。經分析正確值與反應時間之間的分析，通過相關性檢驗，發現正確值與反應時間之間存在著明顯的負相關。正確值越高，反應時間越短，具體表現在三

個方言區內，各個方言來看，本地被試對本地方言的正確值和反應時間均與其他兩組數據表現出明顯的顯著性差異。

隨後本文利用反應時間經過三個步驟計算出時間參數，首先是將被試的真實反應時間通過歸一化處理轉換為在 $[0, 1]$ 區間內的數字，這樣做是為了消除不同被試之間反應時間的實際差異，並且能夠把數字控制在固定的區間內，方便之後的運算和比較。第二步是分類，因為不同實驗材料所需要的反應時間不同，因此本文根據語言材料性質的不同區分了十二個小類。最後一步是參照運算，在這十二個小類中，統計出本地人對該類語言材料的反應時間均值，並以此為參照標準，用該類中所有的歸一化時間除以本地人對該類材料的反應時間均值，得到最終的時間參數。最後利用時間參數與正確值相乘，就得到了該名被試在此題目上的理解度值，可用以進一步統計。

經統計得到，粵方言人聽湘方言的理解度是0.557；粵方言人聽閩南方言的理解度是0.530；湘方言人聽粵方言的理解度是0.509；湘方言人聽閩南方言的理解度是0.455；閩南方言人聽粵方言的理解度是0.482；閩南方言人聽粵方言的理解度是0.496。隨後再將兩個單向的理解度數值合併成一個相互理解度的數值，得到粵方言與湘方言之間的相互可理解度為0.533，粵方言與閩南方言之間的相互可理解度為0.506，湘方言與閩南方言之間的相互可理解度為0.476。

因為理解度數值並不是通常意義上的百分數，其數值造成理解上的不便，所以本文首次引入理解率的概念，在理解度的基礎上進行進一步的歸一化運算，得到一個相對的理解率，這個理解率就可以作為百分數來進行運算和比

較。那麼最終得到粵方言和湘方言之間的相互可理解率是14.68%，粵方言與閩南方言之間的相互可理解度是9.54%，湘方言與閩南方言的相互理解度是3.82%。

本章的最後，對理解度數值的意義和價值進行了簡要的討論。本文認為，理解度的數值正如尺子的讀數。該讀數有兩層含義，第一層屬於比較基礎的淺顯的含義，讀數是多少就代表是多少。根據讀數，可以對其進行比較。例如本實驗中內地非粵方言被試對普通話的理解度就是高於內地非粵方言被試對粵方言的理解度。這一層的含義根據目前的結論是可以得到的。深層的含義則是通過這個讀數對被測事物進行解釋說明甚至歸類。例如身高不超過85釐米的馬就是矮腳馬，身高160釐米的馬就是正常的馬。通過數字對被測物體進行推導的過程隱含著對該物體的類屬有一個常識性的認識，即在140釐米到180釐米的馬就是正常的馬，超出這個範圍的馬就是其他品種。目前理解度數值沒有得到這樣的常模，暫時不能夠對其進行推導。若想達到這個層次，就需要進行越來越多的理解度測試，將現實世界事物的屬性與數字聯繫起來，建立起理解度常模，從而就可以通過理解度數值對被測的語言進行判斷和歸類了。

第四章 粵式普通話理解度研究

4.1 普通話與普通話水平測試

普通話作為中華人民共和國通用語是以北方話為基礎方言，以北京音為標準音，以典範的現代白話文作為語法規範。自從1955年「全國文字改革會議」和「現代漢語規範研討會」確定普通話的明確定義以來，普通話的普及已經隨著大眾媒體和政府的大力推廣之下進入了人們生活的方方面面。普通話已經成為中國人語言生活中必不可少的一部分。2000年10月31日，九屆全國人大常委會第十八次會議通過了《中華人民共和國國家通用語言文字法》，這是我國第一部關於語言文字的專門法律。其中第三條規定國家推廣普通話，推行規範漢字，以及第二十條規定了對外漢語教學應當教授普通話和規範漢字。從法律層面上確定了普通話的應用場合，規定了廣播電台、電視台以普通話為基本的播音用語，以法律手段維護了普通話的地位。普通話也建立了自己的測試體系，形成了獨立的評價標準。根據《普通話水平測試等級標準（試行）》（1997）的內容將普通話的等級分為三級六等。

一級：甲等朗讀和自由交談時，語音標準，詞彙、語法正確無誤，語調自然，表達流暢。測試總失分率在3%以內。乙等朗讀和自由交談時，語音標準，詞彙、語法正確無誤，語調自然，表達流暢。偶然有字音、字調失誤。測試總失分率在8%以內。

二級：甲等朗讀和自由交談時，聲韻調發音基本標準，語調自然，表達流暢。少數難點音（平翹舌音、前後鼻尾音、邊鼻音等）有時出現失誤。詞彙、語法極少有誤。測試總失分率在13%以內。乙等朗讀和自由交談時，個別調值

不准，聲韻母發音有不到位現象。難點音（平翹舌音、前後鼻尾音、邊鼻音、fu-hu、z-zh-j、送氣不送氣、i-ü不分、保留濁塞音和濁塞擦音、丟介音、復韻母單音化等）失誤較多。方言語調不明顯。有使用方言詞、方言語法的情況。測試總失分率在20%以內。

三級：甲等朗讀和自由交談時，聲韻調發音失誤較多，難點音超出常見範圍，聲調調值多不准。方言語調較明顯。詞彙、語法有失誤。測試總失分率在30%以內。乙等朗讀和自由交談時，聲韻調發音失誤多，方音特徵突出。方言語調明顯。詞彙、語法失誤較多。外地人聽其談話有聽不懂情況。測試總失分率在40%以內。

關於普通話測試的研究根據聶丹（2011）統計，已有1203篇，大體分為12個大類。近年來普通話等級測試也從傳統的人工測試（簡稱「人測」）升級成計算機輔助測試（簡稱「機測」），利用計算機技術，使評判結果更加準確，尺度統一，減少了人為因素導致的誤差，同時也大大降低了人力和物力成本。由此也可以看出計算機輔助測試也代表了語言測試的潮流和趨勢，對於計算機輔助測試也已有相關研究對其進行關注。例如：張雷、侯傑泰、何偉傑、文劍冰、王渝光（2001），雷峻（2010），葉軍、王頤嘉、周鳴之、喬麗華（2007）。

目前普通話測試已經是一項成熟的語言測試，理解度測試才剛剛開始，理解度測試與普通話測試的關係是怎樣的，二者之間的差異是什麼，理解度測試可以向普通話測試借鑒哪些內容；對上述問題的討論能夠更好認識到普通話測試與理解度測試的相關性，並且幫助理解度測試更好地應用在普通話研究領

域。

二者之間的差異主要體現在一下三個方面：

（一）參照系統

前文有所論述，理解度測試屬於常模參照型語言測試，而普通話測試的測試性質在學界尚有爭議。仲哲明（1997）在文章中提到普通話測試的性質眾說紛紜，有的說是一種常模參照測試，有的說是目標參照測試，有的說是標準參照測試，有的說是領域參照測試，也有說是常模參照方法同標準參照方法相結合的測試等等。經過作者論證，作者認為普通話水平測試基本上屬於目前比較通行的所謂標準參照性或者說達標性測試的範圍，而不是選拔性的常模參照性測試。因為播音員、教師等工作對普通話等級有所要求，所以從另一個角度上講，作者提出普通話水平測試實際上也是一種資格證書考試。另外張雷、侯傑泰、何偉傑、文劍冰、王渝光（2001）也認為普通話水平測試屬於標準參照測試。王玲玲（2007）認為普通話水平測試屬目標參照性測試。在不同級別、不同語音層次的測試中，音節正誤的標準是以普通話的終端標準為參照的。本文認同將普通話水平測試看作是標準參照測試，因為最終結果是給出一個等級和標準，這符合標準參照測試的基礎定義。不過本文在此不討論普通話水平測試的性質問題。無論普通話的性質是標準參照還是目標參照，與理解度測試的常模參照還是有本質性區別的。

（二）測試內容

理解度測試主要是對被試進行近似聽力測試，主要考察被試對實驗材料的理解能力，而普通話水平測試目前主要是對考生進行口語測試。仲哲明

(1997) 曾經對普通話水平測試是否需要增加聽力測試進行過詳細的討論，其主要觀點是聽力考試主要應用於第二語言水平測試；普通話聽易說難，如果會說了，聽自然不成問題；況且考普通話資格的考生大多都是已經有相當高的普通話水平的教師或播音人員，因此聽力考試沒有必要。田小琳（2001）應和了此觀點，並補充認為在內地聽力考試確實沒有必要，而在香港地區，普通話的聽力對香港考生還是存在一定難度，因此建議在香港應設立普通話聽力測試。即便普通話水平測試設立了聽力測試，與理解度測試也有著本質性的區別。因為普通話水平測試是對一種語言能力的測定，也是對一種經過練習後的語言能力的測定，其結果是標準與不標準的差異；而理解度測試並不需要被試經過訓練，更多的是需要被試呈現出一種自然的狀態。其結果是理解得對與錯，理解對的數量與速度，對一個個體而言並不存在標準和不標準的評判。

（三）測試目的

普通話水平測試的目的相對單一，仲哲明（1997）在文中給出三點測試目的：

1、評定應試人普通話水平所達到的等級，落實上述普及普通話的質的要求，即對不同人提出的不同的等級要求。這是測試的直接目的。

2、通過測試，更好地貫徹新時期推普工作方針，促進普通話的進一步普及，並在普及的基礎上逐步提高全社會的普通話水平，提高現代漢語的規範化程度。這是開展測試的一個重要目的。

3、促進推普工作進一步走上制度化、規範化、科學化的軌道，這是水平測試正常發揮自身功能的必然結果。作者給普通話水平測試總結的三點目的歸

根結底是圍繞著普通話，也僅僅是服務於普通話。而理解度測試的目的就不僅限於單獨一種語言或一種方言，而是將任何一種語言納入到理解度測試的範式中。這也體現了理解度測試具有更強的適配性，其最終結論也具有更強的適用性和解釋力，可以和同類研究在同一平台上相對比。

理解度測試能夠從普通話水平測試中借鑒和吸收以下幾點：

(一) 計算機輔助測試技術已經全面應用在普通話水平測試中，並且已經收到了很好的效果，極大地節約了人力物力，更重要的是利用計算機技術控制了標準和尺度，能夠將各個地區不同考點之間的誤差降至最低。因此計算機輔助技術也是理解度測試的發展方向，也需要借鑒普通話水平測試中的計算機輔助技術，形成一套成規模成體系的軟件系統，從而可以更高效地進行理解度測試。

(二) 普通話水平測試的理論成果豐富，經過數十年眾多學者的研究與討論，普通話水平測試已經成為了一項重要的課題，並且理論成果可以直接指導實踐，實踐結果又可以反過來驗證補充理論，形成了一個良性的循環。理解度理論尚不能達到普通話水平測試這種理論和實踐良性互動的階段，理解度理論和實踐都還需要進一步完善，在逐步完善的過程中能夠對更多的語言現象和語言事實進行探討和解釋。

4.2 方言普通話

4.2.1 方言普通話

經過數十年的普通話推廣工作，普通話推廣的進行達到了一個前所未有的深度和廣度。從深度上看，很多城市中的家庭已經不再用方言作為家庭語言，

因而孩子的第一語言便直接是普通話。越來越多的孩子失去了方言背景，在教育語言中也是直接以普通話作為教學語言，在學校範圍內杜絕方言的存在。從廣度上看，國家語言文字事業「十三五」發展規劃（教育部、國家語委 2016：3）「十二五」期間，特別是黨的十八大以來，國家通用語言文字普及程度進一步提高，普通話普及率達到70%以上。普通話普及情況調查項目組（2011）對河北省普通話的使用進行調查，調查顯示在2000年有52.58%的人可以用普通話交流，到2010年可以使用普通話交流的比例上升到73.70%。從全國的角度來看，已經有將近10億人能夠使用普通話進行交流。

「推普」工作的成功卻又引發出新的問題和現象，推廣普通話這項工作的目的是讓越來越多的人使用普通話，其實很大程度上是讓越來越多的人轉用普通話，從講方言變成講普通話。因此方言似乎和普通話成為了一種莫名的「對立」關係。隨著推普的持續進行，在普通話與方言的接觸的過程中，普通話與方言的結合產生了一種新型的語言變體，各方言區的人在學習和使用普通話的過程中，受自己母方言影響而形成的、帶方言色彩的普通話（鄒遠春2015）。方言向普通話靠攏是「推普」工作的一個重要成就，也反映了漢語走向統一的趨勢，它的發展是社會各種因素發展的結果（郭熙 2001）。從這個角度上看可以說方言普通話是推普的產物。方言普通話的父母應該是普通話和方言。

所謂龍生九子各有不同，方言普通話也是根據不同的與普通話產生出不同的方言普通話，幾個比較有代表性且研究較多的方言普通話有粵式普通話（又稱廣式普通話，港式普通話等）、四川普通話、長沙普通話、上海普通話、山西普通話以及東北普通話等等。隨著方言普通話逐漸形成，有些方言區的方言普通話形成了固化的方言普通話，有些方言區尚未形成固化的方言普通話，牛

芳（2009），勁松，牛芳（2010）所提到的長沙的「塑料普通話」就已經形成了一套定型的，系統穩定的，語感統一，功能明確的普通話變體。這類固化的普通話變體已經形成一種介乎於普通話與方言之間穩定的語言變體，它擁有自己的一套語音系統，社會屬性和交際功能，音系內部可能會有一些自由變體，但是總體上呈現出一個穩定的狀態，外人可以對其進行準確的感知。對於長沙本地人來講，他們可以自由切換長沙方言，塑料普通話及普通話，在不同場合使用不同類型的語言變體。另一類尚未形成一個穩定的方言普通話變體形式。以粵式普通話為例，在粵方言區內尚沒有形成一個普遍接受的粵式普通話的穩定狀態，雖然在語音語法詞彙等方面都表現出一些特點，但是缺乏系統性，自由度高，是在普通話特徵和粵方言特徵中間游走的狀態。對於本地人來說，要麼是講方言，要麼是講普通話，對他們來說並沒有選擇去講粵式普通話這個選項。所以粵式普通話的範圍相對於已經固化成型的方言普通話來說，範圍就更大，邊界更模糊，狀態更不穩定。因此粵式普通話更多是非粵方言區人對於粵方言區人所講的帶有粵方言特色的普通話的一種稱呼。

近三十年來，湧現了大量的方言普通話研究。其中有偏重描寫的本體研究，比如：陳亞川（1987）、姚佑椿（1988，1989）、周潤年（1990）、王文虎（1994）、鄭永輝（1996）、曾子凡（2000）、徐紅梅（2003）、王詠梅（2006）、田皓（2006）、張建強（2008）、陳蒙（2009）、於芳（2009）、葉軍（2012）；有從語言接觸角度進行的接觸研究，比如：顧欽（2007）、陳建偉（2008）、傅靈（2010）、曹曉燕（2012）；有從語言規劃角度研究如何規範方言普通話的問題，比如：陳建民、陳章太（1988）、陳海燕（2001）；有從偏誤分析角度對方言普通話進行認知語言學研究。比如：許長安

(1991)、顧英華(2003)、周萍(2006)、王磊、張穎(2007)、韓玉華(2007)；還有從聲學角度對方言普通話進行實驗語音學的分析研究，比如：李愛軍、王霞、殷治綱(2003)、韓維新(2010)、銀沛月(2010)、李寧、孟子厚、李蕾(2011)；還有一類就是對方言普通話的歸屬以及標準問題進行研究，比如陳亞川(1991)、張建強(2005)、李如龍(1988)、侍建國、卓瓊妍(2013)、徐傑、董思聰(2013)、董思聰、徐傑(2015)。

關於方言普通話的性質歸屬問題已經經過近三十年的討論。例如：李如龍(1988)將方言和普通話之間存在的既不是方言、又不是普通話的語言稱之為過渡語。作者認為過渡語是方言和普通話這一對對立統一物之間的中介，是方言和普通話之間相互影響的中間環節。方言區的人學習普通話，過渡語是必經之道。對於過渡語，作者認為過渡語兼有普通話和方言兩種成分。這兩種成分的構成因地而異，因人而異，並且處在不斷的變化中。但是，在一個方言區的一個時間里，多數人所說的過渡語的兩種成分構成又是大體一致的系統性的。過渡語會受到文化程度等因素的影響。作者認為過渡語具有模糊性，也有系統性，這二者是相互依存的特徵(陳亞川1991)。分析了方言普通話的性質、特徵及其產生根源；方言普通話的研究意義和目的；方言普通話的研究對象和研究方法；以及普通話的名稱選擇等問題。可以說是在國內學術界較早一批詳細並且成系統地審視、討論和分析方言普通話的文獻。文章利用最多的篇幅分析證明瞭該文的核心觀點，即「方言普通話與方言向普通話靠攏不是一回事」。作者認為方言向共同與靠攏事語言演變的問題。而方言普通話是一種中介語現象。並且方言普通話形成的原因是會受到母方言的負遷移、普通話知識的干擾和語言訓練的缺陷等原因。張建強(2005)在梳理了中介語的理論內涵和外延

後認為「方言普通話」的實質就是普通話，只不過是一種帶有地方色彩的不夠標準的普通話，是標準普通話的區域變體。並且還認為把各地域方言區學習普通話看成是一種介於第一語言學習和第二語言學習之間的語言學習，即母語標準語學習。陳遵平（2012）就方言普通話的範圍、類型和規劃落實等問題提出看法。作者認為方言普通話可以區分廣義和狹義之區別。並從母語影響的角度分成四類，帶漢語方言色彩、帶民族語色彩、既帶民族語又帶漢語方言色彩以及帶外語色彩的非標準普通話等四種類型。作者建議從接受過程的角度展開社會語言學語言規劃研究。侍建國 卓瓊妍（2013）提出以北京語音為標準音已經脫離了全民通用語的本質。經過半個多世紀的推普工作，再把通用語和標準語合二為一，不利於通用語的進一步推廣，也不利於漢語的國際推廣。因此文章建議將通用語和標準語並列。重視「約定俗成」的語言規範，理順通用語和標準語的關係。

以上列舉了關於方言普通話的幾位大家之言，總結出「中介語」、「過渡語」、「通用語」等幾個重要概念。根據魯健驥（1984）的定義，中介語指的是由於學習外語的人在學習過程中，對於目的語的規律所做的不正確的歸納與推論，而產生的一個語言系統。這個語言系統既不同於學習者的母語，又區別於他所學的目的語。中間語系統在語音、詞彙、語法、文化等方面都有表現。但它又不是固定不變的，而是隨著學習的發展逐漸向目的語的正確形式靠攏。從這個定義中我們可以看出，中介語是發生在學習外語的過程中產生的語言系統。這個語言系統不是一成不變，是向目標語言不斷靠攏的。中介語的概念在二語習得的領域是非常常用的。不過，一種方言不能視為母語，而只能稱為「母言」（李宇明 2003）。普通話對於中國人來說是民族的共同語，並不是第

二語言，當然並不包括第一語言是少數民族語言的少數民族同胞。即便是在粵方言區，粵方言與普通話的差別比較大；尤其在港澳地區，周遭的語言環境也都是以粵方言為主，香港多數人的母言是粵方言，但其母語仍然應當看作是漢民族共同語（李宇明 2003）。因此，本文並不認為應該將方言普通話看作是中介語。

上文中簡單梳理對於方言普通話的重要研究成果，鄒遠春（2015）在文中曾經指出對於方言普通話的研究還存在一些不足之處：其一，注重對方言普通話的調查描寫，但對其形成機制和動態演變的研究不夠；其二，個案研究多集中於非官話方言區，官話方言區的成果較少；其三，研究過程未能更好地凸顯方言普通話的應用價值。

本文認為對於方言普通話的研究還遠遠不夠，雖然對於來源問題、使用情況等問題沒有太大的爭議，而更多的問題出現在方言普通話的性質問題上，或者說是屬類問題上。方言普通話究竟是應該屬於普通話還是應該屬於方言，還是應該自立一類。這就類似方言和普通話作為方言普通話的父母親，父母親感情不和要離婚，那麼孩子究竟應該歸屬於誰？應該歸屬於爸爸呢，還是應該歸屬於媽媽，還是自己獨立成家。如果簡單來判定，孩子跟誰比較像就歸誰，似乎是一條簡單可行的辦法。然而迎面而來的問題是，方言普通話中的很多特徵既不存在於普通話中，也不存在在方言中。這就比如父親黑眼睛，母親藍眼睛，而孩子是綠眼睛。

判定方言普通話的歸屬問題歸根到底是標準的問題。對於普通話而言，按照目前國家和學界通行的標準應該是以《現代漢語詞典》《新華字典》等辭書

作為依據，規範著普通話的語音系統和詞彙系統。方言的標準則是以各大方言區的方言志或方言研究作為語音、詞彙和句法標準。以上這些標準都有著明確的音系規則、詞彙系統，句法規範，並且無論講話者還是聽話者擁有著相同的心理感知，普通人聽到其他人在講普通話，會自然地生成一個標準或不標準的判斷，可以給出一個「這個人的普通話很標準」或者「這個人的普通話不標準」的結論。方言也是相同，以粵方言為例。通常外來人員初學粵方言，尤其是北方人對於入聲的輔音韻尾不熟悉，經常會將省略輔音韻尾，造成發音的不標準；另外如果對廣東人講：「我今天坐公交車上班。」他們會理解句子的意思，但是並不是地道的粵方言，地道的方言會說：「我今日搭巴士翻工。」；最後如果想表達：「我給你兩百塊錢。」的意思，用粵方言要說：「我畀兩百蚊你。」上述例子就是從語音、詞彙和句法三個方面展示方言的標準，這個標準是很清晰的，明確的。「我畀兩百蚊你。」就是正確的標準的粵方言，而「我畀你兩百蚊。」就是不標準的，聽起來有些怪異感的句子。

當然更科學的方法是通過等級測試進行量化測量，普通話可以通過普通話水平測試獲得相應的等級，粵方言目前官方的等級測試。不過，雖然可以通過水平測試獲得相應的等級，而相應的等級背後意味著什麼，不同的等級究竟有什麼樣的差別，還需要回到當初等級制定的文獻中尋找。陳章太（1997）在文中曾經建議：1、各級標準的描述前面，最好加一句總的描述語，如一級為「標準的普通話」，二級為「比較標準的普通話」，三級為「合格的普通話」。作者在文中回憶對三級普通話的描述從最早的「不很標準的普通話」到「及格的普通話」，後來又提出「不標準的普通話」「不太標準的普通話」「基本標準的普通話」「大體標準的普通話」「基本合格的普通話」「方言普通話」。

最後提出用「合格的普通話」來概括、形容第三級普通話。因為首先它表明這一級是普通話，是基本標準的普通話；第二，在含義和意念上與「比較標準」的二級普通話有一定差別；第三，用「合格」這個詞，跟這一級的總失分率不超過40%相對稱，都表明是及格的。最後作者還建議在三級之外，似可增設一個預備級（至少地方性的普通話水平測試標準可以補充這一級），稱為「不標準的普通話」。總結來看，作者對普通話三個等級的描述加上最後所謂預備級的描述是：

標準的普通話——比較標準的普通話——合格的普通話——不標準的普通話

這四個等級看起來確實構成了一個連續統，普通話水平從高到低。然而卻不能在這四個等級裡面看到究竟標準的線應該劃在哪裡。如果第一級是標準的普通話，那麼二三級和預備級都應該屬於不標準的普通話，然而到了預備級才是不標準的普通話，如果從所謂預備級的描述來看，高於第四級的普通話都應該是標準的普通話。不免看出這裡面對於普通話等級的描述都存在著自相矛盾的現象。也就是說，無法斷定第二等級和第三等級究竟是不是標準的普通話。等級的意義就是用來判定什麼等級是標準的，什麼等級是不標準的，如果等級與標準的對應都還是很模糊，這樣的話等級本身就失去了意義，等級也沒有最終明確地表示出來。

若暫時不考慮所謂的預備級，只看前三個等級。那麼除了第一等級的普通話是標準的普通話外，剩下兩個等級均不是標準的普通話。那麼，普通話作為我國的民族共同語，是十幾億中華兒女的母語，也就是說存在著上億的中國人

不能夠講標準的母語。這一方面在理論上不符合邏輯，另一方面這樣的結論也傷害了大多數人民的情感和民族自信。

普通話的考試是經過專家論證，國家認證的，其結果對普通話的標準問題尚存在灰色地帶。何況其他方言的標準性問題。普通話和方言的標準都是一個不確定因素，那麼方言普通話的標準又該要如何界定依舊是一個棘手的問題。而且如果想要找到方言普通話與普通話和方言之間的關係，需要將三者置於同一平面進行考量，才能夠得到三者之間的關係。目前學界尚沒有找到一個真實有效的方法可以同時衡量這三者。只有將方言、普通話、方言普通話三者放到同一把尺上比較，才能夠得到三者之間的格局，誰遠誰近，一目瞭然。

4.2.2 粵式普通話

普通話與粵方言相接觸產生的方言普通話就是粵式普通話，與其相似的還有廣式普通話和港式普通話等概念。這幾個概念內涵大致相同，指的都是港澳和廣東地區粵方言使用者所講的普通話。港式普通話的外延範圍略小一些，特指在香港地區內香港人所講的普通話。與內地廣東人所講的普通話略有不同的一些地方在於，香港地區的普通話可能會有更大的比例直接在普通話中添加英文詞彙，或借詞。這與香港社會文化有著密切的關聯。本文所研究的對象是方言普通話，因此粵式普通話更能夠體現是帶有粵方言特色的普通話。這個概念也可以包含廣式普通話和港式普通話甚至是澳式普通話，也可以廣泛地應用對廣式普通話和港式普通話的文獻材料。特徵

傅兩賢（1984）曾在《粵方言區人學習普通話趣談》一書中總結了各種粵方言區人學習普通話常見的錯誤，作者將其分為聲母、韻母、聲調、輕聲和兒

化、詞彙、語法諸多章節，每個章節中又有詳細的分類，其中語音類共有五十類容易讀錯的類型，語法部分共有七類容易出現錯誤的地方。全書以趣談的方式詳細描述了粵式普通話的常見問題，事實上就是將粵式普通話的特徵描寫出來。作者對每個問題的描述都使用兩個容易混淆的詞語作為標題，例如在聲母部分就有「老王」和「老黃」，「荒唐」和「方糖」這樣的對立，韻母部分則有「二姐」和「二者」，「通訊」和「通順」這樣的對立，在語法部分則列舉了「走先」，「給一本書我」和「我高過你」這類典型的粵方言句法特點。

曾子凡（2000）較為全面地描寫了「港式普通話」的主要表現：1、完全照搬廣東話；其中包括（1）完全照搬；（2）一字之差詞語誤用；（3）同形異義詞誤用；（4）同義詞吸管的誤用；（5）量詞誤用；（6）聯合雙音詞拆開誤用；（7）次序語序不當；（8）語體不協調；（9）外來詞誤用；（10）貶義詞誤用。2、誤譯、硬譯，3、其他方式的錯誤：（1）多此一舉的對譯；（2）近義詞誤用。

作者在文中還列舉了許多港式普通話的案例，將廣東話和普通話對照來看，港式普通話最突出的表現在於雜糅，將廣東話的特徵雜糅進普通話中。作者對港式普通話的描寫非常全面細緻，讓讀者可以直觀地看到港式普通話和普通話、粵方言之間的差異。

廣東話	港式普通話	普通話
全身無力	全身沒力	渾身/ 全身沒力氣
好似豬咁盆	好似豬一般蠢	像豬那麼笨

件衫好污糟

衣服很骯臟

衣服很臟

石定栩（2006）和石定栩 邵敬敏 朱志瑜（2006）都對港式中文進行了詳細的描述，石定栩（2006）的《港式中文兩面睇》詳細分析了超過150個容易造成誤解的港式詞彙，分別從英式中文、粵方言、文言文、同義詞、同音詞和創新幾個方面來對港式中文的特點進行詳細描述和分析，解釋詞語產生的源頭和用法。在石定栩、邵敬敏、朱志瑜（2006）《港式中文與標準中文的對比》中將香港書面語分成三個等級A級是標準中文，內地人的可懂度為95%以上，少數詞語或個別格式具有地方特色，只不過量控制在5%以下。B級是港式中文，內地人的可懂度為50%-95%，是指具有香港特色的漢語書面語。C級是粵方言中文，內地人的可懂度為50%以下，是指粵方言的詞彙、格式比較普遍，還出現大量粵方言特有的方言漢字。

然而港式中文與粵式普通話之間有著重大的區別，在於港式中文是書面語，而粵式普通話是口語，以往的文獻更多的將關注點放到了書面語上，因為書面語的語料更容易收集，並且可以擁有較為成熟的研究方法對其進行深入地研究。港式中文不能夠代表粵式普通話，因為粵式普通話的特殊性不只是體現在方言詞彙句法上，而是更加直接地體現在語音部分，粵式普通話的語音也是因為雜糅了大量的粵方言的語音特點導致其往往很難被人聽懂，甚至產生歧義。日前網上流傳著這樣的一段話：

「大扎好，我系渣渣輝，探輓懶月，介四里沒有輓過的船新版本，擠需體驗三番鐘，里造會乾我一樣，愛象節款遊戲。絲兄弟，就來啃我。」

這是香港明星張家輝代言的一款網絡遊戲的廣告語。因為明星所講的普通

話並不標準，而被網友用相似的字將其記錄下來。其實內容是：「大家好，我是張家輝，貪玩藍月，這是你沒有玩過的全新版本，只需體驗三分鐘，你就會跟我一樣，愛上這款遊戲。是兄弟，就來砍我。」可見其實原文中並沒有摻雜粵方言的詞彙，也並沒有改變語序，誤用量詞等，而僅僅是語音形式上的差異。並且正是由於這些語音上的差異，才造成了粵式普通話很難聽懂的现象。例如上文所引用港式普通話的例子中，全身沒力，好似豬一般蠢，衣服很骯臟。這些句子如果用普通話的語音系統讀出來，確實不如第三欄讀起來那麼地道流暢，但並不會對理解產生太大的問題。由此可知，粵式普通話的獨特性首先體現在它的語音系統上，其次才是它的詞彙和句法上，同時影響粵式普通話理解的首要因素也是體現在它的語音系統上，其次是它的詞彙和句法。

本文本章是將粵式普通話的理解度作為研究對象，其主要原因是以下幾個方面：

(1) 粵式普通話在語音、詞彙和句法方面都體現著與其他方言普通話明顯的特殊性，因為其包含著豐富的地域特點和歷史因素，並且目前對粵式普通話的研究較為豐富，不過大多停留在介紹，描寫和調查的階段，還未能對其進行深入細緻的考察。其原因在於方言普通話本身的性質特點還有很多不確定因素，尚且沒能得到學界的共識。以理解度作為一個較新的視角去探查這個複雜的系統，可以突破介紹和描寫的表層，使得粵式普通話與心理、理解、認知等層面的內容結合，從而進入一個較為深入的研究層次。

(2) 粵式普通話從目前階段來看尚未能固化成一個定型的，系統穩定的，語感統一，功能明確的普通話變體，還處於一個變化的階段。從此時對其進行

跟蹤研究，也許能夠觀察到粵式普通話固化的過程，或是並不到達固化的階段而一直保持一個動態的狀態。這個歷時的研究需要使用同一種方法在不同的時間里對其進行監測。本文認為可以使用理解度這一理論對其進行監測，觀察粵式普通話的理解度是否會逐年提升還是會常年穩定在一個固定的範圍內不再變化。祝新華、陳瑞端、溫紅博（2012）對十年間（1999-2008）香港大學生普通話水平的發展和變化進行總結，發現十年間大學生的普通話水平呈進步趨勢，並且大學生普通話得分，說話得分表現最好，聆聽次之，朗讀和拼音句中，最差的為普通話基礎知識。在這十年內2004年到2006年所有科目的分數都是最高的，隨後分數都有一定程度的下降，可見其動態性。

（3）粵方言作為一種較為強勢方言，由於上個世紀九十年代香港經濟的發展，文化娛樂市場繁榮，造成粵方言北上的現象。後隨著大陸地區經濟發展，國際影響力的提升，普通話推廣的工作不斷推進，普通話在粵方言區影響力越來越大。越來越多的人開始使用普通話作為工作學習的語言，粵式普通話不僅是在國內的粵方言區是極其普遍的現象，在海外的粵方言區也是極其常見的。大量語言現象的產生需要有相應的語言學研究對其進行歸納總結並且進行深入的探究。正是由於粵方言和粵式普通話有著較為特殊的歷史和地位，也時常會引起廣泛的關注和討論，甚至是調侃。有趣的是對粵式普通話的調侃來自兩方面的調侃，一方面是來自非粵方言區的人，因為粵式普通話較為難懂，故而有「天不怕，地不怕，就怕廣東人講普通話」的戲說；另一方面來自粵方言區內部人的調侃，在網絡上經常將港澳地區的領導人或者明星在正式場合的普通話節選下來配以字幕，用來表示其普通話不標準。例如：「斯靚依DO打詩齋條魚鮮事，弟生屎煙齋生茲澳廉撇KAI淵動會乾乾齋巴絲打你約爺來佬樣瞞姐

速。我哥運動元在比賽搶上玩搶拼波粉蓉真腥，公槌得喇呃斯淥面打煎排，熟把面打現排，呃絲撈面打桶排打齋子。……」這是2016年8月31日里約奧運會國家精英代表團訪問澳門時，澳門特別行政區行政長官崔世安在歡迎晚宴上的致辭。其實其真正內容是：「四年一度的世界體育盛事，第三十一屆夏季奧林匹克運動會剛剛在巴西的里約熱內盧圓滿結束。我國運動員在比賽場上頑強拼搏奮勇爭先，共取得了二十六面的金牌，十八面的銀牌，二十六面的銅牌的佳績。……」這種字幕實際上是利用粵方言的諧音來對所謂不標準的普通話進行注音。因為這種粵式普通話在粵方言區聽起來更容易和粵方言產生聯繫，更容易和自己的方言聲音掛鉤產生笑點。上文中提到的張家輝代言的遊戲廣告的轉寫則是明顯地由普通話讀音的漢字轉寫而來。可見雙方都存在著各種形式地對粵式普通話的關注。這種關注的從側面證明瞭這個問題存在著值得研究的可能性和其中包含的意義。

(4) 前文中對普通話和粵方言分別做了理解度的測試，可以利用前文所做出的測試結果與本次測試的結果進行比較。觀察普通話，粵方言和粵式普通話的理解度，證明相關假設。

4.3 實驗目的

本文對粵式普通話進行理解度測試，計算出粵式普通話的理解度；分析比較不同背景的被試對粵式普通話的理解度差異；並最終與粵方言和普通話的理解度相比較，進而得到粵式普通話在理解度上的格局分布，從而試圖論證粵式普通話的性質歸類等問題。本章所需要解決以下幾個問題：

(1) 粵方言、普通話和粵式普通話的理解度各是怎樣？

(2) 普通話的理解度呈現出何種與粵方言有何性質上的差異？

(3) 粵式普通話與普通話和粵方言的理解度對比，粵式普通話與兩者的關係是怎樣的？

4.4 粵方言與普通話理解度實驗

4.4.1 實驗材料

為了能夠滿足實驗的目的，此次實驗選擇了 24 個詞和 24 個句子作為實驗材料。實驗材料的產生方式是通過第六版《現代漢語詞典》簡單隨機抽樣得到的 48 個詞，然後將 24 個詞造句。唯一的篩選條件是這個詞不可以有同音異型詞，換句話說就是不允許聽聲音產生歧義。每個句子的長度平均值 7.75 字，符合人腦 ± 7 個單位的記憶區間。其中 12 個詞和 12 個句子是使用普通話朗讀，其餘是使用粵方言朗讀的。普通話朗讀者是普通話母語者，男性，普通話等級測試一級乙等。粵方言朗讀者是粵方言母語者，女性，從未長時間離開過廣東地區。音頻採集使用 Praat 軟件，單聲道，44100Hz，48 個音頻總時長 69109ms，平均時長 1439.77ms，極大值 2728ms，極小值 578ms，標準差 556.98。在視覺呈現方面，一半的材料是用圖像呈現，另一半的材料是用文字呈現。圖像資料來自百度圖片關鍵詞搜索，選擇較為靠前且能夠很好傳達意思的圖片。同時在每一類中各安排一個正向答案，一個反向答案。

表格 12 粵方言與普通話實驗材料類型

	詞彙				句子			
	文字		圖像		文字		圖像	
	正向	反向	正向	反向	正向	反向	正向	反向
普通話	3	3	3	3	3	3	3	3
粵方言	3	3	3	3	3	3	3	3

之所以是要選擇普通話和粵方言作為閱讀測試材料的語言是因為以下幾個考量因素：

(1) 普通話作為漢語的共同語，是現今我國最通行的語言，甚至越來越多的人將普通話作為第一語言，可以將普通話認定為是一種被試熟悉且掌握的語言。粵方言作為漢語使用人口最多的方言，其在海內外的影響力也是非常高，圍繞著粵方言也形成了一套成體系的文化系統，音樂、電視劇、網絡流行文化都會見到粵方言的身影。對於絕大多數的非粵方言區的人來說，沒有系統學習接觸過粵方言，初次聽懂還是有相當大的困難。但是粵方言作為漢語的方言之一，又與各大方言以及共同語言之間有著千絲萬縷的聯繫。所以可以將粵方言設定為一種被試不熟悉的語言變體。

(2) 因為對普通話和粵方言的音系差異已經非常清楚，然而對其在心理感知上的差異還並不十分清楚。並且心裡感知上的差異落實到實際的音系上的差異是多大，這種聯繫還沒有建立。所以可以通過對普通話和粵方言之間的數據來建立心理上和音系上的聯繫。

(3) 可以為第三章的實驗做數據基礎，通過此次實驗可以調查到普通話的正確率和反應時間，同時也可以調查到單純粵方言的正確率和反應時間，這樣可以與帶有粵方言特色的普通話的實際理解度進行對比。對於普通話和粵方言的數據就好比是一個區間的兩個端點，帶有粵方言特色的普通話的理解度就會在這兩個端點中間游移，而具體是在哪一個位置，需要通過實驗來證實，從而

可以判斷帶有粵方言特色的普通話究竟是距離哪個端點比較近，這樣可以有助於從實驗數據的角度對其進行定位和界定。

4.4.2 實驗對象

此項的被試共有 278 人，其中中國人 259 名，外國留學生 19 名參與實驗，參與者並非無償進行實驗，實驗者向所有實驗參與者支付相關報酬或等值的禮品以示回饋。中國被試者大多數是在校學生，本實驗先後在澳門大學、天津外國語大學、天津師範大學、北京大學等四所大學以及天津中學高中部進行隨機抽樣，各個年級均有樣本分布，為保證實驗結果的顯著性，在被試的選擇中優先選擇非粵方言區的被試，同時照顧到數據的平衡性，並不是絕對杜絕粵方言區的被試。以下是從性別、籍貫、方言等角度分別對中國和外籍留學生實驗對象背景信息的描寫和描述性統計結果。

4.4.2.1 性別

性別分布女性共有 154 人，佔全部人數的 59.7%；男性共有 104 人，佔全部人數的 40.3%。其中澳門大學參與者有 36 名女性，21 名男性；天津外國語大學、天津師範大學和北京大學參與者共有 107 名女性，60 名男性；天津中學高中部參與者，共有 11 名女生，23 名男生。雖然男女生比例並沒有達到 1:1，但也基本符合大學生源比例。每一組的人數都超過了 100 人可以單獨成組，且有一定的說服力。

4.4.2.2 年齡

由於集中在大學進行調查，年齡主要集中在青年層，主要是考慮到不同年齡層的測試者反應時長會有較明顯的差異，鑒於時間成本，物質成本實驗條件

有限並不能找到全部年齡層合適的被試，因此就著重挑選一個年齡層的被試。年齡的最小值是 15 歲，年齡最大是 52 歲，平均年齡在 21.41 歲，標準差 4.073，方差 16.592，總體水平趨於集中。本測試中所有年齡均為週歲。

男性的平均年齡是 21.44 歲，女性的平均年齡是 21.39 歲，二者之間十分接近，並無組間差距（ $T=0.102, df=256, sig=0.919>0.05$ ）。

4.4.2.3 籍貫

實驗主要是在天津的高校進行，所以近五分之一的被試來自天津，其餘被試均來自大陸其他省份，未有遇到來自青海、上海兩地的被試。港澳也有同學參與本實驗，不過他們祖籍並非港澳，因此就記成相應的祖籍。從表格來看，前 8 位的省份都是北方省份，前八位就佔據了 56.4%可見大多數的被試是來自於中國北方。

表格 13 粵方言與普通話測試被試籍貫統計

省份	频率	百分比	省份	频率	百分比
天津	53	20.5	江蘇	7	2.7
甘肅	18	7.0	云南	6	2.3
山東	16	6.2	贵州	6	2.3
河北	15	5.8	内蒙	5	1.9
河南	14	5.4	广西	5	1.9
陝西	11	4.3	浙江	5	1.9
北京	10	3.9	福建	5	1.9
山西	9	3.5	安徽	3	1.2
廣東	9	3.5	遼寧	3	1.2
新疆	9	3.5	西藏	2	0.8
湖北	9	3.5	重慶	2	0.8
湖南	9	3.5	台灣	1	0.4
四川	8	3.1	寧夏	1	0.4
江西	8	3.1	海南	1	0.4
吉林	7	2.7	黑龍江	1	0.4
			總計	259	100.0

4.4.2.4 第一語言

測試還調查了所有被試的第一語言，向他們詢問了他們最早學會的語言。得到 44% 的被試是以官話為第一語言，其次是以普通話為第一語言的被試，佔到全部的 40.9%。這其中不是僅僅包括北方人，甚至包括很多南方人。除此之外，各大方言區都有相對應的被試。外族語中包含了四名少數民族的被試，他們是四名少數民族的學生，有兩名來自西藏，第一語言是藏語；一名來自內蒙古，第一語言是蒙古語；一名來自吉林，第一語言是朝鮮語。

表格 14 粵方言與普通話測試被試第一語言統計

	頻數	百分比
官話	113	43.8
普通話	106	41.1
粵方言	9	3.5
湘方言	8	3.1
吳方言	6	2.3
外族語	4	1.55
客家話	4	1.55
贛方言	4	1.55
閩方言	4	1.55
總計	259	100.0

下表是在官話中再細分不同的官話區，這是根據被試者填寫的方言進行歸納整理得出的，也包含了所有的官話區的分類。因為本文並不是著重討論方言問題，因此沒有對其進行細緻的劃分，只是按照大致的地域來進行劃分，當然北方官話中還可以再細分出其他子方言片，其他官話區亦然。

表格 15 粵方言與普通話測試被試第一語言方言區統計

	頻數	百分比
北方官話	61	53.50
西南官話	24	21.05
中原官話	16	14.04

西北官話	9	7.90
江淮官話	4	3.51
總計	114	100.0

4.4.2.5 教育背景

此次實驗覆蓋了從高中一年級到博士五年級一共 16 個年級的學生，其中還有一些本科畢業或碩士畢業在學校中復習準備升學考試的同學。高中生共有 34 人，佔總體的 13.1%，本科生及本科畢業生共有 155 人，佔總體的 59.8%；碩士研究生及碩士畢業生共有 55 人，佔總體的 21.3%；博士生共有 15 人，佔總體的 5.8%。數量最多的年級是大三，共有 65 人，佔總體的 25.1%。

表格 16 粵方言與普通話測試被試教育背景統計

年級	頻率	百分比	年級	頻率	百分比
高一	6	2.3	研一	27	10.5
高三	15	5.8	研三	2	0.8
高二	13	5.0	研二	23	8.9
大一	46	17.8	研究生	3	1.2
大三	64	24.8	博一	3	1.2
大二	23	8.9	博三	5	1.9
大五	1	0.4	博二	5	1.9
大四	19	7.4	博五	1	0.4
本科	1	0.4	博四	1	0.4
			總計	258	100.0

4.4.2.6 普通話、粵方言能力自評

對被試者進行背景調查的過程中，增加了兩道題目，讓被試者對自己的普通話語言能力和粵方言的語言能力進行一個自評。其目的是一方面是調查被試自我的語言能力，為後面數據處理提供一些依據；另一方面也是對被試者的語言自信進行調查。題目要求被試者根據自己的感覺給自己的普通話或粵方言語言能力打分，從一分到五分，五分最高，一分最差。雖然在做實驗的過程中交流

全部使用普通話，過程毫無障礙，但是結果發現被試普通話自評成績並沒有很高，平均值僅有 3.89 分，其中有 172 名被試給自己打了 4 分，佔全部人數的 66.4%；有 53 名被試給自己打了 3 分，佔全部人數的 20.5%；有 32 名被試給自己打了 5 分，佔全部人數的 12.4%；有 2 名被試給自己打了 2 分，佔全部人數的 0.8%。結合 3.6.2.4 節調查的第一語言信息，以普通話為母語的被試中居然還有 20 名被試給自己打 3 分，70 名被試給自己打了 4 分，僅僅只有 16 名被試給自己打了 5 分；以官話為母語者的被試中，有 27 名被試給自己打了 3 分，74 名被試給自己打了 4 分，僅有 12 名被試給自己打了 5 分。可見大多數人對普通話並沒有足夠的信心，即便是普通話作為自己的母語，依舊不能有足夠的自信給自己的母語語言能力滿分。

粵方言方面，平均值是 1.32 分，大多數被試是完全不會粵方言的，共有 214 名被試給自己打 1 分，佔全部人數的 82.6%，絕大多數被試表示僅僅聽過粵方言歌，聽過簡單的一兩個詞語，其他並不瞭解；共有 25 名被試給自己打 2 分，佔全部人數的 9.7%，他們表示自己經常看粵方言電視劇，聽粵方言歌等對粵方言有些許瞭解；有 10 名被試自己打 3 分，佔全部人數的 3.9%；有 10 名被試給自己打 5 分，佔全部人數的 3.9%。打五分的這些被試家鄉均來自廣東地區，第一語言均是粵方言。

表格 17 語言能力自評

	總數	最小值	最大值	平均值	標準偏差
普通話能力自評	259	2	5	3.90	0.594
粵方言能力自評	259	1	5	1.33	0.876

另有 19 名外國留學生作為本文的對照試驗組，他們其中有 18 名是天津師範大學本科生，1 名是北京大學比較文學與比較文化研究所三年級的碩士生。他們均是漢語高級學習者，漢語水平考試 (HSK) 等級 12 人為五級，7 人為六級，實驗過程中使用普通話交流完全沒有問題；學習漢語的時間平均有 4.89 年，最長的是已經學習 8 年漢語，最短的也是已經學習了 3 年時間；來中國學習漢語的時間平均有 3.18 年，最長的是來中國 6 年，最短是來中國僅有 1 年。年齡最大是 24 歲，最小是 20 歲，平均在 21.95 歲，與中國學生被試組的年齡 21.41 歲基本相同。他們其中有 6 名男性，13 名女性，15 名被試來自泰國，其餘 4 名分別來自馬達加斯加、贊比亞、薩摩那和美國，來自泰國的 15 名被試第一語言均是泰語，其中有 3 名被試是有華人血統，略懂一些福州話和潮州話，來自贊比亞和美國的被試第一語言是英語，來自馬達加斯加和贊比亞的被試第一語言均為當地語言。他們給自己的普通話打分最高打 5 分，最低打 2 分，平均分在 3.21 分。

本文設置外國留學生對照組主要目的是希望能夠擴大樣本的多樣性，不僅僅局限於特定的人群，通過對外國留學生的理解度測試，也可以瞭解到漢語高級學習者與母語者之間的差異；同時也希望能夠通過對外國留學生的測試探索理解度實驗在二語習得方面的應用。

4.4.3 實驗流程

实验全程大约需要3-5分钟，中间不设置休息时间。实验结束后，会对被试进行一个简单的访谈，首先了解实验状况，是否出现一些非正常的情况，以保证数据真实有效，不是随意做出的反应。然后会对被试的背景做一个简要调查，用李克特等级五级量表格式让被试对自己的普通話和廣東話进行打分。之后询

問被試的第一語言，是否考過普通話等級、籍貫、年齡、年級等問題。所有問題不涉及個人隱私，背景調查也不記錄被試姓名，並口头得到被試的授權允許。調查結束後，會贈送給被試相應的禮品，並請其簽收。具體實驗流程同第三章的程序相同，可參看圖（2）

4.4.4 實驗地點

本實驗先後在四所大學和一所高中進行實驗，這五所學校分別位於澳門、天津、北京三座城市。為保證實驗的順利進行，防止噪音對實驗的干擾。實驗地點都選自安靜的教室、圖書館或語言實驗室。具體地點見下表：

表格 18 粵方言與普通話測試地點

學校	實驗地點
澳門大學	澳門大學人文學院語言實驗室
天津師範大學	天津師範大學津沽學院教室A101 天津師範大學金橋賓館教室808
天津外國語大學	天津外國語大學逸夫樓語音實驗室E601
北京大學	北京大學圖書館自習區
天津中學	天津中學教室A360

4.4.5 實驗結果

全部實驗過程，主要收集被試的反應（Actual Response）和反應時間（Response Time，簡寫成RT）兩項數據，根據正確反應值（Correct Response）與被試的反應相比較得到正確值（Accuracy，簡寫成ACC）。若被試答對，那麼正確值為1，如果答錯，則正確值為0。反應時間的單位是毫秒（ms）；音頻結束後立即顯示文字或圖片，因此反應時間就從音頻結束後開始計算。本文數據分析所用到的基礎數據就是正確值和反應時間。本實驗共有278名被試，每名被試均有48道題目，每道題目記錄兩項數據，有一名被試的實驗結果記錄出現異常，屬於無效結果，因此本實驗共有 $277 \times 48 \times 2 = 26592$ 個有效數據。本實驗

所有的數據分析均使用IBM SPSS Statistics 22版，32位版。

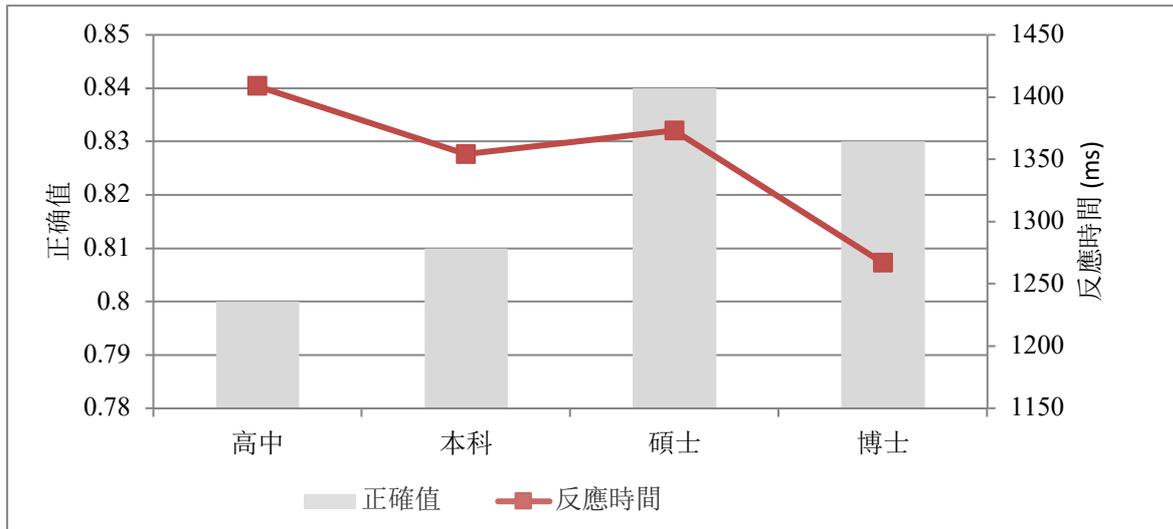
4.4.5.1 總體情況

4.4.5.1.1 性別因素

在正確值方面男性的平均正確值是0.817，標準偏差是0.387；女性的平均正確值是0.804，標準偏差是0.397，男性略高於女性。而在反應時間方面，男性的平均反應時間是1458.242毫秒，女性是1349.679毫秒，女性的反應時間快於男性的反應時間。利用獨立樣本T檢驗的方法對男女兩組進行分析，得到在正確值方面，男女性別之間沒有顯著差異（ $T=1.886, df=13294, p=0.59>0.05$ ），而在反應時間方面，男女性別之間有顯著差異（ $T=7.950, df=13294, p=0.00<0.05$ ）。也就是說在正確值方面，男女性別略有不同，但沒有顯著差異，在反應時間上，女性反應時間顯著快於男性。

4.4.5.1.2 教育因素

教育因素按照被試的教育背景分為四組，分別是高中組，本科組，碩士研究生組和博士研究生組。利用單因素方差分析（ANOVA）分別分析這四組在正確值和反應時間中的差別，這其中不包括留學生樣本。正確值方面，高中組的正確值平均是0.798，標準偏差0.401；本科組平均0.813，標準偏差0.390；碩士研究生組平均是0.839，標準偏差0.367；博士研究生平均是0.830，標準偏差0.375。反應時間方面，高中組的平均反應時長是1408.697毫秒，標準偏差789.12，本科組平均是1362.902毫秒，標準偏差810.63；碩士研究生平均是1373.073，標準偏差751.52，博士研究生平均是1268.831，標準偏差是736.30。



圖表 6 教育背景因素下的正確值與反應時間

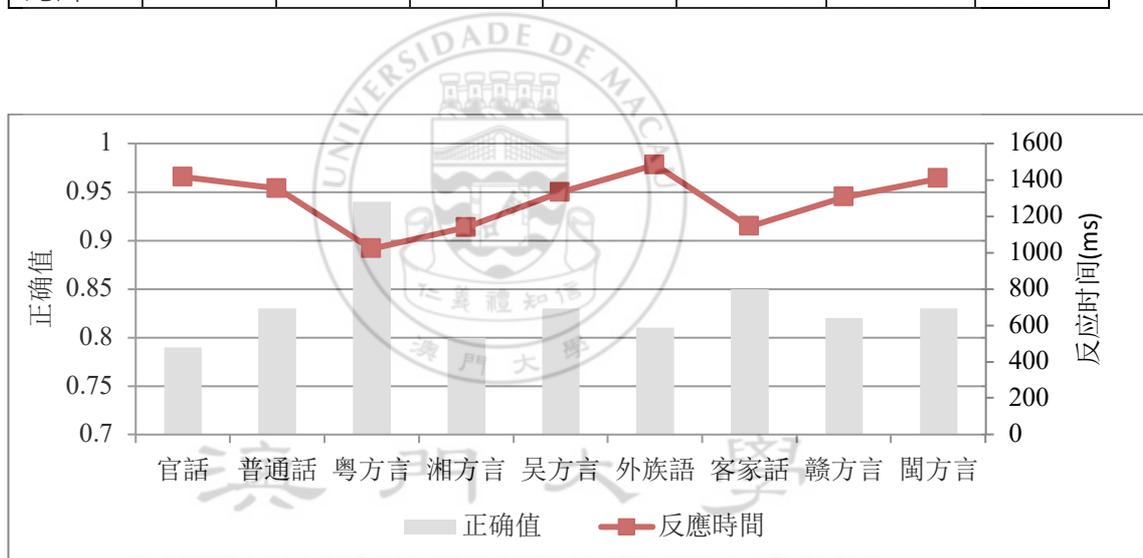
總體上看，碩士研究生的正確率最高，高中組的用時時間最長。根據單因素方差分析，四組之間在正確值方面有顯著差異 [$F(3,12380) = 4.807$, $\text{sig} = 0.002 < 0.05$]；在反應時間方面也具有顯著性差異 [$F(3,12380) = 5.349$, $\text{sig} = 0.001 < 0.05$]。經 Tukey HSD 事後檢驗組間顯著性發現：(1) 正確值方面碩士研究生組與高中組和本科組有顯著性差別 ($\text{sig} = 0.004, 0.014 < 0.05$)，其餘組間並無顯著性差異；(2) 反應時長方面博士研究生組與其他所有組之間均有顯著性差別 ($\text{sig} = 0.000, 0.012, 0.009 < 0.05$)，其餘組間並無顯著性差異。

4.4.5.1.3 第一語言背景

測試按照被試的第一語言背景將被試分為 9 大類，其中是 7 大方言外加普通話和外族語。從數據上看粵方言毫無疑問在正確值中獲得最高的正確率，並且使用了最短的反應時間。其次是客家話，正確率排名第二，反應時間也是略高於粵方言一些；官話區的被試是正確率最低，且用時時間最長的。

表格 19 被試第一語言因素下的正確值和反應時間

	N	正確值			反應時間		
		平均值	標準偏差	標準錯誤	平均值	標準偏差	標準錯誤
官話	5424	0.798	0.402	0.005	1428.563	814.304	11.057
普通話	5088	0.829	0.377	0.005	1351.724	785.416	11.011
粵方言	432	0.937	0.242	0.012	1020.359	643.548	30.963
湘方言	384	0.802	0.399	0.020	1137.625	594.805	30.354
吳方言	288	0.826	0.379	0.022	1331.486	745.969	43.957
外族語	192	0.705	0.456	0.014	1691.253	775.996	23.355
客家話	192	0.849	0.359	0.026	1143.052	789.163	56.953
贛方言	192	0.822	0.383	0.028	1304.922	644.221	46.493
閩方言	192	0.828	0.378	0.027	1406.953	683.061	49.296
總計	12384	0.809	0.393	0.003	1390.982	793.570	6.882



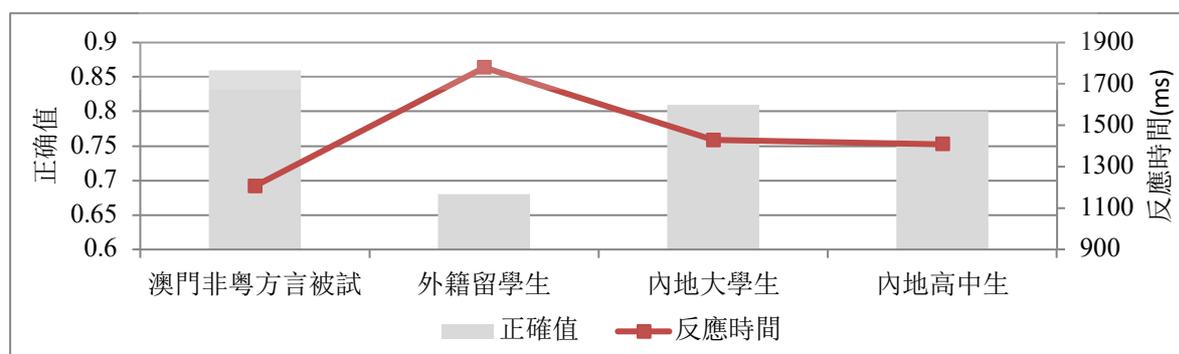
圖表 7 被試第一語言因素下的正確值和反應時間

利用單因素方差分析組間差異得到結果是正確值方面組間具有顯著差異 [F(8, 12423)=9.261, sig=0.000<0.05]，反應時間方面也呈現出顯著性差異 [F(8, 12423)=20.082, sig=0.000<0.05]。經 Tukey HSD 事後檢驗組間顯著性發現在正確值方面，粵方言除了與客家話無顯著性差異之外，與其他所有語言背景的被試均有顯著性差異，除此之外官話背景的被試和普通話背景的被試在正確值上也顯著性差異。在反應時間方面，各個方言呈現出較為混亂的情況，每個方言背景都與個別幾個方言背景之間存在顯著性差異，不過粵方言與其他方言之間的

差異依舊是最多的，除去湘方言和客家話。粵方言組在正確值和反應時間兩個方面都與其他組有顯著性差異這一點並不意外，畢竟題目是由普通話和粵方言組成，對粵方言不瞭解的人必然會答錯和花費更長的時間。因此在後文中討論相關問題的過程中，應將粵方言背景的被試單獨成組，與其他組別分開。

澳門大學是處在粵方言區，因此澳門大學的學生多多少少會接觸到粵方言，因此應考察澳門大學的學生是否有別於大陸就讀的外籍留學生、成年高等學校學生和高中生。若澳門非粵方言被試與其他組差異顯著，則應該單獨考察分析。為了確保樣本的一致性，已排除粵方言區母語者被試。

經計算，在正確值方面，澳門非粵方言被試平均值0.86，標準偏差0.351；外籍留學生平均值0.68，標準偏差0.466；內地成年高等學校學生平均值是0.81，標準偏差是0.396；高中生的平均值是0.80，標準偏差是0.401。在反應時長反面，澳門非粵方言被試平均值1206.23，標準偏差720.618；外籍留學生平均值1779.01，標準偏差818.55；內地成年高等學校學生平均值是1428.98，標準偏差是811.851；高中生的平均值是1408.70，標準偏差是789.124。澳門非粵方言被試在正確值方面得分最高，並且在反應時長方面用時也是最少的。



圖表 8 被試不同組別的正确值和反應時間

經單因素方差分析在正確值方面存在著顯著的組間差異 [F(3,12381)=29.385, sig=0.000<0.05]，在反應時間方面同樣存在著顯著的組間差異 [F(3, 12381)=129.003, sig=0.000<0.05]。經Tukey HSD事後檢驗組間顯著性發現在正確值方面，不出意外地發現，澳門大學與其他所有組別均有顯著性差異，外籍留學生組也與其他所有組別具有顯著性差異，內地高校的成人被試和內地高中生被試之間無顯著性差別；在反應時長方面，有著相同的結果，澳門大學與其他所有組別均有顯著性差異，外籍留學生組也與其他所有組別具有顯著性差異，內地高校的成人被試和內地高中生被試之間無顯著性差別。根據上面的數據，可以將澳門非粵方言被試和外籍留學生分別單獨置於兩組進行分析和研究，內地高校成人和內地高中生可以合併成為一組，可以與澳門非粵方言被試互相比較從而得到有關結論。

4.4.5.2時間參數計算

4.4.5.2.1變量選擇

目前為止，經過試驗數據的初步分析，形成四個數據組，分別是：粵方言母語者組、澳門非粵方言被試組、內地非粵方言被試組以及留學生組。每個組的人數及樣本數量分別是：粵方言母語者11人（528條）、澳門非粵方言被試組49人（2352條）、內地非粵方言被試組198人（9504條）、留學生組19人（912條）。

表格 20 反應時長影響因素及影響量

	III 類平方和	自由 度	均方	F	顯著 性	偏 Eta 平 方
組別	314260209.49	3	104753403.16	176.791	0.000	0.039
詞彙/句子	44878731.883	1	44878731.883	75.741	0.000	0.006

粵方言/普通話	91658606.598	1	91658606.598	154.691	0.000	0.012
圖像/文字	3563073.359	1	3563073.359	6.013	0.014	0.000
正向/反向	6555141.740	1	6555141.740	11.063	0.001	0.001

因此，在計算時間參數時，優先考慮組別因素，粵方言-普通話因素和詞-句子因素，將此三個水平上的因素納入到時間參數的計算中，對變量進行區分。

(1) 組別因素：依照前文中對組別劃分的分析，將所有被試分成四組，在時間參數的計算中，並不會依照組別因素分別處理不同組別的反應時長。因為正式利用時間參數來將所有的被試納入到同一尺度上。同樣是作對一道題，不同被試的反應時間不同恰好也能夠說明他們各自之間的差異。不過在時間參數計算後，依舊會按照組別劃分對其進行分類統計，從而能夠看出組間的異同。

(2) 粵方言-普通話因素：實驗材料是由普通話和粵方言兩種不同的語音進行朗讀的，所以需要進一步探尋四組在不同語音情況下的正確值和反應時長的差異。在普通話方面，澳門非粵方言被試平均正確率和粵方言母語者正確率均約等於0.93，且反應時長也大致相似分別是1015毫秒和1054毫秒；內地非粵方言被試也有0.91，反應時長稍長一些有1213毫秒，而外籍留學生雖然能夠達到普通話高級水平，但是仍然與母語者相距甚遠只有0.783，反應時長長達1678毫秒。除了外籍留學生組之外，澳門非粵方言被試，內地非粵方言被試組和粵方言母語者在普通話方面沒有顯示出非常大的差異。經過單因素方差分析，組間正確值方面呈現出顯著性差異 $[F(3, 6644)=31.313, sig=0.000<0.05]$ ，反應時長方向也呈現出顯著性差異 $[F(3, 6644)=98.514, sig=0.000<0.05]$ 。利用

Tukey HSD事後檢驗，在正確值方面，外籍留學生與其他所有組均有顯著性差異，而其他三組之間並無顯著差異，而在反應時長方面，澳門非粵方言被試和粵方言母語者組之間沒有顯著性差異，其他組間均有顯著性差異。總體上看，在普通話語音材料中，母語者顯示出相對一致的模式，都是在較短的時間內並且較好地完成了任務，外籍留學生則在正確值和反應時長兩個方面都表現出與母語者的差距。

使用粵方言閱讀的實驗材料結果就顯示出了明顯的不同。從正確值上來看，粵方言母語者遙遙領先，正確值依舊保持在高位，0.96比普通話的正確值還要高，澳門大學的學生降到了與內地非粵方言被試差不多的位置，降到0.78，但是依舊好於內地非粵方言被試的0.69，外籍留學生0.581墊底。從反應時間上看，也可以看出明顯的差距，粵方言母語者平均是1083毫秒，與普通話幾乎一樣，澳門非粵方言被試的反應時長稍微長一些到了1399毫秒，內地非粵方言被試由於對粵方言實際日常接觸的機會並不多，因此呈現出和外籍留學生差不多的反應時間1733毫秒，稍微快於外籍留學生的1880毫秒。

經過單因素方差分析，組間正確值方面具有顯著差異 $[F(3, 6644)=53.537, sig=0.000<0.05]$ ，反應時長方向也具有顯著差異 $[F(3, 6644)=86.797, sig=0.000<0.05]$ 。利用Tukey HSD事後檢驗，在正確值和反應時長方面，所有組間均存在這顯著差異，由此可見，粵方言材料是區分不同組性質的決定性因素。

表格 21 不同組別普通話和粵方言的正確值與反應時長

	正確值		反應時長	
	普通話	粵方言	普通話	粵方言
粵方言母語者	0.93	0.96	1054	1083

澳門非粵方言被試	0.93	0.78	1015	1399
內地非粵方言被試	0.91	0.69	1213	1733
外籍留學生	0.78	0.58	1678	1880

在時間參數的計算中，也不會對粵方言和普通話進行單獨考量，恰好被試對粵方言和普通話之間所產生的反應時長的差異是需要被記錄下來的，正因為對於絕大多數被試來說，粵方言是一種不太熟悉、甚至完全不會的方言，粵方言的語音系統在大腦中很難非常快速地與普通話的語音相聯繫，或者很難直接地與大腦中的意義取的聯繫，所以要理解該語音需要耗費比聽熟悉的語音系統更長的時間。

(3) 詞-句子因素：本次實驗中的語音材料有句子和詞的對立，所需要的語音材料本身就有長短的差別，單詞的平均錄音時長是961.79毫秒，句子的錄音時長是1917.75毫秒，平均句子比單詞多一秒。因為句子的信息量一定是大於單詞的，因此對句子的反應時間也會相應長於單詞，在計算時間參數的過程中，有必要將兩種不同類型的材料區分開。總體上看，單詞的平均反應時長是1342.42毫秒，句子的平均反應時長是1519.27毫秒。經獨立樣本T檢驗發現二者之間有顯著的差異 $[T=-12.11, df=13294, sig=0.00<0.05]$ 。句子和詞從正確值上看，句子會稍微比詞的準確值高，單詞的平均準確值是0.78，標準偏差0.416，句子的平均準確值是0.84，標準偏差是0.366。經獨立樣本T檢驗發現二者之間有明顯差異 $[T=-9.268, df=13294, sig=0.00<0.05]$ 。

表格 22 句子與詞彙的反應時長與正確值

		反應時長		正確值	
		普通話	粵方言	普通話	粵方言
句子	粵方言母語者	1071.182	1211.394	0.985	0.977
	澳門非粵方言被試	1052.236	1485.321	0.964	0.828

	內地非粵方言被試	1239.904	1881.058	0.954	0.708
	外籍留學生	1900.075	2009.180	0.864	0.561
詞彙	粵方言母語者	1036.197	955.242	0.871	0.939
	澳門非粵方言被試	978.279	1313.194	0.889	0.741
	內地非粵方言被試	1186.192	1584.428	0.873	0.673
	外籍留學生	1455.118	1751.680	0.702	0.601

由於詞和句子是兩類完全不同類型的實驗材料，並且詞和句子本身實驗材料錄音的長度就存在著顯著差異，同時詞和句子的實驗材料本身所包含的信息量有差別，理解不同信息量的所需要的時間也存在差異；因此，詞和句子的反應時間應該分開計算，分別計算對單詞的反應時間參數和對句子的反應時間參數。

4.4.5.2.2 計算過程

此次試驗對所有被試的所有反應時間進行記錄，因此共有有效反應時間數據 $277 \times 48 = 13296$ 個數據。根據上文所討論的內容將反應時間按照實驗材料的類型進行分開，分別得到詞彙/句子，圖像/文字，正向/反向題目三個變量每個變量兩個水平的類型。依照前文所使用的計算方法，分別進行了歸一化處理，消除被試之間的個體差異。然後根據實驗材料的性質進行分類處理。最後根據統計得到本地人對該語言的平均反應時間，並以此為標準，讓同類型的時間數據與之相除得到最終的時間參數，用其最終參與理解度數值的運算。

4.4.6 理解度分析

根據上文所述，將原始的反應時長通過一系列的計算使其轉化為計算理解度的時間參數。利用這個時間參數，與每一題的正確值相乘，就可以得到該题目的得分，計算同一被試得分的平均分，就可以得到該名被試在整個測驗中的

得分；同理，計算同一道題目的所有被試的得分的平均分，就可以得到該題目在整個測驗中的得分狀況。根據實驗材料的不同屬性以及被試者的不同背景，就可以分析出可理解度的具體數值。得到理解度數值後，就可以根據每名被試的每一道題目的理解度得分計算此次實驗的信度。經計算克隆巴赫系數 $\alpha = 0.819$ ，說明此次測試具有良好的信度，經由測試得到的結論是可信的。

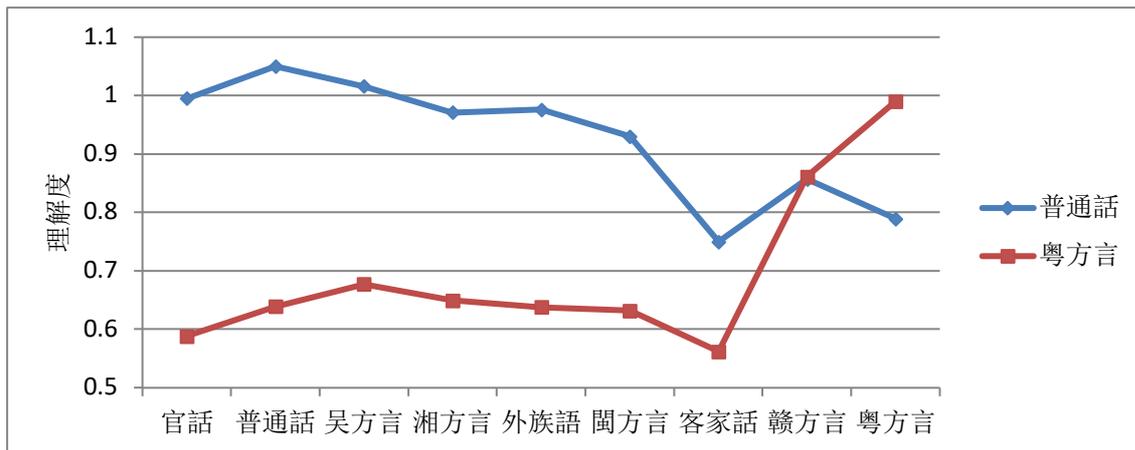
通過計算，共得到理解度數值13296個，整體的理解度平均值是0.806，標準偏差0.657。男性被試的平均理解度是0.808，標準偏差是0.645；女性的平均理解度是0.804，標準偏差是0.664，根據獨立樣本T檢驗，兩性之間並沒有顯著的差異， $[T=0.369, df=13294, sig=0.712 > 0.05]$ 。可見兩性之間在理解度上並沒有太大的差異。男性普通話的理解度均值是1.024，標準偏差0.520，女性普通話的理解度均值為0.960，標準偏差0.555。男性粵方言的理解度均值是0.594，標準偏差0.686，女性粵方言的理解度均值為0.649，標準偏差0.726。經單變量方差分析後，得到性別在普通話和粵方言這兩個水平上均不存在顯著的差異 $[F(1, \quad) = 0.147, sig = 0.701]$ 。

根據被試的不同第一語言背景進行統計，分別得到普通話的理解度和粵方言的理解度，此處統計暫不包括外國留學生組。普通話方面，第一語言是官話的被試理解度平均值是0.995，標準偏差0.540；第一語言是普通話的被試理解度平均值是1.050，標準偏差0.508；第一語言是粵方言的被試理解度平均值是0.789，標準偏差0.561；第一語言是湘方言的被試理解度平均值是0.971，標準偏差0.547；第一語言是吳方言的被試理解度平均值是1.016，標準偏差0.522；第一語言是外族語的被試理解度平均值是0.750，標準偏差0.604；第一語言是客家話的被試理解度平均值是0.857，標準偏差0.607；第一語言是贛方言的被

試理解度平均值是0.930，標準偏差0.551；第一語言是閩方言的被試理解度平均值是0.976，標準偏差0.533。

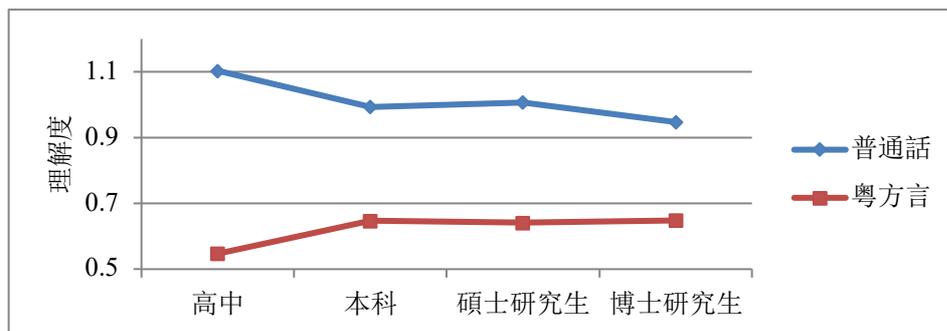
粵方言方面，第一語言是官話的被試理解度平均值是0.588，標準偏差0.695；第一語言是普通話的被試理解度平均值是0.639，標準偏差0.713；第一語言是粵方言的被試理解度平均值是0.990，標準偏差0.670；第一語言是湘方言的被試理解度平均值是0.649，標準偏差0.738；第一語言是吳方言的被試理解度平均值是0.677，標準偏差0.769；第一語言是外族語的被試理解度平均值是0.562，標準偏差0.716；第一語言是客家話的被試理解度平均值是0.861，標準偏差0.754；第一語言是贛方言的被試理解度平均值是0.632，標準偏差0.655；第一語言是閩方言的被試理解度平均值是0.638，標準偏差0.704。

在普通話方面，只有第一語言是客家話和粵方言的被試低於0.90，其餘被試均高於0.90，整體水平平均值是0.985，標準偏差0.542。整體都處於較高的理解度水平。在粵方言方面，第一語言是粵方言的被試的理解度依舊保持在相當高的取值範圍，得到0.990，並且標準偏差較小，說明這組數據是有著相當好的一致性。而其他的被試則呈現出與粵方言被試相當大的差異性，第一語言是外族語的被試僅有0.562的理解度，其餘都是處於0.6左右。整體水平平均值是0.627，標準偏差是0.710。如果將粵方言被試的理解度排除在外，其餘被試的理解度平均值則會下降到0.656。可見其他方言背景的被試呈現出一定的一致性，在對於並不瞭解的語言變體，但又不是完全沒關係的語言，理解度就會下降到0.6左右。



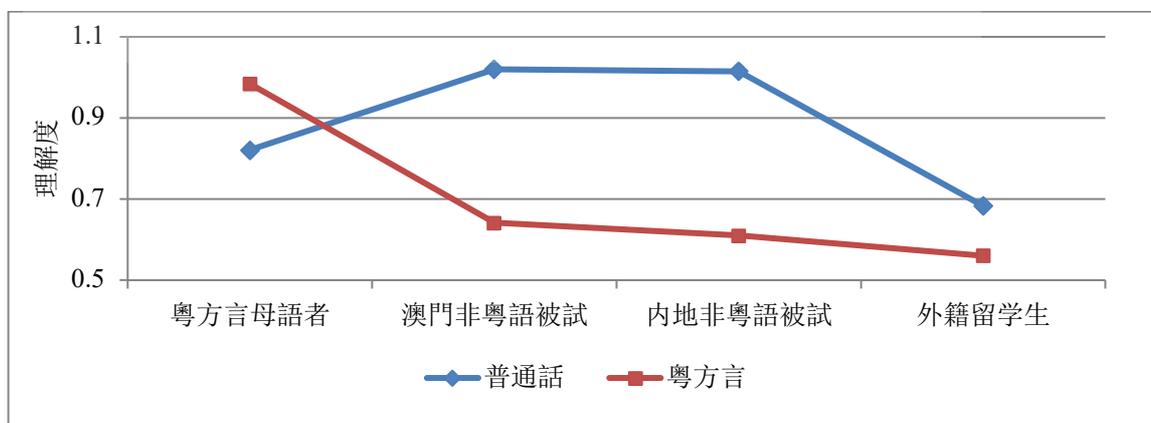
圖表 9 第一語言背景下被試對普通話與粵方言的理解度

按照被試的教育背景進行分析不同教育背景的被試對普通話和粵方言之間的理解度有無差異。在普通話方面，高中學歷背景的被試理解度平均值是 1.103，標準偏差是 0.476；本科學歷背景的被試理解度平均值是 0.993，標準偏差是 0.542；碩士研究生學歷背景的被試理解度平均值是 1.007，標準偏差是 0.522；博士研究生學歷背景的被試理解度平均值是 0.947，標準偏差是 0.543。在粵方言方面，高中學歷背景的被試理解度平均值是 0.547，標準偏差是 0.678；本科學歷背景的被試理解度平均值是 0.647，標準偏差是 0.715；碩士研究生學歷背景的被試理解度平均值是 0.641，標準偏差是 0.708；博士研究生學歷背景的被試理解度平均值是 0.648，標準偏差是 0.704。高中學歷背景的被試在普通話方面理解度是最高的，然而在粵方言的理解度方面理解度是最低的。造成這一現象的原因可能是因為高中生在高中階段主要使用和接觸的普通話，另外高中生生源相對單一，主要是北方學生，因此對於普通話的理解度較高，對於粵方言的接觸較少，所以相比較之下粵方言的理解度最低。經過單因素方差分析，年級變量對理解度沒有顯著性影響， $[F(3)=0.391, sig=0.760]$ ，因此可以將不同年級的被試劃分為同一類型進行考察，不必再分別分析。



圖表 10 教育背景因素下被試對普通話與粵方言的理解度格局

最後根據前文對被試的分組，分別是粵方言母語者組、澳門非粵方言被試、內地非粵方言被試組和外籍留學生組。在普通話方面，粵方言母語者組的理解度平均值是0.821，標準偏差是0.557；澳門非粵方言被試的理解度平均值是1.020，標準偏差是0.509；內地非粵方言被試組的理解度平均值是1.015，標準偏差是0.533；外籍留學生組的理解度平均值是0.683，標準偏差是0.600；總體平均值是0.985，標準偏差是0.542。在粵方言方面，粵方言母語者組的理解度平均值是0.984，標準偏差是0.672；澳門非粵方言被試的理解度平均值是0.641，標準偏差是0.685；內地非粵方言被試組的理解度平均值是0.610，標準偏差是0.712；外籍留學生組的理解度平均值是0.560，標準偏差是0.729；總體平均值是0.627，標準偏差是0.710。

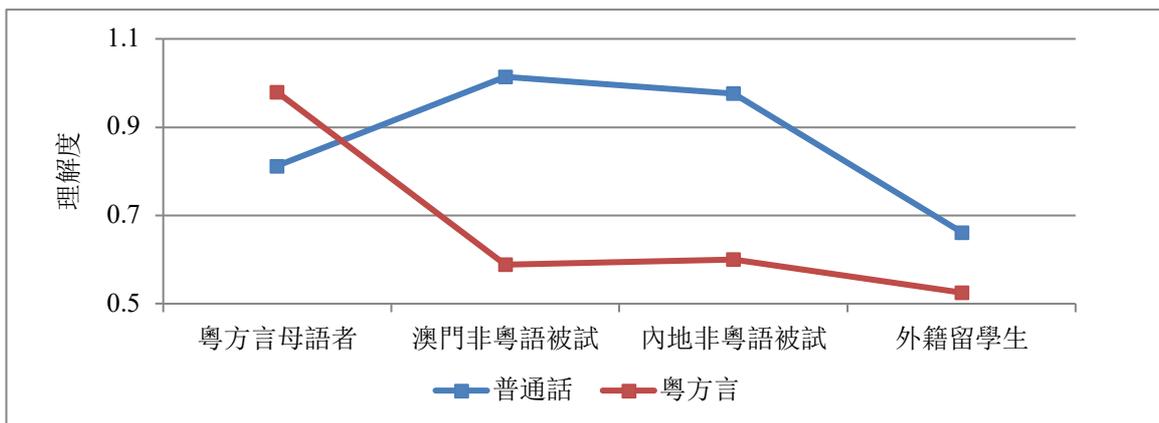


圖表 11 不同組別的普通話與粵方言理解度格局

普通話方面，根據單因素方差分析，組間具有顯著性差異。經過Tukey HSD事後檢驗，澳門非粵方言被試組和內地非粵方言被試組之間無顯著性差異，粵方言母語者組與澳門非粵方言被試組的顯著性系數為 $\text{sig}=0.00$ ，有顯著性差異；粵方言母語者組與內地非粵方言被試組的顯著性系數為 $\text{sig}=0.000$ ，具有顯著性差異；內地非粵方言被試組與澳門非粵方言被試組的顯著性系數為 $\text{sig}=0.991$ ，無顯著性差異。外籍留學生組與所有組別均有顯著性差異。粵方言方面，只有粵方言母語者保留著高度的理解度，其他幾組依次遞減。根據單因素方差分析，組間具有顯著性差異。經過Tukey HSD事後檢驗，粵方言母語者組與其他所有組均有顯著性差異，顯著性系數均為 $\text{sig}=0.00$ ，其他各組之間也都存在顯著性差異，且顯著性系數也均為 $\text{sig}=0.00$ 。

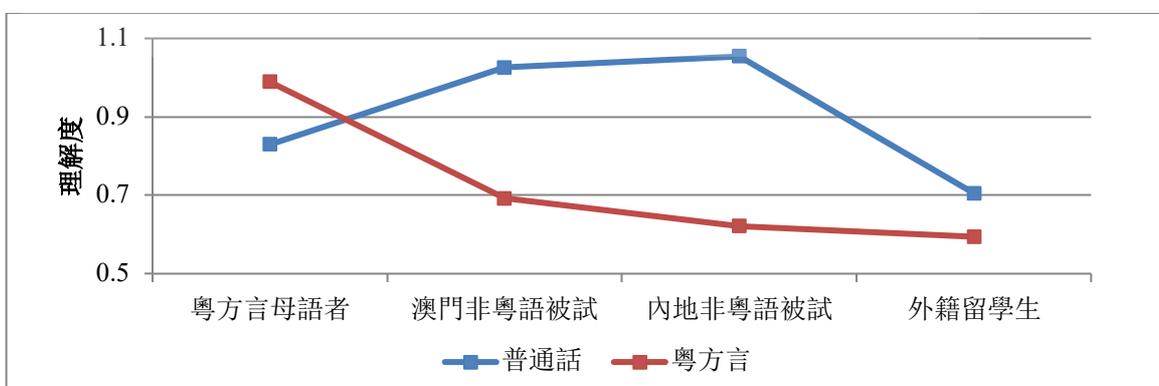
以上是從宏觀角度對理解度進行分析，上文中理解度系數是根據實驗材料信息量的多少進行分別計算，此時可以分別分析不同組別對於不同實驗材料和語言變體之間的理解度。通過分別分析數據規律，可以得到相應的理解度區間。

首先來看詞彙的理解度，普通話方面，粵方言母語者組的理解度平均值是0.812，標準偏差是0.581；澳門非粵方言被試的理解度平均值是1.014，標準偏差是0.553；內地非粵方言被試組的理解度平均值是0.976，標準偏差是0.572；外籍留學生組的理解度平均值是0.662，標準偏差是0.630；總體平均值是0.954，標準偏差是0.579。在粵方言方面，粵方言母語者組的理解度平均值是0.979，標準偏差是0.640；澳門非粵方言被試的理解度平均值是0.589，標準偏差是0.658；內地非粵方言被試組的理解度平均值是0.600，標準偏差是0.696；外籍留學生組的理解度平均值是0.525，標準偏差是0.679；總體平均值是0.608，標準偏差是0.690。



圖表 12 不同組別的普通話與粵方言詞彙理解度格局

其次是句子的理解度，普通話方面，粵方言母語者組的理解度平均值是0.830，標準偏差是0.534；澳門非粵方言被試的理解度平均值是1.026，標準偏差是0.462；內地非粵方言被試組的理解度平均值是1.054，標準偏差是0.488；外籍留學生組的理解度平均值是0.705，標準偏差是0.568；總體平均值是1.016，標準偏差是0.500。在粵方言方面，粵方言母語者組的理解度平均值是0.989，標準偏差是0.704；澳門非粵方言被試的理解度平均值是0.692，標準偏差是0.708；內地非粵方言被試組的理解度平均值是0.621，標準偏差是0.727；外籍留學生組的理解度平均值是0.594，標準偏差是0.776；總體平均值是0.647，標準偏差是0.730。



圖表 13 不同組別的普通話與粵方言句子理解度格局

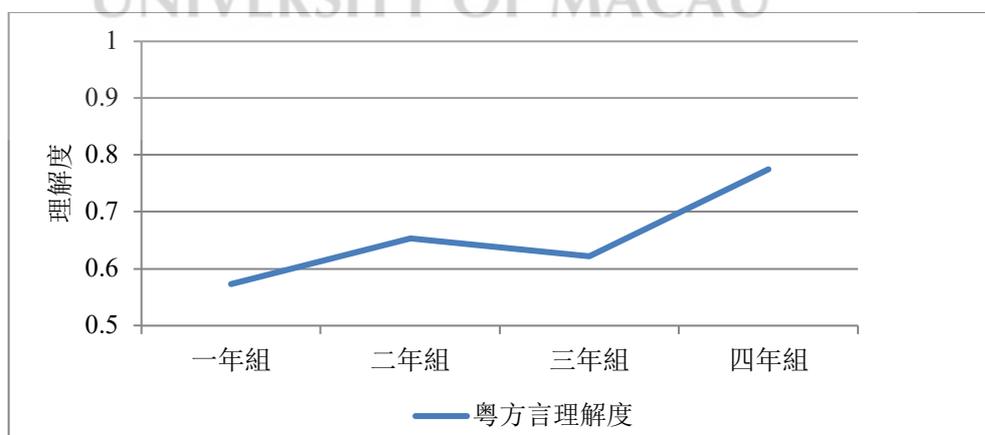
總體上看詞彙和句子呈現出相同的趨勢和走向。普通話方面，除去外國留學生組之外，其他三組無論是詞彙還是句子均保持幾乎相同的得分，這是符合實際情況的結果，可見對於民族共同語有著基本相同的語言理解度，也可以從側面印證實驗數據的有效性，外國留學生組雖然大多數都獲得了HSK高級水平，然而通過理解度測試依舊可以看到他們在普通話方面與本族講話人之間的差距。這個差距一方面是由於正確值低於本族講話人，另一方面也是由於外國留學生的反應時長相對較長，導致他們最終的得分偏低。在粵方言方面，僅有粵方言母語者組保持著高位，並且通過觀察發現，粵方言母語者在粵方言和普通話兩個語言變體的表現是幾乎相同的。也再次體現了對於一門熟悉的語言所表現出的理解度應該是一種什麼狀態。其他組別的理解度則大幅下降，不過大小次序並沒有改變，澳門非粵方言被試組處於內地非粵方言被試和粵方言母語者之間，外籍留學生排在最後。

從細節上看二者之間區別，首先粵方言母語者組和內地非粵方言被試組之間在粵方言上的差距大致上是固定的，差值在3.6-3.8之間。可見這個區間正是二者理解度上面的一個鴻溝，同時也注意到在這一區間內部游離著澳門非粵方言被試，由於澳門非粵方言被試的特殊環境導致了該組成員對粵方言的有著更多的接觸，也有著更多學習的機會。不過他們在詞彙和句子上呈現出不同的狀態：在詞彙方面，澳門非粵方言被試更偏向內地非粵方言被試組，表示他們對粵方言的詞彙理解度並不高；在句子方面，澳門非粵方言被試偏向粵方言母語者組和內地非粵方言被試組的中點，高於他們在詞彙上的理解度。因此可以將澳門非粵方言被試看成是粵方言的學習者，雖然這個學習的過程並不一定是自覺地自發的學習，更大程度上是被動的或者受環境因素影響的。於是可以通過

對澳門非粵方言被試組的内部理解度狀況進行觀察和分析，從中可以瞥見一個相對動態的學習過程在理解度上所反映出的狀態。

在對澳門非粵方言被試的背景調查中比其他組別多設計一個問題，即「你來澳門幾年了？」雖然被試中有本科生、碩士、博士，但是他們來澳門的時間大致都在1-4年內。因此可以按照來澳門的時間長短分成一年組、二年組、三年組和四年組。通過觀察這四個組的理解度變化，可以從内部瞭解到隨著學習時間的增加，理解度是如何變化的。本文此處主要針對澳門非粵方言被試組對粵方言的理解度變化，暫不討論普通話理解度相關方面的問題。

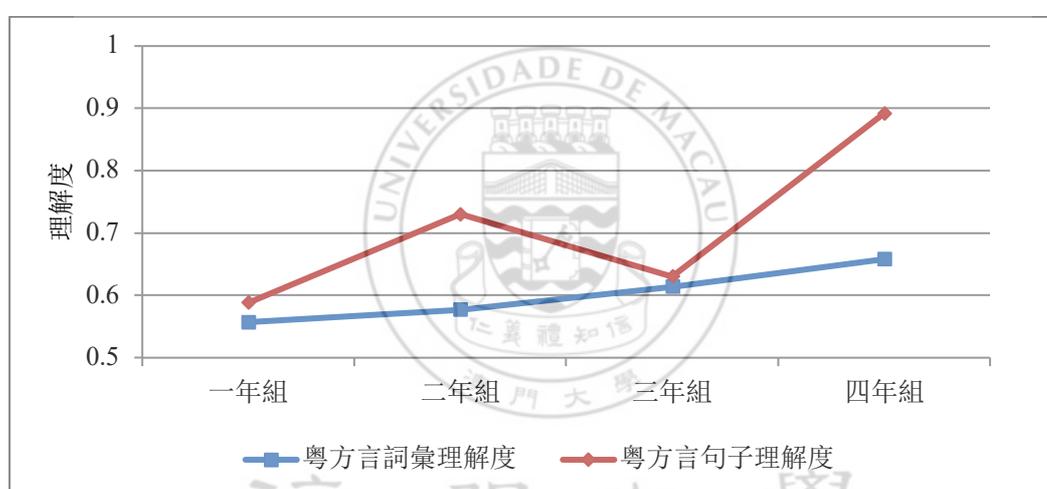
從總體來看，一年組的理解度平均值是0.573，標準偏差是0.674；二年組的理解度平均值是0.653，標準偏差是0.686；三年組的理解度平均值是0.622，標準偏差是0.649；四年組的理解度平均值是0.775，標準偏差是0.711。經過單因素方差檢驗發現組間是存在顯著性差異的。 $[F(3, 1172)=4.391, sig=0.004<0.05]$ 經過Tukey HSD事後檢驗只有一年組和四年組之間存在顯著性差異， $sig=0.002<0.05$ ，其他組別均沒有顯著性差異。



圖表 14 澳門非粵方言被試年級分組粵方言綜合理解度

在詞彙的層面上，一年組的理解度平均值是0.557，標準偏差是0.682；二年組的理解度平均值是0.577，標準偏差是0.619；三年組的理解度平均值是0.614，標準偏差是0.642；四年組的理解度平均值是0.658，標準偏差是0.680。

在句子的層面上，一年組的理解度平均值是0.588，標準偏差是0.667；二年組的理解度平均值是0.730，標準偏差是0.741；三年組的理解度平均值是0.630，標準偏差是0.662；四年組的理解度平均值是0.892，標準偏差是0.724。



圖表 15 澳門非粵方言被試年級分組粵方言句子與詞彙理解度格局

雖然直線在個點之間的斜率有所不同，但總體趨勢卻是大體相同的，是呈現出總體向上的趨勢。一年組的理解度都小於四年組的理解度，可見在組內就是呈現出隨著在澳時間的增加理解度會上升的趨勢，這個上升過程是曲折的，並不是直線上升。綜合來看，一年組是最低的，二年組理解度提高很多，三年組的理解度有所降低，不過仍比一年組要高，最後的四年組是全部組別中最高的。詞彙的理解度和句子的理解度之間比較，依舊是句子的理解度全部高於詞彙的理解度，不過，詞彙理解度的上升勢頭保持得更加穩定，在二年組和三年組並沒有出現大的波折，相反在句子的理解度方面出現了較大的波動，三年組

和四年組之間產生了一個非常大的提升。由此可見詞彙的學習偏重於累積的過程，隨著時間的推進詞彙的理解度會逐漸增加，隨著詞彙的累積到一定程度會激發出句子理解度的激增。

澳門非粵方言被試組本身是處於粵方言母語者組和內地非粵方言被試組之間的一個相對動態的一個組別。通過對澳門非粵方言被試組內部理解度的分析，可以看到在一個相對動態的組中還包含著內部的動態性。這種動態是源自於在澳門大學上學的內地非粵方言被試在一個粵方言環境下對粵方言自覺或不自覺學習和瞭解。從中可以瞥見在一門語言的學習過程中，學生對該語言的理解度的變化。其中包含了從量變到質變的過程，也包含了曲折上升的過程。也可以看出詞彙和句子在不同時間段下的增長模式有所不同。

4.4.7 相互理解度與相互理解率

上文中所計算出的理解度均是單向的理解度，本文的主題是計算語言間相互理解度。囿於技術上的限制，暫時不能夠實現實時的相互理解度的測算。暫時只能通過技術手段將相互對話的過程用電腦分割開，讓對話在不同的時空發生，換句話說就是用兩個單向的理解度拼合成一個雙向的相互理解度。本文將兩個單向理解度的數值取平均值得到相互理解度的數值。

於是首先根據此次測試的實驗結果可以通過內地非粵方言被試與粵方言母語者之間的相互可理解度得到普通話與粵方言之間的相互理解度，內地非粵方言被試講普通話，粵方言母語者將粵方言，雖然內地非粵方言被試組內部的普通話程度也並不統一，但是其對粵方言的理解度是大致相同的，內地非粵方言被試和粵方言母語者之間的相互可理解度是利用內地非粵方言被試對粵方言的

理解度和粵方言母語者對普通話的理解度。內地非粵方言被試對粵方言的理解度是0.641，粵方言母語者對普通話的理解度是0.821。所以相互的可理解度應是：0.731。可見如果內地非粵方言被試講普通話，粵方言母語者講粵方言，成功的概率是很低的，基本上是不能相互溝通的，不能溝通的原因主要來自於內地非粵方言被試對粵方言的理解度偏低。

其次分別從詞彙和句子的角度計算二者之間的相互理解度，從詞彙方面來看，內地非粵方言被試組對粵方言詞彙的理解度是0.589，粵方言母語者對普通話詞彙的理解度是0.812，因此詞彙方面，普通話與粵方言之間的相互理解度是0.7005。從句子方面來看，內地非粵方言被試組對粵方言句子的理解度是0.621，粵方言母語者對普通話句子的理解度是0.830，因此句子方面，普通話與粵方言之間的相互理解度是0.7255。

若計算二者之間的相互理解率，還需要參考第三章中所得到的數據，將以上得到的數據納入到之前的衡量體系中，得到最終粵方言與普通話之間的理解率。非粵方言被試聽粵方言的理解率是34.97%，粵方言母語者聽普通話的理解率是68.84%，因此相互之間的可理解率是51.91%。

4.5 粵式普通話理解度實驗

4.5.1 實驗材料

本文的研究對象是粵式普通話。粵式普通話目前沒有形成一個明確的固化的語言變體，而是從概念上來把握粵式普通話的特徵。總體來說，有幾個要素，把握這些要素，基本上可以確定該語言變體是粵式普通話。首先粵方言母語者是粵式普通話的核心使用者，可以說粵式普通話的創造者是粵方言母語

者，因此以粵方言母語者所講的粵式普通話作為正宗，不排除一些外地人來到廣東地區為了更好地融合進該語言社區，故意給自己的普通話加入一些粵方言元素。其次是要有粵方言的元素特徵，這個特徵可以是語音的，也可以是詞彙的和句法的特徵。這些特徵必須是粵方言僅有的，舌尖擦音和翹舌音不分，亦或是h-f的混用，都不是粵方言獨有的特徵。如果僅有這些特徵，就不能斷定這一定是粵式普通話。最後無論講話人的發音是否標準，講話人在特定的正式場合與特定的人群講話或對話的過程中，講話人自覺選擇非粵方言的普通話變體，那麼就可視其為粵式普通話。

把握住以上的選擇標準，本文在選擇語音材料時秉承上述原則對講話人的身份，講話場合，講話對象以及粵方言特徵等幾個方面進行篩選，除此之外要求在語言材料的選擇上盡量保留足夠多的粵式普通話特徵，以確保實驗材料更能體現粵式普通話的特點。

本文選擇了崔世昌先生於2016年3月12日在全國政協十二屆四次會議第四次全體會議上就大會議題作出有關推動澳門經濟適度多元可持續發展發言。該發言的視頻片段曾在網絡上流傳，作為兩地網民調侃的片段。本實驗選擇該片段中一些句子作為實驗材料，其原因主要有：崔世昌是廣東新會人，是粵方言的母語者，他也是現任澳門特別行政區特別行政長官崔世安的哥哥，也是澳門特別行政區第六屆立法委員；符合講話人的語言和身份要求。其次他在作報告的場合是全國人民政治協商會議全體會議，這種莊嚴肅穆的場合，使用普通話是一個最基本的要求，在這種場合和環境下，沒有理由認為崔世昌是在使用方言，而且一定是盡他所能講出的對他來說最標準的普通話。另外在他講話的錄像中，幾乎每一句都存在粵方言的痕跡，例如演講中的「橫琴」讀出來近似

「杭坑」，「遊客」讀出來是「游冊」，「寬廣」讀出來是「翻廣」等等。可以說是粵式普通話的典型案例。

除此之外，本實驗還選擇了香港一位1952年出生的張姓老人進行訪談並錄音，該老人年輕時由大陸廣東地區遷往香港居住，除去偶爾回老家之外，便不離開香港地區。在錄音的過程中始終使用普通話進行交流。老人現在香港做大廈安保工作，在錄音過程中，老人也介紹了關於他工作的一些情況。該名老人的普通話也符合上文所述對於粵式普通話的相關原則。

本實驗共節選以上兩段音頻中清晰且語義明瞭的句子作為實驗材料，共二十句。

表格 23 粵式普通話實驗例句信息

實驗例句	音頻時長 (ms)	正向/反向
一般一個禮拜放假一天	3080	正向
世界衛生組織在澳門設立傳統醫藥中心	3403	正向
內地旅客赴橫琴長隆國際海洋度假村	3609	正向
澳門成為一帶一路中醫藥交流合作的重要基地	4651	正向
建議在政策方法內創新通關政策	3115	正向
每年赴澳門遊客達三千多萬	2508	正向
澳門為葡語系國家提供清算服務	2778	正向
每天晚上都是十二小時	2751	正向
我的時間很長	1798	正向
夠時間就回去了	1875	正向
在中央和內地的大力支持下	2584	反向
在澳門聯合舉辦一帶一路經濟論壇	3275	反向
澳門經濟適度多元的道路會越走越風光	3355	反向
創新車輛通關機制	1842	反向
為促進兩地旅遊與合作和一同發展	2854	反向
允許澳門大巴進出橫琴口岸接送旅客	3847	反向
允許赴澳內地遊客多次往返	2961	反向
建議人民銀行繼續支持	2452	反向
這個飯不容易吃	1764	反向
肚餓睡不著諗	1557	反向

其中崔世昌講話節選14句，香港張姓大爺錄音節選6句，平均時長是2802.95毫秒。並且隨機抽取10句作為正向素材，剩下10句作為反向素材，在實驗過程中正向素材的聲音與隨後出現的文字相同，反向素材的聲音與隨後出現的文字不同。

本實驗中的實驗材料均是以句子為單位，沒有使用單獨的詞彙作為實驗材料。主要是因為首先句子更接近實際應用，能夠提供一個完整的語義，在實驗中也可以更好模擬對話交流的場景。其次，本實驗研究的是粵式普通話，由於其本身還存在著一些不穩定性，所以很難能夠分離出一個單獨的粵式普通話詞彙，即便從句子中切割出一個詞彙，也很難證明這個詞彙的代表性。最後，由於粵式普通話中存在大量的粵方言混入的情況，一句話中可能會有一兩個詞使用粵方言詞彙。如果僅將該詞彙切分出來，亦或是不保留這類詞彙，都有失偏頗，有人為操控結果的嫌疑。綜上本實驗就只選擇句子為單位的實驗材料，以求試驗結果的真實有效。

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006) 在進行理解度實驗後對被試進行了主觀問卷調查，詢問了實驗錄音的三個屬性，分別是理解度 (Intelligibility)、可懂度 (Comprehensibility) 和口音程度 (Accentedness) 用以和客觀實驗結果進行比較。因此本實驗除去上述實驗材料外，在理解度實驗後，同樣添加了調查問卷。以問卷的方式調查被試口音方面，理解度方面，歸屬問題和理解障礙四個方面的問題。問卷內容如下：

(一) 剛才聽到的錄音，您認為：

1、完全沒有口音；2、有一些口音；3、沒感覺；4、有較重的口音；5、

非常重的口音

(二) 剛才聽到的錄音，您認為：

- 1、完全能理解；2、能理解大部分；3、能理解一半；4、能理解小部分；
- 5、完全不理解

(三) 剛才聽到的錄音，您認為它是：

- 1、標準的普通話；2、不太標準的普通話；3、不太標準的方言；4、標準的方言；5、沒感覺

(四) 您認為對理解剛才聽到的錄音最大的障礙是：

- 1、語音問題；2、詞彙問題；3、句法問題

4.5.2 實驗流程

本實驗同樣使用由卡內基梅隆大學、匹茲堡大學學習研究與發展中心和美國心理學軟件工具公司研發的心理學行為研究實驗軟件E-prime 2.0進行實驗。

實驗共有五部分：

(1) 歡迎屏幕：此頁面上詳細描述了實驗的操作方法：被試根據聽到的語音材料選擇隨後顯示的文字是否與先前所聽到的內容一致，如果聽到的內容與顯示的內容一致，則按鍵盤上的「F」鍵，如果聽到的內容與顯示的內容不一致，則按鍵盤上的「J」鍵。此階段由測試人員為被試進行講解，確認被試掌握實驗方法後，按任意鍵繼續進入下一部分。

(2) 練習部分：此部分目的是為了讓被試能夠直接熟悉操作的方法，進一

步記住按鍵和操作流程。此環節分三個單元。(1)凝視點：屏幕中央上出現凝視點「+」，時長200ms，不需做反應直接進入第二單元；(2)聲音刺激：屏幕上不出現任何信息，被試從耳機中聽到聲音刺激後不需做反應語音播放完畢後直接進入第三單元；(3)顯示文字：屏幕上顯示文字，總時長5000ms，被試需要根據判斷做出選擇，若被試沒有在規定時間內做出選擇，則自動跳轉至下一刺激。練習部分結束後會讓被試進行選擇，如果已經掌握實驗方法，則進入正式測試，如果還需練習則會重新進入練習部分。

(3) 實驗模塊：實驗與練習模塊設置相同，包含凝視點、聲音刺激和顯示文字三個單元，三個單元依次出現，三個單元為一組。不同之處在於，實驗模塊的第三單元顯示文字最長時間是10000毫秒，在第三單元軟件會記錄被試的反應內容，正確率以及反應時長。實驗列表中共有20組刺激組合，會按照軟件的隨機種子隨機播放，並且每一組刺激只會出現一次。20組刺激運行完畢後自動跳入下一部分。

(4) 問卷部分：此部分屏幕依次顯示每道題目，被試可以在鍵盤上輸入他們選擇的答案的相應按鍵。此部分的每道題目作答時間沒有時間限制。答完四道題後自動跳入下一部分。

(5) 結束屏幕：感謝被試的參與，並結束全部實驗。

4.5.3 實驗對象

本次實驗實驗對象與第三章普通話與粵方言理解度實驗的研究對象大致相同，只是沒有使用澳門非粵方言被試參與本次實驗，其餘被試均與前文一致。之所以沒有邀請澳門大學的學生參與本次實驗，主要考慮到因為崔世昌視頻中

所涉及的内容是與澳門相關的。例如「橫琴口岸」，「葡語系國家」，「通關」，「多次往返」，這類詞語在澳門是非常常見和常用的，帶有強烈的地方背景，勢必會對理解度有相應的提高。這種信息背景的優勢就可能與其他組別的被試相比，造成信息不對稱，從而使得數據不精確。故而排除了澳門非粵方言被試作為本實驗的被試。

參與本次實驗的被試共有220名，其中中國人201名，外國留學生19名參與實驗，參與者並非無償進行實驗，實驗者向所有實驗參與者支付相關報酬或等值的禮品以示回饋。中國被試者大多數是在校學生，本實驗先後在天津外國語大學，天津師範大學，北京大學等四所大學以及天津中學高中部進行隨機抽樣，各個年級均有樣本分布。由於本實驗是調查對粵式普通話的理解度，因此優先選擇非粵方言區的被試。因為粵方言區的講話人若是對本方言區內的人講話可以直接使用粵方言，因此粵式普通話就是針對非粵方言區的聽話人所使用的。在實際實驗的過程中，也有粵方言區的被試參與實驗，原則上並不排斥其實驗結果，其結果在分析時會專門分析。

4.5.3.1 性別分布

本次實驗男性被試89名，佔全部人數的40.5%，女性被試131名，佔全部人數的59.5%。天津外國語大學、天津師範大學和北京大學參與者共有107名女性，60名男性。天津中學高中部參與者，共有11名女生，23名男生。

4.5.3.2 年齡

本實驗被試的年齡平均值是21.05歲，標準差3.776，方差16.592。年齡最小值是15歲，最大值是52歲，人數最多的年齡是21歲。此次試驗主要針對的是在

校生，故而年齡主要集中在18歲以上至30歲以下。男性被試的平均年齡是21.30歲，標準偏差5.228；女性被試的平均年齡是20.89歲，標準偏差2.336，總體上略小於男性。經獨立樣本T檢驗，二者之間沒有顯著性差異（ $T=0.805$ ， $df=218$ ， $sig=0.422>0.05$ ）

4.5.3.3籍貫

本實驗共收集到201名中國被試，其中包含了我國29個省市，未能找到來自寧夏回族自治區、青海省和上海市三地的被試。由於實驗地點是在天津市和北京市的學校內，因此學生大多數是來自北方。排在前七位的都是北方的省份，前七位的省份共佔全部被試的60.9%。另外還有19名被試來自國外，其中15名被試來自泰國，其餘4名分別來自贊比亞、馬達加斯加、薩摩亞和美國。

表格 24 粵式普通話實驗被試籍貫信息統計

省份	頻數	百分比	省份	頻數	百分比
天津	52	25.9	貴州	5	2.5
甘肅	16	8.0	內蒙	4	2.0
河北	14	7.0	湖南	4	2.0
山東	13	6.5	四川	3	1.5
河南	11	5.5	福建	3	1.5
北京	8	4.0	安徽	2	1.0
新疆	8	4.0	廣東	2	1.0
江西	7	3.5	江蘇	2	1.0
陝西	7	3.5	西藏	2	1.0
山西	6	3.0	遼寧	2	1.0
雲南	5	2.5	重慶	2	1.0
吉林	5	2.5	台灣	1	0.5
廣西	5	2.5	海南	1	0.5
浙江	5	2.5	黑龍江	1	0.5
湖北	5	2.5	總計	201	100

4.5.3.4第一語言

本實驗調查了所有被試的第一語言，並且明確告知被試是第一個學會的語言。以官話為第一語言的被試數量最多共有94名，佔到了全部人數的42.7%；其次是以普通話為第一語言的被試，共有83名，佔全部人數的37.7%。前兩項加起來就佔到了全體人數的80.4%。其餘方言區均有被試參與。

表格 25 粵式普通話測試被試第一語言信息統計

	頻數	百分比
官話	94	42.7
普通話	83	37.7
外族語	22	10.0
湘方言	5	2.3
閩方言	4	1.8
吳方言	3	1.4
客家話	3	1.4
粵方言	3	1.4
贛方言	3	1.4
總計	220	100.0

對以官話為第一語言的被試進行調查，其中以北方官話為第一語言的被試共有56人，佔以官話為第一語言的人數的59.6%，其次是以西南官話為第一語言的被試共有16人，佔以官話為第一語言的人數的17.0%，再次是以中原官話為第一語言的被試共有13人，佔以官話為第一語言的人數的13.8%，最後是以西北官話為第一語言的被試共有9人，佔以官話為第一語言的人數的9.6%。

表格 26 粵式普通話官話方言區被試頻率統計

	頻數	百分比
北方官話	56	59.6
西南官話	16	17.0
中原官話	13	13.8
西北官話	9	9.6
總計	94	100.0

4.5.3.5教育背景

此次實驗共邀請了從高中一年級到博士五年級，一共 16 個年級的學生參加，其中還有一些本科畢業或碩士畢業在學校中復習準備升學考試的同學以及畢業生。其中高中生共有 34 人，佔總體的 16.9%，本科生及本科畢業生共有 140 人，佔總體的 60.7%；碩士研究生及碩士畢業生共有 39 人，佔總體的 18.9%；博士生共有 7 人，佔總體的 3.5%。數量最多的年級是大三，共有 55 人，佔總體的 27.4%。

表格 27 粵式普通話被試教育信息統計

年級	頻數	百分比	年級	頻數	百分比
高一	6	3.0	研一	26	12.9
高二	13	6.5	研二	8	4.0
高三	15	7.5	研三	1	0.5
大一	35	17.4	研究生	3	1.5
大二	18	9.0	博一	3	1.5
大三	55	27.4	博二	1	0.5
大四	12	6.0	博三	1	0.5
大五	1	0.5	博四	1	0.5
本科	1	0.5	博五	1	0.5
			總計	201	100.0

4.5.4 實驗地點

本實驗三所大學一所高中進行實驗，四所學校集中在天津市內和北京市內。為保證實驗的順利進行，防止噪音對實驗的干擾。實驗地點都選自安靜的教室、圖書館或語言實驗室。具體地點見下表：

表格 28 粵式普通話實驗地點

學校	實驗地點
天津師範大學	天津師範大學津沽學院教室A101 天津師範大學金橋賓館教室808
天津外國語大學	天津外國語大學逸夫樓語音實驗室E601

北京大學	北京大學圖書館自習區
天津中學	天津中學教室A360

4.5.5 實驗結果

實驗過程中，第三部分主要收集被試的反應和反應時間兩項數據，根據正確反應值與被試的反應相比較得到正確值。若被試答對，那麼正確值為1，如果答錯，則正確值為0。反應時間的單位是毫秒（ms）；音頻結束後立即顯示文字，因此反應時間就從音頻結束後開始計算。本文數據分析所用到的基礎數據就是正確值和反應時間。第四部分只收集被試的反應，結果是從1到5的數字，每道題目只收集一次。本實驗共有220名被試，每名被試均有20道題目和4道問卷題目，因此本實驗共有 $220 \times 20 \times 2 + 220 \times 4 = 9680$ 個有效數據。本實驗所有的數據分析均使用IBM SPSS Statistics 22版，32位版。

本實驗的理解度計算方法參考前文所述，利用實驗收集反應時間，經計算獲得時間參數再利用時間參數與正確值相乘獲得該被試在該題的理解度。

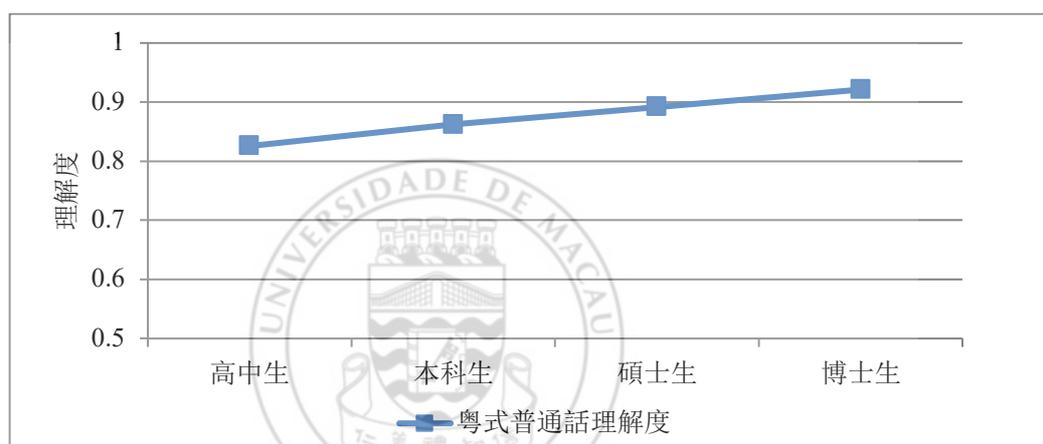
4.5.5.1 性別因素

總體上男性對於粵式普通話的平均理解度是0.833，標準差是0.549；女性的平均理解度是0.841，標準差是0.554。男性的理解度略高於女性，經過獨立樣本T檢驗，二者之間並不存在顯著差異。（ $T = -0.448$, $df = 4841.437$, $sig = 0.625 > 0.05$ ）

4.5.5.2 教育背景

按照被試的教育背景進行分組，將被試分成高中組，大學生組，碩士生組

和博士生組四組，分別對其進行統計，其中不包括外籍留學生。高中組的對粵式普通話的平均理解度是0.826，標準差是0.550；大學生組的對粵式普通話的平均理解度是0.862，標準差是0.547；碩士生組的對粵式普通話的平均理解度是0.892，標準差是0.525；博士生組的對粵式普通話的平均理解度是0.921，標準差是0.520。



圖表 16 不同教育背景下粵式普通話理解度

通過單因素方差分析各組之間不存在顯著性差異。[$F(3, 4036)=2.343$, $\text{sig}=0.071>0.05$]。不過經Tukey HSD事後檢驗可知，一對一單獨比較後，各個組之間並無顯著性差異，互相的差異僅僅是量上的差別，並沒有質的差異。高中生組與本科生組($\text{sig}=0.406>0.05$)、碩士生組並無顯著性差異($\text{sig}=0.093>0.05$)和博士生組($\text{sig}=0.228>0.05$)均不存在顯著性差異。本科生組與碩士生組($\text{sig}=0.543>0.05$)和博士生組($\text{sig}=0.592>0.05$)也不存在顯著性差異、碩士生組和博士生組($\text{sig}=0.937>0.05$)之間也不存在顯著性差異。由此可見教育因素並不能對粵方言普通話產生顯著性影響，不過整體數值也是隨著學歷的提高而提高，因此教育因素會對理解度產生一定量的影響，但是不是顯著的影響因素。

4.5.5.3 第一語言

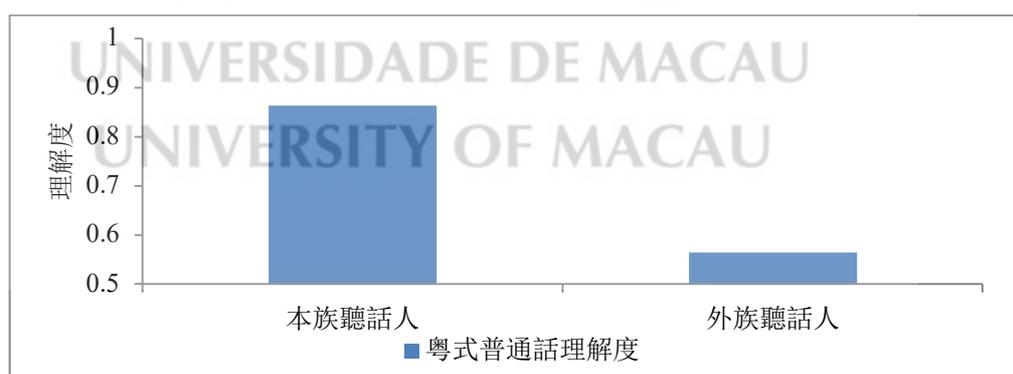
按照被試的第一語言進行分組，可以得到八個組，分別是官話組，普通話組，粵方言組，湘方言組，吳方言組，外族語組，客家話組，贛方言組和閩方言組。以官話作為第一語言的被試對粵式普通話的平均理解度是0.867，標準差是0.549；以普通話作為第一語言的被試對粵式普通話的平均理解度是1.040，標準差是0.487；以粵方言作為第一語言的被試對粵式普通話的平均理解度是0.920，標準差是0.520；以湘方言作為第一語言的被試對粵式普通話的平均理解度是0.859，標準差是0.562；以吳方言作為第一語言的被試對粵式普通話的平均理解度是0.780，標準差是0.503；以外族語作為第一語言的被試對粵式普通話的平均理解度是0.605，標準差是0.577；以客家話作為第一語言的被試對粵式普通話的平均理解度是0.715，標準差是0.518；以贛方言作為第一語言的被試對粵式普通話的平均理解度是0.767，標準差是0.564；以閩方言作為第一語言的被試對粵式普通話的平均理解度是0.834，標準差是0.548。

最高的一組是以普通話為母語的被試，得分是1.040，甚至超過了以粵方言為母語的被試。經過單因素方差分析，各組之間存在顯著性差異 [F(8, 2791)=11.612, sig=0.00<0.05]，經過Tukey HSD事後檢驗後發現，僅有以外族語為母語的被試與其他組有顯著性差異，其餘各組之間均無顯著性差異。可見對於第一語言是漢語的人來說，粵式普通話的理解程度大致相當，並無顯著性差異；而對於第一語言非漢語的人來說，粵式普通話的理解度就與第一語言是漢語的人來說相差非常大。

4.5.5.4分組

根據第三章的分組方法，將本實驗中的被試分為三組，與第三章不同的是

本實驗中沒有澳門非粵方言被試組，僅保留了外國留學生組、內地非粵方言被試組和粵方言母語者組。外國留學生對粵式普通話的平均理解度是0.564，標準差是0.578；內地非粵方言被試對粵式普通話的平均理解度是0.853，標準差是0.543；粵方言母語者組對粵式普通話的平均理解度是0.920，標準差是0.520。經過單因素方差分析，各組之間存在顯著性差異 $[F(2, 4417)=52.715, sig=0.00<0.05]$ ，經過Tukey HSD事後檢驗後發現，外國留學生組與內地非粵方言被試組（ $sig=0.00<0.05$ ）和粵方言母語者組（ $sig=0.00<0.05$ ）均有顯著性差異。而內地非粵方言被試組和粵方言母語者組之間並沒有顯著性差異（ $sig=0.702>0.05$ ）。因此可將內地非粵方言被試組和粵方言母語者組合併成為一組，將整體被試分成兩組，分別為本族聽話人組和外籍聽話人組。本族聽話人對粵式普通話的理解度為0.864，標準差是0.543；外籍聽話人對粵式普通話的理解度為0.564，標準差是0.578。經獨立樣本T檢驗驗證，二者之間存在顯著性差異（ $T=10.237, df=4418, sig=0.000<0.05$ ）。



圖表 17 本族語和外族聽話人對粵式普通話的理解度

4.5.6 主觀評價結果

對被試進行主觀理解度評判一方面能夠從主觀和客觀兩個角度分別看待理解度測試，讓理解度測試的結果更加全面，另一方面也可以直觀地反映出主客

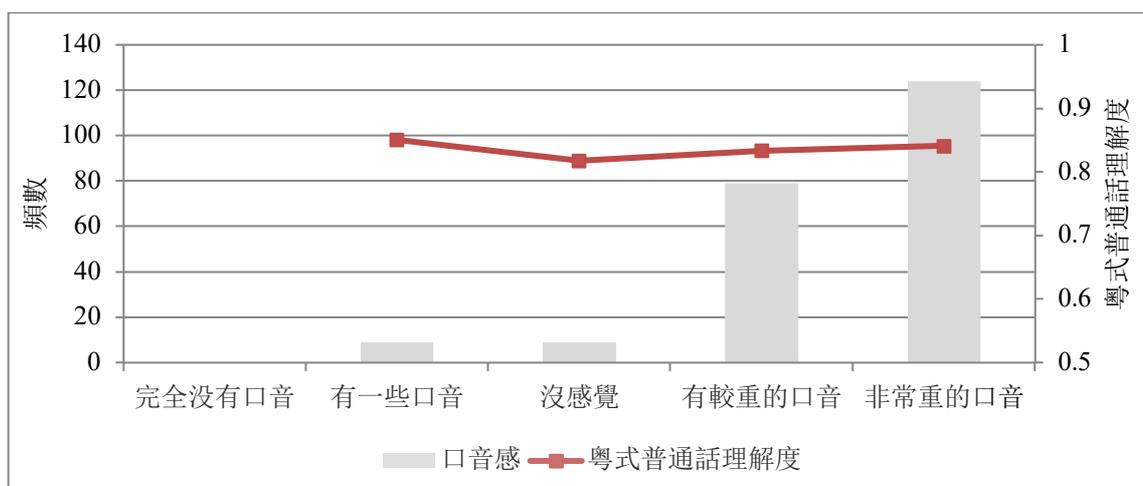
觀的差異，正確認識主客觀評判之間的差異，並不是一定要以客觀結果為主，也不是一味摒棄主觀認知。因為其實理解與沒有理解這個問題本身是一個主觀現象，是因人而異的，有些人可能理解了，有些人可能並沒有理解；有些人理解的快，有些人理解的慢。雖然理解是主觀性的依舊是依靠客觀性的實驗來測試出來，是因為人的主觀判斷並不能給出一個客觀準確的理解度，僅能給出一個大概的範圍，並且通常受到情緒或其他因素的影響。

本實驗中的主觀評價共有四道題，分別是對聽到的粵式普通話的口音度，理解度進行調查。然後調查了被試對聽到的內容的歸類，究竟是屬於普通話還是屬於方言，是不標準的普通話還是不標準的方言。最後是調查了被試認為影響理解錄音內容的最大的障礙是什麼。下文會逐個題目進行分析。

4.5.6.1 口音感

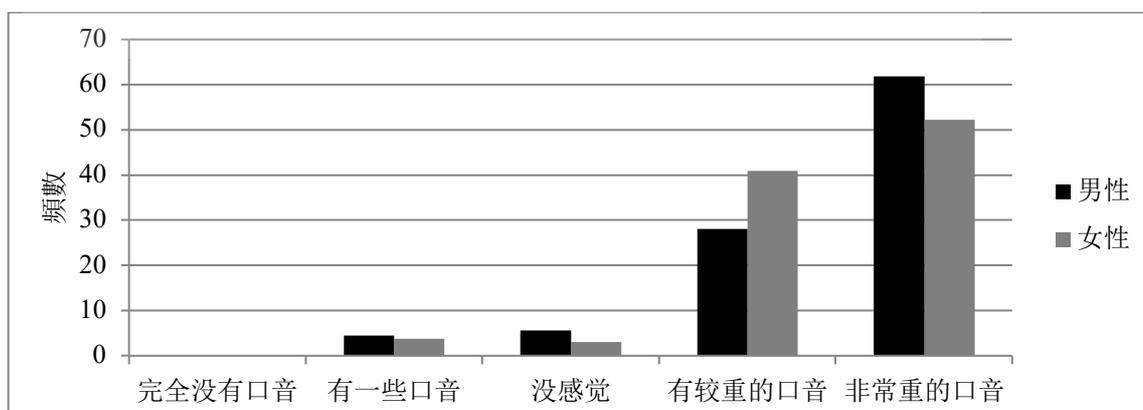
第一題是被試的對聽到錄音的口音進行評價，沒有人認為沒有口音，這一點是很正常的。有9名被試認為有一些口音，佔全部被試的4.07%，有9名被試認為沒什麼感覺，佔全部被試的4.07%，79名被試認為有比較重的口音，佔全部被試的35.75%，124名被試認為有很重的口音，佔全部被試的56.11%。同時統計到選擇「2有一些口音」的被試平均理解度是0.850，標準差是0.133；選擇「3沒感覺」的被試理解度是0.818，標準差是0.213；選擇「4有較重的口音」的被試理解度是0.833，標準差是0.171；選擇「5有非常重的口音」的被試理解度是0.841，標準差是0.171。經單因素方差分析，這四者之間並無顯著性差異。[$F(3, 217)=0.090, sig=0.966>0.05$]，經Tukey HSD事後檢驗表明，四者之間均不存在顯著性差異，所有顯著性系數均大於0.05。可見雖然被試對口音的感知不同，但是

並沒有實際影響理解度。



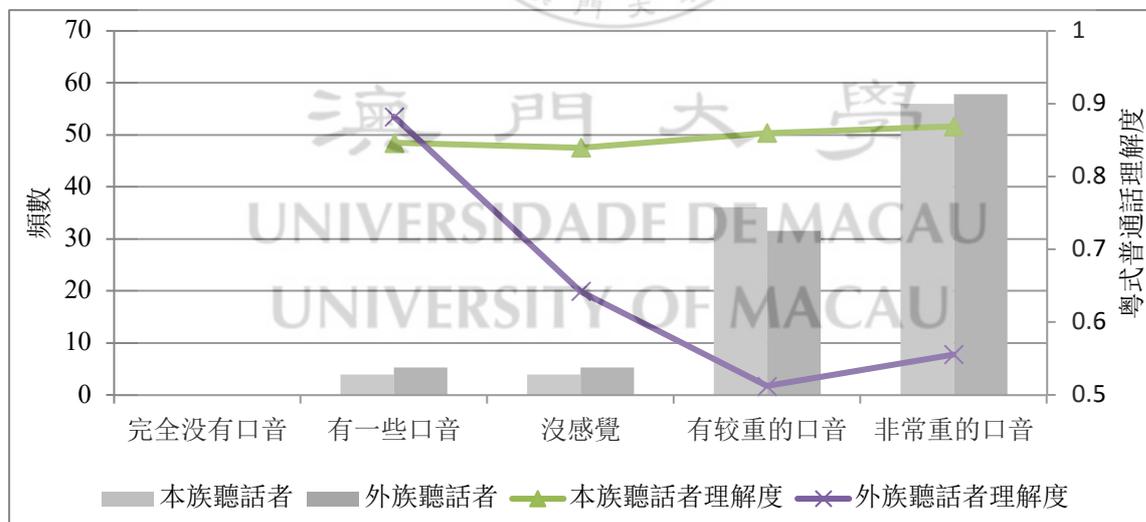
圖表 18 口音感評價與理解度

其中有4.49%的男性被試和3.79%的女性被試選擇「2有一些口音」；有5.62%的男性被試和3.03%的女性被試選擇「3沒感覺」；有28.09%的男性被試和40.91%的女性被試選擇「4有較重的口音」；有61.80%的男性被試和52.27%的女性被試選擇「有非常重的口音」。僅在「有較重的口音」的選項上，女性被試的比例超過男性被試的比例，在其他幾個的選項上，男性被試的比例均高於女性。再此道題目上，男性和女性的最大值都出現在非常重的口音這個選項上，性別因素並沒有對口音感知這個問題上產生差異。



圖表 19 不同性別的口音感評價

與本族聽話者相對的外族聽話者中，有1名被試認為有一些口音，佔全部被試的5.26%，有1名被試認為沒什麼感覺，佔全部被試的5.26%，6名被試認為有比較重的口音，佔全部被試的31.58%，11名被試認為有很重的口音，佔全部被試的57.89%。基本與本族聽話者的感知結果相同。本族聽話者中，選擇有「一些口音」選項的被試的理解度是0.846，標準偏差0.141；選擇有「沒感覺」選項的被試的理解度是0.839，標準偏差0.217；選擇「有較重的口音」選項的被試的理解度是0.860，標準偏差0.147；選擇有「非常重的口音」的被試的理解度是0.869，標準偏差0.140。外族聽話者中，選擇有「一些口音」選項的被試的理解度是0.882；選擇有「沒感覺」選項的被試的理解度是0.642；選擇「有較重的口音」選項的被試的理解度是0.512，標準偏差0.107；選擇有「非常重的口音」的被試的理解度是0.556，標準偏差0.205。



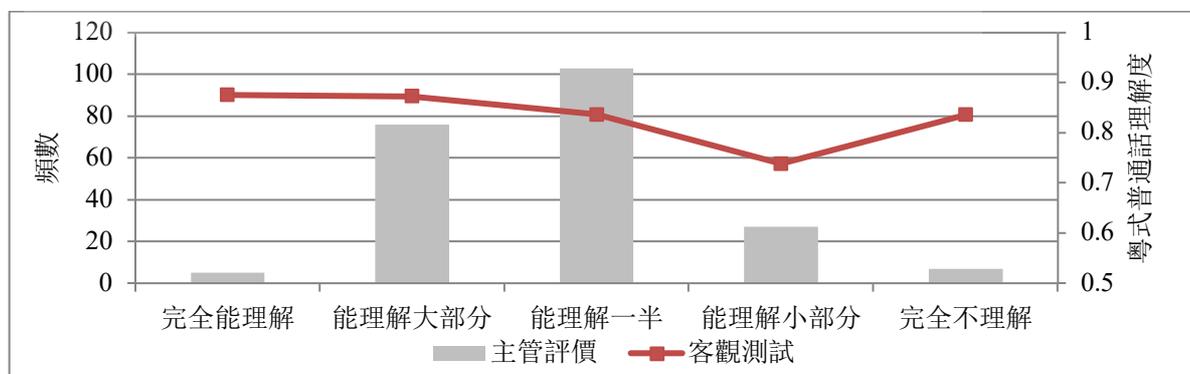
圖表 20 本族和外族聽話者的口音感評價

4.5.6.2 理解度

第二題是關於粵式普通話的理解度問題，有5名被試認為完全理解，佔全部被試的2.26%，有76名被試認為能夠理解大部分內容，佔全部被試的34.39%，103

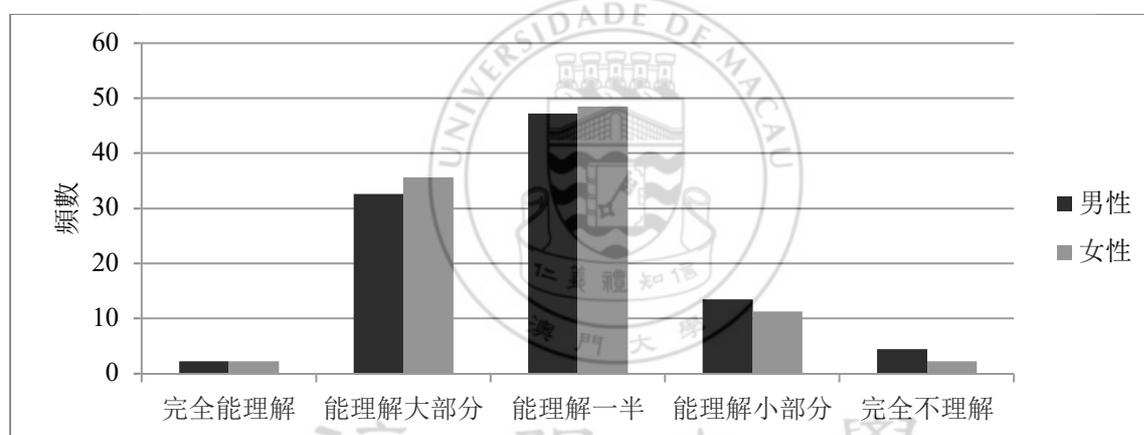
名被試認為能理解一半，佔全部被試的47.96%，有27名被試認為僅能理解一小部分，佔全部被試的12.22%；有7名被試認為完全不理解，佔全部被試的3.17%。

同時統計到選擇「完全能理解」的被試平均理解度是0.876，標準差是0.229；選擇「能理解大部分」的被試理解度是0.873，標準差是0.147；選擇「能理解一般」的被試理解度是0.837，標準差是0.155；選擇「能理解一小部分」的被試理解度是0.738，標準差是0.237；選擇「完全不理解」的被試理解度是0.836，標準差是0.205。經單因素方差分析，這四者之間並無顯著性差異。[F(4, 216)=3.286, sig=0.012<0.05]，經Tukey HSD事後檢驗表明，第一組、第二組、第三組、第四組之間無顯著性差異，顯著性系數均大於0.05，僅有第二組和第四組之間有顯著性差異。總體上看，主觀評價和客觀測試的結果是基本吻合的，認為完全能夠理解的被試有些高估自己的理解度，自己的評價值是高於實際測量值的。解釋這種現象可以有兩種解釋思路，一種是客觀測試的測試機制還需要調整，另外一種是主觀的差距在理解度測試結果上的表現並不是線性的，而是一種非線性的表現。換句話說就是主觀心理理解度的臨界值與客觀測試的理解度值之間是非線性關係。這以現象還有賴於進一步的實驗驗證。也可能出現被試對自己的主觀感受表現出一種謙虛的態度。



圖表 21 理解度主觀評價

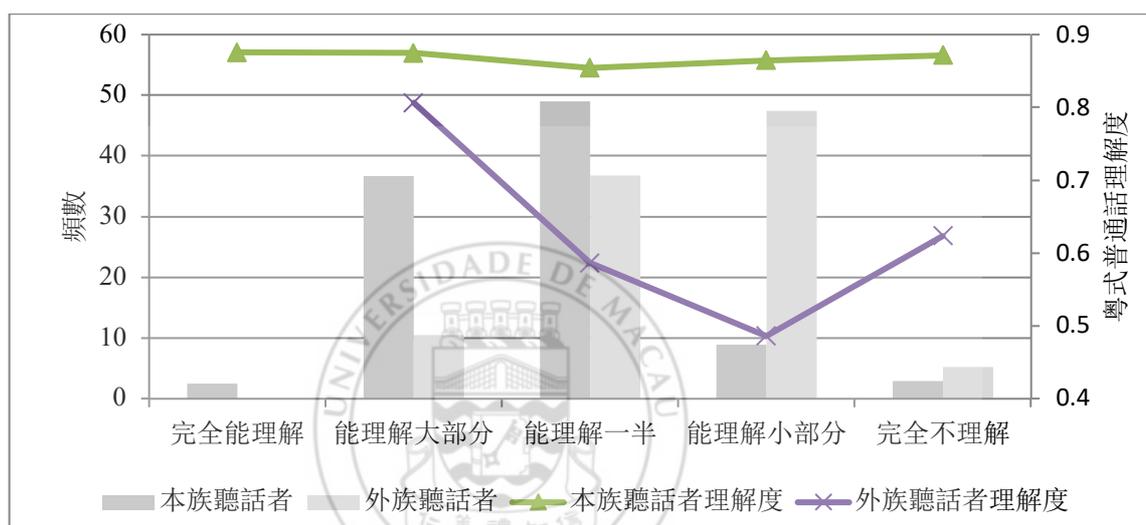
其中有2.25%的男性被試和2.27%的女性被試選擇「完全能理解」；有32.58%的男性被試和35.606%的女性被試選擇「能理解大部分」；有47.19%的男性被試和48.49%的女性被試選擇「能理解一半」；有13.49%的男性被試和11.36%的女性被試選擇「能理解小部分」；有4.49%的男性被試和2.27%的女性被試選擇「完全不理解」。再此題目上，男性與女性展現出同樣的心理反應，不同性別的主觀感受並沒有太大差別，這也側面輔證了之前通過客觀實驗的方法得出的理解度，男性與女性之間的差異並不顯著的結論。



圖表 22 不同性別的理解度主觀評價

外籍聽話者中，沒有有被試認為完全理解，有2名被試認為能夠理解大部分內容，佔全部被試的10.53%，7名被試認為能理解一半，佔全部被試的36.84%，有9名被試認為僅能理解一小部分，佔全部被試的47.37%；有1名被試認為完全不理解，佔全部被試的5.26%。本族聽話者中，選擇有「完全能理解」選項的被試的理解度是0.876，標準偏差0.229；選擇有「能理解大部分」選項的被試的理解度是0.875，標準偏差0.145；選擇「能理解一半」選項的被試的理解度是0.854，標準偏差0.141；選擇有「能理解小部分」的被試的理解度是0.865，標準偏差0.142；選擇有「完全不理解」的被試的理解度是0.871，標準偏差0.200。外族聽話者中，

選擇有「能理解大部分」選項的被試的理解度是0.806，標準偏差0.264；選擇「能理解一半」選項的被試的理解度是0.586，標準偏差0.130；選擇有「能理解小部分」的被試的理解度是0.486，標準偏差0.177；選擇有「完全不理解」的被試的理解度是0.623。

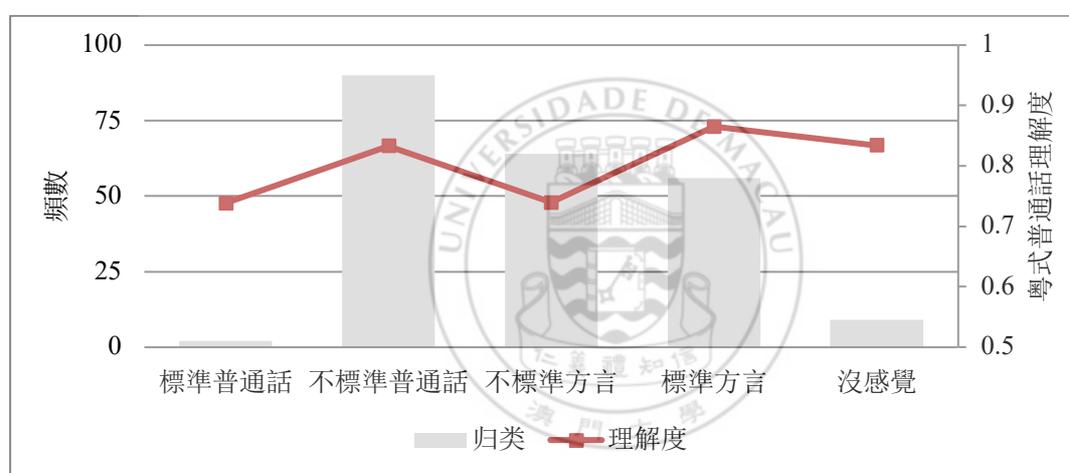


圖表 23 本族與外族聽話者的主觀理解度評價

4.5.6.3 普通話或方言分類

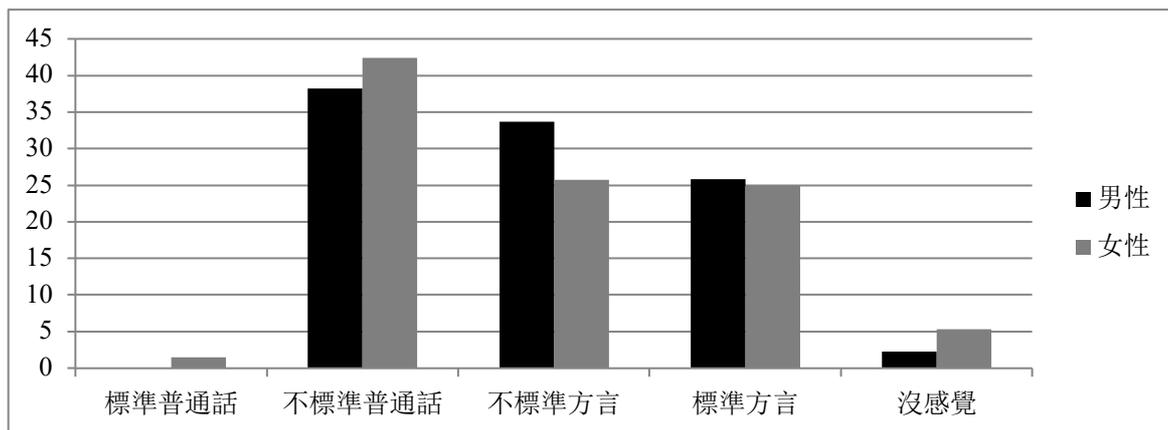
第三題是關於粵式普通話的分類問題，有2名被試認為聽到的錄音是標準普通話，佔全部被試的0.91%，有90名被試認為聽到的錄音是不標準普通話，佔全部被試的40.72%，有64名被試認為聽到的錄音是不標準方言，佔全部被試的28.96%，有56名被試認為聽到的錄音是標準的方言，佔全部被試的25.34%，有9名被試認為沒有對其合適的判斷，佔全部被試的4.07%。同時統計選擇「標準普通話」的被試平均理解度是0.739，標準差是0.395；選擇「不標準的普通話」的被試理解度是0.833，標準差是0.191；選擇「不標準的方言」的被試理解度是0.739，標準差是0.173；選擇「標準的方言」的被試理解度是0.865，標

準差是0.161；選擇「沒感覺」的被試理解度是0.834，標準差是0.135。經單因素方差分析，這五者之間並無顯著性差異。[$F(4, 216)=1.355, sig=0.251>0.05$]，經Tukey HSD事後檢驗表明，幾組之間均不存在顯著性差異，所有的顯著性系數均大於0.05。絕大多數的被試還是將粵式普通話劃分到不標準的普通話這一範疇下，雖然是不標準的普通話也是處於普通話的範疇內，因此還是有更多的人傾向於將粵式普通話劃歸到普通話的範疇下。



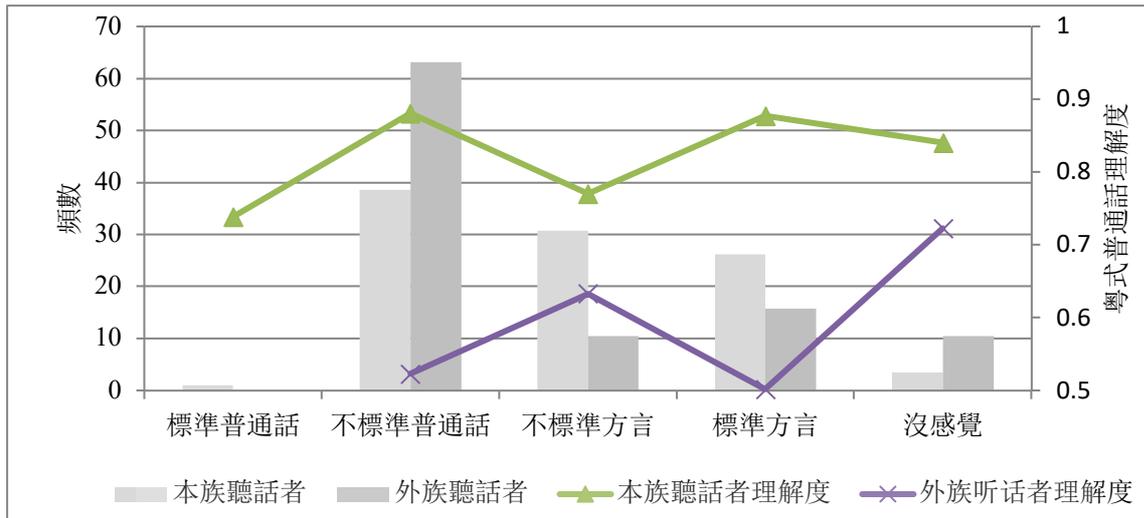
澳門大學
圖表 24 粵式普通話性質主觀評價

其中僅有1.15%的女性被試認為聽到的粵式普通話是「標準的普通話」，沒有男性被試這樣認為；有23.20%的男性被試和42.42%的女性被試認為聽到的粵式普通話是「不標準的普通話」；有33.71%的男性被試和25.76%的女性被試認為聽到的粵式普通話是「不標準的方言」；有25.84%的男性被試和25%的女性被試認為聽到的粵式普通話是「標準的方言」；有2.25%的男性被試和5.30%的女性被試對聽到的粵式普通話沒有合適的分類結果。不同性別的被試大多數都集中認為粵式普通話應屬於不標準的普通話也是將其納入普通話的範疇中。



圖表 25 不同性別的普通話性質主觀評價

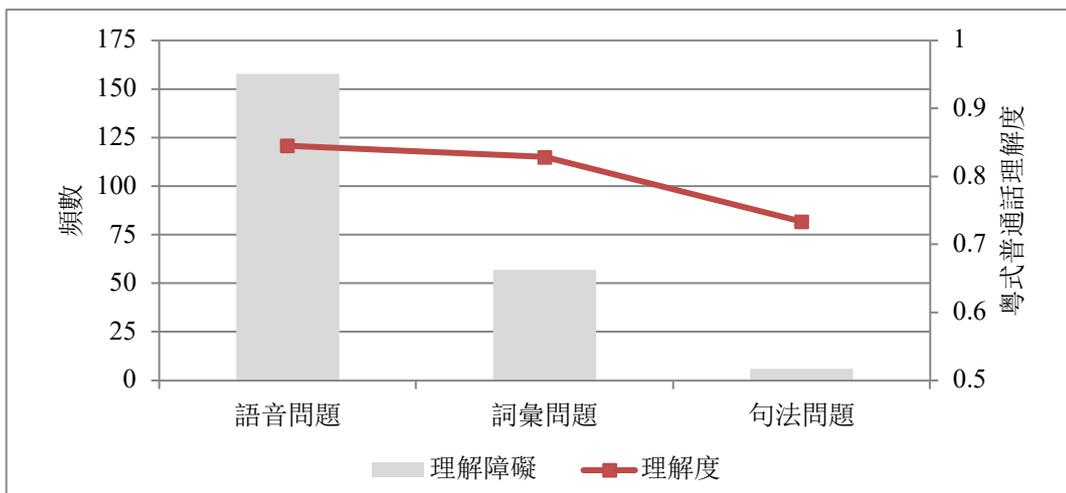
外籍聽話者方面，沒有被試認為聽到的錄音是標準普通話，有12名被試認為聽到的錄音是不標準普通話，佔全部被試的63.16%，有2名被試認為聽到的錄音是不標準方言，佔全部被試的10.53%，有2名被試認為聽到的錄音是標準的方言，佔全部被試的10.53%，有3名被試認為沒有對其合適的判斷，佔全部被試的15.79%。本族聽話者中，選擇有「標準普通話」選項的被試的理解度是0.739，標準偏差0.395；選擇「不標準普通話」選項的被試的理解度是0.881，標準偏差0.144；選擇有「不標準方言」的被試的理解度是0.877，標準偏差0.146；選擇有「標準方言」的被試的理解度是0.840，標準偏差0.128；選擇有「沒感覺」選項的被試的理解度是0.770，標準偏差0.187。外族聽話者中，選擇「不標準普通話」選項的被試的理解度是0.523，標準偏差0.171；選擇有「不標準方言」的被試的理解度是502，標準偏差0.235；選擇有「標準方言」的被試的理解度是0.723，標準偏差0.236；選擇有「沒感覺」選項的被試的理解度是0.633，標準偏差0.146。



圖表 26 本族與外族聽話者粵式普通話性質主觀評價

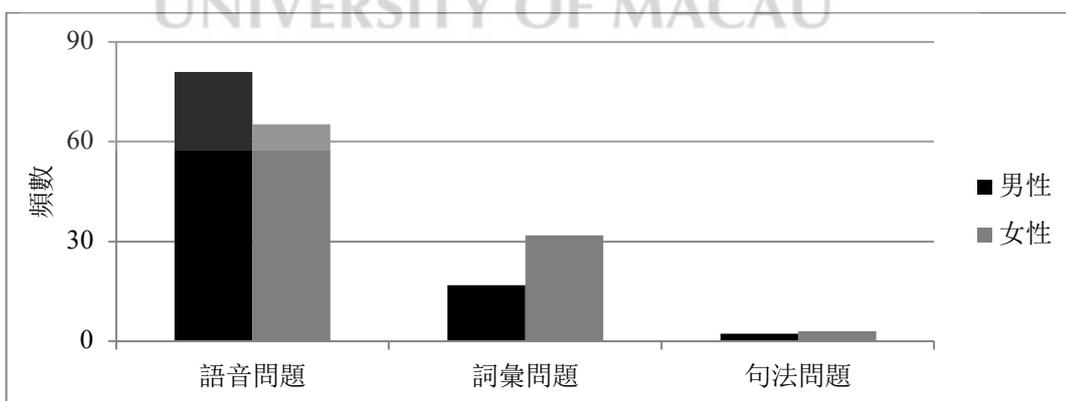
4.5.6.4 理解的障礙

最後一題是調查關於理解粵式普通話最大的障礙，有158名被試認為對理解粵式普通話最大的障礙在於語音問題，佔全部被試的71.49%，有57名被試認為對理解粵式普通話最大的障礙在於詞彙問題，佔全部被試的25.79%，有6名被試認為對理解粵式普通話最大的障礙在於句法問題，佔全部被試的2.71%。同時統計選擇「語音問題」的被試平均理解度是0.845，標準差是0.173；選擇「詞彙問題」的被試平均理解度是0.829，標準差是0.167；選擇「句法問題」的被試平均理解度是0.734，標準差是0.135。經單因素方差分析，這三者之間並無顯著性差異。 $[F(2, 218)=1.350, sig=0.262>0.05]$ ，經Tukey HSD事後檢驗表明，三組之間均不存在顯著性差異，所有的顯著性系數均大於0.05。



圖表 27 理解障礙調查

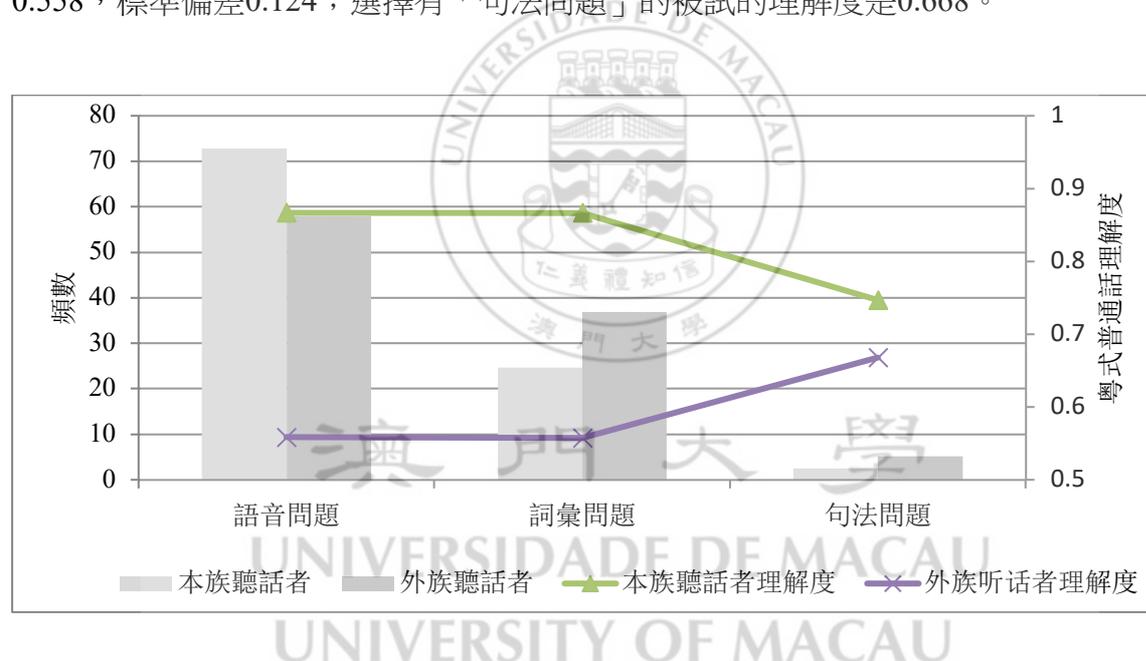
其中有80.899%的男性被試和65.15%的女性被試認為對理解粵式普通話最大的障礙是語音方面的問題；有16.85%的男性被試和31.82%的女性被試認為對理解粵式普通話最大的障礙是詞彙方面的問題；有2.25%的男性被試和3.03%的女性被試認為對理解粵式普通話最大的障礙是句法方面的問題。實際上在本次實驗中使用的實驗材料，並沒有出現粵式普通話中句法上的特點。在語音方面，更多的男性被認為是語音導致的理解障礙，而在詞彙方面，更多的女性被試認為是詞彙問題導致的理解障礙。



圖表 28 不同性別的理解困難點調查

外籍聽話者方面有11名被試認為對理解粵式普通話最大的障礙在於語音問

題，佔全部被試的57.89%，有7名被試認為對理解粵式普通話最大的障礙在於詞彙問題，佔全部被試的36.84%，有1名被試認為對理解粵式普通話最大的障礙在於句法問題，佔全部被試的5.26%。本族聽話者中，選擇有「語音問題」選項的被試的理解度是0.867，標準偏差0.148；選擇「詞彙問題」選項的被試的理解度是0.867，標準偏差0.134；選擇有「句法問題」的被試的理解度是0.747，標準偏差0.146。外族聽話者中，選擇有「語音問題」選項的被試的理解度是0.558，標準偏差0.223；選擇「詞彙問題」選項的被試的理解度是0.558，標準偏差0.124；選擇有「句法問題」的被試的理解度是0.668。



圖表 29 本族與外族聽話者理解障礙調查

4.5.6 分析與討論

本節將會對本章實驗進行總結分析，通過實驗數據的分析得到相關結論。

(一) 本族人對粵式普通話的理解度表現出一致性，外族人表現出與本族人的顯著性差異。首先從理解度的角度上看，本族人的理解度平均值是0.864，標準差是0.145，外籍聽話人的理解度平均值是0.564，標準差是0.179。

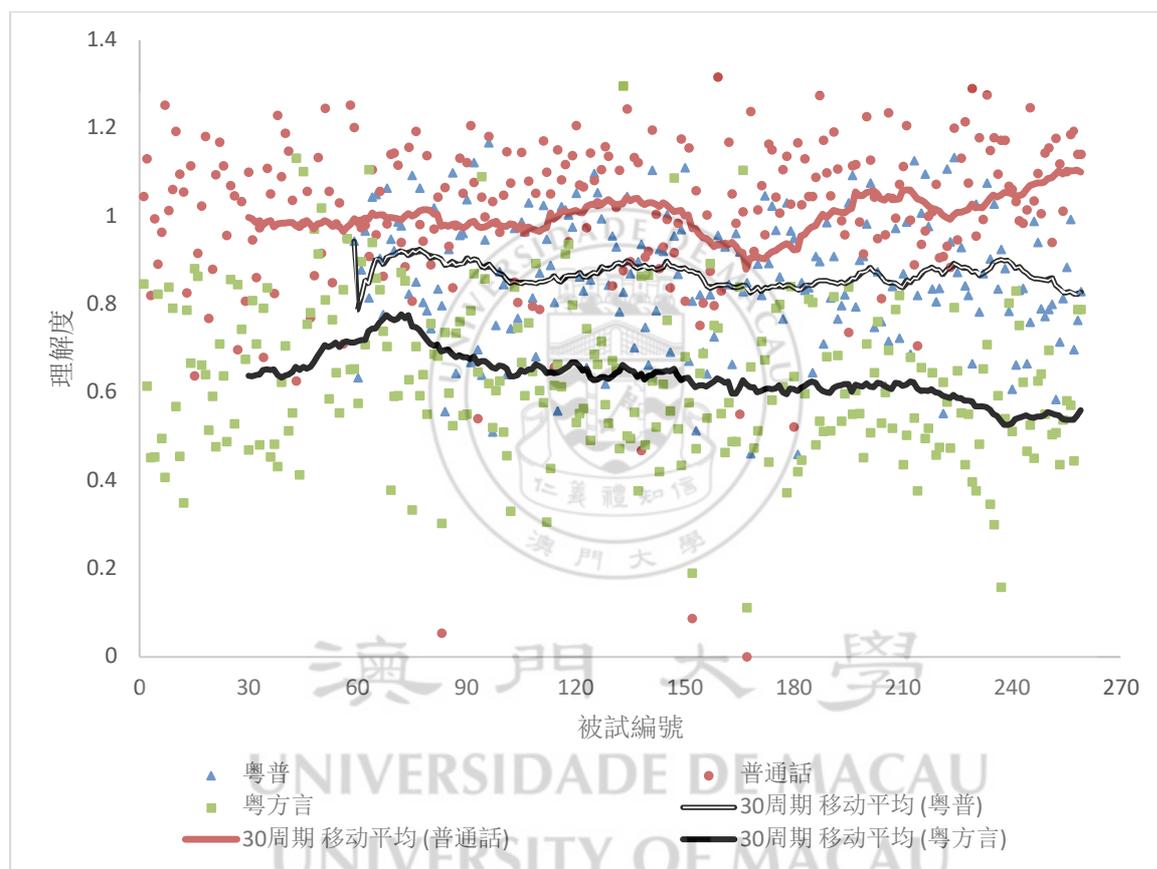
經過獨立樣本T檢驗驗證二者之間有著顯著性的差別（ $T=16.833$, $df=882$, $sig=0.00<0.05$ ）。雖然外籍聽話人已經取得了漢語普通話的高級水平，然而對於粵式普通話卻表現出了較低的水平，而相反，雖然被試身份背景不同，但是同為本族人保持著穩定的一致性，這種差異反映出在本族人頭腦中存在著一個相同的語言理解機制，這種理解機制是外族高級學習這暫時還不具有的。就是這種語言理解的一致性，有助於確定粵式普通話的性質。

（二）主觀判斷和客觀測評基本一致。本實驗採用主客觀結合的方法，將主觀測評結果與客觀評價結果結合；一方面通過主觀測評的結果印證輔證客觀測試的結果，另一方面是可以揭示主客觀的差距，這種差異在本族聽話者處比較明顯，原因是對於本族聽話者在聽本族語的時候，注意力會集中在聽不懂和不理解的地方，因此會將聽不懂，不理解的地方放大，從數量上忽略掉一部分聽得懂的內容。總體來說本實驗的主客觀評價是一致的，也說明客觀測評具有比較充分的心理現實性，基本滿足心理感受，具有一定的可信度。

（三）從被試的性別屬性來看，不同性別的被試對粵式普通話的理解度沒有差異，更加體現出一種粵式普通話在本族人內部的一種共享機制。可見性別對於粵式普通話的理解度並沒有太大的影響。從被試的年齡和教育背景來看，發現這些被試的屬性對粵式普通話均沒有顯著性影響。至於第一語言背景，本文認為會對粵式普通話產生一定的影響，不過由於本文被試所限，並不能夠將每個方言區的被試數量平均，這樣勢必會造成結果上的不平衡。故而本文未將這一問題深入討論。不過值得注意的是，從常識分析，第一語言是粵方言的被試應該對粵式普通話擁有更高的理解度，然而事實上並沒有出現這一結果，由此可見，粵式普通話的理解機制其實已經脫離了粵方言那種被粵方言母語者獨

有的模式，而轉入普通話的共享的理解機制中去。

(四) 根據本章第五節的實驗結果，計算出每名被試的粵方言理解度、普通話理解度和粵式普通話的理解度。以被試的序號為橫軸，以被試的測試的得分作為縱軸，描點得到一個散點圖。



圖表 30 本族人三種語言變體理解度散點圖

圖30中分別畫出每一個被試三種語言變體的理解度得分，並添加每一種語言變體的動態平均趨勢線。由圖可見普通話的理解度得分總體處於圖表的上部，集中在0.9到1之間；粵方言的理解度主要分散在圖表的下半部，主要分布在0.8以下，也有少許高於0.9的是屬於粵方言母語者的得分；粵式普通話的理解度主要分布在0.8到0.9之間。從趨勢線來看，明顯可以看出粵式普通話的趨勢線更加靠近普通話的整體趨勢。

本文採納徐傑、董思聰（2013）中的觀點，將普通話的標準放開，不以北京音為標準音，而是以北京音為基礎音的普通話，即所謂的「寬式普通話」。在這一標準下，將會吸納更多的語言材料進入研究視野，並且能夠將「體系行」、「穩定性」、「可理解性」作為寬式普通話的一項重要標準。在此定義下的普通話更有生命力，更鮮活也更加自然真實。追究普通話的命名之處就在於「普遍通行」，這種寬式普通話從概念的外延上看，包括涵蓋了前人研究中的地方普通話和動態普通話等概念範疇。因為從這種寬式普通話的產生途徑上看，是有方言背景的人在向普通話靠攏的過程中產生的。通過可理解度測試，可以制定出具體的標準。本文著重探討粵式普通話的可理解性的問題，從粵式普通話理解度的角度上，可以掌握粵式普通話的理解度特點。通過本次實驗，可以發現粵式普通話在理解度上最大的特點可以歸納為普遍性。一百餘名隨機挑選的本族被試經過實驗測試，所得到的理解度並不受性別，年齡，教育背景等因素所影響，體現出一種明顯的普遍性的規律。這種普遍性的規律僅僅是本族被試所具有的。與外族被試相比，外族被試對粵式普通話的理解度與本族被試相比有著顯著性差異，其理解度大大低於本族被試。顯然這種僅被本族人所擁有的一致理解度，與前文中對標準普通話的理解度測試結果相同。本族被試對普通話有一種「共享的理解能力」，所有被試對普通話的理解都在一個非常高的水平，也不受性別、年齡、教育背景等背景因素影響。反觀對粵方言的理解能力就處於一個相對低的水平，除去粵方言母語者，其他人都未能享有這種理解能力，並且這種理解能力還會隨著接觸時間的增加而增加。相比較粵式普通話的情況，粵式普通話與普通話的情況相同，雖然理解度略低於普通話的理解度，然而這種普遍性和共享性卻沒有變化，甚至連粵方言母語者對粵式普

通話的理解度也沒有與其他被試有顯著性差異。由於被試對粵式普通話有著與普通話相同的「共享的理解能力」，因此粵式普通話從理解度角度上具有顯著地普通話「普遍通行」的性質，所以可以將粵式普通話歸為普通話的範疇中，實際使用中確實應該注意到粵式普通話的理解度與普通話的理解度有差異，所以將其納入寬式普通話的範疇中。

4.6 小結

本章主要研究了粵方言、普通話以及粵式普通話的理解度，利用客觀實驗和主觀評價的方法對理解度進行了測試。一方面實際應用理解度理論和測試方法解決實際問題，另一方面也利用實驗結果驗證理解度理論和測試方法。在第一節中，對比了普通話水平測試與理解度測試之間的差異，主要體現在參照系統，測試內容和測試目的三個方面。同時還總結了理解度測試應該向普通話水平測試借鑒吸收的地方，主要體現在普通話水平測試利用了計算機輔助技術，形成一套專業的軟件系統，更加高效地進行普通話水平測試，另一方面是針對普通話水平測試的理論成果豐富，理論和實踐之間形成了良性互動。

在第二節中，本文討論了方言普通話的相關研究，對方言普通話的眾多相關研究中，目前對於判定方言普通話的性質問題依舊沒有得到非常好的解答。方言普通話究竟是屬於普通話，還是屬於方言，還是屬於單獨一類，各方都從理論的角度給出了相應的證明，都屬於定性研究的範疇，然而目前還沒有一種利用實驗的手段進行定量研究。因此本文也是希望通過理解度測試對粵式普通話進行分析，以粵式普通話為例，嘗試用定量的方式更好地總結方言普通話的性質。本節還介紹了粵式普通話的相關特點，前人總結偏重於詞彙和句法方面，而參考日常交流以及相關材料的過程中，體現粵式普通話特點的更多的並不是

詞彙和句法方面，而是語音系統。同樣的演講稿用粵式普通話讀出來就是帶有特殊的風格。因此這也確定了本文在對粵式普通話的研究方面，著重於語音部分。

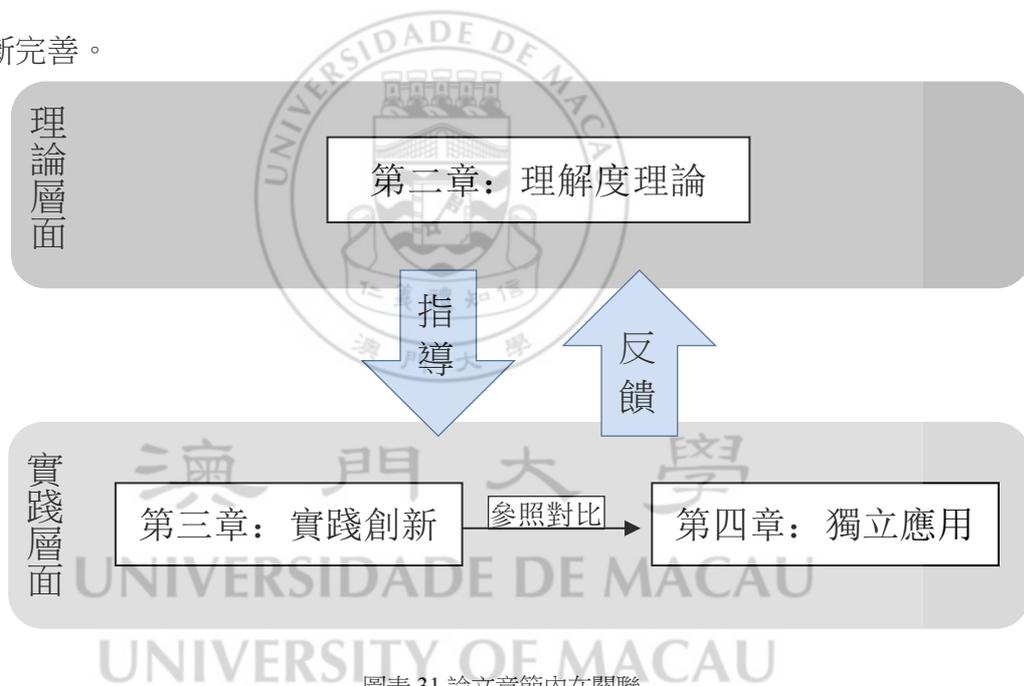
第三節提出了本章實驗的問題，經過後面幾節的實驗論述，可以對第三節的問題進行回答。第四節中對粵方言與普通話進行了理解度測試，共對278名被試進行測試。利用前文中使用的實驗方法與實驗流程，經統計計算得到內地非粵方言被試對粵方言的理解度是0.641，被試對普通話的理解度是0.952。所以相互的可理解度應是：0.797。二者之間的理解率是非粵方言被試聽粵方言的理解率是34.97%，粵方言母語者聽普通話的理解率是93.50%，因此相互之間的可理解率是64.23%。首先粵式普通話的理解度經過實驗驗證，本族聽話人對粵式普通話的平均理解度是0.858，標準差為0.355；可理解率是75.81%；外族聽話人對粵式普通話的理解度為0.595，標準差為0.492，可理解率是26.31%，兩組之間的理解度有顯著性差異，本族聽話人內部對於粵式普通話的理解度表現出較為一致的狀態。性別、教育背景、年齡等因素並沒有對粵式普通話的理解度產生顯著性影響。再次是將粵式普通話的理解度與第三章的數據進行比較分析，對於本族聽話人來說，粵式普通話的理解度更接近於熟悉的語言變體的理解度，換句話說，對於國人來說，相比較於對粵方言的理解度粵式普通話的理解度更接近於普通話的理解度，最後是主觀評判結果與客觀測評的結果關係總體上是一致的，這種一致性可以說明客觀的測評具有一定的心理現實性，能夠正確地反應出被試的心理狀態和實際的理解水平。總體上講，此次試驗達到了實驗預期目標，能夠將需要測試的結果正確測量出來，並且能夠根據結果得到符合實際情況的結論。

最後在此梳理本章研究的研究意義：本章對粵式普通話進行了理解度測試，首次使用理解度測試的方法對粵式普通話進行研究，為地方普通話研究提供了一種新的研究方法，也為粵式普通話增添了新的研究案例。前人研究中對於粵式普通話多集中於描寫性的研究，以及偏誤類型的研究，多是從理論層面對粵式普通話的成因、表現、特點進行詳細闡述，所得出的結論也多數是經驗總結和理論推導，並沒有使用定量研究進行心理語言學等相關感知實驗。因此本文在粵式普通話的研究方法方面提出新的思路，利用定量的方法對於方言普通話的性質進行研究，以定量的結果推導出定性的結論，用定量的方式給出適當的標準，可以更合理地為方言普通話的性質做出界定。其次是為理解度測試提供了新的材料和素材。利用理解度理論和測試對粵式普通話進行實驗，得到了真實且符合實際情況的數據，說明理解度理論和測試是一種有用的、有效的途徑。實驗的細節還需要在今後的研究中不斷完善，用更多的實驗案例來豐富理解度的數據，不斷充實和發展理解度理論。



第五章 結論

本文從理論和實踐的兩個角度詳細論述了理解度的相關問題，並對粵方言、湘方言和閩南方言進行相互理解度測試，此外還對粵式普通話進行理解度測試。本文第二章是理解度的理論核心，第三章是利用理解度理論指導理解度實踐的創新，第四章是獨立運用理解度測試解決實際問題，三者之間相互關聯。理論核心指導實踐的創新和應用，最後獨立應用能夠對理論進行反饋。三者層層遞進又互為首尾，理論和實踐的進步也是在這三者螺旋式上升的模式下不斷完善。



圖表 31 論文章節內在關聯

緒論部分對文章整體的研究對象，研究現狀、研究問題和研究意義進行了初步的討論。本文的研究對象主要是圍繞著理解度展開，分別是理解度、理解度測試、相互理解度以及方言普通話理解度。研究方法包括了理論層面上的推導、歸納和總結，還有實踐層面上的實驗和統計。

研究意義是圍繞著理解度、理解度測試和方言普通話理解度進行討論。在理解度理論方面，繼續擴展了理解度在語言學內的應用，將理解度理論更好地

與語言學研究適配；其次在漢語語言學研究範疇中首次對理解度理論本身進行詳細闡述，一方面搭建理解度研究的理論平台，方便與其他領域接軌；另一方面有利於理論自身的發展，今後可在現有理論的基礎上進一步拓展。最後，理解度理論為語言本體研究和語言應用研究之間搭起一座新的橋梁，因為語言的最基本應用就是讓人理解，讓人聽懂，所以理解度是所有語言應用的基礎。語言的變化在一定程度上會引起理解的差異，語言的本體落腳點應該是語言的應用，語言應用又為語言本體提供研究素材和現象，二者之間都可以借由理解度理論進行溝通。

在理解度測試方面，可以通過理解度測試更準確地描述語言之間的距離，利用理解度測試這把「尺子」衡量語言和語言之間的距離，方言和方言之間的距離，語言變體和語言變體之間的距離。理解度測試也是理解度理論的實踐延伸，二者作為太極的兩級不可分割，並且相互補益。加入時間參數對於理解度測試來說意義深遠，不僅是將理解度測試納入到心理語言學標準測試的模式中，並且能夠使得理解度測試的測試數據更加真實有效，全面適配語言學的精細化研究需要。

在方言普通話理解度方面，理解度測試為方言普通話的研究提供了一條新的思路和研究方法，填補了方言普通話的實驗測試研究的空白。將傳統的定性研究與定量研究相結合，更好地研究和探討方言普通話的性質和特點。

對理解度意義的討論一方面能夠正確認識到理解度理論對語言學研究和發展的重要性，另一方面也能夠為理解度理論提供有效的發展方向和研究使命。

理解度理論從上個世紀二十年代開始逐漸興盛，在各個領域相繼對理解度這一問題進行研究。在語言學研究領域中，心理語言學，社會語言學都對該問題有所涉及，但未見其形成一個完整獨立的理論。理解度理論的構建首先是要對理解度的概念進行界定。本文利用經典五分的通信系統模式來對信息傳播的過程進行描述。依照五分通信系統模式提出理解度的概念：語言信號由說話者說出經傳信系統到達聽話者的這一過程中，聽話者對原信號即時的理解程度，為單向理解度。在對話過程中，聽話者與講話者的身份交替轉變，雙方或多方輪流成為聽話者，對話參與者相互之間的理解度即為相互理解度。並對概念所蘊含的基本假設進行闡述以及概念中的關鍵要素進行詳細說明。

隨後對理解度性質的總結和歸納一方面可以深入瞭解認識理解度本質，另外更重要的是要利用理解度的特性去合理地解決實際問題。本文從即時性、聽覺性、單向性、客觀性、動態性、相對穩定性這六個角度對其進行描述。這六個基本屬性可以較為全面地描述理解度，從這六個角度去把握理解度理論可以更好地指導進一步的理論建設和實驗設計。

理解度研究的瓶頸問題就是在於影響理解度的因素具有種類多、層級多、範圍廣等幾大問題。理解度的變化是受到了眾多影響因素的共同作用，過多的因素阻礙了傳統方法對理解度研究的深入。本文認為影響理解度的眾多因素對理解度的影響是有層次的，並不是每個因素都對理解度產生相同量級的影響，因此對因素進行分級也是必須的。分級處理的方法一方面可以簡化操作，有側重地考慮每個因素的影響力；另一方面可以具體問題具體分析，不同層級不同對待，應對每個層級的影響因素的對策並不相同。試圖做到即不是一刀切，也不是一視同仁。

理解度的測試方法方面，本文先從理論層面來對測試方法進行討論，具體的技術性的測試內容會在後文具體討論。理解度測試理論主要包括實驗材料、實驗方法、實驗對象、和測試工具四個方面。由於理解度測試適用性非常廣，因此會利用理解度測試研究不同具體的問題。所以實驗中的各個要素差別就會非常大。本節的思路是將實驗中可能會出現的變項列出，通過理解度的性質對其進行甄別，留下符合要求的選項。實驗材料的部分，本文認為首選應使用句子作為實驗材料，其次可以使用詞語。不推薦使用字，語篇等單位作為實驗材料。實驗方法部分，本文認為應該將主觀判斷和客觀實驗結合起來，客觀實驗為主，主觀判斷為輔，二者之間相互印證，相互補充；其次使用選擇類的任務類型的方法比使用轉寫類的任務類型更能夠得到真實的數據，且實驗材料錄音只播放一次；最後需要記錄被試的反應時間，用來消除選擇類型任務的誤差。實驗對象的選擇方面盡量保證來源的廣泛性，杜絕利用主觀意識來選擇被試。測試工具的選擇也是本節的重點內容之一，本文分析了傳統測試和改良後的測試的利弊，雖然目前來看，傳統測試的經濟性，便捷性要大大勝過改良後的測試；改良後的測試在數據上的準確性和多樣性上是傳統測試不可比擬的。因此從長遠的角度來看，改良的方法代表著進步的方向，會逐漸成為主流，最終取代傳統的方法。而現階段，應該將二者有機結合起來，發揮二者的優勢，在實踐中不斷完善測試的方法。

本文認為根據理解度的性質和特點，應該使用常模參照標準。這一節首先對常模參照標準和標準參照標準進行了介紹，闡述了二者之間的差異，各自適用的測試背景。接著列舉選擇常模參照標準的原因：首先從理解度的性質出發，同時考慮到理解度是一種語言能力，最後論述了不選標準參照的理由。因

此理解度理論需要使用常模參照標準來設置理論的標準。

對理解度理論意義的討論，是本文的重要核心內容之一。本文從理解度理論的功能角度出發，探討理解度理論能夠為語言學貢獻怎樣的功能。首先是尺子的功能，理解度是作為衡量語言與語言之間距離的尺子，首先建立以理解度作為標準，接著是建立理解度的標準。標準對於一個學科的重要性不言而喻，理解度理論就是這樣的一個標準建立者，也是一個平台的搭建者，在這個平台上可以找到與其他理論的接口，共同協作解決一個問題。其次是理解度理論可以對已有研究的結論進行驗證，驗證的過程中又對原理論進行補充，可以極大促進學術的進步。並且理解度實踐中發現的問題和觀測到的現象，又為傳統理論創造更多可供研究的課題。最後是理解度理論可以提出更多更有意義的問題來引領學科的發展和學術的進步。在實踐中不斷創新，向未知的領域不斷探索。

一個理論一定不是十全十美包打天下的，理解度理論也有它本身的局限，至少是在目前階段屬於它自身天生的不可回避的缺陷。首先是對言外之意的理解，目前的理解度並不能夠處理這種問題的理解度，目前的理解度僅僅是對字面意思的理解，言外之意的理解又屬於更高一個層次的理解範疇，暫時不在本文的研究範圍內。其次是對相互理解度的實時監控，依靠目前的技術手段還不能支撐這一概念的實施。依照目前的手段，只能夠將相互理解度拆分成兩個單向的理解度，之後通過統計手段計算成相互理解度。最後是樣本數量的不足，若想真正建立語言之間的理解度標準，那麼需要對世界上大多數語言進行研究測試，而每一門語言就需要大量的被試來形成基礎性的數據。所以對於構建語言之間的理解度常模，還需要大量的樣本。

本文展望了理解度理論未來的發展方向。本文設想能夠進一步完善和改進人工智能技術，利用理解度理論得到的成果來幫助人工智能技術，測試人工智能的語言能力，優化人工智能的算法。並且希望能夠讓理解度理論與人工智能領域相結合，利用人工智能技術的智能性來模擬人類的理解度模式，從而通過人工智能獲得大量的模擬數據，這樣可以替代傳統測試的方法，極大地提高效率。除此之外，可以應用到機器翻譯的領域，理解度理論可以為機器翻譯改進優化算法，檢測機器翻譯的結果是否能夠被理解，機器理解人類的語言是否是真正的意思等等。最後是理解度理論需要建立在大數據的視角下來考慮問題，宏觀全面地認識問題，要站在上帝視角去審視理解度的問題，就是要充分意識到問題的複雜性，簡單機械的求簡最終會導致失真，承認問題的複雜性，但是並不畏懼複雜，而是用相應的技術和手段去處理複雜，從浩如煙海的數據中找到簡單真實的規律。

第三章是利用一個具體真實的理解度測試，進一步從實踐角度深入闡釋理解度測試，並且對傳統的理解度測試進行改進，在理解度測試中增加時間參數。並且在本章中首先用理論的方式論證時間參數的必要性，另外通過實踐的方式計算出時間參數，最終計算出理解度數值並得出結論。

本文探討了時間參數的概念，時間參數是利用被試的反應時間，經計算後得出的參數，用以參與理解度的計算。在理解度測試中，時間參數是由被試在測試中的反應時長經由一系列計算得來的。需要注意的是，此處原始數據中是反應時長，而並不是理解時間。之所以選擇反應時長而沒有選擇理解時間，本文給出了三條主要原因。第一是理解度實驗目前屬於行為實驗的層級和範疇，還沒有上升到神經實驗的層級。在行為實驗中，提取反應時間是通行做法；第

二是理解時間和反應時間雖然不同，但是具有等效性；最後是考慮到實驗條件的限制選擇了行為實驗而放棄了神經實驗，因此利用反應時間作為時間系數的基礎數據。

添加時間參數的原因本文認為在選擇類任務中，由於被試做出反饋的過程過於簡單，因此有可能出現並不是完全理解，或者乾脆不理解的情況就做出回答，就是通常所說的「蒙」。另外還有一種更加具有迷惑性的現象就是被試不是完全不懂，而是似懂非懂，是通過調動自己的語言知識輔之以聯想及各種提示將題目做對，而這種做對的方式僅僅是題目正確，卻不是理解的結果，而是大腦一些列複雜分析運算的結果。換句話說，本文所謂的理解是「秒懂」，而經過聯想、回憶等一系列複雜運算之後的懂與真正的「秒懂」本質上是有區別的。因為在正常的語流中，只有「秒懂」才能夠更順暢的交流，否則雙方或者一方要需要對每句話每個詞都進行深入分析後才能夠交流恐怕這樣的對話很難進行下去。所以為了區分上述情況，就有必要引入時間參數來進入運算。尤其是在處理兩種語言變體沒有語系上的差異，而是分述不同方言支系這類的情況的時候，引入時間參數的必要性愈加凸顯。並且恰恰是這種在一個語言內部兩種相近的語言變體的理解度往往是爭論的關鍵點，因此更需要引入時間參數來對理解度測試做出精細化地區分和考察。

首先之所以在之前的理解度實踐中沒有引用時間參數是鑒於前人研究的目的和實驗精度並不需要多引入一個維度的數據就可以得到想要的結果，而在語言學的研究中，本文認為時間參數的引入可以對語言研究有更精細化的分析結果，因為多引入了一個維度的數據，就可以將以往簡單的對錯二元對立擴展成二維的矩陣，擴展了對理解度測試的解釋空間。其次是利用時間參數去修正在

實際測試中可能會產生的誤差，從而得到更符合事實的結果。最後是為了將理解度測試納入到科學的測試中，充分借鑒心理學測試的通常範式即採集反應時間這一做法，使得理解度測試本身更具有科學性，同時更好地建立模式化，規範化的理解度測試。

本文利用前文所考察了三個方言之間的相互理解度與相互理解率。這三個方言分別是粵方言，湘方言與閩南方言，此實驗共調查被試122名，其中32名是以粵方言為第一語言，45名是以湘方言為第一語言，45名以閩南方言作為第一語言。女性被試共有66人，佔全部人數的54.1%，男性共有58人，佔全部人數的45.9%。被試的主要群體是在校學生，其中有15名為高中生，均是來自於閩南方言區。本實驗的被試在徵集過程中主要要求是：1、必須是本地人；2、被試的第一語言和家庭語言；3、被試不會其他方言，且沒有長期在外地生活學習。所有被試均是有償招募，實驗者向所有實驗參與者支付相關報酬或等值的禮品以示感謝。

實驗分別選擇了24個詞和24個句子作為實驗材料，實驗共涉及到三個方言分別是：粵方言、湘方言和閩南方言。實驗材料的產生方式是在現代漢語基本常用詞彙表中簡單隨機抽樣得到的48個詞，然後將24個詞造句。最終得到24個句子及24個詞。實驗材料共有三個變量，除了詞彙和句子之外，還分為三種方言和正向與反向題目。三個變量中，除方言變量有三個水平外，其餘兩個變量均有兩個水平。最終共有12種組合，每種組合都包含四個材料。本實驗分別在澳門、長沙和廈門三座城市進行實驗。

本實驗通過E-Prime軟件，主要收集了正確值和反應時間兩個量，因此共

收集了 $122 \times 48 \times 2 = 11712$ 個有效數據，後經過IBM SPSS Statistics進行數據分析。經分析正確值與反應時間之間的分析，通過相關性檢驗，發現正確值與反應時間之間存在著明顯的負相關。正確值越高，反應時間越短，具體表現在三個方言區內，各個方言來看，本地被試對本地方言的正確值和反應時間均與其他兩組數據表現出明顯的顯著性差異。

隨後本文利用反應時間經過三個步驟計算出時間參數，首先是將被試的真實反應時間通過歸一化處理轉換為在 $[0, 1]$ 區間內的數字，這樣做是為了消除不同被試之間反應時間的實際差異，並且能夠把數字控制在固定的區間內，方便之後的運算和比較。第二步是分類，因為不同實驗材料所需要的反應時間不同，因此本文根據語言材料性質的不同區分了十二個小類。最後一步是參照運算，在這十二個小類中，統計出本地人對該類語言材料的反應時間均值，並以此為參照標準，用該類中所有的歸一化時間除以本地人對該類材料的反應時間均值，得到最終的時間參數。最後利用時間參數與正確值相乘，就得到了該名被試在此題目上的理解度值。

經統計得到，粵方言人聽湘方言的理解度是0.557；粵方言人聽閩南方言的理解度是0.530；湘方言人聽粵方言的理解度是0.509；湘方言人聽閩南方言的理解度是0.455；閩南方言人聽粵方言的理解度是0.482；閩南方言人聽湘方言的理解度是0.496。隨後再將兩個單向的理解度數值合併成一個相互理解度的數值，得到粵方言與湘方言之間的相互可理解度為0.533，粵方言與閩南方言之間的相互可理解度為0.506，湘方言與閩南方言之間的相互可理解度為0.476。

本文將得到的理解度數值與前人研究中的溝通度數值進行比較，前人研究中的溝通度是通過語音或詞彙中的相似度進行計算得出的，並非通過心理實驗得到。二者之間的研究方法不盡相同，不能直接進行數據上的比對。整體來看，二者的數據基本都能夠反映出理解度的不平衡性，其結果也基本處於一個大致範圍內。

因為理解度數值並不是通常意義上的百分數，其數值造成理解上的不便，所以本文首次引入理解率的概念，在理解度的基礎上進行進一步的歸一化運算，得到一個相對的理解率，這個理解率就可以作為百分數來進行運算和比較。那麼最終得到粵方言和湘方言之間的相互可理解率是14.68%，粵方言與閩南方言之間的相互可理解度是9.54%，湘方言與閩南方言的相互理解度是3.82%。第三章的最後對理解度數值的意義和價值進行了簡要的討論。本文認為，理解度的數值正如尺子的讀數。該讀數有兩層含義，第一層屬於比較基礎的淺顯的含義，讀數是多少就代表是多少。根據讀數，可以對其進行比較。例如本實驗中內地非粵方言被試對普通話的理解度就是高於內地非粵方言被試對粵方言的理解度。這一層的含義根據目前的結論是可以得到的。深層的含義則是通過這個讀數對被測事物進行解釋說明甚至歸類。例如身高不超過85釐米的馬就是矮腳馬，身高160釐米的馬就是正常的馬。通過數字對被測物體進行推導的過程隱含著對該物體的類屬有一個常識性的認識，即在140釐米到180釐米的馬就是正常的馬，超出這個範圍的馬就是其他品種。目前理解度數值沒有得到這樣的常模，暫時不能夠對其進行推導。若想達到這個層次，就需要進行越來越多的理解度測試，將現實世界事物的屬性與數字聯繫起來，建立起理解度常模，從而就可以通過理解度數值對被測的語言進行判斷和歸類了。

第四章主要研究了粵式普通話的理解度，利用客觀實驗和主觀評價的方法對理解度進行了測試。一方面實際應用理解度理論和測試方法解決實際問題，另一方面也利用實驗結果驗證理解度理論和測試方法。在第一節中，對比了普通話水平測試與理解度測試之間的差異，主要體現在參照系統，測試內容和測試目的三個方面。同時還總結了理解度測試應該向普通話水平測試借鑒吸收的地方，主要體現在普通話水平測試利用了計算機輔助技術，形成一套專業的軟件系統，更加高效地進行普通話水平測試，另一方面是針對普通話水平測試的理論成果豐富，理論和實踐之間形成了良性互動。

針對方言普通話的眾多相關研究中，目前對於判定方言普通話的性質問題依舊沒有得到非常好的解答。方言普通話究竟是屬於普通話，還是屬於方言，還是屬於單獨一類，各方都從理論的角度給出了相應的證明，都屬於定性研究的範疇，然而目前還沒有一種利用實驗的手段進行定量研究。因此本文也是希望通過理解度測試對粵式普通話進行分析，以粵式普通話為例，嘗試用定量的方式更好地總結地方普通話的性質。本節還介紹了粵式普通話的相關特點，前人總結偏重於詞彙和句法方面，而在參考日常交流以及相關材料的過程中，體現粵式普通話特點的更多的並不是詞彙和句法方面，而是語音系統。同樣的演講稿用粵式普通話讀出來就是帶有特殊的風格。因此這也確定了本文在對粵式普通話的研究方面，著重於語音部分。

對粵方言與普通話進行了理解度測試，共對278名被試進行測試。利用前文中使用的實驗方法與實驗流程，經統計計算得到內地非粵方言被試對粵方言的理解度是0.641，被試對普通話的理解度是0.952。所以相互的可理解度應是0.797。二者之間的理解率是非粵方言被試聽粵方言的理解率是34.97%，粵方

言母語者聽普通話的理解率是93.50%，因此相互之間的可理解率是64.23%。首先粵式普通話的理解度經過實驗驗證，本族聽話人對粵式普通話的平均理解度是0.858，標準差為0.355；可理解率是75.81%；外族聽話人對粵式普通話的理解度為0.595，標準差為0.492，可理解率是26.31%，兩組之間的理解度有顯著性差異，本族聽話人內部對於粵式普通話的理解度表現出較為一致的狀態。性別、教育背景、年齡等因素並沒有對粵式普通話的理解度產生顯著性影響。再次是將粵式普通話的理解度與前文的數據進行比較分析，對於本族聽話人來說，粵式普通話的理解度更接近於熟悉的語言變體的理解度，換句話說，對於本族人來說，相比較於對粵方言的理解度粵式普通話的理解度更接近於普通話的理解度，最後是主觀評判結果與客觀測評的結果關係總體上是一致的，這種一致性可以說明客觀的測評具有一定的心理現實性，能夠正確地反應出被試的心理狀態和實際的理解水平。總體上講，此次試驗達到了實驗預期目標，能夠將需要測試的結果正確測量出來，並且能夠根據結果得到符合實際情況的結論。

經過實驗驗證，本文認為粵式普通話的理解度特點與符合普通話理解度特點一致，具有強烈的一致性，這種一致性僅僅體現在本族聽話人的大腦中，在與外族聽話人的理解度相比，可以看出他們的理解度與本族聽話人的理解度有著顯著性差異。從理解度的數值上看也更接近普通話的理解度，且與粵方言的理解度相差較遠，因此本實驗支持將粵式普通話定性為一種「寬式普通話」的結論。

羅馬非一日建成，理解度理論也並不能一步到位，還需要對其進行不斷的加工升級，從而更好地發揮理解度理論的功效。從目前來看，理解度理論還需

要分近期目標與長期目標分階段分層次研究。

近期的研究目標可以縱向深入針對某一語言或語言變體進行深入研究，盡可能包含更多的變量，計算出變量與變量之間的影响因子。例如語音、詞彙、句法三個層面究竟對理解度的影響是怎樣的，具體到語音中，音段音位和超音段音位對理解度的影響是怎樣的等等問題，都可以依照實驗目的修改實驗材料，對其進行有針對性的測試。需要注意的是在做這類的測試時盡量保證實驗規程的統一，這樣得出的結果才能夠具有可比性，否則只能是零散的數據，不能夠得到足夠的信息。另外可以橫向對其他語言或語言變體進行測試，拓展研究的範圍。

同時通過越累越多的實驗累積，進一步驗證理解度測試的有效性，最終確定理解度測試的詳細規程，為進一步量表化、模塊化不斷改進，最終能夠實現理解度實驗全過程的自動化，程式化。同時利用理解度測試的結果豐富參照標準中的常模，通過擴大樣本的數量對理解度數值有更進一步地認識。

長期的研究目標是建立起完整的理解度標準體系，測試體系。可以利用理解度數值對兩種語言或語言變體之間的關係進行判斷，實現根據定量的數值得到定性的結論。明確各個變量與理解度之間的影响變化關係，從而可以形成一套完整的理解度參數，最終可以根據語言本體的數據直接推算出理解度，這樣可以達到從理論上和實踐上的統一。與神經語言學相聯繫，深入探討理解的機制和機理，能夠從神經層面挖掘語言和語言之間理解度差異的機制。

本文對理解度理論進行了較為詳細的論述，理解度理論在語言學世界中還屬於一個較為小眾的研究領域。這也說明尚有很多問題還沒能得到很好地探

究，同時也為研究者提供了更多研究的空間。理解是語言的第一要義，如果一個語言不能夠被理解，那麼這個語言則毫無意義，理解對於語言來說太過重要，卻又被輕易地忽視。面對理解度的理論其實才剛剛開始，還有太多的問題沒能解釋，太多的現象沒有發現。隨著研究的深入才發現研究的問題越多，遇到的問題也就越多。希望可以隨著今後研究的深入，可以將更多的未知轉化為已知。面對無窮的未知世界，本文僅僅是走到了未知的邊緣，向前探出一小步，在探出這一小步的過程中，回過身才發覺已知的世界是如此的渺小。



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

附錄

附錄1：被試信息表

序號	性別	省份	第一語言	年齡	年級
1-1	女	江蘇	官話	26	博三
1-2	女	安徽	官話	27	博二
1-3	男	銀川	普通話	20	大二
1-4	女	廣東	粵方言	19	大二
1-5	女	東北	普通話	24	研一
1-6	男	湖南	普通話	20	大三
1-7	男	甘肅	官話	20	研二
1-8	女	廣東	粵方言	21	研二
1-9	女	順德	粵方言	21	大三
1-10	女	新疆	普通話	19	大一
1-11	女	湖南	湘方言	29	博三
1-12	男	山西	官話	22	大四
1-13	女	山東	普通話	19	大一
1-14	男	河南	官話	30	博三
1-15	男	四川	官話	26	研二
1-16	女	山西	普通話	26	研二
1-17	女	江蘇	官話	24	研二
1-18	女	吉林	官話	20	大二
1-19	男	山東	官話	26	博二
1-20	男	湖南	官話	21	大三
1-21	男	遼寧	普通話	20	大二
1-22	女	廣東	粵方言	24	大三
1-23	女	河南	官話	25	博三
1-24	女	貴州	官話	21	大三
1-25	女	江蘇	吳方言	23	研二
1-26	女	四川	普通話	19	大一
1-27	男	山東	官話	21	大三
1-28	女	陝西	普通話	19	大一
1-29	女	山西	普通話	21	大三
1-30	女	陝西	普通話	19	大一
1-31	男	江西	贛方言	22	大四
1-32	男	北京	普通話	21	大三
1-33	女	廣東	粵方言	48	博二
1-34	女	湖北	官話	23	大一
1-35	女	江蘇	吳方言	28	博二
1-36	女	內蒙古	外族語	23	研二
1-37	女	廣東	客家話	21	大三
1-38	女	陝西	普通話	23	研二
1-39	男	四川	官話	23	研二
1-40	女	四川	官話	22	大四

1-41	女	四川	官話	19	大一
1-42	男	湖南	湘方言	22	大四
1-43	女	甘肅	普通話	19	大一
1-44	女	河北	普通話	20	大二
1-45	女	天津	普通話	24	研三
1-46	女	河南	普通話	23	研二
1-47	男	湖北	官話	23	研二
1-48	女	廣東	粵方言	23	研二
1-49	男	湖北	普通話	19	大一
1-50	男	福建	普通話	19	大一
1-51	女	廈門	普通話	19	大一
1-52	男	江蘇	吳方言	22	大四
1-53	男	北京	普通話	22	大四
1-54	女	宜昌	普通話	24	研二
1-55	男	湖南	湘方言	24	研二
1-56	女	陝西	普通話	24	研二
1-57	男	雲南	官話	22	大四
2-1	男	泰國	外族語	21	
2-2	女	泰國	外族語	20	
2-3	男	贊比亞	外族語	22	
2-4	女	泰國	外族語	22	
2-5	女	泰國	外族語	20	
2-6	男	馬達加斯加	外族語	20	
2-7	女	泰國	外族語	22	
2-8	男	泰國	外族語	24	
2-9	女	泰國	外族語	20	
2-10	男	泰國	外族語	21	
2-11	女	泰國	外族語	21	
2-12	女	薩摩亞	外族語	23	
2-13	女	泰國	外族語	20	
2-14	女	泰國	外族語	21	
2-15	女	泰國	外族語	20	
2-16	女	泰國	外族語	21	
2-17	女	泰國	外族語	21	
2-18	女	泰國	外族語	20	
2-19	男	美國	外族語	23	
3-1	女	江蘇	普通話	23	大四
3-2	女	天津	官話	22	大四
3-3	女	天津	普通話	24	研一
3-4	女	山東	官話	25	研一
3-5	女	河北	普通話	23	研一
3-6	女	山西	官話	25	研一
3-7	女	陝西	官話	23	研一
3-8	女	山西	普通話	24	研一
3-9	女	遼寧	官話	24	研一

3-10	女	天津	普通話	34	研一
3-11	女	河南	官話	23	研一
3-12	女	河南	官話	22	大四
3-13	女	山東	官話	24	研一
3-14	女	黑龍江	官話	26	研一
3-15	女	天津	官話	24	研一
3-16	男	吉林	官話	36	研二
3-17	女	甘肅	官話	25	研一
3-18	女	河北	官話	24	研一
3-19	男	天津	普通話	24	研一
3-20	女	河南	官話	24	研一
3-21	女	陝西	官話	23	研一
3-22	女	吉林	普通話	23	研一
3-23	女	山東	官話	24	研一
3-24	女	江西	贛方言	23	研一
3-25	男	內蒙	官話	52	研究生
3-26	男	天津	普通話	27	研究生
3-27	女	甘肅	官話	21	大四
3-28	女	新疆	普通話	22	大四
3-29	女	廣西	粵方言	22	大四
3-30	女	天津	官話	22	大四
3-31	女	四川	官話	24	大四
3-32	女	天津	官話	22	大四
3-33	男	雲南	官話	22	大三
3-34	男	天津	普通話	22	大三
3-35	女	甘肅	官話	20	大三
3-36	女	甘肅	官話	22	大三
3-37	女	雲南	官話	21	大三
3-38	女	天津	官話	20	大三
3-39	女	甘肅	官話	21	大三
3-40	女	天津	普通話	21	大三
3-41	女	甘肅	官話	21	大三
3-42	女	天津	普通話	21	大三
3-43	女	新疆	普通話	20	大三
3-44	女	甘肅	官話	21	大三
3-45	女	甘肅	普通話	21	大三
3-46	女	廣西	粵方言	22	大三
3-47	女	甘肅	官話	22	大三
3-48	女	天津	官話	20	大三
3-49	女	天津	普通話	21	大三
3-50	女	北京	普通話	21	大三
3-51	女	天津	官話	21	大三
3-52	女	內蒙	官話	22	大三
3-53	女	天津	官話	21	大三
3-54	女	天津	普通話	22	大三

3-55	女	内蒙	普通話	22	大三
3-56	男	贵州	普通話	21	大三
3-57	男	天津	普通話	19	大一
3-58	女	福建	閩方言	19	大一
3-59	女	广西	客家話	19	大一
3-60	男	甘肃	普通話	18	大一
3-61	男	山西	普通話	18	大一
3-62	女	四川	普通話	18	大一
3-63	女	湖北	普通話	19	大一
3-64	男	浙江	普通話	18	大一
3-65	男	甘肃	普通話	18	大一
3-66	男	安徽	普通話	18	大一
3-67	女	甘肃	普通話	19	大一
3-68	女	湖南	湘方言	19	大一
3-69	男	新疆	普通話	19	大一
3-70	男	江西	贛方言	19	大一
3-71	女	天津	普通話	18	大一
3-72	女	陝西	官話	19	大一
3-73	女	贵州	官話	18	大一
3-74	女	江西	普通話	20	大二
3-75	女	山東	官話	19	大二
3-76	女	山東	官話	20	大二
3-77	女	河南	官話	19	大二
3-78	女	陝西	官話	19	大二
3-79	女	山西	官話	19	大二
3-80	女	甘肃	官話	21	大二
3-81	女	山西	官話	20	大二
3-82	女	安徽	吳方言	20	大二
3-83	女	陝西	官話	20	大二
3-84	女	湖南	湘方言	21	大二
3-85	男	西藏	外族語	21	大二
3-86	男	河北	官話	20	大二
3-87	女	重庆	官話	19	大二
3-88	男	山東	官話	24	研一
3-89	女	湖南	湘方言	23	研一
3-90	男	河南	官話	25	研一
3-91	女	天津	官話	20	大一
3-92	女	贵州	官話	20	大一
3-93	女	海南	閩方言	19	大一
3-94	男	西藏	外族語	18	大一
3-95	男	福建	閩方言	19	大一
3-96	男	天津	普通話	19	大一
3-97	男	天津	官話	19	大一
3-98	女	甘肃	官話	19	大一
3-99	女	天津	官話	19	大一

3-100	女	河北	官話	18	大一
3-101	女	湖南	湘方言	18	大一
3-102	女	江西	湘方言	20	大一
3-103	女	广西	客家話	19	大一
3-104	女	天津	官話	21	大三
3-105	女	天津	官話	21	大三
3-106	女	贵州	官話	21	大三
3-107	女	四川	官話	21	大三
3-108	女	天津	普通話	21	大三
3-109	女	吉林	官話	21	大三
3-110	女	天津	官話	21	大三
3-111	女	河北	普通話	20	大三
3-112	女	新疆	普通話	21	大三
3-113	女	天津	官話	20	大三
3-114	女	广西	官話	22	大三
3-115	女	内蒙	普通話	21	大三
3-116	女	天津	普通話	21	大三
3-117	男	吉林	官話	22	大三
3-118	女	河南	官話	21	大三
3-119	女	吉林	官話	21	大三
3-120	女	天津	普通話	21	大三
3-121	女	天津	官話	22	大三
3-122	女	天津	官話	22	大三
3-123	女	甘肃	普通話	21	大三
3-124	女	云南	官話	21	大三
3-125	女	云南	官話	21	大三
3-126	女	天津	官話	21	大三
3-127	女	新疆	普通話	21	大三
3-128	女	天津	官話	21	大三
3-129	男	河北	普通話	20	大三
3-130	女	天津	普通話	20	大三
3-131	女	新疆	普通話	21	大三
3-132	女	天津	普通話	21	大三
3-133	男	浙江	普通話	25	博一
3-134	男	山東	普通話	19	大一
3-135	男	湖北	官話	25	博二
3-136	男	吉林	普通話	26	研三
3-137	男	河南	官話	20	大二
3-138	男	重庆	官話	20	大二
3-139	男	山東	普通話	20	大二
3-140	男	山西	普通話	25	本科
3-141	男	河南	官話	26	博一
3-142	男	江西	普通話	25	研二
3-143	男	台湾	閩方言	25	研二
3-144	男	浙江	普通話	25	研二

3-145	男	浙江	吳方言	25	研二
3-146	男	江西	客家話	27	博一
3-147	男	河南	官話	25	研二
3-148	男	新疆	普通話	23	大四
3-149	男	北京	普通話	28	博五
3-150	男	江西	贛方言	25	研一
3-151	男	北京	普通話	24	大五
3-152	男	陝西	普通話	21	大三
3-153	男	河北	官話	19	大一
3-154	男	廣東	普通話	19	大一
3-155	男	新疆	普通話	19	大二
3-156	男	北京	普通話	18	大一
3-157	男	湖北	官話	23	研二
3-158	男	貴州	官話	21	大四
3-159	男	福建	普通話	30	博三
3-160	男	甘肅	官話	21	大三
3-161	男	陝西	普通話	26	研二
3-162	男	雲南	官話	23	研一
3-163	男	湖北	普通話	17	大一
3-164	男	河南	官話	22	大四
3-165	男	遼寧	官話	33	博四
3-166	男	北京	普通話	22	大三
3-167	男	山東	官話	28	研一
3-168	女	廣東	粵方言	25	研究生
4-1	女	河南	普通話	15	高一
4-2	女	河北	普通話	16	高一
4-3	男	河北	普通話	15	高一
4-4	男	天津	官話	16	高一
4-5	男	天津	外族語	16	高一
4-6	男	北京	普通話	15	高一
4-7	女	山東	官話	17	高三
4-8	男	天津	官話	17	高三
4-9	男	湖北	普通話	17	高三
4-10	女	天津	普通話	17	高三
4-11	男	江蘇	吳方言	18	高三
4-12	男	河北	普通話	17	高三
4-13	男	北京	普通話	17	高三
4-14	男	浙江	普通話	18	高三
4-15	男	河北	普通話	17	高三
4-16	男	天津	官話	17	高三
4-17	女	河北	普通話	17	高三
4-18	女	天津	官話	18	高三
4-19	男	北京	普通話	18	高三
4-20	女	山東	普通話	18	高三
4-21	女	河北	普通話	18	高三

4-22	男	河北	普通話	17	高二
4-23	男	天津	普通話	17	高二
4-24	男	天津	普通話	17	高二
4-25	女	天津	官話	17	高二
4-26	女	山東	普通話	17	高二
4-27	男	天津	官話	17	高二
4-28	男	山東	官話	17	高二
4-29	男	天津	官話	17	高二
4-30	女	天津	普通話	16	高二
4-31	男	天津	官話	17	高二
4-32	男	天津	普通話	17	高二
4-33	男	天津	普通話	16	高二
4-34	男	天津	官話	17	高二
5-1	男	廈門	閩南方言	15	高一
5-2	男	廈門	閩南方言	15	高一
5-3	男	廈門	閩南方言	15	高一
5-4	男	廈門	閩南方言	15	高一
5-5	男	廈門	閩南方言	15	高一
5-6	男	廈門	閩南方言	15	高一
5-7	男	廈門	閩南方言	16	高一
5-8	男	廈門	閩南方言	15	高一
5-9	男	廈門	閩南方言	15	高一
5-10	男	廈門	閩南方言	15	高一
5-11	男	廈門	閩南方言	15	高一
5-12	男	廈門	閩南方言	15	高一
5-13	女	廈門	閩南方言	15	高一
5-14	女	廈門	閩南方言	15	高一
5-15	女	廈門	閩南方言	15	高一
5-16	男	廈門	閩南方言	21	大四
5-17	女	廈門	閩南方言	20	大三
5-18	男	廈門	閩南方言	20	大二
5-19	男	廈門	閩南方言	20	大二
5-20	男	廈門	閩南方言	20	大二
5-21	男	廈門	閩南方言	20	大二
5-22	女	廈門	閩南方言	20	大三
5-23	女	廈門	閩南方言	18	大一
5-24	女	廈門	閩南方言	18	大一
5-25	女	廈門	閩南方言	20	大三
5-26	男	廈門	閩南方言	19	大二
5-27	女	廈門	閩南方言	22	大四
5-28	女	廈門	閩南方言	21	大四
5-29	女	廈門	閩南方言	21	大三
5-30	女	廈門	閩南方言	21	大三
5-31	女	廈門	閩南方言	18	大一
5-32	女	廈門	閩南方言	20	大三

5-33	女	廈門	閩南方言	20	大三
5-34	女	廈門	閩南方言	18	大一
5-35	女	廈門	閩南方言	21	大三
5-36	男	廈門	閩南方言	18	大一
5-37	男	廈門	閩南方言	18	大一
5-38	女	廈門	閩南方言	20	大二
5-39	女	廈門	閩南方言	19	大二
5-40	女	廈門	閩南方言	21	大三
5-41	女	廈門	閩南方言	20	大二
5-42	女	廈門	閩南方言	19	大二
5-43	女	廈門	閩南方言	19	大二
5-44	女	廈門	閩南方言	18	大一
5-45	女	廈門	閩南方言	21	大三
5-46	男	长沙	湘方言	20	大三
5-47	女	长沙	湘方言	18	大一
5-48	男	长沙	湘方言	18	大二
5-49	男	长沙	湘方言	19	大三
5-50	女	长沙	湘方言	19	大三
5-51	女	长沙	湘方言	19	大三
5-52	女	长沙	湘方言	19	大三
5-53	男	长沙	湘方言	20	大三
5-54	女	长沙	湘方言	19	大三
5-55	男	长沙	湘方言	21	大四
5-56	男	长沙	湘方言	25	本科
5-57	男	长沙	湘方言	20	大三
5-58	女	长沙	湘方言	19	大三
5-59	男	长沙	湘方言	19	大二
5-60	女	长沙	湘方言	20	大三
5-61	女	长沙	湘方言	19	大三
5-62	男	长沙	湘方言	20	大三
5-63	女	长沙	湘方言	20	大三
5-64	女	长沙	湘方言	19	大三
5-65	男	长沙	湘方言	21	大三
5-66	女	长沙	湘方言	20	大三
5-67	女	长沙	湘方言	20	大三
5-68	女	长沙	湘方言	20	大三
5-69	女	长沙	湘方言	19	大二
5-70	女	长沙	湘方言	20	大二
5-71	女	长沙	湘方言	19	大二
5-72	男	长沙	湘方言	19	大三
5-73	男	长沙	湘方言	20	大三
5-74	女	长沙	湘方言	20	大三
5-75	女	长沙	湘方言	20	大三
5-76	女	长沙	湘方言	22	大四
5-77	男	长沙	湘方言	21	大四

5-78	男	长沙	湘方言	20	大三
5-79	男	长沙	湘方言	19	大二
5-80	男	长沙	湘方言	18	大一
5-81	女	长沙	湘方言	19	大二
5-82	女	长沙	湘方言	18	大一
5-83	女	长沙	湘方言	18	大一
5-84	女	长沙	湘方言	17	大一
5-85	男	长沙	湘方言	19	大二
5-86	女	长沙	湘方言	18	大一
5-87	男	长沙	湘方言	19	大一
5-88	男	长沙	湘方言	19	大二
5-89	男	长沙	湘方言	19	大二
5-90	男	长沙	湘方言	19	大二
5-91	男	澳門	粵方言	19	大一
5-92	女	澳門	粵方言	26	博四
5-93	女	東莞	粵方言	23	研一
5-94	女	珠海	粵方言	25	大四
5-95	男	澳門	粵方言	30	本科
5-96	男	珠海	粵方言	20	大二
5-97	男	澳門	粵方言	21	大二
5-98	男	澳門	粵方言	20	大一
5-99	男	澳門	粵方言	18	大一
5-100	男	澳門	粵方言	18	大一
5-101	男	澳門	粵方言	18	大一
5-102	女	澳門	粵方言	18	大一
5-103	女	澳門	粵方言	18	大一
5-104	男	澳門	粵方言	23	大四
5-105	女	澳門	粵方言	22	大四
5-106	女	澳門	粵方言	18	大一
5-107	女	澳門	粵方言	19	大一
5-108	男	澳門	粵方言	19	大一
5-109	女	澳門	粵方言	21	大四
5-110	男	澳門	粵方言	19	大二
5-111	女	澳門	粵方言	20	大一
5-112	男	澳門	粵方言	18	大一
5-113	男	澳門	粵方言	18	大一
5-114	男	澳門	粵方言	19	大一
5-115	男	澳門	粵方言	22	大三
5-116	女	澳門	粵方言	21	大三
5-117	女	澳門	粵方言	20	大三
5-118	女	澳門	粵方言	19	大一
5-119	女	澳門	粵方言	21	大三
5-120	女	澳門	粵方言	18	大一
5-121	女	澳門	粵方言	18	大一
5-122	女	澳門	粵方言	18	大一

附錄2：粵方言、湘方言閩南方言理解度實驗材料信息表

实验例句	音频时长	方言	正向/反向
他想去洗手间。	1649	閩南方言	正向
这个是水杯。	1424	閩南方言	正向
这间教室灯光很明亮。	2501	閩南方言	正向
今年天气异常燥热。	2435	閩南方言	反向
他现在在游泳。	1629	閩南方言	反向
我想买新衣服。	1801	閩南方言	反向
那是一个传销组织。	2293	湘方言	正向
你这样太不象话了。	2361	湘方言	正向
商店里的地毯很柔软。	2454	湘方言	正向
他想看祖國的美丽河山。	2857	湘方言	正向
这是我的手表。	1578	湘方言	正向
看台上面空无一人。	2440	湘方言	反向
你不许看电视。	1789	湘方言	反向
他是全家的希望。	1816	湘方言	反向
小张学习很出色。	2382	湘方言	反向
这里房价上涨很多。	2488	湘方言	反向
海浪吞没了小岛。	2083	粵方言	正向
明天我要去买水果。	1997	粵方言	正向
难得一见这璀璨的星空。	2751	粵方言	正向
这里的蚊子很烦人。	1228	粵方言	正向
飘落了一地的花瓣。	2215	粵方言	反向
身上的伤口溃烂了	2328	粵方言	反向
他很担心你的安全。	1939	粵方言	反向
这个公园有些空旷。	2070	粵方言	反向

實驗例句	音頻時長	方言	正向/反向
山羊	1237	閩南方言	正向
能力	907	閩南方言	正向
緊張	1199	閩南方言	正向
混亂	1081	閩南方言	正向
夏天	1137	閩南方言	反向
拖鞋	1153	閩南方言	反向
生活	1157	閩南方言	反向
船头	1121	閩南方言	反向
着急	792	湘方言	正向
挤压	795	湘方言	正向
存折	807	湘方言	正向
鈔票	810	湘方言	正向
眼鏡	778	湘方言	反向
微博	736	湘方言	反向
撫卹	804	湘方言	反向

慘狀	888	湘方言	反向
游客	789	粵方言	正向
探望	807	粵方言	正向
抱枕	1182	粵方言	正向
白菜	1077	粵方言	正向
沙啞	948	粵方言	反向
漆黑	378	粵方言	反向
俯視	972	粵方言	反向
愛情	1042	粵方言	反向

附錄3：普通話粵方言理解度實驗材料信息表

實驗例句	音頻時長	圖像／文字	粵方言／普通話	正向／反向
這回把他憋壞了。	1850	圖像	普通話	正向
他想看祖國的美麗河山。	2006	圖像	普通話	正向
他現在在游蝶泳。	1782	圖像	普通話	反向
他的業務水平過硬。	1865	圖像	普通話	反向
你這樣太不象話了。	1673	圖像	普通話	正向
你要好好唱戲。	1487	圖像	普通話	反向
這只股票今天暴跌。	1880	圖像	粵方言	正向
這個是水杯。	1248	圖像	粵方言	正向
蚊子很煩人。	1227	圖像	粵方言	正向
身上的傷口潰爛了	2328	圖像	粵方言	反向
看台上面空無一人。	2136	圖像	粵方言	反向
今年天氣異常燥熱。	2098	圖像	粵方言	反向
這是我的手錶。	1449	文字	普通話	正向
這個公園有些空曠。	1868	文字	普通話	反向
他是全家的希望。	2011	文字	普通話	反向
商店裡的地毯很柔軟。	2009	文字	普通話	正向
那是一個傳銷組織。	1851	文字	普通話	正向
家裡容不得你造反。	1744	文字	普通話	反向
這間教室燈光很明亮。	2195	文字	粵方言	反向
他很擔心你的安全。	1746	文字	粵方言	反向
飄落了一地的花瓣。	2093	文字	粵方言	反向
難得一見這璀璨的星空。	2669	文字	粵方言	正向
教育需要人文價值導向。	2728	文字	粵方言	正向
海浪吞沒了小島。	2083	文字	粵方言	正向

實驗用詞	音頻時長	圖像／文字	粵方言／普通話	正向／反向
夏天	1207	圖像	普通話	正向
凌亂	1026	圖像	普通話	正向
緊張	927	圖像	普通話	反向

存折	853	圖像	普通話	正向
鈔票	1009	圖像	普通話	反向
博客	900	圖像	普通話	反向
游客	788	圖像	粵方言	正向
沙啞	1248	圖像	粵方言	反向
駕到	890	圖像	粵方言	反向
水果	769	圖像	粵方言	正向
抱枕	1181	圖像	粵方言	正向
爱情	1041	圖像	粵方言	反向
着急	987	文字	普通話	正向
拖沓	867	文字	普通話	反向
殺傷	948	文字	普通話	反向
擠壓	1003	文字	普通話	正向
撫卹	964	文字	普通話	正向
慘狀	1004	文字	普通話	反向
探望	806	文字	粵方言	正向
順眼	1023	文字	粵方言	正向
漆黑	578	文字	粵方言	反向
羚羊	1100	文字	粵方言	正向
俯視	971	文字	粵方言	反向
船头	993	文字	粵方言	反向



澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

參考文獻

- Angel de la Torre, Á., Peinado, A. M., Rubio, A. J., Sánchez, V. E., & Diaz, J. E. (1996). An application of minimum classification error to feature space transformations for speech recognition. *Speech Communication*, 20(3-4), 273-290.
- Alderete, J., Chan, Q., Chan, M., Fan, G., & Nickel, O. Cantonese grammar synopsis.
- Austin, J. L. (1962). *How to do things with words*. Oxford: The Clarendon Press.
- Bansal, R. K. (1969). *The Intelligibility of Indian English: Measurements of the Intelligibility of Connected Speech, and Sentence and Word Material, Presented to Listeners of Different Nationalities*. Ph. D. dissertation. Hyderabad: Central Institute of English.
- Beukelman, D. R., & Yorkston, K. M. (1979). The relationship between information transfer and speech intelligibility of dysarthric speakers. *Journal of communication disorders*, 12(3), 189-196.
- Beukelman, D. R., & Yorkston, K. M. (1980). Influence of passage familiarity on intelligibility estimates of dysarthric speech. *Journal of Communication Disorders*, 13(1), 33-41.
- Blair, F. (1990). *Survey on a shoestring: A manual for small-scale language surveys* (Vol. 96). Summer Inst of Linguistics.
- Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of verbal learning and verbal behavior*, 11(6), 717-726.
- Brennan, E. M., & Brennan, J. S. (1981). Measurements of accent and attitude toward Mexican-American speech. *Journal of Psycholinguistic Research*, 10(5), 487-501.
- Bunton, K. E. (2000). An acoustic item analysis of speech intelligibility. University of Wisconsin-Madison PhD dissertation.
- Casad, E. H. 1974. Dialect intelligibility testing. *Summer Institute of Linguistics: University of Oklahoma, Norman*.
- Comrie, B. (Ed.). (2009). *The world's major languages*. Routledge.
- Crystal, D. (2011). *Dictionary of linguistics and phonetics* (Vol. 30). John Wiley & Sons.
- Catford, J. C. (1950). Intelligibility. *ELT Journal*, (1), 7-15.
- Egan, J. P. (1948). Articulation testing methods. *The Laryngoscope*, 58(9), 955-991.
- Fletcher, H., & Steinberg, J. C. (1929). Articulation testing methods. *Bell Labs Technical Journal*, 8(4), 806-854.
- Fuse, A., Navichkova, Y., & Alloggio, K. (2017). Perception of Intelligibility and Qualities of Non-Native Accented Speakers. *Journal of Communication Disorders*.
- Gooskens, C., & Schneider, C. (2016). Testing mutual intelligibility between closely related languages in an oral society.
- Gooskens, C., van Heuven, V. J., Golubović, J., Schüppert, A., Swarte, F., & Voigt, S. (2018). Mutual intelligibility between closely related languages in Europe. *International Journal of Multilingualism*, 15(2), 169-193.
- Hardman, J. (2010). *The intelligibility of Chinese-accented English to international and American students at a U.S. university*. (Electronic Thesis or Dissertation). Retrieved from <https://etd.ohiolink.edu/>
- Hayes-Harb, R., & Watzinger-Tharp, J. (2012). Accent, Intelligibility, and the Role

- of the Listener: Perceptions of English-Accented German by Native German Speakers. *Foreign Language Annals*, 45(2), 260-282.
- Huong, T. T. (2017). The Intelligibility of the Vietnamese Accented English. *Education and Linguistics Research*, 3(1), 69.
- Kluge, A. (2006). RTT retelling method: An alternative approach to intelligibility testing. *Occasional paper. Dallas: SIL International*.
- Ljubesic, N., Mikelic, N., & Boras, D. (2007, June). Language identification: How to distinguish similar languages?. In *Information Technology Interfaces, 2007. ITI 2007. 29th International Conference on* (pp. 541-546). IEEE.
- Melinger, A. (2018). Distinguishing languages from dialects: A litmus test using the picture-word interference task. *Cognition*, 172, 73-88.
- Metz, D.E., Schiavetti, N., & Sitler, R. (1980). "Toward an objective description of the dependent and independent variables associated with intelligibility assessments of hearing impaired adults." In J. Subtelny (ed.), *Speech Assessment and Speech Improvement for the Hearing Impaired*, 72-81. Washington, DC: A.G. Bell Association for the Deaf.
- Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language learning*, 45(1), 73-97.
- Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in second language acquisition*, 28(1), 111-131.
- Nejjari, W., Gerritsen, M., Van der Haagen, M., & Korzilius, H. (2012). Responses to Dutch-accented English. *World Englishes*, 31(2), 248-267.
- Nelson, C. L. (2012). *Intelligibility in world Englishes: Theory and application*. Routledge.
- Pickering, L. (2006). Current research on intelligibility in English as a lingua franca. *Annual Review of Applied Linguistics*, 26, 219-233.
- PressKent, R. D. (Ed.). (1992). *Intelligibility in Speech Disorders: Theory, Measurement, and Management* (Vol. 1). John Benjamins Publishing.
- Sáfár, A., Meurant, L., Haesenne, T., Nauta, E., De Weerd, D., & Ormel, E. (2015). Mutual intelligibility among the sign languages of Belgium and the Netherlands. *Linguistics*, 53(2), 353-374.
- Schiavetti, N., & Kent, R. D. (1992). Scaling procedures for the measurement of speech intelligibility. *Intelligibility in speech disorders, Intelligibility in Speech Disorders: Theory, Measurement, and Management* (Vol. 1). John Benjamins Publishing. 11-34
- Schüppert, A., & Gooskens, C. (2011). Investigating the role of language attitudes for perception abilities using reaction time. *Dialectologia: revista electrònica*, 119-140.
- Schüppert, A., & Gooskens, C. (2012). The role of extra-linguistic factors in receptive bilingualism: Evidence from Danish and Swedish pre-schoolers. *International Journal of Bilingualism*, 16(3), 332-347.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, 27(3), 379-423.
- Stevens, S. S., Miller, J., Egan, J. P., Waterman, T. H., & Rome, S. C. (1942). *Articulation Testing Methods. OSRD report*, (383).
- Swann, J., Deumert, A., Lillis, T., & Mesthrie, R. (2004). *A dictionary of sociolinguistics*. Edinburgh University Press.
- Swarte, F. H. E. (2016). *Predicting the Mutual Intelligibility of Germanic languages from linguistic and extra-linguistic factors*. Rijksuniversiteit Groningen..

- Szeto, C. (2000). Testing intelligibility among Sinitic dialects. In *Proceedings of ALS2K, the 2000 conference of the Australian Linguistic Society* (pp. 1-10).
- Tang, C., & van Heuven, V. J. J. P. (2008). Mutual intelligibility of Chinese dialects tested functionally.
- Tang, C., & Van Heuven, V. J. (2009). Mutual intelligibility of Chinese dialects experimentally tested. *Lingua*, 119(5), 709-732.
- Tang, C. (2009). *Mutual intelligibility of Chinese dialects: an experimental approach*. LOT, Utrecht.
- Van Kuyk, S., Kleijn, W. B., & Hendriks, R. C. (2017). An instrumental intelligibility metric based on information theory. *IEEE Signal Processing Letters*.
- Varonis, E. M., & Gass, S. 1982. The Comprehensibility of Non-Native Speech. *Studies in Second Language Acquisition*, 4(2), 114-136.
- Voegelin, C. F., & Harris, Z. S. (1951). Methods for determining intelligibility among dialects of natural languages. *Proceedings of the American Philosophical Society*, 95(3), 322-329.
- Wardrip-Fruin, C., & Constantinou, A. (1993). Intelligibility of Mandarin speakers of English: Correlation of acoustic and perceptual measures. *The Journal of the Acoustical Society of America*, 94(3), 1866-1866.
- Weismer, G. 2008. 35 Speech Intelligibility. *The handbook of clinical linguistics*, 568.
- Whaley, L. J., Grenoble, L. A., & Li, F. (1999). Revisiting Tungusic classification from the bottom up: a comparison of Evenki and Oroqen. *Language*, 286-321.
- Yang, H. T. (2005). *Measuring the Systemic Mutual Intelligibility of Five English Speakers: A Socio-Phonological Analysis* (Doctoral dissertation, University of Illinois at Urbana-Champaign).
- Yorkston, K. M., & Beukelman, D. R. (1981). Communication efficiency of dysarthric speakers as measured by sentence intelligibility and speaking rate. *Journal of Speech and Hearing Disorders*, 46(3), 296-301.
- 安妮·安娜斯塔西 蘇珊娜·厄比娜 1997 《心理測驗》，繆小春、竺培梁譯，2011，杭州：浙江教育出版社。
- 曹曉燕 2012 《方言和普通話的語音接觸研究》，博士研究生學位論文，蘇州：蘇州大學。
- 陳 蒙 2009 《濟南普通話語音研究》，碩士研究生學位論文，山東大學。
- 陳海倫 1996 論方言相關度、相似度、溝通度指標問題，《中國語文》(5)：361-368。
- 陳海燕 2001 淺析海南腔普通話在詞彙方面應該規範的問題，《海南師範學院學報(人文社會科學版)》(6)：102-105。
- 陳建民 陳章太 1988 從我國語言實際出發研究社會語言學，《語文建設》(4)：58。
- 陳建偉 2008 《臨沂方言和普通話的接觸研究》，博士研究生學位論文，蘇州：蘇州大學。
- 陳雪清 程佳佳 劉博 劉莎 莫玲燕 孔穎 劉海紅 元貝爾 董瑞娟 李玉玲 王碩 吳燕君 2012 植入年齡對嬰幼兒人工耳蝸植入患者早期言語可懂度的影響，《中國耳鼻咽喉頭頸外科》(10)：529-531。

- 陳亞川 1987 閩南口音普通話說略，《語言教學與研究》(4)：148-155。
- 陳亞川 1991 「地方普通話」的性質特徵及其他，《世界漢語教學》(1)：12-17。
- 陳章太 1997 論普通話水平測試等級標準，《語言文字應用》(3)：12-17。
- 陳遵平 2012 略論地方普通話研究的幾個問題，《遵義師範學院學報》(2)：47-51。
- 董思聰 徐杰 2015 普通話區域變體的特點與普通話差錯的分際，《語言科學》14(6)：598-606。
- 方瑞 趙玉婷 陳小潔 2017 機場等公共交通系統英語播報可理解度研究——以成都平原為例，《英語廣場》(3)：30-32。
- 傅靈 2010 《方言與普通話的接觸研究》，博士研究生學位論文，蘇州：蘇州大學。
- 傅雨賢 1984 《粵方言區人學習普通話趣談》，廣州：中山大學出版社。
- 高超 文秋芳 2012 中國語境中本土化英語的可理解度與可接受度研究，《外語教學》(5)：53-58。
- 高潔 姚勇 2013 維吾爾學生漢語聲調的可理解度實驗研究，《新疆職業大學學報》(3)：60-62。
- 顧欽 2007 《語言接觸對上海市區方言語音演變的影響》，博士研究生學位論文，上海：上海師範大學。
- 顧英華 2003 民族學生漢語口語語音偏誤分析及教學對策，《新疆大學學報(哲學社會科學版)》(S1)：67-68。
- 桂詩春 1991 《實驗心理語言學綱要》，長沙：湖南教育出版社。
- 桂詩春 2011 《什麼是心理語言學》，上海：上海外語教育出版社。
- 郭熙 2001 對新時「推普」的一些思考：以江蘇為例，《南京大學學報(哲學·人文科學·社會科學版)》(2)：117-123。
- 鄭永輝 1996 粵方言影響下的韶關市城區普通話詞彙的特點，《韶關大學學報(社會科學版)》(1)：37-43。
- 韓軍寧 陳雪清 郭倩倩 孟超 尤妍顏 2010 助聽後耳聾兒童音樂能力與聽覺能力、言語可懂度相關性的初步研究，《聽力學及言語疾病雜誌》(1)：58-61。
- 韓維新 2010 粵方言常用入聲字的普通話聲調習得偏誤研究，《中國語言學會語音學分會、中國聲學學會語言、音樂和聽覺專業委員會、中國中文信息學會語音信息專業委員會·第九屆中國語音學學術會議論文集》，中國語言學會語音學分會、中國聲學學會語言、音樂和聽覺專業委員會、中國中文信息學會語音信息專業委員會：中國語言學會語音學分會。
- 韓玉華 2007 從認知語言學的角度淺析香港人普通話中的詞彙偏誤問題，《語言文字應用》(S1)：79-83。
- 何家勇 李珊 何雲亭 2010 EFL英語語音可理解度的影響因素研究，《成都理工大學學報社會科學版》，(4)：102-106。
- 胡珊 毛善平 2006 失語患者的句子理解實驗研究，《卒中與神經疾病》(6)：354-357。
- 黃銳 2012 《標準參照語言測試研究》，廈門：廈門大學出版社。
- 教育部、國家語委 2016 《國家語言文字事業「十三五」發展規劃》。

- 勁松 牛芳 2010 長沙地方普通話固化研究——地方普通話固化的個案調查，《語言文字應用》(4)，41-49。
- 雷峻 2010 計算機輔助普通話測試的問題思考及技術對策，《武漢理工大學學報》(13)：160-163。
- 林素娥 蘇憲萍 2007 方言可懂度計量研究芻議，《甘肅聯合大學學報(社會科學版)》(1)：58-60。
- 李愛軍 2003 漢語普通話和地方普通話的對比研究，《中國語言學會語音學分會、中國中文信息學會語音信息專業委員會、中國聲學學會語言聽覺和音樂分會.第六屆全國現代語音學學術會議論文集(上)》，中國語言學會語音學分會、中國中文信息學會語音信息專業委員會、中國聲學學會語言聽覺和音樂分會:中國中文信息學會。
- 李景娜 王遙 2015 英漢本族語者對中國英語學習者的口音感知及言語理解的對比研究，《外語教學理論與實踐》(1)：50-54+96。
- 李寧 孟子厚 李蕾 2011 山西普通話的語音特徵，《清華大學學報(自然科學版)》(9):1298-1302。
- 李如龍 1988 論方言和普通話之間的過渡語，《福建師範大學學報社會科學版》(2)：62-70。
- 李宇明 2003 論母語，《世界漢語教學》(1)：48-58。
- 魯健驥 1984 中介語理論與外國人學習漢語的語音偏誤分析，《語言教學與研究》(3)：44-56。
- 陸致極 1987 漢語方言間親疏關係的計量描寫，《中國社會科學》(1)：155-166。
- 聶丹 2011 關於普通話水平測試研究走向的思考，《語言文字應用》(2)：92-99。
- 牛芳 2009 《長沙塑料普通話研究——地方普通話固化的個案研究》，博士研究生學位論文，北京：中國人民大學。
- 彭康洲 2013 《基於標準參照語言測試理論的聽理解效度研究》，北京：中國社會科學出版社。
- 普通話普及情況調查項目組 孫曼均 李衛紅 姚喜雙 魏暉 2011 河北省普通話普及情況調查分析，《語言文字應用》(4)：2-12。
- 冉啓斌 索倫·維希曼 2018 怎樣區分語言與方言——基於核心詞彙的距離計算方法探索，《語言戰略研究》(2)：50-58。
- 石林 2010 侗族方言土語間理解度調查，《貴州民族學院學報哲學社會科學版》(4)：32-36。
- 石定栩 邵敬敏 朱志瑜 2006 《港式中文與標準中文的比較》，香港：香港教育圖書公司。
- 石定栩 2006 《港式中文兩面睇》，香港：星島出版有限公司。
- 侍建國 卓瓊妍 2013 關於國家語言的新思考，《語言教學與研究》(1)：80-88。
- 田皓 2006 地方普通話常見詞語失誤分析——以湖南腔普通話為例，《湖南人文科技學院學報》(4):85-87+130。
- 鄒遠春 2015 三十年來「地方普通話」研究綜述，《現代語文:下旬語言研究》(6)：4-6。
- 許長安 1991 大田普通話的普及和偏誤，《語文建設》(7):21-24。

- 王佶旻 2011 《語言測試概論》，北京：北京語言大學出版社
- 王磊 張穎 2007 普通話中介語負遷移及應對策略，《山西師大學報（社會科學版）》（3）：136-139。
- 王玲玲 2007 中介語理論與普通話測試——「目標參照」在普通話測試中的運用，《北華大學學報（社會科學版）》（6）：83-87。
- 王璐 張吉生 2014 吳語互通度與編輯距離之間的關係《語言研究》（2）：65-69。
- 王文虎 1994 四川口音普通話的語音特徵，《四川大學學報：哲學社會科學版》（3）：56-61。
- 王亞同 李文崗 2011 《心理語言學研究》，鄭州：河南大學出版社。
- 王詠梅 2006 龍岩新羅區「地方普通話」的語音特點，《龍岩學院學報》（1）：121-125。
- 魏梅 2017 中國英文報刊中本土化英語的國際可理解度研究，《西安外國語大學學報》（3）：63-68。
- 徐紅梅 2003 從雙語並用看帶方言味普通話，《韶關學院學報（社會科學版）》（2）：108-113。
- 徐杰 董思聰 2013 漢民族共同語的語音標準應微調為「以北京語音為基礎音」，《語言科學》（5）：460-468。
- 亞爾肯·阿吉 許輝 木合塔爾·霍加 2016 齟裂語音理解度評估方法研究進展，《中華實用診斷與治療雜誌》（6）：525-527。
- 楊曉 2011 「言外之意」的理解，《文學界(理論版)》（6）：119-120。
- 姚喜雙 2010 推普工作的重要抓手——談依法推進的普通話水平測試，《語言文字應用》（3）：26-34。
- 姚佑椿 1988 上海口音的普通話說略，《語言教學與研究》（4）：120-128。
- 姚佑椿 1989 應該開展對「地方普通話」的研究，《語文建設》（3）：20-23。
- 葉軍 王頤嘉 周鳴之 喬麗華 2007 計算機輔助普通話水平測試實施研究，《語言文字應用》（S1）：142-148。
- 葉軍 2012 上海地方普通話口音特徵語項動態分析，《語言文字應用》（S1）：130-138。
- 銀沛月 2010 成都口音普通話「去聲+去聲」雙音節詞重音研究之簡介，《中國語言學會語音學分會、中國聲學學會語言、音樂和聽覺專業委員會、中國中文信息學會語音信息專業委員會·第九屆中國語音學學術會議論文集》，中國語言學會語音學分會、中國聲學學會語言、音樂和聽覺專業委員會、中國中文信息學會語音信息專業委員會：中國語言學會語音學分會。
- 游汝傑 楊蓓 1998 粵滬普詞彙接近率研究，鄒嘉彥 主編《漢語計量與計算研究》，香港：香港城市大學語言信息中心。
- 于芳 2009 普通話訓練與測試所要面對的「地方普通話」問題，《武夷學院學報》（3）：57-60。
- 袁海軍 1994 言語可懂度加權平均聽力損失的計算方法及其臨床意義，《聽力學及言語疾病雜誌》（2）：88-89。
- 曾子凡 2000 「港式普通話」剖析，《方言》（3）：278-283。
- 詹伯慧 2003 當前「推普」的一些思考，《學術研究》（12）：148-150。

- 張積家 張鳳玲 2014 熟練粵方言-普通話雙言者聽覺詞的語言表徵，《心理與行為研究》(4)：433-440+468。
- 張建強 2005 「地方普通話」研究芻議，《廣西社會科學》(7)：159-161。
- 張建強 2007 漢語中介語研究的理論思考，《現代語文》(12)：11-13。
- 張建強 2008 鋪門地方普通話語言特徵分析，《賀州學院學報》(2)：74-76。
- 張林軍 2015 聽話者和說話者母語異同及背景噪音對二語學習者漢語言語理解的影響，《語言教學與研究》(4)：1-7。
- 張伶俐 2014 國外可理解性研究述評，《中國外語教育》(7)：47-53。
- 張 雷 侯傑泰 何偉傑 文劍冰 王渝光 2001 普通話測試的錄音評分可行性、信度及經濟效率，《心理學報》33(2)：97-103。
- 鄭錦全 1988 漢語方言親疏關係的計量研究，《中國語文》(4)：87-102。
- 鄭錦全 1994 漢語方言溝通度的計算，《中國語文》(1)：35-43。
- 仲哲明 1997 普通話水平測試若干問題的討論，《語言文字應用》(3)：5-11。
- 祝新華 陳瑞端 溫紅博 2012 十年間香港大學生普通話水平的發展變化——基於香港理工大學的十年考試數據，《語言文字應用》(S1)：58-66。
- 莊小榮 2012 廣東大學生英語語音可理解度的影響因素研究，《南昌高專學報》(2)：58-60。
- 周潤年 1990 析金華「地方普通話」，《浙江師大學報》(3)：110-115。