Learning to Read the Complex Manchu Writing System:

A Mixed Methods Study of Novice Learners in a Chinese University



Bai Li

Department of Education

Wolfson College

University of Oxford

Supervisors

Professor Victoria Murphy

Professor Sonali Nag

Thesis submitted in partial fulfilment of the requirement

for the degree of Doctor of Philosophy in Education

Hilary Term 2022

This thesis is dedicated to my parents, Li Zhaojin and Li Yanjun.

For their endless love, support, and encouragement.

Abstract

This study investigated the association between the Manchu writing system and novice learners' *uju* symbol block decoding performance. The Manchu writing system was examined on two parameters: visual complexity of Manchu symbols and mapping complexity between symbol and sound. Designed as mixed methods research, this study randomly selected undergraduate students (n = 196) in a Chinese university, who were given Manchu uju naming tasks in Test 1 on visual complexity and Test 2 on symbol–sound mapping complexity. Moreover, the instruction on Manchu orthographic knowledge was observed. Theoretically, an area of interest in reading acquisition is the learning about the visual features of a writing system. Symbol visual complexity was previously found to be negatively associated with word reading in a limited number of orthographies such as Kannada (Nag, Snowling, Quinlan, & Hulme, 2014), Arabic (Abdelhadi, Ibrahim, & Eviatar, 2011) and Chinese (Yu, Zhang, Priest, Reichle, & Sheridan, 2018). To measure visual complexity, the present study adopted a multidimensional framework which included pixel count (Nag & Snowling, 2012), perimetric complexity, the number of simple features, the number of connected points and the number of disconnected components (Chang, Chen, & Perfetti, 2018). Results of Test 1 showed that visual complexity significantly and positively associated with participants' uju decoding performance measured by error rate and reaction time. Furthermore, connected point was identified by multiple linear regression as the unique predictor for decoding accuracy, and connected point and disconnected component were found to predict decoding speed. Results of Test 1 suggested that novice readers may use connected point to identify features in visual processing of uju symbol blocks. Secondly, as Manchu orthography is characterized by the complex mixed mapping between symbol and sound (Nag, 2007, 2017a, 2021), it remained unknown what symbol grain size(s) (Ziegler & Goswami, 2005) was/were adopted in Manchu reading. Results of Test 2 showed that

participants' decoding performance was not significantly different when uju were provided in the manipulated "mixed" and "blocked" condition. Yet, their performance was more accurate and rapid in naming uju with all comprising phoneme markers in one-to-one mapping with sound than those containing a target marker in one-to-multiple mapping with sound. The results suggested that phoneme marker may be used for Manchu symbol blocks processing. This finding on learners' choice of the smaller grain size phoneme marker in reading must be seen in the context of the classroom observations on Manchu orthographic knowledge instruction. Although both global and analytic strategies as defined by Nag (2011, 2017a, 2017b) were applied by the teacher in instruction of the uju symbol blocks, even novice learners who received a rather limited ten sessions of Manchu instruction on the orthography were found to be reading analytically by decomposing an uju into component phoneme markers. Finally, in line with theorizing by Nag, this study posits that reading Manchu symbol blocks requires intra-symbol processing, which involves (a) at the feature level, identification and configuration of visual features into phoneme markers, and (b) at the symbol level, conversion of phoneme markers into sounds.

Keywords: Manchu, decoding, uju symbol blocks, novice learners, mixed methods

Acknowledgements

This thesis has marked a notable time in my life, but it would not have been possible without the support and encouragement of many others. Particularly, in my DPhil journey I feel very lucky to have met the following persons, to whom I would like to give my most heartfelt thanks.

First and foremost, I must thank my supervisors, Professor Victoria Murphy and Professor Sonali Nag. Their interview with me in 2017 to discuss my proposed research in Manchu feels like yesterday, and I am deeply grateful to them for agreeing to supervise my DPhil research. Their guidance, inspiration, insight, and critique are invaluable in shaping my thesis, without which this first empirical study on Manchu language would not have been possible. Their erudition and dedication have set an excellent example for my future endeavours in academics. My deepest gratitude to Vicky and Sonali is beyond words.

I owe a significant debt of gratitude to many people who have generously supported me in my pursuit of DPhil studies. I am deeply thankful to Professor Nancy H. Hornberger, Professor Diane Larsen-Freeman, and Professor Yuko Goto Butler at the Graduate School of Education, University of Pennsylvania. Their experiences, dedication, and scholarship ignited my passion for exploring the underexplored. I want to express my deep gratitude to them.

The intellectual stimulation of Oxford is incredible, but I must give special mention to Dr Heath Rose and Dr James Robson, the examiners of my Transfer of Status, Dr Robert Woore and Dr David Mills, the examiners of my Confirmation of Status, and Dr Heath Rose (again) and Dr Li-Yun Chang, the examiners of my *viva voce*, for their provision of the thought-provoking comments and feedback. Special thanks also go to members of the Applied Linguistics and Second Language Acquisition (ALSLA) research group at the Department of Education for the interesting discussion and valuable feedback.

Before beginning my DPhil studies at the University of Oxford, I have spent one and half years doing research related to Manchu language in London, which I also consider an important part of my intellectual journey. I would like to thank Professor Li Wei at the Institute of Education, University College London, Dr Martin Dewey, Dr Simon Coffey, Professor Brian Street, and Professor Constant Leung at King's College London for their guidance to my research. I am particularly grateful to Dr Lars Laamann at the School of Oriental and African Studies (SOAS) University of London, for offering lecture on Manchu orthography, which was truly an academic luxury in the UK. I would also like to thank Professor Peter Austin and Professor Julia Sallabank at SOAS University of London for providing me with the opportunity of conducing real-hands field research on endangered language documentation. My gratitude also goes to Professor Lilija Gorelova, Dr Mikael Thompson, and Dr Wang Qi who have kindly shared their thoughts regarding my questions.

During my residence in Oxford, I considered Wolfson College my "second home" for its tranquillity, peacefulness, and decent food. I am particularly grateful to Sir Tim Hitchens, the President of Wolfson, and Professor Philomen Probert, the interim President, for maintaining the egalitarian, diverse, and international ethos in college. I also want to express deep gratitude to my college advisor Professor Ruben Anderson and University counsellor Dr Silvia Elsner who provided valuable advice and support at the dark time of my life. A special mention also goes to Ms. Catherine Scutt, the librarian at the Department of Education, Oxford, for her professionalism and help with references.

I am also significantly thankful to all the participants in my DPhil research and an almost unending list of people whose friendship and encouragement helped me turn this project into a reality. Specifically, I would like to thank Dr Jin Biao, Ms Sun Cuiyun, Dr Xia Yuzhu, Ms You Gang, Ms Li Mingyi, and Professor Zhao Aping for their kind help during my piloting and fieldwork. Special thanks must go to Dr Wang Shuo, who has graciously

introduced me to the field and participants. I would also like to thank Dr Lian Xi, Dr Samuel Tsang, Dr Fang Zuyi, Dr Zhang Cheng, Dr Ma Biao, and Dr Yan Shuhao at Oxford for all the laughter, joy, and fun that we have shared.

I would like to thank China Scholarship Council (CSC) for providing generous financial support for my DPhil studies. My gratitude also goes to Wolfson College of Oxford University, the Department of Education at Oxford University, the Quantitative Hub of the Department of Education, and the Universities' China Committee in London (UCCL) for their awards of research grants and bursaries.

A special thank you must go to my close friend Dr H Chen, for being a perfect companion whose encouragement and positivity have been inspiring me in my DPhil journey. Finally, the sacrifices made by my parents, Li Yanjun and Li Zhaojin, has been substantial. I owe a significant debt of gratitude to them and am unspeakably grateful for their endless love, support, and encouragement. Thank you, Mom and Dad!

Table of Contents

Abstract *2*

Acknowledgements *4*

Table of Contents *7*

List of Figures *11*

List of Tables *12*

List of Abbreviations *14*

List of Terms *15*

Chapter 1 Introduction *17*

1.1 Chapter Overview **17**

1.2 Focus of the Study **17**

1.3 A Historical Account of Manchu **18**

1.3.1 Manchu as a People 18

1.3.2 Manchu as a Language 19

1.4 Denotation for Symbols and Sound **20**

1.5 Endangerment of Manchu Language **21**

1.6 Manchu Writing System **23**

1.6.1 Definition and Typology of Manchu Writing System 23

1.6.2 Biscripts of Manchu 23

1.6.3 Positional Variation 27

1.6.4 Symbol–Sound Mapping in Manchu 29

1.6.5 Manchu Hergen, Symbol Blocks, and Phoneme Markers 30

1.6.6 Debate on the Typology of Manchu Writing System 32

1.6.7 A "Contextualised" View on Understanding Manchu Writing System 34

1.7 Chapter Summary **34**

Chapter 2 Literature Review *37*

2.1 Chapter Overview **37**

2.2 Visual Complexity and Word Reading **37**

2.3 Symbol–Sound Mapping Complexity and Word Reading **41**

2.4 Instructional Strategies of Orthographic Knowledge across Writing Systems **46**

2.5 Chapter Summary 48

Chapter 3 Methodology 50

3.1 Chapter Overview 50

3.2 A Mixed Methods Research Design 50

3.2.1 Rationale for Using Mixed Methods Research 50

3.2.2 Epistemological and Ontological Considerations on Mixed Methods Research 52

3.2.3 A Multidimensional MMR Framework 54

3.3 Sample and Sampling 55

3.3.1 The Mixed Methods Sampling Model 55

3.3.2 Sampling Frame and Schemes 56

3.3.3 Sample Size 58

3.4 Instruments 59

3.4.1 Measures of Manchu Symbol Block Visual Complexity and Mapping Complexity 59

3.4.2 Measures of Manchu Symbol Block Decoding Performance 62

3.4.3 Validity of Instrument 63

3.5 Procedure of Data Collection 63

3.5.1 Implementation Matrix for Data Collection 63

3.5.2 Test Design 66

3.5.3 Task 71

3.5.4 Procedure 71

3.5.5 Classroom Observation on Manchu Orthography Instruction 72

3.5.6 Fieldwork Timeline 77

3.6 Piloting 77

3.7 Ethical Consideration 78

3.8 Strategies for Data Analysis 79

3.8.1 Separate and Mixed Methods of Data Analysis 79

3.8.2 Analytical Strategies 79

3.9 Chapter Summary 89

Chapter 4 Results for Test 1 on Visual Complexity 92

4.1 Chapter Overview 92

4.2 Introduction to Test 1 92

4.2.1 Purpose of Test 1 92

4.2.2 Items in Test 1 93

4.2.3 Procedure of Test 1 96

4.3 Multiple Imputation of the Missing Data in Test 1 96
4.3.1 Missing Data Mechanism 97
4.3.2 Assumptions Required by Multiple Imputation 97
4.3.3 Steps of Multiple Imputation 98
4.3.4 Results of Multiple Imputation of Missing Data in Test 1 99
4.4 Results of Test 1 101
4.4.1 Participant Demographic Information 101
4.4.2 Participants' Overall Performance in Manchu Symbol Blocks Recognition in Test 1 102
4.4.3 Descriptive Data of Symbol-Level and Outcome-Level Measures 103
4.4.4 Associations among Symbol-Level Measures and between Outcome-Level Measures 106
4.4.5 Association between Symbol-Level and Outcome-Level Measures 107
4.4.6 Predictors for Manchu Symbol Block Decoding Performance 109
4.5 Chapter Summary 117
Chapter 5 Results for Test 2 on Symbol–Sound Mapping Complexity 119
5.1 Chapter Overview 119
5.2 Introduction to Test 2 119
5.2.1 Purpose of Test 2 119
5.2.2 Items in Test 2 120
5.2.3 Control of Item Visual Complexity and Frequency 123
5.2.4 Manipulation of Item Presentation 124
5.2.5 Procedure of Test 2 126
5.3 Multiple Imputation of the Missing Data in Test 2 126
5.3.1 Meeting the Assumptions of Multiple Imputation 127
5.3.2 Results of Multiple Imputation of Missing Data in Test 2 128
5.4 Results for Test 2 133
5.4.1 Assumptions Check for Paired Samples t-Test 133
5.4.2 Participants' Decoding Performance between Blocked and Mixed Conditions 134
5.4.3 Participants' Decoding Performance between Block 1 and Block 2 135
5.4.4 The Reliable Change Index Controlling for Practice Effects 136
5.5 Chapter Summary 140
Chapter 6 Results of Classroom Observation of Manchu Orthography Instruction 142
6.1 Chapter Overview 142
6.2 Introduction to Classroom Observation 142
6.2.1 Purpose of Classroom Observation 142
6.2.2 Focus of Classroom Observation 143

6.3 Data Analysis 143
6.3.1 Data Synthesis Matrix 143
6.3.2 Codes and Subcodes 146
6.3.3 Inter-coder Reliability Check on Coding 146
6.3.4 Symbol Frequency Count 156
6.4 Findings 156
6.4.1 Finding 1: Instruction on the Writing of Uju Symbol Blocks 157
6.4.2 Finding 2: Instruction on the Reading of Uju Symbol Blocks 167
6.5 Chapter Summary 174
Chapter 7 Discussion 176
7.1 Chapter Overview 176
7.2 Key Findings 176
7.3 Discussion 178
7.3.1 Visual Complexity and Manchu Uju Symbol Blocks Processing 178
7.3.2 Symbol-Sound Mapping Complexity and Manchu Uju Symbol Blocks Processing 182
7.3.3 Model for Manchu Uju Decoding 185
7.3.4 Instructional Strategies and Manchu Uju Reading 187
7.4 Chapter Conclusion 190
Chapter 8 Conclusion, Contributions, and Limitations 193
8.1 Chapter Overview 193
8.2 Conclusion 193
8.3 Contributions 194
8.4 Limitations 196
References 198
Appendix A Ethical Application of the Research Project 211
Appendix B Report on Pilot Study 245
Appendix C Questionnaire of Manchu Language Learning Experience 267
Appendix D Frequency of Manchu Phoneme Marker and Uju in Instruction 268

List of Figures

Figure 1 Manchu Symbol Block <SEK> 18

Figure 2 Uju Symbol Blocks, Phoneme Markers, and Latin Letters in the Word <MANJU> 24

Figure 3 Isolated Uju in the First Section of the Twelve Sections of Manchu Syllabary 25

Figure 4 Four Positional Variants of the Uju <MAN> 28

Figure 5 Three Phoneme Markers Corresponding with /a/ 29

Figure 6 A Visually Highly Complex Manchu Word 36

Figure 7 Mapping Complexity Measure on the Target Phoneme Marker 61

Figure 8 Distribution of the Standardized Residual of Error Rate (ER) 115

Figure 9 Standardized Predicted Values and Residual of Error Rate (ER) 115

Figure 10 Distribution of the Standardized Residual of Reaction Time (RT) 116

Figure 11 Standardized Predicted Values and Residual of Reaction Time (RT) 116

Figure 12 Stroke Transformation Applied on Phoneme Marker in Writing the Isolation, Initial, Medial, and Final Forms of <KA> 158

Figure 13 Stroke Segmentation Applied to the Isolation Form of <A> and <E> 159

Figure 14 Segmentation of <NE> into <N-> and <-E> 169

Figure 15 Segmentation of <AM> into <A-> and <-M> 170

Figure 16 Blackboard Presentation of <-S> and -s 172

Figure 17 Connected Points in the Uju <SEK> 181

List of Tables

Table 1 Manchu Speakers' Proficiency at Sanjiazi Village in 2002 22

Table 2 Syllable Distribution of the Isolated Uju in the "Twelve Sections of Manchu Syllabary" 26

Table 3 Symbol–Sound Mapping in Manchu 29

Table 4 Names of the Twelve Basic Strokes in Traditional Manchu Script 32

Table 5 Implementation Matrix for Data Collection 64

Table 6 List of Items in Test 1 on Visual Complexity 67

Table 7 List of Items in Test 2 on Symbol–Sound Mappin Complexity 70

Table 8 Curricular Contents of the Manchu Intangible Cultural Heritage Course 75

Table 9 A Synthesis Matrix for Organising Classroom Observation Data 82

Table 10 Codes, Subcodes, Definitions, and Examples from Classroom Observation 83

Table 11 Symbol-Level Measures of Visual Complexity for the Manchu Uju Symbol Blocks (*n* = 18) in Test 1 94

Table 12 Demographic Information of the Participants 102

Table 13 Mean, Standard Deviation, and Range of Participants' Uju Symbol Block Recognition Scores in Test 1 (Max. Score = 18) 103

Table 14 Outcome-Level Measures of Participants' Decoding Performance for the Manchu Uju Symbol Blocks (*n* = 18) in Test 1 104

Table 15 Means, Standard Deviations, and Ranges of Symbol-Level and Outcome-Level Measures in Test 1 106

Table 16 Correlation among Symbol-Level Measures for Test 1 107

Table 17 Correlation between Symbol-Level and Outcome-Level Measures for Test 1 109

Table 18 Models of Decoding Accuracy from Multiple Linear Regression (Enter) in Test 1 112

Table 19 Models of Decoding Speed from Multiple Linear Regression (Enter) in Test 1 ...113

Table 20 Models of Decoding Accuracy and Decoding Speed from Multiple Linear Regression (Enter) in Test 1 117

Table 21 Measures of Symbol–Sound Mapping Relations, Visual Complexity, and Frequency of the Uju (*n* = 16) in Test 2.....121

Table 22 Results of the Paired-samples t-Tests of the Visual Complexity of the Uju between Block 1 and Block 2 in Test 2.....124

Table 23 Error Rate and Reaction Time of the Items (*n* = 16) Presented in Blocked and Mixed Conditions in Test 2.....131

Table 24 Descriptive Data of the Error Rate (ER) Difference and the Reaction Time (RT) Difference of the Data for Paired Samples t-Tests in Test 2 134

Table 25 Results of Paired Samples t-Test of Participants' Performance in Decoding Items (*n* = 32) Presented in Blocked and Mixed Conditions in Test 2 135

Table 26 Results of Paired Samples t-Test of Participants' Performance in Decoding Items (*n* = 16) in Blocks 1 and 2 in Test 2 136

Table 27 Results of the Reliable Change Index Controlling for Practice Effects (RCI_{PE}) of the Reaction Time of Item Test-Retest in Test 2 139

Table 28 Results of the Inter-coder Reliability Check on Classroom Observation Data.....147

List of Abbreviations

BERA	British Education Research Association
CP	connected point
DC	disconnected component
DREC	Departmental Research Ethics Committee
ER	error rate
FCS	fully conditional specification
MAR	missing at random
MCAR	missing completely at random
MMR	mixed methods research
MNAT	missing not at random
OBH	orthographic breadth hypothesis
ODH	orthographic depth hypothesis
PC	perimetric complexity
PE	practice effects
PX	pixel count
RCI	Reliable Change Index
RQ	research question
RT	reaction time
SED	standard error of the differences
SF	simple feature
UNESCO	United Nations Educational, Scientific, and Cultural Organization
VIF	variance inflation factor

List of Terms

alphabet	A *writing system*[1] denoting all consonants and vowels
grapheme	The basic abstract written unit of an alphabetic language
hergen	A word in Manchu. There are gargata hergen (monosyllabic word, e.g., *ba* "earth" comprising one *uju* <BA>) and holboho hergen (multisyllabic word, e.g., *abka* "heaven" comprising two *uju* <AB-> *and* <-KA>). A word of Manchu origin.
Latin script	An inventory of Latin alphabets adopted to represent vowels and consonants
Manchu phoneme marker	Orthographic symbol used to represent vowels and consonants in Manchu
orthography	Conventional spelling of texts and the principles of spelling
phoneme	The smallest abstract sound unit of a language
script	A particular set of signs or symbols used for spelling sounds

[1] The definition of an italicised term is also available in this list.

syllabary	A *writing system* in which the orthographic symbols represent syllables; a set of symbols representing all syllables in a *writing system*
symbol block	An orthographic unit formed by stacking up *phoneme markers.* A *uju* in the Manchu writing system is a symbol block.
the Twelve Sections of Manchu Syllabary	The twelve divisions of Manchu *Syllabary* developed in the Qing dynasty. “满文十二字头” in Chinese.
traditional Manchu script	A set of orthographic syllables created by adopting Mongolian *script* to represent sounds in Manchu
uju	The basic Manchu orthographic symbols in V, VV, CV, VC, CVV and CVC structures that compose the *Twelve Sections of Manchu Syllabary*. Further decomposable to Manchu phoneme markers (except for the V syllable). A word of Manchu origin.
writing system	A *script* together with an associated *orthography*

Chapter 1 Introduction

1.1 Chapter Overview

This chapter opens the thesis by introducing some fundamental information regarding Manchu. It first presents the focus of this study and provides a historical account of the Manchu people and their language. Following this is the introduction to the biscripts of Manchu, the basic written symbols, and the symbol–sound mapping system of Manchu. Finally, the nature of Manchu writing system is discussed.

1.2 Focus of the Study

This study focused on investigating the ways in which the complexity of Manchu writing system is related to the learners' *symbol block* decoding performance. In this study, symbol block is defined as an orthographic unit formed by stacking up the symbols of *phoneme markers* (Nag, 2007, 2017a), and decoding refers to "the accurate and fast retrieval of the phonological code" for the written symbol blocks (Verhoeven & van Leeuwe, 2009, p. 205). Specifically, this study investigated the two dimensions of visual complexity and mapping complexity of Manchu writing system. With respect to visual complexity, Manchu symbol blocks varying in symbol complexity (e.g., <SEK> in Figure 1 comprises the three symbols of <S->, <-E-> and <-K>) may affect the learners' decoding accuracy and speed. Moreover, regarding mapping complexity, it inspected the extent to which the complex Manchu symbol–sound mapping associates with how accurately and rapidly the novice learners decoded.

The reason why visual complexity and mapping complexity were identified for investigation are twofold. Firstly, they are of theoretical interest. The investigation will contribute to building a more "universal" theory on reading by looking into a writing system that is less studied and different from the much-investigated alphabetic writing systems. Furthermore, this study also has educational implications. It could inform instruction by

suggesting pedagogical approaches that may be adopted as a solution to teach the writing systems featured by complex symbol–sound mapping relations, which could be a challenge for novice learners.

Figure 1

Manchu Symbol Block <SEK>



1.3 A Historical Account of Manchu

1.3.1 Manchu as a People

Geographically, the Manchu people originated from Northeast China, a region which used to be called the Manchurian area. The Manchus had a long and intricate ancestral line, with earliest ancestry traced back to the Qin dynasty (221–206 BC). Generally, Jurchen was identified as the ancestors of Manchu. The Jurchens, by the middle of Ming dynasty (1368–1644), were comprised of the three tribes of Jianzhou Jurchen, Haixi Jurchen, and Yeren Jurchen, all residing in Northeast China. In 1616, Nurhaci, the chieftain of Jianzhou Jurchen, founded the Later Jin by unifying the three tribes. Then, 1635 marked a significant year in Manchu history because Hong Taiji, the successor and son of Nurhaci, renamed the consolidated tribe from Jurchen to Manchu, signifying the official establishment of the ethnic group. In 1644, he renamed Later Jin as Qing (1644–1911), whose armies occupied the capital of Beijing and overthrew Ming. Since then, the Manchu began their ruling of China, a demographically Han-dominated country, until 1911 when Puyi, its last emperor, abdicated

in the wake of Xinhai Revolution. Consequently, the Manchu group lost their privilege as the ruling class but have remained the second most populous minority ethnic groups (out of 55) in current China.

1.3.2 Manchu as a Language

As previously introduced, Manchu people were the immediate descendants of Jurchen, and consequently the Manchu language was primarily derived from that of Jurchen too. It was in the early 1500s that the Jurchen language evolved into what is called Manchu today, but at that time it did not have a script. As a result, the Mongolian script had to be used in communication (An, 1996). Obviously, the discrepancy between spoken Manchu, on the one hand, and written Mongolian, on the other, can be problematic. People had to constantly switch between two linguistic codes that had no established orthography–phonology mapping, which was by no means effective for communication.

Nurhaci, the founding father of Qing dynasty, did not like that. Thus, in 1599 he commissioned two Manchu officials, Erdeni and Gagai, to create a script to spell spoken Manchu. These two learned men, according to *Manzhou Shilu*, found the task so challenging that they turned to Nurhaci for help, who gave the following suggestion,

> “但以蒙古字合我国之语音，联缀成句，即可因文见义矣。例如阿字下合一妈字，非阿妈乎？恶字下合一摸字非恶摸乎?”
>
> “Just use Mongolian letters to denote the spoken language of our state [to obtain the words], and combine the words into sentences, so that the meaning can be seen from the text. For example, isn’t it *ama* if the character of *ma* is connected below *a*? Isn’t it *eme* if the character of *me* is connected below *e*?”[2]
>
> (*Manzhou Shilu [Veritable Records] (Vol. 3)*, 1969, p. 35, my translation from Chinese)

[2] The meaning of *ama* and *eme* is “father” and “mother” respectively.

Following this suggestion, Erdeni and Gagai created the Manchu script ("the Old Manchu") by borrowing and adapting the Mongolian script. While the script did bring significant social, military, and cultural benefits, it was still deficient in some respects, which could result in communicative ambiguity and misunderstanding. For example, the spelling of *aga* ("rain") was identical with that of *aha* ("slave"). To revolve such problems, Hong Taiji, the son of Nurhaci and the second emperor of Qing dynasty, commanded the Manchu official Dahai to update the Old Manchu writing system. Several methods, such as introducing diacritical circle and dot into the writing system, standardising word spelling, and supplementing additional letters for transliteration of loanwords, were employed by Dahai. The updated script was called the "New Manchu" script and the principles of Manchu orthography were established during the reformation. Thus, it can be inferred from the above historical account that a mapping between Manchu syllables and written symbols (e.g., /mɑ/ – <MA>) was set during the creation and reformation of Manchu scripts.

1.4 Denotation for Symbols and Sound

Throughout this thesis, graphemes were written between angle brackets < >, phonemes between slashes / /, and phones between square brackets [ ]. Since no input keyboard is currently available for the macOS system, the traditionally scripted Manchu symbol blocks and phoneme markers were denoted with uppercase Latin letters written between angle brackets (e.g., <MAN> and <F>). Their initial, medial, and final forms were distinguished with a hyphen or hyphens written after, between, and before it, respectively, (e.g., <MAN->, <-MAN->, and <-MAN>), but the isolated form was written independently without hyphen (e.g., <MAN>). The Latin transliteration of traditional symbol blocks and phoneme markers were denoted by italicised lowercase letters (e.g., *man* and *f*).

1.5 Endangerment of Manchu Language

The Manchu language was identified by UNESCO as "critically endangered", next only to "extinct", in the year of 2010. It was estimated that there were between 20 to 100 native speakers of Manchu, all residing in Heilongjiang Province of Northeast China (Guo, 2008; Mosley, 2010; Zhao, Guo, & He, 2013). According to Mosley, as of 2010, ten of them were in Dawujiazi village in Aihui County of Heihe Prefecture, and the other ten at Sanjiazi village in Fuyu County.[3] Zhao, Guo and He (2013) found that 186 Manchu speakers resided at Sanjiazi village. As shown in Table 1, they were categorized into five groups according to Manchu proficiency. While the first three categories (A, B, and C in Table 1) were all capable of using Manchu language, only the speakers in categories A and B were observed to be capable of using Manchu for communicative purpose. However, they were fluent only in daily Manchu but had difficulties producing syntactically complex sentences.

[3] These number, according to Mosley, is a compromise figure based on various sources. However, the sources used are not listed except for a book written in Russian published in 1997. Retrieved from http://www.unesco.org/culture/languages-atlas/en/atlasmap/language-id-453.html; http://www.unesco.org/culture/languages-atlas/en/atlasmap/language-id-454.html.

Table 1

Manchu Speakers' Proficiency at Sanjiazi Village in 2002

Categories [a]	A	B	C	D	E
Number of Manchu speakers	3	15	48	101	19
Percentage in the 186 Manchu-speaking villagers (%)	1.6	8.1	25.8	54.3	10.2

Note. Adapted from Zhao, Guo, & He (2013, p. 100, my translation from Chinese)

[a] A = Very fluent Manchu speaker

B = Understand all spoken Manchu and speak a great deal of Manchu

C = Understand a great deal of spoken Manchu and speak some Manchu

D = Understand some spoken Manchu but can only speak several Manchu sentences

E = Understand several sentences of spoken Manchu but unable to speak it

As for the Manchu speakers in Heihe Prefecture of Heilongjiang Province, Guo (2003) investigated seven villages and counties in that area, including Dawujiazi village which was identified in the above UNESCO research. While the number of Manchu speakers was not reported, it contributed by verifying that besides Dawujiazi village Manchu speakers were also in existence at six other villages and counties in Heihe Prefecture.

Two points are worth noting here. First, as Zhao et al. (2013) noted, for the Manchu speakers in Sanjiazi village, the language is more of a cultural heritage inherited from their ancestors than an actual tool of daily communication. Second, although the above two studies were the latest available sources for the status quo of Manchu speakers, they were conducted nearly two decades ago. Over the past years some changes may have happened. For example, the number of Manchu speakers might have been in further decline due to the high possibility of the elderly speakers passing away. On the other hand, additional members might have joined the Manchu speaking group as a result of learning the language.

1.6 Manchu Writing System

1.6.1 Definition and Typology of Manchu Writing System

Daniels (2018) defines writing system as "a script together with an associated orthography", with script referring to "a particular collection of characters (or signs)" and orthography to "conventional spelling of texts, and the principle therefor" (p. 155). Put together, these definitions can inform us that a writing system encompasses (a) a script (i.e., a set of characters, signs, or symbols), (b) spelling principles (i.e., the conventions or rules of rendering speech sound into writing), and (c) texts (i.e., written words formulated by spelling). Based on the three dimensions, the present study defined Manchu writing system as a set of particular symbols which, based on certain conventions, are used to spell symbol blocks that represent the sound of Manchu. Yet, what typology the Manchu writing system belongs to is still under debate. Among the five fundamental types of writing systems (logography, syllabary, abjad, alphabet, and abugida) identified by Daniels (2018), Manchu was mistakenly placed under abjad—"a writing system denoting consonants only" (p. 156)—despite the fact that both consonants and vowels are denoted in Manchu. In fact, the debate over the nature of Manchu writing system has long been focused on whether it should be regarded as syllabary or alphabet. In later sections of this chapter, detailed discussion will be made on its nature. Yet, prior to that, the fundamental information on Manchu script, symbol–sound mapping, and symbol blocks will be introduced first.

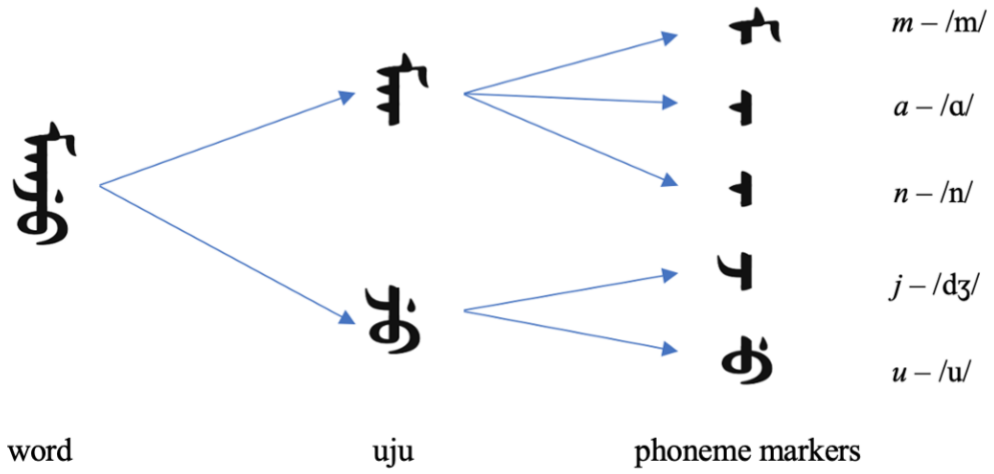
1.6.2 Biscripts of Manchu

Two scripts exist in parallel for Manchu writing system—the traditional script and the Latin script. The traditional script, as previously introduced, was created by adapting Mongolian script in the 1590s and updated in the 1630s. In the traditional script, the basic written symbol is considered *uju*, which is a symbol block in one of the six syllabic structures

of V, VV, VC, CV, CVV, and CVC.[4] For example, in Figure 2, there are two uju <MAN-> and <-JU> that constitute the word <MANJU>.

Figure 2

Uju Symbol Blocks, Phoneme Markers, and Latin Letters in the Word <MANJU>



Approximately, there are over 1,400 uju in Manchu, which were grouped into 12 divisions named *the Twelve Sections of Manchu Syllabary* (满文十二字头). The 1,400ish uju were considered the basic orthographic units for assembling Manchu words and loanwords (An, 1996). The total the number of ujus in Manchu, however, is estimated up to 5,600, because an uju has four variants in isolated, initial, medial, and final forms, if any.[5] The first section of the Twelve Sections of Manchu Syllabary, which consists of six V syllables and around 127 CV syllables, is considered the fundamental section (see Figure 3). The

[4] Uju is a Manchu word, meaning "head", "first", "the first month of one of the four seasons", "a large bead at the beginning of a rosary", "beginning", and "one of the twelve Manchu syllabic headings" (Norman, 2013, p. 386). The present study also uses "uju" for the orthographic unit (i.e., a syllabogram) in Manchu writing system.
[5] The number of 5,600 is an estimated figure because some syllables in Manchu have identical written form with its initial, medial, and final variants. For example, the individual and final forms of the syllable <CA> are graphically identical, and its initial and medial variants share the same written form too.

remaining eleven sections are formulated by connecting <I>, <R>, <N>, <NG>, <K>, <S>, <T>, <B>, <O>, <L> and <M> respectively below each uju in the first section.

Figure 3

Isolated Uju in the First Section of the Twelve Sections of Manchu Syllabary [6]



[6] Retrieved from https://www.omniglot.com/writing/manchu.htm on 13 February 2022.

With respect to syllable types in the Manchu syllabary, the 1,400ish syllables can be categorised into six types of V, VV, VC, CV, CVV and CVC (An, 1996). The distribution of the syllable types is provided in Table 2.

Table 2

Syllable Distribution of the Isolated Uju in the "Twelve Sections of Manchu Syllabary"

Syllable types	Sections												Total No.
	1 –	2 -I	3 -R	4 -N	5 -NG	6 -K	7 -S	8 -T	9 -B	10 -O	11 -L	12 -M	
V	6												6
VV		6								6			12
VC			6	6	6	6	6	6	6		6	6	54
CV	125												125
CVV		118								118			236
CVC			106	118	118	106	106	106	106		106	106	978
Total No. [a]	131	124	112	124	124	112	112	112	112	124	112	112	1411

Note. The table is based on the "Twelve Sections of Manchu Syllabary" in the Qing textbook of《清文启蒙》from the Waseda University library.

[a] This is only the number of uju in their isolated forms, which is but one of the four positional variants an uju may have.

Obviously, the fact that Manchu has an "extensive" inventory size (Nag, 2007) of over 1,400 uju with each having four positional variants is challenging for any learner, which was no exception for the European scholars and visitors to Qing who were more familiar with the Latin alphabet. To address the challenges to learning Manchu, they analyzed uju and found that they are further divisible into sub-units which are in correspondence with vowels and consonants. For example, as Figure 2 shows, the traditionally scripted word <MANJU> comprises two uju <MAN-> and <-JU>, which can be further divided into five components

of <M>, <A>, <N>, <J>, and <U> corresponding with the sounds of /m/, /a/, /n/, /dʒ/, and /u/. Since no term was given to these components in previous literature, the present study named them *Manchu phoneme markers* after similar description for the akshara symbol blocks (Nag, 2007, 2017). By synthesising the previous documentation on Manchu orthography (Gorelova, 2002; He, 2009; Li, 2010; Norman, 2013), this study identified 82 Manchu phoneme markers. The Western scholars adapted the Latin alphabet to transliterate the phoneme markers that denote the sounds of vowels and consonants. This method consequently introduced into Manchu the Latin script, which includes six vowel letters (*a, e, i, o, u, ū*) and eighteen consonant letters (*n, k, g, h, b, p, s, š, t, d, l, m, c, j, y, r, f, w*).

1.6.3 Positional Variation

The Latin letters in Manchu do not have positional variants and the ways in which they are written remain unchanged regardless of its sublexical position. An uju, however, has four positional variants (i.e., the isolated, initial, medial, and final forms) depending on whether it is used in isolation or at the beginning, middle, or end of a word (see Figure 4).

Figure 4

Four Positional Variants of the Uju <MAN>



As for Manchu phoneme markers, since they are graphically different from one another, the present study regarded each individual Manchu phoneme marker as a distinct orthographic unit, although they could also be seen as allographs representing the same phoneme. For example, in Figure 5, although the three phoneme markers all representing the sound of /ɑ/ (i.e., allographs) and could be seen as the positional variants of <A>, this study, however, regarded them as three graphically distinct and individual phoneme markers. In this sense, Manchu phoneme markers, like letters in the Latin script, do not have positional variants, because one of the variants remains the same for that specific position.

Figure 5

Three Phoneme Markers Corresponding with /a/



1.6.4 Symbol–Sound Mapping in Manchu

The symbol–sound mapping between an uju and a syllable is on a four–to–one basis, because an uju has four positional variants. Moreover, the symbol–sound mapping of Manchu phoneme markers is more varied. Out of the 82 Manchu phoneme markers originally identified by this study, six are mapped onto two sounds and two onto three sounds, with the remaining 74 onto one sound only. Regarding the Latin alphabet, it has a one–to–one relation with sound. For instance, the letter *m* is mapped onto /m/ and the letter *a* onto /ɑ/. Thus, its symbol-sound mapping is consistent and transparent. Table 3 summarises the different mapping relations.

Table 3

Symbol–Sound Mapping in Manchu

Scripts	Symbols	Symbol–sound mapping
Traditional script	Uju	uju–syllable: four–to–one mapping
	Phoneme marker	phoneme marker–phoneme: 74 markers in one–to–one mapping 6 markers in one–to–two mapping 2 markers in one–to–three mapping
Latin script	Latin letter	Latin letter–phoneme: one–to–one mapping

1.6.5 Manchu Hergen, Symbol Blocks, and Phoneme Markers

A Manchu *hergen* is a word, which is formulated by stacking up uju or phoneme markers. There are gargata hergen (i.e., monosyllabic word or uju) and holboho hergen (i.e., multisyllabic word). For example, in Figure 3 above, the hergen "MANJU" can be regarded as assembled by an orderly piling of the uju <MAN-> and <-JU>. Alternatively, it can also be seen as assembled by the phoneme markers of <M>, <A>, <N>, <J> and <U>, depending on one's understanding of what counts as the basic symbol in Manchu. Typically, the comprising phoneme markers in an uju or a hergen are connected to each other by vertical lines. But in Manchu *ligatures* (e.g., <BA>, <PE>, <KE>, <GI>, <HU>, etc.), the consonant marker and the vowel marker are written into each other and nested together.

Another perspective to understanding a Manchu hergen is to regard it as a "whole character" as in Chinese orthography. The mnemonic rhyme provided below, which was obtained by the author of this thesis from observing a Chinese university Manchu class in 2016, vividly describes the formation of Manchu characters. It reads

中间一根棍儿，

两边都是刺儿。

加上圈和点儿，

就是满文字儿。

A branch in the middle,

Twigs on both sides.

Add dot and circle,

A Manchu character gets written.

[Translated by Li and Nag from Chinese]

A branch (or 棍儿 “stick”) refers to the vertical “spine” going through the centre of a Manchu hergen, and twigs (or 刺儿 “prickle”) are the comprising Manchu phoneme markers connected to the spine on both sides. The branch and twigs form the main component of a Manchu hergen and when diacritic dots and/or circles are added, a Manchu hergen (or 满文字儿 “Manchu character”) takes form. The name “满文字儿” (i.e., Manchu hergen or “Manchu characters”) indicate that traditionally scripted Manchu were regarded in analogy to Chinese characters which are comprised by strokes. The present study, by drawing on the strokes existing in Chinese and Mongolian, identified 12 basic strokes of *traditional Manchu script*, which are the “head” (头), the “tail” (尾), the “tooth” (牙), the “circle” (圈), the “dot” (点), the “falling” (撇), the “belly” (肚), the “curve” (弧), the “horizontal” (横), the “vertical” (竖), the “bend” (折), and the “hook” (钩) (see Table 4)[7]. In writing a Manchu hergen, all strokes are connected to the “spine”, an invisible line vertically going through the middle of a hergen. Table 4 presents the strokes.

[7] Mikael A. Thompson also contributed by listing the basic strokes in Mongolian (personal communication, July 3, 2018).

Table 4

Names of the Twelve Basic Strokes in Traditional Manchu Script

Stroke name in English	Stroke name in Chinese, transliterated by IPA	Stroke name in Manchu, if any	Examples from Manchu
Crest	头 /tʰoʊ/	*uju*	
Tail	尾 /weɪ/	*uncehen*	
Tooth [a]	牙 /ja/	*a*	
Circle	圈 /tɕʰɥœn/	*fuka*	
Dot	点 /tjɛn/	*kongki*	
Falling [b]	撇 /pʰjɛ/	*šaka*	
Belly	肚 /tu/		
Curve	弧 /xu/		
Horizontal	横 /xɤŋ/		
Vertical	竖 /ʂu/		
Bend [c]	折 /ʈʂɤ/		
Hook [d]	钩 /koʊ/		

[a] There are “short tooth” and “long tooth”. One end of a “tooth” is connected to the spine.

[b] Falling is slightly arched but long tooth is not.

[c] Combined stroke. The finishing of a bend does not taper in writing.

[d] Combined stroke. The finishing of a hook tapers in writing.

1.6.6 Debate on the Typology of Manchu Writing System

The debate on what writing system Manchu belongs to is derived from the controversy over what counts as the basic written symbol in Manchu. A popular argument amongst the Chinese scholars is that the traditional Manchu script uses uju as the basic

symbol. For example, An (1996) argued that uju is the basic orthographic unit in Manchu, which aligns with the idea held by Wang (1996) that uju, which are not further divisible, have been used for hundreds of years in Qing as the basic units in Manchu writing and instruction. What was indicated by Wang should not be misunderstood as suggesting that a Manchu orthographic syllable is not further divisible into phoneme markers per se, but that uju were considered and employed historically by the Qing people as the smallest and thus indivisible unit of orthography. This understanding is in line with the claim that, in the eyes of the Qing people, the traditional script was seen as a set of symbol blocks representing phonological syllables, rather than as blocks with each being made up of a combination of phoneme markers (T. Ma, 2014).

On the other hand, however, the European scholars argued that the Manchu phoneme marker (or in their words the "Manchu letter"), which is not subdivisible into smaller orthographic units and representative of either vowel or consonant sound, is the basic symbol in Manchu writing system (see Saarela, 2015 for a detailed discussion). For instance, Rémusat (1820, p. 106), a French scholar, had made the assertion that "[n]o writing system in the world is composed of simpler and more regular signs [than that of the Manchus]", yet "...this people had no conception of the alphabet, ...waiting instead for people to come from the other side of the world to tell them that they had used letters for two hundred years without knowing it" (cited in Saarela, 2015, pp. 509–510).

What has been introduced so far about the Manchu writing system can be summarised as follows. There is consensus that Manchu language uses biscripts—the traditional script and the Latin script. However, no agreement has yet been reached with respect to which writing system Manchu belongs to when the traditional script is used—some argue that it is syllabary, with the basic written symbol being uju, whereas others maintain that it is alphabet, taking the Manchu phoneme marker as the basic unit for writing.

1.6.7 A "Contextualised" View on Understanding Manchu Writing System

The present study takes a "contextualized" view to understand Manchu writing system. That is, be it syllabary or alphabet, the understanding should rely on the way in which instruction occurs in an authentic classroom context. Being "contextualised" indicates that the nature of a writing system should be conceptualised, first of all, in a time frame in history or at present. Secondly, since this study examines language education in a university setting, all the teaching and learning practices should be understood and analysed in relation to this educational context. Therefore, a contextualized view considers temporal context (i.e., the time in history or at present) and spatial context (i.e., the space such as classrooms).

Following this train of thought, the nature of Manchu writing system could now be examined in two temporal–spatial contexts—the classroom in the Qing period and in present China. In the Qing classroom, it was uju that were regarded as the basic orthographic units in Manchu instruction (T. Ma, 2014). Through this lens, Manchu could be seen as a writing system of syllabary. However, when it comes to the classroom instruction in current China, the picture becomes more complex.

1.7 Chapter Summary

This chapter introduced some fundamental concepts of Manchu that are essential for the literature review and methodology in the subsequent chapters. To begin with, the focus of this study was described as an investigation into the relations between the complexity of Manchu writing system and the novice learners' decoding performance of uju symbol blocks. Following that was a brief historical account of the Manchu people and their language, which, although enjoyed prestigious status in the Qing dynasty (1644–1911), has become critically endangered with less than a hundred native speakers in current China. Then, this chapter continued into a discussion of the nature of Manchu writing system, arguing that it would be too simplistic a view to make an *either/or* division by regarding Manchu as an

alphabet or syllabary, at least not in the case of applied or educational linguistics research. In contrast, it would be more conducive to the present study on Manchu writing system if the understanding and discussion could be framed within an authentic educational context, which proposes a "contextualised" view on Manchu writing system investigation. Moreover, basic terms were introduced in this chapter, including the types of Manchu symbols, symbol–sound mapping, the Manchu phoneme markers, the Manchu symbol blocks called uju, and the Manchu written word called hergen. In terms of symbols, Manchu writing system adopts both the traditional script and the Latin script. The latter contains six vowel letters (*a, e, i, o, u, ū*) and eighteen consonant letters (*n, k, g, h, b, p, s, š, t, d, l, m, c, j, y, r, f, w*). The former, however, includes 1,400ish uju in V, VV, VC, CV, CVV and CVC syllables which are distributed in the "Twelve Sections of Manchu Syllabary". Moreover, the symbol–sound mapping relation in Manchu is complex, including the four–to–one uju–syllable mapping and one–to–multiple phoneme marker–sound mapping. Finally, Manchu hergen can be seen as formulated by stacking the symbol block(s) of uju and each uju as constructed by phoneme markers.

Given the above introduction on the history and orthography of Manchu, it is important to examine this complex writing system so that we can gain a deeper understanding of the world's writing systems by exploring the cognitive mechanisms underpinning reading these systems. The first reason to examine the Manchu writing system is to uncover the processing mechanism of the visually complex Manchu symbol blocks. A polysyllabic word in Manchu such as *bontoholobumbi* (see Figure 6) is made up of a string of symbol blocks with each symbol block (*bon.to.ho.lo.bum.bi*) having a high degree of visual complexity. Thus, it is of theoretical interest to find out the ways in which a visually complex symbol block is processed by the readers so much so that the intricate details entangled and nested together can be distinguished and the word read successfully. In addition, the Manchu writing

system can be theoretically considered as characterized by mixed orthography-phonology mapping (Nag, 2007, 2017a, 2021). That is because the mapping can exist at different levels between uju and syllable (e.g., <NA> – /na/) or alternaitvely between phoneme marker and phoneme (e.g., <N-> – /n/ and <-A> – /a/). Therefore, it will be of great signficance to investigate the cognitive mechanisms of reading Manchu in terms of the orthographic unit(s) adopted for decoding by the novice learners. To achieve these goals, serious endeavour should be given to the under-investigated orthographies among a plethora of human writing systems. Therefore, this study undertook this endeavour and started from examining the visual and mapping complexity of the Manchu writing system.

Figure 6

A Visually Highly Complex Manchu Word



Note. The word is *bontoholobumbi*, meaning “fail to make something happen as planned”.

Chapter 2 Literature Review

2.1 Chapter Overview

This chapter reviewed the relevant theories and studies that the present study has drawn on to understand the ways in which the two perimeters of (a) visual complexity and (b) symbol–sound mapping complexity are related to Manchu word reading. First, the framework of graphic complexity (Chang et al., 2018) is helpful in understanding how the visual complexity of Manchu symbol blocks affects reading. Second, the orthographic depth hypothesis (Frost, Katz, & Bentin, 1987; Katz & Frost, 1992), the psycholinguistic grain size theory (Ziegler & Goswami, 2005), the orthographic breadth hypothesis (Nag, 2007, 2017), and intra-symbol processing (Nag, 2022) provide related perspectives of and beyond symbol–sound mapping to understand learning to read the Manchu orthography. This chapter also identified some gaps in the existing literature, based on which the research questions were developed.

2.2 Visual Complexity and Word Reading

In the existing literature on reading research, no definition was explicitly given to the concept of visual complexity, although it was conventionally referred to as "the level of detail or intricacy" contained in an image (Forsythe, Mulhern, & Sawey, 2008, p. 116). The present study tentatively defined visual complexity as the level of intricacy of the visual features contained in a symbol block that could affect the cognitive demand of readers in the process of reading. While visual complexity was an important dimension appearing to affect reading among beginning and skilled readers, previous studies were carried out on a limited number of orthographies (Daniels & Share, 2018). For example, in a study of the Indic Kannada, Nag, Snowling, Quinlan, and Hulme (2014) reported a moderate, negative correlation between akshara visual complexity and recognition accuracy, suggesting that visual complexity poses challenge in learning to read the akshara orthography. Moreover,

Abdelhadi, Ibrahim, & Eviatar (2011) also reported that the target diacritic was identified slower in Arabic letter than in Hebrew letter, which indicated that the visually more complex Arabic letter results a processing demand in reading. A similar result was also reported in reading Chinese (e.g., Yu, Zhang, Priest, Reichle, & Sheridan, 2018), in which the visually more complex characters took longer for the readers to be identified. All these lines of evidence seemed to suggest that visual complexity has an inhibitory effect on reading accuracy and speed, but it is uncertain whether the findings would be convergent if additional studies are conducted on other visually complex orthographies such as Manchu.

To measure the visual complexity of graphs across writing systems, Chang and colleagues (2018) developed a multidimensional model by combining the measures of perimetric complexity (PC), the number of simple features (SF), the number of connected points (CP), and the number of disconnected components (DC). PC is defined as the ratio of the squared perimeter of the foreground to the foreground area (Pelli, Burns, Farell, & Moore-Page, 2006; Watson, 2012). The measuring formula of PC is

$$PC = \frac{P^2}{A4\pi}$$

where *P* is the summation of the inside and outside perimeters of a foreground and *A* is the area, with 4π as the nominalizing constant. For example, if the Manchu symbol block < ᡮ > /tʃɑ/ has a 1,041-pixel perimeter and a 10,048-pixel area, its PC is 8.576 (= 1,041 × 1,041 / 10,048 / 4π). PC has the advantages of easy calculation, size-invariance, and additivity in measuring the visual complexity of symbols, and was thus applied to various orthographies such as Arabic, Armenian, Chinese, Devanagari, English, and Hebrew (e.g., Liu, Chen, & Wang, 2016; Pelli et al., 2006; Yin & McBride, 2015). However, the assessment on visual

complexity by the objective PC and the subjective human observers was inconsistent, as PC was reported to be only weakly correlated with the observer's rating of the visual complexity of shapes (Arnoult, 1960). Yet, in a study by Nag and colleagues (2014), the measure of pixel count (PX, the amount of "ink" in a printed symbol) was developed as an indicator of the visual complexity of the akshara blocks. They reported a high correlation ($r = .711$, $p < .001$) between PX and the teacher's judgement on akshara visual complexity, suggesting that PX can be a useful index of "visual feature density" (p. 322).

Moreover, the visual complexity of a symbol could also be measured by the number of simple features (SF). In the present study, SF refers to the distinct written element that can be discriminated from others, such as a line segment, a curve, a dot, and a circle (see Table 4). For instance, the Manchu symbol block <  > comprises six SFs, which, from top to bottom, are a "short tooth", a first vertical line, a second vertical line, a slanted line segment, a third vertical line, and a "tail".[8] The measure of SF is sensitive to the number of simple features combined, but not to other properties such as length, width, and thickness (Chang et al., 2018). SF was used to measure visual complexity in studying Chinese characters (e.g., Ma & Li, 2015; Wu, Zhou, & Shu, 1999; Yu et al., 2018) and Japanese kanji (e.g., Tamaoka & Kiyama, 2013). Results of these studies revealed that characters and kanji with higher SF required more time to be recognized, suggesting that more strokes imposed higher cognitive demands for the readers.

Furthermore, closely related to SF is the dimension of the number of connected points (CP), which refers to a joining of at least two simple features (Chang et al., 2018). For example, the number of CP in the Manchu symbol block < > is four.[9] CP was found to be the most important characteristic in letter and word recognition in English (e.g., Fiset et al.,

[8] The "short tooth" is located at the top of the first vertical line. It is less obvious because of the font used here.
[9] The fourth CP is at the conjunction of the third vertical line and the "tail".

2008; Lanthier et al., 2009), which may be due to the information of "continuity" carried by the CP. This explanation can be supported by the neural net model (Hummel & Biederman, 1992), which hypothesised that the simple features responded by the first layer of cells are forwarded to the second layer of vertex detector cells that determine the more complex relation of conjunction between the simple features. Yet, what remains unclear about CP is whether it also affects symbol recognition in other non-alphabetic writing systems.

The final dimension measuring visual complexity is the number of disconnected components (DC). DC is defined as a simple feature that is not linked to other feature in a set (Chang et al., 2018). For example, the Manchu symbol block <  > /tʃɑ/ has one disconnected component while <  > /dʒə/ has two. DC is sensitive to the discontinuity information or "dispersion" (Gibson, 1969).

To briefly summarize, the above review has identified five measures assessing visual complexity, namely, pixel count, perimetric complexity, the number of simple features, the number of connected points, and the number of disconnected components. While they were all reported to be related to the recognition of object, letter, and word, some gaps still existed in previous studies. One is that there was little empirical evidence suggesting how the five measures are associated with word reading in visually complex orthographies. Moreover, it remains unknown how much variability in the reader's Manchu words decoding performance can be predicted by the five measures. Therefore, this study attempted to answer the following two research questions (RQ) with respect to the visual complexity of Manchu orthography.

RQ 1-1: To what extent does the visual complexity of Manchu symbol blocks associate with the learners' decoding accuracy and speed?

RQ 1-2: In what ways do the measures of the visual complexity of Manchu symbol blocks predict the learners' decoding accuracy and speed?

For RQ 1-1, it was hypothesized that the five measures of visual complexity (PX, PC, SF, CP and DC) are negatively associated with novice learners' Manchu symbol blocks decoding accuracy and speed. As for RQ 1-2, the present study hypothesized that SF may be a significant predictor for novice learners' uju symbol block decoding performance. This hypothesis was made based on the classroom observation in piloting that the instruction on Manchu uju writing was focused on the configuration of strokes, which were found to be associated with reader's performance in recognising other visually dense symbol blocks such as Chinese characters (Ma & Li, 2015; Wu, Zhou, & Shu, 1999; Yu et al., 2018) and Japanese kanji (Tamaoka & Kiyama, 2013).

2.3 Symbol–Sound Mapping Complexity and Word Reading

Learning to read is fundamentally a process of matching the orthographic symbols to phonological units in a language, which requires the first step in reading to be the acquisition of symbol-sound mapping (Ziegler & Goswami, 2005). An influential theory in reading acquisition is the orthographic depth hypothesis (ODH; Frost et al., 1987; Katz & Frost, 1992). ODH suggested that orthographies can be seen as a continuum of varying levels of letter–sound consistency, and the reading of the inconsistent (deep) orthography is more difficult than that of the consistent (shallow) orthography. For example, in a cross-linguistic study on European orthographies, Seymour, Aro, and Erskine (2003) found that the reading of deep orthographies (e.g., English) was at a slower pace than those shallow orthographies (e.g., Finnish). In a similar vein, the symbol–sound mapping relations was identified as the "consistency" problem in the psycholinguistic grain size theory (Ziegler & Goswami, 2005). The orthographic consistency of the symbol–sound mapping of an orthography affects

reading acquisition. For example, French children were found to read approximately 87% of a word list correctly (Sprenger-Charolles & Siegel, 1997) and their Italian counterparts read around 94% of the real words correctly (Cossu, Gugliotta, & Marshall, 1995). In contrast, children learning to read English, which has an inconsistent grapheme–phoneme mapping, scored less than 70% in word reading (Frith, Wimmer, & Landerl, 1998). In addition, the psycholinguistic grain size theory also suggested that the orthographic consistency affects granularity, which leads the readers to develop different decoding strategies. The smaller grain sizes in some European languages tend to be relatively more inconsistent; thus, the young learners adopted multiple recoding strategies that use both small and large grains to address their reduced reliability for word decoding. For example, Goswami (1986) reported that English children used, in addition to letters (small grain size), orthographic chunks (large grain sizes) corresponding to rimes to read novel English nonwords. Similarly, Brown and Deavers (1999) also reported that children and adults used both small and large grain sizes to read English nonwords, suggesting that they have developed a sublexical grapheme-phoneme decoding strategies using both grain sizes in parallel. To extend the symbol–sound mapping conceptualisation in the psycholinguistic grain size theory (Ziegler & Goswami, 2005), Nag (2007, 2017a, 2021) introduced the construct of "mixed mapping" which captures the complex nature of orthography–to–phonology correspondence in the akshara writing system. For example, in the Kannada word /koLi/ the akshara–syllable mapping is in the straightforward one–to–one relation (<ko.Li> – /ko.Li/). However, in the word /pakshi/, the akshara <kshi> is mapped onto both the coda of the first syllable /pak/ and the entire second syllable /shi/. This mixed mapping suggests that the knowledge of variability in symbol–sound and sound–symbol correspondence is critical for decoding efficiency (Nag, 2021).

While ODH and psycholinguistic grain size theory are influential frameworks, they are deeply rooted in the Anglophonic view of letter–sound consistency with little attention

giving to the non-alphabetic orthographies; thus they fail to justify the diversity of the world's scripts (Daniels & Share, 2018). In an alternative framework of the orthographic breadth hypothesis (OBH; Nag, 2007, 2017a), orthographies are considered along the continuum of inventory size (the number of symbols in a set), with contained and extensive orthographies at the two ends. Contained orthographies have less than 50 symbols (e.g., English) while extensive orthographies have up to thousands of symbols in a set (e.g., Chinese = 5,000-6,000: Nag, 2022; Manchu = 1,400: Saarela, 2014). An implication of OBH is that there could be a trade-off between the cognitive demands and the inventory size in orthography learning (Nag, 2011). When learning an extensive orthography, the demand on memory is high but there is reduced confusion on symbol–to–sound correspondence. In contrast, in learning a contained orthography, memory demand is relatively low but the complex way that symbols represent sounds could consume more cognitive resources.

Manchu can be regarded as an extensive orthography when the uju (= 524 for the fundamental first section of the syllabary; = 5,600ish for the entire syllabary) is considered the basic written symbol, but a relatively more contained orthography when phoneme markers (= 82) are seen as the basic orthographic unit. Arguably, given the difference in inventory size (i.e., the size of the symbol repertoire) depending on a focus of uju versus phoneme markers, the cognitive demand may be considered to be relatively low for learners when Manchu is learnt as a contained orthography. If Manchu is learnt as an extensive orthography, the demand on memorising the large number of uju ranging from hundreds to thousands is significantly high. In addition, there can be great confusion at the level of the phonology because of the multiple mapping relations between the four positional variants of an uju and a syllable (see Figure 4). On the other hand, if Manchu is learnt as a contained orthography, the demand of remembering the 82 phoneme markers by rote is much lower. Moreover, as a majority of the phoneme markers (74 out of 82) are in a transparent one-to-

one mapping with sound, it would cause less confusion for the learners to decipher the phonology encoded in phoneme markers.

Moreover, in learning extensive orthographies, there may be more instances in which the symbol blocks are not explicitly taught due to the large inventory size (e.g., Chinese: McBride, 2016; Kannada: Nag, 2014), but the learners nevertheless acquire their internal details (Nag, 2007). This acquisition could be accounted for by the model for orthographic learning (Nag, 2011, 2017a, 2017b), in which a global approach (representing a symbol block as an undifferentiated whole), an analytic approach (processing the internal details of a singleton symbol block and decoding it into component parts), and a strategic approach (recognising the phonology encoded by symbol blocks in a word context) were all involved in orthographic learning. It was hypothesised that these approaches provide strategies available for readers in symbol block recognition (Nag, 2017a). The global approach serves for recognising undecomposable symbol blocks, the analytic approach is used to address decoding problems when encountering new, unfamiliar, or low-frequency symbol blocks, and the strategic approach applies the knowledge of basic orthographic unit to disambiguate the phonology encoded in a word.

In addition, the construct of intra-symbol processing is also helpful in explaining the learning process of visually complex and extensive writing systems (Nag, 2022). The intra-symbol processing of symbol blocks presupposes some key architectural principles of the orthography, such as the visual distinctiveness, the positional sensitivity, and the decomposability of symbol blocks (Nag, 2021). These design principles make the symbol blocks analysable in intra-symbol processing, which could gain a subset of symbols where intra-symbol processing facilitates decoding efficiency. An important dimension in intra-symbol processing is the visuo-spatial characteristic of symbols (Nag, 2022). In many writing systems, the symbols are not linearly arrayed along a vertical or horizonal axis (Daniels &

Share, 2018). The vowel markers or diacritics can be systematically positioned above, below, before, after, around, or to either sides of the consonant base (e.g., Kannada: Nag, 2022; Tamil: Nag & Narayanan, 2019; Thai: Winskel, 2009) The nonlinear complex visuo-spatial arrangement of the symbols requires visuo-spatial process for word processing (e.g., Nag, 2014; Vaid & Gupta, 2002). In addition, visuo-spatial skills (i.e., the perception, memorisation, and manipulation of the two-dimensional spatial designs) can be enhanced by reading complex orthographies such as Chinese, because there are strong visual demands for learning Chinese characters (McBride, 2016).

In brief, the above review has identified a multi-writing system informed framework that includes orthographic breath hypothesis (Nag, 2007, 2017a), orthographic learning model (Nag, 2011, 2017a, 2017b), and intra-symbol processing (Nag, 2022). This multi-dimensional framework goes beyond the constructs such as the orthographic depth (Frost et al., 1987; Katz & Frost, 1992) and the psycholinguistic grain size (Ziegler & Goswami, 2005) in accounting for reading acquisition. To investigate the symbol-sound mapping complexity of the Manchu writing system, the present study attempted to examine the process of Manchu uju symbol block reading by answering the following research question (RQ):

> **RQ 2-1**: Do the learners perform differently when they decoded Manchu symbol blocks with phoneme markers in one–to–one and one–to–multiple mapping with sound?

For this research question, the present study hypothesized that the inconsistency in the mapping of phoneme marker to sound has an inhibitory effect on novice learners' Manchu symbol blocks decoding performance. In other words, the uju containing the target phoneme

marker that is mapped onto multiple sounds will be processed with less accuracy and speed than the uju with all comprising phoneme makers corresponding with one sound.

2.4 Instructional Strategies of Orthographic Knowledge across Writing Systems

Further to RQ 2-1 which aimed to identify what orthographic unit(s) (i.e., phoneme marker, uju, or both) was/were adopted by novice learners to decode Manchu uju, the present study also intended to explore the association between grain size choice and the instructional strategies of Manchu orthographic knowledge. Previous studies on alphabetic and alphasyllabic orthographies have found that different instructional strategies were effective for learners of varied literacy abilities. For example, in a study investigating word reading strategies of young native English learners in the US (Treiman & Rodriguez, 1999), children were first taught to pronounce made-up words in three conditions which presented letter-name, letter-sound, and visual cues. Later the children were given a naming task of the nonwords instructed. The results showed that both prereaders and novice readers performed best when the letter name was heard in the spoken word (e.g., BT for "beet"). Furthermore, the novice readers performed better on letter-sound condition (e.g., BT for "bait") than the visual condition (e.g., ʙT for "ham"). The results suggested that readers of English developed an analytic strategy which took advantage of the grapheme–phoneme correspondence in reading rather than memorize words holistically as arbitrary patterns. In a similar study conducted on L2 English children in Hong Kong (McBride-Chang & Treiman, 2003), the children were taught to pronounce made-up English nonwords based on letter-name, letter-sound, and visual cues. The results showed that the second- and third-year kindergartners performed best in reading the nonwords in letter-name conditions, and the third-year children also performed better in letter-sound condition than visual condition. The results suggested that Hong Kong children, as observed in their L1 English counterparts, may also utilise the

analytic strategy by taking advantage of the grapheme–phoneme connection in learning to read English words.

In addition, studies on the alphasyllabic orthographies of Korean and Kannada also provided additional evidence on reading and instructional strategies.[10] For example, Cho, McBride, and Kim (2020) examined the effects of four types of instructional methods (i.e., whole Gulja, alphabet letter, CV-coda and copying) on Hangul reading in emergent, novice, and good readers. They found that the holistic Gulja instruction facilitated emergent readers and the analytic instruction of CV-coda syllabic units benefited novice learners. In a study on Kannada (Nag & Snowling, 2012), the more capable readers were found to perform better on phoneme awareness tasks. Nag and Snowling (2012) proposed that an analytic process may be involved akshara reading, in which the better readers decomposed the symbol blocks into the comprising phoneme markers. Nag (2011, 2017a, 2017b) further developed a model that involves global, analytic, and strategic approaches in decoding. The global approach, which is similar to the holistic method of reading introduced above, treats symbol blocks as singleton wholes without deciphering them into parts. In the analytic approach, symbol blocks are decomposed into constituent markers and decoded based on symbol-sound mappings. However, only applying the global and analytic approaches is not sufficient for symbol blocks reading, and that is when the strategic approach comes in. Being strategic in decoding means to select the more appropriate approach from the global and analytic methods to disambiguate the phonology that an orthographic unit represents within a word context. Application of the approaches depends on the learner's level of knowledge on orthography (Nag, 2011). The global approach is used for familiar symbols that can be recognised by automaticity and the analytic approach for unfamiliar or new words. The

[10] Although Korean writing system is considered alphabetic, Cho, McBride and Kim (2020) claimed that Hangul is called an alphasyllabary because the Hangul syllable is printed in square shape similar to the symbol blocks of the Indic scripts of South Asia.

global and analytic approaches can be merged and used strategically to decode symbol blocks which are visually and syllabically more complex.

The above review highlighted the effectiveness of instructional methods on the strategies employed for reading and writing and across orthographies. While the lines of evidence seemed divergent to some extent, the findings were consistent in a sense that the learners with more advanced literacy abilities appeared to adopt both holistic and analytic strategies in literacy learning. Thus, it was hypothesized that since Manchu symbol blocks were observed to be primarily instructed analytically in class, novice readers of Manchu may take advantage of the analytic strategy in uju reading. It is important to note that although participants of the study have no prior knowledge of Manchu orthography before university and should thus be regarded as beginning readers by definition, it is more appropriate to categorise them as novice readers because they are literate and skilled in reading at least two languages of Chinese and English. To further understand the results of RQ 2-1, this study was also interested in examining the association between the learner's cognitive processing of uju and the teacher's approaches applied in Manchu orthography instruction. Research question (RQ) in this regard was as follows:

> **RQ 2-2**: In what way are the results of the learners' decoding accuracy and speed on Manchu symbol blocks with different phoneme marker–sound mappings related to the instruction on Manchu orthography?

2.5 Chapter Summary

This chapter reviewed relevant theoretical frameworks and studies which are helpful in understanding the association between the two parameters of (a) visual complexity and (b) symbol–sound mapping complexity and Manchu symbol block reading. In terms of visual

complexity assessment, this review identified the five measures of pixel count, perimetric complexity, the number of simple features, the number of connected points, and the number of disconnected components (Chang et al., 2018; Nag et al., 2014), which were reported to be associated with word recognition. Studying these parameters in Manchu is important because, empirical studies examining their association with reading the visually complex orthographies are found in only a limited number of orthographies, such as, Kannada (Nag et al., 2014), Arabic (Abdelhadi et al., 2011), and Chinese (Yu et al., 2018). In addition, while the orthographic depth hypothesis (Frost et al., 1987; Katz & Frost, 1992) and Psycholinguistic Grain Size Theory (Ziegler & Goswami, 2005) have provided a critical perspective of letter-sound mapping across orthographies, their theorising uses data largely from the alphabetic writing systems. Studies carried out on other writing systems and explanation made within alternative theoretical frameworks, such as the orthographic breadth hypothesis (Nag, 2007, 2017) and intra-symbol processing (Nag, 2022), provide a more comprehensive theoretical framework for the study of the Manchu Writing System. In addition, this chapter also reviewed previous studies with respect to the teaching strategies adopted in orthographic knowledge instruction across orthographies such as the alphabetic English (McBride-Chang & Treiman, 2003; Treiman & Rodriguez, 1999) and the alphasyllabic Korean (Cho et al., 2020) and Kannada (Nag, 2011, 2017a, 2017b; Nag & Snowling, 2012). The findings were consistent in that the learners with more advanced literacy abilities appeared to adopt both holistic and analytic strategies in literacy (reading and writing) learning. All in all, a study of the Manchu Writing System promises to provide another line of evidence for a more universal theory of reading is to be built, as reading the world's diverse orthographies may be underpinned by different cognitive mechanisms.

Chapter 3 Methodology

3.1 Chapter Overview

This chapter begins with the discussion on why mixed methods research (MMR) is relevant to the present study. It then considers the epistemological and ontological issues in MMR, and decides to adopt a pragmatic approach (Morgan, 2007) in examining learning to read the Manchu writing system. Afterwards, a multidimensional MMR framework is introduced, in which the present study was conducted in a concurrent multilevel design with quantitative as the principal tool for data collection. Following that is the discussion and decision on sampling strategy and appropriate sample size. Finally, the strategies and techniques for collecting and analysing data are presented and the methodological limitations are discussed.

3.2 A Mixed Methods Research Design

3.2.1 Rationale for Using Mixed Methods Research

Mixed methods research (MMR) is appropriate and relevant to the present study because it fulfilled the needs of answering the research questions and meeting the research objective. MMR is rooted in the premise that research questions determine methods of inquiry (Tashakkori & Newman, 2010). Therefore, decisions should be made about which sorts of research questions are best answered by quantitative research method and which by qualitative method (Bryman, 2012). To recall, the research questions this study attempted to answer are as follows:

RQ 1-1: To what extent does the visual complexity of Manchu symbol blocks associate with the learners' decoding accuracy and speed?

RQ 1-2: In what ways do the measures of the visual complexity of Manchu symbol blocks predict the learners' decoding accuracy and speed?

RQ 2-1: Do the learners perform differently when they decoded Manchu symbol blocks with phoneme markers in one–to–one and one–to–multiple mapping with sound?

RQ 2-2: In what way are the results of the learners' decoding accuracy and speed on Manchu symbol blocks with different phoneme marker–sound mappings related to the instruction on Manchu orthography?

The above research questions are explicitly connected to different research methods. Since RQs 1-1, 1-2, and 2-1 intended to examine the visual and mapping complexity in relation to word decoding performance from the perspectives of correlation, prediction and comparison, quantitative research is more appropriate to address them. However, RQ 2-2 aimed at explaining the results of RQ 2-1 using data obtained from classroom observation, it thus requires a qualitative research method.

With respect to the research objective of this study, it intended to understand the "different facets" (Greene et al., 1989, p. 266) of learning to read the complex Manchu writing system. It used quantitative Tests 1 and 2 on Manchu symbol blocks recognition in naming tasks to examine the visual and mapping complexity of Manchu in relation to participants' word decoding performance respectively, which was combined with a qualitative classroom observation to explain the results of Test 2 on mapping complexity. This mixed method helped to enhance the interpretability by means of elaborating and illustrating the results from one method with results from the other, and the inherent method strengths to be capitalised upon and the intrinsic biases offset somewhat (Greene et al., 1989). However, it is worth noting that the quantitative and qualitative methods are not of equal status in this study, with the former taking relatively more weight in meeting the research objective of examining the association between the complexity of Manchu writing system

and word decoding. Thus, this study is a quantitatively driven design (Johnson & Christensen, 2014), in which the quantitative perspective of conceptualising takes the central role with classroom observation data elaborating the quantitative findings.

3.2.2 Epistemological and Ontological Considerations on Mixed Methods Research

Among the issues arose from combining quantitative and qualitative research, the most critical one is the paradigm argument (Bryman, 2012; Morgan, 1998). This controversy is derived from Kuhn's (1962) concept of paradigm, which refers to a system of beliefs and perspectives held by a community of researchers that influence what they should study, what methods they should use, and how they should interpret the results (Bryman, 1988; Morgan, 2007). Paradigm is featured by its incommensurability, meaning that different paradigms are inconsistent because of their divergent philosophical assumptions and methods (Bryman, 2012). In other words, different underlying epistemological and ontological assumptions would produce conflicts between different paradigms.

With respect to mixed methods research, Morgan (1998) has pointed out that most research that claim to have adopted mixed methods disregard paradigm concerns, thus failing to address the deeper philosophical issues. This could lead to serious consequence because when the quantitative and qualitative approaches are integrated, problems may arise since they "reply on very different assumptions about both the nature of knowledge and appropriate means of gathering knowledge; hence the kinds of information that they produce are often incommensurate" (Morgan, 1998, p. 363). As a result, the inherent incommensurability in epistemological assumptions (e.g., postpositivist versus interpretivist paradigms) and ontological assumptions (e.g., objectivist versus constructivist paradigms) have inevitably resulted in paradigm wars (Tashakkori & Teddlie, 2003), in which the advocates of quantitative and qualitative approaches hold an either-or position that only one approach should be used in research. Moreover, this linkage of methods and philosophical beliefs also

lead to “incompatibility thesis” of using only qualitative or quantitative method (Howe, 1988), and it is against such a backdrop that the newer and alternative thinking on MMR was developed.

To overcome the problem derived from the incompatibility thesis, the present study adopts the view of methodological pluralism (Teddlie & Tashakkori, 2010), which argues that MMR can serve as a “big tent” (p. 9) to encompass a broad range of paradigms which could potentially be utilized in research. More specifically, among the alternative paradigms of MMR (e.g., Biesta, 2010; Greene & Hall, 2010; Maxwell & Mittapalli, 2010; Mertens, Bledsoe, Sullivan, & Wilson, 2010), this study takes the pragmatic approach as the guiding paradigm in research methodology (Morgan, 2007). This approach, transcending the dualism between qualitative and quantitative research by refusing to choose between the two extremes of induction–deduction, subjectivity–objectivity and context–generality, emphasises on the dimensions of abduction, intersubjectivity, and transferability for thinking across methodology.

To apply this pragmatic approach, the present study first refuses the subjective–objective dichotomy in understanding the world. Alternatively, it takes an intersubjective perspective to consider knowledge as “plural, changing and open” (Biesta, 2010, p. 17), and thus the creation of knowledge needs the joint actions of thinking, cooperating and communicating (Greene & Hall, 2010). Therefore, the knowledge of Manchu word reading could be more comprehensively produced if both quantitative and qualitative methods are employed for its examination. Moreover, it would be too absolute to take a pure induction–deduction divide in research design, data collection and analysis. Alternatively, the abductive reasoning is more appropriate because it enables the connection between theory and data. In this regard, this study first connected the quantitative result of Manchu word decoding performance with the orthographic depth hypothesis (Frost et al., 1987; Katz & Frost, 1992),

the psycholinguistic grain size theory (Ziegler & Goswami, 2005), the orthographic breadth hypothesis (Nag, 2007, 2017), and the construct of intra-symbol processing (Nag, 2022), and then explained the results with data from qualitative classroom observation.

3.2.3 A Multidimensional MMR Framework

This study took into consideration three dimensions for an MMR design—(a) tool status (i.e., the principal, subsidiary or equal status of the tools used for data collection and analysis), (b) time order (i.e., whether the quantitative and qualitative research occur simultaneously or sequentially), and (c) sampling (i.e., selecting samples from the same or different groups of population). In this framework, the present study was in a concurrent multilevel design with quantitative principal tool for data collection.

In terms of tool status, Morgan (1998) has suggested that research design needs to select a principal tool for data collection by giving priority to one method of the two candidates (i.e., qualitative and quantitative methods), with the other offering assistance as the subsidiary tool. In this framework, Morgan refused a third possibility of giving equal weight to measure the importance of data collection tools because of his concern over the epistemological incommensurability between quantitative and qualitative methods. However, as this study took a pluralism view towards methodology (Teddlie & Tashakkori, 2010), it therefore considered the tool priority to be encompassing the three possibilities of principal, subsidiary and equal status. The most important criterion for deciding on tool priority is its strengths to achieve research goals (Morgan, 1998). Since the primary objective of this project is to investigate from the quantitative perspective the association between the complexity of Manchu writing system and word decoding, adopting quantitative method as the principal tool is appropriate to collect the data regarding word decoding accuracy and speed. On the other hand, as the classroom observation aimed at complementarily accounting

for the quantitative results of Test 2 on symbol–sound mapping complexity, it was used as the subsidiary tool for data collection.

The second dimension concerning MMR design is the time orientation (Johnson & Christensen, 2014). Time orientation refers to whether the quantitative and qualitative components or phases occur concurrently or whether they are organised sequentially. In this study, the qualitative classroom observation was conducted in a five-week time frame, at the end of which tests on Manchu words decoding were administered to the participants. The time lapse between collecting the qualitative and quantitative data was so minimal that they were conducted approximately at the same point of time. Moreover, there was no separate or second phase for data collection. Therefore, it is self-explanatory that this study took a concurrent time orientation for data collection.

Finally, MMR design also requires making a decision on sampling. After addressing the time order concern, researchers should decide on how to select the units of data collection and analysis. Tashakkori and Newman (2010) suggested that three sampling methods could be used for unit selection: multisample (to sample two related groups of participants from the same population), same/subsample (to select all or part of the participants in the first sample for a second phase of the study), and multilevel sample (to select two samples from different groups of population). This study adopted the multilevel sampling method in which a group of Manchu learners (n = 196) were randomly selected and tested on word decoding. At the same time, a Manchu teacher was observed for the Manchu orthographic knowledge instruction.

3.3 Sample and Sampling

3.3.1 The Mixed Methods Sampling Model

Combining quantitative and qualitative methods in research can increase the complexity and difficulty of sampling strategies. To address this problem, this study adopted

the Two-Dimensional Mixed Methods Sampling Model (Onwuegbuzie & Collins, 2007) for sampling design. This model takes into consideration two dimensions for MMR sampling (a) the time orientation (concurrent vs. sequential) of conducting the quantitative and qualitive components, and (b) the relationship of the qualitative and quantitative samples (identical, parallel, nested, or multilevel). The present study employed the concurrent design using multilevel samples purposively selected from two populations (students and teachers) for the quantitative and qualitive components. In the quantitative component, students who took the Manchu course were sampled and tested on Manchu word decoding. In the qualitative component, a teacher of Manchu in the university was selected and observed for the Manchu orthography instruction.

3.3.2 Sampling Frame and Schemes

The present study is bounded as a semester-long investigation on Manchu symbol blocks decoding by novice learners in a Chinese university conducted from September to December 2019. In the fall semester 2019, the university offered to 1851 freshers the newly developed course *Introduction to Manchu Intangible Cultural Heritage*, which covers a wide range of Manchu-related topics including literature, arts, sports, medicine, folklore and orthography. These students were included into the sampling frame. In terms of the sampling scheme, for the quantitative component of this study, the participants were selected by simple random sample (Bryman, 2012) from a population of 1,851 students taking Manchu course in the university. An important issue of sampling is the homogeneity of population, which for the present study required that none of the students should have had prior experience in learning Manchu. To maintain population homogeneity, a screening was given to exclude any student with prior Manchu learning experience from the sampling frame, which resulted in none. Prior to sampling, the researcher approached the university Manchu teachers to introduce the project and seek their help for participants recruitment. The teachers kindly

agreed to help and strongly encouraged the students to participate in the tests, because that would be beneficial for the students too. As it was advised by a teacher, "preparing and taking part in the tests will be helpful for you [the students] to review the Manchu orthographic knowledge instructed in class". In terms of the random sampling procedure, name lists of the 1,851 students were collected from the teachers and 105 students were randomly selected (see Section 3.3.3 for the decision on sample size). Yet, perhaps due to the teacher's strong encouragement and the university students' curiosity about what research is like, an additional number of students came to the researcher and finally 196 participants took part in the tests.

For the qualitative component, this study took purposive sampling (Bryman, 2012). This non-probability form of sampling aims to strategically select units of analysis (i.e., people, organisations, documents, etc.) based on certain criteria so that the research questions could be answered. This study employed the critical case approach of purposive sampling (Onwuegbuzie & Collins, 2007), which chooses a crucial case (i.e., individuals, groups, or settings) that could provide the researchers with compelling insight about a phenomenon of interest. This study ended up choosing a Manchu teacher in the participating university as the critical case for two reasons. The first reason is that he, with over ten years of Manchu teaching experience, appeared to have developed a very clear teaching strategy in terms of the orthographic unit size (i.e., uju or phoneme markers) to be used in Manchu orthography instruction. This strategy on unit size choice is directly related to RQ 2-2 (In what way were the results of the learners' decoding accuracy and speed on Manchu symbol blocks with different phoneme marker–sound mappings related to the instruction on Manchu orthography?). Moreover, this Manchu teacher was critical because of his willingness to participate in this research project. As an ethnic Manchu, a university lecturer, an active

Manchu educator, and a passionate advocate of Manchu language revitalisation over the past ten years, he was very supportive of studies on Manchu and willing to participate.

3.3.3 Sample Size

The formula provided by Yamane (1967) was adopted to calculate the sample size. A prerequisite of using this formula is the normality of data, which in the present study refers to the normal distribution of the students' scores in Manchu word decoding task. To test the normality, Shapiro-Wilk test (Shapiro & Wilk, 1965) was run on the students' scores of Manchu word decoding task in the pilot study (student n = 15). The results showed a p value of .197, suggesting a normal distribution of the data. Thus, it was assumed that the test scores of the participants in the main study would likely be normally distributed.

For sample size calculation, with an assumed 95% confidence level and 0.5 degree of variability (i.e., the maximum variability in a population), the equation is

$$n = \frac{N}{1 + N(e)^2}$$

where, n = the sample size

N = the population size (1851 learners in the present study)

e = the sampling error (10% in the present study, i.e., a 90% confidence interval)

When the above parameters were applied to the formula, the result was 94.87, meaning 95 participants were required to be included in the sample. This figure was close to the sample size recommended for quantitative search design (Onwuegbuzie & Collins, 2007)—a minimal of 82 participants are required for correlation test (two-tailed) and 64 participants for causal-comparative test (two-tailed), with .80 statistical power at the 5% level of significance. Moreover, to compensate for the students who may decline to participate, an additional 10% participants were selected (Israel, 1992), making the final sample size 105.

For the qualitative sample size, the population of Manchu teachers in the research site was a very limited number of three. This could be partially due to the status of Manchu—it is critically endangered with few people using it for communicative purpose and there are only a few educational programmes in China offering Manchu class. Therefore, after considering whether the teachers have developed a clear strategy of the orthographic unit (uju or phoneme marker) to be adopted in instruction and counselling their willingness to participate, the resulting sample size was one.

3.4 Instruments

3.4.1 Measures of Manchu Symbol Block Visual Complexity and Mapping Complexity

Visual Complexity. This study used the following five dimensions to measure the visual complexity of the Manchu symbol block (n = 18) which were given to the participants for recognition in a naming task. This test is referred to as Test 1.

(a) Pixel count (PX). PX refers to the value of filled pixels of the foreground in an image (Nag et al., 2014). The symbol blocks were in font style Daicing H11, font size 140, and image size 20 kb in a 300 × 400 pixels PNG image, with black foreground (i.e., the symbol block) against white background.[11] The ImageJ software (version 1.52k) for macOS was used to calculate the pixel value statistics of the symbol area in the images.

(b) Perimetric complexity (PC). PC is the ratio of the squared perimeter of the foreground to its area, both of which are measured in pixel count (Pelli et al., 2006; Watson, 2012). Its equation is written as

$$PC = \frac{P^2}{4\pi A}$$

[11] In a binary picture (e.g., a black-and-white picture), foreground is defined as the area(s) of the same colour, and the remainder as background (Watson, 2012, p. 2).

where *P* is the summation of the inside and outside perimeters of a foreground and *A* is the area, with 4π as the nominalizing constant. The ImageJ software (version 1.52k) was used to calculate the pixel value statistics of the perimeters and area of the symbol.

(c) Number of simple features (SF). SF is defined as the distinct written element that can be discriminated from others (Pelli et al., 2006), such as a line segment, a curve, a dot, and a circle. In Manchu orthography, an SF is typically one of the following ten basic strokes in Manchu, namely the head, the tail, the tooth, the circle, the dot, the falling, the belly, the curve, the horizontal, and the vertical. For instance, the symbol block < ᡳ > /i/ comprises four SF of the crest, the vertical line, the long tooth, and the belly.

(d) Number of connected points (CP). CP is an adjoining of at least two simple features (Chang et al., 2018). For example, < ᡳ > contains two CP—one is the connection between the crest and the vertical line, and the other is the contact among the vertical line, the long tooth, and the belly.

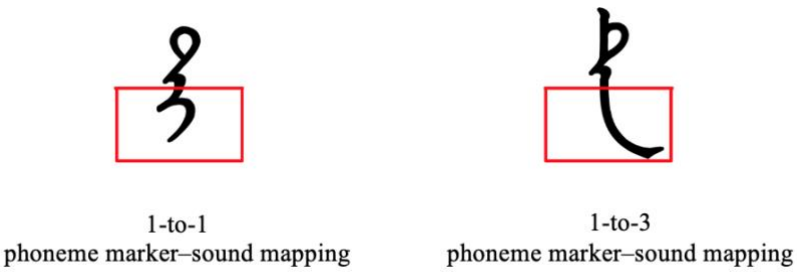
(e) Number of disconnected components (DC). DC is defined as a simple feature that is not linked to other feature in a set (Chang, Plaut, & Perfetti, 2016). If a word is composed of more than one DC, there are spaces among the components (Chang et al., 2018). For example, the Manchu symbol < ᡤᡳ > contains three DC—the two dots and the combination of the remaining simple features.

Mapping Complexity. To examine the association between symbol–sound mapping complexity and symbol block decoding performance, the participants were given two sets of uju blocks (n = 16) for recognition in a naming task. This test is referred to as Test 2. The symbol–sound mapping complexity of the symbol blocks in Test 2 was indexed by the phoneme marker–sound mapping on the "target" phoneme markers. For example, in Figure 7 the target phoneme markers in the red boxes are in 1–to–1 and 1–to–3 mapping with sound,

respectively, with the remaining non-target markers in both having a 1–to–1 mapping with sound. Therefore, the mapping complexity of the two symbol blocks are measured as 1 and 3.

Figure 7

Mapping Complexity Measure on the Target Phoneme Marker



Syllable Complexity. A Test 3 on syllable complexity of the Manchu writing system was originally designed as a part of the present study, but it was dropped out of analysis because of the floor effect of the participant's decoding performance. Test 3 aimed to examine whether there is syllable-frequency effect in Manchu bi-syllabic pseudoword recognition. In a naming task of pseudowords (n = 32) in <$CV_1.CV_2$> structure, the syllable frequency (high vs. low) of CV_1 and CV_2 was manipulated. Initial analysis of the decoding performance (measured by the number of correctly decoded pseudowords) was conducted on 54 participants who were randomly selected from all participants (n = 196). The results revealed that they performed poorly in Test 3—thirty participants scored zero in the naming task (M_{score} = 3.80, SD = 5.87, $Min._{score}$ = 0, $Max._{score}$ = 21). Thus, Test 3 was excluded from data analysis.

3.4.2 Measures of Manchu Symbol Block Decoding Performance

In Test 1 and Test 2, the participants were asked to name the symbol blocks presented to them as accurately and quicky as possible. Their performance in the naming task was measured by error rate for decoding accuracy and reaction time for decoding speed.

(a) Error rate (ER). ER of naming a symbol block was defined as the ratio of the total number of incorrect responses given by the participants to the total number of attempted responses. The unattempted responses were excluded from the calculation of ER. The ER for a symbol block is between 0 (i.e., all correct responses) and 1 (i.e., all incorrect responses). For instance, out of the 196 participants in naming the item <ᡤ> /ɑ/, 53 gave incorrect responses while 141 gave correct ones, with two participants not responding. Therefore, the error rate for naming <ᡤ> is .273 (= 53 / (53 + 141)).

(b) Reaction time (RT). Decoding speed for an item was measured by the average RT (in milliseconds) of the participants who correctly named it. To measure the RT, the entire testing process of a participant was audio-recorded, which included the participant's oral response and the tester's clicking the mouse to scroll slides. The audio files (n = 196) were imported into the WavePad software (Version 10.48) to be displayed in waveform graphs. Six audio files were excluded because they were failed to be played by WavePad. In the remaining data, 43 observations (4.12% of the entire observations) on RT were further deleted due to unclear recording or the participants' self-correction of responses. In the waveform graph, the RT is represented by the distance (i.e., the time duration) between the crest on the wave of the mouse clicking sound (indicating the time point when the item slide was being presented to the participant on computer screen) and the point where the participant was just about to initiate the naming. Time duration between the two points was measured in milliseconds by WavePad too.

3.4.3 Validity of Instrument

Instrument Validity. This study measured participants' Manchu symbol block decoding accuracy and speed indicated by error rate (ER) and reaction time (RT). To ensure measurement validity, concepts must be accurately reflected by the measures (Bryman, 2012). Therefore, the ER and the RT in this study need to really reflect the constructs of decoding accuracy and speed, which requires the assessment of them to be reliable. To achieve this goal, the participants' responses in Tests 1 and 2 were measured for accuracy (correct response = 1, incorrect = 0) and speed (in milliseconds). Subsequently, a repeated measure was made on the responses of 30 randomly selected participants. Finally, Cronbach's α was carried out to measure the intra-rater reliability. In addition, to ensure the internal consistency of the participants' response across the items in both tests, Cronbach's α of their responses was calculated to see whether it reached the critical value of 0.7.

Criterion Validity. It was acknowledged that the measure of Manchu symbol block decoding accuracy and speed could not be easily checked for criterion validity. That is because there was no test previously conducted on Manchu symbol decoding; thus, it was not feasible to compare the tests used by the present study to others.

3.5 Procedure of Data Collection

3.5.1 Implementation Matrix for Data Collection

Since data collection in a mixed methods study aims to answer the research questions (Teddlie & Yu, 2007), it is necessary to keep considering whether the sample and data are truly able to offer answers to the questions. Thus, this study employed the strategy of developing an implementation matrix (Creswell & Plano Clark, 2018) to list side by side the research questions, the samples and the specific data sources (quantitative and qualitative), and the data gathering instruments to answer the research questions. Table 5 presents the connections between these respects.

Table 5

Implementation Matrix for Data Collection

Research questions (RQs)	Samples	Data sources	Instruments	Details
• RQ 1-1: To what extent does the visual complexity of Manchu symbol blocks associate with the learners' decoding accuracy and speed? • RQ 1-2: In what ways do the measures of the visual complexity of Manchu symbol blocks predict the learners' decoding accuracy and speed?	• Randomly selected students taking a Manchu course in a university (n = 196)	• Test 1: Visual complexity of Manchu symbol blocks (quant)	• Pixel count • Perimetric complexity • Number of simple features • Number of connected points • Number of disconnected components	• Software ImageJ used to measure the inside and outside perimeters and the area (all in pixel counts) of a symbol block • Perimetric complexity calculated by the formula $\frac{Perimeter^2}{4\pi Area}$
• RQ 2-1: Do the learners perform differently when they decoded Manchu symbol blocks with phoneme markers in one–to–one and one–to–multiple mapping with sound?		• Test 2: Symbol–sound mapping complexity (quant)	• Phoneme marker–sound mapping on the target phoneme marker	• For 1–to–1 correspondence, mapping complexity = 1 • For 1–to–multiple (M) correspondence, mapping complexity = M

Research questions (RQs)	Samples	Data sources	Instruments	Details
• RQ 1-1 • RQ 1-2 • RQ 2-1		• Tests 1 and 2: Manchu symbol blocks decoding performance (quant)	• Decoding accuracy • Decoding speed	• Decoding accuracy is measured by error rate • Decoding speed is measured by average reaction time (in millisecond)
• RQ 2-2: In what way are the results of the learners' decoding accuracy and speed on Manchu symbol blocks with different phoneme marker–sound mappings related to the instruction on Manchu orthography?	• Purposively sampled teacher who provided Manchu orthography instruction in the university (n = 1)	• Orthographic unit (i.e., uju or phoneme markers) adopted in Manchu orthography instruction (qual) • Frequency of uju and phoneme markers given in observed sessions (quant)	• Nonparticipant classroom observation	• Ten sessions of 45-min instruction on Manchu orthographic knowledge given over five consecutive weeks

3.5.2 Test Design

Rationale for Designing the Test. The design of Manchu symbol blocks recognition test was preconditioned by the fact that, according to a questionnaire of demographic information, all participants (n = 196) are completely novice learners with no prior Manchu knowledge. This fact suggested that the selection of symbol blocks in the tests should be in correspondence with their very limited Manchu proficiency. Thus, the Manchu symbol blocks in Tests 1 and 2 comprise maximumly three phoneme markers (see Tables 6 and 7 for the full item lists). Based on my observation, the instruction of Manchu orthography covered a part of rather than all the 130ish uju distributed in the first section of the Twelve Sections of Manchu Syllabary. These uju instructed were in relatively less complex V, VV, CV and CVC syllables. If the syllabically more complex symbol blocks were to be selected for tests, such as these in CVCV and CVCCV structures, it may result in floor effects which would impair data validity. Such floor effects were observed in my pilot study carried out in 2018 in the same participating university with the same teacher but a different group of participants (see Appendix B for details). The participants (n = 15) in the pilot study, similar to these recruited in the main study, had no prior Manchu knowledge. They were asked to decode 12 Manchu symbol blocks in VV, VCV, CVVC, CVCV, CVCVC, and CVCVCV structures. The results showed that they performed rather poorly in decoding these items (M_{score} = 1.47 out of 12, SD = 0.99, $Min._{\text{score}}$ = 0, and $Max._{\text{score}}$ = 3). Based on this evidence, it seemed reasonable to design the test as a Manchu symbol block naming task, in which the items were controlled to V, CV, and CVC syllables.

Test 1 on Visual Complexity. A list of 18 Manchu symbol blocks were given to the participants in a naming task. Considering the participants had received only 10 sessions of Manchu orthography instruction at the time of testing, the symbol blocks were selected based on three criteria—length, syllable structure, and non-ligature. First, the symbol blocks in Test

1 comprise one to three Manchu phoneme markers, which corresponded with the length of symbol blocks instructed. Second, the symbol blocks were in V, CV and CVC syllables, which were the types of syllables distributed in the first section of the Manchu syllabary. Moreover, symbol blocks written in ligature were excluded from the test.[12] That is because the consonant marker and vowel marker in a CV ligature are nested together, while consonant and vowel markers in a non-ligature CV block are connected by a vertical line segment. This architectural difference in symbol design may affect uju processing. Based on these selection criteria, a list of 18 symbol blocks (V = 5, CV = 8, and CVC = 8) were selected. Table 6 gives the full list.

Table 6

List of Items in Test 1 on Visual Complexity

Items	Latin transcription	Syllabic structures
	*a*	V
	*e*	V
	*i*	V
	*u*	V
	*ū*	V
	*ca*	CV

[12] Li (2010) listed 14 ligatures in Manchu, which are *ba*, *be*, *bi*, *bo*, *bu*, *ke*, *ki*, *ku*, *ge*, *gi*, *gu*, *he*, *hi* and *hu*. This study adds to this group six additional ligatures of *pa, pe, pi, po, pu* and *pū*.

Items	Latin transcription	Syllabic structures
	*du*	CV
	*ha*	CV
	*je*	CV
	*me*	CV
	*ni*	CV
	*so*	CV
	*to*	CV
	*kib*	CVC
	*lan*	CVC
	*sek*	CVC
	*sun*	CVC
	*yor*	CVC

Test 2 on Symbol–Sound Mapping Complexity. A list of 16 Manchu symbol blocks were given in a naming task in Test 2. All were non-ligatures in CV (*n* = 12) or CVC (*n* = 4)

syllables. The symbols were in two blocks, with Block 1 and Block 2 each having 8 items. In Block 1, the phoneme markers constituting a symbol block were all in 1-to-1 mapping with sounds. However, in Block 2, the target phoneme marker in a symbol block had a 1–to–2 or 1–to–3 mapping with sound, with the non-target phoneme markers in a 1–to–1 mapping with sound. The symbol blocks in Blocks 1 and 2 were pair-matched in (a) syllabic structure, (b) disconnected components, and (c) visual complexity (which equals the summation of perimetric complexity, simple features, and connected components). Table 7 gives the list.

Relationship between Test 1 and Test 2. In terms of test administration, Test 1 and Test 2 were given to the participants in the same test session. A participant was first asked to name 18 items in Test 1, which was followed by the immediate administration of Test 2 to name another 16 items. Furthermore, when the thesis was being written, the findings of Test 1 regarding the association between uju visual complexity and learners' decoding performance were applied to examine whether visual complexity of the pair-matched items (e.g., *sū-tu* and *ce-we*) given in Test 2 was effectively controlled. In the design phase of Test 2, no findings was available in terms of the relationship between Manchu symbol visual complexity and readers' decoding performance. As a result, the implementation of visual complexity control of the item pairs in Test 2 was based on the studies on other writing systems (e.g., Chang et al., 2018; Nag et al., 2014). However, after the present study has found the association between Manchu symbol visual complexity and novice learners' decoding performance, this finding was applied to triangulate the results of item visual complexity control in Test 2. A detailed discussion on the triangulation was given in Section 5.2.3.

Table 7

List of Items in Test 2 on Symbol–Sound Mappin Complexity

Block 1		Block 2		Syllabic structure
Items	Latin transcription	Items	Latin transcription	
	*ti*		*te* [a]	CV
	*di*		*de*	CV
	*sū*		*tu*	CV
	*nū*		*du*	CV
	*ce*		*we*	CV
	*cū*		*fū*	CV
	*jūs*		*tus*	CVC
	*tib*		*teb*	CVC

[a] The target phoneme marker in an uju is indicated by a single line for one–to–two mapping with sound and a double line for one–to–three mapping with sound.

The symbols were first typed in Microsoft Word in Daicing H11 font, 140-point size. Then, screenshots were taken to get 300 × 400 pixels images, with black foreground (i.e., the symbol) and white background. Images were then placed at the centre of PowerPoint slides to be presented to the participants in test.

3.5.3 Task

Test 1 and Test 2 used Manchu symbol block naming tasks, which asked the participants to name the symbol blocks one at a time as rapidly and accurately as they can. The participants were asked not to think aloud when attempting to recognise the symbol blocks. Correct naming of an item scored 1 point and incorrect naming scored 0.

3.5.4 Procedure

Prior to taking the test, the participants were provided with an information sheet outlining the purpose of this project and what taking part would involve for them. Consent for participation was also obtained from the participants. Eight testers who have taken extensive Manchu courses for two years at the participating university were recruited. They had received one week training and support. The information sheet and consent form were included in Appendix A.

Each participant was seen individually by a tester in a quiet room in the university. Dell desktop computer with 21.5-inch monitor of 1920 × 1080 resolution was used for presenting the test items. Each item began with the presentation of a fixation cross (+) in the centre of a PowerPoint slide. Then, a Manchu symbol block was presented at fixation and remained there until the participant gave a response, after which the tester paged down to show the fixation cross for the upcoming symbol block.

Each participant received two tests (Test 1 and Test 2) after the practice trials. The practice list included 12 symbol blocks in V, CV, CVC, and CVCV syllables. The aim of the practice was to familiarize the participants with the naming task. Subsequently, Test 1 on

visual complexity presented 18 Manchu symbol blocks to the participants. In the following Test 2 on symbol–sound mapping complexity, the items were given in the manipulated "blocked" and "mixed" conditions. In "blocked" condition, the eight items in Block 1 were given first, followed the eight items in Block 2. In the "mixed" condition, all 16 items were mixed and given randomly. The testing was audiorecorded.

3.5.5 Classroom Observation on Manchu Orthography Instruction

Context. Situated about 140 miles northeast to Beijing is Chengde, a prefecture-level city in Hebei Province. Historically, the city had a close association with the Manchus because of the Mountain Resort, a UNESCO World Heritage Site of the vast imperial palaces and gardens frequently used by the Qing emperors as their summer residence. Besides the magnificent architecture and landscape, another connection that the city keeps with the Manchus is in the less tangible way of teaching Manchu language in a local university, in which the classroom observation was carried out from September to December 2019. The use of Manchu language in the university was first observed on a campus notice board in which the university name was printed multilingually in Chinese, English, and Manchu, as well as on the shades of the lawn lights in which Chinese characters and corresponding Manchu were carved. The university, as introduced on its official website, is a university for nationalities with a rich culture in ethnic minorities activities. It has set up an exhibition centre of minority intangible cultural heritage and offers a wide range of courses related to Manchu. To the best of the researcher's knowledge, it is the only university in China that provides a Manchu language course to a population of over 1,800 graduate students. Although 1,800 is not a substantial number *per se*, considering the typical cohort of around thirty or even fewer postgraduate masters or doctoral students studying in a very limited number of Manchu programmes in China, it is nonetheless noteworthy. Also, workshops for Manchu language were held in the university during summer and winter vacations.

Manchu Course. The course observed was entitled *Manchu Intangible Cultural Heritage*, which was a mandatory elective to all first-year graduate students in the academic year 2019-2020. The general objective of the course as stated in the curriculum is that

> “学生通过该课程的学习能够达到对满族非物质文化遗产的理论和代表性项目有一个而较为系统且全面的初步了解，能够以非遗理论为指导，相对科学合理地分析非遗项目。”
>
> [Through the study of this course, students will be able to achieve a relatively systematic and comprehensive preliminary understanding of the theory and representative items of Manchu intangible cultural heritage, and to use the theory to guide the analysis of intangible cultural heritage in a scientific and reasonable way. My translation.]

The duration of the course was 16 weeks, with two sessions of 45-minute instruction given each week. Table 8 gives the weekly teaching contents. The 32 sessions can be grouped into three themes, which are Manchu culture (sessions 1-16), Manchu orthography (17-28), and curricular practices (29-32). Due to the purpose of this study, the classroom observation was focused on the ten sessions of Manchu orthography instruction.

Classroom Observation. Non-participant observation was conducted on the Manchu instruction in the participating university. Adapted from the construct of component skills of literacy (Nag, Snowling, & Asfaha, 2016), the present focused on observing the more specific *component knowledge of orthography* in Manchu instruction. The component knowledge of orthography included the four dimensions of (a) stroke knowledge (e.g., Jin & Cortazzi, 1998; Shen, 2005), (b) orthographic unit knowledge (e.g., Nag & Narayanan, 2019), (c) symbol–sound mapping knowledge (e.g., Nag & Snowling, 2012), as well as (d) other knowledge describing the orthography instruction not covered by other dimensions. Moreover, the observation also focused on literacy practices (Street, 2000), which included

the teaching and learning activities and interactions between teacher and students in Manchu class. Their literacy practices, drawing on the teacher talk, student talk, and literacy activities in the "teaching practices" in Nag and colleague's (2016) framework, incorporated the three dimensions of (a) explanation (offering relevant orthographic knowledge by teacher), (b) exercises (reading and writing taken by students), and (c) evaluation (assessing whether students have grasped the relevant orthographic knowledge by teacher–students interaction).

Table 8

Curricular Contents of the Manchu Intangible Cultural Heritage Course

Weeks [a]	Topics	Contents
1	General introduction to intangible cultural heritage	Definition, category, values, and preservation of intangible cultural heritage
2	Traditional Manchu arts	Papercut, embroidery, fabric collage, jade carving, and inkstone carving
3	Traditional Manchu music and dance	Musical instruments, dance, play, drama, and folk songs
4	Traditional Manchu sports and games	Pearl ball, ankle bone shooting, archery, Manchu style football, martial arts, and Meljen (i.e., Manchu sports games)
5	Traditional Manchu craft and medicine	Costumes, food and drink, dwelling house, utensils and tools, and medicine
6	Traditional Manchu folk customs	Folk beliefs and rituals, weddings, code of ethics, funerals, and festival customs
7	Manchu folklores (1)	Ulabun (i.e., legend of the tribes)
8	Manchu folklores (2)	Folk tales
9	Manchu orthography (1)	Vowel uju <A>, <E>, <I>, <O>, <U>, <Ū> Uju for diphthongs <AI>, <EI>, <II>, <OI>, <UI>, <ŪI>, <AU>, <EU>, <IU>, <OU>, <UU>, <ŪU> Uju for triphthong <IUI> Consonant phoneme marker <N>

Weeks [a]	Topics	Contents
10	Manchu orthography (2)	Consonant phoneme markers <K1>, <G1>, <H1>, <S>, <Š>, <L>, <M>, <C>, <J>, <Y>, <R>, <F>, <W>, <TS‘>, <DZ>, <Ž> Syllables <CI>, <JI>, <SI>
11	Manchu orthography (3)	Syllables <SY>, <C‘Y>, <JY> Consonant phoneme markers <T>, <D>, <B>, <P>
12	Manchu orthography (4)	Consonant phoneme markers <K2>, <G2>, <H2>, <K’>, <G’>, <H’>, <-K>, <NG>
13	Manchu orthography (5)	Inconsistent syllable-sound mapping in Manchu Stroke ligature in Manchu calligraphy
14	Review of Manchu orthography (no class)	–
15	Visit to the university Exhibition Centre of the Minority Intangible Cultural Heritage	To understand the exhibition items in relation to the course
16	Visit to the Outlying Temples of the Chengde Mountain Resort	To explore the items related to Manchu intangible cultural heritage in the Outlying Temples

[a] The course length was 16 weeks, with two sessions of 45-minutes class given per week.

3.5.6 Fieldwork Timeline

The researcher entered the fieldsite in September 2019 when the freshman came to the university. The researcher started attending the Manchu class in Week 3 when it began and kept observing the class for the following 12 weeks to understand the full cycle of activities (Riemer, 2012). Finally, in Week 15 when the class was to be finished, tests on Manchu symbol blocks recognition were administered to the participants. The tests given to the participants (n = 196) were completed within two weeks' time before Week 17, when students began to prepare for the university examinations.

3.6 Piloting

A pilot study was conducted between September and October 2018, aiming to examine children's Manchu language learning in Chinese educational institutions. The researcher of this study first paid visit to a primary school located in a Manchu Autonomous County and then to a middle school in a city in Northeast China. Some children at both schools were invited for a Manchu symbol block recognition test. The blocks were in V, CV, CVC, VCV, VCCV, CVCV, CVCVC, CVCCV, VCVCV, CVCVCV, and VCCVCCV syllables. It was found that the decoding skills of the children were so limited that their symbol recognition scores were at floor. The mean score of the test for the primary and middle school children was 0 and 2.3 (out of 41) respectively. Moreover, it was also found that there was a discrepancy between the course curriculum and its implementation with respect to class hours and teaching contents. This gap, according to an expert in Manchu education studies, seemed to be quite typical for Manchu courses given in primary and secondary schools. Appendix B reports more details of the school pilot.

Therefore, the alternative research site of the present participating university in North China was identified. This university was selected primarily for three reasons. First, the pilot study carried out in 2018 found that students in the university were capable of naming

Manchu symbols blocks in the relatively less complex V, CV, and CVC syllables. Therefore, it was probable that the incoming students in 2019 could also be able to read these syllables. That would make it possible to collect analysable data to answer the research questions regarding visual complexity and mapping complexity of the Manchu writing system. Second, a population of approximately 2,000 students in the university was taking the Manchu course. Thus, sufficient sample could be drawn from this population. Finally, the university offered a 16-week instruction of Manchu orthographic knowledge (which was unexpectedly truncated to five weeks during the main study). Therefore, observation could be initiated on literacy practices in the classroom so as to answer the research question regarding the relations between instruction and the orthographic unit adopted in symbol blocks decoding.

3.7 Ethical Consideration

The present study followed British Education Research Association's ethical guidelines for educational research (BERA, 2018). To fulfil the research requirement in the Department of Education, University of Oxford, an ethical application was submitted to the Departmental Research Ethics Committee (DREC) for review in September 2018, and a clearance was granted for my project (ED-CIA-18-243). However, for the main study, the subjects were changed from children, as planned in my ethical application, to university students. This is because the Manchu proficiency of the potential participating children was so low, as found by the pilot study, that it was almost impossible to collect valid data from them. This change has been reported to the Chair of the DREC, who confirmed and noted it. In addition, a letter introducing the project was provided to the department head in the participating university to seek permission to collect data. Moreover, information sheets were offered to the potential participants and consent forms were used to record their understanding and agreement to participate in the project.

3.8 Strategies for Data Analysis

3.8.1 Separate and Mixed Methods of Data Analysis

The purpose of data analysis is to answer the research questions. Thus, for RQs 1-1 and 1-2 which aimed at examining the association between visual complexity and symbol block decoding performance, the data collected were analysed quantitatively for correlation test. However, for RQs 2-1 and 2-2 on symbol–sound mapping complexity, a mixed methods data analysis (Creswell & Plano Clark, 2018) was employed to quantitatively analyse the quantitative data and qualitatively analyse the qualitative data before integrating both datasets and results.

3.8.2 Analytical Strategies

Quantitative Data Analysis. The data of the measures of symbol visual complexity and decoding accuracy and speed in Tests 1 and 2 were prepared for IBM SPSS Statistics for macOS (Version 28) for analysis. Then, the data were analysed with inferential statistical tests using SPSS. For Test 1, partial correlation test (with frequency of symbol blocks in Test 1 controlled) was performed to analyse the correlation between the five measures of symbol visual complexity (i.e., pixel count, perimetric complexity, simple features, connected points, and disconnected components) and the error rate and reaction time of the participants' decoding performance. Moreover, a multiple linear regression test (enter) was carried out with the five measures of visual complexity as predictors to build a model predicting the learners' decoding performance in recognising Manchu symbols. For Test 2, paired sample *t*-tests were performed to compare the participants decoding performance in naming the symbol blocks presented in the "blocked" and "mixed" condition, and the decoding performance on symbol blocks between Block 1 and Block 2. Finally, the findings were represented and interpreted to answer RQs 1-1 and 1-2 on symbol block visual complexity and RQ 2-1 on symbol–sound mapping complexity.

Qualitative Data Analysis. The qualitative data obtained from classroom observation was first organised in a synthesis matrix. Afterwards, the dimensions in the synthesis matrix were applied as codes for initial data coding, from which the sub-codes were developed. For data analysis, a thematic extraction was applied to identify what knowledge of Manchu orthography was taught and in what methods they were instructed. Finally, the frequency of uju and Manchu phoneme markers appearing during the ten sessions of instruction was counted.

Qualitative Data Organisation. The qualitative data of Manchu orthography instruction was organised in a synthesis matrix (see Table 9) derived from the meta-ethnography framework developed by Nag, Snowling, and Asfaha (2016) for configuring observation data. Their synthesis framework functions by mapping the observation data of classroom literacy practices onto multiple dimensions under the construct of "component skills of literacy" (e.g., letter and symbol knowledge, manipulation knowledge, reading knowledge, etc.). Each dimension of the components skills of literacy was further mapped onto the observation about "teaching practices" reflected by the four dimensions of teacher talk, student talk, literacy practices, and other features. In application of Nag and colleague's (2016) synthesis matrix, the present study adapted the construct of component skills of literacy to the more specific focus on *component knowledge of orthography* to accommodate to observations made specifically on Manchu orthographic knowledge instruction. Regarding the component knowledge of orthography, previous studies have identified important dimensions on strokes (e.g., Jin & Cortazzi, 1998; Shen, 2005), basic orthographic unit (e.g., Nag & Narayanan, 2019), and symbol–sound mapping (e.g., Nag & Snowling, 2012). Therefore, the three dimensions of (a) writing knowledge, (b) orthographic unit knowledge, and (c) symbol–sound mapping knowledge were used as codes for organising data under the construct of component knowledge of orthography. Additionally, an "other knowledge"

dimension captured the descriptions of Manchu orthography instruction that were not covered elsewhere in the matrix.

The above four dimensions captured the Manchu orthography knowledge instructed, which were further mapped onto the classroom literacy activities. The teaching and learning activities and interactions between teacher and students in Manchu class was conceptualised as literacy practices (Street, 2000). Street acknowledged that the traditional concept of literacy event is helpful in describing what was involved in reading and/or writing in various situations, but the concept of literacy practices transcends mere description by linking activities around literary to other things that people do in a broader social and cultural context. This broadening involves understanding what the nature of the practices is and what meanings are constructed by the practices. Thus, the present study regards the literacy practices of the teacher and students as a social endeavour occurring in the context of educational institutions that attempted to prevent the critically endangered Manchu language from extinction. Their literacy practices, drawing on the dimensions of teacher talk, student talk, and literacy activities in the "teaching practices" in Nag and colleague's (2016) synthesis framework, were mapped onto the three dimensions of (a) explanation of relevant orthographic knowledge by teacher, (b) exercises on reading and writing taken by students, and (c) evaluation by assessing whether students have grasped the relevant orthographic knowledge by teacher–students interaction.

Table 9

A Synthesis Matrix for Organising Classroom Observation Data

	Component knowledge of orthography (What)			
	Writing knowledge	Orthographic unit knowledge	Mapping knowledge	Other knowledge
	a. Stroke combination b. Stroke transformation c. Stroke order	a. Uju b. Phoneme markers	a. Symbol-to-sound b. Sound-to-symbol	a. Linguistics b. History
Literacy practices in classroom (How)				
Explanation (by teacher)				
a. Introduction				
b. Presentation				
c. Demonstration				
d. Manipulation				
e. Transliteration				
Exercise (by students)				
a. Choral reading				
b. Copywriting				
Evaluation (by teacher-students interaction)				
a. With students' response				
b. Without students' response				

Note. Adapted from Nag et al. (2016), p. 54.

Codes and Subcodes. Using the synthesis matrix, the observation data on Manchu orthographic knowledge instruction were initially coded as "writing knowledge", "orthographic unit knowledge", "mapping knowledge", and "other knowledge". Then, each dimension was further coded as "explanation", "exercise", and "evaluation". As coding and recoding continued in the second and third cycles, the codes became more refined and as a result some subcodes emerged (Saldaña, 2021). Table 10 gives codes, subcodes, their definitions, and examples from the observation data.

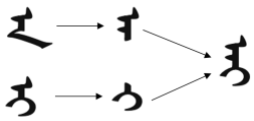
Table 10

Codes, Subcodes, Definitions, and Examples from Classroom Observation

Codes and subcodes	Definition	Examples
Writing knowledge		
Knowledge of combining, transforming, and sequencing strokes in writing Manchu symbols		
Stroke combination	Concatenating basic strokes to form a symbol	"只要在这个辅音字母<K>的第二个短牙右侧加一个'点'，它就变成了对应的<G>。" [If a "dot" is combined to the right of the second "short tooth" of the consonant letter <K>, it will be changed to the corresponding <G>.] (Session 3, 20:50)
Stroke transformation	Writing a stroke by changing the form of a related stroke	(Teacher demonstrating how to write the initial <A-> by transforming the "rightward tail" in the independent <A> to a "short tooth") "向右的'尾巴'等于一个'短牙'，你[在字干下]再写一个'短牙'，这就是<A>的词头型。" [The "rightward tail" is equal to a "short tooth". Write [under the "stem"] a "short tooth", and this is the initial form of *A*.] (Session 2, 22:55)
Stroke order	Sequence of writing the strokes in a symbol	"起笔写一个小线段，然后向右接写一个线段，然后向下转折，然后写一个'短牙'，转折回来写一个'字干'。"

Codes and subcodes	Definition	Examples
		[Begin by writing a “short segment line”, then connect it to a “segment line” written to the right, then bend to the lower left and write a “short tooth”, and bend again to write a “stem”.] (Session 3, 40:01)
Orthographic unit knowledge Knowledge of the positional variants of uju and Manchu phoneme markers		
Uju symbol block	The basic Manchu orthographic symbols in V, VV, CV, VC, CVV and CVC structures	The isolation, initial, medial, and final forms of <NA> (Session 2, 24:23) 
Phoneme markers	Orthographic symbol used to represent vowels and consonants in Manchu	The initial form of <A>, <E>, <I>, <O>, <U> and <Ū> (Session 1, 12:18) 
Mapping knowledge Knowledge of the correspondence between symbol and sound in Manchu		
Symbol-to-sound	Knowledge of spelling a sound	“这个/f/有两种写法：在元音字母<A>, <E>前右上角是出头的，在其他四个元音字母前是不出头的。” [The /f/ has two forms in writing: its top-right stroke protrudes before the vowel letters <A> and <E>, but not before the rest four vowel letters.] (Session 5, 04:45)

Codes and subcodes	Definition	Examples
Sound-to-symbol	Knowledge of reading a written symbol	(With the presentation of the independent and initial forms of <A>, <E>, <I>, <O>, <U>, <Ū> and <IUI> on slides) “我们重点看前六个[字母]。先跟我读一遍，所有人都有啊。/a/, /ə/, /i/, /ɔ/, /u/, /ʊ/。” [Let’s focus on the first six [letters]. Follow me to read them, all together. /a/, /ə/, /i/, /ɔ/, /u/, /ʊ/.] (Session 1, 13:02)
Other knowledge		
Knowledge related to a specific field		
Linguistics	Knowledge related to language	“满文是拼音文字。它和你学的英文，和你学的汉语拼音没有差别，都是 *a, b, c, d* 拼成一个单词 。单词内部没有视觉能看到的空格，... 一个单词内部是连成一串的，单词之间有空格。… 在书写上，从上到下，从左到右。” [Manchu writing is alphabetic. It is no difference from English or *Hanyu Pinyin* you’ve learned, in which *a, b, c,* and *d* are used for word spelling. There is no visual space within a word. … the inner [letters] in a word are connected in a string. But there is space between words. … Regarding writing [direction], [it is] from top to bottom and from left to right.] (Session 1, 09:58)
History	Knowledge related past events	(Teacher presenting the historical development of Manchu script) 唐–有文字（具体来源不详） 1119 年至 1575 年–女真大字，女真小字 1599 年创制–老满文（无圈点满文） 1632 年至今–新满文（圈点满文） Tang dynasty–Script was used for Manchu (no details known for its origin) From 1119 to 1575–Jurchen large script and Jurchen small script

Codes and subcodes	Definition	Examples
		1599–Creation of the old Manchu script (without dots and circles) From 1632 to present–The new Manchu script (with dots and circles) (Session 1, 03:24)
Explanation Making orthographic knowledge more comprehensive by describing, demonstrating, or clarifying		
Introduction	Introducing knowledge orally without referring to visual aids such as slides or writing on blackboard	“满文是拼音文字。它和你学的英文，和你学的汉语拼音没有差别，都是 *a, b, c, d* 拼成一个单词 。” [Manchu writing is alphabetic. It is no difference from English or *Hanyu Pinyin* you’ve learned, in which *a, b, c,* and *d* are used for word spelling.] (Session 1, 09:58)
Presentation	Showing knowledge by PowerPoints slides	Teacher presented the independent and initial forms of <A>, the independent and final forms of <I>, and the combination of A- and -I into AI. (Session 2, 00:01) 
Demonstration	Showing knowledge by writing on blackboard	T demonstrated the transformation of the “rightward tail” in the independent form of <TA> to the “short tooth” in its initial form. (Session 5, 23:50) 
Manipulation	Segmenting or blending orthographic units	<KA> was segmented by boxes into <K_1-> and <-A> in slides; <K_1-> was blended into <-A> to form <KA> on blackboard. (Session 3, 12:55) 

Codes and subcodes	Definition	Examples
Transliteration	Converting a writing symbol from one script to another	<K₁-> and <-A> obtained from segmenting <KA> were transliterated into k and a. (Session 3, 12:55) 
Exercise Activities of drilling		
Choral reading	Reading aloud in unison by the whole class	(Teacher leading the reading aloud of <KA>, <KO>, and <KŪ>) “和我读一下，[qʰa], [qʰɔ], [qʰʊ], [qʰa], [qʰɔ], [qʰʊ]。”[Read after me. [qʰa], [qʰɔ], [qʰʊ], [qʰa], [qʰɔ], [qʰʊ].] (Session 3, 10:34)
Copywriting	Practising writing Manchu symbols in exercise books	“好了，下面给大家两分钟时间，把<KA>的独立，词头，词中，词尾写下去。” [Alright. Now you will have two minutes to copywrite the independent, initial, medial, and final forms of <KA>.] (Session 3, 17:28)
Evaluation Assessment on the progress towards learning objectives		
With Ss’ response	Teacher asking question of taught material and students answering	Teacher: (pointing to <AM> written on blackboard) 这个念啥? [How do you read it?] Students: /am/. (Session 4, 14:10)
Without Ss’ response	Teacher asking question of taught material but students not answering	Teacher: (after writing <ATHA> on blackboard) 这个读啥? [How do you read it?] Students: (silence). (Session 6, 02:45)

Note. In this table, the class session from which an example was taken and the starting time of this example in that session was labelled as Session X, MM:SS. For example, Session 3, 20:50 indicates that the example was taken from Session 3 of the classroom observation and the literacy practices described in the example began at 20:50 of the observation.

Analysis of the Data in Synthesis Matrix. The analysis of the synthesis matrix was made through thematic exaction of the "what" and "how" of the literacy activities captured in the observation (Nag et al., 2016). "What" refers to the Manchu orthography knowledge instructed (e.g., symbol–to–sound mapping knowledge) and "how" refers to the methods of instruction (e.g., copywriting). The analysis revealed the patterns of what orthography knowledge of Manchu was instructed and in what method it was instructed.

Symbol Frequency Count. Frequency of symbols was found to influence reading in a variety of orthographies (e.g., Spanish: Carreiras, Alvarez, & de Vega, 1993; German: Conrad & Jacobs, 2004; Hebrew: Frost et al., 1987; Chinese: Liversedge et al., 2014; Kannada: Nag, 2021; Nag et al., 2014; Korean: Simpson & Kang, 2004; Japanese kanji: Tamaoka & Kiyama, 2013). Thus, the frequency of uju and Manchu phoneme markers was measured and used as a control variable in correlation test. A mini corpus was built to include all the uju and Manchu phoneme markers involved in the literacy practices during the ten sessions Manchu instruction and their frequency was hand-counted. The criterion of counting frequency was whether the learners paid incessant "attention" to a given orthographic unit. If the learners continuously attended to an orthographic unit without any interruption, the frequency of the unit was counted 1. For example, in a teacher-led choral reading of the uju <A>, although the chant went on for multiple times, the frequency of this unit was counted as 1 because the learners' attention has been focused on <A> during this period. However, if the teacher proceeded to teach another orthographic unit after completing the instruction on an orthographic unit, and later he went back to the penultimate unit, the frequency of this unit was counted as 2. For instance, the teacher first instructed the orthographic unit <E> and then proceeded to <I>, after which he guided the students' attention to <E> again. In this scenario, the frequency of <E> was counted 2 because students paid attention to <E> twice. The frequency list is provided in Appendix D.

Data Integration. The mixed methods data analysis consists of the analysis of both quantitative and qualitative data and the integration of both forms of data (Onwuegbuzie & Teddlie, 2003). For mixed methods data analysis, this study employed the framework of integrated data analysis and interpretation within a convergent design (Creswell & Plano Clark, 2018). For the intent of integration, this study combined the quantitative and qualitative results and interpretation to develop a more comprehensive understanding of the Manchu orthography instruction in the university. In the integration procedure, data transformation was used. The quantitative data was first quantitatively analysed and the observation data was qualitatively analysed. After that, the qualitative theme on orthographic units (i.e., uju or phoneme markers) used by the teacher for Manchu orthography instruction was transformed into quantitative frequencies and incorporated into the existing quantitative data for additional analysis. Finally, the mixed methods data analysis advanced interpretation of how the analysis from the transformed data could provide additional insight into the research questions.

3.9 Chapter Summary

This chapter first discussed that mixed methods research (MMR) was appropriate for the present study because it fulfilled the needs of answering the quantitative and qualitative research questions and meeting the research objective. Then, after considering the epistemological and ontological issues on MMR, this study decided to adopt a pragmatic approach (Morgan, 2007) in examining the Manchu writing system.

With respect to the research design, the present study was devised as a concurrent multilevel design with quantitative measurements being the principal tool for data collection. In terms of sampling, the study employed the concurrent design using multilevel samples selected from two populations of students and teachers for the quantitative and qualitive components. For the quantitative component, 196 students were randomly selected for Test 1

on Manchu symbol visual complexity and Test 2 on Manchu symbol–sound mapping complexity. For the qualitative component, one teacher of Manchu was selected for nonparticipant classroom observation.

In terms of the instruments, visual complexity of the symbols in Test 1 was measured by of pixel count, perimetric complexity, number of simple features, number of connected points, and number of disconnected components. Symbol–sound mapping complexity of the symbols in Test 2 was measured by nature of the phoneme marker–sound mapping of "target" phoneme markers. In addition, the participants' decoding performance in Tests 1 and 2 was measured by error rate for accuracy and reaction time for speed.

For data collection, this study adopted a convergent design (Creswell & Plano Clark, 2018), which involved gathering the two strands of quantitative and qualitative data at almost the same time. The implementation matrix (Creswell & Plano Clark, 2018) was applied to list side by side the research questions, the samples and the specific data sources (quantitative and qualitative), and the data gathering instruments to answer the research questions.

Regarding test design, Test 1 and Test 2 were designed as recognition tests, in which the participants were asked to name Manchu symbol blocks. Test 1 (item $n = 18$) examined the association between visual complexity and symbol block recognition. Test 2 (item $n = 16$) investigated symbol–sound mapping complexity in Manchu by what orthographic unit (uju or phoneme marker) was adopted in symbol block reading. In addition, nonparticipant classroom observation was conducted on the ten sessions of Manchu orthographic knowledge instruction. The observation was focused on the orthographic unit (i.e., uju or phoneme markers) adopted by the teacher in instruction and the frequency of uju and phoneme markers given in the observed sessions.

Finally, for quantitative data analysis, partial correlation test (with frequency of symbol blocks in Test 1 controlled) was performed to Test 1 to analyse the association

between the five measures of symbol visual complexity (i.e., pixel count, perimetric complexity, simple features, connected points, and disconnected components) and the error rate and reaction time of the participants' decoding. Moreover, a stepwise multiple linear regression test was carried out with the five measures of visual complexity as predictors to build a model predicting the learners' decoding performance in recognising Manchu symbol blocks. For Test 2, paired sample *t*-tests were performed to compare the participants decoding performance in naming the symbol blocks presented in the "blocked" and "mixed" condition, and the decoding performance on symbol blocks between Block 1 and Block 2. In addition, the qualitative data obtained from classroom observation was first organised in a synthesis matrix. Then, the dimensions in the synthesis matrix were applied as codes for initial data coding, from which the sub-codes were developed. A thematic extraction was subsequently applied to qualitative data analysis to identify what knowledge of Manchu orthography was instructed and in what methods it was instructed. In the integration procedure, the qualitative theme on orthographic units (i.e., uju or phoneme markers) used by the teacher for Manchu orthography instruction was transformed into quantitative frequencies and incorporated into the existing quantitative data for additional analysis.

Chapter 4 Results for Test 1 on Visual Complexity

4.1 Chapter Overview

This chapter presents the results for Test 1, which examined the association between visual complexity of the Manchu uju symbol blocks and the novice learners' decoding performance. The chapter begins by introducing the purpose, the items, and procedure of Test 1. Then, it discusses the application of multiple imputation to dealing with the missing data in Test 1. Following that is the presentation of the results for Test 1, including the association between symbol-level measures (i.e., PX, PC, SF, CP and DC) of uju visual complexity and outcome-level measures of ER and RT in uju decoding. It also presents the models for predicting novice Manchu learners' symbol block decoding performance with respect to accuracy and speed.

4.2 Introduction to Test 1

4.2.1 Purpose of Test 1

Test 1 aimed to examine the extent to which the visual complexity of Manchu symbol blocks associates with novice learners' decoding performance. Previous studies have identified five dimensions related to visual complexity that could lead to difficulty in learning to read in different orthographies. These dimensions include (a) the visual feature density indexed by pixel count (i.e., the "ink" area) in a printed symbol (Nag et al., 2014), (b) the perimetric complexity sensitive to the ratio of perimeter squared over "ink" area of a symbol (Chang et al., 2018; Pelli et al., 2006; Watson, 2012), (c) the number of the simple features sensitive to strokes (Tamaoka & Kiyama, 2013; Wu et al., 1999), (d) the number of connected points in a symbol sensitive to continuity (Lanthier et al., 2009), and (e) the number of disconnected components in a symbol sensitive to discontinuity (Gibson, 1969; Winskel, 2009). All these dimensions of visual complexity were found to be strongly, negatively associated with symbol learning across 131 writing systems (Chang et al., 2016;

Nag et al., 2014). However, little examination has been conducted on the under-investigated Manchu writing system, which is also featured by its visually complex orthography. Therefore, Test 1 of the present study examined the association between visual complexity of Manchu uju symbol blocks and novice learners' decoding performance.

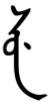
4.2.2 Items in Test 1

Eighteen Manchu uju symbol blocks were given to the participants for recognition in a naming task. Considering the participants had received only 10 sessions of Manchu orthography instruction at the time of testing, the symbol blocks were selected based on three criteria—length, syllable structure, and non-ligature. First, the symbol blocks in Test 1 comprise one to three Manchu phoneme markers, which corresponded with the length of symbol blocks instructed in the Manchu orthography class. Second, the symbol blocks were in V, CV and CVC syllables, which were the types of syllables distributed in the first and most fundamental section of the Twelves Sections of Manchu Syllabary. Yet, symbol blocks in CV syllable written in ligature were excluded from Test 1. In a CV ligature in Manchu, the consonant marker and vowel marker are nested together, while consonant and vowel markers in a non-ligature CV block are connected by a vertical line segment. This architectural difference in symbol design may affect uju processing; thus, ligatured CV uju were excluded. Based on these selection criteria, a list of 18 symbol blocks were selected. Visual complexity of the symbol blocks in Test 1 was manipulated with respect to their PX, PC, SF, CP and DC (Chang et al., 2018; Nag et al., 2014). Finally, frequency of the symbol blocks was also counted based on the classroom observation of the 10 sessions of Manchu orthography instruction. Table 11 provides symbol-level measures of the items in Test 1.

Table 11

Symbol-Level Measures of Visual Complexity for the Manchu Uju Symbol Blocks (n = 18) in Test 1

List of uju		Syllable structures	Measures					Uju frequency [f]
Traditional script	Latin script		PX [a]	PC [b]	SF [c]	CP [d]	DC [e]	
	*a*	V	8292	7.61	3	2	1	12
	*e*	V	8871	9.24	4	2	1	3
	*i*	V	6622	7.22	4	2	1	10
	*u*	V	7740	8.10	4	3	2	7
	*ū*	V	8661	8.15	5	4	1	4
	*ca*	CV	9933	8.69	6	4	1	1
	*du*	CV	10034	9.28	7	6	2	1
	*ha*	CV	12678	12.49	7	3	2	1
	*je*	CV	8651	8.72	4	2	2	0
	*me*	CV	10384	10.60	5	3	2	0
	*ni*	CV	7559	8.49	5	2	2	2

List of uju		Syllable structures	Measures					Uju frequency [f]
Traditional script	Latin script		PX [a]	PC [b]	SF [c]	CP [d]	DC [e]	
	*so*	CV	8320	7.78	6	5	1	0
	*to*	CV	8096	7.98	5	5	1	1
	*kib*	CVC	15804	15.90	8	6	1	0
	*lan*	CVC	11663	10.87	7	5	1	0
	*sek*	CVC	13923	14.02	12	7	2	0
	*sun*	CVC	14982	13.49	9	6	2	0
	*yor*	CVC	14858	13.19	9	8	1	0

[a] PX = pixel count.

[b] PC = perimetric complexity.

[c] SF = simple features.

[d] CP = connected points.

[e] DC = disconnected component.

[f] Frequency of uju symbol blocks was obtained from observing ten sessions of 45-minute instruction on Manchu orthography knowledge in the participating university. See “Symbol Frequency Count” under Section 3.8.2 in Chapter 3 and Section 6.3.4 in Chapter 6 for more details.

4.2.3 Procedure of Test 1

Prior to Test 1, participants were given a trial, in which 12 uju symbol blocks in V, CV, CVC, and CVCV syllables were given for recognition in a naming task. The trial aimed at familiarizing participants with naming task. After that, in Test 1 on visual complexity, a list of 18 Manchu symbol blocks were given to the participants to name. The items were presented in a Dell desktop computer with 21.5-inch monitor of 1920 × 1080 resolution. An item began with the presentation of a fixation cross (+) in the centre of a PowerPoint slide. Then, a Manchu symbol block was presented at fixation and remained there until the participant gave a response, after which the tester continued to present the fixation cross for the upcoming symbol block. Participants were seen individually by a tester in a quiet room in the university. The testing was audiorecorded.

4.3 Multiple Imputation of the Missing Data in Test 1

An initial analysis of the data collected from the Manchu symbol block (item $n = 18$) naming task in Test 1 shows that there are missing values of the participants' decoding performance. Specifically, in Test 1 the participants ($n = 196$) gave 1084 correct responses (30.73%) and 1932 incorrect responses (54.76%), with 512 items unattempted (14.51%). This result suggests that for the estimation of item decoding error rate, 14.51% of the data was missing. For reaction time estimate, a more substantial proportion of data (69.27%) was missing because, in addition to the unattempted items, the incorrect responses were as well excluded from being measured for reaction time. To address this problem, simply deleting the incomplete cases may result in too few observations being left for the analysis. Thus, the missing data, if not handled appropriately, would produce inaccurate parameter estimates which in this study are the Manchu item decoding accuracy and speed, leading to invalid inference of the population of interest (Sinharay, Stern, & Russell, 2001). Although some *ad hoc* methods, such as mean substitution, hot deck imputation, listwise deletion, pairwise

deletion, regression imputation, etc., were helpful in dealing with data missingness, they could produce biased estimates of the mean and the standard error in certain conditions (see Enders, 2010; Graham, Cumsille, & Shevock, 2012; Schafer & Graham, 2002; Van Buuren, 2018; van Ginkel, Linting, Rippe, & van der Voort, 2020). Alternatively, the method of multiple imputation is more recommended (Schafer & Graham, 2002), because it uses multiple values drawn from the posterior predictive distribution of the missing data to complete the dataset and yields an unbiased estimation by averaging these values (Rubin, 1987). Therefore, this study adopted multiple imputation, which is now accepted as "the best general method to deal with incomplete data" (Van Buuren, 2018, p. 25), to address the issue.

4.3.1 Missing Data Mechanism

Rubin (1976) classified the mechanisms of missing data into three categories, which describe the probability of missing values in relation to the data (Enders, 2010; Zhou, Zhou, Liu, & Ding, 2014). The data are missing completely at random (MCAR) if the probability of missing data on a variable is independent of other measured variable(s) and of te values of the variable itself. That is, the lack of data is completely unrelated to the data itself and occurs by chance. A missing at random (MAR) mechanism occurs when the probability of missingness on a variable is related to some other variable(s) but not to the values in the variable itself (e.g., y is missing if the observed $x < 0$). Finally, the data are missing not at random (MNAT) if the probability of missing data on a variable is dependent on other variables as well as on the values of the variable itself (e.g., y is missing if the observed $x < 0$ and the unobserved $y < 0$).

4.3.2 Assumptions Required by Multiple Imputation

Although multiple imputation assumes an MAR mechanism of missing data (Graham et al., 2012), a practical issue with this assumption is that there is no way to confirm that the probability of missing data on a variable is only a function of other measured variables (i.e.,

an MAR mechanism, Enders, 2010). In other words, it is not possible to check whether this assumption is met, nor could it rule out the possibility that the missing data is MNAT without knowing the missing values (Sinharay et al., 2001; Van Buuren, 2018). However, it was suggested that it is a misconception that multiple imputation should be only used on the MAR data (van Ginkel et al., 2020), because even for MNAR data, multiple imputation is "at least as good as" the traditional method such as list deletion (Graham et al., 2012, p. 137), or even better since it gives less biased results than do the old methods (Schafer, 1997). Moreover, it was also suggested that if there are strong associations between the variables, using multiple imputation would usually produce sound results regardless of the missing data mechanism (Rubin, 1996; Sinharay et al., 2001).

In the present study, missing data in Test 1met the requirement of MAR mechanism. By definition, MAR data in a variable is dependent on some measured variable(s) but not the variable itself. First, since word recognition was found to be associated with symbol visual complexity measured by the dimensions of PX, PC, SF, CP and DC (Chang et al., 2018; Nag et al., 2014), it can be inferred that ER and RT of the uju symbol blocks in Test 1 may also be dependent on these variables. Second, the missing data in decoding performance were not dependent on the variable itself. The missing values in ER were due to the unattempted items, which is unrelated to ER itself. Moreover, the missingness of RT data was derived from (a) the incorrect responses and (b) the unattempted items. In either case, the missingness could occur regardless of how long the participants spent responding to the items. For the above reasons the missing data in Test 1 were considered MAR.

4.3.3 Steps of Multiple Imputation

Multiple imputation is a three-step process, which includes the imputation phase, the analysis phase, and the pooling phase (Enders, 2010). In the imputation phase for the missing data in Test 1, this study created multiple copies of imputed data sets for decoding accuracy

(m = 20) and decoding speed (m = 40), with each data set containing different imputed values for every missing accuracy value (correct = 1, incorrect = 0) and speed value (in milliseconds). While it is traditionally suggested that three to five imputations was sufficient, it is recommended that 20 and 40 imputations be needed based on a respective fraction of 30% and 70% missing information as well as a less than 1% tolerance for power falloff (Graham, Olchowski, & Gilreath, 2007). In the analysis phase, these complete data sets were statistically analysed, which yielded 20 and 40 sets of estimates of the participants' mean ER and mean RT to the test items (n = 18). Finally, the pooling phase combined these means into a single set of results, and used Rubin's (1987) formulas to obtain the average mean ER and RT of the items.

4.3.4 Results of Multiple Imputation of Missing Data in Test 1

Following the reporting guidelines (Van Buuren, 2018), this section provides the detailed description of multiple imputation used to deal with the missing values in Test 1.

Reasons for Missing data. The values were missing primarily because of the low Manchu proficiency of the participants, who had received only ten sessions of 45-minute instruction on Manchu orthography before taking the test of uju symbol blocks recognition. The missingness was also related to the outcome measurement. For the estimation of item decoding ER, the 512 unattempted items (14.51%) were excluded from being measured and thus the values were missing. Moreover, when the reaction time was estimated, both the incorrect responses (1932 observations, 54.76%) and the unattempted items (512 observations, 14.51%) were excluded, resulting in a more substantial proportion (69.27%) of missing data.

Method. As previously discussed, the missing data in Test 1 have met the assumption of missing at random (MAR) mechanism, which made it possible to obtain the imputed values using the relationship among the variables from the observed data (Sinharay et al.,

2001). In addition, since the missingness of the outcome variable was considered to be associated with the five dimensions (i.e., PX, PC, SF, CP, and DC) measuring symbol visual complexity, they were included as the independent variables to impute the missing values in the outcome variables of decoding accuracy (correct response = 1, incorrect response = 0) and decoding speed (in milliseconds) for each case.

Software. Incomplete variables were imputed using IBM SPSS Statistics (Version 28). The imputation was conducted under predictive mean matching (PMM), which is arguably the most appropriate method for multiple imputation, and under fully conditional specification (FCS), which is evidenced to yield unbiased estimates (Van Buuren, 2018). Moreover, with respects to the roles played by the variables in the model, PX, PC, SF, CP, and DC were used as "predictors only", and the values of decoding accuracy and decoding speed of each participant was "imputed only".

Number of Imputed Datasets. Multiple copies of imputed data sets were created for decoding accuracy (m = 20) and decoding speed (m = 40), with each data set containing different imputed values for the missing accuracy value (correct response = 1, incorrect response = 0) and the missing speed value (in milliseconds). The number of imputed data sets was based on the recommendation that for a fraction of 30% and 70% missing information (which is close to the percentage of data missingness in Test 1) and a less than 1% tolerance for power falloff, the number of 20 and 40 imputed data sets should be needed (Graham et al., 2007).

Diagnostics. Following the imputation phase, diagnostics were performed to see whether the imputed values were plausible in a sense that they could have been plausibly measured had they not been missing (Van Buuren, 2018). The diagnostics showed that the imputed value of the missing ER value was either 0 or 1, which was consistent with the measure of decoding accuracy (correct response = 1, incorrect response = 0). Moreover, a

number of 2994 (95.59%) observations in the pooled RT of the items (n = 18) of 174 cases in Test 1 was within two standard deviations of the means, which suggested the RT data is appropriate for statistical analysis.

Pooling. The sets of parameter estimates of the mean decoding accuracy (m = 20) and the mean decoding speed (m = 40) of the Manchu items (n = 18) in Test 1, yielded from the analysis phase, were pooled into a single result using Rubin's (1987) formula

$$\bar{\theta} = \frac{1}{m}\sum_{t=1}^{m}\hat{\theta}_t$$

where $\hat{\theta}_t$ is the parameter estimates from data set t and $\bar{\theta}$ is the pooled estimate (Enders, 2010). This equation suggests that for Test 1 of the present study, the multiple imputation point estimate of the mean ER and the mean RT of the Manchu symbol blocks is the arithmetic average of the 20 and 40 estimates, respectively. Table 14 gives the pooled results of ER and RT of the items in Test 1.

4.4 Results of Test 1

4.4.1 Participant Demographic Information

A hundred and ninety-six first year university students (n = 196, M_{age} = 18.79, SD = 0.59, age range = 17.25–20.75 years) were recruited from a population of 1,851 freshmen who took the course *Introduction to Manchu Intangible Cultural Heritage* in the participating university. A screening was made to ensure that the recruited participants were all novice learners with no prior learning experience of Manchu. Table 12 summarises the basic information of the participants. Questionnaire of collecting their basic information was provided in Appendix C.

Table 12

Demographic Information of the Participants

Characteristics	*n*	%
Gender		
Female	176	89.8
Male	20	10.2
Previous learning of Manchu [a]		
Yes	0	0
No	196	100
Extracurricular contact with Manchu in university [b, c]		
Doing homework, revision, or preview	176	93.1
Reading Manchu books	4	2.1
Singing Manchu songs	22	11.6
No contact	13	6.35
Others [d]	28	14.8

Note. Participants n = 196. The average age is 18.79 years (*SD* = 0.59).

[a] Previous learning refers to the learning experience in both formal and informal educational contexts before entering university.

[b] *n* = 189. Seven participants giving contradictory answers to this question were excluded.

[c] Participants were asked to select all applicable answers.

[d] Answers specified by the participants include "watching Manchu instruction video" (*n* = 10), "watching Manchu video" (*n* = 11), "watching video" (*n* = 6), and "using meme stickers containing Manchu words" (*n* = 1).

4.4.2 Participants' Overall Performance in Manchu Symbol Blocks Recognition in Test 1

The data of 196 participants who were assessed for their ability to name 18 Manchu symbols in Test 1 were presented. Their overall performance in symbol naming was described by the mean, standard deviation, and range of the symbol recognition scores (i.e.,

the number of correct symbol naming) in Table 13. The distribution of scores is positively skewed (skewness = .808) and leptokurtic (kurtosis = .137). Sixteen participants (8.16%) scored zero in the test.

Table 13

Mean, Standard Deviation, and Range of Participants' Uju Symbol Block Recognition Scores in Test 1 (Max. Score = 18)

Measure	*n*	*M*	*SD*	Min	Max
Uju recognition score [a]	196	5.53	3.99	0	17

[a] Uju recognition score of a participant equals the number of uju (out of 18) that he or she correctly named.

4.4.3 Descriptive Data of Symbol-Level and Outcome-Level Measures

At the symbol level, the visual complexity of an item was measured by (a) pixel count; (b) perimetric complexity; (c) simple feature; (d) connected point; and (e) disconnected component. At the outcome level, the participants' decoding performance was measured by (a) error rate and (b) reaction time. Table 11 shows the data of the five visual complexity measures as well as frequency of the 18 Manchu symbol blocks. Table 14 shows the participants' error rate and reaction time in symbol block decoding. Table 15 presents the means, standard deviations, and ranges of the symbol- and outcome-level measures.

Table 14

Outcome-Level Measures of Participants' Decoding Performance for the Manchu Uju Symbol Blocks (n = 18) in Test 1

List of uju		Measures		
Traditional script	Latin script	ER [a]	RT [b]	
			M	*SD*
	*a*	.280	3678.59	2336.62
	*e*	.421	4600.32	2846.34
	*i*	.379	4196.26	2909.25
	*u*	.486	5418.97	2890.20
	*ū*	.460	5712.49	3177.32
	*ca*	.715	4952.82	2463.29
	*du*	.816	6241.48	3687.72
	*ha*	.486	5519.39	3187.67
	*je*	.807	5297.85	2877.97
	*me*	.669	4854.37	2398.38

List of uju		Measures		
Traditional script	Latin script	ER [a]	RT [b]	
			M	*SD*
	*ni*	.564	5292.99	3448.45
	*so*	.781	5525.05	2949.62
	*to*	.761	5891.84	3595.92
	*kib*	.821	5805.94	3521.21
	*lan*	.812	5635.15	2914.13
	*sek*	.848	6074.23	4080.03
	*sun*	.806	5833.57	2928.04
	*yor*	.895	5901.35	3342.27

[a] ER = error rate.

[b] RT = reaction time (in ms).

Table 15

Means, Standard Deviations, and Ranges of Symbol-Level and Outcome-Level Measures in Test 1

Measures	*n*	*M*	*SD*	Min	Max
Symbol-level					
PX [a]	18	10392.83	2885.45	6622	15804
PC [b]	18	10.10	2.62	7.22	15.90
SF [c]	18	6.11	2.30	3	12
CP [d]	18	4.17	1.92	2	8
DC [e]	18	1.44	0.51	1	2
Outcome-level					
ER [f]	18	0.656	0.192	0.280	0.895
RT [g]	18	5357.37	674.34	3678.59	6241.48

[a] PX = pixel count.

[b] PC = perimetric complexity.

[c] SF = simple features.

[d] CP = connected points.

[e] DC = disconnected components.

[f] ER = error rate. ER was calculated based on 20 data sets derived from multiple imputation (m = 20).

[g] RT = reaction time (milliseconds per item). RT was calculated based on 40 data sets derived from multiple imputation (m = 40).

4.4.4 Associations among Symbol-Level Measures and between Outcome-Level Measures

Table 16 shows the association among the five symbol-level measures of visual complexity. All measures are significantly correlated with each other, except for disconnected components. Pixel count showed strong positive correlation with perimetric complexity (r = .970, p < .001) and simple features (r = .842, p < .001); and moderate positive correlation with connected points (r = .733, p < .001). Perimetric complexity has

strong positive correlation with simple features ($r = .819$, $p < .001$) and moderate positive correlation with connected points ($r = .641$, $p = .004$). Connected points and simple features are strongly positively correlated ($r = .850$, $p < .001$). Disconnected components showed weak, nonsignificant correlation with the remaining measures: pixel count ($r = .112$, $p = .658$), perimetric complexity ($r = .192$, $p = .444$), simple features ($r = .206$, $p = .413$), and connected points ($r = –.080$, $p = .752$).

Table 16

Correlation among Symbol-Level Measures for Test 1

Measures	PX [a]	PC [b]	SF [c]	CP [d]	DC [e]
PX	—				
PC	.970***	—			
SF	.842***	.819***	—		
CP	.733***	.641**	.850***	—	
DC	.112	.192	.206	–.080	—

[a] PX = pixel count.

[b] PC = perimetric complexity.

[c] SF = simple features.

[d] CP = connected points.

[e] DC = disconnected components.

**$p < .01$, ***$p < .001$.

4.4.5 Association between Symbol-Level and Outcome-Level Measures

As Table 17 presents, when frequency of the uju symbol blocks in Test 1 was not controlled in the correlation test, ER shows strong correlation with CP ($r = .777$, $p < .001$); and moderate correlation with SF ($r = .688$, $p = .002$), PX ($r = .588$, $p = .010$) and PC ($r = .533$, $p = .023$). RT is strongly correlated with CP ($r = .757$, $p < .001$), and moderately

correlated with SF (r = .688, p = .002), PX (r = .519, p = .027) and PC (r = .497, p = .036). DC has weak, nonsignificant association with the outcome-level measures: ER (r = .141, p = .578) and RT (r = .286, p = .251).

However, when frequency of the symbol blocks in Test 1 was controlled in correlation test, CP was found to be the unique symbol-level measure showing significant moderate correlation with ER (r = .688, p = .002) and RT (r = .621, p = .008). The remaining symbol-level measures of PX, PC, SF and DC were not significantly associated with ER (r_{PX-ER} = .274, p = .288, r_{PC-ER} = .173, p = .506, r_{SF-ER} = .427, p = .087, and r_{DC-ER} = –.115, p = .661) or RT (r_{PX-RT} = .194, p = .457, r_{PC-RT} = .169, p = .516, r_{SF-RT} = .454, p = .067, and r_{DC-RT} = .164, p = .529).

Table 17

Correlation between Symbol-Level and Outcome-Level Measures for Test 1

Measures	Symbol-level				
	PX [a]	PC [b]	SF [c]	CP [d]	DC [e]
Outcome-level					
ER [f]					
Uju freq. not controlled	.588*	.533*	.688**	.777***	.141
Uju freq. controlled	.274	.173	.427	.688**	–.115
RT [g]					
Uju freq. not controlled	.519*	.497*	.688**	.757***	.286
Uju freq. controlled	.194	.169	.454	.621**	.164

[a] PX = pixel count.

[b] PC = perimetric complexity.

[c] SF = simple features.

[d] CP = connected points.

[e] DC = disconnected components.

[f] ER = error rate.

[g] RT = reaction time.

$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

4.4.6 Predictors for Manchu Symbol Block Decoding Performance

Multiple linear regression was conducted to determine the best linear combination of the five symbol-level measures of visual complexity (i.e., pixel count, perimetric complexity, the number of simple features, the number of connected points, and the number of disconnected components) for predicting novice Manchu learners' decoding accuracy and decoding speed of uju symbol blocks. First, to predict decoding accuracy, the five independent variables of PX, PC, SF, CP and DC were all entered into the regression equation for analysis. The regression model showed that the combination of these variables significantly predicted participants' decoding accuracy ($F(5, 12) = 4.447$, $p = .016$). The R^2

value was .649, suggesting that almost 65% of the variance in the learners' decoding accuracy was explained by the model. However, there were two issues with this model. The first was that the independent variable of CP was the only significant predictor contributing to the prediction (p = .045), with the remaining four independent variables of PX (p = .894), PC (p = .863), SF (p = .711) and DC (p = .261) insignificant. Second, multicollinearity between PX, PC and SF was also a concern. PX and PC demonstrated possible collinearity which was suggested by the variance inflation factor (VIF) value over the critical value of 10 (VIF_{PX} = 27.99; VIF_{PC} = 25.27; Vittinghoff, Glidden, Shiboski, & McCulloch, 2012). Furthermore, SF was also likely to show collinearity with PX and PC because the correlation coefficient between them ($r_{SF\text{-}PX}$ = .842, $r_{SF\text{-}PC}$ = .819, $r_{PX\text{-}PC}$ = .970) was higher than the threshold value of .8 (Young, 2018).

Second, to predict novice Manchu learners' decoding speed of uju symbol blocks, the five independent variables of PX, PC, SF, CP and DC were all entered into the regression equation for analysis. The regression model showed that the combination of these variables significantly predicted participants' decoding speed ($F(5, 12) = 6.830$, p = .003). The R^2 value was .740, suggesting 74% of the variance in the learners' decoding speed was explained by the model. However, this model was problematic in that the independent variables of PX (p = .250), PC (p = .318), and SF (p = .585) did not make significant contribution to the model, while CP and DC significantly contributed to the prediction (ps = .007 and .047, respectively). Another issue in this model is the multicollinearity existing between PX, PC and SF, as reported above.

To address the issue of multicollinearity, one suggested remedy is to drop one or more highly correlated independent variables, if they measure the same construct for theoretical reasons (O'Brien, 2007). PX and PC are related to each other because they both use "area" of symbol in measuring visual complexity (PX: the perimeter squared over the area of a symbol;

PC: the area of a symbol). In addition, PX and PC were found to be closely related to the number of visual features (Nag et al., 2014; Pelli et al., 2006). Thus, it could be inferred that PX, PC, and SF probably measure the same construct. Based on this reason, PX, PC, and SF was entered, respectively, with CP and DC as the three independent variables in performing multiple linear regression to predict symbol block decoding accuracy and decoding speed. The regression models showed that the three sets of independent variables (PX, CP, and DC; PC, CP, and DC; SF, CP, and DC) all significantly predicted participants' decoding accuracy and speed. However, for decoding accuracy prediction, CP was the only significant predictor in the models. Moreover, for decoding speed prediction, CP and DC contributed significantly to the models. Tables 18 and 19 give detailed statistics of the models.

Table 18

Models of Decoding Accuracy from Multiple Linear Regression (Enter) in Test 1

Models	Predictors	*R*	R^2	*F* (3, 14)	*p*	Estimate	*SE*	*95% CI* [a]		*t* (14)	*p*
								LL [b]	*UL* [c]		
1		.803	.645	8.486	.002						
	Constant					.226	.136	–.066	.517	1.659	.119
	PX [d]					–2.542E–6	.000	.000	.000	–.153	.881
	CP [e]					.082	.024	.030	.134	3.403	.004
	DC [f]					.079	.062	–.054	.212	1.275	.223
2		.803	.645	8.480	.002						
	Constant					.224	.138	–.071	.520	1.629	.126
	PC [g]					–.002	.016	–.036	.032	–.128	.900
	CP					.081	.022	.035	.127	3.767	.002
	DC					.079	.063	–.056	.215	1.252	.231
3		.805	.648	8.606	.002						
	Constant					.215	.117	–.035	.465	1.842	.087
	SF [h]					–.011	.029	–.074	.052	–.388	.704
	CP					.091	.035	.017	.165	2.637	.020
	DC					.091	.070	–.059	.240	1.298	.215

[a] *CI* = confidence interval.

[b] *LL* = lower limit.

[c] *UL* = upper limit.

[d] PX = pixel count.

[e] CP = connected points.

[f] DC = disconnected components.

[g] PC = perimetric complexity.

[h] SF = simple features.

Table 19

Models of Decoding Speed from Multiple Linear Regression (Enter) in Test 1

Models	Predictors	*R*	R^2	*F* (3, 14)	*p*	Estimate	*SE*	*95% CI* [a]		*t* (14)	*p*
								LL [b]	*UL* [c]		
1		.846	.715	11.722	< .001						
	Constant					3766.03	428.67	2846.62	4685.43	8.785	< .001
	PX [d]					–.052	.051	–.160	.057	–1.021	.325
	CP [e]					334.23	75.96	171.31	497.15	4.400	< .001
	DC [f]					509.56	194.97	91.40	927.73	2.614	.020
2		.839	.704	11.110	< .001						
	Constant					3704.51	442.23	2756.02	4652.99	8.377	< .001
	PC [g]					–35.65	51.43	–145.95	74.65	–.693	.499
	CP					308.30	69.17	159.95	456.66	4.457	< .001
	DC					504.260	202.89	69.11	939.41	2.485	.026
3		.841	.707	11.252	< .001						
	Constant					3540.78	374.37	2737.84	4343.73	9.458	< .001
	SF [h]					–73.83	94.44	–276.39	128.73	–.782	.447
	CP					353.28	111.13	114.94	591.62	3.179	.007
	DC					550.90	224.36	69.69	1032.11	2.455	.028

[a] *CI* = confidence interval.

[b] *LL* = lower limit.

[c] *UL* = upper limit.

[d] PX = pixel count.

[e] CP = connected points.

[f] DC = disconnected components.

[g] PC = perimetric complexity.

[h] SF = simple features.

The above models suggested that the independent variable of CP was the unique predictor of decoding accuracy, and the combination of CP and DC significantly contributed to predicting decoding speed. Thus, the independent variables of CP, and CP and DC, which made significant contribution in the regression, were selected and entered into linear regression to produce models for predicting decoding accuracy and decoding speed. Statistical assumptions, such as the normal distribution of residuals and the non-linear correlation between predicted variables and residuals were met in the analysis (see Figures 8, 9, 10 and 11). The regression coefficients of the predictors, 95% confidence interval, and the *p* values are shown in Table 20.

With respect to the decoding accuracy of uju symbol blocks, the variable of CP uniquely predicted ($F(1, 16) = 24.322$, $p < .001$) the outcome variable of the ER made by novice Manchu learners in uju decoding. CP accounted for approximately 60% of the variance of decoding ER ($R^2 = .603$). The regression formulation is given below.

$$\text{Error rate} = 0.322 + .078 \times \text{the number of CP}$$

Moreover, for the prediction of decoding speed, the combination of predictor variables of CP and DC significantly predicted ($F(2, 15) = 17.013$, $p < .001$) the outcome variable of the RT spent by novice learners in uju decoding. CP and DC together explained approximately 69% of the variance of decoding RT ($R^2 = .694$). The regression formulation is given as follows.

$$\text{Reaction time} = 3543.177 + 276.126 \times \text{the number of CP} + 459.463 \times \text{the number of DC}$$

Figure 8

Distribution of the Standardized Residual of Error Rate (ER)

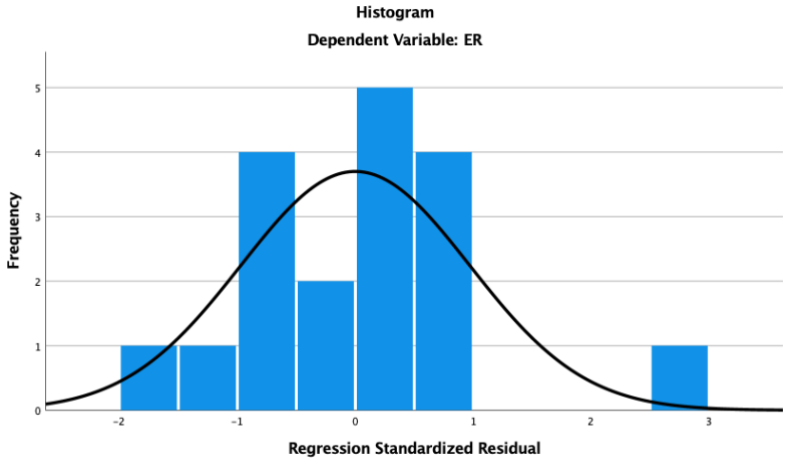


Figure 9

Standardized Predicted Values and Residual of Error Rate (ER)

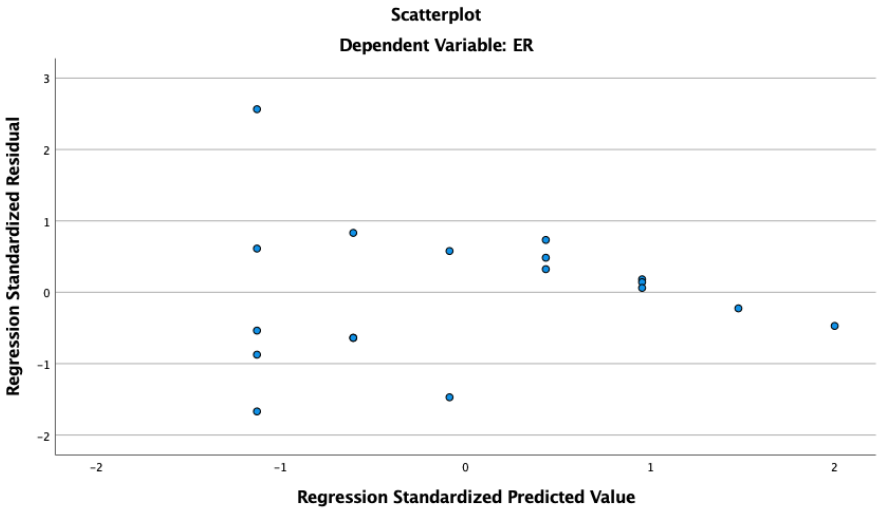


Figure 10

Distribution of the Standardized Residual of Reaction Time (RT)

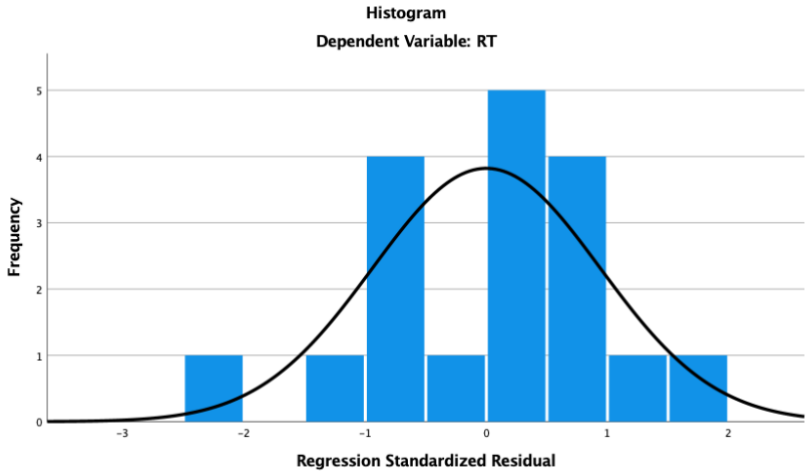


Figure 11

Standardized Predicted Values and Residual of Reaction Time (RT)

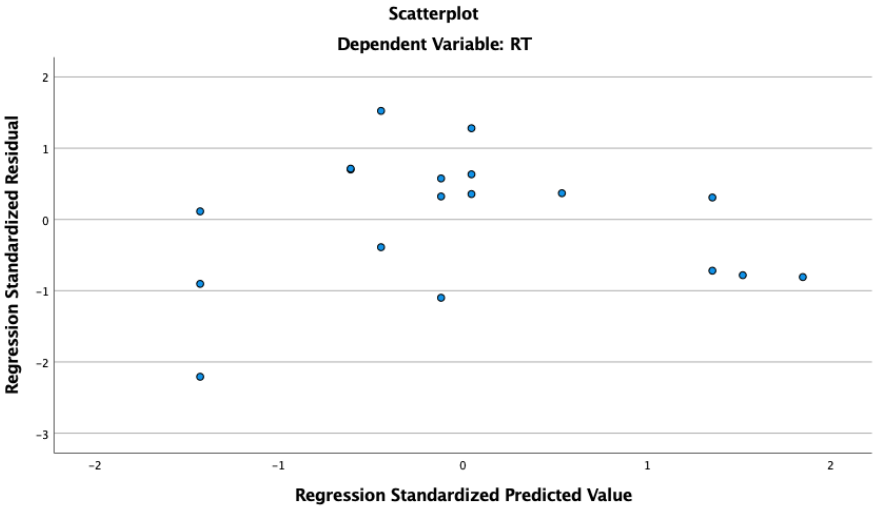


Table 20

Models of Decoding Accuracy and Decoding Speed from Multiple Linear Regression (Enter) in Test 1

Models	Predictors	Estimate	*SE*	95% *CI* [a]		*p*
				LL [b]	*UL* [c]	
Error rate						
	Constant	.322	.072	.180	.485	<.001
	CP [d]	.078	.016	.044	.111	<.001
Reaction time						
	Constant	3543.177	369.474	2755.663	4330.691	<.001
	CP	276.126	50.390	168.722	383.529	<.001
	DC [e]	459.463	188.962	56.700	862.226	.028

[a] *CI* = confidence interval.

[b] *LL* = lower limit.

[c] *UL* = upper limit.

[d] CP = connected points.

[e] DC = disconnected components.

4.5 Chapter Summary

This chapter presents the results of Test 1, which examined the novice learners' performance in decoding visually complex Manchu symbol blocks. Regarding the first research question (RQ) of the present study

> **RQ 1-1**: To what extent does the visual complexity of Manchu symbol blocks associate with the learners' decoding accuracy and speed?

it was found that PX, PC, SF, CP but DC of symbol blocks were positively associated with leaners' decoding ER and RT in Test 1. In addition, when frequency of the symbol blocks

given in Test 1 was controlled, only CP was found to be positively associated with ER and RT in decoding. This finding is interesting in that intuitively CP in an uju symbol block seemed to be a less relevant visual feature than are PX, PC and SF.

In addition, in answer to the second research question (RQ) on the predictors of novice Manchu learner's decoding performance,

> **RQ 1-2**: In what ways do the measures of the visual complexity of Manchu symbol blocks predict the learners' decoding accuracy and speed?

it was found that CP was the unique predictor for learners' decoding accuracy of uju symbol blocks, and moreover, CP and DC were found to be predictors for decoding speed. The models are provided as follows.

$$\text{Error rate} = 0.322 + .078 \times \text{the number of CP}$$

$$\text{Reaction time} = 3543.177 + 276.126 \times \text{the number of CP} + 459.463 \times \text{the number of DC}$$

Chapter 5 Results for Test 2 on Symbol–Sound Mapping Complexity

5.1 Chapter Overview

This chapter presents the results for Test 2, which examined the association between the symbol–sound mapping complexity of uju symbol blocks and the participants' decoding performance. The chapter begins by introducing the purpose, the list of items, item visual complexity control and frequency, manipulation of item presentation, and the procedure of Test 2. After that, it discusses the application of multiple imputation to dealing with the missing data in Test 2. Following that is the presentation of the results for Test 2, including participants' decoding performance on the items repeatedly presented in the "blocked" and "mixed" condition, and their performance in decoding items in Block 1 (all phoneme markers of an item in consistent one–to–one mapping with sound) and Block 2 (the target phoneme marker of an item in less consistent one–to–multiple mapping with sound).

5.2 Introduction to Test 2

5.2.1 Purpose of Test 2

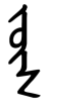
Test 2 aimed to identify what orthographic unit(s)—phoneme marker, uju, or a mixture of both—were adopted by novice learners in Manchu symbol block decoding. Previous studies have found that both letters and syllables were used by readers in decoding symbols across orthographies of varying levels of symbol–sound mapping complexity (e.g., Nag, 2017; Ziegler & Goswami, 2005). Yet, little was known as to what orthographic unit was applied to read the Manchu writing systems, which is also characterized by complex symbol–sound mapping (see Table 3). Therefore, Test 2 of the present study examined the association between mapping complexity of Manchu writing system and the orthographic unit(s) adopted by novice learners in uju symbol blocks decoding.

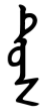
5.2.2 Items in Test 2

Sixteen uju symbol blocks in CV (n = 12) and CVC syllables were given to the participants for recognition in a naming task (see Table 21 for the uju items). The selection of CV and CVC syllables, rather than the more complex CVCV or CVCCV, was due to their correspondence with the length of uju symbol blocks instructed in the Manchu orthography class. The 16 uju given in Test 2 were organised into two blocks (Block 1 and Block 2) based on the symbol–sound mapping complexity. In Block 1, all phoneme markers in an uju were in consistent one–to–one mapping with sound. Yet, in Block 2 an uju was featured by a target phoneme marker which was in correspondence with either two sounds (item n = 5) or three sounds (item n = 3), with the remaining non-target marker(s) in one–to–one symbol–sound mapping. Table 21 presents the symbol–sound mapping of the uju in Test 2. Moreover, ligatured CV symbol blocks were excluded from Test 2. In a Manchu CV ligature, the consonant marker and vowel marker are nested together, while consonant and vowel markers in a non-ligature CV symbol block are connected by a vertical line segment. This architectural difference in symbol block design may affect uju processing; thus, ligatured CV uju were excluded.

Table 21

Measures of Symbol–Sound Mapping Relations, Visual Complexity, and Frequency of the Uju (n = 16) in Test 2

Blocks of items	List of items		Phoneme marker—sound mapping relation	Measures of item visual complexity					Frequency [f]
	Traditional	Latin		PX [a]	PC [b]	SF [c]	CP [d]	DC [e]	
Block 1 (Item *n* = 8)		*sū*	One-to-one	9940	9.06	8	6	1	0
		*ce*		10663	9.85	7	4	2	0
		*di*		8897	9.86	7	4	2	1
		*nū*		9120	9.05	6	4	2	2
		*ti*		8163	8.64	6	4	1	1
		*cū*		9760	8.73	8	6	1	0
		*jūs*		14224	13.47	9	7	1	0
		*tib*		17265	17.32	10	8	1	0

Blocks of items	List of items		Phoneme marker—sound mapping relation	Measures of item visual complexity					Frequency [f]
	Traditional	Latin		PX [a]	PC [b]	SF [c]	CP [d]	DC [e]	
Block 2 (Item *n* = 8)		*tu* [g]	One-to-two [h]	9303	8.11	6	6	1	1
		*we*	One-to-two	11638	10.74	5	3	2	0
		*de*	One-to-three	11827	10.51	7	5	2	1
		*du*	One-to-two	10034	9.27	7	6	2	1
		*te*	One-to-three	11096	9.35	6	5	1	2
		*fū*	One-to-two	10714	9.60	6	5	1	0
		*tus*	One-to-two	15372	13.38	10	9	1	0
		*teb*	One-to-three	17492	16.34	11	8	1	0

Note. The items were matched in pairs (*sū–tu*, *ce–we*, *di–de*, *nū–du*, *ti–te*, *cū–fū*, *jūs–tus*, and *tib–teb*) with respect to visual complexity, syllabic structure, and frequency.

[a] PX = pixel count.

[b] PC = perimetric complexity.

[c] SF = simple feature.

[d] CP = connected point.

[e] DC = disconnected component.

[f] Frequency of the items in Test 2 was counted based on the observation of ten sessions of 45-minute instruction on Manchu orthography knowledge in the participating university. See "Symbol Block Frequency Count" under Section 3.8.2 in Chapter 3 and Section 6.3.4 in Chapter 6 for more details.

[g] The target phoneme marker is underlined.

[h] The mapping relation is between the "target" phoneme marker and sound.

5.2.3 Control of Item Visual Complexity and Frequency

Previous studies have found that five measures of visual complexity—pixel count (Nag et al., 2014), perimetric complexity, the number of simple features, the number of connected points, and the number of disconnected components (Chang et al., 2018)—were associated with symbol reading performance; thus, it was important to control visual complexity of the pair-matched symbol blocks between Block 1 and Block 2 in Test 2. To achieve this goal, in the design phase of Test 2 the present study took into account the above five dimensions to assess whether visual complexity of the pair-match items was effectively controlled. Firstly, disconnected component (DC) was appropriately controlled because each paired uju had identical number of DC (e.g., DC n = 1 for *sū–tu*). Secondly, paired samples *t*-test was performed on pixel count (PX), perimetric complexity (PC), the number of simple features (SF), and the number of connected points (CP) of the paired uju. Table 22 gives the results. The results showed that the paired uju were not significantly different in terms of PC ($t(7) = –.604$, $p = .565$), SF ($t(7) = .753$, $p = .476$) or CP ($t(7) = –1.183$, $p = .275$), but they differed in PX ($t(7) = –2.725$, $p = .030$). The results suggested that visual complexity of the item pairs in Test 2 was effectively controlled if it was assessed by PC, SF and CP, but not if evaluated by PX. However, Test 1 of the present study has identified CP as the unique predictor for uju decoding accuracy, and CP and DC as predictors for uju decoding speed

(see Section 4.4.6 in Chapter 4). Thus, if assessed by the measures CP and DC, visual complexity of the pair-matched items between Block 1 and Block 2 in Test 2 was appropriately controlled.

Frequency of the symbol blocks in Block 1 and Block 2 was also controlled. Based on the classroom observation of 10 sessions of 45-minute Manchu orthography instruction, the frequency count of the uju presented in Test 2 was given in Table 21. The results of paired samples *t*-test performed on the frequency of paired uju between Block 1 and Block 2 suggested that the frequency was under control ($t(7) = –.552$, $p = .598$, $MD = –.125$).

Table 22

Results of the Paired-samples t-Tests of the Visual Complexity of the Uju between Block 1 and Block 2 in Test 2

Measures	Block 1		Block 2		*MD*	*t*(7)	*p*
	M	*SD*	*M*	*SD*			
PX [a]	11004.00	3126.49	12184.50	2803.85	–1180.5	–2.725	.030
PC [b]	10.75	3.08	10.91	2.68	–.17	–.604	.565
SF [c]	7.63	1.41	7.25	2.12	.38	.753	.476
CP [d]	5.38	1.60	5.88	1.89	–.50	–1.183	.275

[a] PX = pixel count.

[b] PC = perimetric complexity.

[c] SF = simple features.

[d] CP = connected points.

5.2.4 Manipulation of Item Presentation

Items in Block 1 ($n = 8$) and Block 2 ($n = 8$) were presented to the participants in "blocked" and "mixed" conditions (Ziegler & Goswami, 2005). In the "blocked" condition, the items were given sequentially by blocks (Block 1 and then Block 2). After that, in the

"mixed" condition, items in Block 1 and Block 2 were mixed up before being presented to participants. The manipulation of item presentation aimed to identify the orthographic unit(s) adopted by the novice learners in decoding Manchu symbol blocks. If a single unit (uju or phoneme marker) was adopted, there should be no difference in participants' decoding performance between "blocked" and "mixed" condition. That is because no switching cost (Ziegler & Goswami, 2005) was incurred by participants' continuing shift between phoneme marker and uju in decoding, which would reduce reading accuracy and speed. In other words, using single orthographic unit in decoding would not affect the participants' performance despite the different conditions of presenting items. On the other hand, if a mixture of uju (for decoding items with more consistent symbol-sound mapping in Block 1) and phoneme markers (for decoding items with less consistent symbol-sound mapping in Block 2) were used, the participants would be expected to perform less accurately and more slowly in the "mixed" condition, because switching between uju and phoneme marker during the process of decoding created difficulties for the learners.

Second, if there is significant difference in decoding performance between items in Block 1 (all phoneme markers in an item are in consistent one–to–one mapping with sound) and items Block 2 (the target phoneme marker in an item is in less consistent one–to–multiple mapping with sound), it would suggest that phoneme marker was used by the novice learners in decoding. As visual complexity of the items in the two blocks were controlled and the target phoneme marker of an item in Block 2 was in a more complex one–to–multiple mapping with sound, the adoption of phoneme marker to decode items in Block 2 would therefore be cognitively more demanding for the participants than would Block 1, thus resulting in poorer decoding performance of lower accuracy and speed.

5.2.5 Procedure of Test 2

In Test 2 on symbol–sound mapping complexity, sixteen uju symbol blocks were given to the participants for recognition. Participants were asked to name the items as quickly and accurately as possible. The items were presented in a Dell desktop computer with 21.5-inch monitor of 1920 × 1080 resolution. An item began with the presentation of a fixation cross (+) in the centre of a PowerPoint slide. Then, a Manchu symbol block was presented at fixation and remained there until the participant gave a response, after which the tester continued to present the fixation cross for the upcoming symbol block. Participants were seen individually by a tester in a quiet room in the university. The testing was audiorecorded.

5.3 Multiple Imputation of the Missing Data in Test 2

An initial analysis of the data collected from the Manchu symbol block (item $n = 32$) naming task in Test 2 showed that a large number of values of the participants' decoding performance were missing. To be specific, in Test 2 the participants ($n = 196$) gave 1224 correct responses (19.52%) and 3876 incorrect responses (61.79%), with 1172 items unattempted (18.69%). This result suggested that for the estimation of item decoding error rate, 18.69% of the data was missing. In addition, for reaction time estimate, a more substantial proportion of data (80.48%) was missing because both the unattempted and the incorrect responses were excluded from being measured for reaction time. To address this problem, simply deleting the incomplete cases would result in too few observations being left for the analysis. Thus, the missing data, if not handled appropriately, would produce inaccurate parameter estimates (i.e., Manchu item decoding accuracy and speed), leading to invalid inference of the population of interest (Sinharay et al., 2001). Although some *ad hoc* methods, such as mean substitution, hot deck imputation, listwise deletion, pairwise deletion, regression imputation, etc., were helpful in dealing with data missingness, they could produce biased estimates of the mean and the standard error in certain conditions (see Enders, 2010;

Graham, Cumsille, & Shevock, 2012; Schafer & Graham, 2002; Van Buuren, 2018; van Ginkel, Linting, Rippe, & van der Voort, 2020). The alternative method of multiple imputation is more recommended (Schafer & Graham, 2002), because it uses multiple values drawn from the posterior predictive distribution of the missing data to complete the dataset and yields an unbiased estimation by averaging these values (Rubin, 1987). Therefore, to address the issue of missing data in Test 2, the present study adopted multiple imputation which is considered as "the best general method to deal with incomplete data" (Van Buuren, 2018, p. 25). As the relevant theories of multiple imputation were previously introduced in Chapter 4 (see Sections 4.3.1 for missing data mechanisms, 4.3.2 for the assumption required by multiple imputation, and 4.3.3 for steps of conducting multiple imputation), the section below focused on discussing the ways in which the missing data in Test 2 met the assumptions of multiple imputation.

5.3.1 Meeting the Assumptions of Multiple Imputation

In the present study, missing data in Test 2 met the requirement of missing at random (MAR) mechanism. First, by definition, MAR data in a variable is dependent on some measured variable(s) but not the variable itself. According to the finding of Test 1 that the participants' decoding accuracy and speed is negatively associated with item visual complexity, it can be inferred that missing values in the dependent variable of decoding performance (measured by error rate and reaction time) was supposed to be related to the independent variable of an item's visual complexity. To be specific, the visually more complex items were more prone to have missing values. Furthermore, the strong association between the variables of decoding performance and visual complexity found by Test 1 also yielded sound multiple imputation results irrespective of the missing data mechanism (Rubin, 1996; Sinharay et al., 2001). Second, the missing data in decoding performance were not dependent on the variable itself. The missing values in error rate measurement were due to

the unattempted items, which is unrelated to error rate itself. Moreover, the missingness of reaction time data was derived from (a) the incorrect responses and (b) the unattempted items. In either case, the missingness could occur regardless of how long the participants spent responding to the items. Therefore, for the above reasons the missing data in Test 2 were considered MAR.

5.3.2 Results of Multiple Imputation of Missing Data in Test 2

Following the reporting guidelines offered by Van Buuren (2018), this section provides a detailed description of multiple imputation used to deal with the missing values in Test 2. With respects to the report of software, number of imputed datasets, and pooling of multiple imputation of Test 2, they can be referred back to in Section 4.3.4.

Reasons for Missing data. The values were missing primarily because of the limited Manchu proficiency of the participants, who had received only ten sessions of 45-minute instruction on Manchu orthography knowledge before taking the test of uju symbol blocks recognition. The missingness was also related to the outcome measurement. For the estimation of item decoding error rate, the 1172 unattempted items (18.69%) were excluded from being measured. Moreover, when the reaction time was estimated, both the incorrect responses (3876 cases, 61.79%) and the unattempted items (1172 cases, 18.69%) were excluded, thus resulting in a more substantial proportion (80.48%) of missing data.

Method. As previously discussed, the missing data in Test 2 have met the assumption of missing at random (MAR) mechanism, which made it possible to obtain the imputed values using the relationship among the variables from the observed data (Sinharay et al., 2001). In addition, since the missingness of the outcome variable was considered to be associated with the five dimensions (i.e., PX, PC, SF, CP, and DC) measuring symbol visual complexity, they were included as the independent variables to impute the missing values in

the outcome variables of decoding accuracy (correct response = 1, incorrect response = 0) and decoding speed (in milliseconds) for each case.

Prior to imputing the missing reaction time values in Test 2, the cases giving no correct responses (i.e., no available reaction time) in the "blocked" (n = 55) and mixed conditions (n = 70) were deleted from the original cases (n = 196), because multiple imputation requires minimally one value in the dependent variable (i.e., the reaction time) before the missing values can be imputed.

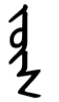
The missing data in the blocked and mixed conditions were imputed separately using the existing values collected from within their respective condition of presentation. That is because in multiple imputation the missing data were imputed by regression using values from the independent variables (i.e., in the present study the five dimensions of visual complexity—PX, PC, SF, CP and DC) and the existing value in the dependent variable (i.e., the error rate and reaction time). Thus, the missing data in the dependent variable in the "blocked" condition should be imputed based on the existing data collected from within this condition, and similarly, the missing data in the "mixed" condition should be predicted from the values obtained when the items were mixed before presentation.

Diagnostics. Following the imputation phase, diagnostics were performed to see whether the imputed values were plausible in a sense that they could have been plausibly measured had they not been missing (Van Buuren, 2018). The diagnostics resulted in the deletion of certain cases (n = 39 in blocked condition and n = 23 in mixed condition) which contained *identical* imputed values for reaction time, because it is practically implausible for a participant to spend the same amount of time reacting to 16 different items, and thus these imputes were considered less valid. Although the deletion reduced the cases to be left for analysis in blocked condition (n = 102) and mixed condition (n = 103), the decreased number of cases did not affect the power analysis, since minimally 94

cases was sufficient for a 1,500 population with a 90% confidence level (Yamane, 1967). The sets of parameter estimates of the mean decoding accuracy (m = 20) and the mean decoding speed (m = 40) of the Manchu items (n = 16) in Test 2 were pooled into a single result. Table 23 gives the results.

Table 23

Error Rate and Reaction Time of the Items (n = 16) Presented in Blocked and Mixed Conditions in Test 2

Blocks of items	Items		Blocked condition		Mixed condition	
	Traditional	Latin	Error rate	Reaction time, in ms	Error rate	Reaction time, in ms
Block 1 (Item *n* = 8)		*sū*	.622	6833.89	.579	6477.27
		*ce*	.678	6958.51	.650	5995.04
		*di*	.631	7927.00	.664	6460.28
		*nū*	.424	6281.21	.448	5753.18
		*ti*	.548	6777.68	.513	6485.83
		*cū*	.551	6615.06	.580	5584.59
		*jūs*	.865	7450.21	.847	6764.22
		*tib*	.748	7521.76	.710	6714.42

Blocks of items	Items		Blocked condition		Mixed condition	
	Traditional	Latin	Error rate	Reaction time, in ms	Error rate	Reaction time, in ms
Block 2 (Item *n* = 8)		*tu* [a]	.635	7659.15	.694	7220.75
		*we*	.878	7850.16	.881	7397.15
		*de* [a]	.807	8050.27	.773	7704.94
		*du*	.685	7282.10	.692	7929.36
		*te*	.801	7599.09	.826	7241.78
		*fū*	.751	7405.73	.731	6975.11
		*tus*	.856	8219.07	.782	7962.80
		*teb*	.915	8043.38	.861	7681.44

[a] Phoneme markers in one–to–two and one–to–three mapping with sound are underlined by single and double straight line, respectively.

5.4 Results for Test 2

In the previous section, multiple imputation was adopted to address the problem of missing values in the participants' decoding performance in Test 2, which ended up yielding estimates of decoding accuracy and decoding speed of the test items (n = 16). This section presents the results of data analysis of Test 2. Two paired samples *t*-tests were performed, with one comparing participants' decoding performance on the items presented in "blocked" and "mixed" conditions, and the other comparing their decoding performance on the items in Block 1 and Block 2. Table 24 provides error rate and reaction time of the items in Block 1 and Block 2, which were presented to participants in "blocked" and "mixed" conditions in Test 2.

5.4.1 Assumptions Check for Paired Samples t-Test

Before conducting paired samples *t*-test, some assumptions were checked to see whether the data met the relevant requirements. The first assumption is that the dependent variable in the *t*-test should be continuous. In Test 2, the dependent variable is the participants' decoding performance measured by error rate (i.e., the ratio of the incorrect responses to the total number of correct and incorrect responses given to an item) and reaction time, both of which produced data of continuous measurement. Secondly, paired samples *t*-test also requires normal distribution of the difference between the paired values. The Shapiro-Wilk test was performed because of its high power to test normality of data (Ghasemi & Zahediasl, 2012). The results showed that two of the *p* values (*p*s = .050 and .044) of Shapiro-Wilk test were lower than the critical value of .05, suggesting that the data were not normally distributed. To address this problem, the robust method of bootstrap (Efron & Tibshirani, 1993) was adopted in performing paired samples *t*-test to reduce the impact of bias incurred by nonnormality. Bootstrap uses samplings with replacement to produce *n* random samples where *n* is the number of data in the sample.

The resampling method of bootstrapping can model the distribution of the parent population (Henderson, 2005).

Table 24

Descriptive Data of the Error Rate (ER) Difference and the Reaction Time (RT) Difference of the Data for Paired Samples t-Tests in Test 2

	Blocked condition – mixed condition		Block 1 – Block 2 (in blocked condition)		Block 1 – Block 2 (in mixed condition)	
	ER difference	RT difference, in ms	ER difference	RT difference, in ms	ER difference	RT difference, in ms
Mean	.014	508	–.158	–718	–.156	–1235
SD	.037	453	.102	276	.114	460
Skewness	–.195	–.354	.998	1.699	.777	–1.163
Kurtosis	–.812	2.844	–.385	3.048	1.407	2.002
p of Shapiro-Wilk test	.050	.805	.071	.044	.585	.230

5.4.2 Participants' Decoding Performance between Blocked and Mixed Conditions

Paired samples *t*-tests were conducted to compare the participants' performance (measured by error rate and reaction time) in decoding uju symbol blocks ($n = 6$) presented in "blocked" and "mixed" conditions in Test 2. Table 25 presents the results. It was found that there was no significant difference between the two conditions with respects to error rate ($MD = .01$, $t(15) = 1.107$, $p = .280$). Yet, a significant difference was found regarding reaction time ($MD = 508$, $t(15) = 4.488$, $p = .002$) between the two conditions. The results suggested that participants did not perform differently in terms of how accurately they decoded the uju presented in the "blocked" and "mixed" conditions, but they performed more rapidly on the uju presented in the "mixed" condition.

Table 25

Results of Paired Samples t-Test of Participants' Performance in Decoding Items (n = 32) Presented in Blocked and Mixed Conditions in Test 2

	Blocked (n = 16)		Mixed (n = 16)		t(15)	p [a]	Cohen's d
	M	*SD*	*M*	*SD*			
Error rate	.712	.139	.702	.127	1.107	.280	.075
Reaction time, in ms	7405	571	6897	747	4.488	.002	.764

Note. In the "blocked" condition, items (n = 16) were presented to the participants sequentially by blocks (Block 1 then Block 2). After that, in the "mixed" condition, items in Block 1 and Block 2 were mixed up before presented to the participants.

[a] The p value was based on a 2000-time bootstrap sampling.

5.4.3 Participants' Decoding Performance between Block 1 and Block 2

Paired samples *t*-tests were also conducted to compare participants' performance in decoding the uju in Block 1 (all phoneme markers of an uju in consistent one–to–one mapping with sound) and Block 2 (the target phoneme marker of an uju in less consistent one–to–multiple mapping with sound). Table 26 presents the results. It was found that there was a significant difference in decoding error rate between Block 1 and Block 2 when the uju were presented in "blocked" condition (MD = –.158, t(7) = –4.381, p = .035) and when they were presented in "mixed" condition (MD = –.156, t(7) = –3.877, p = .008). In addition, it was found that there was a significant difference in decoding reaction time between the two blocks when the uju were presented in "blocked" condition (MD = –718, t(7) = –7.362, p = .008) and in "mixed" condition (MD = –1235, t(7) = –7.590, p < .001).

These results suggested that the symbol–sound mapping complexity in Manchu had inhibitory effects on participants' decoding performance. To be specific, more errors were made and more time was spent when items containing phoneme marker in inconsistent mapping with sound were decoded.

Table 26

Results of Paired Samples t-Test of Participants' Performance in Decoding Items (n = 16) in Blocks 1 and 2 in Test 2

	Conditions of item presentation	Block 1 (n = 8)		Block 2 (n = 8)		t(7)	p [a]	Cohen's d
		M	*SD*	*M*	*SD*			
Error rate	Blocked	.633	.135	.791	.096	–4.381	.035	1.349
	Mixed	.624	.123	.780	.072	–3.877	.008	1.548
Reaction time, in ms	Blocked	7046	543	7764	332	–7.362	.008	1.595
	Mixed	6279	443	7514	359	–7.590	< .001	3.063

Note. The comprising phoneme markers in the uju (n = 8) in Block 1 are all in one–to–one mapping with sound. However, the uju (n = 8) in Block 2 contain a target phoneme marker that is in one–to–two or one–to–three mapping with sound, with the remaining phoneme marker(s) in one–to–one mapping with sound.

[a] The p values were based on a 2000-time bootstrap sampling.

5.4.4 The Reliable Change Index Controlling for Practice Effects

The above result that the participants performed more rapidly in decoding items presented in "mixed" condition than in "blocked" condition (*MD* = 508, in ms) may be due to practice effects, which could be incurred by the repeated presentation of the items in

"blocked" condition before "mixed" condition. Thus, a *post hoc* method of Reliable Change Index which controls for practice effects (RCI_{PE}, Chelune, Naugle, Lüders, Sedlak, & Awad, 1993) was adopted to test whether the item reaction time change across conditions was true or due to measurement error. RCI_{PE} uses the discrepancy score ($T_2 - T_1$) subtracting the mean practice effects from control group ($M_2 - M_1$) as the numerator, and the standard error of the differences scores (SED) as the denominator (i.e., $(T_2 - T_1) - (M_2 - M_1)$/SED). The formula for calculating RCI_{PE} is

$$RCI_{PE} = \frac{(T_2 - T_1) - (M_2 - M_1)}{\sqrt{2S_1{}^2(1 - r_{12})}}$$

where T_1 = score at Time 1;

T_2 = score at Time 2;

M_1 = control group mean at Time 1;

M_2 = control group mean at Time 2

S_1 = standard deviation at Time 1;

r_{12} = correlation between Time 1 and 2 scores.

The resulting RCI_{PE} score, which is regarded as standardized score, is then compared with the cutoff point of ± 1.96. An RCI_{PE} score larger than +1.96 or lower than –1.96 (occurring less than 5% of cases) indicates a statistically significant change with the practice effects controlled (Collie, Maruff, Darby, & McStephen, 2003).

An issue of calculating RCI_{PE} for the present study is that the mean practice effects from the control group (i.e., $M_2 - M_1$) is unknown because no control group was created at the time of testing. To address this problem, this study adopted the method of using the means of

the test and retest reaction time scores (i.e., M_1' and M_2') as substitution to estimate the mean practice effects (e.g., Parsons, Notebaert, Shields, & Guskiewicz, 2009). The justification is as follows. Firstly, the control group test mean (M_1) is assumed to be very close to the test mean (M_1'), because the participants would probably not perform significantly faster or slower if there had been a control group at Time 1, when they could be influenced by practice effects. Moreover, the control group retest mean (M_2) is assumed to be larger than the retest mean (M_2'). That is because the participants in the control group, if there had been one at Time 2, would have received no practice, so that they would probably perform more slowly than the treatment group. Consequently, the mean practice effects estimated from the treatment group (e.g., M_2' – M_1') would be smaller than that from the control group (e.g., M_2 – M_1), resulting in a larger RCI_{PE} score of (T_2 – T_1) – (M_2' – M_1')/SED than to be calculated by the true mean score of (M_2 – M_1). In fact, a larger RCI_{PE} score would increase the possibility of being greater than the cutoff value of ± 1.96 (i.e., a 95% confidence interval), which indicates a statistically true change. Therefore, using (M_2' – M_1') to substitute (M_2 – M_1) would provide a more conservative estimation of the RCI_{PE} score. After substituting (M_2 – M_1) with (M_2' – M_1'), the formula for calculating RCI_{PE} used in this study is

$$RCI_{PE} = \frac{(T_2 - T_1) - (M_2' - M_1')}{\sqrt{2S_1{}^2(1 - r_{12})}}$$

where M_1' = group mean at Time 1, and M_2' = group mean at Time 2.

The Reliable Change Index with practice effects corrected (RCI_{PE}) was calculated for the participants' test-retest reaction time of the items (n = 32) presented in "blocked" and "mixed" conditions in Test 2. Table 27 shows the results that all but one score (RCI_{PE} = 2.54) is larger than the cutoff value of ± 1.96 (i.e., a 95% confidence interval). This result

suggests that after controlling the practice effects, there is no statistically significant difference in the participants' uju decoding speed between the "blocked" and "mixed" conditions in Test 2.

Table 27

Results of the Reliable Change Index Controlling for Practice Effects (RCI_{PE}) of the Reaction Time of Item Test-Retest in Test 2

Blocks	Items	Time 1 RT [a], in ms	Time 2 RT, in ms	RCI_{PE}
Block 1	*sū*	6833.89	6477.27	0.99
(*n* = 8)	*ce*	6958.51	5995.04	–0.47
	di	7927	6460.28	–1.69
	nū	6281.21	5753.18	0.57
	ti	6777.68	6485.83	1.14
	cū	6615.06	5584.59	–0.64
	jūs	7450.21	6764.22	0.19
	tib	7521.76	6714.42	–0.10
Block 2	*tu*	7659.15	7220.75	–0.54
(*n* = 8)	*we*	7850.16	7397.15	–0.58
	de	8050.27	7704.94	–0.27
	du	7282.1	7929.36	2.54
	te	7599.09	7241.78	–0.31
	fū	7405.73	6975.11	–0.51
	tus	8219.07	7962.8	–0.02
	teb	8043.38	7681.44	–0.32

[a] RT = reaction time. At Time 1, the items in Block 1 and Block 2 were presented to participants in "blocked" condition for a recognition test. At Time 2, these items were presented again but in "mixed" condition for a retest.

5.5 Chapter Summary

This chapter presented the results of Test 2, which examined the association between the symbol–sound mapping complexity of uju symbol blocks and the participants' decoding performance. Specifically, it intended to address the following research question (RQ):

> **RQ 2-1**: Do the learners perform differently when they decoded Manchu symbol blocks with phoneme markers in one–to–one and one–to–multiple mapping with sound?

To answer this question, the participants were given Test 2 in which they were asked to name Manchu uju in Block 1 (n = 8) and Block 2 (n = 8) presented in "blocked" and "mixed" conditions. First, it was found that there was no difference between "blocked" and "mixed" condition regarding how accurately participants decoded uju, but their performance was significantly different with respect to how fast they decoded. This difference in decoding speed between the two conditions may be due to practice effects. After the practice effects were statistically controlled, it was found that there was no significant difference in decoding speed between "blocked" and "mixed" condition. This result suggested that the participants used a single orthographic unit (i.e., either Manchu phoneme marker or uju) rather than a mixture of both in decoding. That is because otherwise the cognitive demand incurred by the constant switch between orthographic units in decoding would have reduced participants' accuracy and speed in the "mixed" condition, which was in fact not observed. Furthermore, it was also found that the participants performed more accurately and rapidly when they decoded items in Block 1 (all phoneme markers of an uju in consistent one–to–one mapping with sound) than did in Block 2 (the target phoneme marker of an uju in less consistent one–to–multiple mapping with sound). This result further suggested that the orthographic unit of

phoneme marker was adopted in decoding, because the phoneme marker-sound inconsistency had inhibitory effects on decoding accuracy and speed.

Chapter 6 Results of Classroom Observation of Manchu Orthography Instruction

6.1 Chapter Overview

This chapter presents the results of classroom observation conducted on the Manchu orthography instruction offered in the participating university. The chapter first introduces the purpose and focus of observation. Following that is the report on data analysis methods, including the synthesis matrix for organizing the observation data, the codes and subcodes for coding data, the inter-coder check on coding reliability, and symbol frequency count. Finally, findings of the observation were provided with respect to the instruction on writing and reading Manchu uju symbol blocks.

6.2 Introduction to Classroom Observation

6.2.1 Purpose of Classroom Observation

The previous Chapter 5 has reported the results of Test 2 in which the participants were asked to name two blocks (Block 1 and Block 2) of Manchu uju. In Block 1, all phoneme markers in an uju were in one–to–one mapping with sound (e.g., <SŪ> and <TIB>), but in Block 2, the target phoneme marker in an uju was in one–to–multiple mapping with sound (e.g., <TU> and <TEB>).[13] The results of Test 2 showed that the inconsistency in phoneme marker–sound mapping had inhibitory effects on participants' decoding performance, in comparison to the phoneme markers that are in consistent mapping with sound. As visual complexity and frequency of the uju in Block 1 and Block 2 were matched in pairs, the inhibitory effects could be attributed to the phoneme marker in inconsistent mapping with sound in Block 2. This result suggested that phoneme marker may be adopted as the basic orthographic unit by novice Manchu learners in uju decoding. To explain this finding of Test 2, the present study further explored from a qualitative

[13] The phoneme marker < ᡠ > represents /u/ after <T> and <D>, but /ɔ/ after other consonant phoneme markers. Similarly, the phoneme marker < ᡝ > in <TEB> is mapped to /ə/, but it can also be mapped onto /ɑ/ or /n/ depending on the preceding phoneme marker.

perspective the association between the orthographic unit choice of the novice learners and the Manchu orthography knowledge instruction offered to them. In this regard, the following research question (RQ) was asked.

> **RQ 2-2**: In what way are the results of the learners' decoding accuracy and speed on Manchu symbol blocks with different phoneme marker–sound mappings related to the instruction on Manchu orthography?

The answer to this question will be complementary to the understanding of the quantitative results in Chapter 5. It will provide additional qualitative evidence gleaned from the observation of ten sessions of Manchu orthography instruction, which was an integral part of the *Manchu Intangible Cultural Heritage* course offered in the university.

6.2.2 Focus of Classroom Observation

Non-participant observation was conducted on Manchu orthography instruction offered in the participating university. Adapted from the construct of component skills of literacy (Nag et al., 2016), the focus of observation was on the more specific *component knowledge of orthography* in Manchu instruction. Moreover, the observation also concentrated on classroom literacy practices (Street, 2000), which included the teaching and learning activities and interactions between teacher and students in Manchu class.

6.3 Data Analysis

6.3.1 Data Synthesis Matrix

The qualitative data of Manchu orthography instruction was organised in a synthesis matrix (see Table 9) derived from the meta-ethnography framework developed by Nag, Snowling, and Asfaha (2016) for configuring observation data. Their synthesis framework functions by mapping the observation data of classroom literacy practices onto multiple

dimensions under the construct of "component skills of literacy" (e.g., letter and symbol knowledge, manipulation knowledge, reading knowledge, etc.). Each dimension of the components skills of literacy was further mapped onto the observation about "teaching practices" reflected by the four dimensions of teacher talk, student talk, literacy practices, and other features. In application of Nag and colleague's (2016) synthesis matrix, the present study adapted the construct of component skills of literacy to the more specific focus on *component knowledge of orthography* to accommodate to observations made specifically on Manchu orthographic knowledge instruction. Regarding the component knowledge of orthography, previous studies have identified important dimensions on strokes (e.g., Jin & Cortazzi, 1998; Shen, 2005), basic orthographic unit (e.g., Nag & Narayanan, 2019), and symbol–sound mapping (e.g., Nag & Snowling, 2012). Therefore, the three dimensions of (a) writing knowledge, (b) orthographic unit knowledge, and (c) symbol–sound mapping knowledge were used as codes for organising data under the construct of component knowledge of orthography. Additionally, an "other knowledge" dimension captured the descriptions of Manchu orthography instruction that were not covered elsewhere in the matrix.

The above four dimensions captured the Manchu orthography knowledge instructed, which were further mapped onto the classroom activities and interactions. The activities and interactions between teacher and students in Manchu class was conceptualised as literacy practices (Street, 2000). The present study regards the literacy practices of the teacher and students as a social endeavour occurring in the context of educational institutions to prevent the critically endangered Manchu language from extinction. Their literacy practices, drawing on the dimensions of teacher talk, student talk, and literacy activities in the "teaching practices" in Nag and colleague's (2016) synthesis framework, were mapped onto the three dimensions of (a) explanation of relevant orthographic knowledge by teacher, (b) exercises

on reading and writing taken by students, and (c) evaluation by assessing whether students have grasped the relevant orthographic knowledge by teacher–students interaction.

The analysis of the synthesis matrix was made through thematic exaction of the "what" and "how" of the literacy activities captured in the observation (Nag et al., 2016). "What" refers to the Manchu orthography knowledge instructed (e.g., symbol–to–sound mapping knowledge) and "how" refers to the methods of instruction (e.g., copywriting). The analysis revealed the patterns of what orthography knowledge of Manchu was instructed and in what method it was instructed.

6.3.2 Codes and Subcodes

Using the synthesis matrix, the observation data on Manchu orthographic knowledge instruction were initially coded as “writing knowledge”, “orthographic unit knowledge”, “mapping knowledge”, and “other knowledge”. Then, each dimension was further coded as “explanation”, “exercise”, and “evaluation”. As the iterative process of coding and recoding continued in the second and third cycles, the codes became more refined and as a result some subcodes emerged (Saldaña, 2021). Table 10 gives codes, subcodes, their definitions, and examples from the observation data.

6.3.3 Inter-coder Reliability Check on Coding

To ensure the observation data were reliably coded, a second coder was given a sub-data set comprising 10% of the data for independent coding (O’Connor & Joffe, 2020). The sub-data set was created by randomly extracting a clip around 5 minutes from each of the ten video-recorded sessions, with a total length of about 50 minutes. After the two coders independently coded the clips, they compared their coding results. For any disagreements, discussions were carried out until consensus was reached. Table 28 provides details.

Table 28

Results of the Inter-coder Reliability Check on Classroom Observation Data

Sessions	Duration	Agreed codes and data	Disagreed codes and data	Reason for disagreement	Solution
1	15:15-20:20	Demonstration–Stroke order: T [a] demonstrated how to write the independent and initial forms of <A> on blackboard. Copywriting–Uju: T asked Ss [b] to copywrite the independent and initial forms of <A>. Demonstration–Stroke transformation: T demonstrated by applying the rule of "rightward tail" = "short tooth" in writing A- from A on blackboard.	Demonstration–Stroke: T introduced the strokes of "short tooth", "rightward masculine tail", and "stem" during demonstration of stroke order.	Only this piece of data concerned "Stroke" and the rest data coded as "Stroke" were in fact about stroke combination.	This piece of data was coded as Other knowledge–Linguistics, and a more accurate code "Stroke combination" was developed for coding the rest data previously coded as "Stroke".
2	24:23-29:07	Copywriting–Uju: T asked Ss to copywrite the independent, initial, medial, and final forms of <NA>.	Demonstration–troke order: T demonstrated the stroke sequence of the four forms of <NA> by writing them on blackboard.	Coder #2 coded the data as "Demonstration–Uju".	"Demonstration–Stroke order" was kept because the teaching purpose here was to demonstrate the stroke order in writing, rather than to demonstrate the symbol blocks of uju.

Sessions	Duration	Agreed codes and data	Disagreed codes and data	Reason for disagreement	Solution
			Transliteration–Uju: T wrote *na, ne, ni, no, nu,* and *nū* on blackboard before writing down the four forms of <NA>, <NE>, <NI>, <NO>, <NU> and <NŪ>.	Coder #2 thought it is more appropriate to use the codes of "Transcription" or "Sound-to-symbol mapping" instead of "Transliteration".	"Transliteration–Uju" was kept because the observation aligned with the definition of transliteration (i.e., converting a writing symbol from one *script* to another *script*).
3	12:55-18:25	Presentation–Stroke transformation: T presented the "rightward tail" in <KA> was transformed to "short tooth" in writing <KA->. Demonstration–Stroke transformation: T transformed "rightward tail" in <KA> to a "short tooth" in writing <KA->. Demonstration–Stroke order: T demonstrated how to write the four variants of <KA> on blackboard. Manipulation–Uju: <KA> was segmented by boxes into <K_1-> and <-A> in slides. Manipulation–Phoneme markers: <K_1-> was blended into <-A> to form <KA> on blackboard.	Coding was missed for the observation data during 16:55-17:10.	Coder #2 noted the piece of observation during 16:55-17:10 was not coded.	This piece of observation was coded "Evaluation (with Ss' response)–Stroke transformation".

Sessions	Duration	Agreed codes and data	Disagreed codes and data	Reason for disagreement	Solution
		Copywrite–Uju: T asked Ss to copywrite the independent, initial, medial, and final forms of <KA>.			
4	05:40-10:42	Demonstration–Stroke order: T demonstrated how to write the four variants of <LA> on blackboard.	Transliteration–Uju: T wrote *la* on blackboard before writing the four forms of <LA>.	Coder #2 thought it is more appropriate to use the codes of "Transcription" instead of "Transliteration".	"Transliteration–Uju" was kept because the observation aligned with the definition of transliteration (i.e., converting a writing symbol from one *script* to another *script*).
		Copywrite–Uju: T asked Ss to copywrite the independent, initial, medial, and final forms of <LA>.			
5	23:50-28:55	Demonstration–Stroke transformation: T transformed "rightward tail" in <TA> to a "short tooth" in writing <TA->.	No disagreement		
		Demonstration–Stroke order: T demonstrated how to write the four variants of <TA> on blackboard.			
		Manipulation–Uju: <TA> was segmented by boxes into <T_1-> and <-A> in slides.			

Sessions	Duration	Agreed codes and data	Disagreed codes and data	Reason for disagreement	Solution
		Manipulation–Phoneme markers:<-A> was blended into <T_1-> to form <TA> on blackboard. Transliteration–Uju: *ta*, *te*, *ti*, *to*, and *tu* were written before <TA>, <TE>, <TI>, <TO>, and <TU> Copywrite–Uju: T asked Ss to copywrite the independent, initial, medial, and final forms of <TA>.			
6	15:40-21:00	Presentation–Uju: T presented the four positional variants of <BA>. Presentation–Phoneme markers: T presented <B->, <-B->, and <-B>. Presentation–Symbol-to-sound mapping: T presented the leftward tail in <BA> was only decorative, and it was not mapped onto any sound.	Presentation–Stroke: During 15:40-17:10, T presented the "short tooth" in <BA> was not a part of <B->. Instead, it was <-A> in <BA>. <B-> and <-A> were written as ligature.	Coder #2 thought it is less appropriate to use the code of "Stroke" for this piece of observation data because it concerned the relations among strokes.	This piece of data was coded using the new subcode "Stroke combination" as it is a ligature.

Sessions	Duration	Agreed codes and data	Disagreed codes and data	Reason for disagreement	Solution
		Demonstration–Stroke transformation: T transformed "leftward tail" in <BA> to a "stem" to write <BA->. Demonstration–Stroke order: T demonstrated how to write the four variants of <BA> on blackboard. Manipulation–Phoneme markers: <B-> and <-A> were blended into <BA>, which was then segmented into <B-> and <-A> with a leftward tail. Transliteration–Uju: *ba, be, bi, bo, bu,* and *bū* were written before <BA>, <BE>, <BI>, <BO>, <BU>, and <BŪ>. Transliteration–Phoneme marker: *b* was written for <B-> and *a* for <-A>. Copywrite–Uju: T asked Ss to copywrite the independent, initial, medial, and final forms of <BA>.			

Sessions	Duration	Agreed codes and data	Disagreed codes and data	Reason for disagreement	Solution
7	08:50-13:52	Presentation–Stroke combination: In <KE>, <K_2-> and <-E> is written as ligature. Demonstration–Stroke transformation: T applied the rule "leftward tail" = "stem" to transform <KE> into <KE->. Demonstration–Stroke order: T demonstrated how to write the four variants of <KE> on blackboard. Manipulation–Uju: <KE> was segmented into <K_2->, <-E>, and a leftward tail. Manipulation–Phoneme markers: <K_2->, <-E>, and a leftward tail were blended into <KE>. Transliteration–Uju: *ke*, *ki*, and *ku* were written for the four variants of <KE>, <KI>, and <KU>.	Demonstration–Stroke: the <-E> in <KE> was written without a "dot".	Coder #2 thought it is less appropriate to use the code of "Stroke" for this piece of observation data because it was not about how a stroke was written.	This piece of data was coded as "Demonstration–Phoneme marker". The <-E> without a "dot" in KE was regarded as a special case of the phoneme marker <-E> which typically carries a "dot".

Sessions	Duration	Agreed codes and data	Disagreed codes and data	Reason for disagreement	Solution
		Transliteration–Phoneme markers: *k* and *e* were written for <K_2-> and <-E> respectively. Copywrite–Uju: T asked Ss to copywrite the independent, initial, medial, and final forms of <KE>.			
8	28:35-33:30	Manipulation–Uju: ANGGA was segmented by T into <A->, <-NG->, <-GA>, ENGGE into <E->, <-NG->, <-GE>, MANGGI into <MA->, <-NG->, <-GI>, ONGGO into <O->, <-NG->,< -GO>, UNGGU into <U->, <-NG->, <-GU>, and TANGGŪ into <TA->, <-NG->, <-GŪ>. Transliteration–Uju: T used *ga* for <-GA>, *ge* for <-GE>, *gi* for <-GI>, *go* for <-GO>, *gu* for <-GU>, *gū* for <-GŪ>, *ma* for <MA->, and *ta* for <TA->. Transliteration–Phoneme markers: T used *a* for <A->, *e* for <E->, *o* for <O->, and *u* for <U->.	No disagreement		

Sessions	Duration	Agreed codes and data	Disagreed codes and data	Reason for disagreement	Solution
		Copywriting–Uju: T asked Ss to copywrite <A->, <-NG->, <-GA>, <E->, <-NG->, <-GE>, <MA->, <-NG->, <-GI>, <O->, <-NG->, <-GO>, <U->, <-NG->, <-GU>; <TA->, <-NG->, <-GŪ> with transliteration. Evaluation (with Ss' response)–phoneme marker: T: (pointing to <A-> in ANGGA on blackboard) 这个是啥？这个肯定认识。 [What is it? I'm sure you know it.] Ss: (some Ss) /*a*/. Evaluation (without Ss' response)–Uju: T: (pointing to <-GE> in ENGGE on blackboard) 这个是什么？刚学的今儿个。 [What is it? We just learned it today.] Ss: (only a couple of Ss murmuring without giving correct response) T: (with an emphasizing tone) 这个是*ge*。[It's *ge*.]			

Sessions	Duration	Agreed codes and data	Disagreed codes and data	Reason for disagreement	Solution
9	36:25-41:00	Demonstration–Symbol-to-sound mapping: *-iy-* was mapped onto /i/, *-iyei-*, *-iye-* onto *pinyi* /ye/, i.e., [jɣ], *siyei* onto "谢" [ɕiɛ]. The "*w*" in *ow*, *uw*, *uw*, *w* is silent, *iui* onto *pinyi* /y/, i.e., [y], *niui* onto *pinyi* /nü/, i.e., [ny]. Choral reading–Sound-to-symbol mapping: T-led reading of *biya*, *boa*, *huesi*, and *gūa*.	No disagreement		
10	14:30-19:25	Demonstration–Symbol-to-sound mapping: *iuwei*, *-iuwe* was mapped onto *pinyin* /üe/, i.e., [yɛ], and *siuwei* onto [syɛ].	No disagreement		

[a] T = Teacher

[b] Ss = Students

6.3.4 Symbol Frequency Count

Frequency of symbols was found to influence reading in a variety of orthographies (e.g., Spanish: Carreiras, Alvarez, & de Vega, 1993; German: Conrad & Jacobs, 2004; Hebrew: Frost et al., 1987; Chinese: Liversedge et al., 2014; Kannada: Nag, 2021; Nag et al., 2014; Korean: Simpson & Kang, 2004; Japanese kanji: Tamaoka & Kiyama, 2013). Thus, the frequency of uju and Manchu phoneme markers was measured and used as a control variable in correlation test. A mini corpus was built to include all the uju and Manchu phoneme markers involved in the literacy practices during the ten sessions Manchu instruction and their frequency was hand-counted. The criterion of counting frequency was whether the learners paid incessant "attention" to a given orthographic unit. If the learners continuously attended to an orthographic unit without any interruption, the frequency of the unit was counted 1. For example, in a teacher-led choral reading of the uju <A>, although the chant went on for multiple times, the frequency of this unit was counted as 1 because the learners' attention has been focused on <A> during this period. However, if the teacher proceeded to teach another orthographic unit after completing the instruction on an orthographic unit, and later he went back to the penultimate unit, the frequency of this unit was counted as 2. For instance, the teacher first instructed the orthographic unit <E> and then proceeded to <I>, after which he guided the students' attention back to <E>. In this scenario, the frequency of <E> was counted 2 because students paid attention to <E> twice. The frequency list is provided in Appendix D.

6.4 Findings

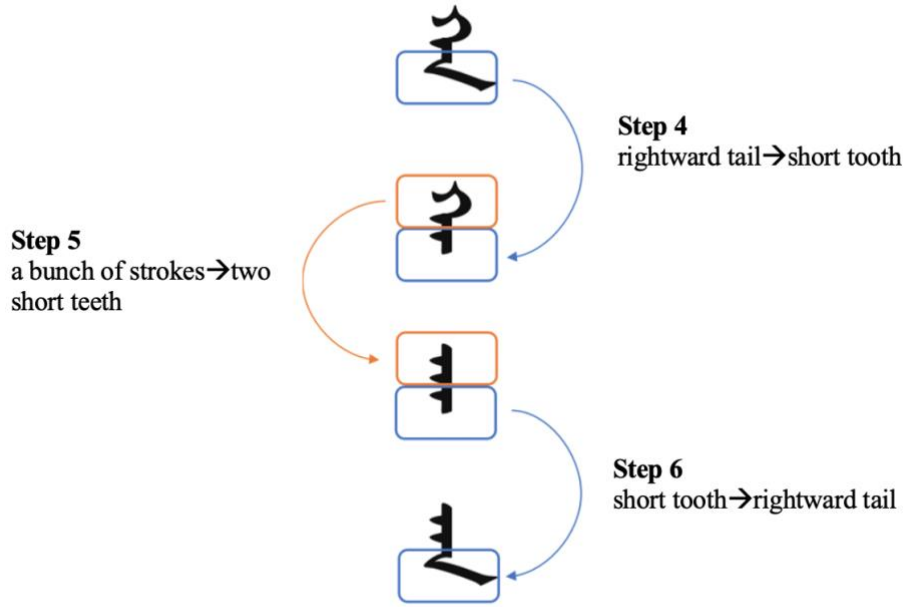
This section presents the findings obtained from observing ten sessions of Manchu orthography instruction in the participating university. Two findings with respect to writing and reading instruction on uju symbol blocks were identified from the thematic extraction of the observation data.

6.4.1 Finding 1: Instruction on the Writing of Uju Symbol Blocks

As prescribed in the curriculum, one objective of the Manchu course is to instruct the writing of the four positional variants (i.e., the isolation, initial, medial, and final forms) of an uju. Classroom observation revealed that a six-step instructional procedure was taken in uju writing instruction. In Step 1, the teacher presented in slides the four positional variants of a target uju (e.g., <KA>, <KA->, <-KA>, and <-KA>, see Figure 12), accompanied by their Latin transliteration (*ka*). In Step 2, the teacher demonstrated on blackboard the stroke order of writing the isolation form (e.g., <KA>). Then, in Step 3, the strokes constituting the isolation uju were segmented by a box. For bisyllabic uju, strokes segmentation was carried out based on phoneme markers (e.g., <KA> was segmented into a bunch of strokes constituting <K-> and a rightward "tail" of <-A>). Next, in Step 4, stroke transformation was applied to the final-position stroke in the isolation form to produce the initial form (e.g., based on the transformation rule, the rightward "tail" in <KA> was transformed to a short "tooth" to produce <KA->). Similarly, Step 5 produced the medial form by transforming the relevant stroke in the initial form (e.g., a bunch of strokes constituting <K-> were transformed into two short "teeth" of <-K->). Finally, Step 6 yielded the final form of an uju by applying again the stroke transformation to the final-position stroke in the medial form (e.g., the short "tooth" in <-KA-> was transformed to a rightward "tail" to produce <-KA>). Figure 12 describes the stroke transformation applied in Steps 4-6 between the isolation, initial, medial and final forms of uju <KA>.

Figure 12

Stroke Transformation Applied on Phoneme Marker in Writing the Isolation, Initial, Medial, and Final Forms of <KA>



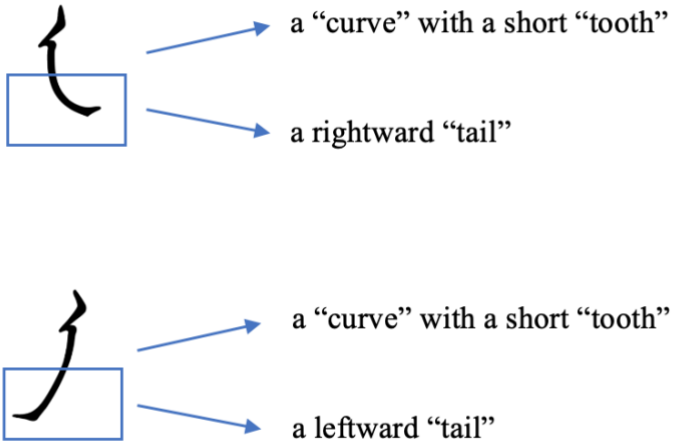
Note. The four uju symbol blocks from top to bottom are the isolation, initial, medial, and final forms of <KA>.

According to the classroom observation, most of the uju instructed in class were CV comprising two phoneme markers and thus this six-step instructional procedure was applied. However, for the six syllabically V uju (<A>, <E>, <I>, <O>, <U>, and <Ū>), instruction was given on how to write the isolation and initial forms; thus, only the first four steps in the procedure were taken. Moreover, unlike CV uju which were segmented into the comprising phoneme markers in Step 3, uju in the syllabic structure of V was segmented into a "curve" with a short "tooth" plus the remaining strokes (e.g., <A> into a "curve" with a short "tooth" plus a rightward "tail", and <E> into a "curve" with a short "tooth" plus a leftward "tail". See

Figure 13). Vignette 1 provides an example of applying stroke transformation into the writing instruction on the four forms of <KA> uju.

Figure 13

Stroke Segmentation Applied to the Isolation Form of <A> and <E>



Vignette 1 [Session 3, 12:58-21:25]

The teacher presented in the slides the isolation, initial, medial, and final forms of <KA> with the Latin transliteration *ka*. Among them, the independent form <KA> was segmented by boxes into <K$_1$-> and <-A>, with the transliteration *k* and *a* written next to them. In the subsequent demonstration of the stroke order of the isolation <KA>, the teacher said,

Teacher: 第一笔从左上到右下，… 写一个点。第二笔，… 兜一个类似于半圆的结构。 … 继续兜然后写一个短牙。…然后顿笔下来。这是这个辅音字母、阳性的 K 的写法。如果我们写 KA，是不是在下面接一个向右的尾巴就行了啊？ [The first stroke is from upper left to lower right …

to write a “dot”. The second stroke… is like a “curve”. … Keep writing another “curve” followed by a “short tooth”. Then, write downward [a “stem”]. This is how the consonant letter masculine <K> is written. To write <KA>, shouldn’t we just connect a “rightward tail” from below?]

Teacher: (在<K>下面接写向右的尾巴) … 然后我们分析一下：我们要把独立型变成词头型，是不是只需要把最后一个笔画变型啊。向右的尾巴等于什么？[(After writing a “rightward tail” connected to K from below) … Let’s make an analysis. If we want to change the isolation form [of <KA>] to the initial form, shouldn’t we just need to transform the final stroke [of the “rightward tail”]? What is a “rightward tail” equal to?]

Students: (一些学生一起回答)一个短牙。[(some students in unison) A “short tooth”.]

Teacher: 非常好啊。我们写一下。[Very good. Let’s write it down.]

Teacher: (教师在黑板写出<KA->) … 然后要把词头变词中是不是得变前半截了啊？…我们来分析一下。这个字母是不是带了一个短牙了啊。然后它有一个翘起来的结构看到了吗？上边这个[翘起来的结构]相当于一个短牙，… 那么这个辅音字母等于几个短牙？[(After writing the initial <KA> on blackboard) … Now if the initial form is to be changed to the medial form, shouldn’t the first part [of <K>] be transformed? Let’s make an analysis. This letter [<K>] contains a “short tooth”, right? Can you see it also has a curved stroke? This [curved stroke] on the top is equal to a “short tooth”. … Then, how many “short teeth” does this consonant letter [<K>] have?]

Students: (有些回答)三个。[(Some students) Three.]

Students: (另一些大声强调回答)两个。[(Other students in a louder and emphasising tone) Two.]

Teacher: 非常好。它就相当于两个短牙。那这个元音字母<A>是一个短牙，所以在词中的<KA>是？[Very good. It is equal to two “short teeth”. Since the vowel letter <A>has one “short tooth”, the medial <KA> has [how many “short teeth”]?]

Students: (一齐回答)三个短牙。[(In unison) Three “short teeth”.]

Teacher: (教师在黑板写出<KA>的词中型)好。… 我们既然知道词中的<KA>是三个短牙，并且我们能分析出前两个短牙是阳性的辅音字母<K>，第三个短牙是元音字母 A，那么它的词尾咋写？[(After writing the medial <KA> on blackboard) Good. … Now that we knew the medial <KA> has three “short teeth”, and by analysis we know the first two “short teeth” is the masculine consonant letter <K> and the third “short tooth” is the vowel letter <A>, then how should its final form be written?]

Students: (无回答) [(silence)]

Teacher: 把第三个短牙变成什么？[What should the third “short tooth” [in the medial KA] be transformed to?]

Students: (几个学生)阳性字尾。[(Several students) “Masculine tail”.]

Teacher: 向右的阳性的字尾。那还剩几个短牙了？[“Rightward masculine tail”. So, how many “short teeth” were left [in the final KA]?]

Students: (一齐回答)两个。[(In unison) Two.]

Teacher: (在黑板写出 KA 的词尾型) 给大家两分钟时间抄写。[(After writing the final form of KA) Now you’ll have two minutes to copywrite them.]

Vignette 1 showed that stroke transformation (e.g., the rule that a “rightward tail” is equal to a short “tooth” in transformation) was applied to teaching how to write the initial and final forms of <KA>. Moreover, Vignette 1 also revealed that manipulation (i.e., segmenting and blending) on phoneme markers was adopted by the teacher in uju writing instruction. For example, when transforming the initial <KA-> to the medial <-KA->, the teacher asked, “How many ‘short teeth’ does this *consonant letter* [<-K->] have (emphasis added)?”, after which he continued, “Since the *vowel letter* <-A-> has one ‘short tooth’, the medial <-KA-> has [three ‘short teeth’] (emphasis added).” In other words, the consonant letter <-K-> (two “short teeth”) was blended into the vowel letter <-A-> (one “short tooth”) to produce the medial <-KA->. Similarly, in observing Session 2 of the Manchu course in which the trigraph <IUI> was instructed, the teacher asked, “<I>用什么型？<U>用什么型？… 然后第二个<I>用什么型？… <I>, <U>, <I>分别变成相应的形态，摞在一起，OK 了吧。” [What form should <I> take? What form should <U> take? … And what form should the second <I> take? … Transform <I>, <U>, and <I> into their respective [positional] form and stack them together. Then, it’s OK.] This observation suggested that the phoneme markers of <I->, <-U->, and <-I> were blended together to write the uju <IUI>.

Yet, unlike Vignette 1 in which segmentation of an uju was based on its comprising phoneme markers, Vignette 2 given below revealed that segmentation was made to the strokes in an uju (i.e., segmenting <A> into a “curve” with a “short tooth” and a “rightward masculine tail”, see Figure 13).

Vignette 2 [Session 1, 15:10- 24:16]

The teacher began the instruction by presenting in slides the independent and initial forms of the vowel uju <A>, <E>, <I>, <O>, <U>, <Ū> and <IUI> accompanied by their Latin transliteration and a choral reading of them. Then, he wrote on the

blackboard the Latin transliteration *a, e, i, o, u,* and *ū* and demonstrated the stroke order of the isolation <A>.

Teacher: (写出一个略带弧形的线段)这不是直的，有点弧吧。然后折返，… 这都是弧线吧。别停笔，从左到右，画一条水平的线。… 然后转折，从上到下，写一小段。这一段有多小？越小越好，竖直的。然后继续向下，… 弧度越来越大，… 顿一下笔，然后水平画一个小线段，略微抬起。[(After writing a slightly curved segment line) It's not straight but somewhat curved. Then turn around [the direction and keep writing], … [this stroke is] also a curve. Keep going. From left to right, draw a horizontal line. … Then bend [the writing direction], from top to bottom, write a short segment line. How short should it be? The shorter, the better. [It's] vertical. And keep writing downward, … the curve becomes increasingly bigger, … pause the pen, draw a short horizontal segment line, and lift [the final stroke] up a little bit.]

The teacher then introduced the three strokes comprising the isolation <A>—a "short tooth", a "stem", and a "rightward masculine tail". Afterwards, the students were asked to practise copywriting the isolation <A> before the teacher continued to teach the initial <A->.

Teacher: 下边我们要开始把它[isolation <A>]变型了。… 还记得最底下这个方框[里面的笔画]叫什么吗？[Now we're going to change its [the isolation <A>'s] form. … Still remember what the [stroke in the] box is called?]

Students: (一齐回答)阳性的字尾。[(In unison) "Masculine tail".]

Teacher: 好啦。阳性的字尾它是不是个笔画的名称？… 阳性的字尾这种笔画永远等于一个短牙。跟它表示哪一个字母没有关系。只要是向右的阳性的字尾， 永远等于一个短牙。我们要把<A>的独立型变成词头型。…独立型的尾巴都翘出去了没有办法连别人了，所以我们要把这个向右的尾巴变个型，让它能连别的字母。这个独立型原来有几个短牙？

[Alright. The “masculine tail” is a stroke name, right? … The stroke of a “masculine tail” is always equal to a “short tooth” [in transformation], regardless of what [Manchu] letter it represents. As long as it is a “rightward masculine tail”, it’s forever equal to a “short tooth”. We’re going to change the isolation form of <A> into its initial form. … Since the tail of the isolation [<A>] is curved up, it can’t be connected to other [letters from below]. So, we’ll have to change the form of the “rightward tail” so that it can be connected to other letters. How many “short teeth” does the isolation form [of <A>] have?]

Students: (一齐回答)一个。[(In unison) one.]

Teacher: 非常好。… 那我们照原样把它写下来。[Very good. … We just write it down as it is.]

Teacher: (教师在黑板写出 crest)向右的尾巴等于什么？[(After writing the stroke of “short tooth” with a “crest” on top). What is the “rightward tail” equal to?]

Students: (一齐回答)一个短牙。[(In unison) A “short tooth”.]

Teacher: 一个短牙。你再写一个短牙。[A “short tooth”. [So] you write another “short tooth”.]

Teacher: (教师在 “短牙”和“字冠”下连写短牙) 这就是<A>的词头型的来源。

[(After writing an additional “short tooth” connected to the “short tooth” with a “crest” from below) This is the way in which the initial form of <A> is derived.]

Finally, the students were asked to practise writing the initial <A->.

Vignette 2 showcased that Steps 1-4 in the procedure of uju writing instruction was taken to teach the isolation and initial forms of the <A> uju. The process involved Step 1 the presentation of the target uju (i.e., the isolation and initial <A>), Step 2 the demonstration of the stroke order of isolation <A> (i.e., “curves” followed by a “short tooth” and a “rightward masculine tail”), and Step 3 the manipulation of strokes (i.e., the somewhat implicit segmentation of isolation <A> into a “curve” with a “short tooth” and a “rightward masculine tail” when the teacher asking “Still remember what the [stroke in the] box is called?”). In addition, the observation revealed that in Step 4 the teacher applied stroke transformation (e.g., a “rightward tail” = a “short tooth”) to the isolation <A> to produce the initial <A->. Applying stroke transformation in teaching how to write the positional variants of an uju was a repeatedly observed strategy. The following pithy formula for stroke transformation was observed to be reviewed multiple times by students’ choral reading in Sessions 3, 5, 7 and 9 of the Manchu course.

阳性字尾等于一个短牙!

阴性字尾等于一个字干!

两个元音字母相连第二个元音字母只能是 i 或 u,

如果是 u，则 u 没有点!

A masculine “tail” is equal to a short “tooth”![14]

A feminine “tail” is equal to a “stem”!

If two vowel phoneme markers are adjacent, the second one can only be *i* or *u*,

If it is *u*, it is written without a “dot”! [My translation.]

The pithy formula, especially the first two lines, was observed to be extensively applied to vowel phoneme markers in writing instruction on CV uju, such as <NA> and <NE> (Session 2), <SA> (Session 3), <TS‘I> and <DZI> (Session 4), <TA> and <TE> (Session 5), <AT> (Session 6) and so on. This observation suggested that phoneme marker was adopted by the teacher as the basic orthographic unit in uju writing instruction. The choice of phoneme marker presupposed the teacher’s belief that Manchu is an alphabet writing system, and consequently, combining consonant and vowel phoneme markers (or “stack them together”) will be able to produce uju symbol blocks. This belief can be reflected by the observation at the very beginning of the class in Session 1, when the teacher made the claim that “满文是拼音文字, 它和你学的英文和汉语拼音没有差别, 都是 *a*, *b*, *c*, *d* 拼成的一个个单词” [Manchu is alphabetic writing, which is no difference from English or *Hanyu Pinyin* you’ve learnt. The words are written by combining [like in the *pinyin*] *a*, *b*, *c*, [and] *d*.] Then the teacher continued to introduce “…满文有二十多个字母” [Manchu has over twenty letters] and “满文有六个元音字母” [Manchu has six vowel letters]. The statements explicitly showed that the teacher regarded Manchu phoneme markers (or “letters”) as the basic orthographic units. In the subsequent sessions throughout the course, Manchu phoneme markers were repeatedly adopted by the teacher in uju writing instruction. For example, when teaching <AI> in Session 2, the teacher said, “有一个单词<AI>由两个字母组成, 是<A>和

[14] A masculine “tail” is a rightward “tail”, and a feminine “tail” is a leftward “tail”.

<I>, 对吧? 其中元音字母<A>要用什么型?…元音字母<I>用什么型?" [The word <AI> comprises two letters, <A> and <I>, right? What form should the vowel letter <A> take? ... What form should the vowel letter <I> take?] In this example, the Manchu phoneme markers (or in the teacher's term "letters") of <A> and <I> were regarded by the teacher as the basic orthographic units that constitute the uju <AI>.

In addition, taking phoneme markers as the basic unit in writing instruction can also be evidenced by the observation that the teacher presented a matrix listing 17 consonant phoneme markers in their isolation, initial, medial, and final forms, if any. According to the teacher, using them "能把一般辅音字母和元音字母相拼的所有形式给它写出来" [could in general write all the [four] forms [of an uju] by combining consonant phoneme markers with vowel phoneme markers.]

6.4.2 Finding 2: Instruction on the Reading of Uju Symbol Blocks

The present study found that the knowledge of sound–to–uju mapping and sound–to–phoneme marker mapping was instructed for the novice learners to read Manchu symbol blocks. The sound–to–uju mapping knowledge was taught in a global approach (Nag, 2017a, 2022) of uju choral reading. More importantly, the sound–to–phoneme marker mapping knowledge was instructed in an analytic approach (Nag, 2017a, 2022) of uju segmentation. In addition, it was found that the sound of phoneme markers was taught with the mediation of Latin letters. Moreover, the question-posing strategy was found to be frequently used by the teacher to assess students' decoding ability on uju.

Sound–to–Uju Mapping Knowledge. Regarding instruction on the sound mapped onto uju, the only literacy practices (Street, 2000) observed to engage students into actual "sounding out" an uju was the teacher-led choral reading. For example, in Session 1, the teacher asked students to read after him multiple times /ɑ/, /ə/, /i/, /ɔ/, /u/ and /ʊ/, while pointing to the isolation and initial forms of <A>, <E>, <I>, <O>, <U> and <Ū> and their

Latin transliteration *a*, *e*, *i*, *o*, *u*, and *ū* presented in slides [Session 1, 13:07]. Similarly, in Session 3, the teacher led students to repeatedly read /kɑ/, /kɔ/, /kʊ/, /gɑ/, /gɔ/, /gʊ/, /hɑ/, /hɔ/, and /hʊ/, with the slide presenting "*k* 阳 + *a*/*o*/*ū*", "*g* 阳 + *a*/*o*/*ū*" and "*h* 阳 + *a*/*o*/*ū*" [Session 3, 11:00].[15] The observation suggested a global approach (Nag, 2017a, 2022) of uju reading was applied to instruction, during which students were supposed to learn the sound mapped onto an uju (e.g., /a/ – <A>).

However, there was a potential issue with teaching sound–to–uju mapping knowledge in reading instruction. As the students were novice learners with no prior knowledge of Manchu orthography, they probably may not be able to memorize the sounds mapped onto the visually complex uju symbol blocks only through several rounds of teacher-led choral reading. A closer examination on the observation data revealed that the sound–to–phoneme marker mapping knowledge was also taught in uju reading instruction.

Sound–to–Phoneme Marker Mapping Knowledge. The finding that uju reading was made possible by applying sound–to–phoneme marker mapping knowledge seemed to some extent surprising at first. That was because no actual "sounding out" of phoneme markers was observed throughout the ten sessions of Manchu orthography instruction. The sound–to–phoneme marker mapping knowledge was instructed in an analytical approach (Nag, 2017a, 2022) of uju segmentation, which could be evidenced by the following observation.

Segmentation of Uju into Phoneme Markers. The classroom observation revealed that in uju symbol block reading instruction, the teacher analysed an uju by segmenting it into comprising phoneme markers. Vignettes 3 and 4 below provide examples.

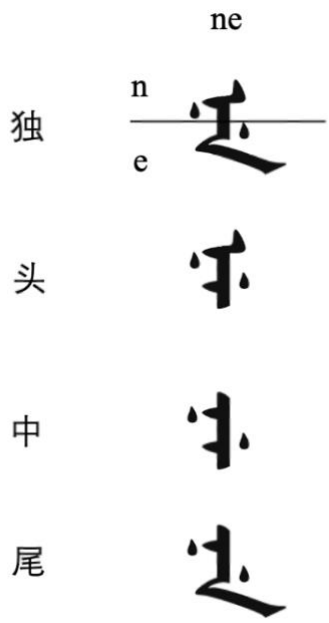
[15] The 阳 ("masculine") form of *k*, *g*, *h* are only followed by *a*, *o*, and *ū*. In contrast, the 阴 ("feminine") form of *k*, *g*, *h* are only followed by *e*, *i*, and *u*.

Vignette 3 (Session 2, 27:15)

When teaching the uju <NE>, the teacher wrote the isolation form <NE> on blackboard and drew a horizontal line to segment it into <N-> and <-E> (see Figure 14), saying, “我切一刀，上面是辅音字母<N>，下面是元音字母<E>” [I knife-cut it. The upper is the consonant letter <N>, and the lower is the vowel letter <E>.] He then wrote *n* and *e* next to the respective phoneme marker. After that, he continued to demonstrate the writing of the initial, medial, and final forms of <NE>.

Figure 14

Segmentation of <NE> into <N-> and <-E>



Note. The Chinese characters indicate the isolation, initial, medial, and final forms.

Vignette 4 (Session 4, 13:10)

After the teacher explained and demonstrated how to write the final position <-M>, he wrote <AM> on blackboard (see Figure 15).

Teacher: … 这个念啥？ [How do you read it?]

Students: (in unison) /am/.

Teacher: (segmenting <AM> into <A-> and <-M> with a horizontal line) 一定要会切，是吧。前边这两个“牙”是/a/，下边这个是/m/。[[You] must know how to cut it, right? The first two “short teeth” is /a/, and the one below is /m/.]

Figure 15

Segmentation of <AM> into <A-> and <-M>



Vignette 3 showed that in teaching the isolation <NE>, the teacher segmented (“knife-cut”) the uju and analysed that it consists of the two comprising phoneme markers <N-> and <-E>. Similarly, in Vignette 4, the teacher also “cut” <AM> into two segments and explained that it comprises the phoneme markers of <A-> and <-M>. The observation revealed that the teacher applied an analytical approach (Nag, 2017a, 2022) in reading instruction by segmenting an uju into its comprising phoneme markers. This finding suggested that phoneme marker was regarded as the basic orthographic in uju reading instruction.

However, it is worth noting that not all uju were segmented into its comprising phoneme markers by the teacher during instruction. For example, the observation revealed that segmentation was not applied to the uju of <CA>, <JA>, <YA>, <RA>, <FA>, <FI>, <WA>, <TS‘A>, <DZA>, <ŽA>, <TS‘I> and <DZI>. As most of these uju were in C+*a* structure which the learners were familiarized with during instruction, a strategic approach (Nag, 2017a, 2022) was probably applied by learners to segment the uju (e.g., <CA> into <C-> and <-A>) based on the instruction they have received.

Latin Letter-Mediated Instruction on Phoneme Marker Sound. It was observed that for the phoneme markers resulted from uju segmentation the teacher instructed the sound mapped onto them with the mediation of Latin letters. In Manchu orthography, the script of Latin letters typically functions to transliterate the traditionally scripted uju and words. For example, it was observed in Session 1 that before demonstrating how to write the uju <A>, <E>, <I>, <O>, <U>, and <Ū>, the teacher first wrote on blackboard the Latin letters *a, e, i, o, u* and *ū* and introduced that “我上面写的这些拉丁字母是拉丁满文, 在手机、电脑上大家用这个东西觉得比较方便” [These Latin letters written by me are Latinised Manchu. People feel it’s more convenient to use them in mobile phones and on computers.] (Session 1, 15:06). This statement indicated that Latin letters were used as a more accessible alternative to the traditional Manchu script in writing, suggesting that Latin letters function to transliterate the traditional Manchu. This function of Latin letters can be exemplified by Vignette 3, in which the teacher wrote *n* and *e* next to the phoneme markers <N-> and <-E>.

More importantly, in uju reading instruction, the Latin letters were used by the teacher as a mediator to phonologically notate the sound of the phoneme markers. Vignettes 5 and 6 offered evidence of the teacher’s Latin letters-mediated instruction on sound–to–phoneme marker mapping knowledge.

Vignette 5 (Session 3, 44:01)

After the teacher demonstrated the stroker order of writing <-S> on blackboard, he wrote *-s* next to <-S> on blackboard (see Figure 16) and said,

Teacher: 这个就是写在单词尾的/s/。[This is the word-final /s/.]

Figure 16

Blackboard Presentation of <-S> and -s



Vignette 6 (Session 4, 13:10)

After the teacher demonstrated the stroker order of writing <-M> on blackboard, he said,

Teacher: (while simultaneously writing *-m* on blackboard) 注意，这是/m/。

[Attention. This is /m/.]

Vignettes 5 and 6 showed that the teacher adopted the Latin letters to mediate the mapping between phoneme marker and sound. In Vignette 5, after demonstrating how to write the phoneme marker <-S>, the teacher wrote next to it the Latin letter *s* while simultaneously uttered /s/. This observation suggested that the letter *s* was used as mediation to map the phoneme marker <-S> with its sound /s/ (i.e., <-S> – *s* – /s/). So was it for the letter *s* in Vignette 6 which mediated the sound–to-phoneme marker knowledge instruction by connecting the phoneme marker <-M> to the sound /m/. It is worthwhile noting that in the two vignettes the letters *s* and *m* should not be misunderstood as transliteration for <-S> and

<-M>, because otherwise the teacher should have read their letter name (i.e., /es/ and /em/) rather than letter sound (i.e., /s/ and /m/).

The observation suggested that Latin letters were used to mediate the teaching of sound–to–phoneme marker mapping knowledge. Although the representation of sound conventionally makes use of symbols such as the International Phonetic Alphabet, the teacher seemed to notate the sound of Manchu uju and phoneme markers in his own way. His loose but pragmatic use of symbols can also be evidenced by the observation of his employment of *Hanyu Pinyin* and Chinese characters in teaching sounds mapped onto uju symbol blocks. For example, *pinyin* /ü/ was used to phonologically notate <IUI> (Session 1, 16:02), and the Chinese characters "七" [tɕʰi], "鸡" [tɕi], and "西" [ɕi] was applied to represent the sounds of <CI>, <JI> and <SI>, respectively (Session 4, 44:08).

Question-Posing Strategy of Assessment on Decoding Ability. The classroom observation revealed that the teacher frequently adopted the question-posing strategy to assess the students' uju decoding ability. For example, in Vignette 4 the teacher asked the students "这个念啥" [How do you read it?] after writing the uju <AM> on blackboard to evaluate whether they were able to decode it. Vignette 7 provides an additional example. The teacher raised a question on decoding <AL> by asking "这个念啥" [How do you read it?] to students, who responded by answering /al/.

Vignette 7 (Session 4, 08:00)

The teacher demonstrated how to write the word-final <L>. Then, he gave an example by writing <AL> on blackboard.

Teacher: 这个念啥？[How do you read it?]

Students: (several students) /al/.

A final note on the observation of sound–to–phoneme marker knowledge instruction. The above methods (i.e., segmentation of uju into phoneme markers, Latin letter-mediated instruction on phoneme marker sound, and the question-posing strategy of assessment on decoding ability) were often observed to be applied collectively in the instruction of sound–to–phoneme marker mapping knowledge. Vignette 8 provides an example.

Vignette 8 (Session 6, 02:00)

Teacher: (after demonstrating <AT> on blackboard) 这个念啥？如果根据刚才学的知识，它一定读作/at/。我们切一下。[How do you read it? According to the knowledge we just learned, it must be /at/. Let's cut it.]

Teacher: (segmenting <AT> into <A-> and <-T> and writing *a* and *t* next to them) 上面是/a/，底下是/t/，是吧？/at/。 [The upper is /a/, [and] the lower is /t/, right? /at/.]

6.5 Chapter Summary

This chapter reported the results of classroom observation, which explored from a qualitative perspective the Manchu orthographic knowledge instruction offered in a Chinese university. Test 2 on symbol–sound mapping complexity of Manchu has found that phoneme marker was adopted by novice learners as the basic orthographic unit in uju decoding. To complementarily understand this finding, this chapter addressed the following research question (RQ) regarding the association between the learners' choice of phoneme marker in decoding and the Manchu orthography knowledge instruction offered to them.

RQ 2-2: In what way are the results of the learners' decoding accuracy and speed on Manchu symbol blocks with different phoneme marker–sound mappings related to the instruction on Manchu orthography?

Firstly, the classroom observation found that regarding uju writing instruction, stroke transformation was occasionally applied to the final-position stroke in syllabically V uju (e.g., the rightward "tail" in <A>), but it was primarily applied to phoneme markers in CV uju (e.g., <-A> in <KA>). Taking phoneme marker as the basic orthographic unit in writing instruction was based on the teacher's belief that Manchu is an alphabet writing system. Therefore, combining consonant and vowel phoneme markers (or "stack them together") will be able to produce uju symbol blocks.

Secondly, with respect to uju reading instruction, it was found that the knowledge of sound–to–uju mapping and sound–to–phoneme marker mapping was instructed for the novice learners to read Manchu symbol blocks. The sound–to–uju mapping knowledge was taught in a global approach (Nag, 2017a, 2022) of uju choral reading. More importantly, the sound–to–phoneme marker mapping knowledge was instructed in an analytic approach (Nag, 2017a, 2022) of uju segmentation. In addition, it was found that the sound of phoneme markers was taught with the mediation of Latin letters. Moreover, the question-posing strategy was found to be frequently used by the teacher to assess students' decoding ability on uju.

Chapter 7 Discussion

7.1 Chapter Overview

This chapter discusses the findings of Test 1 on uju symbol blocks visual complexity, Test 2 on symbol–sound mapping complexity, and classroom observation on Manchu orthographic knowledge instruction. The chapter first reiterates the research questions and presents the key findings. Afterwards, it discusses the findings in relation to previous theories and studies of reading. Finally, it posits that decoding Manchu uju symbol blocks requires intra-symbol processing (Nag, 2022), which involves (a) at the feature level, identification and configuration of visual features into phoneme markers, and (b) at the symbol level, conversion of phoneme markers into sounds.

7.2 Key Findings

The overarching goal of the present study was to expand the understanding of how complexity of Manchu writing system affects novice learners' decoding performance on Manchu uju symbol blocks. The goal was addressed by answering the following four research questions (RQs) regarding visual complexity and symbol-sound mapping complexity of Manchu.

RQ 1-1: To what extent does the visual complexity of Manchu symbol blocks associate with the learners' decoding accuracy and speed?

RQ 1-2: In what ways do the measures of the visual complexity of Manchu symbol blocks predict the learners' decoding accuracy and speed?

RQ 2-1: Do the learners perform differently when they decoded Manchu symbol blocks with phoneme markers in one–to–one and one–to–multiple mapping with sound?

RQ 2-2: In what way are the results of the learners' decoding accuracy and speed on Manchu symbol blocks with different phoneme marker–sound mappings related to the instruction on Manchu orthography?

Firstly, regarding visual complexity of Manchu uju symbol blocks, it was found that when frequency of the uju given in Test 1 was not statistically controlled, the measures of pixel count (PX), perimetric complexity (PC), simple features (SF), connected points (CP) but disconnected components (DC) are positively associated with each other, and with novice learners' decoding accuracy and speed measured by error rate (ER) and reaction time (RT), respectively. However, after uju frequency was controlled, only CP remains to be positively associated with ER and RT. Moreover, CP was found to be the unique predictor for decoding accuracy of Manchu symbol blocks, and CP and DC were found to be predictors for decoding speed. The models are provided as follows.

$$\text{Error rate} = 0.322 + .078 \times \text{the number of CP}$$

$$\text{Reaction time} = 3543.177 + 276.126 \times \text{the number of CP} + 459.463 \times \text{the number of DC}$$[16]

Secondly, with respect to symbol–sound mapping complexity of the Manchu writing system, no difference was found in novice learners' decoding accuracy or speed when uju in Block 1 (all comprising phoneme markers of an uju in consistent one–to–one mapping with sound) and these in Block 2 (the target phoneme marker of an uju in less consistent one–to–multiple mapping with sound) were presented in the manipulated "blocked" condition (i.e., uju in Blocks 1 and 2 were provided sequentially) and "mixed" condition (i.e., uju in Blocks

[16] Reaction time was measured in milliseconds.

1 and 2 were randomly mixed up and provided). However, more accurate and rapid performance was found in decoding uju in Block 1 than was in Block 2.

Finally, the classroom observation on Manchu orthography instruction revealed that stroke transformation was applied to phoneme markers in uju writing instruction. In addition, it also revealed that the global and analytic approaches (Nag, 2011, 2017a, 2017b) of uju choral reading and uju segmentation were applied to Manchu symbol blocks reading instruction.

7.3 Discussion

7.3.1 Visual Complexity and Manchu Uju Symbol Blocks Processing

Pixel Count, Perimetric Complexity and Simple Features. Test 1 found that the levels of visual complexity of uju affect Manchu reader's symbol block processing accuracy and speed. The visual complexity effects were suggested by the association between uju decoding performance and PX, which was used as an indicator of symbol visual complexity in examining akshara of Kannada writing system (Nag et al., 2014). The present study confirms that PX is a useful measure of visual complexity based on its association with the number of errors made and the amount of time spent in Manchu symbol blocks processing. In a similar vein, PC, a measure that captures the overall configural complexity of symbols by the perimeter squared over the ink area (Pelli et al., 2006; Watson, 2012) also demonstrated inhibitory effects on uju decoding accuracy and speed. This result is consistent with that of Pelli and colleagues' study (2006), in which PC was found to be inversely proportionate to graph identification efficiency. In addition, SF, an intuitively more straightforward measure of visual complexity, also inhibited the processing accuracy and speed of Manchu symbol blocks. Similar inhibitory effects were identified in processing Chinese characters (Yu et al., 2018) and Japanese kanji (Tamaoka & Kiyama, 2013). To explain their finding, Yu and colleagues (2018) speculated that visual features (i.e., individual strokes) were configured

into a hierarch of orthographic codes (i.e., radicals and characters) before they are converted into abstract lexical codes. This speculation suggests that character with more visual features requires more time to be converted into the lexical codes. Manchu symbol blocks, like Chinese characters, can be regarded as comprised of visual features such as the "crest", the "tail", the "belly", the "tooth", etc. Thus, the configural process of the visual features may generate Manchu phoneme markers, which are subsequently assembled into a symbol block before being converted into lexical code. Since Manchu symbol blocks comprising greater visual features require more visual-spatial discrimination to disambiguate and longer time to generate the lexical codes, SF may therefore affect the processing accuracy and speed.

Moreover, in the multiple linear regression analysis, the independent variables of PX, PC and SF were excluded from the models for predicting uju decoding accuracy and speed. Statistically, the exclusion was due to multicollinearity existing between PX, PC and SF (see Section 4.4.6 in Chapter 4). In other words, an increase in the number of SF in an uju would result in greater values of PX and PC. From a theoretical perspective, the collinearity between PX and SF suggested that they may measure the same construct of "visual feature density" (Nag et al., 2014, p. 322), as the "ink" area captured by PX would expand with the increase of SF in a printed symbol. However, with respect to PC (i.e., the inside and outside perimeters of a binary black-and-white image, squared, divided by the image area and 4π), it is a Gestalt measure which considers how every part affects every other parts in shape perception and their relationship is described "as a whole" by the relationship of P^2 / A (Attneave & Arnoult, 1956, p. 467). The circle was considered the visually least complex shape with a PC value of 1 (Watson, 2012). Based on the Gestalt approach to perception, a circle is a visual shape of "goodness" due to its properties such as symmetry (Pelli et al., 2006). Thus, the distribution of PC values greater than 1 indicates the levels of deviation of visual shapes from the Gestalt "goodness", which poses possible difficulties in shape

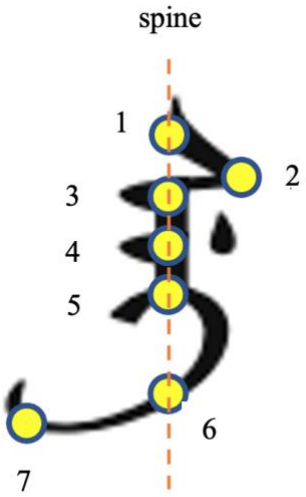
perception. Therefore, theoretically, the measure of PC in visual complexity indicates the whole figural "goodness" of an uju in Manchu symbol blocks processing.

Connected Points. An interesting finding in Test 1 is that CP showed inhibitory effects on processing accuracy and speed of uju symbol blocks, whether frequency of the uju given in the test was controlled. Moreover, CP was also identified as the unique predictor for uju decoding accuracy. CP were previously found to be critical in the recognition of objects (Biederman, 1987) and letters and words (Lanthier et al., 2009; Szwed, Cohen, Qiao, & Dehaene, 2009). Following Szwed and colleagues (2009), the present study conceptualised CP as a junction where two or more simple features adjoin. For example, in Figure 17, the CP indexed by 5 is where the vertical line, the long "tooth" and the "curve" adjoin. When reading a printed Manchu symbol block, a reader needs to discriminate all the adjoining visual features from each other based on their geometric characteristics such as relative orientation, relative lengths, and shapes (Changizi, Zhang, Ye, & Shimojo, 2006). This visual processing of an uju leads to the identification of individual visual feature, which may be facilitated by the CP functioning to mark the termination of the visual features adjoining together. As a result, when visually processing an uju with more CP, readers may spend longer time discriminating the visual features connected together, and they may be more prone to make errors in distinguishing small features, such as stroke length (e.g., a short "tooth" <  > vs. a long "tooth" < >) and stroke orientation (e.g., a downward line segment in < > /m/ vs. an upward line segment in < > /l/). In other words, the junction of visual features occurring at CP inhibited the accuracy and speed of uju processing. Using CP to discriminate the adjoining visual features is perhaps but one of its possible functions in Manchu symbol blocks processing—the visual feature discrimination function. Moreover, CP may also function to connect visual features into the recognisable pattern of Manchu

phoneme markers. This is called the visual feature configuration function which will be discussed in Section 7.7.3.

Figure 17

Connected Points in the Uju <SEK>



Disconnected Components. In Test 1 on uju visual complexity, no significant correlation was found between the number of DC and participants' decoding performance. From a statistical perspective, the result can be attributed to *range restriction* (Sackett & Yang, 2000), meaning that there is only a limited amount of variability in the data. With other things being equal, the small variability will result in lower value of *r* (Goodwin & Leech, 2006). As the DC values of the uju (n = 18) given in Test 1 ranged between 1 and 2, the correlation between DC and decoding performance was thus low (when uju frequency not under control $r_{\text{DC-ER}} = .141$, $r_{\text{DC-RT}} = .286$; when uju frequency under control $r_{\text{DC-ER}} = -.115$, $r_{\text{DC-RT}} = .164$). The restricted variability in DC was due to the architectural design of Manchu symbol blocks. The DC in an uju comes from the phoneme markers stacked together (i.e., the main architecture) and the diacritical dot or circle systematically positioned to the left or right

of a phoneme marker (see <SEK> in Figure 17). Since only a limited number of Manchu phoneme markers take diacritics, the uju of V, CV, and CVC structures given in Test 1 had maximumly two DC.

However, the low correlation between DC and decoding performance did not indicate that DC is not relevant in uju reading. In fact, DC and CP were identified as predictors in the model for uju decoding speed, which suggested that theoretically DC should be explained in combination with CP in understanding uju processing speed. Historically, the purpose of systematically adding diacritical dot and circle to phoneme markers was for disambiguation. For example, the <ᡳ> corresponding with both /ɑ/ and /ə/ was disambiguated by adding a dot to its right <ᡳᠸ> which represent /ə/ only. Similarly, <ᡴ> [q^h], <ᡵ> [q], and <ᡶ> [χ] were distinguished by the placement of null diacritical mark, a dot, and a circle, respectively. The nonlinear complex arrangement of the visual features of dot and circle requires visuo-spatial processing for printed symbols (Nag, 2014; Vaid & Gupta, 2002). Thus, the predictors of CP and DC in the model for uju reading speed suggested that decoding an uju comprising two DC may require readers to process the visual features adjoining at the CP (see discussion on CP above). In addition, it may also involve processing the visual features of dot and circle placed to the left or right of a phoneme marker. Consequently, this nonlinear visuo-spatial processing of the features adjoining at CP and discontinued from the main architecture may collectively consume cognitive resources.

7.3.2 Symbol-Sound Mapping Complexity and Manchu Uju Symbol Blocks Processing

Basic Orthographic Unit in Uju Processing. Test 2 found that novice learners of Manchu did not perform differently when decoding uju presented in the manipulated "blocked" and "mixed" conditions.[17] The fact that the decoding performance was not affected

[17] In the "blocked" condition, uju in Block 1 (all comprising phoneme markers in an uju are in one–to–one mapping with sound) and Block 2 (the target phoneme marker in an uju is in one–to–multiple mapping with

by blocking manipulation suggested that orthographic unit of the same size (either the large grain size uju or the small grain size phoneme marker) may be applied in Manchu symbol blocks processing. Otherwise, applying a mixture of orthographic units would have incurred switching cost in decoding uju presented in the "mixed" condition, which was in fact not observed in Test 2. Furthermore, when uju decoding performance in Block 1 (all phoneme markers of an uju in consistent one–to–one mapping with sound) and Block 2 (the target phoneme marker of an uju in less consistent one–to–multiple mapping with sound) was compared, the learners were found to make more errors and spend longer time in responding to the uju presented in Block 2. As the uju between Block 1 and Block 2 were pair-matched with respect to their visual complexity, syllabic structure, and frequency, the inhibitory effects in uju decoding may be attributed to the inconsistent mapping between the target phoneme marker and sound in Block 2. This result suggested that phoneme marker may be applied as the basic orthographic unit in processing Manchu symbol blocks.

Mixed Symbol-Sound Mapping and Uju Decoding. Manchu has a mixed mapping (Nag, 2007, 2017a, 2021) in terms of its orthography–to–phonology correspondence. Manchu uju are in four–to–one mapping with sound (i.e., syllable) because an uju has four positional variants (see Figure 4). When uju are further segmented, the phoneme markers produced are mostly (74 out of 82) corresponded with one phoneme (see Table 3 for details). The complex symbol–sound mapping relations may pose challenges for novice learners, who were found to address this issue by adopting phoneme markers in uju decoding. This finding seemed to confirm the orthographic depth hypothesis (Frost et al., 1987; Katz & Frost, 1992), which posited that for a transparent orthography, the phonology of a printed symbol is prelexically assembled by grapheme–phoneme correspondence. In other words, the consistent mapping

sound) were presented sequentially. Yet, in the "mixed" condition, all uju in Blocks 1 and 2 were randomly mixed up and presented.

between phoneme marker and sound in Manchu determines that it would be cognitively less demanding for novice learners to read symbol blocks by applying the strategy of generating phonological information through phoneme marker–to–sound conversion. In addition, the psycholinguistic grain size theory (Ziegler & Goswami, 2005) suggested that the orthographic consistency affects granularity, which leads the readers to develop different decoding strategies. As the smaller grain sizes in some European languages tend to be relatively less consistent, learners adopted multiple strategies that use both small and large grains to address their reduced reliability for decoding. Yet, contrary to these European orthographies, the small grain phoneme marker in Manchu orthography is more consistent with sound than is the larger grain uju. Therefore, the novice learners adopted the decoding strategy of using the small grain phoneme markers in processing uju symbol blocks. However, it is worth noting that the uju given in Test 2 were in the relatively simple CV and CVC syllables. It still remains unknown whether alternative decoding strategies (using large grain uju, or a mixture of both small and large grains of phoneme marker and uju) would be applied in reading syllabically more complex Manchu words, such as *katun* "strong" <CV.CVC>, *bithe* "word, book" <CVC.CV>, *agūra* "utensil" <V.CV.CV>, *muduri* "dragon" <CV.CV.CV>, *amgambi* "to sleep" <VC.CVC.CV>, etc.

Orthographic Breadth and Uju Decoding. While the complex mixed symbol–sound mapping affects uju processing, learning to read Manchu symbol blocks is also constrained by the orthographic breadth of its inventory size (Nag, 2007, 2017a). Manchu is an extensive orthography when uju (= 5,600ish for the entire syllabary; = 524 for the fundamental first section of the syllabary) is considered the basic written symbol, but a contained orthography when phoneme marker (= 82) is regarded as the basic symbol. The orthographic breadth implies that there could be a trade-off between the cognitive demands and the inventory size in orthography learning (Nag, 2011). When learning an extensive orthography, the demand

on memory is high but there is reduced confusion on symbol–to–sound correspondence. In contrast, in learning a contained orthography, memory demand is relatively low but the complex way that symbols represent sounds could consume more cognitive resources. Yet, the extensive Manchu orthography goes hand in hand with opaque symbol–to–sound mapping (i.e., four positional variants of an uju mapped onto one sound), and the contained Manchu orthography goes along with a relatively more transparent symbol–to–sound mapping (i.e., 74 out of 82 phoneme markers mapped onto one sound). As orthography instruction is a matter of "efficiency" in matching its phonology to orthography (Katz & Frost, 1992, p. 67), the contained-transparent association determines that using phoneme marker in Manchu symbol blocks processing would be more efficient and cognitively less demanding for the novice learners, due to the symbol–sound mapping invariance and smaller inventory size.

7.3.3 Model for Manchu Uju Decoding

Taken together, this study posits that decoding Manchu uju symbol blocks requires intra-symbol processing (Nag, 2022), which involves (a) at the feature level, identification and configuration of visual features into phoneme markers, and (b) at the symbol level, conversion of phoneme markers into sounds.

First, at the feature level, decoding a Manchu uju may demand the readers to identify the constituent visual features in a symbol block. A basic question in printed symbol identification is whether they are perceived by features or as a global whole (Fiset et al., 2008). The present study suggested that Manchu uju is perceived by visual features (Pelli et al., 2006). As discussed above, the visual processing of uju results in the identification of individual visual feature, which is facilitated by the CP functioning to mark the termination of the visual features adjoining together. For example, in Figure 17, CP 3 indicates the termination of a slanted segment line and a short "tooth". However, visual feature

discrimination is not sufficient for processing uju. The identified visual features characterised by relative orientation, relative lengths, and shapes (Changizi et al., 2006) need to be configured into Manchu phoneme markers by visual-spatial processing. In this process, the visual features are connected into the recognisable and distinguishable pattern of individual Manchu phoneme marker through the cognitive operation of organisation (Feuerstein, Falik, & Feuerstein, 2006). For instance, in Figure 17, CPs 5, 6, and 7 connect a long "tooth", a curve, a slanted line segment, and a short line segment, which are configured into the phoneme marker < > <-K>. The visual-spatial processing occurs top to bottom along the spine within the configuration of an uju, but it is nonlinear. When processing diacritical dots or circles systematically positioned to the left or right of the main architecture of an uju, the visual-spatial processing may deviate from the spine to combine diacritics to relevant phoneme markers (e.g., in Figure 17 the diacritical dot placed to the right of < > /ɑ/ transforms it to /ə/).

In addition to the identification and configuration of visual features into Manchu phoneme markers, decoding uju symbol blocks also involves applying the sound–to–symbol mapping knowledge to convert phoneme markers into sounds (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). The basic question here is whether uju is read by assembled phonology (generated by symbol–sound mapping rules) or addressed phonology (accessed as a whole from lexicon through the presentation of the printed word) (Katz & Frost, 1992). Since the orthographic unit of phoneme marker rather than uju was applied in Manchu symbol block reading, the present study speculates that uju may be decoded by assembled phonology. To be specific, the process initiates by considering the first phoneme marker and converts it into corresponding sound, then the first two phoneme markers, and then the first three phoneme markers. Until all the markers are translated into sounds, the uju symbol block is decoded (Coltheart, 2005). For example, to decode the uju in Figure 17, the first phoneme

marker <> <S-> is converted into /s/ based on the sound–to–symbol mapping between them. Then, the second phoneme marker <> <-E-> is converted to /ə/ so that /sə/ is obtained. Finally, with the conversion of the phoneme marker <> <-K> into /k/, the uju is decoded as /sək/. This process of uju decoding can be evidenced by the higher accuracy and faster speed in reading the uju constituted by phoneme markers that are all in consistent one–to–one mapping with sound. In Manchu orthography, some symbol–sound mapping rules are context-sensitive. For example, <> corresponds with /ɑ/ in most symbol blocks, but /ə/ after <T> and <D> and /n/ after a vowel phoneme marker. Consequently, when applying sound–to–symbol conversion in decoding, processing the context-sensitive phoneme marker would result in conflicting activation at the phoneme level. Resolution of the conflicts takes time and readers are more inclined to make errors during the process, which is responsible for the lower accuracy and speed in decoding. Moreover, another line of evidence that assembled phonology may be applied comes from the successful decoding of the nine uninstructed uju in CV and CVC syllables provided in Test 2, which would perhaps not be possible if the addressed phonology was applied.

7.3.4 Instructional Strategies and Manchu Uju Reading

The classroom observation found that both holistic and analytic strategies, in the forms of uju choral reading and uju segmentation, were applied in Manchu reading instruction. Yet, the novice learners in the present study appeared to only apply the analytic approach in reading by decomposing uju into its component phoneme markers. This finding confirms previous studies on reading (Cho et al., 2020; Nag & Snowling, 2012) which found that novice learners with relatively advanced literacy ability adopted the analytical approach to read. However, unlike the novice readers in previous studies, the participants in the present study who are also novice learners were not found to take the holistic method to read uju as undifferentiated whole symbol blocks. Not using holistic method in reading Manchu may be

resulted from the accessibility problem (Ziegler & Goswami, 2005), which means that some phonological units in Manchu were not available to the readers during instruction. For example, among the 16 items given in Test 2, five CV uju (*sū*, *ce*, *cū*, *we* and *fū*) and all four CVC uju (*jūs*, *tib*, *tus* and *teb*) were not at all instructed in class. Since application of the approaches to reading depends on the learner's level of knowledge on orthography (Nag, 2011), unfamiliarity with these uju made it impossible for the novice learners to read it holistically. Alternatively, the participants analysed the internal details of the uju and decoded it by applying the phoneme marker-sound correspondence.

Moreover, the fact that Manchu has an "extensive" inventory size (Nag, 2007) of over 1,400 uju with each having four positional variants provides another reason why the analytic approach was applied in reading uju symbol blocks. The holistic instructional method which treats each uju as a singleton whole symbol block is a less effective method for teaching Manchu because it was implausible for the teacher to instruct the phonological information encoded in 1,400ish uju within ten sessions of class. This limitation suggests that the learners will not be able to access the phonology of the *un*instructed uju. However, the analytic strategy of instruction splits uju into the comprising phoneme markers (which is in a relatively much smaller number of 82 in Manchu compared with the 1400ish uju) and teaches the sounds encoded in them. As a result, the learners will be able to decode both the instructed and uninstructed uju symbol blocks by applying the phoneme marker-sound mapping rules. Applying the analytic approach was found in learning other extensive orthographies with a large inventory size such as Chinese (McBride, 2016) and Kannada (Nag et al., 2014), in which the symbols were not explicitly taught but the learners nevertheless acquire their internal details (Nag, 2007).

Another possible reason why the novice readers utilised the analytic strategy rather than the holistic strategy or both of them is that they may have acquired the alphabetic

principle of Manchu. The participants are skilled readers in English who have been taking English courses for over ten years beginning from Grade 3 in primary school as prescribed by the National English Curriculum. They may thus be able to transfer specific knowledge from English (as an L2 or foreign language) to Manchu. Specifically, they may have acquired the alphabetic principle in English and transferred the symbol-sound mapping knowledge to reading Manchu, which can arguably be regarded as an alphabetic writing system (see Section 1.6.6 for a more detailed discussion on the typology of Manchu writing system).

Finally, only applying holistic and/or analytic approaches may not be sufficient for reading Manchu writing system characterised by mixed symbol-sound mapping. Thus, the strategic approach (Nag, 2011, 2017a, 2017b), which applies orthographic knowledge to disambiguate the phonology encoded depending on a word context, could facilitate Manchu reading. For example, in the Manchu word *abka* "sky, heaven", the first uju <> (<AB->) can be regarded as mapped onto the syllable /ab/, or alternatively, onto the two phonemes of /a/ and /b/. Similarly, the second uju <> (<-KA>) can be seen as in correspondence to the syllable /ka/ or two phonemes /k/ and /a/. If the two constituent uju <> (<AB->) and <> (<-KA>) in *abka* are processed holistically when reading *abka*, it will potentially produce two problems. The first one is that the two uju are visually complex and thus cognitively demanding for the novice readers to recognise. Secondly, as the uju <> (<AB->) was not included in the course curriculum, the learners would not be able to learn its pronunciation from classroom instruction, which is probably the only channel for learning Manchu orthographic knowledge. On the other hand, however, if the word *abka* is read analytically by decomposition into the four component phoneme markers of <> (<A->), <> (<-B->), <> (<-K->), and <> (<-A>), <> and <> are relatively easy to read because they each represent one sound only (i.e., /ɑ/ and /b/). Yet, decoding <> (<-A>) is cognitively

demanding for the readers as it is in multiple mapping with the sounds /a/, /ə/ and /n/. Furthermore, while <ᡤ> is only mapped onto /k/, graphically it consists of two "short teeth" <ᡨ>, which in its own right is a Manchu phoneme marker encoding /a/, /ə/ and /n/. All these complex mapping relations will cause great confusion for novice learners of Manchu. The strategic approach, compared with holistic and analytic methods, seems to be a more effective option. For the first uju <ᡝᡳ> (<AB->, /ɑb/), the readers can take advantage of the alphabetic principle to analyse and segment it into the phoneme markers of <ᡝ> (<A->) and <ᡳ> (<-B->), which are in the transparent one–to–one mapping with sound /ɑ/ and /b/, respectively. Moreover, the second uju <ᡤᡝ> (<-KA>), which corresponds only with the phonological syllable /ka/, is better treated as a whole unit than being processed analytically by decomposition. The transparent <-KA> – /ka/ mapping will reduce the consumption of cognitive resources for the novice readers. Therefore, taking the strategic approach to disambiguate the phonology of /ɑbkɑ/ encoded in *abka* appears to be more effective for the novice readers. In this regard, the mixed orthography-phonology mapping in Manchu would be to the novice learners' advantage if handled appropriately, as it could provide the possibility of flexibly adopting the basic orthographic unit in reading Manchu.

However, it is important to note that as the participants only took a limited ten sessions of instruction focusing on V, CV and CVC syllables in Manchu, they may not yet be able to develop a strategic approach in reading due to their unfamiliarity with the orthography. In other words, they may not be aware that applying the orthographic units of both phoneme marker and uju can facilitate reading Manchu.

7.4 Chapter Conclusion

This chapter first discussed the key findings on the association between visual complexity and the novice learners' decoding performance. In term of the visual processing of Manchu uju symbol blocks, pixel count and simple features both capture the "visual

feature density" (Nag et al., 2014, p. 322), which have an inhibitory effect on uju decoding accuracy and speed. Moreover, perimetric complexity (PC) measures uju visual complexity by Gestalt approach to perception. The PC value greater than 1 (i.e., the value of the visually least complex shape of circle) indicates the levels of deviation of visual shapes from the Gestalt "goodness", which poses possible difficulties in shape perception. In addition, connected point in an uju is a junction where two or more simple features adjoin. In visual processing an uju, a reader needs to discriminate all the adjoining visual features from each other based on their geometric characteristics such as relative orientation, relative lengths, and shapes (Changizi et al., 2006). Finally, disconnected components should be accounted for in combination with CP in understanding the visual processing of uju. The nonlinear visuo-spatial processing of the visual features adjoining at CP and the DC discontinued from the main architecture of an uju may collectively consume cognitive resources.

In addition, this chapter also discussed the key findings on the symbol–sound mapping complexity of the Manchu writing system. The inhibitory effect of the inconsistent phoneme marker–sound mapping on the novice learners' decoding performance suggested that phoneme markers may be applied as the basic orthographic unit in processing Manchu symbol blocks. Moreover, the consistent mapping between Manchu phoneme marker and sound determines that it would be cognitively less demanding for the novice learners to use the small grain phoneme markers in processing uju symbol blocks. Adopting phoneme marker in decoding is also associated with the contained Manchu orthography, which goes along with a relatively more transparent symbol–to–sound mapping (i.e., 74 out of 82 phoneme markers mapped onto one sound). Thus, the contained-transparent association determines that using phoneme marker in Manchu symbol blocks processing would be more efficient and cognitively less demanding for the novice learners, due to the symbol–sound mapping invariance and smaller inventory size. Finally, using phoneme marker as the basic

orthographic unit in uju decoding was found to be associated with the teaching approaches adopted in Manchu orthography instruction. Compared with the global approach of uju choral reading, the analytic strategy of uju segmentation is a more sustainable method which allows the learners to decode the uju whose phonology had not been instructed in class.

Finally, taken together, this study proposes a two-level model of Manchu symbol blocks reading, which requires intra-symbol processing (Nag, 2022). At the feature level, Manchu symbol blocks decoding may demand the readers to identify the individual visual features and configure them into phoneme markers. Moreover, at the symbol level, decoding symbol blocks involves applying the sound–to–symbol mapping knowledge to convert the generated phoneme markers into sounds.

Chapter 8 Conclusion, Contributions, and Limitations

8.1 Chapter Overview

This chapter concludes the present study by summarising the key findings in relation to the research aim and questions. It also discusses the contributions and values thereof. Finally, it reviews the limitations of the study and proposes recommendations for future research.

8.2 Conclusion

This study investigated the association between the Manchu writing system and novice learners' decoding performance on uju symbol blocks. Manchu writing system was examined on two parameters: visual complexity of Manchu symbol blocks and mapping complexity between symbol and sound. Firstly, regarding visual complexity, connected points (CP) may function to mark the termination of the adjoining visual features that are of various orientations, relative lengths, and shapes in an uju. The identified visual features are configured into the recognisable and distinguishable pattern of Manchu phoneme marker(s) through the cognitive operation of organisation. In addition, the predictor of disconnected component in the reading model suggests that reading an uju may also involve processing the visual features of diacritical dot and circle systematically placed to the left or right of a phoneme marker. Consequently, this nonlinear visuo-spatial processing of the visual features that both adjoin at CP and deviate from the main architecture of an uju are involved in Manchu reading.

Secondly, with respect to symbol–sound mapping complexity, the finding that learners' decoding performance was not affected by blocking manipulation suggested that orthographic unit of the same size (either the large grain uju or the small grain phoneme marker) may be applied in Manchu symbol blocks processing, because applying a mixture of orthographic units would have incurred switching cost in decoding uju presented in the

"mixed" condition. Furthermore, the inhibitory effect of the inconsistent phoneme marker–sound mapping on the participants' decoding performance suggested that phoneme markers may be applied as the basic orthographic unit in processing Manchu symbol blocks.

Thirdly, using phoneme marker as the basic unit in uju decoding was further corroborated by the classroom observation that the sound–to–phoneme marker mapping knowledge was instructed in analytic approach of uju segmentation (e.g., <NA> → <N-> + <-A>. The finding indicated that using phoneme marker as the basic unit in Manchu orthography instruction would be pedagogically more effective and cognitively less demanding for the novice learners, due to the symbol–sound mapping invariance and smaller inventory size of Manchu phoneme markers (i.e., 74 out of the entire 82 phoneme markers are mapped onto one sound). The finding also suggested that compared with the global approach of reading instruction, the analytic strategy of uju segmentation benefits the learners by allowing them to read unfamiliar and new uju.

Taking the findings together, this study posits a two-level model of Manchu symbol blocks reading, which requires intra-symbol processing (Nag, 2022). At the visual feature level, Manchu symbol blocks decoding may demand the readers to identify the individual visual features and configure them into phoneme markers. Further, at the phoneme marker level, decoding symbol blocks involves applying the sound–to–symbol mapping knowledge to convert the generated phoneme markers into sounds.

8.3 Contributions

The present study, to the best of the author's knowledge, is the first empirical research conducted on Manchu writing system. Theoretically, it found the important function of connected points (CP) in visual symbol processing. CP may function to mark the termination of the visual features adjoining together and it may also facilitate the readers to configure the visual features into individually identifiable phoneme markers in reading. Based on the

evidence, this study also proposes a two-level reading model which contributes to uncovering the visual and cognitive mechanism of symbol blocks reading. This model suggests that CP was adopted to process the visual simple features connected together in a symbol block, and then the phoneme marker–sound conversion was applied to reading. This model adds up to the existing theories of reading which were derived from a relatively narrow range of empirical studies on English and other European orthographies (Daniels & Share, 2018). Another contribution of this study is that it provides empirical evidence to support the development of theorising of reading. In addition to orthographic depth hypothesis (Frost et al., 1987; Katz & Frost, 1992) and psycholinguistic grain size theory (Ziegler & Goswami, 2005), the constructs of orthographic breadth (Nag, 2007, 2017) and intra-symbol processing (Nag, 2022) are also important dimensions that need to be considered in understanding reading.

This study also has pedagogical implication for endangered languages instruction. The revitalisation of endangered languages requires a comprehensive understanding of the fields including linguistics, education, and pedagogy (Hinton, 2011). Findings of this study could inform literacy instruction on endangered languages with respect to decisions on the optimal grain size and instructional approach. Compared to the long class hours and heavy workload required for instructing the larger grain size uju (524 for the fundamental first section of the syllabary; 5,600ish for the entire syllabary), teaching the 82 smaller size phoneme markers would be pedagogically more effective for the teachers. As observed by this study, it allowed the teacher to complete instructing Manchu orthographic knowledge within a very short period of ten class hours. From a cognitive perspective, learning phoneme markers would be less demanding for the novice learners due to the transparent symbol–sound mapping and smaller inventory size. This may help maintain the learners' interest in learning Manchu, a language practically not used at all in everyday context. Moreover, with

respect to the instructional strategies of reading, the present study suggested that the analytic strategy of decomposing uju into comprising phoneme markers has the advantage of allowing learners to read unfamiliar and new uju. This is a relatively more sustainable approach compared with the global/holistic method of uju choral reading because the learners would be able to read uju by analysing the internal details instead of relying on the teacher for learning uju pronunciation. As time is the most pressing issue in preserving endangered languages, the application of appropriate teaching approaches will increase the effectiveness and efficiency of maintaining the languages facing extinction.

8.4 Limitations

The present study is limited in that mixed methods were applied to address research questions (RQs) 2-1 and 2-2 but not RQs 1-1 and 1-2, which quantitatively investigated the association between visual complexity and symbol blocks decoding by correlation tests and multiple linear regression. The non-use of qualitative research in examining symbol visual complexity led to the impossibility of triangulating findings at the level of research questions.

Moreover, this study also lacks a pre-test which would have allowed to control for pre-existing Manchu knowledge as a covariate. However, the participants were asked to self-report in a questionnaire whether they had any experience of learning Manchu before university, and the results showed that none of the participants reported such learning experience.

In addition, the study lacks the use of advanced statistical methods that could have shed more light on data analysis. For instance, the results of Test 1 showed that there was multicollinearity existing between the visual complexity measures of pixel count, perimetric complexity and simple features. To address this issue, these measures were entered into regression models separately. In fact, this issue could have been dealt with using different

procedures of regression such as Support Vector Regress (SVR), a method belonging to the machine learning model which is currently beyond the researcher's grasp.

Another limitation of this study is the provision of only ten sessions of 45 minutes instruction on Manchu orthography to the novice learners. However, this limitation on class hours did not negatively affect validity for the tests on visual and symbol-sound mapping complexity, because the orthography of Manchu was fully covered during instruction. In other words, the orthographic knowledge necessary for naming the symbol blocks in V, CV, and CVC structures had been instructed to the participants. Yet, the limited sessions of instruction did make the more challenging CVCV symbol blocks given in the intended Test 3 on syllabic complexity beyond their ability to name.

Finally, the length of Manchu symbol blocks given in the tests was limited to 1 to 3 phoneme markers in Test 1 on visual complexity and 2 or 3 in Test 2 on symbol–sound mapping complexity. As a result, findings of this study on visual and mapping complexity may not be generalizable to longer Manchu symbol blocks. If a comprehensive theory of Manchu symbol reading is to be built, more complex Manchu symbol blocks should also be examined in the future.

References

Abdelhadi, S., Ibrahim, R., & Eviatar, Z. (2011). Perceptual load in the reading of Arabic: Effects of orthographic visual complexity on detection. *Writing Systems Research*, *3*(2), 117–127. https://doi.org/10.1093/wsr/wsr014

An, S. (1996). manwen shier zitou wo jian [My thoughts on the twelve sections of Manchu syllabary]. In Z. Wang (Ed.), *Manchu history and culture* (pp. 233–243). China Minzu University Press.

Arnoult, M. D. (1960). Prediction of perceptual responses from structural characteristics of the stimulus. *Perceptual and Motor Skills*, *11*, 261–268.

Attneave, F., & Arnoult, M. D. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin*, *53*(6), 452–471. https://doi.org/10.1037/h0044049

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115–147. https://doi.org/10.4324/9781351156288-24

Biesta, G. (2010). Pragmatism and the philosophical foundations of mixed methods research. In Abbas Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 95–118). SAGE Publications.

Brown, G. D. A., & Deavers, R. P. (1999). Units of analysis in nonword reading: Evidence from children and adults. *Journal of Experimental Child Psychology*, *73*(3), 208–242. https://doi.org/10.1006/jecp.1999.2502

Bryman, A. (1988). *Quantity and quality in social research*. Routledge.

Bryman, A. (2012). *Social research methods* (4th ed.). Oxford University Press.

Carreiras, M., Alvarez, C. J., & de Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language*, *32*, 766–780.

Chang, L. Y., Chen, Y. C., & Perfetti, C. A. (2018). GraphCom: A multidimensional measure of graphic complexity applied to 131 written languages. *Behavior Research Methods*,

50(1), 427–449. https://doi.org/10.3758/s13428-017-0881-y

Chang, L. Y., Plaut, D. C., & Perfetti, C. A. (2016). Visual complexity in orthographic learning: Modeling learning across writing system variations. *Scientific Studies of Reading*, *20*(1), 64–85. https://doi.org/10.1080/10888438.2015.1104688

Changizi, M. A., Zhang, Q., Ye, H., & Shimojo, S. (2006). The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. *The American Naturalist*, *167*(5), E117–E139. https://doi.org/10.1086/502806

Chelune, G. J., Naugle, R. I., Lüders, H., Sedlak, J., & Awad, I. A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology*, *7*(1), 41–52. https://doi.org/10.1037/0894-4105.7.1.41

Cho, J.-R., McBride, C., & Kim, B. (2020). Effective teaching instructions for Hangul learning among Korean kindergartners. *Reading and Writing*, *33*, 1791–1808. https://doi.org/10.1007/s11145-019-09985-0

Collie, A., Maruff, P., Darby, D. G., & McStephen, M. (2003). The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test-retest intervals. *Journal of the International Neuropsychological Society*, *9*, 419–428. https://doi.org/10.1017/S1355617703930074

Coltheart, M. (2005). Modeling reading: The dual-route approach. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 6–23). Blackwell. https://doi.org/10.1002/9780470757642.ch1

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256.

Conrad, M., & Jacobs, A. M. (2004). Replicating syllable frequency effects in Spanish in

German: One more challenge to computational models of visual word recognition. *Language and Cognitive Processes*, *19*(3), 369–390. https://doi.org/10.1080/01690960344000224

Cossu, G., Gugliotta, M., & Marshall, J. C. (1995). Acquisition of reading and written spelling in a transparent orthography: Two non parallel processes? *Reading and Writing*, *7*, 9–22. https://doi.org/10.1007/BF01026945

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.). Sage.

Daniels, P. T. (2018). *An exploration of writing*. Equinox Publishing Ltd.

Daniels, P. T., & Share, D. L. (2018). Writing system variation and its consequences for reading and dyslexia. *Scientific Studies of Reading*, *22*(1), 101–116. https://doi.org/10.1080/10888438.2017.1379082

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. Chapman & Hall.

Enders, C. K. (2010). *Applied missing data analysis*. The Guilford Press.

Feuerstein, R., Falik, L. H., & Feuerstein, R. (2006). The learning potential assessment device. In R. Feuerstein (Ed.), *Feuerstein's theory & applied systems: A reader* (pp. 51–97). ICELP Publications.

Fiset, D., Blais, C., Éthier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008). Features for identification of uppercase and lowercase letters. *Psychological Science*, *19*(11), 1161–1168. https://doi.org/10.1111/j.1467-9280.2008.02218.x

Forsythe, A., Mulhern, G., & Sawey, M. (2008). Confounds in pictorial sets: The role of complexity and familiarity in basic-level picture processing. *Behavior Research Methods*, *40*(1), 116–129. https://doi.org/10.3758/BRM.40.1.116

Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading*, *2*(1), 31–54.

https://doi.org/10.1207/s1532799xssr0201

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographic depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(1), 104–115.

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, *10*(2), 486–489. https://doi.org/10.5812/ijem.3505

Gibson, E. J. (1969). *Principles of perceptual learning and development*. Meredith.

Goodwin, L. D., & Leech, N. L. (2006). Understanding correlation: Factors that affect the size of r. *Journal of Experimental Education*, *74*(3), 251–266. https://doi.org/10.3200/JEXE.74.3.249-266

Gorelova, L. M. (2002). *Manchu grammar*. Brill.

Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology*, *42*, 73–83. https://doi.org/10.1016/0022-0965(86)90016-0

Graham, J. W., Cumsille, P. E., & Shevock, A. E. (2012). Methods for handling missing data. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Research methods in psychology (Vol. 2)* (2nd ed., pp. 109–141). John Wiley & Sons.

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, *8*(3), 206–213. https://doi.org/10.1007/s11121-007-0070-9

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, *11*(3), 255–274. https://doi.org/10.3102/01623737011003255

Greene, J. C., & Hall, J. N. (2010). Dialectics and pragmatism: Being of consequence. In

Abbas Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 119–144). SAGE Publications.

Guo, M. (2003). heihediqu manyu shiyong xianzhuang diaochayanjiu [Investigation on the status quo of the Manchu language use in Heihe area]. *Manchu Studies*, *17*(2), 82–88.

Guo, M. (2008). manyu binwei yuanyin tanxi [A discussion on the causes of the Manchu language endangerment]. *Manchu Studies*, *22*(2), 30–36.

He, R. (2009). *manyu 365 ju [365 sentences of Manchu]*. Liaoning Ethnic Press.

Henderson, A. R. (2005). The bootstrap: A technique for data-driven statistics. Using computer-intensive analyses to explore experimental data. *Clinica Chimica Acta*, *359*, 1–26. https://doi.org/10.1016/j.cccn.2005.04.002

Hinton, L. (2011). Revitalization of endangered languages. In P. K. Austin & J. Sallabank (Eds.), *The Cambridge handbook of endangered languages* (pp. 291–311). Cambridge University Press.

Howe, K. R. (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher*, *7*, 10–16.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, *99*(3), 480–517.

Israel, G. D. (1992). *Determinig sample size*. Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida. https://doi.org/10.4039/Ent85108-3

Jin, L., & Cortazzi, M. (1998). Dimensions of dialogue: Large classes in China. *International Journal of Educational Research*, *29*, 739–761. https://doi.org/10.1016/S0883-0355(98)00061-5

Johnson, R. B., & Christensen, L. (2014). *Educational research: Quantitative, qualitative, and mixed approaches* (5th ed.). SAGE Publications.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographis: The orthographic depth hypothesis. In L. Katz & R. Frost (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 67–84). Elsevier.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.

Lanthier, S. N., Risko, E. F., Stolz, J. A., & Besner, D. (2009). Not all visual features are created equal: Early processing in letter and word recognition. *Psychonomic Bulletin & Review*, *16*(1), 67–73.

Li, G. R. (2010). *Manchu: A textbook for reading documents* (2nd ed.). National Foreign Language Resource Center University of Hawai'i.

Liu, D., Chen, X., & Wang, Y. (2016). The impact of visual-spatial attention on reading and spelling in Chinese children. *Reading and Writing*, *29*(7), 1435–1447. https://doi.org/10.1007/s11145-016-9644-x

Liversedge, S. P., Zang, C., Zhang, M., Bai, X., Yan, G., & Drieghe, D. (2014). The effect of visual complexity and word frequency on eye movements during Chinese reading. *Visual Cognition*, *22*(3), 441–457. https://doi.org/10.1080/13506285.2014.889260

Ma, G., & Li, X. (2015). How character complexity modulates eye movement control in Chinese reading. *Reading and Writing*, *28*, 747–761. https://doi.org/10.1007/s11145-015-9548-1

Ma, T. (2014). shier zitou yu qingdai manwenyu xue [The "twelve heads" and Manchu language studies in the Qing period]. *The Qing History Journal*, (3), 1–11.

Maxwell, J., & Mittapalli, K. (2010). Realism as a stance for mixed methods. In Abbas Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 145–168). SAGE Publications.

McBride-Chang, C., & Treiman, R. (2003). Hong Kong Chinese kindergartners learn to read Englisy analytically. *Psychological Science*, *14*(2), 138–143.

McBride, C. A. (2016). Is Chinese special? Four aspects of Chinese literacy acquisition that might distinguish learning Chinese from learning alphabetic orthographies. *Educational Psychology Review*, *28*, 523–549. https://doi.org/10.1007/s10648-015-9318-2

Mertens, B. D. M., Bledsoe, K. L., Sullivan, M., & Wilson, A. (2010). Utilization of mixed methods for transformative purposes. In Abbas Tashakkori & C. Teddlie (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research* (2nd ed., pp. 193–214). SAGE Publications.

Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative Health Research*, *8*(3), 362–376.

Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, *1*(1), 48–76. https://doi.org/10.1177/2345678906292462

Mosley, C. (Ed.). (2010). *Atlas of the world's languages in danger* (3rd ed.). UNESCO Publishing. Retrieved from http://www.unesco.org/culture/en/endangeredlanguages/atlas

Nag, S. (2007). Early reading in Kannada: The pace of acquisition of orthographic knowledge and phonemic awareness. *Journal of Research in Reading*, *30*(1), 7–22. https://doi.org/10.1111/j.1467-9817.2006.00329.x

Nag, S. (2011). The akshara languages: What do they tell us about children's literacy learning? In R. Mishra & N. Srinivasan (Eds.), *Language-Cognition: State of the Art* (pp. 272–290). Lincom.

Nag, S. (2014). Akshara-phonology mappings: The common yet uncommon case of the consonant cluster. *Writing Systems Research*, *6*(1), 105–119. https://doi.org/10.1080/17586801.2013.855621

Nag, S. (2017a). Learning to read alphasyllabaries. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of Reading Development* (pp. 75–98). John Benjamins

Publishing Company. https://doi.org/10.1075/swll.15.05nag

Nag, S. (2017b). Learning to read Kannada and other languages of South Asia. In L. Verhoeven & C. Perfetti (Eds.), *Learning to read across languages and writing systems* (pp. 104–126). Cambridge University Press.

Nag, S. (2021). How children learn to use a writing system: Mapping evidence from an Indic orthography to written language in children's books. *Written Language & Literacy*, *24*(2), 284–302.

Nag, S. (2022). Reading the akshara writing system. In M. J. Snowling, C. Hulme, & K. Nation (Eds.), *The science of reading: A handbook* (2nd ed., pp. 372–389). John Wiley & Sons.

Nag, S., & Narayanan, B. (2019). Orthographic knowledge, reading and spelling development in Tamil: The first three years. In M. Joshi & C. McBride (Eds.), *Handbook of literacy in akshara orthography* (pp. 55–83). Springer. https://doi.org/10.1007/978-3-030-05977-4_4

Nag, S., & Snowling, M. J. (2012). Reading in an Alphasyllabary: Implications for a Language Universal Theory of Learning to Read. *Scientific Studies of Reading*, *16*(5), 404–423. https://doi.org/10.1080/10888438.2011.576352

Nag, S., Snowling, M. J., & Asfaha, Y. M. (2016). Classroom literacy practices in low- and middle- income countries: an interpretative synthesis of ethnographic studies. *Oxford Review of Education*, *42*(1), 36–54. https://doi.org/10.1080/03054985.2015.1135115

Nag, S., Snowling, M., Quinlan, P., & Hulme, C. (2014). Child and symbol factors in learning to read a visually complex writing system. *Scientific Studies of Reading*, *18*(5), 309–324. https://doi.org/10.1080/10888438.2014.892489

Norman, J. (2013). *A comprehensive Manchu-English dictionary*. Harvard University Asia Center.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity*, *41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

O'Connor, C., & Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods*, *19*, 1–13. https://doi.org/10.1177/1609406919899220

Onwuegbuzie, A. J., & Collins, K. M. . (2007). A typology of mixed methods sampling designs in social science research. *The Qualitative Report*, *12*(2), 281–316. Retrieved from https://nsuworks.nova.edu/tqr/vol12/iss2/9

Onwuegbuzie, Anthony J., & Teddlie, C. (2003). A framework for analyzing data in mixed methods research. In Abbas Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 351–383). Sage.

Parsons, T. D., Notebaert, A. J., Shields, E. W., & Guskiewicz, K. M. (2009). Application of reliable change indices to computerized neuropsychological measures of concussion. *International Journal of Neuroscience*, *119*(4), 492–507. https://doi.org/10.1080/00207450802330876

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*, 4646–4674. https://doi.org/10.1016/j.visres.2006.04.023

Riemer, F. J. (2012). Ethnographic research. In S. D. Lapan, M. T. Quartaroli, & F. J. Riemer (Eds.), *Qualitative research: An introduction to methods and designs* (pp. 163–188). Jossey-Bass.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581–592.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Wiley.

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association*, *91*, 473–489.

Saarela, M. S. (2015). *Manchu and the study of language in China (1607–1911)*. Princeton University.

Saarela, S. M. (2014). shier zitou jizhu (Collected notes on the twelve heads): A recently discovered work by Shen Qiliang. *Saksaha: A Journal of Manchu Studies*, *12*, 9–31. https://doi.org/10.3998/saksaha.13401746.0012.003

Sackett, P. R., & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology*, *85*(1), 112–118. https://doi.org/10.1037/0021-9010.85.1.112

Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.). Sage. https://doi.org/10.1002/9780470975220.ch1

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. Chapman & Hall.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177.

Seymour, P., Aro, M., & Erskine, J. (2003). Foundations literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–174.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3/4), 591–611.

Shen, H. H. (2005). An investigation of Chinese-character learning strategies among non-native speakers of chinese. *System*, *33*, 49–68. https://doi.org/10.1016/j.system.2004.11.001

Simpson, G. B., & Kang, H. (2004). Syllable Processing in Alphabetic Korean. *Reading and Writing: An Interdisciplinary Journal*, *17*, 137–151. https://doi.org/10.1023/b:read.0000013808.65933.a1

Sinharay, S., Stern, H. S., & Russell, D. (2001). The Use of Multiple Imputation for the Analysis of Missing Data.pdf. *Psychological Methods*, *6*(4), 317–329.

Sprenger-Charolles, L., & Siegel, L. S. (1997). A longitudinal study of the effects of syllabic structure on the development of reading and spelling skills in French. *Applied Psycholinguistics*, *18*(4), 485–505. https://doi.org/10.1017/S014271640001095X

Street, B. (2000). Literacy events and literacy practices: Theory and practice in the New Literacy Studies. In M. Martin-Jones & K. E. Jones (Eds.), *Multilingual literacy: Reading and writing different worlds* (pp. 17–29). John Benjamins Publishing Company. https://doi.org/10.1075/swll.10.08str

Szwed, M., Cohen, L., Qiao, E., & Dehaene, S. (2009). The role of invariant line junctions in object and visual word recognition. *Vision Research*, *49*(7), 718–725. https://doi.org/10.1016/j.visres.2009.01.003

Tamaoka, K., & Kiyama, S. (2013). The effects of visual complexity for Japanese kanji processing with high and low frequencies. *Reading and Writing*, *26*(2), 205–223. https://doi.org/10.1007/s11145-012-9363-x

Tashakkori, A, & Newman, I. (2010). Quantitative and qualitative approaches to research – Integration. In P. L. Peterson, E. Baker, & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., pp. 514–520). Elsevier.

Tashakkori, Abbas, & Teddlie, C. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In Abbas Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 3–50). Sage.

Teddlie, C., & Tashakkori, A. (2010). Overview of contemporary issues in mixed methods research. In Abbas Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 1–42). SAGE Publications.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, *1*(1), 77–100.

Treiman, R., & Rodriguez, K. (1999). Young children use letter names in learning to read

words. *Psychological Science*, *10*(4), 334–338.

Vaid, J., & Gupta, A. (2002). Exploring word recognition in a semi-alphabetic script: The case of Devanagari. *Brain and Language*, *81*, 679–690. https://doi.org/10.1006/brln.2001.2556

Van Buuren, S. (2018). *Flexible imputation of missing data* (2nd ed.). CRC Press.

van Ginkel, J. R., Linting, M., Rippe, R. C. A., & van der Voort, A. (2020). Rebutting Existing Misconceptions About Multiple Imputation as a Method for Handling Missing Data. *Journal of Personality Assessment*, *102*(3), 297–308. https://doi.org/10.1080/00223891.2018.1530680

Verhoeven, L., & van Leeuwe, J. (2009). Modeling the growth of word-decoding skills: Evidence from dutch. *Scientific Studies of Reading*, *13*(3), 205–223. https://doi.org/10.1080/10888430902851356

Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, S. E. (2012). *Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models* (2nd ed.). Springer.

Wang, Q. (1996). manwen yinjie zimu de yinsu fenxi [An analysis on the phonemes in the syllables of Manchu]. In Z. Wang (Ed.), *Manchu history and culture*. China Minzu University Press.

Watson, A. B. (2012). Perimetric complexity of binary digital images Notes on calculation and relation to visual complexity. *The Mathemarica Journal*, *14*, 1–41. https://doi.org/10.3888/tmj.14-5

Winskel, H. (2009). Reading in Thai: The case of misaligned vowels. *Reading and Writing: An Interdisciplinary Journal*, *22*, 1–24. https://doi.org/10.1007/s11145-007-9100-z

Wu, N., Zhou, X., & Shu, H. (1999). Sublexical processing in reading Chinese: A development study. *Language and Cognitive Processes*, *14*(5/6), 503–524.

https://doi.org/10.1080/016909699386176

Yamane, T. (1967). *Statistics: An introductory analysis* (2nd ed.). Harper and Row.

Yin, L., & McBride, C. (2015). Chinese kindergartners learn to read characters analytically. *Psychological Science*, *26*(4), 424–432. https://doi.org/10.1177/0956797614567203

Young, D. S. (2018). *Handbook of regression methods*. Chapman and Hall/CRC.

Yu, L., Zhang, Q., Priest, C., Reichle, E. D., & Sheridan, H. (2018). Character-complexity effects in Chinese reading and visual search: A comparison and theoretical implications. *Quarterly Journal of Experimental Psychology*, *71*(1), 140–151. https://doi.org/10.1080/17470218.2016.1272616

Zhao, A., Guo, M., & He, X. (2013). *binweiyuyan manyu hezheyu gongshi yanjiu [Endangered languages: A synchronic study of the Manchu and Hezhe languages]*. Social Sciences Academic Press (China).

Zhou, X., Zhou, C., Liu, D., & Ding, X. (2014). *Applied Missing Data Analysis in the Health Sciences*. Wiley.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Appendix A

Ethical Application of the Research Project

The University of Oxford places a high value on the knowledge, expertise, and integrity of its members and their ability to conduct research to high standards of scholarship and ethics. The research ethics clearance procedures have been established to ensure that the University is meeting its obligations as a responsible institution.

They start from the presumption that all members of the University will take their responsibilities and obligations seriously and will ensure that their research on human subjects is conducted according to the established principles and good practice in their fields and in accordance, where appropriate, with legal requirements. Since the requirements of research ethics review will vary from field to field and from project to project, the University accepts that different guidelines and procedures will be appropriate.

- Please check **"Where and how to apply for ethical review"** and the **CUREC flowchart** first to see if you need ethics approval.
- Please complete this form using a word processor and email it, together with your supporting documents, to your **Departmental Research Ethics Committee (DREC)** (**if applicable**). If you don't have a DREC please email this form to **ethics@socsci.ox.ac.uk** using your official **ox.ac.uk email** address. **Only emailed applications will be accepted.**

WHAT THIS CHECKLIST IS DESIGNED FOR

This **CUREC 1A checklist** is designed largely for research that falls within the Divisions of Social Sciences and Humanities where ethical issues are relatively few and straightforward. Interviews, field work and oral history are also included in the CUREC process.

The **full CUREC 2 application** is only required where certain project characteristics (e.g. type of participants, or procedures) result in a more complex set of ethical issues. It is expected that only in a limited number of cases will it be necessary for researchers to complete a CUREC 2 application. The checklist below will direct you to a CUREC 2 application if needed.

WHAT THIS CHECKLIST WILL NOT ASSESS

This checklist does not cover research governance, satisfactory methodology, or compliance with the requirements of publishers when administering their tests or questionnaires. As principal researcher, it is your responsibility to ensure that requirements in these areas are met.

CUREC does not review studies classed as **audit** (see Glossary and Decision Flowchart for CUREC on our website).

If your study involves **NHS patients, NHS staff / data / facilities, or human tissue**, please check the Decision Flowchart for NHS approval and contact the Clinical Trials and Research Governance (CTRG) team in the first instance.

Further information on the University's research ethics procedures is available from the CUREC website.

SECTION A: Filter for CUREC2 application

This section determines whether your study raises more complex issues which require the completion of a full application for ethical review, known as the CUREC 2 application.

(Please mark 'X' in the Yes/No column as appropriate to indicate your response.)

1.	Are research participants classed as **people whose ability to give free and informed consent is in question**? (This may include those under 18 (though see "competent youths"), prisoners, or adults "at risk".) Your attention is drawn to the University's **Safeguarding Code of Practice** and its implications for researchers involving children or adults at risk, including the need for the work to be risk assessed and for researchers to undertake related training. (**Note**: If any of your participants are aged 16 or under, please answer 'Yes' here **and** also answer question 5 below.)	Yes X	No
2.	By taking part in the research, will participants be at serious risk of criminal prosecution (e.g. by providing information on drug abuse or child abuse)?	Yes	No X
3.	Does the research involve the **deception** of participants?	Yes	No X
4.	Does your research raise issues relevant to the **Counter-Terrorism and Security Act (the Prevent duty)**, which seeks to prevent people from being drawn into terrorism? Please see advice on this on our **Best Practice Guidance web page**.	Yes	No X
If you have answered 'No' to all of the questions above please go to **Section B**. If you have answered 'Yes' to any question above continue to question 5 below.			
5.	Is your project covered by a CUREC **approved procedure** (formerly known as "CUREC Protocols")?	Yes X	No
If yes, please give research procedure number(s): 15, 25			

If you answered '**Yes'** to **ANY** of questions **1-4**, **and** answered '**No'** to question **5**, **please stop completing this checklist and do not submit it for ethical review**. Instead, please complete the CUREC 2 application form from the CUREC website. Then submit the CUREC 2 form for ethical review.

If you answered 'Yes' to ANY of questions 1-3, **and** answered 'Yes' to question 5, please go on to **Section B**.

SECTION B: Contact details and project description (NB: must be typed not handwritten)	
Contact details:	
1. **Principal researcher/**supervisor (title and name) (if student research):	Bai Li / Professor Victoria Murphy, Dr. Sonali Nag
2. Name of student (if student research):	Bai Li
3. Degree programme, e.g. DPhil, BA, MPhil, BSc, MSc (if student research):	DPhil
4. Department or Institute name:	Education
5. Address for correspondence (if different from above):	Wolfson College, Linton Road, Oxford, OX2 6UD
6. University e-mail (**not** private email) and telephone:	bai.li@education.ox.ac.uk, 07713 868845
7. Name and status of others taking part in the project, e.g. third year undergraduate; postdoctoral research assistant:	N/A

SECTION B continued	
Project description:	
8. Title of research project:	Complexity of the Manchu Writing System and Visual Word Recognition: A study on beginning level young learners in China
9. List of location(s) where project will be conducted:	A primary school in a Manchu Autonomous County in Northeast China
10. If your research involves overseas travel or fieldwork, by the time the research starts, will you have completed and returned a travel risk assessment form? (This has to be approved by your department before you travel. If you are travelling overseas, you are strongly advised to take out University travel insurance.)	Yes X No N/A
11. Anticipated duration of research project overall:	4 years

12. Anticipated start and end dates of the research project involving human participants:	From: (01/10/2017) To: (01/10/2021) Please note that you will need ethics approval **before** you start your research. CUREC1As may take up to 30 days to process.
13. External organisation funding the research (if applicable): China Scholarship Council	
14. Title and very brief and simple lay description of **research** (about 150 words), plus description (about 200 words) of the nature of participants.	
a) Title, brief description of research (150 words) in lay language. When describing the research, please include your methodology, how you are applying professional guidelines, and the use to which results/data will be put. **Please also declare any conflicts of interest here.**	

Complexity of the Manchu Writing System and Visual Word Recognition: A study on beginning level young learners in China

This study investigates Manchu word recognition at a primary school located in a Manchu Autonomous County in Northeast China. This local school provides Manchu classes as a subject to children in Grades 3-5 (9- to 11-year-olds) as a way to preserve and maintain the critically endangered Manchu language, which currently has an estimate of only 20-100 native speakers. This project thus intends to investigate in what ways the complexity of the Manchu writing system influences visual word recognition by the children in this primary school setting.

Methodologically, both quantitative and qualitative methods will be applied for data collection. To collect quantitative data, the participating children will be tested for their non-verbal IQ, Mandarin vocabulary proficiency, Manchu/English/Mandarin items speeded pronunciation, phonological awareness, and attitude towards Manchu. Qualitatively, the Manchu teaching and learning in the classroom will be observed, and semi-structured interview will be conducted either individually with participating children, teaching staff, and parents, or with a focus group. In addition, questionnaire will be used to collect some basic information of the participants (e.g., the age, ethnic identity, attitude towards Mandarin, English, and Manchu, etc.).

Specifically, the tests will be given in two separate sessions on different days, with each session being approximately 45 minutes:

1) Session One
    - Questionnaire of basic personal information and attitude towards Manchu language

- Non-verbal IQ test
- Mandarin vocabulary proficiency test

2) Session Two
   - Speeded pronunciation of Manchu, English, and Mandarin words
   - Phonological awareness test (approximately 20 minutes)

BERA guidelines will apply.

There is no conflict of interest.

SECTION B continued

b) Description of participants and **obtaining informed consent** (200 words). When describing participants, please include

- criteria for inclusion/exclusion
- method of recruitment
- processes for consent to participate

Please ensure you attach as separate documents (if applicable, in English translation):

- your recruitment and advertisement material e.g. a poster or brief invitation letter/ email
- information for participants to read (or hear) before they agree to take part e.g. written information sheets or (only if applicable) oral information scripts.
- a document to record informed consent. Templates for written consent forms and/or oral information scripts (in case of an oral consent process) are available from the CUREC website
- a guide to interview questions (this may be a list of questions to be asked, or a preliminary scope of questions), or a sample of other instruments (such as a sample questionnaire)
- (if relevant) debriefing document after participants have taken part

Recruitment will primarily be through the primary school located at the Manchu Autonomous County. The intended participants are the school children, the teaching staff, and the children's parents. Three classes of children (9- to 11-year-olds) in Grades 3, 4, and 5 respectively will be invited for this project using convenience sampling. The potential participating parents will be introduced by the class teachers to the researcher and then invited by the latter.

The school's agreement to participate will be obtained first. Then, consent for participating in the project will be in two strands. For the adults (teaching staff and parents), it will be through the written information sheet and written consent form. The purpose of the study will be explicitly stated and their right to withdraw without any consequences will be stated clearly too, under Procedure 15. For the children, they will be offered the information sheet. Since they are under the age of 13 and personal data will be obtained and processed in the course of the study, their parents or guardians will be given information sheets and opt-in consent form to review, under Procedure 25. Also, the children will be asked to give their assent to participate and any child who does not give their assent, even if their parents agreed, will not be in the study. The study purpose and the children's right to withdraw at any time will also be stated explicitly.

For the interview with the potential participants (including teaching staff, children, and parents), questions primarily concern their Manchu teaching and learning experiences, their attitudes towards Manchu, Mandarin, and English, etc. All the interviews with the children will be conducted at the participating school to ensure their safety. For the adult participants, the location of interviewing will be discussed before conducted.

15. What are the ethical issues connected with your research and what steps have you taken to address them? Please do not answer 'none'. The committee needs to see evidence that you have identified potential ethical issues with respect to your research and have taken steps to address them. These issues could relate to:

- your own physical and psychological safety as a researcher (please see the University's and Social Science Division's Safety in Fieldwork guidance
- participant burdens and/or risks, and
- data protection/ confidentiality.

For more guidance on ethical issues, please see http://researchsupport.admin.ox.ac.uk/governance/ethics/resources

Semi-structured interview with parents, school teaching staff, and children will be undertaken.

For parents, the main ethnical issue is the time commitment they will make participating in interviews, which will be resolved by the consent made between interviewer and interviewee in terms of the amount of time they would like to spend on interviews. In addition, what the interviewees said will be kept anonymous.

The main ethical issue for the teachers and school children are as follows:

1. Pressure derived from the teaching staff's expression of own opinions on Manchu teaching and learning in the school, which may conflict with the ideas of the school authorities. This issue will be resolved through clear consent and anonymity;
2. Ensuring appropriate consent, without making the process itself burdensome, thus those who are 16 years old and over are considered competent, and opt-out is used for those under the age of 16;
3. Ethnic identity of all the participants will be included in the questionnaire for descriptive purpose. This will not bring up ethical concern;
4. Pressure for children taking the tests of non-verbal IQ, vocabulary proficiency, speeded pronunciation, and phonemic awareness and taking part in the interviews. The researcher will treat the test scores confidentially to minimize their children's pressure. Burdens on participants will be minimised by ensuring that interview and focus group schedules are kept short. Focus group, if any, will provide opportunities for them to share their views, reducing the likelihood of distress or discomfort. Audio recording will be used only with the explicit agreement of the participants.
5. To prevent the fatigue that the children may feel, the tests will be conducted in two separate sessions of 45 minutes each, to be administered on different days.

Section B continued		
16. Will you obtain <u>informed consent</u> according to CUREC guidelines and good practice in your discipline before participation?	Yes X	No
If you have marked '**No**', please give a brief explanation and justification for this decision here:		

17. Will your research involve discussing sensitive issues? This could be information relating to race or ethnic origin, political opinions, religious beliefs, physical/mental health, trade union membership, sexual life or criminal activities.	Yes	No X

If you have marked '**Yes**', please make sure that you have included some **supporting information** (as directed in question 14 of this section) showing the range of questions covering these issues.

18. Will you ensure that **personal data** collected directly from participants or via a **third party** is held and processed in accordance with the provisions of the Data Protection Act?	Yes X	No

19. How will you ensure that any personal and/or sensitive data are captured, transferred and stored securely?

In particular if data are to be captured electronically, please consult with the University's research data team (**researchdata@ox.ac.uk**) and your local IT department and, with respect to University IT security policies, please comment on how you will capture such data in the first instance, how you will transfer them over networks or via portable media and how, where and how long data will be stored. For more information please see the University's web pages on research data management.

All possible steps will be taken to ensure confidentiality and anonymity for participants. Nothing will be said or written that enables any school or individual to be identified. Participants will be reassured that data collected are for research purposes only. All data will be kept confidential including storing audio recordings on a secure university server; labelling participants with anonymised codes which are stored separately to other identifying details; and removing potentially identifying details in publicly disseminated materials (thus complying with the DPA). The data will be stored for as long as they are of academic value.

SECTION C: Methods and procedures to be used

Method used: Please ensure you have addressed any potential ethical issues related to these methods in Section 14 and in your Participant Information Sheet	**Please mark 'X'**
1. Analysis of existing records	
2. Snowball sampling (recruiting through contacts of existing participants)	
3. Use of casual or local workers e.g. interpreters	

4.	Participant observation	X
5.	Covert observation	
6.	Observation of specific organisational practices	
7.	Participant completes questionnaire in hard copy	X
8.	Participant completes online questionnaire or other online task	
9.	Using social media	
10.	Participant performs paper and pencil task	X
11.	Participant performs verbal or aural task (e.g. for linguistic study)	X
12.	Focus group	X
13.	Interview	X
14.	Audio recording of participant (you will generally need specific consent from participants for this)	X
15.	Video recording of participant (you will generally need specific consent from participants for this)	
16.	Photography of participant (you will generally need specific consent from participants for this)	
17.	Others (please specify):	

SECTION D: Professional guidelines and training		
In this section, please mark 'X' against at least one of the following professional guidelines you aim to adhere to. You should use the principles listed in your chosen guideline(s) in conducting your own research. **Note**: this is not an exhaustive list.		**Please mark 'X'**
Research specialism/ methodology	**Association and guidance document**	
Anthropology	Association of Social Anthropologists of the UK and Commonwealth	
Criminology	British Society of Criminology: Code of Ethics for Researchers in the Field of Criminology	
Education	British Educational Research Association Ethical Guidelines for Educational Research	X
Geography	Association of American Geographers Statement on Professional Ethics Royal Geographical Society: Research Ethics and Code of Practice	
History	Oral History Society of the UK Ethical Guidelines	
Internet-based Research	British Psychological Society: Conducting Research on the Internet Association of Internet Researchers Ethics Guide Also see our Best Practice Guidance on internet-based research	
Law (Socio-Legal)	Socio-Legal Studies Association: Statement of Principles of Ethical Research	
Management	Academy of Management's Professional Code of Ethics	
Political Science	American Political Science Association (APSA) Guide to Professional Ethics in Political Science	
Politics	Political Studies Association. Guidelines for Good Professional Conduct	
Psychology	British Psychological Society Code of Ethics and Conduct British Psychological Society: Conducting Research on the Internet Also see "Internet-based Research" guidance above	
Social Research	Social Research Association: Ethical Guidelines	
Sociology	The British Sociological Association: Statement of Ethical Practice	
Visual Research	ESRC National Centre for Research Methods Review Paper: Visual Ethics: Ethical Issues in Visual Research	
Other professional guidelines. Please specify the other guidelines used here*:*		

Please indicate what training in research ethics the researchers involved with this study have received, e.g. the title of the course and date completed (online training available at http://researchsupport.admin.ox.ac.uk/support/training/ethics). If no formal training has been undertaken, please indicate any discussions of research methodology between researchers and supervisors here.
Mark has received training in research ethics at the Department of Education, University of Oxford. He previously applied for and was granted research ethics for his study of Manchu learning in university context in China.

SECTION E: Signatures

- 'Electronic signatures' sent as email confirmations from a University of Oxford email address can be accepted. Separate emails should come from each of the relevant signatories as outlined below, indicating acceptance of the relevant responsibilities.
- If you have obtained handwritten (wet-ink) signatures, please scan them and the rest of the checklist pages to create a single PDF document and email through.

Please ensure this checklist is signed by:

For staff research:	**For student research**:
1. **Principal researcher**	1. **Principal researcher (project supervisor)**
2. **Head of Department (or nominee)**	2. **Head of Department (or nominee)**
	3. **Student researcher**

1. **Principal researcher signature/supervisor signature (if student research)**

I understand my responsibilities as **principal researcher** as outlined in the CUREC glossary and guidance on the CUREC website.

I declare that the answers above accurately describe the research as presently designed, and that a new checklist will be submitted should the research design change in a way which would alter any of the above responses so as to require completion of CUREC 2 (involving full scrutiny by an IDREC). I will inform the relevant IDREC if I cease to be the principal researcher on this project and supply the name and contact details of my successor if appropriate.

Signature: ……..

Date: 27.08.18

Print name (block capitals): Victoria Murphy

2. **Departmental endorsement signature**

I have read the research project application named above. On the basis of the information available to me, I:

(i) consider the principal researcher to be aware of her/his ethical responsibilities in regard to this research;
(ii) consider that any ethical issues raised have been satisfactorily resolved or are covered by relevant professional guidelines and/or CUREC approved procedures, and that it is appropriate for the research to proceed (noting the principal researcher's obligation to report should the design of the research change in a way which would alter any of the above responses so as to require completion of a CUREC 2 full application);
(iii) am satisfied that: the proposed project design and scientific methodology is sound; the project has been/will be subject to appropriate **peer review**; and is likely to contribute to existing knowledge and/or to the education and training of the researcher(s) and that it is in the **public interest**.

Signed by Head of Department or nominee *(example nominees for student research include the Director of Graduate Studies/ Director of Undergraduate Studies)*:

Signature: ……..

Print name (block capitals): ……..... **Date**: ……..

3. **Student signature (if student research)**

I understand the questions and answers that have been entered above describing the research, and I will ensure that my practice in this research complies with these answers, subject to any modifications made by the principal researcher properly authorised by the CUREC system.

Signed by student: ……Bai Li…….

Date: ……26 July 2018……..

Print name (block capitals): ……BAI LI……..

SECTION F: SUBMITTING THE COMPLETED CHECKLIST	Please mark 'X'
	X
1. Check you have completed all sections (A-E)	X
2. Ensure your application is signed by you, your supervisor (if student) and department	X
3. **Please attach all supporting documents (see section B, question 14b for details).** If the appropriate supporting documentation is not included with your application, you will then be asked to provide this separately. **This may well delay the ethical review process, and thus the start of your research.**	X X X
4. Ensure you have declared conflicts of interest (if any) in Section B, question 14a.	X
5. If your department has a Departmental Research Ethics Committee (DREC), submit this checklist and supporting information to the appropriate departmental officer.	
6. If your department does not have a DREC, submit the checklist and supporting information to the SSH IDREC (email ethics@socsci.ox.ac.uk).	
7. Applications must be sent by email from your official ox.ac.uk email account. Please do not send applications by post.	

CUREC Documents List

1. Letter to headmaster
2. Parent/guardian consent form
3. Information sheet for children 6-10
4. Information sheet for children 11-15
5. Information sheet for parent/guardian
6. Semi-structured interview guidelines
7. Questionnaire to children

University of Oxford

Department of Education

15 Norham Gardens, Oxford OX2 6PY

Director: Professor Jo-Anne Baird

Tel: +44 (0)1865 274002 Fax: +44 (0)1865 274027

[Head teacher name]

Yongling Manchu Primary School

1 September 2018

Dear [Head teacher name],

I am writing to enquire about conducting some research in your school in the next academic year. I am a doctoral research student at the University of Oxford, supervised by Professor Victoria Murphy and Dr Sonali Nag. In my research project, *Complexity of the Manchu Writing System and Visual Word Recognition: A study on beginning level young learners in China*, I will explore in what ways the Manchu writing system influences visual word recognition by the children in a primary school setting.

The research will take place in the Manchu class, with the target grades of 3, 4 and 5. I am looking for teacher who gives the Manchu class in your school. I am not aiming to change what or how the teacher chooses to teach and will not be making any judgements about teaching. My research focus is on students' Manchu word recognition and teacher's teaching practices in the classroom.

By participating in the research, your school would be contributing to a project that will deepen our understanding of how children's word recognition is related to the complexity of Manchu writing system, and so contribute towards developing ways of improving attainment for similar students in the future.

The commitment from the school would be to allow me into Manchu lessons over the course of one academic year (i.e., two terms), from the beginning of the Autumn Term (1 September 2018). I would observe the teaching and learning practices in the classroom and take notes, give and audio-record some tests administered to the students, interview and audio-record teachers, students, and parents with regards to the practice of Manchu teaching and learning and attitudes towards Manchu. Any collected data will not be seen by anyone else or used for research without further explicit permission.

The University of Oxford has strict ethical procedures on conducting ethical research with teachers and young people, consistent with current British Educational Research Association guidelines. Before beginning the research, I would inform parents and guardians about the research and offer the students, parents and guardians the opportunity to refuse to participate. Throughout the research students, parents and guardians will be able to refuse to participate at any time.

All participants, including students, teacher, parents, and the school, would be made anonymous in all research reports. The data collected would be kept strictly confidential, available only to my supervisors and myself, and not used other than specified without the further consent of all involved being obtained. All audio recording would be stored for as long as they are of academic value and kept in a secure university server.

If you feel you would like to take part in the study, or need more information about what is involved, please contact me. Whether or not you feel it would be appropriate for your school to participate, I would be grateful if you would complete the pro-

form below and return it to me in the stamped addressed envelope enclosed in this letter.

Thank you for your time and attention. I look forward to hearing from you.

Yours sincerely,

Bai Li

Complexity of the Manchu Writing System and Visual Word Recognition: A study on beginning level young learners in China

Bai Li
University of Oxford, Department of Education

Yongling Manchu Primary School

[School address]

[Head teacher name]

☐ We do not wish to participate in this project.

☐ We would like to find out more about this project.

☐ We would like to take part in this project.

If you would like further information, or are interested in taking part, please give the name of a contact person for your school, and details of the best way to contact him or her.

Contact name: _____

Contact email: _____

Contact telephone number: _____

Please return this form in the stamped addressed envelope enclosed with this letter.

Thank you for your help.

Department of Education

University of Oxford

15 Norham Gardens

Oxford OX2 6PY

phone +44 (0)1865 274002

fax +44 (0)1865 274027

Director: Professor Jo-Anne Baird

Project: Complexity of the Manchu Writing System and Visual Word Recognition: A study on beginning level young learners in China

Researcher: Bai Li

Parent/guardian consent form

Please circle as appropriate

- ❑ I have read the information contained in the leaflet
- ❑ I would like more information about the research, please contact me (give phone or email)
- ❑ I want my child to take part in this research
- ❑ I understand I can change my mind later

Child's name (block capitals) ……

Parent/guardian name (block capitals) ……

Parent/Guardian signature: ……Date: ……

Researcher name: ……Bai Li……

Researcher signature……Date: ……

Thank you very much for your help.

Please return this form to your child's Manchu teacher.

Department of Education

University of Oxford



Bai Li

Department of Education, University of Oxford

15 Norham Gardens, Oxford, OX2 6PY, UK

Direct Line: 01865 274024

E-Mail: bai.li@education.ox.ac.uk

http://www.education.ox.ac.uk

Complexity of the Manchu Writing System and Visual Word Recognition: A study on beginning level young learners in China

INFORMATION SHEET FOR CHILDREN AGED 6 TO 10

To be shown and read by parent/carer if required

Ethics Approval Reference: [Insert]

My name is Bai Li and I work at the University of Oxford. I am doing some research and would like you to join in.

Research is a way we try to find out the answers to questions. I want to see how children recognize Manchu words at school.

You can talk to your family, friends, or the researchers if you want to before you agree to join in.

Why have I been asked?

I am asking you if you would like to take part because you are between 9 and 10 years old.

I am asking all the children in your class to help me.

Your parent/carer has said it is OK for you to join in.

Do I have to join in?

No, you don't have to if you don't want to! You can ask questions before choosing whether you want to join in.

You can change your mind at any time by telling the researcher or your parent/carer. You don't have to say why.

If you decide to stop, no one will be upset with you.

What will happen?

You will be asked to fill in a form with some basic questions about yourself (around 10 minutes);

You will be invited to play a puzzle game [for IQ test] (around 20 minutes);

You will be invited to do a task to show how many Chinese words you can recognize (around 20 minutes);

You will be invited to read aloud some Manchu, English, and Mandarin words (around 20 minutes);

You will be invited to play a game with me focusing on the sounds of words in Manchu, English, and Mandarin (around 20 minutes);

You will be invited to tell me your ideas about learning Manchu at school (around 20 minutes);

You will not miss any school for taking part in all of above task;

All the above tasks and games will be done at your school;

I love Manchu language, so I will also be learning Manchu with you in class.

Will anything about the research upset me?

Your performance during the tasks and games will will not be shown to anyone but my supervisors and me and it will not have any bearing on your education. So, don't worry and take it easy!

Will joining in help me?

The study will not help you but it might help other children in the future.

Will anyone else know I'm doing this?

The people in our research team will know you are taking part. No one will know that you have helped us with this research - unless, of course, you tell them yourself!

What happens to what the researchers find out?

When I collect information from you I will keep it in a safe place and only the people doing the research, or helping with the research, can look at it. We will use the information to write a really long paper, thanks to your help!

Is this study OK to do?

Before any research involving people happens it has to be checked by a group of people known as a Research Ethics Committee to make sure that it is fair. This study has been checked by the Ethics Committee at the University of Oxford.

What if there is a problem or something goes wrong?

If you are not happy because of something that happened in the study, please talk to your parent/carer who will let the researcher know.

Thank you for reading – please ask us any questions.

Department of Education

University of Oxford



Bai Li

Department of Education, University of Oxford

15 Norham Gardens, Oxford, OX2 6PY, UK

Direct Line: 01865 274024

E-Mail: bai.li@education.ox.ac.uk

http://www.education.ox.ac.uk

Complexity of the Manchu Writing System and Visual Word Recognition: A study on beginning level young learners in China

INFORMATION SHEET FOR YOUNG PEOPLE AGED 11 TO 15

Ethics Approval Reference: [Insert]

I am inviting you to join in a research study. My name is Bai Li and I work at the University of Oxford in the Department of Education.

Before you decide if you would like to join in, it's important to understand what the study is about, why we're doing it and what it would involve for you. Please read and think about this leaflet carefully. Please feel free to talk to your family, friends, or the researchers about it if you want.

If anything isn't clear or you have more questions you can ask your parents/carer to give us a call and we can discuss it with you and your parents/carer.

Why are we doing this research?

I am doing this research to understand how Chinese children recognise Manchu words at school, like what they found easy or difficulty to read recognise Manchu words.

Why have I been invited to take part?

We are inviting you to take part because you are a young person, aged between 11 and 12 years. We are inviting all the young people in your class to take part.

Do I have to take part?

No - It is up to you. We will give you a copy of this information sheet and your parents and guardians will decide whether you will take part. You are free to stop taking part at any time during the research without giving a reason, by telling the researcher. If you decide to stop, we will use the information we have already collected from you.

What will happen to me if I take part?

You will be asked to fill in a form with some basic questions about yourself (around 10 minutes);
You will be invited to play a puzzle game [for IQ test] (around 20 minutes);
You will be invited to do a task to show how many Chinese words you can recognize (around 20 minutes);
You will be invited to read aloud some Manchu, English, and Mandarin words (around 20 minutes);
You will be invited to play a game with me focusing on the sounds of words in Manchu, English, and Mandarin (around 20 minutes);
You will be invited to tell me your ideas about learning Manchu at school (around 20 minutes);
You will not miss any school for taking part in all of above task;
All the above tasks and games will be done at your school;
I love Manchu language, so I will also be learning Manchu with you in class.

What happens to the results of the study?

Results for each young person taking part are kept strictly confidential. Only a number will be used to identify you, and all information and results are kept in the server in the University.

If you agree to participate in this study, the research will be written up as part of my doctoral thesis.

We also aim to publish our findings in scientific journals, but this may take some time from the end of the study.

All research data and records will be stored for a minimum retention period of 5 years after publication or public release of the work of the research. Third parties may be given access to research data for monitoring and/or audit of the study, or for data storage purposes.

What are the advantages / disadvantages of taking part?

There is no risk of your taking part.

By participating in the research, you would be contributing to a project that will deepen our understanding of how children's word recognition is related to the complexity of Manchu writing system, and so contribute towards developing ways of improving attainment for similar students in the future.

Will anyone else know I'm doing this?

We will keep your information private. This means we will only tell those who have a need or right to know, such as the research team and your parents/carer. Wherever possible, we will only share information that has your name and address removed.

What if I don't want to take part in the research anymore?

Just tell your parents/carer and the people carrying out the research that you don't want to take part. You don't have to give a reason and no one will be annoyed with you. It is YOUR choice.

Who is organising and funding the research?

The research is organised by Bai Li of University of Oxford, who is a DPhil student. The study is being supported by the China Scholarship Council.

Who has reviewed the study?

Before any research involving people can start, it has to be checked by a Research Ethics Committee to make sure that it is OK for the research to go ahead. This study has been approved by the University of Oxford Central University Research Ethics Committee.

What if there is a problem or something goes wrong?

Please tell us if you are worried about any part of this study, by contacting the researcher (bai.li@education.ox.ac.uk). You may also talk to your teacher/parent/carer who will let the researcher know. If you are still unhappy or wish to make a complaint, either you or your teacher/parent(s)/carer(s) can contact the chair of the Research Ethics Committee at the University of Oxford (Chair, Social Sciences) Inter-Divisional Research Ethics Committee; Email:

ethics@socsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD).

Contact details

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

Bai Li
Department of Education, University of Oxford
15 Norham Gardens, Oxford, OX2 6PY, UK
Direct Line: 01865 274024
E-Mail: bai.li@education.ox.ac.uk

Thank you for reading this – please ask any questions if you need to.

Department of Education

University of Oxford



Bai Li

Department of Education, University of Oxford

15 Norham Gardens, Oxford, OX2 6PY, UK

Direct Line: 01865 274024

E-Mail: bai.li@education.ox.ac.uk

http://www.education.ox.ac.uk

Complexity of the Manchu Writing System and Visual Word Recognition: A study on beginning level young learners in China

INFORMATION SHEET FOR PARENTS / GUARDIANS

Ethics Approval Reference: [Insert]

In partnership with researchers at the University of Oxford, Yongling Manchu Primary School has agreed to take part in a study investigating in what ways children's Manchu word recognition is influenced by the complexity of Manchu writing system. We would like to invite your child to be part of this study. We very much hope you would like your child to take part, but before you decide, it is important that you understand why the study is being done and what it will involve.

What are we trying to find out?

This project is looking at how children recognise Manchu words. We are especially interested in how children's visual word recognition is affected by the complexity of Manchu writing system. Manchu is identified by UNESCO as a "critically endangered language" in China and thus it is facing the serious crisis of extinction. However, as a method to preserve Manchu, Manchu classes are currently offered by a number of schools in Northeast China to the students. Unfortunately, we know very little about

how children learn this language. Thus, this project will help us to understand how children recognise Manchu words. We hope that our findings will allow us to identify ways of helping children learn Manchu more effectively.

More information about the project can be obtained by contacting me via email bai.li@education.ox.ac.uk.

Why has my child been invited to take part?

We are inviting your child to take part because they are a young person, aged between [9 and 11 years], attending Yongling Manchu Primary School.
We are inviting all the students in Grades 3, 4, and 5 to take part.

Does my child have to take part?

No. You can ask questions about the study before deciding whether or not to allow your child to participate. If you do agree to participation, you may withdraw your child [and their data] from the study at any time, without giving a reason and without penalty, by advising the researchers of this decision.

What will happen if my child takes part?

- Your child will be involved in the research for two terms;
- The research will start on 1 September 2018 and goes on to 1 July 2019. But if it takes longer, you child will be invited again for participation;
- Your child will be observed during Manchu class, will fill in a basic information questionnaire (e.g., name, gender, age, etc.), will be given IQ and language ability tests, and will be interviewed. All research will take place at Yongling Manchu Primary School;
- For the classroom observation, your child will do what she or he typically does during a Manchu class; nothing will be changed. His or her learning process of Manchu will be observed and noted down;
- Your child will take tests in two separate sessions, with each session being around 45 minutes. In session one, they will fill in a questionnaire of basic information and their attitudes toward Manchu language. They will also be given an IQ test and a Mandarin vocabulary proficiency test. In session two, they will be asked to read some written words in Manchu, English, and Mandarin, and will play a game focusing on the sounds of the above languages;

- Your child will not have to do anything about the research at home;
- Your child may be invited for some follow-up research after the current project is completed.

What are the advantages / disadvantages of taking part?

There will be no financial compensation, intended benefit, or risk for your child to take part in the research. But participating in the research will give your child the opportunity to understand his or her language ability.

What happens to the results of the study?

Results for each child are kept strictly confidential. Only a number or a pseudo name will be used to identify children, and all information and results are kept in the online servers in the University. Third parties may be given access to research data for monitoring and/or audit of the study, or for data storage purposes. Summaries of our findings will be given to the school and will be available to interested families.

If you agree to participate in this project, the research will be written up as a thesis. On successful submission of the thesis, it will be deposited both in print and online in the University archives, to facilitate its use in future research. The thesis will probably be published with open access.

We also aim to publish our findings in scientific journals, but this may take some time from the end of the study.

All research data and records will be stored for a minimum retention period of 5 years after publication or public release of the work of the research.

What will happen to any samples taken from my child?

We may take some written work of Manchu from your child as samples. We will use your child's samples as data for analysis.

Who is conducting this research?

The research project is organized by Bai Li of University of Oxford, who is a DPhil student. He is being supervised by Professor Victoria Murphy (victoria.murphy@education.ox.ac.uk) and Dr Sonali Nag (sonali.nag@education.ox.ac.uk) The research is funded by the China Scholarship Council. This study

has been reviewed by, and received ethics clearance through, the University of Oxford's Central University Research Ethics Committee, [insert CUREC reference number].

What if there is a problem?

If you have a concern about any aspect of this project, please speak to the relevant researcher (Bai Li, Tel： 07713868845, Email: bai.li@education.ox.ac.uk) who will do his best to answer your query. The researcher should acknowledge your concern within 10 working days and give you an indication of how he intends to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter in a reasonably expeditious manner.

Chair, **Social Sciences & Humanities Inter-Divisional Research Ethics Committee**; Email: ethics@socsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

What should I do next?

Please fill in the enclosed form and return it to your child's class teacher if you would not like your child to take part in this study. Please remember that you may withdraw your child at any time, without penalty and without giving a reason, by notifying the researcher. If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

Bai Li
Department of Education, University of Oxford
15 Norham Gardens, Oxford, OX2 6PY, UK
Tel: 01865 274024
Email: bai.li@education.ox.ac.uk

Semi-structured interview guidelines with teachers

I. Teaching experience

1. How long have you been working in this school?
2. How long have you been teaching Manchu?
3. How many hours of Manchu class do you usually give each week?

II. Manchu textbook & curriculum

4. What textbook is in use for the Manchu class right now?
5. Why was this textbook chosen?
6. Do children have access to other printed Manchu books besides the textbook?
7. Is there any curriculum developed for the Manchu class?

III. Manchu script

8. Which script, Romanised Manchu or traditional Manchu, was used in teaching Manchu? Why was it/were they chosen?
9. Did the traditional Manchu script pose any difficulties for the student's word recognition? In what ways?
10. What are the shared difficulties among children in recognising Manchu letter and words?

IV. Attitudes toward Manchu

11. Are you ethnically Manchu?
12. Do you like teaching Manchu?
13. Do you think it is important for children to learn Manchu? Why?
14. Do you think Manchu education at primary school would help preserve the language?
15. Do you think Manchu is still useful in China today? If yes, in what ways? If no, why?

Questionnaire to children

Part I Personal Information

1. Name:
2. Name in *pinyin*:
3. Date of birth: (MM/YYYY)
4. Gender: ☐ Male ☐ Female
5. Ethnicity: ☐ Han ☐ Manchu ☐ Mongolian
   ☐ Other (Please specify_____)
6. Grade: ☐ Grade 3 ☐ Grade 4 ☐ Grade 5
7. Languages: ☐ Mandarin ☐ Manchu ☐ English ☐ Japanese
   ☐ Other (Please specify_____)

Part II Attitudes towards learning Manchu[18]

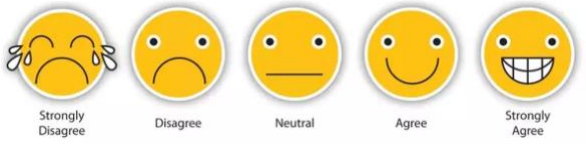
1. Learning Manchu is really great.



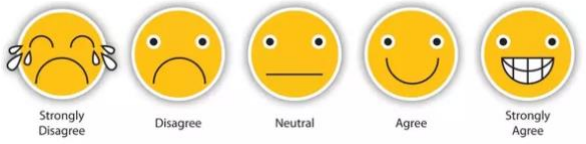
2. I plan to learn as much Manchu as possible.



3. I hate Manchu.



4. French is an important part of the school programme.



[18] Adapted from Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. Arnold.

5. I think that learning Manchu is dull.



6. I really enjoy learning Manchu.



7. I would rather spend my time on subjects other than Manchu.



8. Learning Manchu is a waste of time.



9. I love learning Manchu.



10. When I leave school, I shall give up the study of Manchu entirely because I am not interested in it.



Thank you!



Email from Dr Liam Gearon confirming the change of project participants

Dear Mark,

Thank you – we can note the change, please keep this as a record. I am happy if your supervisor is?

Kindest

Liam

From: Bai Li
Sent: 13 February 2019 15:02
To: Liam Gearon
Subject: Inquiry about CUREC approval

Dear Dr Gearon,

I hope this email finds you will. My name is Bai Li and I am a second year DPhil student in the Department of Education. I am writing this email to see if I could inquire about my approved CUREC.

I was granted a clearance about my DPhil project on Manchu language education, with the reference No. ED-CIA-18-243 (please see the email sent from Dr Heath Rose below). The potential participants in my CUREC application were children in primary school in China. However, during a pilot study that I have made to the primary school (after I was granted the CUREC approval) I found that the Manchu language proficiency of the children were so low that it was basically impossible to collect valid data from them. Thus, I had to change the potential participants for my main study from children to university students in China, and this is the only change made in my CUREC. In this case, could you please confirm whether it is acceptable to use the CUREC already granted to me for my main study?

Thank you very much. I look forward to your reply.

Best regards,
Mark

Appendix B

Report on Pilot Study

During September and October 2018, following the suggestion of my supervisors I conducted a pilot study in China (Departmental Research Ethics Committee reference number ED-CIA-18-243), during which I paid visits to three educational institutions – a primary school located in the Xinbin Manchu Autonomous County, a middle school in the City of Shenyang, and a university in the City of Chengde. This report presents and analyses the data collected from the pilot study and discusses the implication it brings for my main study.

1. Aim of study

The aim of the pilot study was threefold. Firstly, it attempts to conduct classroom observation and interviews with teachers to understand the status quo of Manchu education in China. Moreover, it plans to test the validity of the tasks used to collect data of the subjects' literacy skills in Manchu as well as in Mandarin and English. Finally, it also aims to test a questionnaire assessing children's attitude towards Manchu language learning.

2. Site identification

Since my DPhil project intends to investigate Manchu vocabulary decoding by Chinese children, I contacted two informants in China to see whether they were able to introduce me into a primary or a middle school for a pilot first. One informant, Mr Aisin,[19] an activist in Manchu language and culture preservation, recommended a primary school in the Xinbin Manchu Autonomous County in Northeast China,[20] which, according to him, provides the *best* Manchu teaching in that region. This claim was further supported by a teacher in Manchu at the school, Ms Guang, describing the Manchu classes offered to students in Grades 3, 4 and 5 as regular and well-structured. In addition, I, too, assumed that Manchu language may be well preserved in that school, not only because of its location in the county

[19] All names in this report are pseudonyms.

[20] Northeast China, an area sometimes referred to as "Manchuria", consists of the three provinces of Liaoning, Jilin, and Heilongjiang.

where the historical Manchu people originated, but also the large population of ethnic Manchus there.[21]

The other informant, a professor with the expertise in Manchu and Xibo studies in a university, offered to introduce me into a K-12 Manchu School (hereafter as "the middle school") in the City of Shenyang because of her connections with the headmaster there.

As for the university, it was not identified until after my visits to the above two sites, where I found that the Manchu literacy skill of the participating children is significantly low. Therefore, I had to reach out to Mr Sure, a lecturer in Manchu Studies in a university in Chengde, to see whether he could give me any help. Fortunately, he told me that a population of around 1,600 first year students were taking Manchu class in the university and I was welcomed to go for a visit; thus, the university was identified as the third site.

Figure 1

Location of the Three Pilot Sites



[21] Over seventy percent of the population in the Xinbin Manchu Autonomous County is registered as ethnic Manchus. Brief introduction to Xinbin. (2015, March 4). Retrieved from http://www.xinbin.gov.cn/ins.asp?t=1&s=1&i=2358.

3. Participants

A total number of 64 participants were recruited for the pilot study, all of whom were randomly selected by their class teachers. Table 1 gives detailed information of the participants, including their year of study, weekly Manchu class hours, weeks for Manchu course per semester, and course objectives.

Table 1

Participant Demographic Information

	Primary school		Middle school	University	
Year of study	Grade 4 students (N=20)	Grade 5 students (N=10)	Grade 8 students (N=10)	Year 1 students (N=15)	Year2 students (N=9)
Weekly Manchu class hours	45 mins	45 mins	45 mins	90 mins	90 mins
Scheduled weeks for Manchu course per semester	10-12 weeks	10-12 weeks	N/A [a]	13 weeks	13 weeks
Implemented weeks for Manchu course per semester	6 weeks**	6 weeks**	N/A [a]	12-13 weeks	12-13 weeks
Course objectives [b]	Romanised Manchu letters and vocabulary	Manchu ujus and vocabulary	N/A [a]	Manchu ujus and vocabulary	N/A***

[a] The information is missing because the school Manchu teacher was in lactation and was not at school during the pilot. ** This is an estimated figure based on my interviews with the children during tests. It was not verified by the Manchu teacher because of her reluctance to take part in the pilot. ***This information was not collected during the pilot and was thus missed.

[b] The objectives listed under Grades 4 and 5 participants are what they have learned a year ago when they were in Grade 3 and Grade 4 respectively, whereas the objective for the Year 1 university participants is what they were supposed to learn in the semester when I visited them.

For the middle school participants (N=10) in Grade 8, no Manchu class was offered to them during the past three years—the latest Manchu class was given to them when they were in Grade 5. The Year 1 university participants (N=15), along with all other first-year peers (approximately 1,600), are required to take Manchu as a mandatory course in the first

semester. The Year 2 participants (N=9), all majoring in Chinese Literature, are taking Manchu in the Experimental Manchu Programme, an initiative programme offered by the staff in the Manchu Language and Culture Group. This initiative offers one module in Manchu language or Manchu culture each semester for five consecutive semesters, and the Year 2 participants were taking their third module during the pilot.

Moreover, in the primary school there is a discrepancy between the planned teaching weeks per semester (10-12 weeks) and its implementation (6 weeks in estimation). This is because, according to the participating students, the Manchu class was frequently appropriated by the class teachers for the instruction of major subjects, such as Chinese Literacy and Mathematics.

4. Data collection

Four methods were adopted for data collection. Classroom observation and interview with teachers were conducted for the understanding of Manchu teaching and learning in the sites. In addition, tests in Manchu, English and Mandarin assessed the participant's literacy skills in the languages. Furthermore, questionnaire was used for the measurement of the children's attitudes toward Manchu language learning.

4.1 Classroom observation

The planned observation in the primary school had to be cancelled during the pilot due to the Manchu teacher's reluctance to give me access into the classroom, despite my efforts to approach her by the recommendation of the school head teacher. For the middle school, since the Manchu teacher was in lactation and was thus assigned to teach the less demanding Music course, no Manchu class was being offered during my pilot, resulting in the cancellation of the classroom observation too.

In the university, unobtrusive observation was adopted with the observer sitting at the back of the lecture hall, where the class was given to over a hundred first year students. My observation focused on what the teacher and the students did in class in terms of Manchu teaching and learning. Specifically, I am interested in identifying 1) the ways in which romanised and traditional Manchu scripts are used in class; 2) the focus of instruction (i.e.,

what is instructed in class) and the session objective (i.e., what the students should learn during this session of Manchu). For the students, I intend to find what they do in class to learn Manchu. Before the observation, I was informed by Mr Sure that the contents of the parallel sessions offered to the students that week are the same; thus, one session is sufficiently representative of what a typical Manchu class looks like in the university.

4.2 Tests for literacy skills

Eight tasks of English word recognition, Chinese character naming, and Manchu symbol and word naming were given to the subjects to assess their literacy skills. A *word recognition task* requires the subject to sound out the printed word presented to him or her and give its Chinese meaning, whereas a *naming task* only asks the participants to pronounce the printed items presented to them.

Test purposes. The tests were designed for two purposes. First, they aim to identify whether the participants have obtained the basic knowledge in Manchu, such as its orthography and phonology, before the main study in Manchu word decoding could be done with them. In addition, the tests were to examine the validity of the test items with respect to their graded levels of difficulty.

Items selection. The items in Tests 1 were selected from the student's textbook for English, with the item difficulty being controlled by the number of the letters in a word (ranging from 2 to 8). For the Chinese characters in Tests 2, the first five characters in each group (e.g., 夫, 印, 争, 坛, 易 and 止, 辽, 尘, 吵, 枚), chosen from the textbook used in the previous academic year, should have been instructed to the students in class and were thus considered less difficult. The remaining five characters were selected from the textbook currently used, and therefore, they were regarded as more difficult because the students may not have learned them. In addition, the character structure, including single component, left-right, top-bottom, and enclosed, was also taken into consideration for the character selection.

The items for Manchu literacy in Tests 3 and 6 and Tests 4 and 7 are identical except that they were written in romanised and traditional Manchu script respectively. Tests 3 and 6

aim to examine the subject's knowledge in Manchu letters and phonology, thus they include all the Manchu phonemes (six vowels and 23 consonants) except that of /ŋ/, with the consonants being presented in the C + *a* syllable. Tests 4 and 7 include real Manchu words of various phonological structures, including V(diphthong), CV, CVC, VCV, VCCV, CVCV, CVCVC, CVCCV, VCVCV, CVCVCV, and VCCVCCV. Test 8 comprises eight Manchu ligatures (*be, bo, ge, ba, bu, ke, bi, he*) and seven ujus that contain diacritic dots and circles. The eight ligatures are in Test 5 too but are spelt in romanised letters. All the test items are listed in Table 2.

Materials*.* In the task of English word recognition, the twelve words were listed in four rows of three and printed in an A4 size paper with Times New Roman 32 points. In the Mandarin naming task, the ten Chinese characters were presented in two rows of five and printed in an A4 size paper with DengXian Regular 28 points. For the tests on Manchu, each item of letter, symbol or word was printed individually at the centre of an A4 paper, with romanised script in Times New Roman 60 points and traditional script in Abkai Xanyan 150 points.

Test administration. Each test was conducted on a one-to-one basis between the experimenter and the subject. At the beginning of the test, the subject was informed of what the test is about and instructed on what they need to do for the test, as well as given a trial. No time limit was set for the subject's reaction in recognising or naming the printed items. The process of testing was audio recorded.

Table 2

List of Test Items

Test No.	Tasks	Subjects	Test items
1	English word recognition (12 items)	Grade 4 students	my, pen, blue three, school, brother in, man, tall, fruit, family, twenty
		Grade 5 students	man, tall, fruit family, find, quiet dinner, computer, sunny pretty, weather, expensive
2	Mandarin characters naming (10 items)	Grade 4 students	夫，印，争，坛，易 止，辽，尘，吵，枚
		Grade 5 students	止，辽，尘，吵，枚 占，夹，折，昆，挖
3	Romanised Manchu letter naming (29 items)	All participants	*a, e, i, o, u, ū* *na, ka, ke, ga, ge, ha, he, ba, pa, sa, ša, ta, te, da, de, la, ma, ca, ja, ya, ra, fa, wa*
4	Romanised Manchu word naming (11 items)		*ai, bi, duin, ehe, abka, gebu, katak, bithe, agūla, muduri, amgambi*
5	Romanised Manchu ligature naming (8 items)	University participants	*be, bo, ge, ba, bu, ke, bi, he*
6	Traditional Manchu uju naming (29 items)	All participants excluding the Grade 4 students	*a, e, i, o, u, ū* *na, ka, ke, ga, ge, ha, he, ba, pa, sa, ša, ta, te, da, de, la, ma, ca, ja, ya, ra, fa, wa*
7	Traditional Manchu word naming (11 items)	Grade 8 students and university participants	*ai, bi, duin, ehe, abka, gebu, katak, bithe, agūla, muduri, amgambi*
8	Traditional Manchu ligature & diacritics naming (15 items)	University participants	*be, bo, ge, ba, bu, ke, bi, he* *ta, de, ka, da, ga, te, ha*

4.3 Interview

Semi-structured interviews with respects to course objectives, curriculum, and textbook were conducted with teachers to catch a glimpse of the status quo of Manchu teaching in the sites. Also, for the emerging interesting topics during interview, additional questions were given to the interviewees to probe for more details.

4.4 Questionnaire

The questionnaire aims at measuring the children's attitude towards Manchu language learning. The first part of the questionnaire gathers demographic information of the participants, such as their name, gender, ethnicity, etc. The second part has ten statements giving opinions towards Manchu language learning, and the participants are expected to choose from a five-point Likert scale with categories "Strongly Disagree", "Disagree", "Neutral", "Agree", and "Strongly Disagree". The subjects were given instructions before filling in the questionnaire, with explanation specifically given to the Likert scale.

5. Results and discussion

5.1 Classroom observation

In response to my first question for observation, while both romanised and traditional scripts were used by the teacher in the university class, they varied with respects to *frequency* and *function*.[22] Romanised script was used four times – two for spelling the ujus (Fieldnote Entries 3 and 6), one for presenting the text of a Manchu conversation (Entry 2), and one for spelling the Manchu words in a transcribing drill (Entry 14). The traditional script, however, was more frequently employed – they were used 12 times, six for presenting the ujus (Entries 3, 4, 6, and 8) and the remaining six for Manchu words (Entries 4, 5, 7, 9, 10, and 12).

The observation data suggest that the romanised and traditional scripts served two different functions in class, which are named in this report the "medium function" and the "target function". The romanised script functions primarily as the medium or tool of spelling or presenting the ujus and words, which could probably reduce the visual and mapping

[22] As introduced in Section 4.1, classroom observation was conducted only in the university site because of access issue.

complexity for the students. The traditional script, on the other hand, is the objective of learning. That is, the teacher intends to teach and the students were supposed to learn the ujus, including their orthography and phonology. Thus, the teacher made efforts to design and gave a number of classroom activities, such as ujus copying, chorus reading-aloud, word segmentation, and word transcription, to help students familiarise and memorise the symbols and the orthographic rules.

The above discussion also partially provides answer to the second question, that is, what is the course focus (i.e., what is instructed and learnt in class) and objective (i.e., what the students are supposed to achieve upon the completion of the class). Based on my observation, the course focused on instructing the Manchu orthography (i.e., the traditional script and the orthographic rules, especially the grapheme-to-phoneme mapping rules),[23] which ultimately aims at training the student's skills of decoding the traditionally scripted words.

With respect to what the students do in class to learn Manchu, their activities involved speaking (greeting the teacher in Manchu),[24] writing (practising copying the ujus and transcribing the romanised words),[25] and reading (sounding out ujus and words and naming the traditional Manchu words).[26] However, while the classroom practise involved three skills, the focus is on word decoding, which is in line with the course objective.

Apart from the above three questions, other topics of potential interest also emerged from my classroom observation. For example, would the word segmentation drill influence the student's *phonological awareness*? If it would, could that correlate with their *word decoding* ability? Moreover, whether or not was the *corrective feedback* provided by the teacher effective for improving the student's skills in Manchu *word decoding*?

[23] The grapheme-to-phoneme mapping rule is trained by the drill of transcribing the romanised Manchu words using the traditional script (Entry 14).
[24] See Fieldnote Entry 2.
[25] See Fieldnote Entries 3, 6, and 14.
[26] For sounding out, see Entries 3-6, and 8; for naming, see Entries 5, 9, 10, and 13.

5.2 Tests for literacy skills

Two methods were used for analyzing the data of literacy skills. The first method runs a task-by-task analysis by calculating the average score of the tasks (hereafter as “task average score” or “TAS”) taken by different groups of participants. The second is an item-by-item analysis, in which the average test score of an item (hereafter as “item average score” or “IAS”) is calculated, so as to understand 1) whether the items in a test have graded difficulty, and 2) whether and why certain item features (e.g., phonology, ligature, and diacritics) impact the participant’s performance.

The TASs are presented in Table 3. The most outstanding result is that the Year 2 university students outperform all other groups in Tests 3, 4, 7, and 8. Three reasons may account for this result. Firstly, as young adults the university students are cognitively more developed than the children; therefore, they performed better in the tasks. In addition, they, as members in the Experimental Manchu Programme, have passed the rigorous admission process by taking test in linguistic aptitude (e.g., phonological awareness), attitude towards Manchu language and culture, admission interview, etc.; thus, they are probably linguistically more capable and affectively more motivated in learning Manchu. Most importantly, the course length (five semesters in all with two already taken) and class frequency (90 minutes every week) of the Manchu initiative programme ensured that they have been able to take Manchu classes in a regular and continuous manner, which may also contribute to their outstanding performance.

Table 3

Task Average Score (TAS) of the Participants

Test No.	Tasks	Total points	Task average score (TAS)				
			Primary school participants		Middle School participants	University participants	
			Grade 4 (N=20)	Grade 5 (N=10)	Grade 8 (N=10)	Year 1 (N=15)	Year 2 (N=9)
1	English word recognition	24	12.7	9.4	–	–	–
2	Mandarin characters naming	10	9.5	9.8	–	–	–
3	Romanised Manchu letter naming	29	7.6	7.7	22.2	21.3	26.8
4	Romanised Manchu word naming	12	1.8	1.7	5.5	6.2	11.1
5	Romanised Manchu ligature naming	8	–	–	–	7.1	8
6	Traditional Manchu ujus naming	29	–	0	1.8	5.0 (out of 17) [a]	23.9
7	Traditional Manchu word naming	12	–	–	0.5	1.5	8.8
8	Traditional Manchu ligature & diacritics naming	15	–	–	–	2.9	12.3

[a] At the time of my pilot, the Year 1 participants were given only five sessions of Manchu, which covered the first 17 items (rather than all the 29) in this test. Thus, the average score of 5.0 was out of 17 possible points.

For the Year 1 university participants, they performed well in Test 3 (romanised Manchu letter naming) and Test 5 (romanised Manchu ligature naming). The fact that the vowels and C+*a* structure in Test 3 and the CV structure in Test 5 have their phonological counterparts in English and *pinyin* may be helpful for the participants to name the items, considering they are linguistically capable in Mandarin and English. Similarly, this explanation could also be applied to the good performance of the Grade 8 students in Test 3, who began learning *pinyin* in Grade 1. Furthermore, this account is also supported by the interview data. When the eighth graders were asked "How do you consider these test items –as Manchu letters,

pinyin, or English letters?", about half regarded them as *pinyin* and thirty percent saw them as English letters or letter combinations.

The Year 1 university students, however, did less well in Tests 6-8, with an average score of 5.0 (out of 17), 1.5 (out of 12), and 2.9 (out of 15) respectively. While the scores are low, it will be less surprising if the fact is considered that they are beginning learners who have taken only five sessions of Manchu before the pilot.

In the primary school group, the children in both Grades 4 and 5 performed poorly in the Manchu naming tests. The fourth graders were given the naming tasks on romanised Manchu letters and words, which, according to Ms Guang, is the planned contents for Grade 3 students in the school Manchu course curriculum.[27] This test result could be attributed to the appropriation of Manchu class hours and the poor curriculum implementation. During the tests I was informed by the children that the Manchu class were often appropriated by their class teacher either for the instruction of Mandarin Chinese and Mathematics or for giving examinations. As a result of that, they only took a very limited number of Manchu classes in Year 3. More importantly, Ms Guang introduced in the interview that the instruction for Year 3 students has focused on children's songs, games, and basic conversation, rather than on Manchu literacy. Consequently, not much was taught or learnt on Manchu letters and words.

For Grade 5 students, they began taking Manchu class two years ago in Grade 3. Although, according to Ms Guang, the curriculum requires the romanised Manchu script to be instructed in Grade 3 and the traditional script in Grades 4, in fact, the Grade 5 students, like their Grade 4 counterparts, were offered limited hours of Manchu during the past two years.[28] What is significantly surprising is that they were unable to recognise *any* traditional Manchu ujus in Test 6. This is probably due to the fact that few ujus had been instructed to them in Grade 4,[29] which is not consistent with the course curriculum information provided by Ms Guang during the interview.

[27] The fourth graders were tested in the pilot on what they have learned in the previous academic year.
[28] This information was gathered from the Grade 5 participants during test.
[29] This information was provided by the Grade 5 participants during test.

In a nutshell, the above data suggest that the university participants performed better than other groups in the Manchu naming tasks. In addition, the limited Manchu class hours offered to the children and the discrepancy between the curriculum and its implementation could also account for the young learner's poor test performance.

The following analysis, adopting an item-by-item method, attempts to examine whether the items in the tests of Manchu literacy bear graded difficulty. Moreover, it also aims at understanding whether and why certain item features (e.g., phonology, ligature, and diacritics) impact the participant's performance. Note that the following analysis only uses the test data of the university participants.

The total possible score for a participant in Tests 3-8 is 93 points.[30] The test scores of the Year 1 participants (N=15) range between 29 to 59, and the scores of Year 2 students (N=9) are from 71 to 93. Figure 2 shows that the number of participants in each score range falls into two curves of normal distribution, with the curve on the left being the Year 1 participant group and the one on the right Year 2 group. Therefore, the data of the two groups were treated individually as two clusters instead of one cluster with the 24 sets of data mixed together. Thus, when calculating item average score (IAS), the average test score of an item in each cluster was calculated first, and then the two scores were averaged again for the IAS.

The data in Table 4 suggest that the test items have cross-sectional graded difficulty, with the traditionally scripted items in Tasks 6-8 more challenging than the romanised ones in Tasks 3-5 for the university participants to name. This result may be attributed to their extensive language learning experiences in *pinyin* and English and the consequent familiarity with the inventory of Latin letters.[31] The traditional Manchu script, on the contrary, is a set of completely new writing symbols that they have begun learning very recently.

[30] Although originally there were 29 items in Test 6, only the data of the first 17 items were included in the analysis because the remaining 12 items were not yet instructed to the Year 1 participants during the pilot and they were thus excluded from data analysis.

[31] In Chinese educational system, children began learning *pinyin* and English in Grade One and Grade Three respectively. In addition to the language instruction received at school, a large number of children in China take extracurricular English classes after school.

Moreover, the items in each individual task also demonstrate graded difficulty. the following analysis focuses on the outlier items. In Task 3, the outliers are Item 17 *ša* (.256), Item 24 *ca* (.167), and Item 27 *ra* (.278), which appear to be more difficult for the participants to name than others. For *ša* and *ca*, the difficulty probably comes from the allographs for /ʃ/ – <*sh*> was instructed in class while <*š*> was used for the test item. So could it be applied to <*ca*> and <*cha*>. As for *ra*, it poses challenge for the participants because there is no alveolar trill /r/ in their L1 Mandarin Chinese or L2 English, thus making it almost impossible to transfer the pronunciation from Mandarin or English to Manchu.

The difficulty in sounding out /r/ may also associate with the participant's poor performance in Item 11 *muduri* (.478) in Task 4. In fact, the word *muduri* is supposed to be phonologically less challenging because of its structural components of three CVs, which is typical in Mandarin Chinese phonology. In addition, the easiness of naming CV structure is corroborated by the significantly high IAS (.946) of Task 5, in which all the eight test items are in CV structure, and by the high IASs of *bi* with CV structure (.933) and *gebu* with CVCV structure (.867) in Task 4. Thus, one possible reason for the low score of *muduri* could be that the participants had difficulty in pronouncing /r/, thus resulting in the failure in naming the entire item. A qualitative analysis on the types of errors made by the participants on *muduri* could provide a more accurate answer.

Figure 2

Distribution of the Number of Participants in Each Score Range (n = 24)

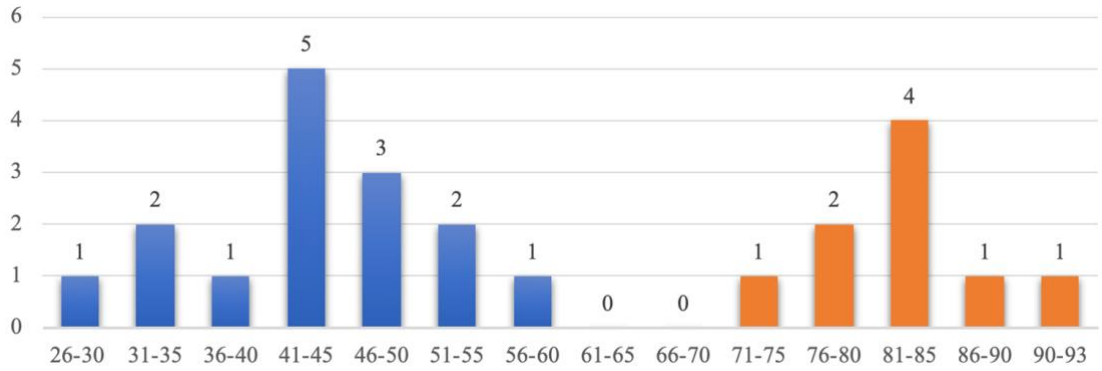


Table 4

Item Average Score (IAS) of the University Participants

Test item No.	Task 3 Romanised Manchu letter naming (N=29)	Task 4 Romanised Manchu word naming (N=12)	Task 5 Romanised Manchu ligature naming (N=8)	Task 6 Traditional Manchu ujus naming (N=17)	Task 7 Traditional Manchu word naming (N=12)	Task 8 Traditional Manchu ligature & diacritics naming (N=15)
1	1.000	1.000	0.867	0.622	0.589	0.667
2	1.000	0.933	0.900	0.611	0.667	0.656
3	1.000	0.933	0.933	0.733	0.833	0.444
4	0.911	0.667	1.000	0.789	0.389	0.744
5	0.967	0.733	1.000	0.678	0.444	0.711
6	0.544	0.589	0.933	0.656	0.167	0.500
7	1.000	0.867	1.000	0.656	0.533	0.733
8	0.967	0.511	0.933	0.400	0.111	0.444
9	0.933	0.567	–	0.389	0.500	0.500
10	1.000	0.733	–	0.422	0.333	0.278
11	0.933	0.478	–	0.444	0.389	0.433
12	1.000	0.644	–	0.567	0.167	0.222
13	0.933	–	–	0.500	–	0.311
14	1.000	–	–	0.700	–	0.389
15	1.000	–	–	0.578	–	0.567
16	1.000	–	–	0.578	–	–
17	0.256	–	–	0.567	–	–
18	0.900	–	–	–	–	–
19	0.867	–	–	–	–	–
20	0.833	–	–	–	–	–
21	0.800	–	–	–	–	–
22	0.800	–	–	–	–	–
23	0.900	–	–	–	–	–
24	0.167	–	–	–	–	–
25	0.567	–	–	–	–	–
26	0.833	–	–	–	–	–
27	0.278	–	–	–	–	–
28	0.833	–	–	–	–	–
29	0.800	–	–	–	–	–
Ave.	**0.828**	**0.721**	**0.946**	**0.582**	**0.427**	**0.507**

For Item 9 *bithe* (.567) and Item 6 *abka* (.589) in Task 4, the difficulty could probably be attributed to their phonological structure too. The two adjacent consonants ("*th*" and "*bk*") in them make it demanding for the participants to segment the phonemes into syllables of CVC/CV and VC/CV before the words could be correctly named. Similarly, this difficulty in phonemes-to-syllables segmentation could also account for the low item test score of Item 8 *katak* (.511), which can phonologically be segmented into either CV+CVC (i.e., ka/tak) or CVC+VC (i.e., *kat/ak), with all the structures of CV, CVC, and VC legitimate in Manchu phonology.

The following section discusses the IAS of Tasks 6-8, in which the items were traditionally scripted. One obvious result is that the IASs of Tasks 6-8 (.582, .427, and .507) are lower than these of Tasks 3-5 (.828, .721, and .946), indicating that the traditional Manchu script poses more challenge for the participants in item naming. In Task 6, the four lowest IASs go to Item 9 *ke*^{T} (.389), Item 8 *ka*^{T} (.400), Item 10 *ga*^{T} (.422) and Item 11 *ge*^{T} (.444).[32] This result corresponds with the Task 8 data which, too, show that these items are among those of the lowest IASs, suggesting that they are of significant difficulty for the participants to name. A possible reason could be that *ka*^{T} and *ga*^{T} are orthographically identical except the fact that *ga*^{T} takes an additional diacritic dot on its right, and such orthographic similarity may have caused confusion for the participants when they attempted to name the two items. So could this explanation be applied to the low IASs of *ke*^{T} and *ge*^{T}.

Task 7, the traditional Manchu word naming, is the most challenging task for the participants. Significantly low are the IASs for Item 8 *katak*^{T}(.111), Item 6 *abka*^{T} (.167), and Item 12 *amgambi*^{T} (.167), together with the other two low-scoring items of *agūla*^{T} (.333) and *muduri*^{T} (.389). Different reasons may account for the item difficulty. For *katak*^{T}, while the ujus of *ka*$_{1}$ and *ta*$_{2}$ are of high frequency in Manchu,[33] *k*$_{3}$ is a much less frequently used item in word and thus the unfamiliarity with this item brought difficulty in its decoding. The word *abka*^{T}, however, poses a different challenge – the ligature of *b*^{T} and *ka*^{T} (i.e., integrating *ka*^{T} into *b*^{T} in writing) makes it difficult to segment the word into *ab*$_{1}$ and *ka*$_{3.}$

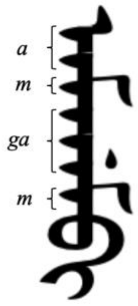
[32] Since there is no Manchu input keyboard available for the macOS system, the Manchu uju and the traditionally scripted Manchu words are indicated by a superscript capitalized letter "T", for example, *ka*^{T} and *abka*^{T}.

[33] Subscript numerals are adopted as indicators of the positional variations of a uju, with 0 indicating the isolation variation, 1 the initial variation, 2 the medial variation, and 3 the final variation.

As for *amgambi*^{T}, its difficult may be derived, firstly, from the seven graphemic "short tooth" embedded in the word, which could be visually complex and thus intimidating for the participants to name. Secondly, the seven short teeth carry grapheme-to-phoneme mapping complexity, in that a single short tooth in Manchu is in correspondence with /a/ and /n/ and a cluster of three short teeth corresponds with /ka/. Closely related to this mapping complexity is the boundary identification between ujus. For example, the second short tooth needs to be considered in combination with the graphemic "crest" above it (part of which is a short tooth) and the short tooth (and the downward hook) underneath it to identify the uju of *am*. Similarly, the fourth, fifth and sixth short teeth should be jointly considered as an entity, and with the diacritic dot to its right, to be identified as the ujusof ga_2. It is until all of these were considered by a reader that a boundary could be identified between the six short teeth (together with other elementary graphemic component, such as the crest and the downward hook) in *amgambi*^{T}. That is, it is essential for the readers to identify the *boundaries* between ujus so that a Manchu word be correctly decoded. Without doubt, the grapheme-to-phoneme mapping complexity increases the difficulty for boundary identification.

Figure 3

Grapheme-to-Phoneme Mapping Of AMGAMBI



5.3 Interview

As planned, semi-structured interview was conducted with Ms Guang in the primary school and Mr Sure in the university. Moreover, an additional interview was carried out with the head teacher in the middle school, although it was not a part of my original plan.

According to Ms Guang, Manchu classes are offered to students in Grades 3, 4 and 5 in the school. The instruction for Grade 3 students, mediated by the romanised Manchu script, focuses on children's songs and very basic conversation, which involved the words of animals, colours, numbers, and titles. For Grade 4 students, Ms Guang gave the information that they are supposed to learn the traditional Manchu ujus. However, the result of Task 6 demonstrated that the Grade 4 participants were unable to name any traditionally scripted words. They told the experimenter that in fact only a few sessions of Manchu class were given during Grade 4 and little instruction was on traditional Manchu uju. For Grade 5 students, the plan in the curriculum is to make up for what was left untaught and unfinished in the previous two years.

The textbook in use is *Primary School Manchu*, but, according to Ms Guang, it does not fit the children's needs because finishing it requires more class hours than what the school could provide. Moreover, a curriculum for Manchu course is said to have been developed; however, Ms Guang was reluctant to show it to me.

In the university the course objective for Manchu, according to Mr Sure, is to learn the elementary Manchu orthographic knowledge, including the ujus, the symbol-to-sound mapping, and the decoding of Manchu words.

During the interview in the middle school with the headmaster, Mr Zang, two interesting topics emerged. One is the priority of the Manchu-related activities in the school. Mr Zang introduced that the school gives priority to the activities concerning Manchu culture and custom, which probably suggests that Manchu literacy teaching is not among the top priorities. This policy is in line with my observation in the primary school where photos of students engaging with and winning Manchu calligraphy contest, Manchu sports meeting, Manchu dancing, etc. are hung on the walls in the building corridors to show the honours. The Manchu literacy class, however, did not receive that much of attention at the school.

Secondly, Mr Zang listed the four big challenges of Manchu teaching and learning that the school is facing. The biggest challenge is the lack of teachers in Manchu and other subjects. Although it seems that this issue can be easily addressed by recruiting more teachers, the problem is that the school is not entitled to employ teachers as they like, because the school teacher quota is allocated by the government. Even if additional quota is given, the priority will typically be given to recruiting teachers of the major subjects, such as Mandarin Chinese, Mathematics, and English, rather than Manchu. Second, the parent's lukewarm attitude towards Manchu does not help increase their children's enthusiasm in learning it. While the parents are not against the offer of Manchu course to their children at school, they do not care much about it. Instead, they are more interested in their children's performance in major subjects which are to be assessed in the high-stake examinations for high school and college entrance. This may have an unhelpful, if not negative, influence on children's Manchu learning. Third, the children are holding a positive attitude toward Manchu as a language, which could be further supported by the questionnaire data, but still they do not see any practical benefit of learning it. Finally, no linguistic environment exists for the children to use Manchu for real communication.

5.4 Questionnaire

A total number 40 participants (20 fourth graders, 10 fifth graders, and 10 eighth graders) were invited to complete the questionnaire measuring their attitude towards Manchu language learning. Because of age difference, the data was groups into two clusters – the primary school cluster (30 participants) and the middle school cluster (10 participants) – for analysis and the item score is the average of the cluster averages.[34] The result of the questionnaire is given in Table 5.

[34] A Likert scale of 1-5 points was used in correspondence with categories "Strongly Disagree", "Disagree", "Neutral", "Agree", and "Strongly Agree" in the questionnaire.

Table 5

Scores of the Questionnaire Items

Item No.	Questions	Primary school cluster item average score	Middle school cluster item average score	Item score
1	Learning Manchu is really great.	4.07	4.20	4.13
2	I plan to learn as much Manchu as possible.	3.83	4.50	4.17
3	I hate Manchu.	2.10	1.70	1.90
4	Manchu is an important part of the school programme.	3.53	4.00	3.77
5	I think that learning Manchu is dull.	1.87	1.70	1.78
6	I really enjoy learning Manchu.	4.03	4.40	4.22
7	I would rather spend my time on subjects other than Manchu.	2.63	2.30	2.47
8	Learning Manchu is a waste of time.	1.73	1.80	1.77
9	I love learning Manchu.	3.97	4.50	4.23
10	When I leave school, I shall give up the study of Manchu entirely because I am not interested in it.	2.00	1.70	1.85

Note. Adapted from Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. Arnold.

The data show that Questions 1, 2, 4, 6 and 9, which give statements describing positive attitude towards Manchu learning, score over four points (“Agree”) only with the exception of Question 4 (3.77). This suggests that the children are interested in learning Manchu, but not very certain about its importance among the school subjects. For the five negative descriptors of Manchu, the four items of 3, 5, 8 and 10 scored between 1 and 2 points, demonstrating that the participants hold a strong opinion against the idea that Manchu language and learning is boring or a waste of time. However, when asked in Question 7 whether they would like to give priority to learning other subjects rather than Manchu, while the data (2.47) present that the participants did not think they would do so, they still show uncertainty towards the importance of Manchu when it is compared with other school subjects. Overall, the questionnaire result is in line with the interview data from Mr Zang

that the students are interested in Manchu as a language, but they are not quite certain about its importance among the school subjects.

6. Findings

Based on the above data, this section presents the findings regarding the Manchu course provided by each site and the test items adopted in the tasks. In the primary school, while the participating children have a strong interest in learning Manchu, their Manchu word decoding skill is rather limited, especially in that of the traditionally scripted items. Moreover, a discrepancy exists between the course curriculum and its implementation with respects to the class hours and the teaching contents. In the site of the middle school, the offer of Manchu course to the students was not regular, partially due to the availability of the teachers of Manchu in a semester. However, in the university, the curriculum was well implemented – the Manchu classes were regularly and continuously offered each week and the classroom instruction fulfilled the course objective of being able to decode the traditionally scripted words. Not only did the university participants outperform other two groups in the tasks of word naming and recognizing (of course, their age difference should be taken into consideration), the number of participants in each score range is also normally distributed (cf. Figure 2).

With respect to the test items, the six sets present inter-sectional graded difficulties, with the traditionally scripted items more challenging to name than the romanised ones. In addition, the items within a set demonstrate intra-sectional graded difficulties. An analysis of the outlier items (i.e., the items with low IASs), suggests that a number of reasons could account for the difficulty, including allographs for a phoneme (e.g., <sh> and <š> for /ʃ/), the failure of phonological transfer between Mandarin and/or English to Manchu (e.g., the alveolar trill /r/), the phonological structure of two adjacent consonants in a word (e.g., CVCCV and VCCV), orthographic similarity (e.g. *ke*^{T} and *ge*^{T} with the only difference being a diacritic dot), frequency of a uju in Manchu vocabulary (e.g., the variation of *k*^{T} in the final position), visual complexity (e.g., seven “short teeth” in a word), and grapheme-to-phoneme mapping complexity (e.g., one grapheme in correspondence with multiple phonemes).

7. Implication for my main study

The above findings bring three implications for my main study. First of all, the university located in the City of Chengde would be a more ideal site for my research project. Secondly, while the pilot study gives the average score for each test item, what yet remains unknown is what factors have led to the errors or mistakes in the naming tasks, which might be identified by quantitative measures (e.g., phonological awareness test) or qualitative interviews with the participants. Finally, the data and discussion also shed some light on my main study, which attempts to investigates the relations between the complexity of Manchu writing system (i.e., visual complexity, phonological complexity, and symbol-to-sound mapping complexity) and Manchu word decoding.

Appendix C

Questionnaire of Manchu Language Learning Experience

Your responses to this survey will be kept confidential and anonymous.
Please answer all questions HONESTLY.

1. Birthday: Gregorian Year_____Month_____
2. Sex (Please tick): ☐ Female ☐ Male
3. Have you learned any Manchu **before** entering university? (Please tick)

   ☐ Yes (→ Continue to Question 4)

   ☐ No (→ Skip to Question 5)
4. In what ways did you learn Manchu? (Please tick all that apply)

   ☐ Attending Manchu classes or interest group

   ☐ Learning with Manchu native speakers or people who are Manchu-literate

   ☐ Self-study

   ☐ Others (Please specify)_____
5. In addition to the 90-minute Manchu instruction you are receiving every week, do you have any **extracurricular** contact with Manchu? (Please tick all that apply)

   ☐ I do homework, revision, or preview for the Manchu course.

   ☐ I read Manchu books.

   ☐ I listen to Manchu songs.

   ☐ I do **NOT** have extracurricular contact with Manchu.

   ☐ Others (Please specify)_____
6. You took the National College Entrance Examination (NCEE) in the Province of _____.
7. Your Chinese NCEE score is _____points.
8. Your English NCEE score is _____points.

THANK YOU FOR YOUR COORPERATION!

Appendix D

Frequency of Manchu Phoneme Marker and Uju in Instruction

1. Manchu symbols written in traditional script are denoted by uppercase letters (e.g., NA), and their Latin transliteration are by lowercase letters (e.g., na).
2. The isolation form of a Manchu symbol is written independently (e.g., NA), and its initial, medial, and final forms were written before, between and after a hyphen or hyphens ("-") (e.g., NA-, -NA-, and -NA).
3. The K that is only followed by A, O, and Ū is denoted as K_1 (e.g., K_1A), and the K only followed by E, I, and U is K_2 (e.g., K_2E). This denotation is also applicable to G and H.
4. The T that is only followed by A, I and O is denoted as T_1 (e.g., T_1A), and the T only followed by E and U is T_2 (e.g., T_2E).
5. The D only followed by Ū is denoted as D_3 (i.e., D_3Ū). D_3 and T_2 are identical in writing.
6. The final form of T is denoted as T_3 (e.g., AT_3), and the T at the end of a syllable which is not at the end of a word is denoted as T_4 (e.g., AT_4-).
7. The U with and without a dot on the right is denoted as U and U_1 respectively.
8. The medial form of "I" with one long tooth was written as -I_1- , and the one with two long teeth was denoted -I_2-, the possessive case of I is I_2.
9. The medial form of N (i.e., -N-), if at the beginning of a syllable, is denoted -N_1- (e.g., AN_1A); if at the end of a syllable, is -N_2- (e.g., NAN_2N_1A). The final form of N in EN is denoted -N_1.
10. The final form of B is denoted as -B_1 (e.g., AB_1).

Orthographic units of phoneme marker and uju	Frequency count
A	12
A-	20
-A-	4
-A	7
E	3
E-	5
-E-	2

Orthographic units of phoneme marker and uju	Frequency count
-E	3
I	10
I-	5
-I-	4
-I	8
-I_1-	9
I_2	1
O	5
O-	5
-O	3
U	7
U-	3
-U	1
U_1-	1
-U_1-	6
-U_1	5
Ū	4
Ū-	3
IU_1I	1
IU_1I_1-	1
AI	1
AI_1-	1
AU_1	1
AU_1-	1
EU_1-	1
-IY-	3
OU	1
AB-	2
AB_1	2
AL	1
AM-	1

Orthographic units of phoneme marker and uju	Frequency count
AM	2
AN	2
ANG	1
ANG-	1
AR	2
AS	1
AS-	1
AŠ	1
AŠ-	1
AT_3	1
AT_4-	1
-AT_4-	3
B-	6
-B-	3
BA	3
BA-	2
–BA–	2
-BA	3
BE	2
BE-	1
–BE–	1
-BE	2
BI	2
BI-	1
-BI-	1
-BI	1
BO	2
BO-	1
-BO-	1
-BO	1
BU	1

Orthographic units of phoneme marker and uju	Frequency count
BU-	1
-BU-	1
-BU	1
BŪ	1
BŪ-	1
-BŪ-	1
-BŪ	1
C’Y	1
C’Y-	1
-C’Y-	1
-C’Y	1
CA	1
CA-	1
-CA-	1
-CA	1
CI	1
D_1-	1
D_1A	1
D_1A-	1
-D_1A-	1
-D_1A	1
D_1I	1
D_1I-	1
-D_1I-	1
–D_1I	1
D_1O	1
D_1O-	1
-D_1O-	1
-D_1O	1
D_2E	1
D_2E-	1

Orthographic units of phoneme marker and uju	Frequency count
-D$_2$E-	1
-D$_2$E	1
D$_2$U	1
D$_2$U-	1
-D$_2$U-	1
-D$_2$U	1
D$_3$Ū	1
D$_3$Ū-	1
-D$_3$Ū-	1
-D$_3$Ū	1
DZA	1
DZI	1
DZI-	1
-DZI-	1
-DZI	1
EN$_3$	1
ENG	1
-F-	1
FA	1
FA-	1
-FA-	1
-FA	1
FI	1
FI-	1
-FI-	1
-FI	1
G$_1$-	1
-G$_1$-	1
G$_1$A	1
G$_1$A-	1
-G$_1$A-	2

Orthographic units of phoneme marker and uju	Frequency count
-G_1A	5
–G_2E	1
–G_2I	1
–G_1O–	1
–G_1O	2
G_2U	2
–G_2U	1
–G_1Ū–	1
–G_1Ū	2
H_1-	1
-H_1-	1
H’Ū	1
H’Ū-	1
-H’Ū-	1
-H’Ū	1
H_1A	1
H_1A-	1
-H_1A-	1
-H_1A	4
-H_1O	1
H_2U	1
H_2U-	1
H_2UK_2-	2
IL-	1
ING	1
-J-	5
JA	1
JA-	1
-JA-	1
-JA	1
JI	1

Orthographic units of phoneme marker and uju	Frequency count
JY	1
JY-	1
-JY-	1
-JY	1
K_1-	4
-K_1-	5
-K_1	4
K_2-	4
-K_2-	4
-K_2	4
K’-	2
K’A	3
K’A-	1
-K’A-	1
-K’A	1
K’O	1
K’O-	1
-K’O-	1
-K’O	1
K’Ū	1
K’Ū-	1
-K’Ū-	1
-K’Ū	1
K_1A	3
K_1A-	5
-K_1A-	5
-K_1A	4
K_2E	3
K_2E-	2
-K_2E-	2
-K_2E	1

Orthographic units of phoneme marker and uju	Frequency count
K_2I	2
K_2I-	1
-K_2I-	2
K_2U	3
K_2U-	2
-K_2U-	1
-K_2U	1
K_1AI	1
K_1AI-	1
-K_1AI-	1
-K_1AI	1
K_1AN	1
K_1AN-	1
-K_1AN-	1
-K_1AN	1
K_1AU_1	1
K_1AU_1-	1
-K_1AU_1-	1
-K_1AU_1	1
-K_2EI_1-	1
-L	2
LA	1
LA-	1
-LA-	1
-LA	1
-M	3
MA	1
MA-	2
-MA-	1
-MA	1
N-	2

Orthographic units of phoneme marker and uju	Frequency count
-N	5
-N_1-	1
-N_2-	3
NA	5
NA-	8
-NA-	4
-NA	7
NE	3
NE-	3
-NE-	1
-NE	2
-NG-	8
-NG	2
NI	2
NI-	2
-NI-	1
-NI	1
NO	2
NO-	2
-NO-	2
-NO	2
NU	2
NU-	1
-NU-	1
-NU	1
NŪ	2
NŪ-	1
-NŪ-	1
-NŪ	1
NAN-	1
OR-	1

Orthographic units of phoneme marker and uju	Frequency count
PA	2
PA-	1
-PA-	1
-PA	1
PE	1
PE-	1
-PE-	1
-PE	1
PI	1
PI-	1
-PI-	1
-PI	1
PO	1
PO-	1
-PO-	1
-PO	1
PU	1
PU-	1
-PU-	1
-PU	1
PŪ	1
PŪ-	1
-PŪ-	1
-PŪ	1
-R	2
RA	1
RA-	1
-RA-	1
-RA	1
-S	2
SA	5

Orthographic units of phoneme marker and uju	Frequency count
SA-	2
-SA-	2
-SA	2
SY	2
SY-	2
-SY-	2
-SY	2
-Š	2
ŠA	1
ŠA-	1
-ŠA-	2
-ŠA	1
-ŠE-	1
-ŠEN	2
T_1-	1
T_2-	1
-T_3	2
-T_4-	2
T_1A	3
T_1A-	2
-T_1A-	2
-T_1A	2
T_1I	1
T_1I-	1
-T_1I-	1
-T_1I	1
T_1O	1
T_1O-	1
-T_1O-	1
-T_1O	1
T_2E	2

Orthographic units of phoneme marker and uju	Frequency count
T₂E-	2
-T₂E-	2
-T₂E	2
T₂U	1
T₂U-	1
-T₂U-	1
-T₂U	1
TS’A	1
TS’I	1
TS’I-	1
-TS’I-	1
-TS’I	1
WA	1
WA-	1
-WA-	4
-WA	1
YA	1
YA-	1
-YA-	1
-YA	1
ŽA	1
ŽA-	1
-ŽA-	1
-ŽA	1