Linguistik

**Lu Wang**

# Linguistic Distance and Mutual Intelligibility Among Five Wu Dialects

Shaker Verlag
Aachen 2015

**Bibliographic information published by the Deutsche Nationalbibliothek**
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Copyright Shaker Verlag 2015
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publishers.

Printed in Germany.

ISBN 978-3-8440-3790-6
ISSN 1613-4532

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen
Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9
Internet: www.shaker.de • e-mail: info@shaker.de

**This book is dedicated to my daughter Yuhua Yang.**

# Acknowledgements

This book could not be completed withoud the help of many people. The author would like to thank Zhang Jisheng, Carlos Gussenhoven, Haike Jacobs, Jereon van de Weijer, Chen zhongmin, Zhu Yiner, Yang Wei, Li Susu, Wang huafu for their help with this book.

# 内容摘要

汉语不同方言间几乎不能互通，这是众所周知的现象，甚至同一方言内部间的互通度也很低。汉语不同方言间极低的互通度是汉语方言的一大特点。影响方言间互通度的因素很多，就语言本身而言，包括语音、音系、词汇、句法以及文化等因素；就语言外的因素而言，可能涉及人的文化背景、社会背景、身体和心理状况、年龄与性别，以及地理环境等因素。但最重要的因素还是语言间的不同的音系（Gooskens, 2007）。

吴方言是仅次于北方方言（普通话）的中国第二大方言。无论是吴方言学家或是吴语为母语的人都有这样的共识：吴语内部的互通性很低，尤其是南北吴语很难互通，而南部吴语间差别很大，互通度很低。本研究选择了吴方言内部五个方言（吴江，黄岩，永康，衢州，温州）点，通过听力实验的方法测量吴方言内部五个方言间的互通度，并通过计算方言间的语言距离（linguistic distance），分析解释吴方言内部互通度与语言距离之间的关系。本文提出如下研究假设：吴语间的互通度与各方言间的同源词之间的语言距离密切相关。

语言学家们对于互通度的研究可以追溯至上世纪五十年代。Voegelin 和 Harris（1951）提出了“确定自然语言间互通度的方法”，开创了关于测试方言间互通度研究的先河。他们提出的测量方言互通度的两种方法，至今仍被广泛地用于测量各语体间的互通度，即 “询问受试者的方法”和 “测试受试者的方法”。van Heuven（2008）曾通过“测试受试者的方法测量了真正的互通度”。我们的研究目的也是为了要量化吴语各方言间的互通度，因此本研究采用了听力实验的方法来测试受试者，以此来测试吴语内部方言间的互通度。

本研究主要分两大部分：互通度测试和语言距离计算。互通度实验是在两个层面上进行的，包括 30 个三音节的词和 20 个句子。分别用五个不同吴方言录制测试语料，到五个方言点对五种不同方言的语料进行听力测试，每个方言点选取 15 位符合要求的受试人。测试采用听写的方法，通过取所有参加实验的被试者的平均值，以及两个层面互通度测试结果的平均值，我们得出了一个相对合理的

数据作为方言间的相对互通度值，从而得出了这五个方言间的互通程度，其结果为（“>”表示“互通度大于”）：

黄岩-衢州, 黄岩-吴江 > 黄岩-永康, 吴江-衢州, 衢州-永康 > 吴江-永康, 温州-吴江, 黄岩-温州, 温州-永康, 温州-衢州。

那么基于其中任意一个方言的测试语料被其他四个方言的受试者所听懂的比例，我们得出了五个方言的被听懂的难易度，即黄岩, 吴江>衢州, 永康>温州（“>”表示“更易被听懂”）。这一结论与我们平常对于这些方言的难易度的印象基本一致，即温州话最难懂，吴江话好懂。而黄岩话因为兼具了南北吴语的特点，因此也同样较容易被南北吴语区听懂。

在互通度值的计算上，我们还重新检验了调整后的听懂度（Pierce, 1952），结果表明，原始的听懂度分值与调整后的听懂度分值并无统计学意义上的区别。

关于语言距离的测量方法，语言学家们也尝试了各种方法，主要可分为感知语言距离和客观语言距离两种方法。感知语言距离的测试是依靠受试者的感觉，而客观的语言距离的测量方法是基于语言本身的差异。Kessler于1995年第一次将编辑距离（edit distance）作为测量爱尔兰方言间的语言距离。编辑距离指的是字符串A转化为字符串B所需的最少编辑数。那么相应地应用到语言学中, 一个语言变体的一串语音表达可以相应地对应到另一个语言变体的一串语音表达。语音表达的对应通常是基于词的音节结构, 元音对应元音, 辅音对应辅音。那么基于任何一个同源词的不同方言的语音表达间的编辑距离，不同方言间的语言距离就可以被计算出来了。有不少研究将编辑距离应用于测量西日尔曼语，斯堪的纳维亚语言等拼音语言。编辑距离被证明是研究西方拼音语言之间的语言距离的有效方法。

基于编辑距离的方法，我们以五个方言的 30 个三音节词和 20 个句子的语音表达作为计算五个方言间的语言距离的语料。我们计算出了五个方言所构成的 10 个方言对之间的语言距离。此外，基于句子的语料，我们还提出对编辑距离的两种应用方法，被称为方法 1 和方法 2。方法 1 是将每句话的所有测试点看作是一个字符串。而方法 2 是将每句话中的每个测试点看作是一个字符串，两个方

言间的基于句子的编辑距离就是所有测试点的编辑距离的平均值。通过分析，我们发现方法 1 更为有效，因为这种方法将句子的语境也考虑在内，因此与互通度的相关更高。通过对 10 组方言对之间语言距离的计算，它们相互间语言距离的排序如下（“<”表示“语言距离小于”）：

黄岩-衢州，黄岩-吴江 < 黄岩-永康，黄岩-温州，温州-永康, 吴江-衢州 < 温州-吴江, 衢州-永康，温州-衢州, 吴江-永康。

我们把与互通度可能相关的各种客观距离，包括语言距离，结构距离，地理距离，旅行距离，与互通度值做了统计学上的相关分析。结果发现唯一与互通度值相关的是语言距离，但并没有发现其他的客观距离与互通度存在有意义的关联。

本研究发现通过编辑距离的计算得出的语音距离是可以预测五个吴方言互通度的最好的预测因子。研究结论证明，通过同源词的音段编辑距离测出的语言距离与五个方言间的互通度有密切相关性：语言距离越大，其互通度越小；反之，语言距离越小，其互通度越大。

**关键词：** 吴方言，互通度，语言距离，编辑距离

# Abstract

It is well known that different Chinese dialects are not mutually intelligible. Even within one dialect family, e.g., the Wu dialect, mutual intelligibility between different Wu subgroups is very low. There are many factors that may affect mutual intelligibility between dialects. The linguistic factors may be related to phonetics, phonology, vocabulary, syntax, etc. The extra-linguistic factors may be related to speakers' education, socio-cultural background, age, gender, psychological condition, geography, contact, etc. However, phonological differences always have the greatest impact on mutual intelligibility between dialects or closely related languages, according to Gooskens (2007).

According to the Language Atlas of China (1987), the Wu dialect family is the second largest dialect family in China, second to the Northern dialect family (Mandarin). There is a general impression that different Wu dialects are not mutually intelligible, particularly between Southern Wu and Northern Wu, and even within Southern Wu. The present study aims to investigate mutual intelligibility among five Wu dialects, including Wujiang, Huangyan, Yongkang, Quzhou and Wenzhou (representing the five Wu dialect groups) and to calculate the linguistic distances among the five Wu dialects, in order to explain the possible reasons for the different degrees of mutually intelligibility. The study is based on the research hypothesis that the different degrees of mutual intelligibility between different Wu dialects correlate with the different linguistic distances between each pair of dialects.

The study of mutual intelligibility can be traced back to the 1950s. Voegelin and Harris (1951) first proposed a methodology for determining intelligibility among dialects of natural languages. The two methods they proposed, the 'ask the informant' method and the 'test the informant' method, had great impact on the study of intelligibility. These two methods have been applied, with refinements, to intelligibility tests between dialects of languages ever since.

Van Heuven (2008) proposed that 'functional intelligibility tests measure the real thing'. Since the purpose of the present study is to measure 'the real thing', the method of testing the informant is applied in the experiment in which the subjects are exposed to another dialect (or their own), and are required to write down what they think they have heard.

Testing mutual intelligibility and measuring linguistic distances among the five Wu dialects are two major tasks of the present study. Our mutual intelligibility experiment is conducted at both the word and the sentence level. Assessment of the degree of intelligibility is done on the basis of the recognition scores of 30 trisyllabic words and 20 sentences. The mean value of the results of the two level tests between two dialects indicates the degree of mutual intelligibility between any two dialects, which resulted in the following ranking of mutual intelligibility among the five dialects (where '>' means 'more mutually intelligible than'):

Huangyan-Quzhou, Huangyan-Wujiang > Huangyan-Yongkang,

Wujiang-Quzhou, Quzhou-Yongkang > Wujiang-Yongkang, Wenzhou-Wujiang,

Huangyan-Wenzhou, Wenzhou-Yongkang, Wenzhou-Quzhou.

Based on the percentage of each dialect that was correctly recognized by speakers of the other four dialects, the degree of intelligibility of the five dialects was also established, viz., Huangyan, Wujiang > Quzhou, Yongkang > Wenzhou (where '>' means 'more intelligible than'). The result is in accordance with the general impression that Wenzhou is the most difficult Wu dialect to understand, and Wujiang (a Northern Wu dialect) is easier. A reasonable explanation for the fact that Huangyan is most intelligible is probably due to the fact that Huangyan, linguistically in between Northern Wu and Southern Wu, shares more linguistic similarities with both.

We also reexamined the concept of 'adjusted percentage of information transfer', proposed by Pierce (1952). The results show that there is no significant difference between raw mutual intelligibility scores and adjusted mutual intelligibility scores,

which strongly supports our experimental design.

Linguists have also tried different means to measure linguistic distance between varieties, by making use of either perceived linguistic distance or objective linguistic distance. The perceived linguistic distance is measured on the basis of subjects' perception, while the objective linguistic distance is calculated on the basis of the differences between varieties themselves. To measure linguistic distance among the five dialects, edit distance (Levenshtein distance) was applied, which is a widely used measure in various fields of research. Kessler (1995) first applied this method to measure linguistic distance between Irish dialects, and it has also been applied to measure linguistic distances between West Germanic languages, Scandinavian languages and other alphabetic languages. Edit distance or Levenshtein distance between two strings is defined as the minimum number of edits needed to transform one string into the other. When words are treated as phonetic transcription sequences, the distance between two words is simply the number of the minimum number of edits in phonetic transcription that are needed to transform one transcription into the other between the two languages in question. In this way, two different pronunciations of any cognate in two different dialects can be compared. The distance between any two dialects can be calculated based on the distances of cognates.

Linguistic distances among the five Wu dialects were carefully calculated based on the 30 trisyllabic words and the 20 sentences that were recorded. In the process of calculating the edit distances based on the sentences, we motivated the choice of the specific way, referred to as Method 1, to calculate those distances. We found that Method 1 is a more efficient measure than Method 2, because the former takes the sentence context into consideration. Based on the calculations of 10 dialect pairs, the linguistic distances among them can be ranked as follows (where '<' means 'is smaller in linguistic distance'):

Huangyan-Wujiang, Huangyan-Quzhou, < Huangyan-Yongkang,

Huangyan-Wenzhou, Wenzhou-Yongkang, Wujiang-Quzhou <

Wenzhou-Wujiang, Quzhou-Yongkang, Wenzhou-Quzhou, Wujiang-Yongkang.

We also explored the different correlations that may exist between the mutual intelligibility scores and objective distance measures, viz., linguistic distance, structural distance, geographical distance and travel distance. We did not find any significant correlations between the mutual intelligibility scores and structural distance, geographical distance or travel distance. The only significant correlation is between the mutual intelligibility scores and the linguistic distances.

On the basis of this analysis, the conclusion can be drawn that phonetic distance calculated by measuring edit distance is the best predictor for mutual intelligibility among the five Wu dialects and that linguistic distance is closely related to mutual intelligibility among the five Wu dialects, i.e., the larger the linguistic distance, the lower the mutual intelligibility, and vice versa.

**Keywords:** Wu Dialect; mutual intelligibility, linguistic distance, edit distance

**Contents**

**Acknowledgements** ..... i

**内容摘要** ..... ii

**Abstract** ..... v

**List of Tables** ..... xv

**List of Figures** ..... xx

**List of Abbreviations** ..... xxi

**1 One Introduction** ..... 1

1.1 Rationale of the Study ..... 1

1.2 The Choice of Dialects for the Study: the Five Wu Dialects ..... 3

1.3 Relevant Notions ..... 10

1.3.1 Linguistic Distance ..... 10

1.3.2 Mutual Intelligibility ..... 12

1.3.3 Language, Dialect, Accent and Variety ..... 12

1.4 The Aim of the Study ..... 15

1.5 Structure of the book ..... 16

**2 Previous Studies on Mutual Intelligibility** ..... 18

2.1 The Early Studies on Mutual Intelligibility ..... 18

2.2 Objective Testing versus Subjective Testing ..... 22

2.3 Studies on Mutual Intelligibility of Chinese Dialects ..... 26

2.4 Summary ..... 31

**3 The Phonology of Five Wu Dialects** ..... 34

3.1 General Introduction of the Previous Studies ..... 34

3.2 The Phonology of Huangyan ..... 38

3.2.1. Initials ..... 38
3.2.2 Finals ..... 41
3.2.3 Tones ..... 44
3.3 The Phonology of Wenzhou ..... 45
3.3.1 Initials ..... 45
3.3.2 Finals ..... 49
3.3.3 Tones ..... 51
3.4 The Phonology of Quzhou ..... 53
3.4.1 Initials ..... 53
3.4.2 Finals ..... 54
3.4.3 Tones ..... 55
3.5 The Phonology of Yongkang ..... 56
3.5.1 Initials ..... 56
3.5.2 Finals ..... 58
3.5.3 Tones ..... 59
3.6 The Phonology of Wujiang ..... 60
3.6.1 Initials ..... 60
3.6.2 Finals ..... 61
3.6.3 Tones ..... 62
3.7 A Brief Comparison of the Phonological Structures of the Five Dialects ..... 64
3.7.1 Initials ..... 66
3.7.2 Finals ..... 68
3.7.3 Tones ..... 72
3.7.4 Syllable Structure ..... 73
3.8 Summary ..... 74
**4 Experimental Analysis of Mutual Intelligibility Among Five Wu Dialects ..... 75**
4.1 Introduction ..... 75

4.2 Stimulus Material ..... 76

4.3 Recording of the Material ..... 79

4.3.1 Speakers of the Stimuli ..... 79

4.3.2 Sound Recordings ..... 80

4.3.3 Sound Editing ..... 80

4.4 Listeners for the Test ..... 83

4.5 Testing ..... 84

4.6 The Results ..... 86

4.6.1 Results of Self-self Pairs ..... 86

4.6.2 Results between Different Dialect Pairs ..... 89

4.6.3 Comparisons between the Results of Self-self Pairs and those of Dialect Pairs ..... 90

4.6.4 Mutual Intelligibility Scores ..... 92

4.6.5 The Intelligibility Degree of the Five Wu Dialects ..... 94

4.6.6 Adjusted Intelligibility Scores ..... 96

4.7 Summary ..... 99

**5 Linguistic Distances Among Five Wu Dialects ..... 101**

5.1 Edit Distance ..... 102

5.1.1 Introduction ..... 102

5.1.2 Edit Distance: Phonetic Transcription Based ..... 105

5.1.3 Matches of Segments ..... 106

5.1.4 Operation Weights ..... 107

5.1.5 Non-normalized and Normalized Edit Distances ..... 108

5.1.5.1 Non-normalized Edit Distance ..... 108

5.1.5.2 Normalized Edit Distance ..... 109

5.2 Linguistic Distances of the Ten Dialect Pairs ..... 112

5.2.1 Reliability of the Sample ..... 112

5.2.2 Edit Distances Calculated Based on the Data of Trisyllabic Words ..... 115

5.2.3 Edit Distances Calculated Based on the Data of Sentences ..... 137

5.2.4 Linguistic Distances Calculated Based on the Data of Words and Sentences ..... 156

5.2.5 Tonal Distances based on the Data of Trisyllabic Words ..... 157

5.3 Visualizing Aggregate Differences: Multidimensional Scaling ..... 160

5.3.1 Visualizing the Phonetic Distances ..... 160

5.3.1.1 Visualizing the Non-Normalized Phonetic Distances ..... 160

5.3.1.2 Visualizing the Normalized Phonetic Distances (Method 1) ..... 162

5.3.1.3 Visualizing the Normalized Phonetic Distances (Method 2) ..... 164

5.3.2 Visualizing the Vowel Distances ..... 166

5.3.2.1 Visualize the Non-normalized Vowel Distances ..... 166

5.3.2.2 Visualizing the Normalized Vowel Distances (Method 1) ..... 167

5.3.2.3 Visualizing the Normalized Vowel Distances (Method 2) ..... 169

5.3.3 Visualizing the Consonant Distances ..... 171

5.3.3.1 Visualizing the Non-normalized Consonant Distances ..... 171

5.3.3.2 Visualizing the Normalized Consonant Distances (Method 1) ..... 172

5.3.3.3 Visualizing the Normalized Consonant Distances (Method 2) ..... 174

5.3.4 Visualizing the Mutual Intelligibility Distances ..... 175

5.3.4.1 Visualizing the Mutual Intelligibility Distances (Based on the Word Level and Sentence Level Tests) ..... 176

5.3.4.2 Visualizing the Word Mutual Intelligibility Distances ..... 177

5.3.4.3 Visualizing the Sentence Mutual Intelligibility Distances ..... 179

5.4 Summary ..... 181

**6 Predictors for Mutual Intelligibility Among Five Wu Dialects ..... 182**

6.1 Introduction ..... 182

6.2 Linguistic Distance as a Predictor ..... 183

6.2.1 Correlations between the Linguistic Distances and the Raw Mutual Intelligibility Scores ..... 184

6.2.2 Correlations between the Linguistic Distances and the Adjusted Mutual Intelligibility Scores ..... 188

6.2.3 Correlations between the Tonal Distances and the Mutual Intelligibility Scores 192

6.3 Structural Distance as a Predictor ..... 193

6.3.1 Structural Distances Among the Five Dialects ..... 196

6.3.2 Correlations Between the Structural Distances and the Mutual Intelligibility Scores ..... 200

6.4 Geographical Distance as a Predictor ..... 202

6.5 Summary ..... 204

**7 Conclusion ..... 206**

7.1 Summary of the Major Findings ..... 206

7.1.1 Word Level Test vs. Sentence Level Test ..... 206

7.1.2 The Degree of Mutual Intelligibility Among Five Wu Dialects ..... 207

7.1.3 Raw Intelligibility Scores vs. Adjusted Intelligibility Scores ..... 208

7.1.4 Non-normalized Linguistic Distance vs. Normalized Linguistic Distance ..... 209

7.1.5 Method 1 vs. Method 2 ..... 210

7.1.6 Vowels vs. Consonants ..... 211

7.1.7 Predictors for Mutual Intelligibility ..... 212

7.2 Limitations ..... 213

7.3 Suggestions ..... 214

**References ..... 216**

**Appendix I ..... 230**

**Appendix II ..... 232**

**Appendix III ..... 240**

**Appendix IV ..... 241**

# List of Tables

Table 3.1 The Initials of the Huangyan Dialect (according to Qian, 1992) .....40

Table 3.2 The Finals of the Huangyan Dialect (according to Qian, 1992).....42

Table 3.3 The Tonal System of the Huangyan Dialect (according to Qian, 1992) .....44

Table 3.4 The Tonemes of the Huangyan Dialect .....45

Table 3.5 The Differences between Qian (1992), You and Yang (1998), You (2002) and Scholz (2012) .....45

Table 3.6 The Initials of the Wenzhou Dialect (mainly according to Qian, 1992) .....49

Table 3.7 The Finals of the Wenzhou Dialect (according to Qian, 1992).....51

Table 3.8 The Tonal System of the Wenzhou Dialect (Qian, 1992) .....51

Table 3.9 The Tonal System of the Wenzhou Dialect (Chen, 2000; Zhengzhang, 1980).....52

Table 3.10 The Tonemes of the Wenzhou Dialect (according to Chen, 2000) .....52

Table 3.11 The Initials of the Quzhou Dialect (according to Qian, 1992) .....54

Table 3.12 The Finals of the Quzhou Dialect (according to Qian, 1992) .....54

Table 3.13 The Tonal System of the Quzhou Dialect .....55

Table 3.14 The Tonemes of the Quzhou Dialect.....55

Table 3.15 The Initials of the Yongkang Dialect (according to Qian, 1992) .....57

Table 3.16 The Finals of the Yongkang Dialect (according to Qian, 1992) .....58

Table 3.17 The Tonal System of the Yongkang Dialect .....59

Table 3.18 The Tonemes of the Yongkang Dialect.....59

Table 3.19 The Initials of the Wujiang Dialect (according to Qian, 1992).....61

Table 3.20 The Finals of the Wujiang Dialect (according to Qian, 1992) .....61

Table 3.21 The Tonal System of the Wujiang Dialect (Qian, 1992).....62

Table 3.22 Numeric descriptions of Wujiang Citation Tones by Six Scholars (Wang, 2010: 311).....63

Table 3.23 The Tonemes of the Wujiang Dialect.....64

Table 3.24 The Number of the Phonological Segments of the Five Dialects.....65

Table 3.25 The Occurrence of the Initials in the Five Dialects ..... 66

Table 3.26 The Occurrence of the Finals in the Five Dialects ..... 68

Table 3.27 The Occurrence of Tonemes in the Five Dialects ..... 72

Table 4.1 The Pretest Scores of Speakers of Huangyan (percentages of correct responses) ..... 81

Table 4.2 The Pretest Scores of Speakers of Wenzhou (percentages of correct responses) ..... 81

Table 4.3 The Pretest Scores of Speakers of Quzhou (percentages of correct responses) ..... 82

Table 4.4 The Pretest Scores of Speakers of Yongkang (percentages of correct responses) ..... 82

Table 4.5 The Pretest Scores of Speakers of Wujiang (percentages of correct responses) ..... 82

Table 4.6 Background of the Listeners from the Five Dialects ..... 84

Table 4.7 Examples of Grading for the Trisyllabic Word '黄梅天' ..... 85

Table 4.8 Examples of Grading for the Sentence '下雪天爬山有点危险' ..... 86

Table 4.9 The Word Intelligibility Scores of Self-self Pairs ..... 87

Table 4.10 The Sentence Intelligibility Scores of Self-self pairs (percentage) ..... 88

Table 4.11 The Overall Intelligibility Scores of Self-self Pairs ..... 88

Table 4.12 The Word Intelligibility Scores between 20 Dialect Pairs ..... 89

Table 4.13 The Sentence Intelligibility Scores between 20 Dialect Pairs ..... 89

Table 4.14 The Overall Intelligibility Scores between 20 Dialect Pairs ..... 90

Table 4.15 Mutual Intelligibility Scores between the Ten Dialect Pairs ..... 92

Table 4.16 The Overall Mutual Intelligibility Scores of the Ten Dialect Pairs ..... 93

Table 4.17 The Intelligibility Degree of the Five Wu Dialects ..... 94

Table 4.18 The Adjusted Word Intelligibility Scores between 20 Dialect Pairs ..... 97

Table 4.19 The Adjusted Sentence Intelligibility Scores between 20 Dialect Pairs ..... 97

Table 4.20 The Adjusted Mutual Intelligibility Scores between the Ten Dialect Pairs ..... 97

Table 4.21 The Overall Adjusted Mutual Intelligibility Scores between the Ten Dialect Pairs ..... 98

Table 5.1 Edit Distance Algorithm Between 'Bag' and 'Big' ..... 103

Table 5.2 The Edit Distance Between [kaːdɪf] and [kæːdɪf] ..... 105

Table 5.3 The Non-normalized Edit Distance Between [dag] and [bagz] ..... 109

Table 5.4 The Normalized Edit Distance Between [daɡ] and [baɡz] ..... 110

Table 5.5 Cronbach's Alpha and the Degrees of Internal Consistency ..... 114

Table 5.6 Edit Distances Between the Two Pronunciations of the Trisyllabic Word'八仙桌' in Huangyan and Wenzhou Dialects ..... 115

Table 5.7 Segmental Edit Distances Between Huangyan and Wenzhou Based on the 30 Trisyllabic Words ..... 117

Table 5.8 Segmental Edit distances Between Huangyan and Quzhou Based on the 30 Trisyllabic Words ..... 119

Table 5.9 Segmental Edit Distances Between Huangyan and Yongkang Based on the 30 Trisyllabic Words ..... 121

Table 5.10 Segmental Edit Distances Between Huangyan and Wujiang Based on the 30 Trisyllabic Words ..... 122

Table 5.11 Segmental Edit Distances Between Wenzhou and Quzhou Based on the 30 Trisyllabic Words ..... 124

Table 5.12 Segmental Edit Distances Between Wenzhou and Yongkang Based on the 30 Trisyllabic Words ..... 126

Table 5.13 Segmental Edit Distances Between Wenzhou and Wujiang Based on the 30 Trisyllabic Words ..... 127

Table 5.14 Segmental Edit Distances Between Quzhou and Yongkang Based on the 30 Trisyllabic Words ..... 129

Table 5.15 Segmental Edit Distances Between Quzhou and Wujiang Based on the 30 Trisyllabic Words ..... 131

Table 5.16 Segmental Edit Distances Between Yongkang and Wujiang Based on the 30 Trisyllabic Words ..... 132

Table 5.17 Segmental Edit Distances (non-normalized and normalized) Between 10 Dialect Pairs Based on the Trisyllabic Words ..... 134

Table 5.18 Edit Distances Between Huangyan and Wenzhou Based on the Sentence ..... 138

Table 5.19 Edit Distances of the Item '夏天' Between Huangyan and Wenzhou .....139

Table 5.20 Edit Distances of the Item '蚊子' Between Huangyan and Wenzhou .....139

Table 5.21 Edit Distances of the Item '多' Between Huangyan and Wenzhou .....139

Table 5.22 Edit Distances of '夏天屋里蚊子太多' Between Huangyan and Wenzhou .....140

Table 5.23 Edit Distances Between Huangyan and Wenzhou Based on the 20 Sentences.....141

Table 5.24 Edit Distances Between Huangyan and Quzhou Based on the 20 Sentences .....142

Table 5.25 Edit Distances Between Huangyan and Yongkang Based on the 20 Sentences .....144

Table 5.26 Edit Distances Between Huangyan and Wujiang Based on the 20 Sentences .....145

Table 5.27 Edit Distances Between Wenzhou and Quzhou Based on the 20 Sentences.....146

Table 5.28 Edit Distances Between Wenzhou and Yongkang Based on the 20 Sentences.....148

Table 5.29 Edit Distances Between Wenzhou and Wujiang Based on the 20 Sentences .....149

Table 5.30 Edit Distances Between Quzhou and Yongkang Based on the 20 Sentences .....150

Table 5.31 Edit Distances Between Quzhou and Wujiang Based on the 20 Sentences .....152

Table 5.32 Edit Distances Between Yongkang and Wujiang Based on the 20 Sentences .....153

Table 5.33 Edit Distances Calculated by Using Method 1 and Method 2.....155

Table 5.34 Linguistic Distances Between the Ten Dialect Pairs.....156

Table 6.1 Correlations Between Linguistic Distances and Mutual Intelligibility Scores at the Word Level Test .....184

Table 6.2 Correlations Between Linguistic Distances and Mutual Intelligibility Scores in the Sentence Level Test .....185

Table 6.3 Correlations Between Linguistic Distances and the Raw Overall Mutual Intelligibility Scores .....187

Table 6.4 Correlations Between Linguistic Distances and Adjusted Mutual Intelligibility Scores at the Word Level Test .....189

Table 6.5 Correlations Between Linguistic Distances and Adjusted Mutual Intelligibility Scores at the Sentence Level Test.....190

Table 6.6 Correlations Between Linguistic Distances and the Adjusted Overall Mutual Intelligibility

Scores .....191

Table 6.7 Correlations Between the Tonal Distances and the Mutual intelligibility Scores.....193

Table 6.8 The Structural Distances Based on the Inventories of Initials of the Five Wu dialects .....196

Table 6.9 The Structural Distances Based on the Inventories of Finals of the Five Wu Dialects.....197

Table 6.10 The Structural Distances Based on the Inventories of Tonemes of the Five Wu dialects...199

Table 6.11 The Structural Distances Among the Five Wu Dialect .....201

Table 6.12 Correlations Between the Structural Distances and the Mutual Intelligibility Scores .....201

Table 6.13 Correlations Between Geographical Distances and Linguistic Distances.....203

# List of Figures

Figure 1.1 The Wu Dialect Area in China ..... 4

Figure 1.2 Divisions of Northern Wu and Southern Wu (Fu et al., 1986) ..... 6

Figure 1.3 Geographical Map of the Wu Dialect Area (Language Atlas of China, 1987) ..... 8

Figure 3.1 The Traditional Account of Chinese Syllable Structure (Wang and Smith, 1997: 7) ..... 38

Figure 3.2 The Hierarchical View of Syllable Structure (Treiman, 1993: 19) ..... 39

Figure 3.3 The Chinese Syllable Structure in Modern Phonological Terminology (Cheng, 1966: 136) 39

Figure 4.1 Ten Dialect Pairs of the Five Wu Dialects ..... 92

Figure 5.1 MDS of Non-normalized Phonetic Distances Among the Five Wu Dialects ..... 162

Figure 5.2 MDS of Normalized Phonetic Distances (Method 1) Among the Five Wu Dialects ..... 163

Figure 5.3 MDS of Normalized Phonetic Distances Among the Five Wu Dialects (Method 2) ..... 165

Figure 5.4 MDS of Non-normalized Vowel Distances Among the Five Wu Dialects ..... 167

Figure 5.5 MDS of Normalized Vowel Distances Among the Five Wu Dialects (Method 1) ..... 168

Figure 5.6 MDS of Normalized Vowel Distances Among the Five Wu Dialects (Method 2) ..... 170

Figure 5.7 MDS of Non-normalized Consonant Distances Among the Five Wu Dialects ..... 172

Figure 5.8 MDS of Normalized Consonant Distances Among the Five Wu Dialects (Method 1) ..... 173

Figure 5.9 MDS of Normalized Consonant Distances Among the Five Wu Dialects (Method 2) ..... 175

Figure 5.10 MDS of Mutual Intelligibility Distances Among the Five Wu Dialects ..... 177

Figure 5.11 MDS of Word Intelligibility Distances Among the Five Wu Dialects ..... 178

Figure 5.12 MDS of Sentence Mutual Intelligibility Distances Among the Five Wu Dialects ..... 180

Figure 6.1 MDS of Structural Distances Based on the Inventories of Initials of the Five Wu Dialects 197

Figure 6.2 MDS of Structural Distances Based on the Inventories of Finals of the Five Wu Dialects 198

Figure 6.3 MDS of Structural Distances Based on the Inventories of Tones of the Five Wu Dialects 200

# List of Abbreviations

| | |
|---|---|
| CASS | Chinese Academy of Social Sciences |
| CD | Consonant Distance |
| Del. | Deletion |
| DOC | Dictionary on Computer |
| ECNU | East China Normal University |
| HY | Huangyan |
| Ins. | Insertion |
| MDS | Multidimensional Scaling |
| MIS | Mutual Intelligibility Scores |
| NCD | Normalized Consonant Distance |
| NNCD | Non-normalized Consonant Distance |
| NNVD | Non-normalized Vowel Distance |
| NNPD | Non-normalized Phonetic Distance |
| NPD | Normalized Phonetic Distance |
| NVD | Normalized Vowel Distance |
| Oper. | Operation |
| PD | Phonetic Distance |
| SMI | Systemic Mutual Intelligibility |
| Sub. | Substitution |
| VD | Vowel Distance |
| WZ | Wenzhou |