Chinese EFL Learners' Acquisition of Phonology:

A Comparative Analysis of the Influence of Two Dialects

(Northeastern and Cantonese)

Xinliang Jiang

Thesis submitted for the award of Doctor of Philosophy

Faculty of Humanities and Social Sciences

School of English Literature, Language and Linguistics

October 2019

Abstract

Chinese and English belong to two distinct language families. Given that English is a *lingua franca*, there are millions of learners of English who speak one of the varieties of Chinese. Due to the substantial phonological differences between the two languages, Chinese learners of English may encounter difficulties when they communicate in English; developmental issues may also play a role. In addition to Mandarin, Chinese has many dialects/languages and, because these dialects/languages differ from each other, speakers of these Chinese dialects/languages pronounce English differently. It is reasonable to suppose that this behaviour is phonologically governed. The influence from the first language (L1) on second language (L2) phonology has long been viewed as an important factor (Lado, 1957; see also Anderson, 1987; Corder, 1967; Ellis, 1994; Fisiak, 1978, 1991; Gass, 1979; Odlin, 1989; Rasier and Hiligsmann, 2007; Young-Scholten, 1985). Moreover, a large number of relevant research studies on the influence of L1 on L2 have been conducted on the effects of L1 Chinese - typically Mandarin - on L2 English phonology (Li, 2006; Wang, 2007) and to a lesser extent on the influence of other L1 Chinese dialects (Chen, 2010).

Why do different Chinese dialects/languages generate differences in the non-target production of English? How do different Chinese dialects/languages influence L2 English and what features, error types and specific errors do different L1 dialect/language speakers make? Questions like these can be answered partially by consulting the literature, but also need further exploration. Moreover, comparative research analysing one language (or language group) but two dialects/languages with L2 English is limited. Therefore, this study explores the phonological differences between two L1 dialects/languages to see what different effects they have on L2 phonology, and thus it contributes to filling this gap in the literature. In so doing, Lado's Contrastive Analysis Hypothesis (CAH, 1995) and Flege's Speech Learning

Model (SLM, 1995) are applied as theoretical frameworks. The CAH and SLM involve contradictory notions concerning categories of 'similar' and 'new' in terms of which is more difficult for the learner. Both are addressed in the present thesis to determine which can best account for the difficulties L2 English learners have in their oral production of L2 English. It needs to noted that as Groves and Mair (2008) said, Chinese situation is unique because mutual intelligibility principle is not sufficient to determine whether Chinese varieties are dialects or languages, thus, I will refer to Harbinese Mandarin and Guangzhou Cantonese as dialects as they are conventionally referred to.

This research firstly compares Harbinese Mandarin and Guangzhou Cantonese which fall under the umbrella of Chinese, and to do so with respect to segments, syllable structure and stress, and their different effects on learners' acquisition of English phonology, followed by the proposal of hypotheses based on Flege's idea of L1-L2 similarity-based degree of difficulty in SLM.

Data was collected to test these hypotheses from 65 participants from three schools at different educational levels (middle school, high school and university) from Harbin and Guangzhou. Auditory analysis, together with acoustic analysis and a native speaker's spot check, was used to guarantee the validity of the analysis and the reliability of the results. In addition, independent-samples t-tests were carried out to check the significance of the differences in L2 English production between the two Chinese groups.

The results indicate that the influence of L1 Chinese dialects/languages on L2 English is found everywhere in the sample, including in segments, syllable structure and stress, and that this influence is statistically significant. Different error types and patterns made by Harbinese and Cantonese learners of English were found. Mandarin is also spoken by the participants and its influence can be detected from the Cantonese results. The hypotheses in the category of ‘similar’ were generally rejected and in the category of ‘new’ were completely rejected. These findings indicate that Flege’s SLM model suggesting that L1-L2 differences that are ‘similar’ are more difficult than ‘new’, cannot be supported in this context. On the contrary, Lado’s CAH, where ‘new’ differences are predicted for the difficulties L2 learners may have, was supported. In addition, the varieties of English used by Harbinese and Cantonese speakers were also checked. It seems that Harbinese speakers tend to speak American English and Cantonese speakers speak British English, but the difference is not strongly significant. Thus, it is suggested that the variety they speak may be influenced by the similarity between L1 dialects/languages and English varieties; that is to say, the dialect/language more similar to a variety of English influences oral production. With respect to the hypothesis that increased length of exposure leads to reduced error rates, the results are not completely supportive because high school subjects score best among the three levels. This may be due to factors relating to the recent evolution of English teaching in China.

Dedication

To my dear parents,

Mr Lianfang Jiang and Mrs. Chunjuan Zheng,

and my dear wife,

Yue Zhuo

Acknowledgements

After four years of arduous and diligent work, I have completed my research as scheduled and successfully written my PhD thesis. Over the past four years, I have received continuous support from my family and have missed them tremendously. I have been inspired by encouragement and support from my supervisors and friends when I have struggled with the work. I have tasted the satisfaction of obtaining knowledge although I have felt frustrated when my goals were not achieved. In short, there have been both relaxed and happy times as well as nervous moments of restlessness. I have run the full gamut of emotions while working on this thesis.

During the past four years, many people have given me help, encouragement, support and care. Here, I express my most sincere gratitude to them!

First of all, I would like to express my sincere gratitude to my two supervisors: Professor Martha Young-Scholten, and Dr. S. J. Hannahs. Many thanks to Dr. Hannahs for guiding me through the world of phonology when I did my MA in linguistics, and I thank him for offering me the chance to study for a PhD at Newcastle University, and I also thank him for supervising and supporting my PhD research. What's more, great thanks go to my first supervisor, Professor Martha Young-Scholten, who has guided me to the successful completion of this thesis. Without her insistence, I would never have attended an international conference in 2016 when I was a first-year PhD student, and I would never have met so many leading researchers in second language acquisition (SLA). This experience broadened my horizons immensely. If she had not kindly pointed out my problems in pronunciation, I would never have been aware of them, and might never have improved my pronunciation. Without her encouragement and support, I would never have participated in committees in my school nor taken on responsibilities in those committees. Without her patient guidance, I would not have been able to complete my doctoral thesis. I also thank all tutors and staff who have imparted knowledge to me, and especially Dr. Danielle Turton, who were a great help in my research. Thanks also go to Dr. Ella Dzelzainis, Dr. Heike Pichler, and Mr. Scott Burdon who gave me support and encouragement when I worked as secretary for the postgraduate staff-student committee. I also express special thanks to Dr. Alex Ho-Cheong Leung and Dr. Gary Taylor-Raebel for agreeing to be my examiners.

Secondly, I wish to express my greatest appreciation to my family, and especially to my dearest parents: my father, Mr. Lianfang Jiang and my mother, Mrs. Chunjuan Zheng. Without their love and their strong and endless support emotionally and financially, there would not have been the opportunity to study at such a wonderful university and I would not have completed my PhD. Also, great thanks also go to my dearest wife, Mrs. Yue Zhuo, who gives me endless love, strong support and encouragement. Although this research study has meant we are living in two different countries, our love for and thoughts of each other are deep-rooted. In addition, many thanks are given to my parents-in-law, Associate Professor Ou Zhuo and Mrs. Fengxia Wu, and my uncles, aunties, and cousins for their endless support and encouragement.

Thirdly, I would like to send my great thanks to all the people who have given me great help, care and support over the past four years. I am very grateful to my old friend, Philip Dove, who proofread my long research texts without complaint and invited me for Christmas dinner every year. I am also very thankful to my Iraqi sister, Ekhlas Ali Mohsin, and Iraqi brother, Hayder Mohsin, who gave me not only such strong support and help, but also provided me with the warmth and care of the family. Great thanks are given to my English friends, Layla Embleton, George Harrison, Matthew Brindley, Zac Mabbs, and Caitlin Halfacre, for their support in the transcription of some of my data, to Dr. Yang Long and Dr. Simon Kometa for their support in the use of statistical software, to Mr. Bill Gibson and Dr. Daniel Bell for their support in proofreading part of my writing and to Mr. Len Shaw for professionally proofreading the final draft of my thesis. Thanks also go to my colleagues Dr. Dongyan Chen, Dr. Rebecca Ishaku Musa, Dr. Chanuwan Phetkla, Dr. Enas Filimban, Dr. Chutinan Noobutra, Khansaa Martakush, Azza Al-Kendi, Vean Al-Saka, Dr. Maha Jasim, Dr. Bashayer Alotaibi, Dr. Suaad Ambu Saidi, Dr. Ourooba Shetewi, Andreas Krug, Bilal Alsharif, and Andrew Latimer for their encouragement and care, and to my friends, Zhen Gong, Dr. Feng Li, Shikun Deng, Dr. Egle Mocciaro, Jules Kielmann, Neslihan Köroğlu, Martin Enilov, Kaveh Nobari, Filippo Corbelli, Phil Franklin, Wei Jiang and Quan Gao, for their friendship and encouragement, to Mr. Eddy Tang and Mrs. Alice Tang for their cooking of delicious food for us in the Liddle Court family church on Fridays, to Mr. John Smith and Mrs. Joan Smith, Mr. Tom Longhurst and Mrs. Rosemary Longhurst, Mr. Bruce Coles and Mrs. Sandra Coles and many other friends in the Liddle Court family church for their support and encouragement, especially when I was feeling low during the writing of my thesis, to my Chinese teaching colleagues, Meiping Zhu and Chunfang Wang for their kindness and invitations for dinner at home. Thanks to all my dear colleagues and friends, with you I no longer feel lonely and helpless in a foreign country. Also, many thanks are devoted to my previous colleagues in Harbin, Professor Xiuying Cai, Mrs. Liying Shen, and Mrs. Ling Yang, for their long support and encouragement since 2005.

I am especially grateful to those who helped me collect the data. They are Dan Zhao from Harbin 70th Middle School, Miaomiao Sun, Yanli Sun, and Hongyan Yan from Harbin 13th High School, Yanchun Yang and Ying Wu from HIP Harbin Institute of Petroleum, Jing Liu and Yimin Zheng from Guangzhou Yucai Middle School, Jingqi He from Guangzhou 6th High School, and Dr. Jianbo Chen, Wenyu Hu and many other teachers from Nanfang College at Sun Yat-sen University. In addition, I would like to thank the great participants from the above schools in Harbin and Guangzhou for their strong support for the data collection for up to one hour per person. Without their full help, there would be no data today, and there would be no findings in this thesis.

I am also thankful to those who should have been listed here for any encouragement, support and help given to me, but whose names I have forgotten to include.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Affr --- Affricate

Alv --- Alveolar

AoL --- Age of Learning

Appr --- Approximant

Asp --- Aspirated

Bilab --- Bilabial

C --- Consonant

Ch --- Chapter

C_N --- nasal consonant

C_P --- plosive consonant

CA --- Contrastive Analysis

CAH --- Contrastive Analysis Hypothesis

C&CP --- Character and Cantonese Pinyin

Den --- Dental

Diph-rhymes --- Diphthongal rhymes

EA --- Error Analysis

EFL --- English as a Foreign Language

F-F --- full-full (disyllable)

F-L --- full-light (disyllable)

F0 --- fundamental frequency

F1 --- the first formant

F2 --- the second formant

Fric --- Fricative

G --- Glide

GA --- General American

GH1 --- participant 1 from a high school in Guangzhou

Glt --- Glottal

GM1 --- participant 1 from a middle school in Guangzhou

GU1 --- participant 1 from a University in Guangzhou

GZ --- Guangzhou

GZC --- Guangzhou Cantonese

HH --- Heavy and Heavy (syllables)

HH1 --- participant 1 from a high school in Harbin

HL --- Heavy and Light (syllables)

HM1 --- participant 1 from a middle school in Harbin

HU1 --- participant 1 from a university in Harbin

IL --- Interlanguage Theory

IPA --- International Phonetic Alphabet

L1 --- First Language

$L1_C$ --- First Language Cantonese

$L1_H$ --- First Language Harbinese

L2 --- Second Language

$L2_E$ --- Second Language English

Lab-dent --- Labiodental

Lab-vel --- Labial-velar

LH --- Light and Heavy (syllables)

LoR --- Length of Residence

MCCD --- Multi-function Chinese Character Database

MDH --- Markedness Differential Hypothesis

MEPRC --- Ministry of Education of the People's Republic of China

ND --- Northeastern Dialect

OM --- Ontogeny Model

OT --- Optimality Theory

Pal --- Palatal

PAM --- Perception Assimilation Model

Q1 --- Question 1

RP --- Received Pronunciation

S --- Syllable

SW --- Strong and Weak (syllables)

SCPT --- Standard Chinese Proficiency Test

SLA --- Second Language Acquisition

SLM --- Speech Learning Model

T1 --- Tone one

T2 --- Tone two

T3 --- Tone three

T4 --- Tone four

T5 --- Tone five

T6 --- Tone six

T7 --- Tone seven

T8 --- Tone eight

T9 --- Tone nine

Unasp --- Unaspirated

V --- Vowel

Vel --- Velar

WS --- Weak and Strong (syllables)

Chapter 1. Introduction

1.1 Background

Chinese languages and particularly its variety, Mandarin Chinese, are spoken by the largest numbers of people as first language in the world (Enthlogue, 2020). However, English as a global lingua franca has been used both as a native language and is learned and spoken as a second or even a third language by huge numbers of people across the world. English is an important and compulsory subject in China and has been taught to students across all educational levels from nursery to university. Thus, there are considerable numbers of learners of English who speak Mandarin and other varieties of Chinese.

In second language learning, the interference of the first language (L1) on the second language (L2) is inevitable and significant (Weinreich, 1953; Lado, 1957; Corder, 1967; Fisiak, 1978, 1991; Gass, 1979; Young-Scholten, 1985; Anderson, 1987; Odlin, 1989; Ellis, 1994; Rasier and Hiligsmann, 2007). Lado (1957) claimed that, when learning an L2, if it is very similar to the learner's L1 it will be easier to acquire. On the other hand, if the L2 is very different from the L1, L2 learners will find it more difficult to learn. Because Chinese languages and English belong to two different language families, Chinese learners of English will, according to Lado's claim, encounter many difficulties in acquiring L2 English. Lado's Contrastive Analysis Hypothesis (1957) provides a powerful analytical method to compare the differences between two languages and to predict the difficulties L2 learners will have. A huge amount of relevant research has been carried out to compare different languages worldwide. Many researchers (Shi, 1996; Kong and Wang, 2001; Li, 2006; Wang, 2007; Yan, 2007; Gong and Tian, 2008; Peng, 2009; Yu, 2009) have also conducted research on the influence of L1 Chinese (usually but not always Mandarin) on L2 English phonology. Chinese has many dialects (or sublanguages) and each dialect/language has some similarities with, but many differences from others, especially in pronunciation. Thus, in studies of the influence of Chinese on English phonology it is essential to specify which Chinese dialect/language is involved. Due to the different influence of each dialect/language on English phonology, Chinese dialect/language speakers will pronounce English differently. It is reasonable to suppose that this behaviour is phonologically governed. Some research has been carried out to identify the differences in the influence of L1 dialects/languages on English. However, much of this research has been introductory and produces only summative descriptions of the production errors dialect/language speakers make. Few studies have conducted thorough and detailed experimental analysis.

However, in the early 1970s, a number of researchers and teachers (Whitman, 1970; see also Lightbown & Spada, 2006) criticised Lado's CAH (1957) and found that it could not predict all of the difficulties learners have and not all errors could be explained by L1 transfer. Flege (1995) then proposed an influential theoretical framework, the Speech Learning Model (SLM), which holds a position directly opposed to Lado's CAH on the likely difficulties experienced due to the similarities and differences between an L1 and L2. The SLM (Flege, 1995) predicts that the more similar L2 sounds are to L1 sounds, the more difficult it will be for L2 learners to perceive them, and thus they will have greater difficulties in acquiring those L2 sounds, and vice versa. That is to say, Flege (1995) thinks that 'new' is easy but 'similar' is difficult. This idea is supported by a series of empirical studies by Flege (1987, 1995, 1997) and Piske (2007, 2008). In addition, Flege emphasises the importance of phonetic perception. The SLM claims (Flege, 1995:238) that, without accurate perception of L2 speech, there will be no accurate production of L2 speech sounds. Moreover, the SLM (1997:17) posits that "interlingual identification occurs at the phonetic rather than the phonemic level". Therefore, Flege's SLM (1995) is used to predict learners' difficulties through the phonetic comparison of languages rather than Lado's phonemic comparison.

Consequently, a series of questions occur concerning whether similar sounds in two languages are easier to learn or more difficult. If it is easier, as Lado (1957) claims, what degree of similarity results in the learning of L2 phonology being easier? And if it is more difficult, as Flege claims, to what extent is similarity more difficult? However, few published studies can be found to answer these questions. Therefore, further research needs to be conducted in this area. The findings of such research can be used to determine whether the CAH or SLM is more helpful in predicting the potential difficulties L2 learners may have. Thus, it is necessary to conduct this type of research to attempt to fill this gap in knowledge.

In addition, personal curiosity prompted me to conduct research comparing these two Chinese dialects/languages with L2 English phonology. While teaching in different regions of China for approximately ten years, I observed that the students in the area where my own Northeastern dialect is spoken pronounce English differently from students in Cantonese-speaking areas. After consulting the literature, I found that differences in pronunciation can most probably be attributed to their influence.

Before we continue, it needs to be noted that the question of whether Cantonese and Harbinese Chinese varieties are dialects or languages. This is not straightforward. Here two principles are used to identify languages: one is mutual intelligibility and the other is group self-identification (Paolillo and Das, 2006:4). Mutual intelligibility is regarded as the more linguistically meaningful criterion (ibid, 2006:13). Relevant to the principle of mutual intelligibility, Cantonese and Mandarin differ for 80% of their phonology, 40% of their vocabulary and 20% of their grammar (Groves and Mair, 2008:11) can be indeed regarded as two languages. However, this principle can be problematic if we consider two types of examples. On the one hand, there are examples such Urdu and Hindi, generally considered to be two different languages despite being mutually intelligible in their spoken form (Paolillo et al., 2006:4). On the other hand, there are examples of varieties within one language which are not mutually intelligible: Newcastle, England and Atlanta, USA speakers of English "would have a great deal of difficulty to understand one another" (ibid). Paolillo and

Das (2006) also mention Arabic whose everyday spoken varieties are not in general mutually intelligible, and the formal variety, Modern Standard Arabic, is used in many countries. Linguistic factors such as mutual intelligibility are not the only means used to decide the boundaries between languages and dialects (Groves and Mair, 2008). Political, religious, literacy (writing systems) and social factors are involved, an example of which is Urdu, whose speakers are Muslim and which uses the Perso-Arabic script, and Hindi which uses the Devanagari script, whose speakers are Hindu. All varieties of Chinese use the same writing system, Hanzi. This isn't completely straightforward because closely related varieties of German and Dutch and less closely related standard German and Swiss German also use the same writing systems. Yet Dutch is considered a language while Swiss German is considered a variety/dialect. This points to the situation with varieties of Chinese: they are considered dialects rather languages --- essentially for political and socio-cultural reasons. They (ibid, 2008:9) state “Chinese situation is unique in that it represents a different kind of exception to the mutual intelligibility principle”.

In general, Chinese is regarded as one general language representing the same national identity, the unity of Chinese culture and the same nature of the Chinese Han race. In this case, Cantonese and other varieties of Chinese are always considered as its dialects. However, if Chinese is considered as a Chinese language group with a collection of languages, its varieties can be called different languages.

Comparing Cantonese and Harbinese as dialects of the same language was the starting point of this thesis, but whether they are dialecsts or languages is actually not crucial in looking at the the impact of the differences between Harbin Mandarin and Guangzhou Cantonese, which are sufficiently similar to see how their similarities to and differences from English result in different errors. Since the two are conventionally referred to as dialects, when it is not possible to refer to them in a more neutral way, I will continue in the present thesis to refer to them as such.

The present thesis seeks to answer this question: what special features in English are exhibited by speakers of closely-related Cantonese and Harbinese-Mandarin, that is, what error patterns do they display and what specific errors do they make? Although some literature exists on these two Chinese varieties, they tend not to be compared but researched separately. Also, unfortunately, less attention has been paid to the Northeastern dialect. There is very little research conducting comparative analysis of one language (or language group), namely Chinese but two dialects (Northeastern dialect and Cantonese) with L2 English phonology. Thus, again, such research needs to be carried out to fill this gap.

There are regional differences whereby a language varies linguistically, especially in phonology. Thus, the choice of research locations for data collection in this thesis was considered carefully. Harbin and Guangzhou were selected because these two cities are capitals of their respective provinces, they are also economic centres in their regions, and each has a large population. The dialects/languages used in the two cities are representative of the dialect groups (language groups) in the regions. Also, the two cities are from the north and the south of China, and are separated by a distance of approximately 3320 kms (measured by Google Map, 2019), which adds to the interest in determining the extent of differences between these two distant dialects/languages. Thus, the sound systems of the dialects/languages under study are based on the dialects/languages spoken in these two cities, namely, Harbinese and Guangzhou Cantonese.

1.2 Aims of the research

In the light of the background to the study, research needs to be conducted to uncover what is going on. As mentioned above, CAH has been criticised since the 1970s. Although the CAH itself (Lado, 1957) is now considered not perfect, contrastive analysis as a research method is still useful and beneficial for teachers (Klein, 1986:26; Mitchell and Myles, 2004:31-32) and is widely applied in different domains by many researchers (Ajmer, Alenberg and Johansson, 1996; Belz, 2003; Blyth, 1995; Johansson and Oksefjell, 1998; Kramsch, 2007; Toury, 1995 and Venuti, 2000, cited by Kramsch, 2007:245). Thus, to avoid the limitations associated with each theoretical framework, this research combines Lado's CAH (1957) with Flege's SLM (1995). In addition, it also takes into account comparisons between L2 and L1 sounds at both the phonetic and phonemic levels.

Therefore, the research first determines the differences between two Harbinese Mandarin and Guangzhou Cantonese and then relates the differences between the two to L2 English production in terms of segments, syllable structure and stress. Based on the prediction from Flege's SLM (1995) that 'similar' is difficult and 'new' is easy, hypotheses are proposed concerning the predictions of difficulties encountered by L2 English learners. The testing instrument include word translations and Chinese reading were designed to investigate these difficulties. 174 items were given to participants to produce in English. The testing instrument aimed to find out the answers of the first two bulleted questions below, and the details can be seen in section 3.5.3 Testing Instrument of Chapter 3. Flege's SLM is applied in this thesis to the second two bullet points below.

- how the speakers of Guangzhou Cantonese and Harbinese Mandarin acquire English phonology;
- what features and error types speakers of each group make;
- which is more difficult between L1-L2 'similar' or 'new';
- which theoretical framework provides more help to L2 English learners and teachers.

1.3 Contributions of the research

Research on the influence of the L1 on L2 has been conducted by many researchers (Lado, 1957; see also Anderson, 1987; Corder, 1967; Ellis, 1994; Fisiak, 1978, 1991; Gass, 1979; Odlin, 1989; Young-Scholten, 1985), but research into the influence of different dialects in the same language on the L2 is quite limited. Therefore, this thesis can first counter the lack of research in this domain especially on the influence of Chinese dialects on the acquisition of English phonology. Moreover, only a limited amount of research can be found which compares predictions of Lado's 'similar' as easy vs. Flege's 'similar' as difficult. The study of two Chinese dialects/languages and L2 English production not only enriches the relevant literature for each dialect/language and research into second language acquisition (SLA), but can also be used to provide pedagogical guidance in the decision of future educational strategies and assistance for teachers and learners in local regions.

1.4 Layout of the thesis

This thesis is composed of six chapters including this introduction.

Chapter two introduces the linguistic background of two target Chinese dialects/languages, makes comparisons between Harbinese and Guangzhou Cantonese in terms of segments, syllable structure and stress, and selects potential target sounds for the study. Afterwards, comparisons of each Chinese variety with English are conducted so as to predict the potential difficulties Harbinese and Guangzhou Cantonese speakers may have. In addition, the literature on L2 acquisition and its development as well as the theoretical frameworks of the CAH and SLM are reviewed and previous relevant research is discussed.

Chapter three describes the methodology used in the study. It introduces the classification criteria for targets, and proposes hypotheses. Based on the differences between the two dialects/languages and English, predictions are made in the form of hypotheses using Flege's difficulty criteria. Detailed accounts are also given of the selection of participants, instrument design, and data collection and analysis.

Chapter four presents the results of the study. It first displays results of the analysis of questionnaire data, which help in the testing of certain hypotheses and help in the response to the use of an English variety. Subsequently, the full results for the production of all targets are presented. Summaries of the features of the error types made by Harbinese and Guangzhou Cantonese speakers are also presented. The findings are then used to answer all hypotheses in this study.

Chapter five comprises a series of discussions. It first discusses the influence of the L1 on L2 and relates the findings of this study with those of previous research. The influence of Mandarin on Cantonese speakers of L2 English phonology is also considered. The application of Lado's CAH and Flege's SLM is then discussed, and an explanation is provided of why Flege's prediction of difficulty in the SLM does not apply in this research. The chapter also discusses the inconsistent conclusions of hypotheses in the same categories and the difficulty different types of 'similar' may have. Finally, the significance of the English varieties used by different dialect/language speakers is considered and pedagogical implications of the thesis are listed.

Chapter six presents the conclusions of the whole thesis and the limitations of the study as well as giving recommendations for further research which needs to be conducted in the future.

Chapter 2. Background and Literature Review

This chapter explains the linguistic background of two target Chinese dialects/languages, Harbinese and Guangzhou Cantonese, and compares them in terms of segments, syllable structure and stress in order to select potential target sounds. It also compares each Chinese variety with English to provide predictions of potential difficulties Harbinese speakers and Guangzhou Cantonese speakers may have. In addition, it reviews the literature on L2 acquisition, highlighting two theoretical frameworks (CAH and SLM), and describes previous relevant research.

2.1 Introduction to the two sound systems

2.1.1 The sound system of Harbinese

The Northeastern dialect (ND) is one of the varieties of Mandarin belonging to one of the seven major dialect groups in China, and is widely spoken across the Northeastern part of China in the provinces of Heilongjiang, Jilin and Liaoning, as well as parts of Inner Mongolia. On the whole, it is close to Standard Mandarin in pronunciation, but it has its own characteristics in vocabulary and in the tones of some specific words. There are also variations in the Northeastern dialect associated with particular cities and towns, but they all share substantial similarities and have common features from a linguistic perspective. The Harbin dialect is one of the sub-varieties of the Northeastern dialect. The sound system of the Northeastern dialects demonstrated in this thesis is based on that of the Harbin dialect because Harbin is a representative city in the Northeast of China and its dialect is a typical variety of the Northeastern dialect. It is also a typical representative variety in the Mandarin dialect group. For convenience, the Harbin dialect is called Harbinese in the rest of this thesis.

The sound systems of all Chinese varieties are composed of what are called *initials* and *finals*, and *tones*. Unlike the conventional linguistic terms (*onset* and *coda*), *initials* and *finals* are traditional Chinese terms and are still used and taught in this way at school.

Initials are initial consonants of syllables and *finals* are the remainder of the syllables after *initials*. *Finals* are composed of the *final* head, *final* body and *final* tail. The *final* head is treated as a vowel in the traditional Chinese sound system, and can also be called a medial vowel or a glide in formal linguistics. The *final* body is the main vowel of *finals*, which in formal linguistics is identified as the nucleus. The *final* tail can be a vowel or a consonant. If the *final* tail is a consonant, it is called coda linguistically. The *final* head and *final* tail are optional in Chinese syllables but the *final* body, that is, the nucleus, is structurally obligatory. There are, however, parallels between *initials* and onsets and *finals* and vowels plus codas, i.e. rhymes. The sound system of Harbinese is illustrated likewise.

2.1.1.1 Consonants

Twenty-three consonants are allowed to occur in the initial position in Harbinese if the two semi-vowels /j/ and /w/ are included. There are 3 bilabial initials, /p/, /pʰ/, /m/, 1 labio-dental, /f/, 7 alveolar, /t/, /tʰ/, /t͡s/, /t͡sʰ/, /s/, /n/, /l/, 4 retroflex, /t͡ʂ/, /t͡ʂʰ/, /ʂ/, /ɻ/, 4 palatal, /t͡ɕ/, /t͡ɕʰ/, /ɕ/, /j/, 3 velar /k/, /kʰ/, /x/ and 1 labio-velar initial /w/. As observed by Lin (2007) and Duanmu (2007), Chinese plosives, fricatives, and affricate consonants do not contrast in voicing but in aspiration, and thus they are all voiceless consonants. It is important here to note the transcription of the consonant *r* in Mandarin/Harbinese. This is differently transcribed in IPA by researchers as /ɹ/, /r/, /ʐ/ and /ɻ/ in the literature, but /ʐ/ and /ɻ/ are used more often. The symbol /ɹ/ is often used to represent the *r* consonant in English. Clearly, the *r* sound in English is different from that in Harbinese/Mandarin, and thus it is not appropriate to use the symbol /ɹ/ for the *r* sound in Harbinese/Mandarin. In addition, Harbinese/Mandarin has no trill consonant, so /r/ cannot be used here either. As Lin and Duanmu note, Mandarin does not contrast the voicing in stops, fricatives, and affricates but only contrasts the aspiration, which means stops, fricatives, and affricates in Mandarin are all voiceless; thus, /ʐ/ as a voiced retroflex fricative paired with /ʂ/ in voicing in IPA chart violates this constraint. Therefore, I think the approximant /ɻ/ is the best symbol to transcribe the Harbinese/Mandarin retroflex *r* sound rather than /ʐ/. In this research the retroflex *r* sound is represented by /ɻ/.

The consonant sound system of Harbinese is the same as Standard Mandarin. All these onsets can be seen clearly in Table 2.1 adapted slightly from Lin (2007:41).

Table 2.1 Consonants of Harbinese/Standard Mandarin

		Bilabial		Labio dental		Alveolar		Retroflex		Palatal		Velar	
Plosive	Unasp	p				t						k	
	Asp	pʰ				tʰ						kʰ	
Affricate	Unasp					t͡s		t͡ʂ		t͡ɕ			
	Asp					t͡sʰ		t͡ʂʰ		t͡ɕʰ			
Fricative				f		s		ʂ		ɕ		x	
Nasal			m				n		ŋ				
Approximant			(w)				l		ɻ		(j)		(w)

(Unasp: unaspirated; Asp: aspirated)

Source: adapted from Lin (2007:41)

2.1.1.2 Vowels

As a Mandarin dialect, Harbinese has significant similarities with Standard Mandarin, even in its sound system. Therefore, they are closely related. In addition, given its official status Standard Mandarin has been amply researched, which provides sufficient background for an introduction to Harbinese. Thus, the vowel inventory of Standard Mandarin is introduced first before describing Harbinese.

Standard Mandarin is the official variety of Chinese used in China, and it has fewer vowel phonemes than Cantonese. According to Lin (2007:82), the phonemic vowel inventory of Standard Mandarin has only five vowels: the three high vowels /i, y, u/, a mid-vowel /ə/, and a low vowel /a/ as seen in Figure 2.1.

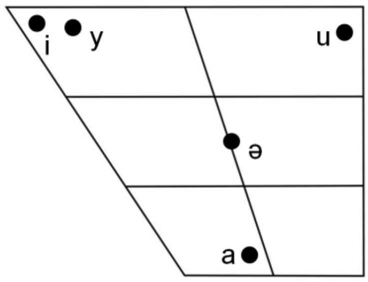


Figure 2.1 Phonemic vowels in Standard Mandarin

Lin (ibid.) stated that in Standard Mandarin the three high vowel phonemes not only have the allophones [i], [y], [u], but can as also having the glides [j], [ɥ], and [w] which can only be followed by a mid or low vowel. In this situation, they are allophones of /i, y, u/. Examples are shown in (2.1) as follows:

(2.1)

a. High vowel phonemes /i, y, u/ as vowels [i, y, u]

Character	Pinyin	IPA	Meaning
弟	dì	[t**i**T4]	'younger brother'
句	jù	[tɕ**y**T4]	'sentence'
湖	hú	[x**u**T2]	'lake'

b. High vowel phonemes /i, y, u/ as glides [j, ɥ, w]

Character	Pinyin	IPA	Meaning
跳	tiào	[t^h**j**ɑuT4]	'jump'
选	xuǎn	[ɕ**ɥ**ɛnT3]	'choose'
瓜	guā	[k**w**aT1]	'melon'

Meanwhile Lin claimed (2007:74-76) that Standard Mandarin has only one mid-vowel phoneme, and that is /ə/ which has four vowel allophones which are [e], [ə], [ɤ] and [o], occurring in complementary distribution. The vowel allophone [e] always occurs in a diphthong preceding [i], [ə] is only present before the nasal coda consonant [n]/[ŋ], and [ɤ] can only occur as a nucleus in a (C)V (Consonant-Vowel) syllable, while the rounded vowel allophone [o] can be followed by [u] in a diphthong or be preceded by a glide [w] in the final position of a syllable. Examples of these allophonic vowels for the

Standard Mandarin mid-vowel phoneme /ə/ are indicated in (2.2):

(2.2)

a. Allophone [e] in a diphthong preceding [i]

Character	Pinyin	IPA	Meaning
北	běi	[p**e**iT3]	'north'
内	nèi	[n**e**iT4]	'inside'

b. Allophone [ə] before the nasal consonant [n]/[ŋ]

Character	Pinyin	IPA	Meaning
恩	ēn	[**ə**nT1]	'favour'
灯	dēng	[t**ə**ŋT1]	'light, lamp'
拼	pīn	[p^hj**ə**nT1]	'to spell'
名	míng	[mj**ə**ŋT2]	'name'
吞	tūn	[t^hw**ə**nT1]	'to swallow'
翁	wēng	[w**ə**ŋT1]	'light, lamp'

c. Allophone [ɤ] as a nucleus in (C)V syllables

Character	Pinyin	IPA	Meaning
鹅	é	[**ɤ**T2]	'goose'
特	tè	[t^h**ɤ**T4]	'special'

d. Allophone [o] in diphthong followed by [u] or in syllable final position preceded by a glide [w]

Character	Pinyin	IPA	Meaning
豆	dòu	[t**o**uT4]	'bean'
猴	hóu	[x**o**uT2]	'monkey'
火	huǒ	[xw**o**T3]	'fire'
桌	zhuō	[ʈʂw**o**T1]	'desk'

In addition to these phonemes, in Standard Mandarin there is a low vowel phoneme, as Lin mentioned (2007:77), which is /a/. This has three different phonetic surface representations in terms of the position where it occurs. The phonetic vowel [ɑ] appears before a high vowel [u] in a diphthong or before the velar nasal coda [ŋ] . The vowel occurring between glides [j]/[ɥ] and the dental nasal coda [n] is another phonetic representation [ɛ]. However, in other contexts the phonetic form of /a/ is [a]. The phonetic low vowels [ɑ], [a] and [ɛ] are allophones of the low vowel phoneme /a/ represented in a complementary distribution. The examples in (2.3) indicate the three allophones of the low vowel phoneme /a/ occurring in different distributions.

(2.3)

a. Allophone [ɑ] before high vowel [u] or before the velar nasal coda [ŋ]

Character	Pinyin	IPA	Meaning
桃	táo	[tʰ**ɑ**uT2]	'peach'
鸟	niǎo	[nj**ɑ**uT3]	'bird'
昂	áng	[**ɑ**ŋT2]	'to raise'
房	fáng	[f**ɑ**ŋT2]	'house'

b. Allophone [ɛ] between glides [j]/[ɥ] and the dental nasal coda [n]

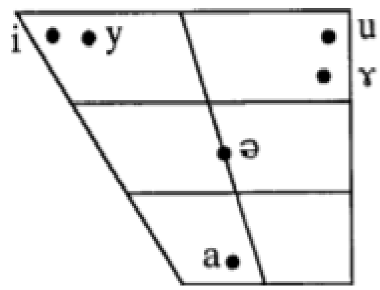
Character	Pinyin	IPA	Meaning
烟	yān	[j**ɛ**nT1]	'cigarette'
田	tián	[tʰj**ɛ**nT2]	'field'
远	yuǎn	[y**ɛ**nT3]	'far'
劝	quàn	[tɕʰy**ɛ**nT4]	'to persuade'

Lin (2007:78) also claims that Standard Mandarin has only four diphthongs, [ai], [ei], [ɑu] and [ou], which are composed of variants of Standard Mandarin's five vowel phonemes.

However, it seems that researchers have not arrived at consensus concerning Standard Mandarin's vowel system, because it is represented differently by different researchers. According to Lee and Zee (2003:110), the vowel sound inventory in Standard Mandarin consists of six monothongs, /i, y, u, ɤ, ə, a/, eleven diphthongs, /ai, au, ou, uo, ei, ye, ie, ia, ua, uə, iu/, and four triphthongs, /iau, iou, uai, uei/, as seen in Figure 2.2.

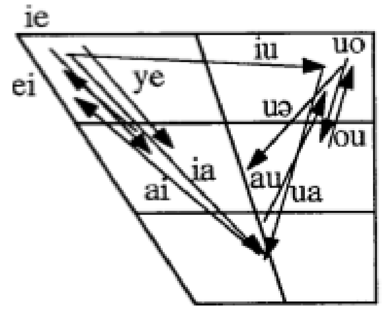
a.

Standard Mandarin Monophthongs



b.

Standard Mandarin Diphthongs



c.

Standard Mandarin Triphthongs

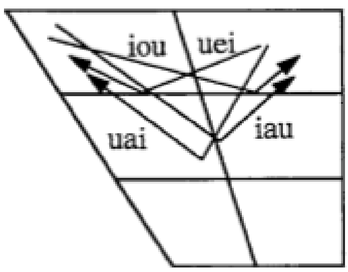


Figure 2.2 Vowel charts of Standard Mandarin

It can be seen that this classification is possibly based on the Chinese Pinyin system, which is a letter-based writing system for Chinese pronunciation, because the six monophthongs /i, y, u, ɤ, ə, a/ accurately correspond to the six vowel letters in Chinese Pinyin, 'i, y, u, e, o, a'. However, [ɤ] should be an allophone of /ə/ in terms of Lin's analysis because [ɤ] and [ə] occur in complementary distribution. [ɤ] can only be a rhyme in (C)V syllables and also it can never occur in the position of [ə] to distinguish meanings as a phoneme. Therefore, Lin's description is more convincing because [ɤ] is an allophone but not a phoneme.

Lee and Zee (2003:110) state that there are 11 diphthongs in Standard Mandarin. This is much more than the number Lin suggests, with the ratio of 9 vs. 4. Four diphthongs [ai, au, ou, ei] coincide with Lin's diphthongs, and these are indisputable in all research into the Standard Mandarin sound system. However, in Standard Mandarin, under the influence of feature backing when the phoneme /a/ connects with the back vowel [u] and back velar nasal [ŋ], the /a/ is assimilated by them to the back low vowel [ɑ]. Therefore, the diphthong [au] here presented by Lee and Zee is not very accurate phonetically. The correct transcription for that diphthong should instead be [ɑu] in Standard Mandarin. Similarly, for the phoneme /a/ connecting with the front vowel [i], assimilation occurs and the front [a] should be adopted to form the diphthong [ai]. This is accurately described by Lee and Zee. Apart from these four diphthongs, the status of the following five [ie, ia, ua, uo, ye] is debatable. These diphthongs share the characteristic that they are diphthongs with glides [i, u, y] preceding the nucleus vowels. Lin (2007:80) does not regard them as diphthongs because she thinks that both vowels do not occupy the nucleus position. To take [ɑu] as an example, the two vowels together occupy the nucleus position from Lin's perspective but in [ia] they do not. In addition, Lin claims that the /i, y, u/ before the nucleus are glides [j], [ɥ], [w], which are considered by her as parts of the onset rather than one of the vowels in the syllable rhymes. As opposed to Lin's idea, although [j], [ɥ], [w] are glides, some think that they are still vowels which belong to one of the parts of rhymes, and therefore the glides [j], [ɥ], [w] are always represented in transcription as [i], [y], [u] by these researchers. The status of /i, y, u/ as glides to onset or rhyme is still controversial, with different understandings generating distinct claims for the Standard Mandarin vowel phonetic inventory as mentioned earlier in the discussion. The last two diphthongs [uə, iu] mentioned by Lee and Zee are not called diphthongs by other researchers. These two seem to be extracted from the *finals* or so-called rhymes [uən]/[uəŋ] and [iuŋ]. They may not be considered as diphthongs because, phonologically, the nuclei [ə] and [u] in these *rhymes* connecting with its post nasal-coda [n] or [ŋ] as a constituent closer than to their preceding glides. Therefore, that these two are diphthongs seems not to be reasonable and authentic in Standard Mandarin.

Lee and Zee (2003:110) also describe four triphthongs in Standard Mandarin, which is in accordance with the traditional view. Lin disagrees, and does not think four triphthongs exist in Standard Mandarin because she puts the glides [i, y, u] to the side of the onset.

The Standard Mandarin vowel inventory presented in the *Standard Chinese Proficiency Test* (Meng et al., 2001:28) is also different from those in the two studies mentioned above. This suggests instead that the sound system of Standard Mandarin has 39 of what are referred to in the literature as *finals*, i.e. as rhymes. These consist of 10 single rhymes, 13 compound rhymes and 16 nasal rhymes, as seen in Table 2.2. The 10 single rhymes are monophthongs, and the 13 compound rhymes are vowel cluster rhymes including 9 diphthongs and 4 triphthongs. The 16 nasal rhymes here refer to those ending with a nasal coda, which are not discussed further here because of their irrelevance to the vowel system.

Table 2.2 Rhymes in Standard Mandarin

Mouth Shape / Internal Structure	Open-Mouth Non- **[i] [u] [y]**	Close-Teeth i [i]	Round-Lip u [u]	Tense-Lip ü[y]	Final Head / Final Tail
Single Rhymes	(z)i [ɿ] (zh)i [ʅ]	i [i]	u [u]	ü [y]	Non-Final Tail
	a [a]	ia [ia]	ua [ua]		
	o [o]		uo [uo]		
	e [ɤ]				
	ê [ɛ]	ie [iɛ]		üe [yɛ]	
	er [ɚ]				
Compound Rhymes	ai [ai]		uai [uai]		Vowel Final Tail
	ei [ei]		uei (ui) [uei]		
	ao [ɑu]	iao [iɑu]			
	ou [ou]	iou (iu) [iou]			

Nasal Rhymes	an [an]	ian [iɛn]	uan [uan]	üan [yɛn]	Nasal Final Tail
	en [ən]	in [in]	uen (un) [uən]	ün [yn]	
	ang [ɑŋ]	iang [iɑŋ]	uang [uɑŋ]		
	eng [əŋ]	ing [iŋ]	ueng [uəŋ]		
	ong [uŋ]	iong [iuŋ]			

Source: adapted from *Standard Chinese Proficiency Test* (Meng et al. 2001:31)

Table 2.2 adapted from *Standard Chinese Proficiency Test* is a conventional way to introduce the Chinese vowel system. It can be seen from Table 2.2 that the ten monophthongs of Standard Mandarin are: [ɿ], [ʅ], [i], [u], [y], [a], [o], [ɤ], [ɛ] and [ɚ], and its nine diphthongs and four triphthongs are respectively: [ai], [ei], [ɑu], [ou], [ia], [ua], [uo], [iɛ], and [yɛ] and [iɑu], [iou], [uai], and [uei]. Obviously, they are phonetically represented, which is distinct from the phonemic descriptions given by other researchers such as Lin (2007). The four monophthongs [i], [u], [y], [a] also occur in Lin's description as well as Lee and Zee's , while [o] and [ɤ] are regarded by Lin as allophones of /ə/ and [ɛ] as one of the allophones of /a/. However, the rest of the monophthongs are not included by Lin and Lee in their vowel sound system of Standard Mandarin. The two monophthongs [ɿ] and [ʅ] in the *Standard Chinese Proficiency Test* can only occur phonotactically, respectively after alveolar fricative and affricate initials [ts], [ts^{h}], [s] for [ɿ] and after the retroflex initials [ʈʂ], [ʈʂ^{h}], [ʂ] and [ɻ] for [ʅ] as rhymes of open syllables. However, Lin and Duanmu do not think they are vowels in that position but consider them to be syllabic consonants. Also, Duanmu (2000:36) uses the transcriptions [z] and [r] to replace [ɿ] and [ʅ] respectively although [r] indicates a trill which is not the case, while Lin (2007:72) only exploits the syllabic consonant [ɻ̩] instead. Examples are shown in (2.4) for clearer illustrations of the differences between these researchers.

(2.4)

Character	Pinyin	Meaning	SCPT IPA	Duanmu IPA	Lin IPA
字	zì	'character'	[tsɿT4]	[tszT4]	[tsɹ̩T4]
词	cí	'word'	[ts^{h}ɿT2]	[ts^{h}zT2]	[ts^{h}ɹ̩T2]
死	sǐ	'to die'	[sɿT3]	[szT3]	[sɹ̩T3]
纸	zhǐ	'paper'	[ʈʂʅT3]	[ʈʂrT3]	[ʈʂɹ̩T3]
吃	chī	'to eat'	[ʈʂ^{h}ʅT1]	[ʈʂ^{h}rT1]	[ʈʂ^{h}ɹ̩T1]
是	shì	'is/am/are'	[ʂʅT4]	[ʂrT4]	[ʂɹ̩T4]
日	rì	'day'	[ʐʅT1]	[rrT1]	[ɹɹ̩T1]

Apart from these nine monophthongs, the last one is transcribed as [ɚ] in the *Standard Chinese Proficiency Test*, but it is also transcribed as [ər] or [ar], or even the consonant [r], by other researchers (Duanmu, 2000:41; Lin 2007:80-81) who call it a retroflexed or rhotacised vowel. It occurs in a zero-onset syllable rhyme or it suffixes words to replace the syllable codas. In the former situation, it is always transcribed as [ər], [ar] or [ɚ] but as a suffix its transcription is always abbreviated to [r] instead as seen in the examples in (2.5a). In addition, a special case is that *èr* in Chinese has only one corresponding character, 二, meaning the number 'two'. In Standard Mandarin, it has only one pronunciation, which is [arT4]; therefore, the transcriptions [ɚT4]/[ərT4] for *èr* 'two' are not accurate. They may exist in other Mandarin dialects but not in Standard Mandarin. Apart from [ar] only for *èr* 'two', all other *er* syllables can be transcribed as [ər] or [ɚ] as seen as (2.5b). However, as a Standard Mandarin speaker, I think [ər] is a better transcription because [ər] has a clear transition from [ə] to [r]. As mentioned above about the transcription of *r* sound in Mandarin (Harbinese), it is transcribed as [ɻ] in this thesis, thus, [ər] should be replaced by [əɻ] and all examples with *r* sound for Mandarin (Harbinese) below are transcribed as [ɻ].

(2.5)

a. As Zero-onset syllable rhyme

Character	Pinyin	Meaning	IPA	Rhotacised
儿	ér	'son'	[əɻT2]	[ɚT2]
耳	ěr	'ear'	[əɻT3]	[ɚT3]
二	èr	'two'	[aɻT4]	([ɚT4])

b. As a suffix

Character	Pinyin	Meaning	Syllabic Con IPA
花儿	huār	'flower'	[xuaɻT1]
孩儿	háir	'child'	[xaɻT2]

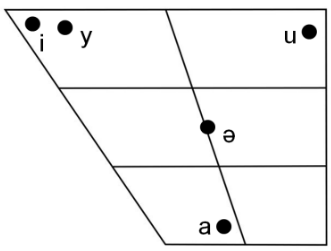
Duanmu stated (2000:42) that, the abbreviated retroflex vowel [r] he used and the consonant [r] are phonetically different but they are the same phonologically because they are in complementary distribution: the former occupies a coda position while the latter is in onset position. Therefore, he reckoned they are allophones of the same phoneme /r/.

The compound *rhymes* in the *Standard Chinese Proficiency Test* are composed of nine diphthongs and four triphthongs which are conventionally classified into three types based on sonority as front sonorant, back sonorant, or central sonorant compound rhymes. The front sonorant means that the most sonorant vowel in compound rhymes is in the front, such as in the diphthongs [**a**i], [**e**i], [**ɑ**u] and [**o**u]. The back sonorant has the most sonorant vowel at the back as in the diphthongs [i**a**], [i**ɛ**], [u**a**], [u**o**] and [y**ɛ**]. The central sonorant has the most sonorant vowel in the centre, for instance, in [i**ɑ**u], [i**o**u], [u**a**i], [u**e**i]. Front and back sonorant compound rhymes refer to diphthongs, but the central ones represent triphthongs because only if there are at least three vowels can there be a middle one. It is reasonable when explaining diphthongs and triphthongs to note that Standard Mandarin is conventionally considered to have not only front diphthongs [**a**i], [**e**i], [**ɑ**u] and [**o**u], but also the back diphthongs [i**a**], [i**ɛ**], [u**a**], [u**o**], [y**ɛ**], which Lin (2007:79) called rising diphthongs and which should be represented by [j**a**], [j**ɛ**], [w**a**], [w**o**], [ɥ**ɛ**] in Lin's description, and triphthongs [i**ɑ**u], [i**o**u], [u**a**i], [u**e**i] whose glides /i, y, u/ are in front of the nucleus. However, unlike the conventional classification, some researchers (Lin, 2007:78-80; Duanmu, 2000:42) do not consider back sonorant diphthongs and triphthongs to be real diphthongs and triphthongs in Mandarin because they do not regard the glides /i, y, u/ in back sonorant diphthongs and triphthongs as vowels.

So, there is no unanimously accepted Standard Mandarin vowel system, and the principal discrepancy is the belonging of glides to onsets or rhymes. This is still a controversial issue. Therefore, I will adopt the merits of each view which are reasonable for this study. I assume that there are five phonemic vowels in Standard Mandarin but at least ten phonetic vowels if the debatable two vowels [ɿ] and [ʅ] and the retroflex vowel [ɚ] (which can also be transcribed as ([ər]) are not taken into consideration. They are illustrated in Figure 2.3.

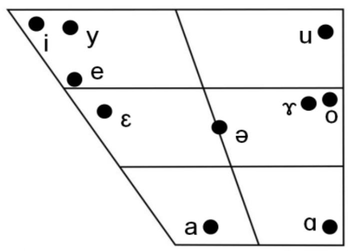
a.

Phonemic Vowel Inventory of Standard Mandarin:



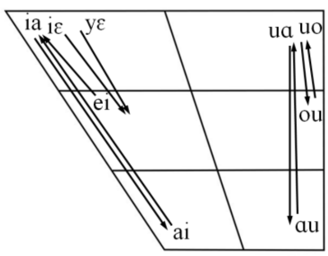
b.

Phonetic Vowel Inventory of Standard Mandarin:



c.

Standard Mandarin Diphthongs:



d.

Standard Mandarin Triphthongs:

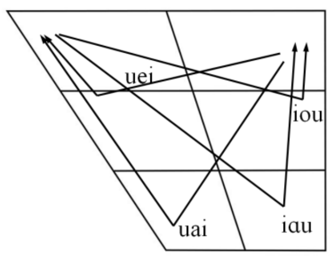


Figure 2.3 Vowel inventory of Standard Mandarin

Harbinese is a variety of the Northeastern dialect which belongs to the Mandarin dialect group. It is commonly acknowledged that Harbinese is the closest variety of Mandarin dialects to Standard Mandarin, much closer even than the Beijing Mandarin dialect (Nie, 2005:35). The vowel system of Harbinese is very similar to that of Standard Mandarin. The only difference between the two is that, in Harbinese, the vowel reflex of the Mandarin [o] is [ɤ], as indicated in the examples in (2.6).

(2.6)

[ɤ] replacing [o] in Harbinese as an open syllable rhyme

Character	Pinyin	Meaning	S-Mandarin	Harbinese
波	bō	'wave'	[poT1]	[pɤT1]
破	pò	'broken'	[p^hoT4]	[p^hɤT4]
摸	mō	'to touch'	[moT1]	[mɤT1]
佛	fó	'budda'	[foT2]	[fɤT2]

However, [o] exists in the diphthongs [ou] and [uo] when connecting with the rounded [u]. Some researchers, nevertheless, have proposed (Nie, 2005:67) that [ou] and [uo] should be [ɤu] and [uɤ] in Harbinese because there is no [o] monophthong but instead [ɤ]. As a native speaker of Harbinese, I do not agree. Moreover, with the constraint of the assimilation process of feature backing, it is true that [ou] exists in Harbinese, but [uo] sounds to be slightly inclined to [uɤ] in perception. Because it is extremely difficult to judge whether it is [uo] or [uɤ]. Perhaps, it can be resolved with acoustic analysis, but it is not within the scope of my research. Therefore, in this research I consider it to be [uo] rather than [uɤ] with the support of backing assimilation in Standard Mandarin. In conclusion, Harbinese has very similar phonemic and phonetic vowel sound system to Standard Mandarin; that are nine monophthongs, nine diphthongs and four triphthongs in Harbinese. However, [o] is replaced by [ɤ] as the rhyme of an open syllable in Harbinese, which is distinct from Standard Mandarin. Other differences between Harbinese and Standard Mandarin are not relevant to the vowel inventory, and are not discussed here.

2.1.1.3 Tones

Although tone is not one of the factors investigated in the present thesis, it is useful to know the role it plays in the varieties of Chinese. In Harbinese, the number of tones is identical with those in Standard Mandarin, with four lexical tones and a neutral tone which originates from the loss of the lexical tones. The differences between Standard Mandarin and Harbinese concern the pitch height of the tones. According to Nie (2005), all pitch values of the four tones in Harbinese are comparatively lower than in Standard Mandarin, which is a typical characteristic of Harbinese distinguishing it from Standard Mandarin. The four tones are called *Yingping*, the high level tone (abbreviated as T1), *Yangping*, the rising tone (T2), *Shangsheng*, the falling-rising tone (T3), and *Qusheng*, the falling tone (T4). The pitch height of each tone can also be seen in Table 2.3 extracted from Nie (2005).

Table 2.3 Tones of Harbinese

Tone Category	Yinping Level	Yangping Rising	Shangsheng Falling-rising	Qusheng Falling
Tone Code	1	2	3	4
Pitch Height	44	24	213	42

Source: Nie (2005)

2.1.2 The sound system of Guangzhou Cantonese

Cantonese is a dialect/language in one of the seven major Chinese dialect/language groups mainly spoken in Guangdong province and some parts of Guangxi province and in Chinese communities overseas. The term Cantonese first developed from Canton (the old English name for Guangzhou city) along the Pearl River Delta thousands of years ago and Guangzhou Cantonese is considered to be the prestige variety, the standard Cantonese, since ancient times. Therefore, the sound system of Cantonese is based on Guangzhou Cantonese.

2.1.2.1 Consonants

Guangzhou Cantonese has 19 consonant phonemes as shown in Table 2.4 (Gui, 2005:65). It consists of three bilabial initials, /p/, /p^h/, /m/, one labio-dental, /f/, seven alveolars, /t/, /t^h/, /ts/, /ts^h/, /s/, /n/, /l/, one palatal, /j/, three velars /k/, /k^h/, /ŋ/, three labio-velars, /k^w/, /k^{hw}/, /w/, and one glottal /h/. Like Standard Mandarin and Harbinese, Cantonese as one of varieties of Chinese does not contrast voicing in plosive, fricative and affricative sounds but in aspiration because all plosives, fricatives and affricatives in Cantonese are voiceless (ibid.). Each of the Cantonese dental phonemes /ts/, /ts^h/, /s/ have a pair of allophones which are respectively dental allophones [ts], [ts^h], [s] and palatalised allophones [tʃ], [$tʃ^h$] and [ʃ] (Bauer and Benedict, 1997). These pairs of allophones are in complementary distribution. Song and Yu (2007:1-2) claim that [ts], [ts^h], [s] are followed by [a:], [ɔː], [ɛː], [ɐ], [œː] and [uː] but [tʃ], [$tʃ^h$] and [ʃ] arise in front of [i:], [y:].

Table 2.4 Initials of Guangzhou Cantonese

		Bilabial		Labiodental		Dental/ alveolar		Retro -flex		Palatal		Velar		Labialised velar		Glottal	
Plos	Unasp	p				t						k		k^w			
	Asp	p^h				t^h						k^h		k^{hw}			
Affr	Unasp					ts				*tʃ							
	Asp					ts^h				*$tʃ^h$							
Fricative				f		s				*ʃ						h	
Nasal			m				n						ŋ				
Approximant			w				l				j				w		

* allophones of /ts, ts^h, s/

Source: adapted from Gui (2005:65)

2.1.2.2 Vowels

In literature, researchers present the Cantonese vowel sound system in different ways. Zee (1999:59) states that Cantonese has 11 monophthongs and 10 diphthongs which are respectively: /a, ɐ, ɛ, e, i, ɔ, o, œ, ɵ, u, y/ and /ai, au, ɐi, ɐu, ei, iu, ou, ɵy, ɔy, uy/. These are shown in Figure 2.4.

a: Cantonese Monophthongs

b: Cantonese Diphthongs

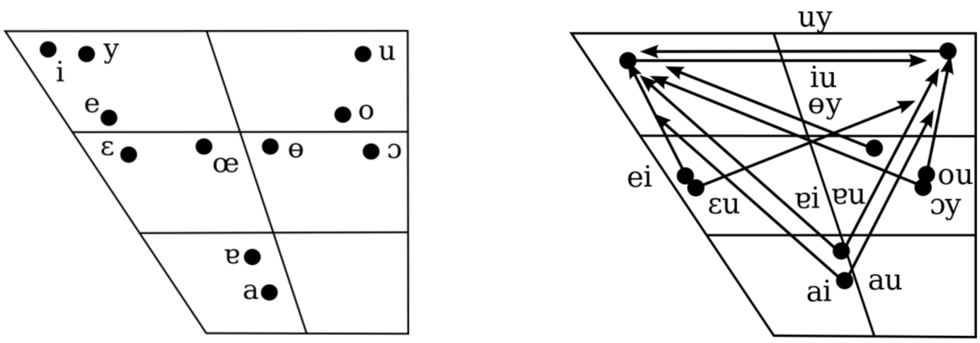


Figure 2.4 Vowel charts of Cantonese

Source: Zee (1999:59)

However, Chan and Li (2000:72) argue that there are only eight phonemic vowels in Cantonese, which are /a:, a, ɛ, i, u, y, ɔ, œ/. Among these monophthongs, /a:/ and /a/ are the only phonemic pair which contrast in length with the long length for /a:/ and short for /a/, which are always transcribed as /a:/ and /ɐ/ respectively in the literature for easy recognition. Therefore, I will use /a:/ and /ɐ/ to conform to the conventions used in the literature. In addition, Chan and Li (2000:72) stated that, among these eight vowel phonemes, five vowels /ɛ, i, u, ɔ, œ/ have long and short allophone pairs, but they occur in complementary distribution. For instance, /i/ has two allophones, [i:] which occurs before labial and alveolar consonants such as [t^{h}i:n55] 'day' and [ɪ] which only occurs before velars such as [pɪŋ55] 'ice'. That is to say, some monophthongs such as [e] [o] [ɵ] in Zee’s description are allophones of the phonemes /ɛ, ɔ, œ/ according to Chan and Li. Actually, based on Zee's perspective, there should be two more monophthongs [ɪ] [ʊ] used before a velar coda; for example, [sɪk2] 'eat' and [lʊk2] 'six'. However, these are not included by Zee in the Cantonese vowels. However, they are treated as allophones of the phonemes /i, u/ by Chan and Li.

The vowel system of Cantonese constructed in this traditional way is shown in Table 2.5 as adapted from the *Guangzhou Cantonese Pronunciation Dictionary* (Zhan, 2004:967-969). This is another approach to show the Cantonese vowel sound system.

Table 2.5 Rhymes of Guangzhou Cantonese

Pinyin	Monophthong Rhymes	Diphthong Rhymes			Nasal Rhymes			Plosive Rhymes		
	-	i	u	y	m	n	ng	p	t	k
aa	aː	aːi	aːu	-	aːm	aːn	aːŋ	aːp̚	aːt̚	aːk̚
a	-	ɐi	ɐu	-	ɐm	ɐn	ɐŋ	ɐp̚	ɐt̚	ɐk̚
e	ɛː	ei		-		-	ɛːŋ		-	ɛːk̚
i	iː	-	iːu	-	iːm	iːn	ɪŋ	iːp̚	iːt̚	ɪk̚
o	ɔː	ɔːi	ou	-	-	ɔːn	ɔːŋ	-	ɔːt̚	ɔːk̚
oe	œː	-	-	ɵy	-	ɵn	œːŋ	-	ɵt̚	œːk̚
u	uː	uːi	-	-	-	uːn	ʊŋ	-	uːt̚	ʊk̚
y	yː	-	-	-	-	yːn	-	-	yːt̚	-
Syllabic	-	-	-	-	m̩	-	ŋ̍	-	-	-

Source: adapted from the *Guangzhou Cantonese Pronunciation Dictionary* (Zhan, 2004:967-969)

From the table, we can clearly see that there are seven monophthongs and ten diphthongs which stands in contrast to Zee's identification of Cantonese monophthongs; however, the number of diphthongs is agreed upon. In Cantonese, all monophthongs are long, as shown in the above table, but only /ɐ/ is short. The short vowel /ɐ/ cannot itself be a rhyme of a syllable according to Cantonese constraints, and that is why the short vowel /ɐ/ is missing from the monophthong column. It is easy to think that the eight phonemic vowels Chan and Li mentioned include this short /ɐ/ because it contrasts with

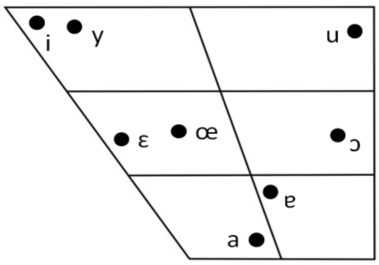
/a:/ phonemically. These seven long monophthongs in the above table are phonemes, five of which, /ɛ, i, u, ɔ, œ/, do have allophones as indicated in the table, but they only occur in a phonotactic way constrained by certain codas or vowels. This generally represents a very similar statement to that of Chan and Li. In addition, it is seen from Table 2.5 that the five phonemic monophthongs /a:, i:, u:, y:, ɔ:/ and three allophones [e, o, ɵ] are tense vowels, while the two phonemes /ɛ: œ:/ and three allophones [ɐ, ɪ, ʊ] are lax vowels. It needs to be emphasised that, in Cantonese, there is only one pair of phonemic tense-lax contrast, which is /a:/ and /ɐ/, and others differ allophonically such as in [i:] and [ɪ] of the phoneme /i:/. In addition, it is noted that the phonemes are always long, and thus transcriptions of Cantonese phonemic vowels can be represented without colons to show the length, as seen in Chan and Li's research (2000:72).

Nevertheless, all these descriptions of Cantonese vowels are reasonable, in that they demonstrate the vowels from different perspectives; that is, phonetically, phonemically or as traditionally stated.

Cantonese has eight phonemic vowels. The above-mentioned eight monophthongs /a, ɐ, ɛ, i, u, y, ɔ, œ/ described by Chan and Li refer to them. The vowels [e] [o] [ɵ] [ɪ] [ʊ] are assuredly allophones of /ɛ, ɔ, œ, i, u/ respectively. Chan and Li merely talk about the vowels from a phonemic and allophonic perspective, whereas Zee adopts a phonetic perspective. The Cantonese phonetic vowel inventory contains 13 vowels which generate 10 diphthongs. Note also that Zee omits the two phonetic vowels ([ɪ] [ʊ]) from the description of Cantonese vowels, but two diphthongs [ɔy] [uy] are included which do not exist in Cantonese. These two should instead be [ɔi] and [ui] respectively in Cantonese. In this study, for the sake of a clear and accurate demonstration of the Cantonese vowel inventory, they are illustrated both phonemically and phonetically as in Figure 2.5.

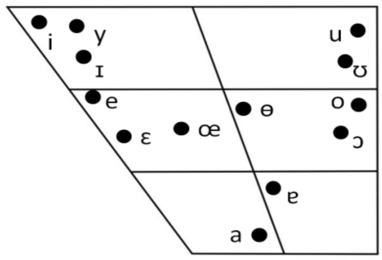
a.

Cantonese Phonemic Vowel Inventory:



b.

Cantonese Phonetic Vowel Inventory:



c.

Cantonese Diphthongs:

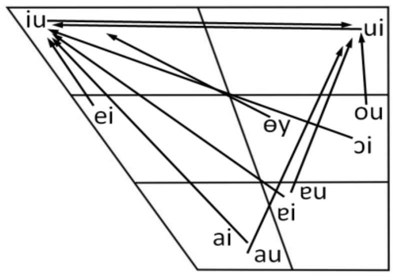


Figure 2.5 Vowel inventory of Cantonese

2.1.2.3 Tones

Cantonese has a more complicated tone system than Harbinese and Mandarin. It has 6 main lexical tones but 9 tones in all, having perfectly preserved the tonal system of Middle Chinese phonology. Tones in the sound system of Middle Chinese were classified into 4 categories called Ping, Shang, Qu and Ru representing level, rising, departing and entering respectively. Also, each tone category was developed and split into two tone registers, according to the syllabic initials, which are called Yin and Yang in Chinese. Syllables with voiced onsets tended to be pronounced with Yang tones means a lower pitch while syllables with voiceless onsets tended to be pronounced with Yin tones, that is, in a upper pitch. Therefore, four tone categories were developed into eight tones in the Middle Chinese period. These are Yin Ping (high level) vs. Yang Ping (low level), Yin Shang (medium rising) vs. Yang Shang (low rising), Yin Qu (medium level) vs. Yang Qu (low level), and Yin Ru (high entering) and Yang Ru (low entering). However, modern Cantonese has further developed Yin Ru into Shang Yin

Ru or Yin Ru (high entering) and Xia Yin Ru or Zhong Ru (medium entering). Therefore, Cantonese actually has nine tones in all, as shown in Table 2.6, but six lexical tones from T1 to T6. The three Ru tones (T7, T8 and T9) overlap with T1, T3 and T6 respectively in pitch height, but the length is shorter. As a result, Ru tones are short tones, corresponding to pitch heights of 5, 3, and 2. In addition, these Ru tones are only marked with plosive rhymes ending with [p̚], [t̚], [k̚] in syllable structure. Although voicing is lost in most modern varieties of Chinese except the Wu dialect/language group and some Gan dialects, Cantonese preserves the tone system of Middle Chinese.

Table 2.6 Tones of Guangzhou Cantonese

Tone Category	Yin ping	Yin shang	Yin qu	Yang ping	Yang shang	Yang qu	Yin ru	Zhong ru	Yang ru
Tone Code	1	2	3	4	5	6	7	8	9
Pitch Height	55/53	35	33	21/11	13	22	5	3	2

Source: adapted from Song and Yu (2007:3) and Zhan and Gan (2012:25)

2.1.3 Comparison of the sound systems between Harbinese Mandarin and Guangzhou Cantonese

From the above introduction to the two chosen dialects/languages of Chinese, it can be seen that their sound systems are very different. They are compared in the following sections to clearly demonstrate the similarities and differences between the dialects/languages in terms of consonants, vowels and tones.

2.1.3.1 Comparison of consonants

Table 2.7 clearly illustrates the differences between consonants in the two dialects/languages. Firstly, the numbers of consonants differ. Harbinese has 23, including /j/ and /w/, while Guangzhou Cantonese comprises only 19 consonants. Those in bold in Table 2.7 indicate the different consonant phonemes in Harbinese and

Guangzhou Cantonese. The special consonants in Harbinese are the four retroflex sounds /ʈʂ/, /ʈʂʰ/, /ʂ/, and /ɻ/, which are absent in Cantonese. Harbinese also has three palatal fricatives and affricates, [tɕ], [tɕʰ], [ɕ], but Cantonese lacks these consonants. However, Cantonese has its own special labio-velar sounds, /kʷ/ and /kʰʷ/, and the nasal sound /ŋ/. Although Harbinese also has the velar nasal /ŋ/, it can only occur in the coda position while the Cantonese /ŋ/ can occur as an onset or coda. The glides /j/ and /w/ are also different in the two dialects/languages. In Harbinese, /j/ and /w/ can act as onsets or as glides of syllables. However, /j/ and /w/ in Guangzhou Cantonese can only act as onsets. Because of these different functions, /j/ and /w/ are parenthesised in the table to indicate the difference between Harbinese and Cantonese.

Table 2.7 Comparison of consonants

		Harbinese	Guangzhou Cantonese
Consonants	Labial	p pʰ m f (w)	p pʰ m f w
	Alveolar	t tʰ n l ts tsʰ s	t tʰ n l ts tsʰ s
	Retroflex	ʈʂ ʈʂʰ ʂ ɻ	
	Palatal	(j) tɕ tɕʰ ɕ	j
	Velar	k kʰ ŋ **x**	k kʰ **ŋ** **kʷ** **kʰʷ**
	Glottal		h

2.1.3.2 Comparison of vowels

Harbinese has five phonemic vowels and at least 10 phonetic vowels, excluding the controversial /ɿ, ʅ, ɚ/, while Cantonese has eight and 13 respectively as seen in Table 2.8. They have four similar phonemes which are the three high vowels /i, y, u/, and a low vowel /a/, but they differ in vowel length, that is, Cantonese contrast in length but Harbinese does not. Cantonese has four phonemes /ɛː, œː, ɔː, ɐ/ which are absent in Harbinese as phonemes, whereas Harbinese has an allophone [ɛ]. Thus, the other three phonemes /œː, ɔː, ɐ/ are specific to Cantonese and distinguish it from Harbinese. Harbinese has the special phoneme /ə/ which does not exist in Cantonese. Cantonese has one phonemic pair (/a/ vs. /ɐ/) and all other phonemes except /y/ have a tense and a lax allophone, depending on their phonological environment. However, Harbinese has no phonemic contrast and no tense-lax allophones. Four diphthongs from each share similarities. They are [ai] [ɑu] [ei] [ou] from Harbinese and [aːi] [aːu] [ei] [ou] from Cantonese. The subtle difference between them is that the Cantonese /a/ has long vowel length. The main difference in features between Harbinese and Cantonese vowels involves a phonemic contrast in length, as shown in detail below.

Table 2.8 Comparison of vowels

	Harbinese	Guangzhou Cantonese
Phonemic vowels	i y u ə a	iː yː uː ɛː œː ɔː ɐ aː
Phonetic vowels	i y u e ɛ ə ɤ o a ɑ (ɿ) (ʅ) (ɚ)	i-ɪ y u-ʊ ɛ-e œ-ɵ ɔ-o a-ɐ
Diphthongs	[ai] [ɑu] [ei] [ou] [ia] [iɛ] [ua] [uo] [yɛ]	[aːi] [aːu] [ei] [ou] [ɐi] [ɐu] [ɔːi] [ɵy] [iːu] [uːi]

2.1.3.2.1 Contrast in tenseness and laxness

Tense vowels are longer than their counterpart lax vowels. English is a good example to show tense and lax phonemic contrast, which means that substituting a tense vowel with its counterpart lax vowel in a word will generate a new word representing a different meaning (Lin, 2007:61). For example, the tense vowel [i] in [bit] 'beat' replaced by its counterpart lax vowel [ɪ] will produce a new word 'bit' [bɪt], which changes the meaning.

Unlike English, Cantonese has a phonemic contrast between tense and lax but only for one pair among eight vowel phonemes; however, Harbinese/Mandarin has none. The phonemic contrast in Cantonese involves the tense /aː/ and the lax /ɐ/, which can be represented not only in closed syllables but also in syllables with diphthongs (Matthews and Yip, 1994:18). Whether this tense-lax pair occurs in open or closed syllables, the actual contrast is between the [aː] and [ɐ] of the phonemes /aː/ and /ɐ/. For example, the closed syllables [ma:n22] *maan* 'slow' and [mɐn22] *man* 'to ask' contrast with each other primarily because the syllables contain items of the main contrasted phonemic pair /aː/ and /ɐ/. Because of the constraint that the short vowel cannot alone be a syllable rhyme in Cantonese, /ɐ/ cannot be used as an open syllable, but the long vowel /aː/ can. Therefore, there is indeed no monophthong [a:] and [ɐ] contrast in open syllables but there are contrasts between diphthongs or between rhymes with nasal codas. In this case, there are a total of eight pairs demonstrating such a contrast primarily between /aː/ and /ɐ/, as seen in the examples of minimal pairs in (2.7) retrieved from the MCCD (Multi-function Chinese Character Database, 2014):

(2.7)

a. contrast between [a:] and [ɐ] in (C)V syllables: which does NOT EXIST

C&CP	IPA	Meaning		C&CP	IPA	Meaning
妈 maa	[ma: 55]	'mother'	vs	*ma	*[mɐ55]	*
花 faa	[fa: 55]	'flower'	vs	*fa	*[fɐ55]	*

(C&CP stands for Character and Cantonese Pinyin)

b. [a:] and [ɐ] contrasted in closed syllables:

C&CP	IPA	Meaning		C&CP	IPA	Meaning
三 saam	[sa:m55]	'three'	vs.	心 sam	[sɐm55]	'heart'
蓝 laam	[la:m21/11]	'blue'	vs.	林 lam	[lɐm21/11]	'woods'
山 saan	[sa:n55]	'hill'	vs.	新 san	[sɐn55]	'new'
但 daan	[ta:n22]	'but'	vs.	炖 dan	[tɐn22]	'to stew'
耕 gaang	[ka:ŋ55]	'to plow'	vs.	汤 gang	[kɐŋ55]	'soup'
盲 maang	[ma:ŋ21/11]	'blind'	vs.	盟 mang	[mɐŋ21/11]	'alliance'
甲 gaap	[ka:p3]	'fingernail'	vs.	鸽 gap	[kɐp3]	'pigeon'
峡 haap	[ha:p2]	'gorge'	vs.	盒 hap	[hɐp2]	'box'
卡 kaat	[kʰa:t5]	'card'	vs.	咳 kat	[kʰɐt5]	'cough'
滑 waat	[wa:t2]	'to slide'	vs.	核 wat	[wɐt2]	'core'
轭 aak	[a:k5]	'yoke'	vs.	抓 ak	[ɐk5]	'to grasp'
赫 haak	[ha:k5]	'bright'	vs.	黑 hak	[hɐk5]	'black'

c. [a:] and [ɐ] contrasted in syllables with diphthongs:

C&CP	IPA	Meaning		C&CP	IPA	Meaning
街 gaai	[ga:i55]	'street'	vs.	鸡 gai	[gɐi55]	'chicken'
快 faai	[fa:i33]	'fast'	vs.	肺 fai	[fɐi33]	'lung'
煮 ngaau	[ŋa:u11/21]	'to boil'	vs.	牛 ngau	[ŋɐu11/21]	'cow'
炒 caau	[ts^ha:u35]	'stir-fry'	vs.	丑 cau	[ts^hɐu35]	'ugly'

Some may be unsure whether [a:] and [ɐ] in Cantonese are two phonemes or two allophones of one phoneme /a/. The answer is that they are phonemes, as verified in examples of the minimal pairs above because they contrast with each other while allophones are in complementary distribution.

Unlike in Cantonese, in Standard Mandarin and Harbinese there is no phonemic contrast between tense and lax (Lin, 2007:62; Lai, 2010:157). Although Lin (2007:68) thought that in Standard Mandarin the diphthongs /ai/ /ei/ /ɑu/ /ou/ end with tense vowels [i] [u], the actual phonetic diphthongs do not reach the heights of tense vowels but to the heights of lax vowels [ɪ] [ʊ]. But since no tense-lax contrast exists in Harbinese/Standard Mandarin therefore the tense vowels [i] [u] in diphthongs are mostly adopted for the transcriptions of the endings of diphthongs instead of using the method employed with lax vowel endings such as /ai/ /ei/ /ɑu/ /ou/. In addition, Lin (2007:62) also mentioned that the mid front and back vowels [e] and [o] in the diphthongs /ei/ and /ou/ in actual pronunciation are closer to [ɛ] and [ɔ] due to the lack of a tense-lax contrast in Standard Mandarin to differentiate meaning in words, and so the tense vowels [e] and [o] are preferred in transcription.

In addition to the main pair of phonemic tense-lax contrasts /aː/ vs. /ɐ/ ([aː] vs. [ɐ]) in Cantonese, there are some other tense-lax vowel pairs which are allophones occurring in complementary distribution, especially in closed syllables. This is unlike the main tense-lax vowel pair mentioned above in Cantonese and tense-lax vowels in English which are phonemes. However, Harbinese/Mandarin does not have this tense-lax contrast although it also has allophonic variants.

The five phonemes /i, ɛ, u, ɔ, œ/ in Cantonese have tense and lax allophones which occur in complementary distribution. For example, [i:] and [ɪ] are the two allophones of the phoneme /i/, but they are in complementary distribution: the tense [i:] occurs before the labials /m, p/, alveolars /n, t/ and vowel /u/ while the lax [ɪ] can only occur before the velars /ŋ/ and /k/. The same is the case with the allophones of /u/, [u:] and [ʊ], where the tense [u:] can occur before the alveolar /n, t/ and the vowel /i/ while the lax [ʊ] only occurs before the velars /ŋ/ and /k/. The two allophones of /œ/ are the tense [ɵ] and lax [œ:]. The tense [ɵ], a centralised mid vowel, occurs before the alveolars /n, t/ and the vowel /y/ while the lax [œ:] occurs only before the velars /ŋ/ and /k/. The allophones of /ɛ, ɔ/ are [ɛ:] and [e] and [ɔ:] and [o] respectively. The tense vowels [e] and [o] are rare, only occurring in diphthongs when respectively combining with [i] and [u], but both their counterpart lax vowels [ɛ:] and [ɔ:] can occur before the velars /ŋ/ and /k/, and also [ɔ:] can appear before the alveolars /n, t/ and the vowel /i/. All of these tense and lax contrasts can be seen clearly in Table 2.9 showing Cantonese rhymes with the dotted line separating them.

Table 2.9 Rhymes of Cantonese

Phonemes	Allophones	Mono-rhymes	Diphthong-rhymes			Nasal-rhymes			Plosive-rhymes		
		-	i	u	y	m	n	ng	p	t	k
/a:/	[a:]	[a:]	[a:i]	[a:u]	-	[a:m]	[a:n]	[a:ŋ]	[a:p̚]	[a:t̚]	[a:k̚]
/ɐ/	[ɐ]	-	[ɐi]	[ɐu]	-	[ɐm]	[ɐn]	[ɐŋ]	[ɐp̚]	[ɐt̚]	[ɐk̚]
/i:/	[i:]	[i:]	-	[i:u]	-	[i:m]	[i:n]		[i:p̚]	[i:t̚]	
	[ɪ]							[ɪŋ]			[ɪk̚]
/u:/	[u:]	[u:]	[u:i]	-	-	-	[u:n]		-	[u:t̚]	
	[ʊ]							[ʊŋ]			[ʊk̚]
/œ:/	[ɵ]		-	-	[ɵy]	-	[ɵn]	[œ:ŋ]	-	[ɵt̚]	
	[œ:]	[œ:]	-	-		-		[œ:ŋ]	-		[œ:k̚]
/ɛ:/	[e]		[ei]								

	[ɛː]	[ɛː]			-		-	[ɛːŋ]		-	[ɛːk̚]
/ɔː/	[o]			[ou]							
	[ɔː]	[ɔː]	[ɔːi]		-	-	[ɔːn]	[ɔːŋ]	-	[ɔːt̚]	[ɔːk̚]
/yː/	[yː]	[yː]	-	-	-	-	[yːn]	-	-	[yːt̚]	-

Source: adapted from Barrie (2003:5)

From the above information, it can be observed that all Cantonese lax vowels can occur before the velar codas /ŋ/ and /k/ and all phonemes are long vowels except for the lax phoneme /a/ ([ɐ]). In addition, all high and low tense vowels such as [aː], [iː], [uː] can be monophthongal rhymes but the mid tense [ɵ], [e], [o] cannot. Instead, their counterpart lax vowels can be syllable rhymes alone but must be lengthened, such as in [œː], [ɛː] and [ɔː].

2.1.3.3 Comparison of tones

Although the two varieties of Chinese both belong to a tonal language, the tone system varies considerably between these two. Harbinese has a much simpler tone system than Guangzhou Cantonese. It has 4 tones while the number of tones in Guangzhou Cantonese reaches up to 9, including 6 main lexical tones, and each has a different tone pitch contour value as seen in Table 2.10. T1 and T2 in Harbinese are the same as those in Guangzhou Cantonese. They are both high level tones, and only the pitch height slightly differs, with the pitch values of T1 and T2 in Harbinese being lower than those in Guangzhou Cantonese. According to Chen (2011:4), T4 and T6 in Guangzhou Cantonese are similar to the falling part of T3 in Mandarin, while T5 in Guangzhou Cantonese is equivalent to the rising part of T3 in Mandarin. He still mentions that the T3 of Guangzhou Cantonese can be considered as like the neutral tone in Mandarin. However, Harbinese has a tone system extremely similar to that of Standard Mandarin, with the only exception being values of tone height or contour. Therefore, it is likely that T4 and T6 in Guangzhou Cantonese generally correspond to the falling part of T3 and T5 to the rising part of T3 in Harbinese as well as the Guangzhou Cantonese T3 corresponding to the neutral tone of Harbinese. The neutral tone is not a real tone in Mandarin and Harbinese but a phenomenon of tone change among the 4 tones, and therefore the neutral tone is not listed in Table 2.10 as a tone. The Ru tones T7, T8 and T9 actually overlap with T1, T3 and T6 respectively in Guangzhou Cantonese in pitch height, but their duration is half as long; therefore, their pitch values are represented by only one digit rather than two. In addition, Ru tones are only marked in syllables with plosive rhymes. The high falling tone, like T4 in Harbinese, does not arise in the Guangzhou Cantonese tone system.

Table 2.10 Comparison and contrast in tones

	Harbinese		Guangzhou Cantonese	
Tone Code	Tone Category	Pitch	Tone Category	Pitch
T1	Yinping (Level)	44	Yin ping	55/53
T2	Yangping (Rising)	24	Yin shang	35
T3	Shangsheng (Falling-rising)	213	Yin qu	33
T4	Qusheng (Falling)	42	Yang ping	21/11
T5			Yang shang	13
T6			Yang qu	22
T7			Yin ru	5
T8			Zhong ru	3
T9			Yang ru	2

2.1.3.3.1 Contrasts in pitch height of tones

Mandarin/Harbinese and Cantonese are all contour tone varieties of Chinese. Mandarin and Harbinese have only one level tone with respective pitch height values of 55 and 44[1]. Meanwhile Cantonese has three level contour tones (Mathews and Yip, 1994:21): T1 (abbreviated from Tone one), the high level tone at 55; T3, the mid level tone at 33; and T6, the low level tone at 22[2]. T4 is a low falling tone in Cantonese (Mathews and Yip, 1994:22), which has a pitch height which is commonly 21 but sometimes is at a

---

[1] According to Nie (2005:113), the tone pitch values of Harbinese are respectively T1 (tone one) 44, T2 24, T3 213, T4 42, while the values for Mandarin are T1 55, T2 35, T3 214, and T4 51 (from Yang, 2010:99).

[2] Pitch height values of Cantonese tones are T1 55/53, T2 35, T3 33, T4 21/11, T5 13, T6 22, T7 5, T8 3 and T9 2 (Song and Yu 2007:3; Zhan and Gan 2012:25).

lower level of 11. If the pitch value of 11 for T4 is taken into consideration, then there would be four level tones in Cantonese. However, there is only one level tone in Mandarin/Harbinese. No matter how high or low the T1 level tone is, it has the same meaning in Mandarin/Harbinese. However, Cantonese differs in contrasting pitch height in level and rising tones. For example, the T1 level tone as an example, it exists in all these three varieties of Chinese but with different tone contour pitch height value. It is 55 in Standard Mandarin but 44 in Harbinese, which is one degree lower. Despite the difference in T1 level tone pitch, speakers of Mandarin and Harbinese still perceive them as having the same meaning because there is only one level tone. For example, in Standard Mandarin the syllable /si/ is transcribed as [sʅ55] 'to think' with T1 (55) but in Harbinese it is [sʅ44] 'to think' with T1 (44). However, both indicate exactly the same meaning to Mandarin speakers, although pitch height of the Harbinese T1 is a little lower. But the situation is entirely different in Cantonese. Although a syllable is the same, changing the pitch height of its level tone will generate different meanings in Cantonese. For example, the syllable [si:] allocated with different level pitch height values such as 55, 33, 22 and 11 produces different meanings. The syllable [si:] with a level tone pitch height of 55, [si:55] means 'poem' while with a pitch of 33 [si:33] represents 'try', of [si:22] 'thing' and of [si:11] 'time'. The same case also applies to the T2 (35), mid rising tone, and T5 (13), low rising tone, which have only a 2-degree gap between them. To take the same syllable [si:] as an example, the rising tones applied to pitch height value give the syllables [si:35] and [si:13] which are distinct in meaning in Cantonese because T2 is 2 degrees higher than T5. The syllable [si:35] means 'history' while [si:13] denotes 'city'. However, there are no differences in meaning between Standard Mandarin and Harbinese. For example, the syllable [ma] with T2 in Standard Mandarin is [ma35] but [ma24] in Harbinese; however, both express the same meaning of 'flax or hemp'. Therefore, in Cantonese meanings are distinguished by changing the pitch height values of level and rising tones, but in Mandarin and closely-related dialects such as Harbinese have no contrast in pitch height.

The different characteristics[3] of tone contours in Mandarin/Harbinese and Cantonese tones are shown in Figure 2.6 in terms of the values of tone pitch height from Yang (2010:99) for Mandarin, Nie (2005:113) for Harbinese, and Song and Yu (2012:49) for Cantonese.

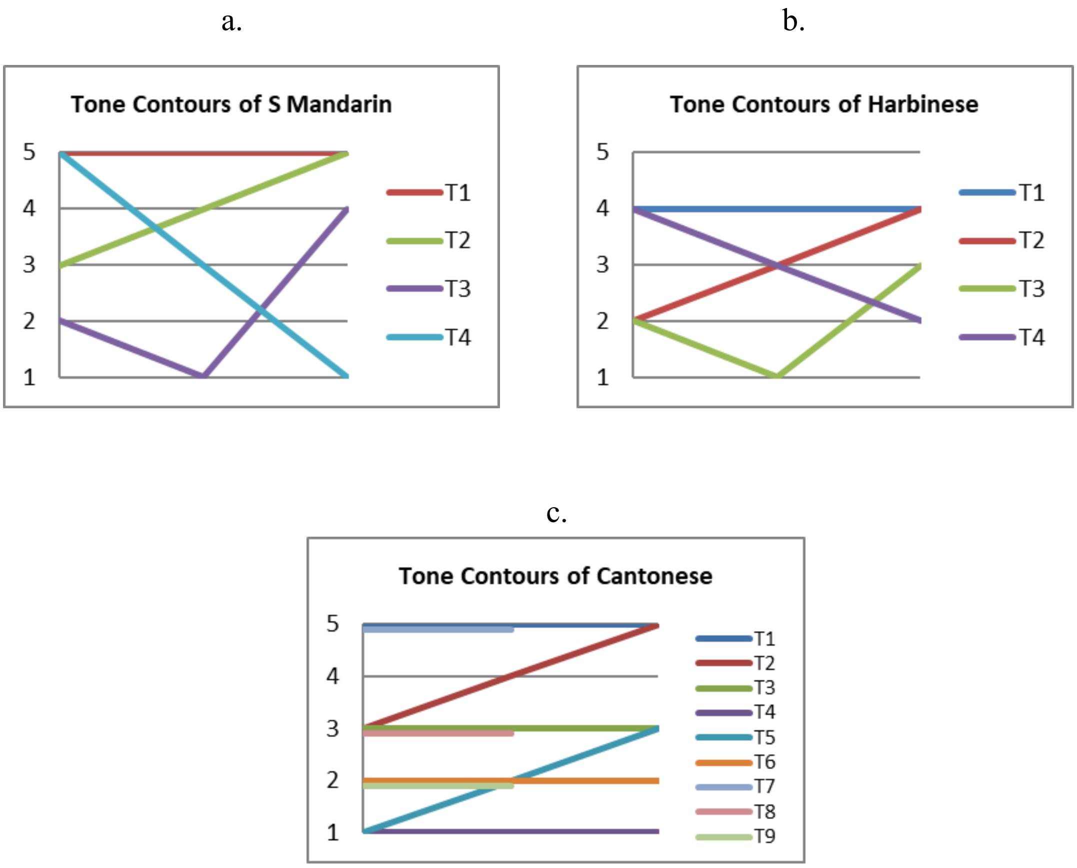


Figure 2.6 Tone contours of Mandarin, Harbinese and Cantonese

2.1.3.4 Comparison of stress

2.1.3.4.1 Differences between Mandarin and the Harbinese and Guangzhou Catonese in stress

Since Chinese is a tonal language, its varieties do not show stress as noticeably and

[3] Cantonese is very sensitive to variation in pitch height but Mandarin and Harbinese are not because there is only one level tone. English is a stress-timed language in which stress plays a significant role due to the syllable with the main stress having a pitch height higher than other syllables within a word, or across words. Although this would have been interesting to investigate, it was beyond the scope of this thesis.

clearly as in English. Some linguists think Chinese is a language without stress (Hyman, 1977; Selkirk and Shen, 1990:315). However, Duanmu (2000:144) argues that there is stress in Standard Mandarin. Duanmu (2000:129) claims that one obvious such stress difference can be detected in disyllabic words with a neutral tone. For example, the disyllabic word bàba [pa51pa] 'father' is composed of a full syllable *bà* [pa51] (with a lexical tone) plus a light syllable ba [pa] (with a neutral tone, which is not a lexical tone but derived from losing its original tone). Here it is clear that the full syllable bears the stress but the light one does not. This is because, as stated by Duanmu (2007:130), the full syllable is "louder and has greater duration and amplitude than the light syllable". The full syllable with a lexical tone is stressed and is heavier and more prominent than the light one with a neutral tone. In this case, it is clear that there is stress in Standard Mandarin.

Some disyllabic words in Standard Mandarin/Harbinese have contrasts in the neutral tone to distinguish between meanings. For example, if the Chinese disyllable word 编辑 *bianji* in (2.8) has two full syllables, where both of the syllables have lexical tones such as in *biānjí* [pian55 tɕi35], then it is a verb meaning 'to edit'. However, if the second syllable *jí* in *biānjí* [pian55 tɕi35] is shifted from the lexical tone T2 to a neutral tone T0 as *ji* [tɕi] (where the neutral tone is represented by the number 0), then the disyllabic word becomes the full-light word *biānji* [pian55 tɕi] which is a noun and means 'editor'. The same applies in the second example 大意 *dayi.* When it consists of full-full syllables, it means the 'main idea' but when composed of full-light syllables, it means 'careless'. As mentioned the full syllable attracts the stress in a full-light disyllable, it is obvious that the stress is attached to the full syllables of each of these two full-light disyllabic words; in *biān* [pian55] in *biānji* [pian55 tɕi0] 'editor', and dà [ta51] in dàyi [ta51 i0] 'careless'. It is noted that the neutral tone may have minor functions to contrast some words but not for all, so it could not be strictly speaking a phonemic tone. It may not show the stress directly like English stress but it shows a heavy-light foot in metrics. Moreover, the key factor for stress is pitch which overlaps with the pitch of tones in some way.

(2.8)

Characters & Pinyin	F - F (full-full disyllable)	F - L (full-light disyllable)
编辑 bianji	[pian55 tɕi35] 'to edit'	[pian55 tɕi0] 'editor'
大意 dayi	[ta51 i51] 'main idea'	[ta51 i0] 'careless'

Nevertheless, in the above full-full group, it is rather difficult to identify which syllable should bear the stress. The most important phonetic factor for English stress is F0 (fundamental frequency) according to Fry (1958) and the next most important cues are duration and then intensity. Mandarin tones are acoustically manifested by different F0 contours. Similarly, F0 is also the primary parameter to distinguish meanings (Liu, 1924; Howie, 1976; Wu, 1986; Kuo, Rosen and Faulkner, 2008). Although F0 is the primary cue for both stress and tone, it acts in different ways. The most important factor of F0 in stress is taken to represent tones in different contour variations in Mandarin in differentiating meanings; therefore, stress is not prominently and freely manifested by F0 (Lin, 2007:225) in Mandarin. Therefore, whether or not there is stress in full syllables in Standard Mandarin/Harbinese is unclear, but some linguists (Hoa, 1983; Duanmu, 2007:Ch6) have demonstrated that stress works phonologically in full syllables but is difficult to detect phonetically.

In studies of stress, Duanmu (2000; 2007) adopts the foot structure to explain the stress existing in Standard Mandarin. The foot is a prosodic domain representing the alternation between a stressed (full or strong) and an unstressed syllable (light or weak), with prominence given to the left-stressed syllable. The discussed full-light syllables can be represented in the foot structure model as in example (2.9).

(2.9)

	Left-prominent foot:	F-L (full-light disyllable)
Head (stress)	x	x
Foot	(S S)	(S S)
Pattern	1 - 0	1 - 0
IPA	[pian55 tɕi0]	[ta51 i0]
Meaning	'editor'	'careless'

Here, S stands for each syllable of the disyllabic word and two Ss constitute a foot with the parenthesis indicating the word boundary. X indicates the strong beat of the foot which has a higher pitch and is called stress or the head. 1-0 is the pattern of this disyllabic word with 1 for the strong beat and 0 for the weak. Duanmu (2000:125) states that Chinese words have initial stress such as that for full-light disyllables, and the basic stress pattern is a syllabic trochee. However, in a subsequent study (2007:141), he changed his mind and agreed with Hoa's (1983) idea that there are three general patterns of word stress in disyllabic words in Standard Mandarin: “1-0 (heavy-light), 1-2 (heavy-heavy, where the first syllable has more stress) and 2-1 (heavy-heavy, where the second syllable has more stress)”. He gives the examples in (2.10):

(2.10)

a.	b.	c.
(HL)	(HH)	H(HØ)
X	X	X
(S S)	(S S)	S (S Ø)
1 - 0	1 - 2	2 - 1
[paa.pa]	[tɕii.x^{w}aa]	[s^{w}uu. ʂɤɤ]
HL-Ø	HL-HL	HL-HL
'dad'	'plan'	'dorm'

In these examples, H in the top line means heavy and L light and Ø means an empty beat combined with the preceding syllable to constitute a foot. The ratio number indicates which syllable attracts more stress. The number 1 is stronger than 2 which is also stronger than 0 where the latter is often used to mark a syllable with a neutral tone. From these patterns, it can be observed that Standard Mandarin can have initial stress (a,b) and final stress (c) in disyllabic words.

Xu (1982, cited by Duanmu, 2007:141) classified a total of 20,000 Chinese common disyllabic words and compounds in these three word stress patterns and found that the final stress H(HØ) pattern accounts for the largest proportion (70%), followed by the (HH) (22.5%) and (HL) (7.5%) patterns. However, Duanmu (2007:142) noted that stress will shift from 2-1 to 1-2 or 2-3 in the pattern of H(HØ) when it follows a word. Here the empty beat is replaced by a syllable, but the foot is reconstituted by the disyllabic word because of the left-prominence rule. The first syllable then has more stress than the second one as in the examples in (2.11) which derives from (2.10 c) plus a syllable [lou35].

(2.11)

a.	b.
H(HØ)	(HH) H
	X
X	X X
S (S Ø)	(S S) S
2 - 1	2 - 3- 1
[s^wuu. ʂɤɤ]	[[s^wuu. ʂɤɤ] lou]
HL-HL	HL-HL-HL
'dorm'	'dorm building'

However, it can be seen that there is still a final stress pattern. In simple words with three or more full syllables, such as loanwords of foreign names and places, the stress pattern 2-X-1 was proposed by Chao (1968, cited in Duanmu, 2007:142) but Hoa (1983) presented the very detailed stress pattern as seen in Table 2.11 (extracted and adapted from Duanmu 2007:142-143).

Table 2.11 Stress patterns in Standard Mandarin

Length	Hoa's pattern	Hoa's pattern in SW	Duanmu's foot structure
2 syllables	2-1	WS	H(HØ)
3 syllables	2-x-1	SWS	(HH)(HØ)
4 syllables	2-x-x-1	SWWS	(HH)H(HØ)
5 syllables	2-x-3-x-1	SWSWS	(HH)(HH)(HØ)
6 syllables	2-x-3-x-x-1	SWSWWS	(HH)(HH)H(HØ)
	3-x-2-x-x-1	SWWSWS	(HH)H(HH)(HØ)
7 syllables	3-x-2-x-3-x-1	SWSWSWS	(HH)(HH)(HH)(HØ)
	2-x-3-x-3-x-1	SWWSWWS	(HH)H(HH)H(HØ)

In Hoa's stress pattern, 1 is strongest, followed by 2 and then 3, and x is the weakest. S and W in Hoa's pattern in SW stands for strong and weak respectively and then converted in terms of Hoa's pattern by Duanmu. Ø is used as an empty beat to constitute a foot with its previous heavy syllable H. In Standard Mandarin, the neutral tone cannot occur on an initial syllable but mostly on the final; therefore, Duanmu (2007:143) thinks that the final syllable may sound stronger because it seems longer with the possession of an empty beat as a foot.

As a result of the fact that Harbinese is the dialect most similar in pronunciation to Standard Mandarin among all Mandarin dialects, and also because their sound systems are very similar, (with only the single exception of the lack of the monophthong [o] in Standard Mandarin), Harbinese and Standard Mandarin have generally similar stress patterns. However, Harbinese still has characteristics distinct from Standard Mandarin, and there may be other slight differences due to tonal differences in the Harbinese colloquial vocabulary and Harbinese Russian loanwords. For historical reasons, Harbinese has Russian loanwords absent from Standard Mandarin and other Mandarin dialects and even other Northeastern dialects. These words are still frequently used in the daily life of Harbin people. It seems that Harbinese has a special stress pattern different from Standard Mandarin, as in loanwords and stress pattern in Harbinese shown in Table 2.12.

Table 2.12 Special stress patterns in Harbinese

Length	Pinyin	IPA	Meaning	Pattern	Origin
2 syllables:	mǎshén	[maT3ʂənT2]	'machine'	2-1	маши̇на
3 syllables:	bǔláji	[puT3laT2tɕiT0]	'dress'	2-1-0	пла́тье
	mǎdámu	[maT3taT2muT0]	'madam'	2-1-0	мада́м
	gěwási	[kɤT3waT2srT0]	'a Russian drink'	2-1-0	квас
	bālízi	[paT1liT2tsrT0]	'jail'	2-1-0	поли̇ция
	wěideluór	[veiT3tɤT0luorT2]	'water bucket'	2-0-1	ведро́
	hālashào	[xaT1laT0ʂɑuT4]	'good'	2-0-1	хорошо́

In the stress patterns of special Russian loanwords in Harbinese, disyllabic loanwords have the same pattern 2-1 as Standard Mandarin, but there seems to be a special stress pattern of 2-1-0 in trisyllabic loanwords with the prominent stress on the second syllable instead of Standard Mandarin's stress-final 2-3-1 pattern. This is because the final syllable in Harbinese is often an unstressed syllable due to the neutral tone. This can be observed in the comparison of 格瓦斯 *gewasi* 'the Russian style drink' in Harbinese and Standard Mandarin in (2.12). This loanword only develops a corresponding Standard Mandarin pronunciation in contrast to its Harbinese version due to the fact that it has become popular and has spread to the whole China as a new type of drink. Although it has the same Chinese character 格瓦斯, pronunciations of the drink's name are different. All three syllables of *gewasi* have lexical tones in Standard Mandarin but in Harbinese the tones change, with a neutral one on the final syllable as seen in (2.12).

(2.12)

Charater	Harbinese	Standard Mandarin	Russian
格瓦斯	gěwási [kɤT3waT2srT0]	géwǎsī[kɤT2waT3srT1]	квас [kvas]
	2-1-0	2-3-1	

Different word stress patterns and pronunciations in Harbinese and Standard Mandarin for this loanword occur because Harbinese retains the original stress of the Russian word *квас* [kvas] 'a Russian drink' when the word is loaned, while Standard Mandarin only borrows the Chinese character 格瓦斯 from Harbinese but not the pronunciation; therefore, it is pronounced differently outside the Harbin area. Although *квас* [kvas] was loaned and converted into Harbinese as a trisyllabic word in conformity with the Chinese convention of the wellformedness of syllable structure, where a consonant cluster is not allowed in Chinese syllables. So the stress on the second syllable [wa] is still in accordance with that of the nucleus [a] of the Russian *квас* [kvas].

It can be observed that the stress is retained in Harbinese for the Russian loanword not only for the word *gěwási* 'a Russial drink' but also other loanwords such as *wěideluó,* ведрó, 'water bucket', and *hālashào,* хорошó, 'good'. The final stress is kept in Harbinese by neutralising the previous syllable in accordance with the stress on the final syllable in Russian words. However, for trisyllables, the word stress pattern in Harbinese retaining the Russian 2-0-1 stress is analogous to the stress of the Standard Mandarin 2-3-1. Both have final stress although not in exactly the same way. Therefore, the stress pattern of 2-1-0 is specific to Harbinese.

Generally speaking, if there is stress in Chinese, Standard Mandarin and Harbinese have different stress patterns depending on word length and the presence of full or light syllables. The most frequently used word stress patterns in these two varieties of Chinese are final stress in full syllables followed by initial stress for full-light syllables, although there are also some other less common patterns.

Compared with Standard Mandarin and Harbinese, Cantonese has not been argued to mark stress (Yip, 1990:27). The neutral tone does not exist in Cantonese, and therefore there is no full-light distinction between syllables. However, Yip (1990:27) claims that the only place where a stress difference can be found in Cantonese is where bi-syllables are formed by attaching a prefix [a33] to a family name or by suffixing a high tone to that name. Yip's examples in (2.13) indicate this type of bi-syllabic form, with the more prominent second syllable bearing stress.

(2.13)

Family name	Prefix [a33]	+	Family name
[ji:p22]	[a33]		[ji:p25]
[tsʰɐn21]	[a33]		[tsʰɐn25]
[tsœːŋ53]	[a33]		[tsœːŋ5]

The prefix [a33] is always attached to a family name in Cantonese to indicate a close relation. When it is prefixed to a family name, the tone is changed from original low tone to a high tone pitch contour. After the prefix [a33] attached to a family name, the end pitch of the tone on the name syllable rises to high from low and reaches a high level tone value of 5. For example, the family name [ji:p22] 'Yip' has a low level tone T6 with a pitch of 22, but its end pitch is switched from a low pitch 2 in a low level tone 22 to a high pitch value of 5 when it is preceded by the prefix [a33]. In this case, Cantonese has prominent stress on the final syllable in disyllabic words. As to other full-full syllables, no other studies indicate stress patterns in Cantonese.

2.1.3.4.2 Differences in prosody

Yip (1992:25) provides evidence that all Cantonese syllables are bi-moraic in loanword phonology. This is attributed to the distribution of long and short vowels in Cantonese. In closed syllables the vowels can be long or short, but in open syllables the vowels can only be long monophthongs or diphthongs. Also, she states that prefixes in Cantonese family names or those discussed above in relation to Cantonese stress can represent not only a final stress pattern in Cantonese but also a prosodic output template which is in an iambic foot (Yip, 1992:27). Examples in (2.14) from (2.13) are re-illustrated here in terms of foot structure.

(2.14)

(LH)	(LH)	(LH)
X	X	X
(S S)	(S S)	(S S)
2 - 1	2 - 1	2 - 1
[a33 ji:p25]	[a33 ts^{h}ɐn25]	[a33 tsœːŋ55]
HL- HL	HL- HL	HL- HL
'Yip'	'Can'	'Zoeng'

LH represents light and heavy syllables, and X is the stress or head of the foot composed of two syllables, with 1 for strong and 2 for weak. Yip mentions that the second syllable is lengthened while the first is shorter because it is mono-moraic as a prefix here. Therefore, in this LH pattern (2-1), the final syllable is stressed and is much stronger than the first short prefixed unstressed syllable and it forms the iambic foot.

Hoa's three general patterns of word stress in disyllabic words in Standard Mandarin/Harbinese mentioned above are HL (1-0) (heavy-light), HH (1-2) (heavy-heavy where the first syllable has more stress) and H(HØ) (2-1). According to this, Standard Mandarin and Harbinese can have different prosodic templates for the foot. The examples in (2.10) are used again in (2.15) to explain the foot template in Standard Mandarin/Harbinese (Duanmu, 2007:141):

(2.15)

a.	b.	c.
(HL)-7.5%	(HH)-22.5%	H(HØ)-70%
X	X	X
(S S)	(S S)	S (S Ø)
1 – 0	1 - 2	2 - 1
[paa.pa]	[tɕii.x^{w}aa]	[s^{w}uu. ʂɤɤ]
HL-Ø	HL-HL	HL-HL
'dad'	'plan'	'dorm'

In the findings from Xu (1982) mentioned above, both Standard Mandarin and Harbinese have a high tendency (70%) to have final stress (H(HØ)) in full-full disyllables, followed by the initial stress pattern (HH) (22.5%) in full-full syllables and the initial stress pattern (HL) (7.5%) in full-light syllables with a final neutral tone. From these stress patterns, it can be seen that Mandarin/Harbinese has two main prosodic templates: the iambic foot, accounting for 70% of 20,000 commonly used disyllabic words; and the trochaic foot, for the 30% with initial stress.

2.1.4 The sound system of the English Language (RP and GA English)

English is spoken as a native or a foreign language by vast numbers of people. Unsurprisingly, there is much variation among English dialects and English accents. Among these varieties, British Received Pronunciation (RP) and General American English (GA) are commonly regarded as standard English and taught and learnt as a foreign language. As noted by Brown (1991), Received Pronunciation is the most prestigious accent of English and is widely used in EFL teaching and learning throughout the world. In China, most Chinese learners of English learn RP English in primary school, and most textbooks are based on its pronunciation. However, with the great popularity of American films and dramas in recent decades, American English has become increasingly influential. Although RP is adopted and used in most school textbooks, Chinese learners of English are often exposed to American English from American media in their daily lives. This leads to a mixed accent in Chinese learners of English. Griner (2014:1) notes that Chinese English learners have a mixed understanding of English pronunciation because they were taught RP in primary school but learned more from GA afterwards. Since Chinese learners of English mix RP and GA, not only is the English sound system of RP English considered in this research but also that of GA and the typical differences between these two varieties are taken into consideration.

2.1.4.1 Consonants

2.1.4.1.1 Consonants in RP

According to Roach (2004), RP English has 24 consonants shown in Table 2.13, classified by manner and place of articulation with voiceless consonants on the left. In terms of the former, the consonants are composed of 6 plosives, 2 affricates, 9 fricatives, 3 nasals and 4 approximant sounds. However, in terms of place of articulation there are 3 bilabial consonants, 2 labio-dentals, 2 dentals, 7 alveolars, 5 palatal-alveolars, 3 velars, 1 labio-velar and 1 glottal.

Table 2.13 Consonants in RP English

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Palatal	Velar	Labialvelar	Glottal
Plosive	p b			t d			k g		
Affricate					tʃ dʒ				
Fricative		f v	θ ð	s z	ʃ ʒ				h
Nasal	m			n			ŋ		
Approximant				ɹ		j		w	
Lateral approximant				l					

Source: Roach (2004:240)

2.1.4.1.2 Consonants in GA

There are few differences in consonant phonemes between the two varieties of English. But significant differences in allophonic distributions are discussed below. The only differences between the two varieties of English concern phonological position.

2.1.4.2 Vowels

2.1.4.2.1 Vowels in RP

Received Pronunciation has 20 vowel phonemes, including 12 monophthongs and 8 diphthongs, as illustrated in detail in (2.16) (Roach, Hartman and Setter, 2011:viii).

(2.16)

The 12 RP monophthongs:
/iː/, /ɪ/, /e/, /æ/, /ɜː/, /ə/, /uː/, /ʊ/, /ʌ/, /ɑː/, /ɔː/, /ɒ/.

The 8 RP diphthongs:
/eɪ/, /aɪ/, /aʊ/, /ɔɪ/, /əʊ/, /ɪə/, /eə/, /ʊə/.

According to the vowel traits of height, backness, rounding and length, the 12 monophthongs can be classified as illustrated in Table 2.14.

Table 2.14 Monophthongs in RP English

		Front		Central		Back	
		Long	Short	Long	Short	Long	Short
Close /High	Unrounded	/iː/	/ɪ/				
	Rounded					/uː/	/ʊ/
Mid	Unrounded		/e/	/ɜː/	/ə/		
	Rounded					/ɔː/	
Open /Low	Unrounded		/æ/		/ʌ/	/ɑː/	
	Rounded						/ɒ/

Source: adapted from Roach (2004:242)

It can be seen from the Table 2.14 that, in terms of length, RP has five long vowels /iː, ɜː, uː, ɔː, ɑː/ and seven short vowels /ɪ, e, æ, ə, ʌ, ʊ, ɒ/; or in terms of vowel height there are four high vowels /iː, ɪ, uː, ʊ/, four mid vowels /e, ɜː, ə, ɔː/ and four low vowels /æ, ʌ, ɑː, ɒ/; or, classified by vowel roundedness, four rounded vowels /uː, ʊ, ɔː, ɒ/. It is standard practice to represent the monophthongs in a quadrilateral chart as in Figure 2.7 (Roach, Hartman and Setter, 2011:viii). The vowels in the chart are located according to relative place of articulation, which may not be absolutely accurately defined (ibid.).

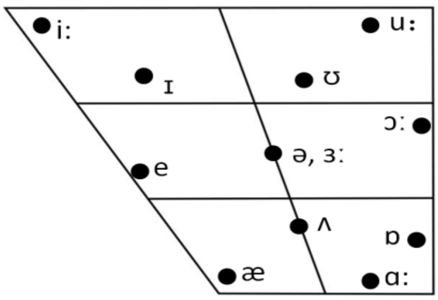


Figure 2.7 Monophthongs in RP English

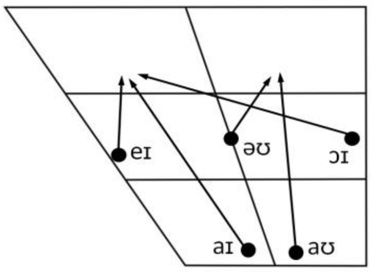
Received Pronunciation has 8 diphthongs as in Table 2.15 which can be classified by the height of the final vowel into categories of closing diphthongs (ending in high vowels /ɪ, ʊ/) or centring diphthongs (ending in mid vowel /ə/). There are five closing diphthongs /eɪ, aɪ, ɔɪ, aʊ, əʊ/ and three centring diphthongs /ɪə, eə, ʊə/, as seen in Figure 2.8.

Table 2.15 Diphthongs in RP English

Closing	/eɪ/ /aɪ/ /ɔɪ/ /aʊ/ /əʊ/
Centring	/ɪə/ /eə/ /ʊə/

Source: Roach (2004:242; 2013:17)

Closing diphthongs of RP English



Centring diphthongs of RP English

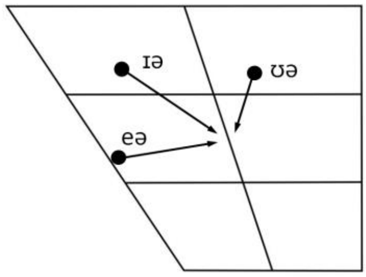


Figure 2.8 Diphthongs of RP English

2.1.4.2.2 Vowels in GA

The vowel system of General American English is generally described using the terms tense and lax, unlike British RP which is commonly described using the terms long and short (Roach, Hartman and Setter, 2011:ix). Tense vowels are, however, longer than lax vowels, but vowel length in GA is not as important as in RP and is primarily indicated by phonological constraints (ibid.). Therefore, transcriptions of tense vowels in GA do not include length [:], as in RP. However, the Cambridge English Pronouncing Dictionary (ibid.) retains the length diacritic [:] in transcriptions of GA tense vowels in order to more clearly reveal the relationship between the vowel systems of GA and RP (ibid.). In this research, transcriptions of English words with GA pronunciation abide by the rules of the Cambridge English Pronouncing Dictionary. The vowel inventory of GA is shown in (2.17) (ibid: x):

(2.17)

GA tense monophthongs:

/iː, ɜː, uː, ɔː, ɑː/

GA lax monophthongs:

/ɪ, e(ɛ)[4], æ, ə, ʌ, ʊ/

---

[4] /e/ is always used to transcribe the vowel in the word *red* in British English, but /ɛ/ is often used in American

GA diphthongs:

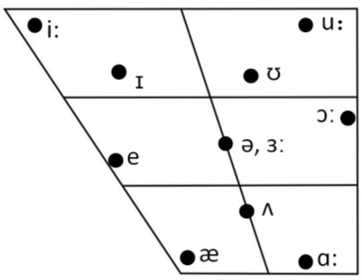
/eɪ, aɪ, ɔɪ, aʊ, oʊ/

GA retroflex vowels:

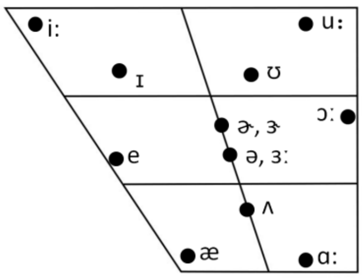
[ɚ, ɝ]

There are 11 monophthongs in GA, with five tense and six lax vowels. Compared with RP, GA lacks the short back vowel /ɒ/. Moreover, GA has only five closing diphthongs ending in the high vowels /ɪ, ʊ/, while the three RP centring diphthongs /ɪə, eə, ʊə/ ending in the mid schwa /ə/ are absent in GA because the /ɹ/ is realised and pronounced in all positions; therefore, the three centring diphthongs /ɪə, eə, ʊə/ in RP are not diphthongs but instead are pronounced as [ɪɹ, eɹ, ʊɹ] in GA. In addition, when the vowels /ə, ɜː/ precede /ɹ/, they are co-articulated as the retroflex vowels [ɚ, ɝ]. Thus, the vowels /ə, ɜː/ in GA have two special retroflex allophones [ɚ, ɝ]. The vowel system of GA is represented in Figure 2.9.

Phonemic monophthongs of GA:



Phonetic monophthongs of GA:



English. Thus, both are provided here and /ɛ/ in parenthesis.

Closing diphthongs of GA:

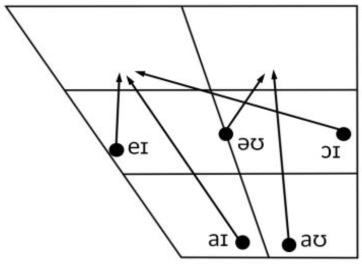


Figure 2.9 Vowel charts of GA English

2.1.4.3 Stress

Stress is an important phonological phenomenon in many languages. It refers to the measure of the prominence assigned to a certain syllable in a word. Among a sequence of syllables in a word, that which bears more prominence than others carries stress. This syllable is always articulated with more muscular effort and is perceived as louder and with longer duration and more pitch variation (Davenport and Hannahs, 2014:78). This stress on a syllable within a word is called lexical or word stress. The stress is not always fixed to a particular syllable in a word in all languages with stress, but its placement is variable; it could be on an initial, final or penultimate syllable within a word. Unlike Finnish and Czech with stress placed on the initial and French on the final syllables in words, English places stress on different syllables in words. For example, there is initial stress in words like *'**be**.tter'* and *'**bro.**ther'*, penultimate stress in *'re.**mem**.ber'* and *'Oc.**to**.ber'*, and final stress in *'em.ploy.**ee**'* and *'ba.**lloon**'*. Stress tends to be placed on the syllable counted from the right-hand edge of the word (ibid.). Therefore, the stress placed on the ult (the last syllable) such as in *'em.ploy.**ee**'* and *'ba.**lloon**'* is called final stress, or on the penult (the syllable before the ult) called penultimate stress. Therefore, the above examples with initial stress, *'**be**.tter'* and *'**bro.**ther',* are usually said to have penultimate stress, as are *'re.**mem**.ber'* and *'Oc.**to**.ber'*. Also, stress placed on the antepenult (the syllable before the penult) is called antepenultimate stress as in the words *'**oc**.to.pus'* and *'e.co.**no**.mi.cal'*.

In addition, English stress occurs not only at the level of words, but also at the level of larger units such as phrases or sentences. In the present thesis, only word stress was considered. Fry (1958:126) argues that F0 (pitch), duration, and intensity are the three most important acoustic features of English word stress, among which intensity is considered the least important.

The prominence of a syllable within an English word is not simply confined to two possibilities of stressed (strong) or unstressed (weak). In longer English words, it can be recognised that the degree of prominence taken by syllables differs. For example, in the word ***e**co**no**mical* [ˌiːkə'nɒmɪk(ə)l], the syllable *no* **[**nɒ] is pronounced stronger than the syllable *e* [i:], but *e* [i:] is also pronounced stronger than other syllables. Then, stress within a word has different levels and the syllable *no* **[**nɒ] in ***e**co**no**mical* [ˌiːkə'nɒmɪk(ə)l] is strongest. Therefore, the strongest stress in a word is called primary stress, indicated by the superscript ', while stress less prominent such as the syllable [i:] is called secondary stress, always represented by the subscript ˌ, and other syllables are unstressed without symbolic indication. According to Roach (2009:75), stress could also be further classified at a third level called tertiary stress. However, this is not relevant to this study and is not discussed further. It is worth mentioning that, in English vowels such as /ə, ɪ/ or a syllabic consonant in unstressed syllables sound weaker than other vowels in unstressed syllables (ibid:75).

2.1.4.4 Comparison of sound systems in RP and GA

2.1.4.4.1 Phonological differences between RP and GA in consonants

Although the two varieties of Standard English share the same phonemic consonants, there are still some major phonological differences between RP and GA in their realisation, perception and allophonic distribution.

a. rhotic /ɹ/

A typical distinguishing feature between RP and GA is the realisation and pronunciation of the consonant /ɹ/. In RP, /ɹ/ exists only in the onset of syllables, while in GA it can occur in all positions. Examples are seen in Table 2.16.

Table 2.16 Differences between RP and GA in rhotic /ɹ/

Word	RP	GA	Position
red	[ɹed]	[ɹed]	Initially
bark	[bɑ:k]	[bɑɹk]	Finally
car	[kɑ:]	[kɑɹ]	Finally
more	[mɔː]	[mɔːɹ]	Finally

From the table, it is clear that words with *r* post-vocalically in each syllable such as *car* are pronounced without a final /ɹ/ so very differently in RP and GA. The word *car* is pronounced as [kɑ:] in RP and [kɑɹ] in GA. Thus, RP is characterised as 'non-rhotic' and GA as 'rhotic'.

b. intervocalic /t/

Both RP and GA have the consonant phoneme /t/, but its allophonic distribution in these two varieties of English differs. When /t/ occurs intervocalically (between vowels) in individual words or across word boundaries, it is often pronounced as a flap with voicing in GA and is transcribed as [ɾ]. However, the /t/ in RP remains unvoiced. For example, /t/ in the word *butter* is phonetically [t] in [ˈbʌtə] in RP. Examples showing differences in the intervocalic consonant /t/ between RP and GA are shown in Table 2.17.

Table 2.17 Differences between RP and GA in intervocalic /t/

Word	RP	GA	Position (V for vowel)
butter	[ˈbʌtə]	[ˈbʌɾɚ]	V+/t/+V in individual word
water	[ˈwɔːtə]	[ˈwɑːɾɚ]	V+/t/+V in individual word
letter	[ˈletə]	[ˈleɾɚ]	V+/t/+V in individual word
get it	[ˈget it]	[ˈgeɾ it]	V+/t/ V across word boundaries
out of	[ˈaʊt əv]	[ˈaʊɾ əv]	V+/t/ V across word boundaries

Nevertheless, not all instances of /t/ occurring between vowels must be voiced (Dimitrova, 2013): it is not voiced in GA if the preceding vowel is not stressed, such as in *attach* [ə'tætʃ] and *retain* [rɪ'teɪn] in Table 2.18. The vowels preceding the intervocalic /t/ are not stressed, and so the /t/ is aspirated rather than voiced.

Table 2.18 Exceptions of intervocalic /t/ between RP and GA

Word	RP	GA	Position
attach	[ə'tætʃ]	[ə'tætʃ]	V+ /t/ + 'V ('V means V is stressed)
retain	[rɪ'teɪn]	[rɪ'teɪn]	V+ /t/ + 'V

The examples from Tables 2.17 and 2.18 clearly indicate the different realisations of /t/ in British and American varieties of English. The flapped and voiced /t/, i.e. [ɾ], is another important feature of GA.

c. glide /j/

Another dissimilarity between RP and GA relates to the occurrence of the glide /j/ between certain consonants and a long vowel /u:/. This glide before /u:/ in a stressed syllable commonly occurs as [ju:] in RP, but /j/ is always absent in GA when it occurs between the alveolar consonants /t, d, n/ and /u:/. This is illustrated in Table 2.19.

Table 2.19 Differences between RP and GA in glide /j/

Word	RP	GA	Position
tube	/t**j**u:b/	/tu:b/	/t_u:/
due	/d**j**u:/	/du:/	/d_u:/
new	/n**j**u:/	/nu:/	/n_u:/

d. /w/

The grapheme *wh-* is always pronounced as [w] in RP; however, in GA it is generally pronounced as [ʍ], as seen in Table 2.20 (He, 1998:99; She, 2011:326). The grapheme *wh-* in the *Cambridge English Pronouncing Dictionary* is transcribed as /hw/ for both RP and GA accents, but a /h/ sound before /w/ is, actually, not pronounced in RP. Therefore, the transcriptions for *wh-* words in this research adopt /w/ rather than /hw/

for RP. This also constitutes a slight difference between British and American English accents with regard to consonants.

Table 2.20 Differences between RP and GA in /w/

Word	RP	GA
where	[weə]	[ʍeɹ]
when	[wen]	[ʍen]
wheat	[wi:t]	[ʍi:t]
whip	[wɪp]	[ʍɪp]

f. /ʃ/ or /ʒ/

Another slight difference between the two varieties is the tendency in RP to use /ʃ/ instead of /ʒ/ in some words (Dimitrova, 2013). In RP, certain words are more typically pronounced using the voiceless fricative palatal-alveolar consonant /ʃ/, but in the same contexts they tend to be pronounced using the voiced /ʒ/ in GA, as shown in Table 2.21.

Table 2.21 Differences between RP and GA in /ʃ/ or /ʒ/

Word	RP	GA
Asia	[ˈeɪʃə] / [ˈeɪʒə]	[ˈeɪʒə]
Persia	['pə:ʃə]	['pɝ:ʒə]
version	['vɜ:ʃ(ə)n]	['vɝ:ʒ(ə)n]

Consonants and their phonetic realisations in RP and GA are otherwise generally analogous to each other with the exception of the above discussed differences. In addition, it is noted that nuances exist between two varieties of English such as /θ, ð/ being interdental in GA but dental in RP. Although this bears more investigation, the differences across English varieties were not investigated in the present thesis.

2.1.4.4.2 Differences between RP and GA in vowels

2.1.4.4.2.1 Monophthongal differences

RP has 12 phonemic monophthongs while GA has 11. Most of the monophthongs in the two varieties are similar; only the short back rounded vowel /ɒ/ in RP does not exist in GA, as mentioned above. Monophthongs in GA described as tense and lax correspond to monophthongs in RP described as long and short (though the correspondence is not exact).

Although monophthongs are similar in number in RP and GA, the actual pronunciation of vowels for the same words does not fully correspond. When considering the same words, the mid and low vowels in RP can be seen to have shifted forward in certain contexts in GA (Griner, 2014:6), as shown in detail in Table 2.22.

Table 2.22 Vowel shifts from RP to GA

Word	Vowel Shift	RP	GA
caught	/ɔ:/ → /ɑ:/	[kɔ:t]	[kɑ:t]
pot	/ɒ/ → /ɑ:/	[pɒt]	[pɑ:t]
what	/ɒ/ → /ʌ/	[wɒt]	[wʌt]
fast[5]	/ɑ:/ → /æ/	[fɑ:st]	[fæst]
Barry	/æ/ → /e/	[bæɹi]	[beɹi]

Source: adapted from Griner (2014:6)

a. British /ɒ/ pronounced as [ɑ:] in GA

Another clear difference between the two varieties of English is that the British short rounded back vowel /ɒ/ is missing in GA. In consequence, this sound occurring in words in RP is always replaced by the back unrounded tense vowel /ɑ:/ in GA. For example, the negation word *not* is pronounced [nɒt] in RP but [nɑ:t] in GA. Other

[5] It has to be mentioned that *fast* is an exception under the title in Table 2.23. RP and GA both developed from a common source where [fæst] was the original pronunciation. [fæst] was shifted to [fɑ:st] in RP but remained in GA.

examples of vowel variation between /ɒ/ and /ɑ:/ in the two varieties can be seen in Table 2.23.

Table 2.23 Vowel shift from RP /ɒ/ to GA /ɑ:/

/ɒ/ in RP → /ɑ:/ in GA		
Word	RP	GA
dog god offer doctor opposite	[dɒg] [gɒd] ['ɒfə] ['dɒktə] ['ɒpəzɪt]	[d**ɑ:**g]/[dɔːg] [g**ɑ:**d] ['**ɑ:**fɚ] ['d**ɑ:**ktɚ] ['**ɑ:**pəzɪt]

In the above examples, the word *dog* has to be mentioned because it is also possible to pronounce it as [dɔːg]. That means the short low back vowel /ɒ/ in RP is pronounced mainly as /ɑ:/ and can also be pronounced as /ɔː/ in some specific words in GA, such as *log*, *fog* and *hog*.

b. British /ɒ/ sometimes pronounced as [ʌ] in GA

Besides the replacement of the British short /ɒ/ by the unrounded long /ɑ:/ in GA, the British short vowel /ɒ/ can also be pronounced as the central unrounded /ʌ/ in certain words in GA, such as in the word *what* shown in Table 2.24 which is pronounced [wɒt] in RP but [ʍʌt] in GA. Compound words with *what* always have this slight variation in GA.

Table 2.24 Vowel shift from RP /ɒ/ to GA /ʌ/

/ɒ/ in RP → /ʌ/ in GA		
Word	RP	GA
what whatever	[wɒt] ['wɒtevə]	[ʍ**ʌ**t]/[ʍɑ:t] ['ʍ**ʌ**ɾevɚ]/[ʍɑ:ɾevɚ]

c. British /ɔː/ often pronounced as [ɑː] in GA

Americans seem to prefer to pronounce unrounded back vowels instead of rounded ones. The long back mid rounded vowel /ɔː/ in British pronunciation is often pronounced using the low back unrounded long vowel /ɑː/ in GA, (although the original pronunciation /ɔː/ is also possible), as in the examples in Table 2.25.

Table 2.25 Vowel shift from RP /ɔː/ to GA /ɑː/

/ɔː/ in RP → /ɑː/ in GA		
Word	RP	GA
water	['wɔːtə]	['w**ɑː**ɾɚ]/['wɔːɾɚ]
fall	[fɔːl]	[f**ɑː**l]/[fɔːl]
taught	[tɔːt]	[t**ɑː**t]/[tɔːt]
autumn	['ɔːtəm]	['**ɑː**ɾəm]/['ɔːɾəm]

d. British /ɑː/ pronounced as [æ] in GA

From the above rule *a*, it can be predicted that *lost* is pronounced [lɒst] in RP and [lɑːst] in GA. However, the American sound [lɑːst] can sound like the word *last* to British ears if there is no specific context. Meanwhile, the word *last* is always pronounced [læst] in American English rather than British [lɑːst]. That is, the long back vowel /ɑː/ without a following /ɹ/ sound in a British accent tends to be pronounced further forward as a front /æ/ in the American accent. Examples are shown in Table 2.26.

Table 2.26 Vowel shift from RP /ɑː/ to GA /æ/

/ɑː/ in RP → /æ/ in GA		
Word	RP	GA
can't	[kɑːnt]	[k**æ**nt]
after	['ɑːftə]	['**æ**ftɚ]
bath	[bɑːθ]	[b**æ**θ]

e. British /æ/ pronounced as [e] ([ɛ]) in GA

The front low vowel /æ/ in a British accent in most words is pronounced similarly by

General American English speakers. However, a slight vocalic difference between RP and GA is that, if /ɹ/ occurs after /æ/, the pronunciation of /æ/ in some words in a British accent is sometimes shifted up to the mid vowel /e/ in the American variety (/**ɛ**/ is used for transcription in some dictionaries, but I adopt /e/ in accordance with the transcriptions in the *Cambridge English Pronouncing Dictionary*). For example, the name *Barry* is pronounced ['bæɹi] by British speakers with Received Pronunciation and ['beɹi] by speakers of General American English. However, the British pronunciation /'bæri/ is also an acceptable alternative for American speakers. Other examples are shown in Table 2.27.

Table 2.27 Vowel shift from RP /æ/ to GA /e/ (/ɛ/)

/æ/ in RP → /e/ (/ɛ/) in GA		
Word	RP	GA
carry	['kæɹi]	['k**e**ɹi]/['kæɹi]
barrister	['bæɹɪstə]	['b**e**ɹɪstɚ]/['bæɹɪstɚ]
barrow	['bæɹəʊ]	['b**e**ɹoʊ]/['bæɹoʊ]

f. R-colouring

The vowels preceding the consonant /ɹ/ are impacted by the postvocalic consonant /ɹ/. This phenomenon of vowels being influenced by the following /ɹ/ is termed r-colouring. All vowels in GA feature r-colouring when /ɹ/ occurs after vowels (Dimitrova, 2013), but this does not apply in the British accent, in the examples in Table 2.28.

Table 2.28 Differences between RP and GA in r-colouring

/V/ in RP → /V/+/ɹ/ in GA		
Word	RP	GA
smart	[sma:t]	[sm**a:**ɹt]
dark	[da:k]	[d**a:**ɹk]
pork	[pɔːk]	[p**ɔ:**ɹk]
court	[kɔːt]	[k**ɔ**:ɹt]

Among the vowels featuring r-colouring, the mid central vowels /ə/ and /ɜː/ are particularly strongly influenced. Due to the impact of the post-vocalic /ɹ/, the vowels /ə/ and /ɜː/ are often co-articulated with the following /ɹ/ as retroflex vowels in GA and transcribed as /ɚ, ɝ/. Although they are transcribed as /əɹ/ and /ɜːɹ/ in some dictionaries, the transcription /ɚ, ɝ/ is used in this research, based on the Cambridge English Pronouncing Dictionary. The r-colouring in vowels is a very prominent characteristic of GA. Examples with(out) r-colouring for the mid vowels /ə/ and /ɜː/ are seen in Table 2.29.

Table 2.29 Vowel shifts from RP /ə, ɜː/ to GA /ɚ, ɝ/

/ə, ɜː/ in RP → /ɚ, ɝ/ in GA		
Word	RP	GA
Peter	['piːtə]	['piːɾɚ]
pepper	['pepə]	['pɛpɚ]
hurt	[hɜːt]	[hɝt]
herd	[hɜːd]	[hɝd]
birth	[bɜːθ]	[bɝθ]

2.1.4.4.2.2 Diphthongal differences

The vocalic differences between RP and GA not only exist in monophthongs but also in diphthongs.

g. British /əʊ/ pronounced as /oʊ/ in GA

The diphthong /əʊ/ is pronounced in RP starting from the central schwa /ə/ to the high back lax rounded vowel /ʊ/, while in GA both elements are more fully back. Therefore, the British diphthong /əʊ/[6] is pronounced /oʊ/ in GA, gliding from a back mid vowel to a back high vowel, as shown in Table 2.30.

[6] It is worth noting that in RP /əʊ/ is usually (now) pronounced phonectically like [əʉ], so that the second element is more central. The pronunciation with a back [ʊ] is considered very conservative now. In order not to complicate it, only the phonological difference between RP /əʊ/ and GA /oʊ/ is introduced in the text above.

Table 2.30 Vowel shift from RP /əʊ/ to GA /oʊ/

/əʊ/ in RP → /oʊ/ in GA		
Word	RP	GA
over	['əʊvə]	['oʊvɚ]
boat	[bəʊt]	[boʊt]
snow	[snəʊ]	[snoʊ]
hope	[həʊp]	[hoʊp]

h. British diphthongs /ɪə, eə, ʊə/ are pronounced as /ɪɹ, eɹ, ʊɹ/ in GA

The letter *r* is always pronounced in GA, as mentioned above; therefore, due to the influence of /ɹ/, the three British centring diphthongs /ɪə, eə, ʊə/ are not present in GA. Instead, they are pronounced as [ɪɹ, eɹ, ʊɹ] with /ɹ/ replacing its original central mid vowel /ə/, as seen in the examples in Table 2.31. This is also an important difference in pronunciation between RP and GA.

Table 2.31 Vowel shifts from RP /ɪə, eə, ʊə/ to GA /ɪɹ, eɹ, ʊɹ/

/ɪə, eə, ʊə/ in RP → /ɪɹ, eɹ, ʊɹ/ in GA		
Word	RP	GA
dear	[dɪə]	[dɪɹ]
beer	[bɪə]	[bɪɹ]
bear	[beə]	[beɹ]
tear	[teə]	[teɹ]
tour	[tʊə]	[tʊɹ]
cure	[kjʊə]	[kjʊɹ]

2.2 Syllable structure

A syllable is a unit of a phonological groupings of segments (Davenport and Hannahs, 2013:245), and is considered to be the most basic unit of words. It is typically composed of a nucleus (often a vowel segment in most languages) preceded or followed by segments which are respectively called onsets or codas which are sometimes optional. The onset is the initial consonant(s) of the syllable and the coda can be the ending consonant(s). In different languages, the structures of syllables are distinctly constrained.

2.2.1 Syllable structure of Harbinese

In Chinese, each character corresponds to a syllable. The structure of a Chinese syllable generally comprises initials and finals (as addressed traditionally by Chinese linguists), which currently may also be termed onsets and rhymes by Western linguists. Initials of Chinese syllables must be consonants but can be optional. A syllable without an initial is called a zero initial or zero onset. A final can consist of medial glide plus rhyme, or just a rhyme, while rhymes can also be further divided into a nucleus (main vowel) and an ending which either may be a nasal consonant or a vowel. However, the nucleus is a compulsory segment in Chinese Mandarin and Mandarin dialects and the other components are optional. The structure of a syllable in Chinese can be generally illustrated in the form of CGVX (Lin, 2007:107), where C stands for consonant, G for glide, V for vowel, and X for C or V. The syntactic tree in Figure 2.10 clearly represents the Chinese syllable.

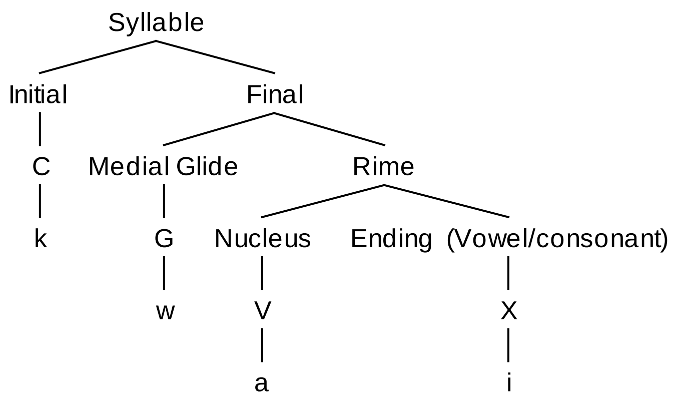


Figure 2.10 Syntactic tree of Mandarin syllable structure

All possible types of Standard Mandarin syllable structures are presented in (2.18) (Lin 2007:107):

(2.18)

1) V	e	[ɤ]	goose
2) CV	ba	[pa]	father
3) VC	an	[an]	safety
4) VV	ai	[ai]	love
5) CVC	dan	[dan]	egg
6) CVV	cai	[tsai]	vegetable
7) GV	wa	[wa]	frog
8) GVC	wan	[wan]	bowl
9) GVV	yao	[jɑʊ]	want
10) CGV	jia	[tɕja]	home
11) CGVC	duan	[twan]	short
12) CGVV	zhuai	[ʈʂwai]	pull

It is a conventional way which adequately represents the syllable structure of Mandarin. However, in this study, I am comparing two varieties of Chinese with English. I need to compare them using the same terminology. Thus, to be consistent in representing the syllable structure applied for different languages, CCVX is used to indicate the syllable structure in Mandarin/Harbinese. Therefore, there are nine types of syllable structure in Mandarin/Harbinese in my study shown in (2.19) below:

(2.19)

1) V	e	[ɤ:]	goose
2) CV	ba	[pa]	father
3) VC	an	[an]	safety
4) VV	ai	[ai]	love
5) CVC	dan	[dan]	egg
6) CVV	cai	[tsai]	vegetable
7) CCV	jia	[tɕja]	home
8) CCVC	duan	[twan]	short
9) CCVV	zhuai	[ʈʂwai]	pull

2.2.2 Syllable structure of Cantonese

Although Cantonese is one of the varieties of the Chinese language, it is different from Standard Mandarin (or Harbinese) in syllable structure. Compared with Harbinese/Standard Mandarin, Cantonese syllable structure is easier. The general pattern of Cantonese syllable structure is CVX. Like the syllable structure in Harbinese, C stands for consonant, V for vowel, and X for V or C. In Cantonese syllable structure, V is generally compulsory except for some words such as *ng* [ŋ̩] 'five', and others are optional. Thus, all the possible types of Cantonese syllable structure are shown in (2.20) (Wan, 2008:46):

(2.20)

1) V	i	[iː]	clothes
2) CV	faa	[faː]	flower
3) VC	aap	[aːp̚]	duck
4) VV	iu	[iu]	waist
5) CVC	gam	[kam]	gold
6) CVV	gai	[kɐi]	chicken

2.2.3 Comparison of syllable structure of the Harbinese and Guangzhou Cantonese

From the above two groups of syllable structures, it can be seen that the first six types exist in both Harbinese and Guangzhou Cantonese, but Guangzhou Cantonese lacks the types with the CC, i.e. types from 7) to 9), CCV, CCVC and CCVV. Thus, the big difference between the two concerns the part of prenucleus, onset. Mandarin allows onset clusters but Cantonese does not. Due to the phonotactic constriction in Mandarin/Harbiniese, the onset clusters are only constrained to C+glide type. To clearly show this specific phonotatic restriction, the CCV type in Mandarin/Harbinese could be further classified as CC_GV, where C_G stands for the glides restricted in this position.

In Mandarin and Harbinese, the coda is phonotactically constrained to only two nasal segments, [n] and [ŋ], but it can be not only the nasal consonants [m], [n], [ŋ] but also plosive consonants [p̚], [t̚], [k̚] in Cantonese. If the coda C is marked by the phonotactic restictions, such as C_N (nasal consonant) and C_P (plosive consonant), a CVC type can be further classified as CVC_N in Harbinese and CVC_N and CVC_P. Thus, Cantonese has two special types of syllable structures, VC_p and CVC_p, which Harbinese does not have. The main differences between Harbinese and Guangzhou Cantonese in syllable structure can be seen in (2.21) below:

(2.21)

Harbinese/Mandarin:

(1) CC_GV	duo	[xwai]	many
(2) CC_GVV	huai	[pʰjan]	bad
(3) CC_GVC	pian	[two]	cheat

Cantonese:

(4) VC_P	ap	[aːp̚]	duck
(5) CVC_P	sek	[sɛːk̚]	stone
(6) (C)VC_N [m]	sam	[sɐm]	heart

The general syllable structure can be concluded in the form of (C)(C_G)V(V/C_N) in Harbinese (or Mandarin) but (C)V(V/C_N/C_P) in Cantonese.

2.2.4 Syllable structure in English

English syllable structures are distinct from those in Chinese. An English word can have either one or many syllables, respectively called a monosyllable or a multi-syllable. An English syllable can be made up of an onset and rhyme. The rhyme can consist of a nucleus and coda. The nucleus is the main vowel in an English syllable. It can be a monophthong or a diphthong. The onset is the consonant(s) preceding the nucleus, which can have up to three consonants. The coda refers to the consonant(s) following the nucleus, with the maximum number up to four (Fatemi, Sobhani and Abolhassani, 2012:71). Taking the words *spring* [sprɪŋ] and *twelfths* [twelfθs] as examples, the onset of [sprɪŋ] is a complex of consonants comprising three [spr] while the coda of [twelfθs] is composed of four consonants [lfθs]. However, the onset and coda can be optional in an English syllable, whereas the nucleus is generally obligatory except in some particular onomatopoeic words such as *psst* [pst]. All possible types of English syllable structure are illustrated in (2.22), partially taken from Fatemi et al. (2012:71).

(2.22)

1) V	a	[ə]
2) VC	in	[ɪn]
3) VCC	opt	[ɒpt]
4) VCCC	asked	[ɑːskt]/[æskt]
5) CV	bee	[biː]
6) CVC	bull	[bʊl]
7) CVCC	help	[help]
8) CVCCC	text	[tekst]
9) CVCCCC	sixths	[sɪksθs]
10) CCV	spy	['spaɪ]
11) CCVC	sport	[spɔːt]
12) CCVCC	grand	[gɹænd]
13) CCVCCC	spends	[spendz]
14) CCVCCCC	twelfths	[twelfθs]
15) CCCV	stew	[stjuː]
16) CCCVC	stress	[stɹes]
17) CCCVCC	springs	[spɹɪŋz]
18) CCCVCCC	scripts	[skɹɪpts]
19) CCCVCCCC	strengths	[stɹeŋkθs]

2.3 Comparison of the sound systems of Chinese Harbinese and Guangzhou Cantonese and English, and possible difficulties for L2 learners

2.3.1 Comparison of consonants and possible difficulties

In section 2.1, the sound systems of Harbinese, Cantonese and English have been introduced. However, in order to conveniently see the differences between them, the consonant inventories of these languages/dialects are combined in Table 2.32.

Table 2.32 Comparison of English, Harbinese and Cantonese Consonants

Comparison of English, Harbinese and Cantonese in Consonants

Manner of articulation	Langs	Place of articulation																	
		Bilabial		Labiodental		Dental		Alveolar		Postalveolar/ Retroflex		Palatal		Velar		Labio-velar		Glottal	
Plosives /stops	E	p	b					t	d					k	g				
	H	p^h p						t^h t						k^h k					
	C	p^h p						t^h t						k^h k		k^{hw} k^w			
Fricatives	E			f	v	θ	ð	s	z	ʃ	ʒ							h	
	H			f	v			s		ʂ		ɕ		x					
	C			f				s										h	
Affricates	E									tʃ	dʒ								
	H							ts^h ts		tʂ^h tʂ		tɕ^h tɕ							
	C							ts^h ts											
Nasals	E		m						n						ŋ				
	H		m						n						ŋ				
	C		m						n						ŋ				
Lateral	E								l										
	H								l										
	C								l										
Approximants	E		w						ɹ				j						
	H		w								ɻ		j						
	C		w										j						

Note: The consonants to the left of the dotted lines are voiceless, and those on the right are voiced. The English consonant inventory is shown in bold, Harbinese in red, and Cantonese in blue. The English consonants with yellow shading do not have corresponding counterparts in either Chinese variety.

According to Lado's CAH, differences between two languages can lead to difficulties for L2 learners and the differences between these Chinese dialects/languages and how this is reflected in their L2 English are the focus of this research. Therefore, English sounds with complete correspondence with both Chinese dialects/languages and English sounds which are absent from both Chinese dialects/languages are not considered as research targets because they will be likely to generate the same production in both dialects/languages and speakers of both dialects/languages may have the same difficulties. Therefore, the focus of this research is on English consonants which correspond to the consonants of either dialect/language but not both.

It can be seen from Table 2.32 that 11 English consonants correspond to their counterparts in both Chinese dialects/languages. These include three voiceless stops /p, t, k/, one voiceless labio-dental fricative /f/, one voiceless alveolar fricative /s/, three nasals /m, n, ŋ/, one lateral /l/ and two approximants /j, w/. Therefore, these 11 consonants are excluded from being targets. Apart from these English consonants, most of the others have no correspondence in either dialect/language. As mentioned above, there is no point researching consonants absent from both varieties. However, only the English /h/ meets the criteria. It is identical with the Cantonese /h/. Wang (2007) found that Chinese orthographic Pinyin letters influence learners' English pronunciation. The Pinyin letters which are the same as English letters, such as *ch*, *sh*, *h*, *r*, lead learners of English to assume that they are pronounced in the same way or very similarly. The letter *h* is a velar fricative sound /x/ in Harbinese, but it is a glottal fricative /h/ in English. Therefore, the English /h/ sound is identical with Cantonese but only similar in Harbinese. Similarly, the letter *r* /ɻ/ in Harbinese is considered to be similar to /ɹ/ in English.

The English /ʃ, tʃ, dʒ/ are lacking in both Harbinese and Cantonese, but Kong and Wang (2001) found that the Harbinese/Mandarin retroflex /ʂ, tʂʰ, tʂ/ are perceived similarly to be the English /ʃ, tʃ, dʒ/. In addition, due to Harbinese and English sharing the same letters *sh, ch*, the fricatives /ʂ, tʂʰ, tʂ/ are always considered to sound similar to the English /ʃ, tʃ, dʒ/. Although Cantonese does not have the same phonemes as the English /ʃ, tʃ, dʒ/, Bauer and Benedict (1997) noted that Cantonese has the palatalised allophones [tʃ], [tʃʰ] and [ʃ] of the phonemes /ts/, /tsʰ/, and /s/. In this research, the English /ʃ, tʃ, dʒ/ are regarded as similar to items in each dialect/language but in different ways.

The English /v/ is absent in both dialects/languages. However, Shi (1996) found that Mandarin speakers often confuse /v/ and /w/ and often pronounce /v/ as [w]. Gong and Tian (2008) and Chen (2010) confirmed this difficulty and found English learners in the

Northeast to substitute /w/ and /v/ for each other. Although Harbinese does not have the English phonome /v/, it has an allophonic [v] of the phoneme /w/. In addition, Cantonese speakers are reported to use the different sounds [f] or [w] to substitute for the English /v/ (Chan and Li, 2000). Therefore, it is worth checking the production of the consonant /v/ in this research.

The consonants /θ, ð, ʒ/ are absent from both dialects/languages. However, it was reported by Li (2006) that L2 English learners had great difficulty in pronouncing these consonants. Moreover, although both Harbinese and Cantonese lack these two consonants, they pronounce them differently. Chan and Li (2000:79) mentions that Cantonese learners of English commonly substituted [t] or [f] for /θ/ and [d] or [f] for /ð/, while Chen (2012:12) and Gong and Tian (2008:31) found that Harbinese speakers often substituted [s] and [ts] for /θ/ and /ð/. As Eckman (1977) claimed, marked sounds will result in difficulties in acquisition. The English /θ/ and /ð/ are the most marked sounds and are rare cross-linguistically (Wells, 1982:96). Differential substitutions of /θ/ and /ð/ have attracted close attention among researchers. Therefore, these three consonants are investigated in this thesis.

The nine English consonants /v, ʃ, ʒ, tʃ, dʒ, h, ɹ, θ, ð/ are of greatest interest for the present research. Apart from these consonants, another three consonants are also worthy of attention. They are the English consonants /w/, /n/, and /l/. Although they exist in both dialects/languages, they cause problems even in their own dialects/languages. For example, it has been noted (Matthews and Yip, 1994:16; Chan and Li, 2000:71; Gui, 2005:37) that Cantonese speakers, especially young speakers, confuse the initial consonants /n/ and /l/ and neutralise them, and they found that /n/ tends to be pronounced as [l]. However, the Harbinese can differentiate them. Therefore, it is worth taking the English nasal /n/ and lateral /l/ into consideration in this research in terms of possible distinctions between Harbinese and Cantonese speakers in their acquisition of English. Interestingly, Cantonese speakers have difficulties distinguishing /l/ not only from /n/ but also from /ɹ/, as mentioned by Chan et al. (2000:80). Some Cantonese learners of English substitute /l/ or /w/ for /ɹ/. However, there is no evidence that Harbinese speakers cannot distinguish /l/ from /ɹ/ in English, except in some cases where some Harbinese speakers may substitute the Chinese /l/ for /ɻ/. For example, /l/ in [ləŋ] is substituted for /ɻ/ in [ɻəŋ] 'throw', although the Chinese /ɻ/ is different from the English /ɹ/ but with some similarity. Thus, the English /ɹ/ is shown here in comparison with /n/ and /l/. Another point to note is that the consonant /1/ in English has two allophones: the clear [l] used as a prevocalic onset and the dark [ɫ] used as a postvocalic coda, but /l/ in Chinese has no dark [ɫ] allophone. Therefore, the dark [ɫ] can be difficult for speakers of both dialects/languages to acquire.

The English /w/ also exists in both dialects/languages. However, Harbinese speakers tend to use /v/ to replace /w/ in Harbinese. In research on the initial /w/, Wang (2011) found that Harbinese speakers have a tendency to pronounce it as [v]. This implies that the Harbinese /w/ has two allophones: [w] and [v]. However, no study has identified this phenomenon in Cantonese. Thus, there are differences here between Harbinese and Cantonese.

The consonants discussed above are therefore selected as research targets to see how speakers of the two dialects/languages produce them and what difficulties they experience. These consonant targets are presented in Table 2.33 to clearly show the differences between the two dialects/languages and English.

Table 2.33 English consonant research targets

a

English	/v/	/w/	/ʃ/	/tʃ/	/dʒ/	/h/	/ʒ/	/θ/	/ð/
Harbinese	[v]	/w/ [w][v]	/ʂ/	/ʈʂʰ/	/ʈʂ/	/x/			
Cantonese		/w/	[ʃ]	[tʃʰ]	[tʃ]	/h/			

b

English	/n/	/l/ [l] - [ɫ]	/ɹ/
Harbinese	/n/	/l/ [l]	/ɻ/ [ɻ]->[l]
Cantonese	/n/ [n]->[l]	/l/ [l]->[n]	(English: [ɹ]-> [l] /[w])

As mentioned above, English /l/ has a clear [l] occurring word initially and a dark [ɫ] which occurs finally. It clearly indicates that the learning difficulties of L2 segments may be caused by different positions in a syllable or in a word, for example, there is no word-finally dark [ɫ] in Harbinese Mandarin and Guangzhou Cantonese. In addition, in discussing the reasons for the error type of substitution, Lombardi (2003) claims that the substitution is not led by the phonetic sounds around the L2 segment but by the position of the L2 target. Zampini (2008) agrees with her and mentions that the transfer of L1 syllable constraints also brings the difficulties to the acquisition of L2 segments in L2 acquisition. She thinks that German can be a good example to confirm this because L1 German speakers of Engligh often devoice the English word-finally voiced stops, which would be influenced by L1 syllable constraints. Hansen also (2004) notes that L1 transfer has an important impact on the production of segments in different positions of a syllable. Broselow, Chen and Wang (1998) found evidence that Mandarin speakers always insert a schwa after a stop, or delete the stop consonants, or devoice the word-final voiced stops in English production. Whether Cantonese speakers would make the same errors or error types in word-final position like Mandarin speakers? Also, Cantonese speakers will netrulise word-initial /n/ and /l/ mentioned above, does that mean Harbinese Mandarin speakers would do the same? The segments in different word positions should be an interest in this thesis for further exploration. Therefore, position is considered in this study by looking at allophones for segments, and onset and coda for syllables.

2.3.2 Comparison of vowels and possible difficulties

2.3.2.1 Monophthongs

The same types of comparison as for consonants is given here for vowels. As stated by Yavas (2016:19), "identically transcribed vowels from different languages may not be the same phonetically". Therefore, it should be noted that the phonemes shown in Figure 2.11 are not phonetically identical. Therefore, Figure 2.12 gives their phonetic realisations. The same transcribed vowels are shown together to show the differences between languages and dialects in vowels.

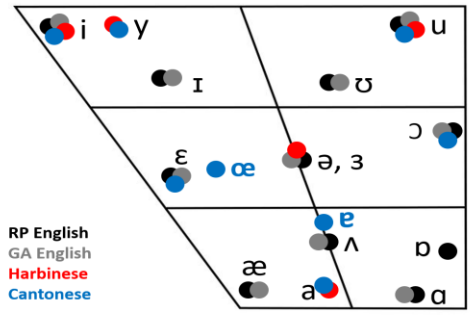


Figure 2.11 Comparison of RP English, GA English, Harbinese and Cantonese in phonemic vowels

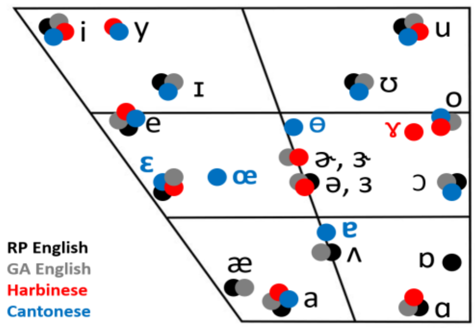


Figure 2.12 Comparison of RP English, GA English, Harbinese and Cantonese in phonetic vowels

Because the English vowel systems in RP and GA vary, as introduced in section 2.1, two different colours are used to represent the different varieties of English, with black for RP and grey for GA. Red and blue are for Harbinese and Cantonese respectively. Figure 2.11 indicates that there are few phonemic differences between the two English varieties. The only difference is that GA lacks the phoneme /ɒ/. Various phonetic differences of representations of vowels are shown in Figure 2.12.

Phonemically, Harbinese has five vowels while Cantonese has three more. The two Chinese dialects/languages have the same four vowel phonemes which are the high vowels /i, y, u/ and a low vowel /a/. Among these four, the two unrounded high vowels /i, u/ overlap with English vowel phonemes and are the only two phonemes existing in all dialects/languages. The distinctions between Harbinese and Cantonese vowels lead us to predict differences in difficulties in the pronunciation of English. Harbinese has a phoneme /ə/ which occurs in the English vowel system but not in Cantonese. In this case, the English schwa /ə/ is predicted to be difficult for Cantonese learners to acquire but not for Harbinese learners. Similarly, Cantonese has four vowels /ɐ, ɛ, œ, ɔ/ which do not exist in Harbinese. One of these, /ɔ/, corresponds to the English /ɔ/; thus, the acquisition of English /ɔ/ may be more difficult for the Harbinese than the Cantonese. The Cantonese lax or short vowel /ɐ/ is a low central vowel, resembling the English central lax /ʌ/ (Matthews and Yip, 1994:18). Therefore, some researchers (Gui, 2005:52-53) transcribe it as /ʌ/ but others (Hashimoto, 1972:90; Flynn, 2001:3) still use /ɐ/ instead for the Cantonese phoneme. Due to the similarity between the Cantonese /ɐ/ and English /ʌ/, Cantonese learners of English are predicted to have less difficulty than Harbinese learners in learning /ʌ/.

The other six English vowel phonemes in the Figure 2.11, /ɪ, ʊ, e, æ, ɒ, ɑ/, do not occur in either Chinese dialect/language. Tense and lax vowels such as /i/ vs. /ɪ/ and /u/ vs. /ʊ/ are contrasted in English, while neither Chinese dialect/language contrasts them. However, as seen in Figure 2.12, the lax vowels [ɪ] and [ʊ] exist in Cantonese as allophones but they do not exist in Harbinese even as allophones. In that case, it seems that it may be comparatively easier for Cantonese English learners to acquire lax vowels than Harbinese learners even though these phonemes are absent in both Chinese dialects/languages.

Besides the high lax /ɪ/ and /ʊ/, the mid /e/ and three other low vowels /æ, ɒ, ɑ/ are missing in both Chinese dialects/languages. The low lax vowels /æ/ and /ɒ/ do not have corresponding counterparts in Chinese even as allophones. Therefore, these are predicted to be difficult for both dialectal learners of English. Although neither the English tense vowels /e/ or /ɑ/ occur in either dialect/language, they are allophones of Chinese varieties as seen in the Figure 2.12. Both Harbinese and Cantonese have the allophone [e], which is always realised as a part of the diphthong [ei] in these two dialects similarly to the English [eɪ], but [e] as an allophone arises from different phonemes in Harbinese and Cantonese. In Harbinese, [e] is one of the allophones of the phoneme /ə/, while it is an allophone of the phoneme /ɛ/ in Cantonese. Similarly, [o] as shown in Figure 2.12 is an allophone of the phoneme /ə/ in Harbinese but of the phoneme /ɔ/ in Cantonese. Distinct from English, Harbinese represents /ɑ/ as an allophone ([ɑ]) of the phoneme /a/ and [ɑ] is always represented by the combination with [u] or [ŋ] as the diphthong [ɑu] and a rhyme [ɑŋ], while Cantonese does not have [ɑ] even as an allophone.

If the phonetic realisations of vowels in their L1 dialects can be considered as a factor assisting Chinese learners of English to acquire the English vowel inventory, these English vowel phonemes /ɪ, ʊ, ə, ɔ, æ, ʌ, ɒ, ɑ/ will be somewhat difficult for Harbinese or Cantonese learners. Cantonese does not have /ə, æ, ɒ, ɑ/, and they may be difficult for Cantonese learners; /ɪ, ʊ, ɔ, æ, ʌ, ɒ/ are absent in Harbinese and may be difficult for Harbinese learners. Apart from the lack of tense vowels /ɑ/ in Cantonese and /ɔ/ in Harbinese, it is predicted to be difficult for Chinese learners of English to acquire English lax vowels. However, it may be more difficult for Harbinese learners to acquire them because Harbinese does not contrast tense and lax in vowels but Cantonese does, although the only main contrasted pair is /a:/ and /ɐ/. In addition, Cantonese has lax vowels [ɪ] and [ʊ] as allophones.

2.3.2.2 Diphthongs

RP has eight diphthongs but GA has five. The three centring diphthongs in RP do not exist in GA because of the realisation of rhotics in GA mentioned in section 2.14. In addition, neither Chinese dialect/language has the centring diphthongs of RP; therefore, only five English diphthongs are targeted in this research. The targeted diphthongs with two Chinese dialects/languages are compared in Table 2.34.

Table 2.34 Correspondences of English diphthongs in Harbinese and Cantonese

English	/eɪ/	/aɪ/	/aʊ/	/əʊ/ (/oʊ/)	/ɔɪ/
Harbinese	/ei/	/ai/	/ɑu/	/ou/	
Cantonese	/ei/	/aːi/	/aːu/	/ou/	/ɔːi/

The English diphthongs [eɪ], [aɪ], [aʊ] and [əʊ] ([oʊ] in GA) are similar in perception to [ei], [ai], [ɑu] and [ou] in Harbinese and to [ei], [aːi], [aːu] and [ou] in Cantonese. Also, Cantonese has one more similar diphthong [ɔːi] that corresponds to the English diphthong [ɔɪ] but Harbinese does not have this.

The main difference in diphthongs between the two dialects/languages and English is that English diphthongs end with the lax vowels [ɪ] and [ʊ] while the height of ending vowels in both dialects/languages' diphthongs can reach as high as a tense vowel [i] or [u]. Irrespective of this, these Cantonese diphthongs have greater similarity with English diphthongs in number and similarity than those in Harbinese. Cantonese contrasts tense and lax vowels, a distinction only seen in [a:] and [ɐ], and also they are distinct in length, with [a:] being longer. Besides these two points, all six of the other Cantonese vowel phonemes have long and short allophone pairs. The long allophones of phonemes in Cantonese are always transcribed with a colon diacritic to show the length. Therefore, apart from [ei] and [ou], other diphthongs with the diacritic indicate they have longer length and the first vowel of the diphthong is pronounced longer than in English and Harbinese diphthongs. Therefore, Cantonese speakers are predicted to pronounce English diphthongs as longer. Although there is a similar correspondence between the English [aʊ] and [oʊ] in Harbinese, the Harbinese diphthong [ɑu] has a back low vowel [ɑ] distinct from English diphthong [aʊ] with a central low [a] due to the feature of backing. Also, due to the preference of the substitution of [ɤ] for [o] in the Harbinese the use of [ɤu] to replace [ou] is slightly different from the GA English [oʊ] in perception. Thus, Cantonese speakers are predicted to have fewer difficulties than Harbinese in acquiring English diphthongs.

Predictions: Based on the above comparison of vowels in English and the two groups of Chinese speakers, we see that Cantonese has more vowel phonemes and comparatively more vowels similar to English than Harbinese; therefore, it may be predicted that Cantonese speakers, but not Harbinese, will be more successful in their acquisition of English vowels.

2.3.2.3 Tense and lax

English contrasts tense and lax vowels such as /i:/ vs. /ɪ/ and /u:/ vs. /ʊ/ and Cantonese only contrasts one phonemic pair /aː/ vs. /ɐ/; however, Harbinese does not (see Table 2.35). In addition, Cantonese has other allophonic tense-lax vowel pairs but Harbinese does not.

Table 2.35 Comparison of vowel tenseness between English, Cantonese and Harbinese

Length / Language	tense	lax
English	/i:, ɜ:, u:, ɔ:, ɑ: eɪ, oʊ/	/ɪ, e(ɛ), æ, ə, ʌ, ʊ, ɒ/
Cantonese	/aː/ and [i:, u:, ei, ou, ɔ:]	/ɐ/ and [ɪ, ʊ, œ:, ɛ:, ɵ]
Harbinese	No contrast: /i, y, u, ə, a/	

Prediction: In the light of the differences between Cantonese and Harbinese in vowel tenseness, it is predicted that Cantonese speakers will acquire English tense-lax vowels but Harbinese speakers will not acquire this distinction.

2.3.2.4 Vowel length

The length of English vowels varies, with longer length for tense vowels and shorter length for lax vowels even though the tense-lax and long-short dimensions are not the same. RP English has five long vowels which are always transcribed with a colon diacritic to show length. Cantonese has long and short vowels which are not only represented by one main phoneme pair, as seen in /a:/ and /ɐ/, but also by allophone pairs of the other six phonemes shown in Table 2.36. However, Harbinese has no contrastive phonemes in length.

Table 2.36 Comparison of vowel length in English, Cantonese and Harbinese

Length / Language	long	short
English	/i:, ɜ:, u:, ɔ:, ɑ:/	/ɪ, e(ɛ), æ, ə, ʌ, ʊ, ɒ/
Cantonese	/a:/ and [ɛ:, i:, ɔ:, œ:, u:, y:]	/ɐ/ and [e, ɪ, o, ɵ, ʊ]
Harbinese	No contrast: /i, y, u, ə, a/	

Although, as mentioned above, Cantonese and Harbinese have similar correspondences to English diphthongs such as /aɪ/ and /aʊ/, the first vowels in these diphthongs are articulated differently in terms of duration. Cantonese diphthongs are transcribed with length diacritics such as [aːi] and [aːu] to indicate that the first vowel of diphthongs is pronounced for longer, but English and Harbinese diphthongs seem not to be pronounced for such a long perioid. The duration of the first vowel in English and Harbinese diphthongs is also different, and in the English [tʰaɪ] *tie*, the first vowel is articulated as longer than in the Harbinese [tʰai] *tai* 'very'.

Prediction: From the above discussion, we know Cantonese contrasts the vowel length and the duration of same diphthongs are different. Based on these patterns of tense vowels, I predict that Cantonese speakers will acquire English vowel length but not

Harbinese speakers.

2.3.3 Comparison of syllable structure and possible difficulties

From examples (2.22) in Section 2.2.4 Syllable structure in English, there are up to 19 types of English syllable structure, which is much richer than Harbinese and Cantonese. However, it can be observed that four types of English syllable structure, 1) V, 2) VC, 5) CV and 6) CVC, overlap with those both in Harbinese and Cantonese, and another two types, 10) CCV and 11) CCVC overlap with Harbinese.

The most substantial difference between English and Chinese in syllable structure is that English allows complex consonant clusters in onsets and codas. This is prohibited in Chinese except C+glide (CC_G) onset clusters in Mandarin/Harbinese. Therefore, consonant clusters should be very difficult for speakers of either Chinese variety except CC_G for Harbinese.

In addition, the number of English syllables in a word mainly depends on the number of vowels; therefore, a monosyllable includes one vowel represented by the capital letter V in English syllable structure. However, the V in English syllable structure can be a monophthong or a diphthong. The diphthong is considered to be a constituent of an English syllable while in Chinese it is counted as two segments because syllables in Chinese do not rely on the number of vowels. Therefore, English syllables can have one vowel position which can be a diphthong or a monophthong, but one vowel position in Chinese syllable peaks can only be a monophthong. This means a diphthong in Mandarin has to be represented by VV. For sake of clarity and consistency between Chinese and English in this study, the segment V is denoted by V(V) for the diphthong in illustrations of English syllable structure.

Table 2.37 combines the general form of syllable structures among Harbinese, Cantonese and English and Table 2.38 illustrates their specific syllable structures.

Differences among English and Chinese dialects/languages are shown.

Table 2.37 General syllable structures in Harbinese, Cantonese and English

Languages/Dialects	General Syllable Structures
Harbinese	(CC) + V + (X)
Cantonese	(C) + V + (X)
English	(CCC) + V + (CCCC)

Parenthesised segments are optional.

Table 2.38 Specific syllable structures in Harbinese, Cantonese and English

Languages/Dialects	Specific Syllable Structures
Harbinese	(C)(C_G) + V + (V/C_N)
Cantonese	(C) + V + (V/C_N/C_P)
English	(C)(C)(C) + V(V) + (C)(C)(C)(C)

Parenthesised segments are optional.

Difference 1: Differences among Harbinese, Cantonese and English syllable structures

Harbinese has three glide /j, w, ɥ/ but Cantonese has two /j, w/. The glides in Harbinese can only occur on the second C of CCV(C) while Cantonese does not allow any consonant before a glide, thus, Cantonese does not have a CC complex onset. The C+glide (CC_G) onset is the obvious distinction between Harbinese and Cantonese syllable structures. English has only two glides: /j/ and /w/ and the onset can have up to three consonants. Therefore, English pattern (C)(C)V(C) can correspond to the Harbinese syllable pattern (C)(C)V(C), but the second C in the English pattern must be constrained to /j/ or /w/ to conform to the glide restrictions in Harbinese, that is, /C+j+V+(C)/ (mostly in RP) or /C+w+V+(C)/. Although Cantonese has glides /j/ and /w/, it does not have the structure type with CC onset. Thus, it is predicted that it will be difficult for Cantonese speakers to acquire English syllables with glides, i.e. /C(C_G)V(C)/.

All consonants in English except /h/ and /j, w/ (+/ɹ/) can be codas but limited consonants can occur in coda position in Chinese. Due to the different constraints on syllable coda positions in Harbinese and Cantonese, differences in production are predicted when pronouncing English with codas. As mentioned above, Harbinese syllable codas can only have [n] and [ŋ] while in Cantonese six are possible [m], [n], [ŋ], [p], [t], and [k]. Therefore, it should be easier for Cantonese than Harbinese to acquire English words with the syllable structure pattern CVC_N and CVC_P because Cantonese can have more consonants in the coda. However, Cantonese speakers will have difficulties in pronouncing English with the CVC_P pattern. They do not release stops (Jia, 2011) and listeners might not perceive there is a stop. This is influenced by negative transfer from Cantonese where stop codas are unreleased. Although Cantonese speakers will have difficulties with the CVC_P syllable pattern, they will not have such difficulty with the CVC_N pattern with [m] as a coda. Therefore, Harbinese speakers are still predicted to experience more difficulty in pronouncing English with the CVC_N and CVC_P patterns.

Difference 2: Similarities between Harbinese and Cantonese syllable structures but which are distinct from English syllable structure

The vowel in the English syllable structure is represented by a single V , and this can be a monophthong or a diphthong, as mentioned above. This is different from Harbinese and Cantonese syllable structure, in which V only represents a monophthong and a diphthong is indicated by VV. These different representations are led by the phonotactic constraints on Chinese syllable structure, that when there is a diphthong in a syllable it must be an open syllable (no coda is allowed), so the pattern is (C)V((V)/(C)), whereas a diphthong in an English syllable can have a coda so the pattern is (C)V(V)(C). To be consistent between English and Chinese, V(V) is used to represent a diphthong in this study, as mentioned above. I predict that Chinese speakers will have difficulties in the acquisition of this type of English syllable structure, i.e., a diphthong plus a consonant.

In addition, English can have rich consonant clusters in onsets and codas but Harbinese and Cantonese cannot except that Harbinese only allows C+glide onset clusters. Therefore, it is very difficult for both dialect/language speakers to pronounce complex English onsets and codas. All possible types of combinations of consonants as onsets and codas are illustrated in Table 2.39.

Table 2.39 Possible phonotactic sequences in English

Onset(s)	V	Coda(s)
All single consonant phonemes except /ŋ/	/iː/, /ɪ/, /e/, /æ/; /ɜː/, /ə/, /ʌ/; /uː/, /ʊ/, /ɑː/, /ɔː/, /ɒ/	The single consonant phonemes except /h/, /w/, /j/ and /ɹ/ in RP
/pl/, /bl/, /kl/, /gl/, /pr/, /br/, /tr/, /dr/, /kr/, /gr/; /tw/, /dw/, /gw/, /kw/, /pw/; /fl/, /sl/, /θl/; /fr/, /θr/, /ʃr/; /hw/, /sw/, /θw/, /vw/; /pj/, /bj/, /kj/, /gj/, /mj/, /fj/, /vj/, /hj/; /sp/, /st/, /sk/; /sm/, /sn/; /sf/, /sθ/; /spl/, /skl/, /spr/, /str/, /skr/, /skw/, /smj/, /spj/, /skj/; /sfr/; In RP: /tj/, /dj/, /nj/, /θj/, /sj/, /stj/, /zj/, /lj/, /stj/.	/eɪ/, /aɪ/, /ɔɪ/; /aʊ/, /əʊ/; In RP: /ɪə/, /eə/, /ʊə/.	/lp/, /lb/, /lt/, /ld/, /ltʃ/, /ldʒ/, /lk/; /lf/, /lv/, /lθ/, /ls/, /lʃ/; /mp/, /nt/, /nd/, /ntʃ/, /ndʒ/, /ŋk/; /mf/, /mθ/, /nθ/, /ns/, /nz/, /ŋθ/; /ft/, /sp/, /st/, /sk/, /fθ/; /pt/, /kt/; /pθ/, /ps/, /tθ/, /ts/, /dθ/, /ks/; /mpt/, /mps/, /ndθ/, /ŋkt/, /ŋks/, /ŋkθ/; /ksθ/, /kst/; In GA (rhotics): /rp/, /rb/, /rt/, /rd/, /rtʃ/, /rdʒ/, /rk/, /rg/; /rf/, /rv/, /rθ/, /rs/, /rz/, /rʃ/; /rmθ/, /rpt/, /rps/, /rts/, /rst/, /rkt/.

Predictions: Except the C+glide onset in Harbinese, Cantonese and Harbinese can be generally regarded as no consonant clusters, thus, speakers of both dialects/languages may have difficulties with English consonant clusters.

2.3.4 Comparison of stress and possible difficulties

As introduced above in section 2.1.4.3, we know that English stress patterns are variable, including final stress, penultimate stress, and antepenultimate stress, but its main foot pattern is trochaic. According to the discussion on Mandarin/Harbinese stress in section 2.1.3.4, Harbinese stress is final in 70% of cases with an iambic foot, whereas initial stress pattern represents 30% with a trochaic foot. However, there is insufficient evidence to indicate that Cantonese has any stress patterns.

Predictions: Given the Harbinese stress patterns, it can be stated that Harbinese has two main prosodic templates: an iambic foot and a trochaic foot, while English is a language with mainly trochaic feet. Therefore, I predict that Harbinese speakers will impose iambic stress on disyllabic English words more than Cantonese speakers.

2.4 The second language acquisition of phonology

Since the first contact of those speaking different languages, understanding a foreign language and learning a new language have been works in progress. In the modern period of globalisation, learning a second or even a third language is not uncommon. In the process of learning and teaching a new language, many issues concerning L2 acquisition have arisen, and been enthusiastically researched, such as L1 influence, or transfer. Transfer is "an effect of previously learned languages on subsequently learned languages", and its examination play a significant role in theory construction and pedagogical developments in L2 phonology (Edwards and Zampini, 2008:2). In the process of analysing second language phenomena and accounting for second language learning, second language acquisition has gradually developed into a separate academic discipline.

The study of second language acquisition as an independent discipline began in the late 1960s and early 1970s (Gass, Behney and Plongsky, 2013). The construction of theoretical systems was based mainly on the description of the second language acquisition process and its characteristics. In the last 40 years, the acquisition of second language phonology has been widely researched. Comparisons of sounds in the L1 and L2 were focused on in research on the similarity of L1 and L2 sounds. Subsequently, the research focus has varied from language transfer to developmental factors in second language acquisition under the framework of universal grammar. A great number of important theoretical insights into second language acquisition in phonology have emerged; for example, Lado's Contrasive Analysis Hypothesis (CAH) (1957), Corder's Error Analysis (EA) (1967), Selinker's Interlanguage Theory (IL) (1972), Eckman's Markedness Differential Hypothesis (MDH) (1977), Major's Ontogeny Model (1987) and Ontogeny Phylogeny Model (OPM) (2001), Best's Perception Assimilation Model (PAM) (1993), Flege's Speech Learning Model (SLM) (1995) and Optimality Theory (Prince and Smolensky, 1993). All of these ideas, hypotheses, and models are devoted to the development of the discipline and the ability to account for important issues in the process of L2 acquisition.

Observations of transfer had existed long before (see Thomas 2013), but it was not formalised until the seminal writings of Fries (1945), Weinreich (1953), and Lado (1957). Since then, transfer has come to be seen as most prevalent in the area of L2 phonology where it is considered a dominant influence. Among these three works, the most influential one was Lado' Contrastive Analysis Hypothesis (1957). It aimed to find similarities and differences between an L1 and L2 so that predictions could be made that L1 sounds that are similar to L2 sounds will be easy to learn but different sounds will be difficult. Larsen-Freeman and Long (1991:53) frame this as, "where two languages [are] similar, positive transfer [will] occur; where they [are] different, negative transfer, or interference, [will] result". According to the CAH, all errors made by L2 learners result from L1 transfer and the CAH could be used to predict all difficulties. However, this hypothesis was critisised by an increasing number of researchers and teachers in the early 1970s because they found that the CAH could not predict all learners' difficulties and not all errors were caused by L1 transfer.

Since the CAH is unable to predict all of the errors that L2 learners make in the process of L2 acquisition, the research focus shifted in the late 1960s and the early 1970s from predictions of difficulty to the analysis of learners' language errors (Gass et al., 2013; Hou, 2011). The systematic analysis of the language errors made by L2 learners revealed the nature of the acquisition process of a second language (ibid.). To further explain and validate Lado's CAH, Corder (1967) proposed Error Analysis which can be said to be an alternative to Lado's CAH. This approach stressed that learners' errors are crucial in offering evidence for understanding not only how languages are acquired but also for the strategies and procedures used by learners. Corder (ibid.) also believed that errors in second language acquisition arise not only from L1 interference in the sense of interlingual transfer from L1, but also of intralingual transfer from the target language.

However, Error Analysis also encountered limitations where the method looked at only where learners made errors but ignoring where they did not err (Gass, 2013; Fan and He, 2012:113). In addition, avoidance through communication strategies is more common in second language acquisition than first language acquisition. EA cannot explain the avoidance phenomenon (Schachter, 1974:212). She found that if the L2 was similar to L1, native language strategies were transferred to the target language while, if the L2 was extremely different, L2 learners may reject, or not attempt to produce, targets they find difficult, and then there will be no errors to be analysed. Apart from these problematic areas, Gass et al., (2013:94, 96) also pointed out that EA has problems to determine 'what an error is an error of', i.e. whether the source of an error is the L1 or based on development/universals.

Because both CA and EA have theoretical and practical deficiencies, researchers have continued to seek new research methods to account for issues in L2 acquisition. The focus shifted from contrasting two languages and analysing errors to the language of language learning itself, or 'interlanguage'. Such researchers (Selinker, 1969, 1972; Corder, 1971; Nemser, 1971) think that during the process of the acquisition of the L2 learners undergo a series of language transitions starting from the L1 to the end of the successful acquisition of the target language (L2). The language L2 learners speak in the transitional phases of second language development is called interlanguage (IL), which differs from both the native language (L1) and the target language (L2), but which is a transitional language between the mother tongue and the target language. Different scholars have different names for interlanguage, for example "learners' languages" or "idiosyncratic dialects" (Corder, 1971), and "approximative system" (Nemser, 1971), but the concept of IL was first formally proposed by Selinker (1972). Selinker (1992) stressed that the IL is an observable speech output and a highly structured and independent language. Ellis (1994) pointed out that interlanguage theory is an appropriate starting point for the study of second language acquisition. It was the first theory to provide an explanation of second language acquisition. Therefore, the concept of interlanguage gives the field of second language acquisition its own independent research direction.

In contrast to Lado's CAH, Best (1993, 1994) and Flege (1995) thought that the more similar L2 sounds are to those of L1, the more difficult it will be for the L2 sounds to be perceived by L2 learners, and thus the accurate acquisition of L2 sounds will not be attained. Since the 1980s, many studies of L2 speech perception have been conducted (Best, 1993; Flege, 1995; Kuhl, 1995; Strange, 1995) to explain the perception of non-native speech and its language experience effect. In research into speech perception, two significant models have been proposed, namely the Perceptual Assimilation Model (PAM) (Best, 1993) and Speech Learning Model (SLM) (Flege, 1995).

Best's PAM (1993) mainly focuses on the effect of the native language experience of listeners on the perception of non-native speech. The model proposes that when a sound in unfamiliar non-native speech is heard, the listener will perceive it according to the closest sounds in his/her native language, and eventually the non-native sound is assimilated to a similar phonetic sound in the native language. However, if the non-native sound cannot be perceived with a close native sound, and it is perceived as very different from the native language sound, then the sound cannot be assimilated into the native sound system and so it will be attributed to a non-assimilable sound. In their research on Zulu click discrimination by English-speaking adults and infants, Best and her colleagues (1988) found that the click in Zulu is regarded in the same way as non-speech sounds by English native speakers. The PAM (Guion, Flege, Akahane-Yamada, and Pruitt, 2000:2712-2713) points out that listeners may assimilate the non-native phonetic perception to the mother tongue by detecting similarities to native language sounds based on perceived articulatory properties including articulators, blockage location, and degree of constriction. The main factors affecting the perceived outcome, and differences in discriminability, are the similarity and perceived distance between the non-native sounds and the native sounds (ibid. 2000:2713). In other words, if a foreign sound is perceived as very similar to a sound in the native language, or the perceived phonetic distance between this foreign sound and the native sound is very small, the foreign sound is easier to assimilate into the closest sound of the native language; thus it is rather difficult to accurately perceive the foreign sound and this new sound could not be accurately produced.

Based on Best's PAM (1993), Flege (1995) proposed the Speech Learning Model (SLM). This focuses on the ultimate attainment of L2 pronunciation and claims that, if there is no accurate L2 speech perception, there will be no accurate production of L2 speech sounds. This means that L2 speech perception affects pronunciation. The proposal of the SPM enables the study of second language speech to achieve cross-language integration. Moreover, it can be used to analyse phonetic differences and similarities between the L1 and L2 from the perspectives of both perception and production. A comprehensive introduction to this model is given below.

To increase explanatory power, cross-linguistic influence and universal factors are combined to account for issues in SLA. Eckman (1977) attempted to combine the influence of markedness and cross-linguistic factors, thereby forming a new theoretical idea called the Markedness Differential Hypothesis (MDH). The MDH (Eckman, 1977) claims that difficulties in the process of L2 acquisition mainly result from the elements in L2 which are distinct from those in the L1. Therefore, these L2 elements are more marked to the L1 and so it is more difficult for L2 learners to master these L2 elements.

A further development which needs to be mentioned is Major's (1986) Ontogeny Model (OM), which combines cross-linguistic factors and universal factors, i.e. focuses on the relationship between transfer and development process. This model (Major, 1986:453) claims that "transfer processes decrease over time, while developmental processes increase and then decrease." Major's research on the L2 acquisition of the Spanish *r* fits his model in that, with an increase of exposure time, the transfer of the English /ɹ/ to the Spanish /r/ and /ɾ/ is reduced and rates of developmental errors display upward and then downward trends (Major, 1986). In his Ontogeny Phylogeny Model (Major, 2001), which is a new version of the OM, the phonological relationships between L1 and L2, as well as universals are discussed and both transfer and markedness are regarded as key factors which affect the acquisition of L2 phonology. He (ibid.:82) states the pattern of development for IL in his OPM is that "L2 increases, L1 decreases, and U (universals) increases and then decreases".

A powerful recent theory in research into phonology is Optimality Theory (OT, Smolensky and Prince, 1993), which "dispenses with phonological rules and proposes that the relationship between an underlying form and its surface realisation is not derivational in nature. It proposes instead that underlying forms are linked directly to surface forms by means of a set of constraints" (Davenport and Hannahs, 2013:244). OT was first proposed by Alan Prince and Paul Smolensky at an Arizona Phonology Conference in 1991 (Archangeli and Langendoen, 1997:1). Although OT originated from the tradition of generative grammar, it is very distinct from traditional derivational phonology. As a newly-developed theory for phonological research, OT can explain previously unsolvable phonological problems in some languages through the analysis of constraint interaction which traditional derivational phonology cannot. OT has since increasingly prevailed not only when applied to phonology, but also in second language acquisition of syntax, morphology, and pragmatics.

The above developments have made great contributions to the understanding of second language acquisition, providing explanations from different angles and perspectives. CAH and SLM are the main theoretical frameworks employed in the present research; and these two frameworks are discussed in detail below.

2.4.1 Contrastive Analysis Hypothesis (CAH)

The comparison of languages has existed ever since there were language contacts, but it was not until the 1940s that Whorf (1941:713) in his book *Languages and Logic* first proposed the concept of 'contrastive linguistics' to distinguish his approach from the 'comparative linguistics' which had prevailed in the West since the 19th century. This paved the way for the introduction of Lado's Contrastive Analysis Hypothesis. As mentioned above, Lado was the first to propose the concept of contrastive analysis and established its theoretical framework. Fry (1945) also helped lay the foundations for CAH. In his famous book *Linguistics Across Cultures*, Lado (1957:1) cited Fry's statement that: "the most effective materials are those that are based upon a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner". Inspired by Fry, Lado proposed his fundamental assumption:

> *the student who comes in contact with a foreign language will find some features of it quite easy and others extremely difficult. Those elements that are similar to his native language will be simple for him, and those elements that are different will be difficult. (1957:2)*

The Contrastive Analysis Hypothesis (Lado, 1957) aims to accurately predict the difficulties experienced by L2 learners by comparing their native language and the target language, where the differences between two languages are believed to be sources of difficulties for learners and they become obstacles leading to resistance to the successful acquisition of the L2 (James, 1980). By comparing the L2 with L1, such differences can be identified and difficulties predicted. According to the CAH (1957), similarities between the native and target languages are easy for learners, but differences are difficult. The CAH (1957) was applied comprehensively by Lado, including not only in linguistics domains such as pronunciation, grammatical structure, vocabulary, and writing systems, but also in terms of cultures. However, in comparing two sound systems, Lado (1957) compared the phonemic differences between two languages and systematically analysed the phonemic segmentation and pronunciation of sounds, and Kramsch (2007:243) has pointed out that Lado thought that an L2 learner does not actually hear an L2 phoneme, but his own phoneme.

Nevertheless, because CAH can help teachers to predict difficulties among students learning a second language, and because it has wide application in comparisons of languages, it has attracted extensive interest in the language teaching and research community. Correspondingly, a large number of relevant studies have been conducted by researchers.

However, in the late 1960s, the psychological and linguistic foundations of CAH began to be challenged by many researchers (Burt and Dulay, 1974; Ellis, 1986; Wardhaugh, 1970; Whitman and Jackson, 1972). One of the criticisms is that the CAH equates linguistic differences with psychological difficulties, which actually belong to two different domains. In addition, many empirical studies have shown that CAH cannot completely predict all errors students make when learning a target language, and furthermore not all errors made by students can be explained by L1 transfer. Whitman and Jackson (1972) tested English syntax among 2500 Japanese students to predict the L2 learners' levels of difficulty in learning English syntactic patterns. They found that contrastive analysis could not sufficiently predict the errors. Burt and Dulay (1974) analysed English errors made by Spanish learners and found that only 3% could be attributed to the interference of their native language, while 85% were developmental errors and the remaining 12% belonged to other categories. Ellis (1986) also pointed out that errors caused by native language interference generally account for only about 30% of all errors.

The CAH has been developed into different versions. The strong version is Lado's original claim (1957) that the interference of the mother tongue is the cause of all difficulties and errors in learning a foreign language, and the differences between two languages cause the difficulties. After many doubts and criticisms, Wardhaugh (1970) proposed a weaker version based specifically on the differences between the mother tongue and the target language. Rather than emphasising prediction, the weak version stresses the analysis and explanation of errors, finding their sources and thereby aiming to correct the errors. However, the weak form is still also limited to the question of interlingual barriers, completely ignoring intralingual barriers from within a language system. Wardhaugh (1970:123) also criticised both forms: "the strong version of the hypothesis is untenable and even the weak version creates difficulties for linguists." In response to criticism, a more moderate version of the hypothesis was proposed by Oller and Ziahosseiny (1970) which claims:

The categorization of abstract and concrete patterns (including time sequenced events) according to their perceived similarities and differences is the basis for learning;

therefore, wherever patterns are minimally distinct in form or meaning in one or more systems, confusion may result. Conversely, where patterns are functionally or perceptually equivalent in a system or systems correct generalization may occur. (Oller and Ziahosseiny, 1970:185-186)

This moderate form suggests that neglect of the most subtle distinctions between two languages, or from within either of the languages, in the form of sounds, sequences or meanings is more likely to lead to learning difficulties. It also emphasises that both interlingual and intralingual errors can now be accounted for.

Although CAH has been doubted and criticised and has fallen out of favour, its contribution to research in linguistics and language acquisition, as well as to foreign language teaching, is undeniable and indelible. Moreover, it has been revitalised and applied to a wider range of fields, such as translation (Toury, 1995, Venuti, 2000, cited by Kramsch, 2007:245), computer corpora (Ajmer, Alenberg and Johansson, 1996; Johansson and Oksefjell, 1998, ibid.), and L2 classroom teaching (Blyth, 1995; Belz, 2003, ibid.). Kramsch (2007:246) praised CA's role in the social and cultural studies of today's applied linguistics, saying that it is preeminent.

The present study tends to determine how Chinese speakers of two dialects/languages acquire English phonology, and contrastive analysis is employed since it is a very substantial and useful method for identifying differences between languages.

2.4.2 Speech Learning Model (SLM)

An influential model proposed by Flege (1995) is the Speech Learning Model (SLM). As mentioned previously, Flege's opinion differs from Lado's as to the effects of similarities and differences between the L1 and L2. Flege (1995) thinks the more similar L2 sounds are to L1 sounds, the more difficult it will be for L2 learners to perceive them, and the more difficulty they will experience in acquiring those L2 sounds. On the other hand, the more different L2 sounds are from L1 sounds, the more easily L2 learners can perceive and establish the contrast, and thus the more likely it is that they can acquire the L1 sounds. Put simply, Flege thinks that 'new' is easy but 'similar' is difficult, which is completely opposite to Lado's CAH. Flege (1987) supported his idea in an empirical study. He examined the French vowels /y/ and /u/ produced by native English speakers, and the results indicated that English speakers were more successful in producing the French vowel /y/, which was completely new to them, than the French vowel /u/ which is closer to the English /u/.

Flege agrees with Best that phonetic perception of the L1 and L2 is very important. Based on Best's (1993) PAM, Flege proposed the SLM. Flege (1995) believes that, if there is no accurate perception of L2 sounds, there will be inaccurate production. Therefore he claims that phonetic perception influences the establishment and acquisition of the listener's L2 sounds. He further mentions that many L2 speech production errors can be attributed to the mistaken perception of the representations of properties which are specific to L2 sounds (Wester, Gilbers and Lowie, 2007:479). He thinks that is the reason why English native speakers have greater difficulty in acquiring the French /u/ than the French /y/, because they cannot perceive the difference between French /u/ and English /u/ and thus cannot produce the different sounds accurately. However, the SLM (Flege, 1995:238) does not claim that "all L2 production errors are perceptually motivated."

Unlike Lado's CAH, Flege's SLM (1997:17) posits that "interlingual identification occurs at the phonetic rather than the phonemic level". The main reason that he thinks this is because "L1 phonology filters out features (or properties) of L2 sounds that are important phonetically but not phonologically, or both" (Flege, 1995:238). Moreover, Flege and his colleagues (2000:2713) think the acquisition of phonetic categories can make L2 sound perception more native-like because it enables learners to perceive L2 speech input without interference from previous learning. This is also supported by

Trubetzkoy's assumption (1932/1969) that acoustic differences which are phonemically irrelevant in the L1 could be filtered out by L2 learners' L1 phonology.

Flege (1997:16) mentions that it is useful to classify an L2 sound as 'identical', 'similar' and 'new' to represent the relationship between L1 and L2 sounds. He uses a classification of equivalence to categorise target sounds and three classification criteria are used:

1) IPA symbols;
2) acoustic measurements;
3) listeners' perceptual judgements of sounds in L1 and L2 (Flege, 1997:17).

Therefore, the 'identical' category is classified in SLM based on firstly, L2 sounds which are represented by the same IPA symbols in L1 and, secondly, that no significant acoustic difference is found between the L2 sound and its counterpart in the L1, and thirdly that there are no perceptual differences between the L2 and L1 sounds detected by listeners. Flege mentions that there is no accepted metric to measure the phonetic distance between sounds in the two languages. Flege's 'similar' category is classified according to rules such as that an L2 sound uses the same IPA symbol as an L1 sound, but acoustic measurements show significant differences, and audible differences between the sounds in L2 and L1 can also be detected. The criterion for the category of 'new' (ibid.) is that the IPA symbol of an L2 sound is different from that of any sound in the L1, and moreover the L2 sound should differ acoustically and perceptually from the L1 sounds which are closest to it. For the classification of new vowels, one more rule is required (Bohn and Flege, 1992, cited by Flege, 1997:18), namely that it is new "only if most of its realizations occupy a portion of the acoustic phonetic vowel space that is unoccupied by the realizations of any L2 vowel. This implies that few of the vowels in an L2 will be new for learners whose L1 has a large vowel inventory."

However, Wester et al. (2007:479) pointed out that this classification is problematic because it is hard to distinguish which sounds should be 'similar' and which should be 'new', such as with the dental fricatives produced by Dutch learners of English. These learners use acoustically similar L1 sounds to substitute for these L2 sounds, which satisfies the criterion of newness that IPA symbols are different, but it violates the criterion that L2 sounds should differ acoustically from L1 sounds. There is the same problem if two dental fricatives are classified as 'similar' because it cannot meet the requirement that their IPAs should be the same.

Flege (1995:264) thinks that two important variables which play a key role in influencing development according to SLM are Age of Learning (AOL) and perceived cross-language phonetic distance. He believes that "the greater the perceived distance of an L2 sound from the closest L1 sound, the more likely it is that a separate category will be established for the L2 sound" (Flege, 1995:264), and thus, it will be easier for L2 learners to eventually acquire this L2 sound. In addition, he also asserts that "the earlier L2 learning commences, the smaller the perceived phonetic distance needed to trigger the process of category formation" (ibid.). Flege thinks that there are no objective methods to assess the degree of perceived phonetic distance. Although Best (1993, cited by Flege, 1995:264) employs in his PAM the "spatial proximity of constriction locations and active articulators, and similarities in constriction degree and gestural phasing", Flege thinks that this is too difficult to apply. However, both hold the view that cross-language mapping experiments can be used to empirically gauge perceived phonetic distance (Guion, Flege, Akahane-Yamada, and Pruitt, 2000:2713).

2.4.3 Application of CA and SLM in this research

My study explores how speakers of two different Chinese dialects/languages acquire English phonology. Contrastive analysis is a powerful tool to identify the differences between two L1 Chinese varieties and between the two and L2 English. Although the CAH cannot predict and explain all errors, as mentioned in the above section, contrastive analysis as an analytical method for language research is feasible and effective. As Fan and He (2012:113) mention, IL rejects the CAH, but does not deny that CAH is a useful research method, and they point out that Selinker also used comparative analysis to analyse the sounds and grammar in the native and target languages and interlanguage.

Therefore, the main method used in this study is contrastive analysis. To make up for the deficiencies of contrastive analysis, this study also attempts to combine contrastive analysis with Flege's SLM model in order to make predictions of the difficulties L2 English learners may encounter.

As mentioned above, CA contrasts the sound inventories of the L1 and L2 phonemically, while SLM compares the language sounds at the phonetic level. The two theoretical frameworks operate at different levels. To truly succeed in acquiring an L2 phonology, the learner must not only successfully acquire the phoneme, but also its phonetic realisations or allophones. Therefore, this study not only compares the phonemic systems of Cantonese and Harbinese t with English, but also the phonetic variants of phonemes in both with English. Moreover, as Flege (1995:238) thought, L1 phonology may filter out some important phonetic features of L2 sounds; therefore, it seems sensible to also contrast the phonetic realisations of phonemes. Although the two methods seem to be incompatible, the comparison of inventories phonemically and allophonically (phonetically) makes it possible to combine the two theoretical frameworks. This also creates an opportunity to see from a new perspective how learners of L2 English with different yet very similar L1s acquire English phonology, and it also provides a chance to test which framework is more applicable in the current research.

In addition, it is also worth noting that Flege's SLM (1995) seems not to be as widely applied as Lado's CAH, mentioned in above sections 2.4.1 and 2.4.2. Flege's SLM is only used to study L1-L2 segments in isolation or segments in different positions of syllables but nothing beyond. However, apart from segments, this study will also apply Flege's SLM to superasegmentals to see if it works in, for example, the comparison of L1-L2 syllable structure and stress.

Moreover, the SLM seems to be mostly directed at highly experienced learners or bilinguals who are not beginners (Flege, 1995:238). Although Flege never says so, in looking at subsequent work , this seems to be the case. Also, his speakers are immigrants, not L2 classroom learners. Thus, I wish to see if SLM can be extended to less advanced learners (i.e. beginner and intermediate learners) especially in the context of L2 classroom learning.

2.4.4 Exposure length in second language acquisition

In the process of L2 acquisition, a variety of factors affect the L2 learners' performance. In addition to non-linguistic factors such as personality, motivation, aptitude and gender play, variables which are considered to have greater influence on the degree of L2 foreign accent as shown in production and perception are age, input, exposure length, language use and training (Piske, 2007:303).

Age of learning (AoL, initial exposure to the L2) is a variable most frequently examined in learners' performance (Piske, MacKay and Flege, 2001; Bialystok, 1997; Dulay, Burt and Krashen, 1982; Scovel, 2000; Singleton & Lengyel, 1995; Mayo, 2003). Researchers have found that age of learning has a significant impact on L2 learners' performance and supports the view that the earlier a learner starts to learn an L2, the better his or her pronunciation could be (Seliger, Krashen and Ladefoged,1975; Oyama, 1976; Flege, 1988; Thompson, 1991; Flege, Munro and

MacKay, 1995; Moyer, 1999).

Many researchers (Flege et al., 1995; Guion, Flege and Loftin, 2000; Piske et al., 2001; Piske, Flege, MacKay and Meador, 2002) have also shown that language use is an important variable. In Flege et al. (1995) on the production of short English sentences by two groups of early Italian immigrants with similar ages of learning but different L1 use they found that the group with frequent L1 use had stronger foreign accents in English production than the group with little use.

Finally, among the variables affecting L2 learners' production, and relevant to the present thesis, exposure length is the second most frequently examined factor (Piske, MacKay and Flege, 2001). It is also regarded as a significant predictor in L2 performance by e.g. Hammer, Komaroff, Rodriguez, Lopez, Scarpino and Goldstein (2002). Exposure length or amount of L2 experience is also called 'length of residence' (LoR) in L2 studies on immigrants, which refers to the amount of time L2 learners are exposed to the L2 as dominant native language environment. As to the importance of exposure length on L2 learners' performance, Piske et al. (2001) mentioned that previous studies had different results. Some researchers found what is expected: that the longer exposure L2 learners have, the more native-like L2 performance they have e.g. Flege and his colleagues (1992, 1995, 1999) as well as Burstall (1975) and Mayo (2003).

However, other studies (Flege, 1988; Piske et al., 2001; Thompson, 1991; Moyer, 1999) did not show a positive effect of exposure length on L2 pronunciation. Flege (1988) researched foreign accents from two groups of Taiwanese adult learners of L2 English with length of residence in the USA of 1.1 years and 5.1 years and found that the results of accent ratings did not show a significant difference. Flege and Fletcher (1992) made research on two groups of late Spanish-English bilinguals with 0.7 and 14.3 years of LoR in production of English sentences and found that LoR is less important than AoL as a predictor to examine L2 prounciation. In addition, Flege *et al*. (1995), Riney and Flege (1998) and Meador, Flege and MacKay (2000) concluded from their research that "the size of LoR effects depends on whether learners are still in an early phase of L2 learning or not" (Piske, 2007:198). They found that highly experienced L2 learners, additional amount of exposure to L2 is unlikely to decrease the degree of foreign accent.

Whether there are effects of length exposure on L2 learners' performance in an immigrant setting is not clear. In addition, in looking at the literature, we can find that most studies were carried out to investigate the L2 performance of immigrants but few studies were done to check the effects of length exposure in L2 classrooms. The situations of L2 immigrants are completely different from the L2 learners in foreign language settings (Cook, 1999; Piske, 2007). In the former the learner is not only more frequently able to use L2, but also he or she is exposed over a longer period of time and also to native speakers than learning in the classroom.

Let's now look at L2 exposure in the classroom. Mayo (2003:106) conducted research on Basque-Spanish bilingual learners of different age groups from EFL settings to deal with the issues of grammaticality judgements and her results confirmed her hypothesis 'the longer the exposure to the language, the better performance becomes'. Also, he (ibid.) found 'an earlier start does not produce significantly better results in a situation of instructed foreign language acquisition'. Mayo (1999) compared younger and older learners and found the latter performed better in the environment of L2 as a foreign language and found that the amount of exposure of the younger group did not seem to be sufficient to achieve the level of the older group in the foreign language settings. Ouyang (2018) found that learners at a higher level in school had better results than those at lower levels in both perception and production of dental fricatives by Chinese learners of L2 English in the classroom.

2.5 Previous studies of English phonology acquisition

2.5.1 Influence of language and dialect on English phonology

Many researchers have studied the influence of the L1 on the L2 or transfer from the L1 to L2, especially in phonology (Lado, 1957; see also Anderson, 1987; Corder, 1967; Ellis, 1994; Fisiak, 1978, 1991; Gass, 1979; Odlin, 1989; Rasier and Hiligsmann, 2007; Weinreich, 1953; Young-Scholten, 1985). Mukattash (1981) said that interference from the L1 seems to play a significant role at the phonological and the lexical levels especially if the source and target languages are not related. Research into the factors affecting English reading performance in children with Chinese as a first language provides a growing body of evidence on the cross-language transfer of phonological processing in the L2 learning of EFL learners (Gottardo, Yan, Siegel and Wade-Woolley, 2001). In addition, Cook (2004) claimed that L2 learners carry over certain features of their first language to the second, especially in pronunciation. Major (2013) discussed the reasons why a foreign accent is so difficult to overcome, and concluded that the effect of a foreign accent is governed by many interrelated factors, among which one of the most important is L1 transfer. As mentioned by Major (2013), the Portuguese /v/ is pronounced with more friction than the English /v/, so an English /v/ made by a Portuguese sounds like a Spanish accent to Brazilians.

From the above, it can be seen that the L1 does influence L2 learning, especially in phonology. In addition to phonological studies of effects of the L1 on L2, the impact has been explored of specific languages or regional dialects on English phonological learning. For example, Wiltshire and Harnsberger (2006) researched the influence of two local Indian dialects on Indian English and revealed a clear transfer effect with the back vowels of Gujarati English, Tamil English rhotics and certain rising and falling pitch accents in Gujarati English. Hung (2000) claimed that Hong Kong English was influenced by the much simpler vowel system in Cantonese, and Liang (2010) demonstrated that the characteristics of Singapore English were derived from the influence of the Chinese Min and Yue dialects. The main phonological influence was mainly reflected in vowels, such as the disappearance of the lengthening of English vowels and the substitution of English diphthongs by Singlish monophthongs.

2.5.2 Influence of Mandarin Chinese on English phonology

In recent decades, research into the influence of the L1 on L2 phonology has also been widely conducted in China. A large number of researchers have focused on the influence of the Chinese L1 on the learning of other languages, and research regarding the influence of Chinese or Mandarin on English phonology is flourishing (Gong and Tian, 2008; Kong and Wang, 2001; Li, 2006; Peng, 2009; Shi, 1996; Wang, 2007; Yan, 2007; Yu, 2009).

Shi (1996) researched the influence of Mandarin on the English production of Chinese English beginners in terms of segments, syllable structure and stress, and even sound liaison. She found that Mandarin speakers have problems with some English segments such as the English dental fricatives /θ/ and /ð/ which are often substituted by [s] and [d], the English /æ/ and /ɔ/ by [ai] and [au], the English /v/ by [w] and confusion between /v/ and /w/. For English syllable structure and consonant clusters, she found that Mandarin speakers often insert a vowel between consonants and after a coda in English closed syllables due to their lack of consonant clusters and the preference for open syllables, because only /n/ and /ŋ/ can be codas in Chinese. Due to the disallowed linking in Chinese of a coda with a subsequent word starting with a vowel, she found that Chinese speakers cannot break a word boundary to resyllabify English words for a sound liaison.

Kong and Wang (2001) mentioned that /ɪ/, /ʊ/, /ʌ/, /æ/ were the most difficult vowels for them to teach because they are lacking in Chinese, and their students often used [i], [u], [a] and [ɛ] respectively to replace them. They found that transfer to their students' L2 English was from not only Mandarin but also their dialects. For example, a confusion between /n/ and /ŋ/ as codas exists in their dialect and both sounds were produced as /ŋ/; thus the English word /an/ will be produced as [ɑŋ] instead. For consonants, they found that the English dark /l/ is substituted by [ɹ] or [ə], /ʃ / and /ʒ/ by the Mandarin [ʂ] and [ɻ], /w/ by [v]. They pointed out that the syllable structure /VVC_n/ was much more difficult for students; for example, /aʊn/ was often replaced by [ɑŋ] because Chinese does not allow a syllable structure with a diphthong plus a nasal but only allows /VC_n/, such as in /an/ and /aŋ/. Gong and Tian (2008) described similar problems with English vowels and syllable structure.

Li (2006) compared Chinese with English in terms of segmental phonemes and suprasegments. She found that some English vowels like /ɪ/, /æ/, /ʌ/, and /ɒ/ cannot correspond to Chinese vowels, and that interference from Chinese in English therefore occurred; for example, resulting in confusion between /i/ and /ɪ/. Non-correspondence also exists for English consonants like /θ/, /ð/, /v/, /ʒ/, /tɹ/ and /dɹ/, causing many mispronunciations. Due to differences in syllable structure between the two languages, Chinese students tended to insert a vowel into English consonant clusters or omit certain consonant(s). Chinese speech is pronounced syllable by syllable because of the correspondence of each character to one syllable, triggering effects on the pronunciation of connected English speech without liaison, assimilation, or omission.

Wang (2007) studied negative transfer from Chinese to English. She found that students exhibited confusion between consonants such as /n/ vs. /l/, /n/ vs. /ŋ/, /h/ vs. /f/ because of the influence of Chinese. She also found that Chinese orthographic Pinyin letters influenced learners' English pronunciation. For example, the Pinyin letter *i* leads to the confusion between /i/ vs. /ɪ/, and for letter *a* between /ɑ/ vs. /a/, and *o* between /o/ vs. /ɔ/, as well as the consonants in Pinyin *ch*, *sh*, *h*, *r* which have the same spellings as the English sounds /tʃ/, /ʃ/, /h/, /ɹ/.

Peng (2009) investigated the effects of Chinese on English from the perspective of positive transfer. Many similar pronunciations between consonants and vowels can help Chinese English learners acquire English phonology. For example, the pronunciations of the Chinese /i/, /a/, /u/ are similar to the English /i/, /ɑ/, /u/, the Chinese diphthongs /ai/, /ei/, /ɑu/, /ou/ are similar to the English /ai/, /ei/, /aʊ/, /oʊ/, and the Chinese consonants /p/, /m/, /f/ are similar to /p/, /m/, /f/.

2.5.3 Influence of Chinese dialects/languages on English phonology

Many researchers have examined the influence of Chinese dialects/languages on English phonology. Cantonese interference in English phonology has been investigated in various studies (Chen, 2013; Jia, 2011; Li, 2008; Li, 2009; Li and Chen, 2007; Liu and Wang, 2007; Yang, 1997; Zhang, 2017). Liu and Wang (2007) compared Cantonese and English phonemes, syllables, stress, tones and intonations and found many influences of Cantonese on English. The effects of transfer on English consonants can be seen from the confusion between /n/ and /l/ in Cantonese which causes the same problems with the English /n/ and /l/ (Liu and Wang, 2007; Chen, 2013; Liao, 2014). Liao (2014) found that /n/ substituting for /l/ is more common than /l/ substituting for /n/, and /l/ is often nasalised but /n/ is sometimes omitted.

In addition, the English consonant /ɹ/ is substituted by [l] , /v/ by [f] or [w], /θ/-/ð/ by [s]-[z], and /ʃ/ by [s] (Liu and Wang, 2007). Unlike Liu and Wang's findings for /θ/-/ð/, Yang (1997), Liu and Guo (2013) and Zhang (2017) found that Cantonese speakers often used [s] or [f] to substitute for /θ/, [d] for /ð/ ([d] and [z] for /ð/). Interestingly, Deterding, Wong and Kirkpatrick (2008) found that Hong Kong Cantonese speakers often substituted [f] for /θ/. Chen (2013) found that negative transfer of Cantonese occurred with the English consonants /ɹ/ and /l/, which she thought was the most typical example.

Transfer affecting English vowels is comparatively less significant (Liu and Wang, 2007). It is likely that Cantonese has more vowels similar to those in English. However, difficulties may be caused by the lack of the phonemic contrast in length and tenseness in Cantonese, although Cantonese has a pair of vowels with length contrast. Thus English long (tense) vowels are often pronounced as their counterpart short or lax ones. Yang (1997) agreed and claimed that Cantonese speakers have less difficulty in learning English vowels than northern Mandarin speakers. Zhang (2017) found that /æ/ and /ɛ/ are confused by Cantonese speakers, and 54% of subjects pronounced /ɛ/ as [æ]. Although /æ/ does not exist in Cantonese, it is pronounced better.

Influences on English syllable structure have been identified (Liu and Wang, 2007). The absence of consonant clusters in Cantonese leads to the insertion of vowels into English clusters, and unreleased voiceless plosive codas trigger their omission in English. Jia (2011) compared the plosive codas /p/, /t/, /k/ ([p̚], [t̚], [k̚]) in Cantonese with the English voiceless plosive codas /p/, /t/, /k/ and found that the latter in Cantonese codas without the release of plosion resulted in the omission of the corresponding codas in English. Yang (1997) mentioned that unreleased codas in Cantonese even influence the release of the English /b/, /d/, /g/ as codas. In addition, a special phenomenon is where Cantonese speakers often insert a voiced stop after words ending with vowels or with a lateral, such as then [smɔ:l] 'small' and [flaʊə] 'flower' are pronounced as [smɔ:ld] and [flaʊəd] (Liu and Guo, 2013; Yang, 1997). This phenomenon was described as 'overgeneralisation'. Yang (1997) thought that some differences between Cantonese and English were similar to points between Mandarin and English, and thus she thought that improvements in Mandarin pronunciation would have a positive influence on English pronunciation. Yu (2009) agreed, comparing the similarities between Mandarin and English in her study entitled *The Influence of Speaking Good Mandarin on Learning English Pronunciation*.

Much research on the influence of the Northeastern dialect on the learning of English phonology has also been conducted (Chen, 2010; Yan and Yang, 2011; Yang, 2014; Yin and Li, 2014; Zhang, 2012). Northeastern dialect is a variety in the Mandarin group and therefore its influence on English pronunciation is very similar to the influence of

Mandarin on segments and syllable structure. To avoid repetition, this section describes research findings for Northeastern Chinese which are not the same as those for the influence of Mandarin. Apart from the substitution of [w] for the English /v/ in research into Mandarin influence on English phonology, the English /w/ is also frequently pronounced as [v] by Northeastern dialect speakers (Yin and Li, 2014) because Northeasterners are not used to pronouncing segments with rounded lips. In fact, Gong and Tian (2008) and Chen (2010) found a mutual substitution of /w/ and /v/ in learners of English speaking the Northeast dialect. In addition, some Northeastern dialect speakers have problems in distinguishing the apical fricative and affricative sounds /s/, /ts^{h}/, /ts/ from the postalveolar fricative and affricative sounds /ʂ/, /ʈʂ^{h}/, /ʈʂ/ in the Northeastern dialect and thus transfer is seen in their English production between /s/, /ts/, /dz/ and /ʃ/, /tʃ/, /dʒ/ (Chen, 2010; Yan and Yang, 2011). It should also be noted that the English dental fricatives /θ/ and /ð/ are often substituted by [s] and [ts] (Chen, 2010; Zhang, 2012).

Yan and Yang (2011) pointed out that because of the lack of tenseness and length contrasts in the Northeastern dialect, the tense/long English vowels are often pronounced as their counterpart lax/short ones; for example, /i:/ and /u:/ are pronounced as [ɪ] and [ʊ]. Yan and Yang also mentioned that students in the Northeast confuse and often substitute /aɪ/ for /æ/. Chen (2010) found that, although Mandarin/Northeastern dialect diphthongs /ai/, /ei/, /ou/, /au/ are similar to those in English, the duration of the vowels' pronunciation is different. Diphthongs in Mandarin/Northeastern dialect are pronounced as gliding from the main vowel to the end vowel rapidly without an obvious transition, while English diphthongs are pronounced with clear transitions in gliding from one to the other. Wang and Liu (2017) conducted an acoustic experimental study of the influence of the Northeastern dialect on the English monophthongs /ɑ/, /ɜ/, /i/ and /u/ and found that the English /ɑ/ and /ɜ/ are pronounced as lower and more advanced than by native English speakers but /i/ and /u/ are pronounced lower and more retracted.

Zhang (2012) investigated the English pronunciation of students from three provinces of the Northeast. He found that Northeastern speakers make very common English pronunciation errors but also errors specific to the areas in the Northeast. For example, the English /s/ and /z/ are confused with /ʃ/ and /ʒ/ by students in Liaoning province and Jixi and Hulin cities but not in Heilongjiang and Jilin provinces. Common errors can be seen in the consonants /θ/, /ð/ substituted by [s] and [ts], /w/ by [v], and in the vowels /e/, /æ/, /ʌ/ substituted by [ei], [ai] and [a].

As in the difficulties experienced by Mandarin speakers, a vowel like [ə] or [u] is often inserted after the coda in English closed syllables and among consonant clusters due to the influence of the Northeastern dialect (Chen, 2010). Like Mandarin and Cantonese, Gong and Tian (2008) also found that the English syllable pattern /VVC_n/ is difficult for Northeastern dialect speakers. This English structure would be difficult for all Chinese native speakers because it does not exist in the Chinese language.

Due to the absence of word stress in Mandarin and the Northeastern dialect, Chen (2010) found that English stress is often mistakenly placed by students in the Northeast. Zhang (2012) researched the Northeastern dialectal influence on the intonation of English phonology and found that students from Liaoning read both statement and question sentences with rising tones, affected by the negative transfer from their local Northeastern dialect.

Apart from the above dialects/languages, the influence of many other Chinese dialects/languages on English has been researched , such as Hakka (Chen, 2014), Henan Dialect (Niu Jie, 2011; Ping, 2001), Hubei (Chen, Teng and Song, 2010) and Anhui (Zhang, 2013).

2.6 Summary

This chapter started by describing the differences between the two varieties of Chinese and between these varieties and English. It introduced sound systems of Harbinese Mandarin, Guangzhou Cantonese and English (RP and GA) in detail to compare and contrast these two varieties of Chinese with English with respect to consonants, vowels and syllable structure. Then, in presenting ideas on the development of a second language phonology, starting with the CAH (Lado, 1957) and moving on to the SLM (Flege, 1995), I turned to the aim of the present thesis, of looking at L1-L2 similarities vs. differences on the basis of two very similar dialects.

Browsing the literature, it was found that little research on these two Chinese varieties (Harbinese and Cantonese) and their influence on the acquisition of L2 English phonology exists, and then even less attention has been paid to the Northeastern dialect (Harbinese).

Through the comparison of Harbinese and Cantonese with English, difficulties specific to each Chinese variety are predicted, namely that Harbinese speakers will have difficulties in acquiring English consonants /w, ʃ, tʃ, dʒ, h, ʒ, θ, ð, ɹ/, vowels /ɪ, ʊ, ɔ, æ, ʌ, ɒ, aʊ, ɔɪ/, and syllable structure /CVC_N, CVC_P/, while Cantonese speakers will have difficulties in consonants /v, n, l, ʒ, θ, ð, ɹ/, vowels /ə, æ, ɒ, ɑ/ and syllable structure /CVC_P, CjV(C), CwV(C)/. Apart from these, it is predicted that Cantonese will acquire the English tense-lax vowels distinction but Harbinese will not. Both Harbinese and Cantonese will have difficulties in acquiring consonant clusters and syllable structure /$CVVC_N$/. Finally, Harbinese speakers will apply iambic stress whereas Cantonese speakers will not.

Chapter 3. Methodology

In this chapter, I first classify the selected English targets into three categories and then present hypotheses and research questions based on the criteria of difficulty. Afterwards, the methods used for data collection and data analysis are introduced, including how participants were recruited, and how the hypotheses were tested and the research questions answered.

3.1 Classification of categories

By comparing the two Chinese dialects/languages with English using the contrastive analysis method (Lado, 1957), differences in segments, syllable structures and stress between the two dialects/languages which might lead to difficulties in L2 English have been explored thoroughly in the literature review chapter. Those that might cause difficulties in L2 are the essential targets in my research. It is worth mentioning that when the segments and syllable structures in the two dialects/languages and English are the same this may largely produce the same production of English, which is not of interest to the research. Therefore, only differences between the two Chinese dialects/languages and English are targeted.

As noted by Flege (1997:16), it is useful to differentiate whether L2 sounds are 'identical', 'similar' or 'new' to those in the L1 and to check the degree of difficulty L2 learners have according to those categories. Therefore, all English targets are classified into the three categories. However, how they are classified is extremely important, and there should be clear criteria. In this study two theoretical frameworks are applied, including Lado's CAH (for the selection of research targets) and Flege's SLM. As mentioned in chapter 2, CA is very effective and can be used to contrast many elements in linguistic domains including sounds, syllable structures, syntax, vocabulary, and even suprasegmental factors such as stress. Thus, it is a very good tool to compare different L1s with an L2. However, its predictive ability has been criticised as mentioned in section 2.3.2. Therefore, to compensate for the deficiencies of the CAH, SLM is used in this study. As introduced in the literature, the CAH (Lado, 1957) contrasts L2 and L1 phonemically and proposes that difference is difficult to learn but similarity is easy, while the SLM (Flege, 1995) is mainly used to study only sounds comparing L2 with L1 sounds at the phonetic level, predicting that similarity is difficult to learn and difference is easy. It can be clearly seen that these two frameworks make completely opposing predictions. However, I have combined them in this research and hope thereby to answer a series of questions, such as which idea is more suitable for this kind of research? What is the relationship between these two theoretical frameworks and why do they lead to different predictions?

For the purpose of classifying the targets, as mentioned, CAH (Lado, 1957) compares the L2 and L1 phonemically, identifying the phoneme inventories of the L1 and L2. If the L2 has a phoneme which is the same in the L1, it is considered to be 'identical'. If an L2 phoneme is similar, but not identical, to a sound in the L1 it is regarded as 'similar'. If an L2 phoneme is absent in the L1 that means it is 'new'. SLM (Flege, 1997) uses relatively detailed classification criteria based on the IPA symbols, acoustic measurements, and native speakers' perceptions, (see section 2.3.3). However, Wester et al. (2007:479) have pointed out that this is also problematic (also see section 2.3.3). According to Flege's criteria, segments deemed 'identical' in CAH may be placed in the category of 'similar' in SLM and the 'similar' could be 'new' because if a 'phoneme' is identical it does not mean that allophones of the phoneme must be the same. The problem for both frameworks is that CAH may have many identical pairs but SLM may have many new pairs. To avoid these classification problems this study requires its own classification criteria. SLM is limited in application to sounds, whereas this research also focuses on syllable structure and stress. Therefore, I mainly use the CAH method but SLM is also taken into account for the classification of sounds. That is to say, if an L2 phoneme is completely identical to an L1 phoneme, including having the same allophones, it is classified in the category of 'identical'.

Additionally, if a phoneme in the L2 is absent from the L1, it is put into the 'new' group. In this way, the category of 'similar' in my research should represent a large number of items.

The criteria for classification are as follows:

1. Mainly based on the comparison of phonemes and allophones in the L2 and L1.

The comparison between L1 and L2 is mainly based on phonemes. If L1 and L2 share a phoneme with the same IPA symbol, they are considered to share that phoneme. If the phoneme in L1 and L2 has also the same allophones, this will be regarded as completely identical in the L1 and L2 and categorised as 'Identical'. It is true that there are also nuance between two languages which share the same phonemic and phonetic symbols. For example, /d/ [d] in English and French are not identical in the feature of place. English /d/ [d] is pronounced alveolar but in French dental. These more nuanced features between two languages is not considered to classify segments in the present thesis.

If an L1 segment is phonemically identical to the L2 one but not allophonically (and vice versa), it is not considered identical but similar. If an L1 segment does not exist in the L2, it is considered a new sound and classified as such. In addition, if an L1 sound is phonemically different in the L2, it is classified as 'similar' not 'new'. It has to be mentioned why an L2 segment non-existent in the L1 is categorised 'new' but not 'similar'. This is because an L2 learner may use a similar L1 segment to represent the L2 segment but which does not exist in L1. One might also ask why two phonemically different segments in two languages are categorised as 'similar' but not 'new'. The issues of whether these segments are new or similar cannot be clarified clearly by only criteria 1. To dispel doubts, other criteria need to be set up. Criteria 2 and 3 illustrated below will solve such issues and clarify the classification in detail.

2. Taking previous research into account, which category a segment belongs to also takes account of the perception of L1 native speakers as described in the literature concerning the L2 segment.

Whether an L2 segment is new or similar cannot be determined only by criteria 1, the actual perception and recognition of L2 learners to L2 sounds also have to be taken into account and, so does the possible impact of the orthography on L2 pronunciation (the impact of orthography will be discussed in criteria 3 below).

English dental fricatives are absent in many languages. Thus, they are new to languages such as Chinese. However, some people may doubt that they are substituted by the closest L1 sounds, and thus, they could be assumed to be similar as mentioned above. However, according to Eckman's Markedness Differential Hypothesis (1977), those sounds are the most marked, and so they should be the most difficult to learn. Thus they should be considered as new because of their absence in the L1. Although substituted by the closest sounds in the L1, L1 speakers still consider them as 'new' in their language. Chen (2010) and Zhang (2012) confirmed that Mandarin speakers often substitute /θ/ and /ð/ with the closest Mandarin sounds [s] and [z], but they both mentioned that the new sounds are difficult for L2 English learners. We can see here that, although they used the closest sounds to replace those two sounds, they still referred to /θ/ and /ð/ as new sounds. In this case, we should not classify them as 'similar'. In addition, the spellings of English /θ/ and /ð/ and the closest Mandarin sounds [s] and [z] are 'th' for /θ/ and /ð/ and 's' and 'z' for [s] and [z]. They do not share the same orthography. Therefore, /θ/ and /ð/ are categorised into 'new' but not 'similar'.

3. If two languages share the same orthographic representation, and the sounds are perceived similarly, they should be deemed 'similar' but not 'new'. An example is /x/ and /h/ sharing the same letter *h* as mentioned in section 2.5.2 of chapter 2.

Another example is that the Mandarin /x/ is different from the English /h/ in place of articulation. /x/ and /h/ should have been categorised into the group ‘new’; however, they are considered to be phonemically different but similar in this study because they share the same orthographic representation of letter *h* in Mandarin and English. Mandarin speakers consider Mandarin /x/ to be ‘similar’ with English /h/ as mentioned in relation to positive transfer from the L1 by Wang (2007). In this case, they are classified as ‘similar’ instead of ‘new’. Therefore, although the segments in the L1 and the L2, English, are phonemically different, they share the same orthographic representation and the L2 sound is perceived as a similar sound to the L1 by L2 English learners, and they would be put into the ‘similar’ category.

Classification of categories

1. Identical: completely identical phonemes and allophones.
2. Similar:
    a. Allophonically identical but phonemically different.
    b. Phonemically identical but allophonically different.
    c. Phonemically different but similar.
3. New: Phonemically different.

Note that if a phoneme has only one allophone, allophones do not need to be compared. In the same example of the Mandarin /x/ and English /h/, each phoneme in the respective language only has one allophone, and thus to compare allophones is the same as comparing the phonemes.

In the present research, all research targets for segments, syllable structure and stress were classified into three categories to test Flege’s ideas concerning ‘identical, similar and new’ and hypotheses were proposed based on these three categories for the predictions of difficulty. For the first two categories, this term applies in one dialect/language but not the other and a hierarchy is followed in which the category ‘identical’ has priority over ‘similar’ and ‘similar’ over ‘new’. This means that if the sound in the L2 is identical with one in the L1 dialect/language, irrespective of the relationship of the L2 sound to the other L1 dialect/language (‘similar’ or ‘new’), predictions would be made based on the priority one. The detailed rules for the categories are described below.

Category 1: 'Identical' means that segments, syllable structures and stress in the L1 and L2 are completely identical, and the results of the comparison are that it is 'identical in dialect/language x but *not* dialect/language y', but the '*not*' can be either 'similar' or 'new'. That is to say, an English segment is identical in Harbinese, and not in Cantonese, but it can be either similar or new in the dialect/language, and vice versa.

Category 2: 'Similar' refers to segments, syllable structures and stress in L1 and L2 pairs. The main criterion for this category is when these are 'similar in L1 and L2 in dialect/language x but *not* dialect/language y'. If they are not similar this means they are *new* rather than identical, since that situation falls under category 1. Specifically, if an English segment is similar to one in Harbinese, then it is not identical in Cantonese, but rather it is new or another type of 'similar'.

Category 3: The term 'new' means that it does not exist in either dialect/language with respect to the L2 (see Chapter 2 for a discussion of Flege's and others' definition of similar vs. new). (This category is included because there is evidence of dialect/language-specific differences in error patterns for learners of English. The sources of these differences are explored in Chapter 5.)

Placement of segments, syllable structure and stress in categories 1, 2 and 3.

Before I carry out the classification, I have to mention that this does not depend on their categorization as dialects rather than closely-related languages. However, in the learners' and teachers' minds, Cantonese and Harbinese-Mandarin are dialects and whatever feelings of being speakers of a shared language these learners might have influences their approach to learning (and for teachers, teaching) English. What the specific influence is, however, is beyond the scope of the present thesis. The targets are placed into the three categories of 'identical', 'similar' and 'new' based on the principles and criteria described above, as shown in the examples below.

3.1.1 Classification of Consonants

All 12 target English consonants /v, w, ʃ, ʒ, tʃ, dʒ, h, ɹ, θ, ð, n, l/ were placed into the three categories.

'Identical' consonants

The English /h/ is a glottal fricative which is completely identical in Cantonese, but Harbinese only has the similar /x/. According to criterion 1, one segment is completely identical in one dialect/language but not identical or could be similar in the other. Therefore, /h/ was placed in the category of 'identical'.

The English phoneme /w/ exists in both Harbinese and Cantonese. However, in the literature it is mentioned that Harbinese speakers often pronounce /w/ as [v]. Actually, the Harbinese /w/ has two allophones [w] and [v]. However, Cantonese has only one allophone which is the same as in English. Thus, the English phoneme /w/ is phonemically identical but allophonically different in Harbinese, so it is identical in Cantonese but only similar in Harbinese. According to the rules of category placement, /w/ is categorised as 'identical'.

'Similar' consonants

This category is complicated because various segments are assigned into this group, but the types of similarity are different. There are three types of similarity as shown earlier in section 3.1.

The English /v/ is new in Cantonese but it is an allophone in Harbinese. Thus, the English /v/ is allophonically identical but phonemically different in Harbinese. Based on the rules of category placement, the English /v/ is classified into the category 'similar'. It has to be emphasised here why English /v/ is new in Cantonese but not similar. The main reason is because /v/ is an english phoneme, but there is no such phoneme in Cantonese. Thus, although Cantonese may pronounce it as [w], but it is a new phoneme to them. To produce the English phoneme, the prediction is that they will recruit a similar sound from their own language. Therefore, English /v/ is new to Cantonese speakers.

The English /n/ and /l/ [l] (/l/ [ɫ] is absent in both Chinese dialects/languages thus it will not be classified here) exist in both dialects/languages, however, it is mentioned in the literature that Cantonese speakers neutralise /n/ and /l/ in their dialect/language, which reveals that the sounds are not totally identical with the English counterparts. Therefore, /n/ and /l/ are also classified in the category of 'similar'.

The English /ɹ/ is similar to the Harbinese retroflex approximant /ɻ/, but it is absent in Cantonese. In addition, English and Harbinese share the same orthography r. Therefore, it was assigned to the category of 'similar'.

The English /dʒ, tʃ, ʃ/ are similar to the Harbinese retroflex segments /ʈʂ, ʈʂʰ, ʂ/. Thus the English is phonemically similar to the Harbinese. The English /dʒ, tʃ, ʃ/ are closely similar to the Cantonese allophones [tʃ, tʃʰ, ʃ]. The only difference is that Cantonese does not contrast in voicing but aspiration. So, although they share the same allophones, they are still considered to be 'allophonically identical but phonemically different'. There are two different types of 'similar' in this group. They conform to the requirements of category 2; therefore, the English /dʒ, tʃ, ʃ/ are defined as 'similar'.

'New' consonants

This category can be easily classified. Segments which are absent in both Chinese dialects/languages are placed in this category. Among the 12 targeted consonants, /ʒ, θ, ð/ and dark /l/ [ɫ] do not exist in either dialect/language and were assigned to the category of ‘new’.

3.1.2 Classification of vowels

The 12 target English vowels /i, ɪ, u, ʊ, ɔ, ɒ, e, æ, ə, ɜ, ʌ, ɑ/ were also placed in the three categories.

‘Identical’ vowels

Cantonese has more diphthongs than English and Harbinese. The English diphthongs /eɪ, aɪ, aʊ, oʊ (GA), ɔɪ/ exist in Cantonese, so they are considered as ‘identical’ with English disregarding nuances between them such as vowel duration. However, Harbinese has a different /ɑu/ and lacks /ɔɪ/; thus, Harbinese is similar to English in these diphthongs as a whole. Therefore, English diphthongs as a whole are classified into category 1, ‘identical’. According to the placement rules, the English diphthongs /eɪ, aɪ, aʊ, oʊ (GA), ɔɪ/ are classified as ‘identical’.

‘Similar’ vowels

The English /ɔ/ exists in Cantonese but is absent from Harbinese. However, Cantonese phoneme /ɔ/ has two allophones [ɔ] and [o]. English /ɔ/ is phonemically identical but allophically different in Cantonese. Thus, the relationship of the English /ɔ/ with each dialect/language is respectively ‘similar’ and ‘new’. Based on the rule of categorisation, the English /ɔ/ is classified in the ‘similar’ category.

The English /ɛ/ is phonemically identical but allophonically different in Cantonese because Cantonese /ɛ/ has two allophones [ɛ] and [e], but English /ɛ/ is only allophonically identical in Harbinese. Thus, the relationship of English /ɛ/ with each dialect/language is ‘similar’ vs. ‘similar’ which belonging to different types of ‘similar’. The English /ɛ/ is placed into the category as ‘similar’.

The English /ə/ is phonemically identical but allophonically different in Harbinese because the Harbinese phoneme /ə/ has four allophones. Thus, the relationship between the English /ə/ and Harbinese /ə/ is 'similar'. However, it is new in Cantonese due to its absence in the dialect/language. According to the rules of category placement, the English /ə/ should then be classified as 'similar'.

The low back unrounded English phoneme /ɑ/ works as the allophone [ɑ] in Harbinese but is lacking in Cantonese. Thus, English /ɑ/ is allophonically identical but phonemically different in Harbinese while it is new in Cantonese. Therefore, /ɑ/ should be classified in the category of 'similar'.

The English central phoneme /ʌ/ is similar to /ɐ/ in Cantonese, and some researchers transcribe Cantonese short /ɐ/ as /ʌ/ (Gui, 2005:52); but it is new in Harbinese. Thus, /ʌ/ is filed in the category of 'similar'.

It is noted that the tense English /ɜ/ only exists in RP, but it occurs in GA in the rhoticised form /ɝ/. The English vowel /ɜ/ or /ɝ/ is phonemically similar to the Harbinese /ɚ/, but new in Cantonese. Thus, the English vowel /ɜ/ or /ɝ/ is put into the category of 'similar'.

The English /ɒ/ is new in Harbinese but phonemically similar to the Cantonese /ɔ/, and thus belongs to the category of 'similar'.

Harbinese does not have the tense-lax and length distinction. It has the two tense vowels /i/ and /u/ which are identical in English, but it lacks the lax counterparts. Cantonese has the tense-lax and length distinction, according to the literature, and has the tense /i/ and /u/ and allophones [ɪ, ʊ]. Thus, Harbinese is similar to English to some extent but Cantonese is more similar to English in this case. These are two different types of 'similar'. According to the placement rules, the English /i/ and /u/ vs. the English /ɪ, ʊ/ as contrast pairs are defined as 'similar'.

‘New’ vowels

The English vowel /æ/ does not exist in either dialect/language. Thus, it is categorised as ‘new’.

3.1.3 Classification of syllable structure

Chinese syllable structure is very simple compared to English. Chinese dialects/languages do not allow consonant clusters except C+glide onset cluster in Harbinese. Also, constraints are implemented on codas so that only a limited number of segments can work as codas, as mentioned in chapter 2. There are slight differences in syllable structure between the two target Chinese, Harbinese and Cantonese, which may lead to differences in L2 English. Thus, the different syllable structure between L1s are targeted in the research in order to detect any differences in production between different dialect/language speakers. All the patterns of syllable structure are displayed in Table 3.1 and are classified into the different categories to test whether Flege’s SLM works for syllable structure.

Table 3.1 Specific patterns of English syllable structure

No.	Structures	Patterns
1	/C+j+V+(C)/ (mostly in RP)	C+/j/+V+(C)
2	/C+w+V+(C)/	C+/w/+V+(C)
3	$CVC_{N:}$	CV/m/
4	$CVC_{P:}$	CV/p/ CV/t/ CV/k/
5	$(C)V(V)C_{N:}$	(C)/aɪn/ (C)/eɪn/ (C)/aʊn/ (C)/əʊn/ (C)/ɔɪn/

‘Identical’ syllable structure

The syllable structures /CjV(C)/ and /CwV(C)/ exist in English (usually /CjV(C)/ in

RP English) and Harbinese, but are absent in Cantonese which does not have C+glide onset. Thus, the onsets /Cj/ and /Cw/ in the syllable structure are the targets. Therefore, /CjV(C)/ and /CwV(C)/ are filed in category 1, 'identical'.

The syllable structure /CVm/ can occur in Cantonese and English but not in Harbinese, because /m/ cannot be a coda in the Harbinese syllable structure. Thus, /m/ is a target and /CVm/ is treated as identical because English and Cantonese have it but it is new in Harbinese, so it is filed as 'identical'.

'Similar' syllable structure

The difference between Cantonese and English is that the voiceless stop codas in Cantonese are the unreleased /p̚, t̚, k̚/. Thus, Cantonese is similar to this structure in English. Harbinese lacks this syllable structure because it does not allow any other consonants occurring in the coda position apart from /n/ and /ŋ/. Therefore, this group of syllable patterns conforms to the rule of category 2 and is assigned as 'similar'.

'New' syllable structure

The two /V/s in the syllable structure /(C)VVC_N/ represent the two segments of diphthongs and /C_N/ refers to the coda constrained to /n/. The syllable structure is /(C)VVn/. The specific patterns of this structure can be /aɪn/, /eɪn/, /aʊn/, /əʊn/, and /ɔɪn/ in English. Both dialects/languages prohibit a diphthong plus a nasal, so this English syllable structure is new for both dialects/languages. Thus, this group of structures goes into the category 'new'.

3.1.4 Classification of stress

'Similar' stress

Harbinese has two main stress patterns in disyllabic words: final stress (70%) and initial stress (30%), as mentioned in the literature review. These patterns in Harbinese can also be converted into two main prosodic templates of an iambic foot and a trochaic foot. However, there is no strong evidence that Cantonese has stress and it has been argued that there is no stress in this dialect/language (Gui, 2005; Bauer and Benedict, 1997; Hashimoto, 1972). Thus, the idea that Cantonese has no stress is adopted here. English is a language with many different stress patterns, mainly using the trochaic foot. Therefore, it is similar in Harbinese but new in Cantonese. Thus, stress is categorised as 'similar'.

3.2 Criteria for the degree of difficulty of categories

Does each category have the same degree of difficulty when learning the L2? The answer to this question is absolutely not. Questions as to which category is the most difficult and which is the least should be clearly answered. Both the CAH (Lado, 1957) and SLM (Flege, 1995) agree in predicting that the category of 'identical' is easy to learn. The main disagreement between these two theories concerns the categories of 'similar' and 'new' and which is more difficult to learn. I apply Flege's SLM throughout this study for consistency in predicting difficulties in the hypotheses and to test the accuracy of the SLM. Therefore, the category 'similar' in this study is predicted to be difficult and 'new' to be easy according to Flege's SLM (1995). It has to be mentioned that although 'new' is also considered to be easy, easy also has to apply to the category 'identical', i.e. it should be ranked the same because, as pointed out both by Lado and by Flege. In fact the category of 'identical' has a higher ranking since there is nothing to acquire assuming transfer is possible.

The category of 'similar' in my study is a huge group, and includes different types of similarity. As mentioned in section 3.1, there are three types of similarities in this study, which are 'allophonically identical but phonemically different', 'phonemically identical but allophonically different' and 'phonemically similar'. As to these three types of similarities, which type of similarity that is more difficult than the other types cannot be measured here because there are no sources in the literature that can be implemented to predict the different types of 'similar'. Flege's SLM (1997:17) considers that “interlingual identification occurs at the phonetic rather than the phonemic level”, and thus it can be assured that phonetic acquisition is more important than phonemic acquisition in the L2 to some extent. Based on that, I posit that ‘allophonically (phonetically) identical’ should be much closer than the ‘phonemically identical’. If they are much closer, this means that they are more similar. Therefore, according to Flege’s SLM (1995), the more similar it is, the more difficult it will be to learn. So I predict that the difficulties of learning different types of ‘similar’ are that: the allophonically identical is more difficult than the phonemically identical, which is in turn more difficult than the phonemically similar:

1. Allophonically identical but phonemically different

   IS MORE DIFFICULT THAN

2. Phonemically identical but allophonically different

   IS MORE DIFFICULT THAN

3. Phonemically different but similar

The criteria clearly shows how to place items in the categories with different degrees of difficulty for ‘identical’, ‘similar’ and ‘new’. All of the following hypotheses are proposed based on these criteria for difficulty. To clearly represent the relation, the categories with the different difficulties were shown in a logical way from the least difficult to the most difficult.

1. Category of ‘identical’: **the least difficult** (most easy)
2. Category of ‘new’: **somewhat more difficult** (not so easy)
3. Category of ‘similar’: **difficult**
   a. phonemically different but similar: **quite difficult**
   b. phonemically identical but allophonically different: **very difficult**
   c. allophonically identical but phonemically different: **the most difficult**.

According to the above more detailed criteria for difficulty, categories will be placed with those degrees of difficulty in all hypotheses. Therefore, the hypotheses for segments and syllable structures in the category of 'identical' are proposed to involve the least difficulty, or as easy. The hypotheses for the targets in the category of 'new' are proposed to involve less difficulty or to be less easy. The hypotheses for the remaining targets placed in the category of 'similar' are predicted to be difficult. As seen above, the category of 'similar' is a huge group with three different types of 'similar', so each type of 'similar' has been allocated different degrees of difficulty. If one hypothesis for a target involves two types of 'similar', the difficulty expected in the hypothesis would be predicted as shown above. Because hypotheses must be predicted as absolutes, in the hypotheses for the category of 'similar', any dialect/language with similarity to English being more difficult to learn than the other should be phrased so that the former dialect/language cannot acquire the English but the other can.

Note that this classification disregards nuances between the L2 and L1 such as having the same IPA but articulated in different places, for example as with the English alveolar [t] and French [t]. This is because it would make the study unnecessarily complex to take those nuances into account. A second reason is that, if the closest type of 'similar' in this research, such as 'allophonically identical but phonemically different', can eventually be confirmed to be difficult, it would imply that nuanced differences should be even more difficult. Moreover, if the closest type of 'similar' is classed as easy, it can indicate that nuances might be the only type of similarity leading to it being more difficult, thus it may be that similarity being difficult could be mainly attributed to subtle differences between sounds in L2 and L1. That could be a subject for further research.

According to Flege's SLM, segments, syllable structures and stress in the category of 'similar' are predicted to involve more difficulty in learning. However, anything that is ‘new’ in the L2 would pose fewer difficulties, and leaners would make fewer errors than for ‘similar’.

3.3 Hypotheses

3.3.1 Hypotheses for consonants

These hypotheses are based on Table 3.2.

Table 3.2 Hypotheses for consonants

Categories	No.	Harbinese ($L1_H$)	Cantonese ($L1_C$)	English (L2)	Hypotheses	Predictions
Identical	H1.1	/x/	/h/	/h/	The English consonant phoneme /h/ is identical in Cantonese but similar to /x/ in Harbinese, so Harbinese speakers will make errors.	Harbinese will make errors.
	H1.2	/w/ [w] and [v]	/w/	/w/	The English /w/ is identical in Cantonese but similar in Harbinese because Harbinese has /w/ with two allophones [w] and [v], so the English /w/ is phonemically identical but allophonically different in Harbinese. Thus, the Harbinese will make errors with /w/ .	Harbinese will make errors.
	H1.3	/n/ and /l/[l]	Neutralise /n/ and /l/[l]	/n/ and /l/[l]	The English /n/ and /l/ [l] exist in Harbinese and Cantonese but are sometimes neutralised in Cantonese, thus the English /n/ and /l/ [l] are identical with Harbinese but only similar in Cantonese. Therefore, the Cantonese will make errors.	Cantonese will make errors with /n/ and /l/ [l].
Similar	H1.4	[v] /w/	/	/v/	The English /v/ is new in Cantonese but exists as an allophone in Harbinese where it is similar. Therefore, according to the degree of difficulty for the category of 'similar', it is hypothesised that Harbinese speakers will make errors.	Harbinese will make errors.
	H1.5	/ɻ/	/	/ɹ/	The English /ɹ/ is similar to the Harbinese retroflex /ɻ/ but absent in Cantonese. According to the difficulty scale, Harbinese speakers will make errors but Cantonese speakers will not.	Harbinese will make errors.
	H1.6	/ʈʂ, ʈʂʰ, ʂ/	[tʃ, tʃʰ, ʃ]	/dʒ, tʃ,	The English /dʒ, tʃ, ʃ/ are phonemically similar to the retroflex /ʈʂ,	a. Both should

			/ts, tsʰ,s/	ʃ/	ʈʂʰ, ʂ/ in Harbinese. However, /dʒ, tʃ, ʃ/ are comparatively allophones in Cantonese, so they are allophonically identical but phonemically different in Cantonese. According to the criteria of degree of difficulty in the category of 'similar', it is hypothesised that a) both should be difficult but b) the Cantonese will make errors.	be difficult. b. Cantonese will make errors.
New	H1.7	/	/	/θ, ð/, /ʒ/, (/l/) [ɫ]	The English /θ, ð, ʒ/ and (/l/) [ɫ] are absent from both dialects/languages. According to Flege's SLM, 'new' is easy; therefore, a) neither will make errors with these segments; b) if they do make errors, there will be no differences between the two dialects/languages.	a. Neither will make errors. b. No differences.

3.3.2 Hypotheses for vowels

These hypotheses are based on Table 3.3.

Table 3.3 Hypotheses for vowels

Categories	No.	Harbinese ($L1_H$)	Cantonese ($L1_C$)	English (L2)	Hypotheses	Predictions
Identical	H2.1	/eɪ, aɪ, ɑu, oʊ (GA)/–	/eɪ, aɪ, aʊ, oʊ (GA), ɔɪ/	/eɪ, aɪ, aʊ, oʊ (GA), ɔɪ/	The English diphthongs /eɪ, aɪ, aʊ, oʊ (GA), ɔɪ/ exist in Cantonese but are only similar in Harbinese because Harbinese has a different /ɑu/ and lacks /ɔɪ/. Harbinese	Harbinese will make errors.

					English learners will therefore make errors with these diphthongs.	
Similar	H2.2	/	/ɔ/ [ɔ] and [o]	/ɔ/	Harbinese lacks /ɔ/ while Cantonese has this English phoneme but with two allophones. Thus, English /ɔ/ is phonemically identical but allophocially different in Cantonese but new in Harbinese. It is hypothesised that Cantonese speakers will make errors.	Cantonese will make errors.
	H2.3	[ɛ] /a/	/ɛ/ [ɛ] and [e]	/ɛ/	The English /ɛ/ is phonemically identical and allophonicially different in Cantonese but allophonically identical and phonemically different in Harbinese. Thus, the relation is 'similar' vs. 'similar'. Based on the rules of degree of difficulty, the type of similarity with 'allophonically identical' is predicted to be more difficult than that with 'phonemically identical'. Therefore, it is hypothesised that a) both should be difficult but b) Harbinese will make more errors.	a. Both should be difficult. b. Harbinese will make more errors.
	H2.4	/ə/ [ə,e,o,ɤ]	/	/ə/	Harbinese is phonemically identical but allophonically different from the English phoneme /ə/, while Cantonese does not have this phoneme. Thus the Harbinese will make errors.	Harbinese will make errors.
	H2.5	[ɑ]	/	/ɑ/	The English vowel phoneme /ɑ/ is allophonically identical but phonemically different in Harbinese. Thus, the English /ɑ/ is similar in Harbinese but new in Cantonese. According to Flege's SLM, the 'similar' is more difficult than the 'new', and so Harbinese speakers	Harbinese will make errors.

					will make errors and the Cantonese will not.	
	H2.6	/	/ɐ/	/ʌ/	Harbinese does not have the English /ʌ/ but Cantonese has the phonemically similar /ɐ/, so Cantonese speakers will make errors with /ʌ/.	Cantonese will make errors.
	H2.7	/	/ɔ/	/ɒ/	Harbinese does not have the English /ɒ/ but Cantonese has the phonemically similar /ɔ/, so Cantonese speakers will make errors with /ɒ/.	Cantonese will make errors.
	H2.8	/ɚ/	/	/ɜ/-/ɝ/	Cantonese does not have the English /ɜ/ or /ɝ/ but Harbinese has the phonemically similar /ə/ (/ɚ/), so Harbinese speakers will make errors with the English /ɜ/ (/ɝ/).	Harbinese will make errors.
	H2.9	/i, u/ /	/i, u/ [ɪ, ʊ]	/i, u/ /ɪ, ʊ/	Harbinese does not contrast tense-lax but Cantonese does. Both Chinese dialects/languages have the English tense vowels /i, u/ while Harbinese does not have the lax counterparts /ɪ, ʊ/ but Cantonese has [ɪ, ʊ] as allophones. Therefore, each dialect/language is similar to English in a different way. According to Flege's SLM where 'similar' is difficult, it is predicted that both are difficult but the Cantonese will make more errors.	Cantonese will make errors with tense-lax /i, u/ and /ɪ, ʊ/.
New	H2.10	/	/	/æ/	The English vowel phoneme /æ/ is absent from both dialects/languages. According to Flege's SLM, 'new' is easy; therefore, a) neither will make errors with this segment; b) if they do make errors, there will be no differences between the two dialects/languages.	a. Neither will make errors. b. No differences.

3.3.3 Hypotheses for syllable structure

These hypotheses are based on Table 3.4.

Table 3.4 Hypotheses for syllable structure

Categories	No.	Harbinese ($L1_H$)	Cantonese ($L1_C$)	English (L2)	Hypotheses	Predictions
Identical	H3.1	/CjV(C)/ (mostly in RP) or /CwV(C)/	/	/CjV(C)/ (mostly in RP) or /CwV(C)/	Cantonese does not have C+glide onset but Harbinese and English have in the syllable /CC_GV(C)/. Thus, Cantonese but not Harbinese speakers will make errors with English syllables with glides such as in /CjV(C)/ (mostly in RP) and /CwV(C)/.	Cantonese will make errors.
	H3.2	/	/CVm/	/CVm/	Harbinese does not allow /m/ in the coda position while Cantonese does. Therefore, Harbinese but not Cantonese speakers will make errors with the English syllable structure /CVC_N/ (/CVm/).	Harbinese will make errors.
Similar	H3.3	/	/CVp̚, t̚, k̚/	/CVp, t, k/	Cantonese has a similar syllable /CVC_P/ (/CVp̚, t̚, k̚/) to English but the syllable structure is new in Harbinese, so Cantonese speakers but not Harbinese speakers will make errors with the English syllable pattern /CVC_P/ (/CVp, t, k/).	Cantonese will make errors.
New	H3.4	/	/	/(C)VVn/	Harbinese and Cantonese prohibit the pattern of a diphthong plus a nasal coda, especially with /n/, such	a. Neither will make errors.

					as /(C)VVC$_N$/ in their L1s. Therefore, according to Flege's SLM, 'new' is easy and it is predicted that a) neither will make errors with the English syllable structure /(C)VVC$_N$/; b) if they do make errors, there will be no differences between the two dialects/languages.	b. No difference.

3.3.4 Hypothesis for stress

This hypothesis is based on Table 3.5.

Table 3.5 Hypothesis for stress

Categories	No.	Harbinese (L1$_H$)	Cantonese (L1$_C$)	English (L2)	Hypothesis	Predictions
Similar	H4	Two stress patterns with trochaic but main iambic foot	/	Various stress patterns with trochaic foot	English has various stress patterns with trochaic feet but Harbinese has two main stress patterns mainly with the iambic foot but also with the trochaic foot as mentioned in the literature review. Harbinese is similar to some extent to English in the prosodic template for stress. However, Cantonese has no stress, so it is new in Cantonese.	Harbinese will make errors.

					Therefore, the stress pattern should be classified into the category of ‘similar’ and, it is hypothesised that Harbinese speakers will make errors with English stress pattern with iambic foot but the Cantonese will not.	

3.3.5 Hypothesis for exposure length

Hypothesis 5

L1 based errors will decrease with increased exposure to English.

3.4 Participants

In this research, participants were recruited in the two dialect/language-speaking areas of Harbin and Guangzhou, because the Harbin Northeastern dialect and Guangzhou Cantonese are both representative dialects in these areas. Moreover, both Harbin and Guangzhou are provincial capitals with a certain economic and cultural influence in their regions and with comparatively greater populations. Therefore, participants were recruited from speakers of the Harbin Northeastern and Guangzhou Cantonese dialects.

In China, formal education can generally be classified into four periods. These include primary school, ranging from Year 1 to Year 5 or 6, a secondary school period from Year 6 or 7 to Year 9, a 3-year high school period from Year 10 to Year 12 and a 4-year higher education period from Year 1 to Year 4 with the exception of some specific disciplines such as medicine which involve five years of study.

To explore the acquisition of English phonology, participants were recruited from various levels of English instruction. Students at primary school are at the low beginning level, students at middle school are high beginners, students at high school are low intermediate, and university non-English-major students are at high intermediate level. Advanced level is for English major upper year students. Owing to the limited English vocabulary that primary school students have mastered, this period of schooling was excluded although it would have been very interesting to investigate the development of English at this level. In addition, English major students at advanced levels were not included in the present research because the duration and intensity of their English exposure is distinct from that of non-English-major students. Therefore, in this research participants were recruited from three main periods of schooling: middle school, high school and university. This corresponds to the three levels of English exposure at high beginner, low intermediate and high intermediate levels.

Before the recruitment of participants, other factors had to be explicitly considered. Variations exist in the duration of English exposure, as well as disparities between schools, regions, school teaching quality, and even family backgrounds. Generally, formal English teaching formally starts in primary school in Year Three (MEPRC, 2001), but this may vary slightly from school to school and may depend on levels of local economic development. For example, some primary schools in regions or cities with good economic development may provide English teaching earlier, but most schools across the country follow the guidance of the MEPRC. In addition, the actual duration of participants' exposure to English may also depend on family background although they receive the same teaching at English school. A student may be exposed to English in pre-school at a very young age if the parents have a strong awareness of how English language mastery is important and beneficial to their children's future. If a child is born into a wealthy family, then bilingual schooling may be used to provide a sound early English education. However, no matter what age the students start to learn English or their length of exposure to English, they eventually take the same national examination in English for entrance to university in Year 12. Moreover, students from the same school year may differ in English proficiency, and so will their English teachers, as will the teaching methods used, and so the quality of English education across schools and cities also varies.

If these dimensions were comprehensively taken into account, the present research could not have been carried out because it would be impossible to control for all factors. Therefore, the factors mentioned above were not taken into account as influential variables but instead the number of years of exposure to English through schooling is taken as a reference criterion. However, in attempting to guarantee the validity and reliability of the data collected from the participants, certain factors were considered seriously as main criteria for the selection of participants in the research, including the quality of schools and the dialects spoken by participants.

Participants in the study were restricted to those who were born and raised in either Harbin or Guangzhou and who had naturally acquired Harbinese or Guangzhou Cantonese as their main Chinese dialect/language. Because Mandarin is the only official language in China and is dominant and widespread, it was deemed unrealistic to look for monolingual dialect/language speakers who could not speak Mandarin, especially among younger generations. However, speaking dialects other than Harbinese and Cantonese (excluding Mandarin) alongside their first dialect was a factor used to exclude participants so as to avoid the influence of other dialects.

In addition, the background of the schools and universities was considered as a selection criterion to ensure that participants came from similar school backgrounds in terms of teaching quality and school ratings in both cities. Schools are ranked in the relevant cities mainly based on teaching quality, which can be generally classified in terms of provincial schools, city schools and common schools while universities are ranked across the whole country and classified as top universities (Level A), Level B universities and common universities (Level C). Which level of university Chinese students can enrol in mainly depends on their scores in the national exams for university entrance. Therefore, the selection of participants at schools and universities at the same level in both cities reduced disparities between the samples of dialect/language speakers.

Following these criteria, in the present thesis, city-level middle schools and high schools and Level C universities in both cities were chosen to recruit participants. Originally, a 3-year gap between two adjacent levels was selected; that is to say, it was planned to recruit the participants from Year 9, Year 12 and Year 3 at university; however, Year 12 is the final year in high school for participants who have to focus their utmost effort and time on preparing for the national examination for university entrance. Their schools did not permit anyone to disrupt their learning at this crucial time. Thus, the planned Year 12 had to be changed to Year 11 after contacting the relevant schools for permission for data collection. Eventually, 15 male and female participants, were recruited from each Year 9 at middle school, Year 11 at high school and Year 3 from university, resulting in 90 participants in all, as shown in Table 3.6 below. One facet of the present study has to be mentioned that it was of classroom learners, and at various levels, based on their year in school/at university, whereas Flege's studies have been of immigrants. This means the study had the potential to make a contribution to what we know about similar vs. new in the context of a different learner group.

Table 3.6 Participant information

School Years	Harbin Group				Guangzhou Group			
	Number	Age range	Age mean	Std. Deviation	Number	Age range	Age mean	Std. Deviation
Year 9	15	14-16	14.40	0.632	15	14-15	14.33	0.488
Year 11	15	16-17	16.33	0.488	15	16-17	16.47	0.516
Year 3 in Uni	15	20-22	21.00	0.756	15	20-21	20.27	0.458
Total	45	14-22	17.24	2.870	45	14-21	17.02	2.527

3.5 Data collection

3.5.1 Participant information sheet and consent form

According to the requirements of Newcastle University concerning ethical issues in research, all research involving human participation should obtain ethical approval prior to conducting data collection and the researcher should provide an information sheet and consent form to participants to guarantee their understanding of the relevant research information and the procedures for participation and their consent to participate in the research. It can also guarantee the legality of the participants' permission for the use of the data.

Ethical approval for the present research was obtained from the University before data collection and a formal written form to give participants information and gain their consent was produced with the University Crest as a letterhead and text stating the project title, basic information on the study, and an invitation and expression of gratitude for their participation as well as key points concerning consent seen in Appendix 1.1 and Appendix 1.3. Participants were provided with a Chinese version of the information and consent form for ease of understanding and for their convenience (see Appendix 1.2 and Appendix 1.4). To obtain authentic data without any influence of the information on the consent forms, I debriefed them afterwards about the real purpose of the study, so that they would not correct their pronunciation with too much care and clear awareness when speaking.

3.5.2 Questionnaire

The questionnaire was designed to select appropriate participants and to obtain more information about them. It contained 28 questions covering topics in the areas of personal background, language learning history and dialect use, and understanding of the varieties of English (see Appendix 2.1 Questionnaire (English version).

Basic demographic information from the questionnaire included name, gender, age and hometown (birthplace). The rest of the questionnaire consisted of multiple-choice questions for participants to choose the best answer and open questions for them to provide their ideas and opinions. Question 1 (Q1) was designed to ascertain the participants' educational level and to know whether they were appropriate participants from the targeted school years. Participants from middle school had to choose 'B' and 'Year 9', those from high schools 'C' and 'Year 11' and university participants 'D' and 'Year 3'. Q2 to Q7 were designed to elicit personal information about where participants and their parents were born and raised. This information was essential so as to exclude those who did not meet the requirements such as those not born and raised in the targeted city, those not using the local dialect as their first dialect, and those influenced by parents who spoke other dialects. With regard to language use, questions 8-14 were included to find out which dialects/languages participants and their parents spoke and used in their daily lives. Among these, Q9 and Q11 were especially significant in discovering if the targeted dialects were their first dialects and also to know whether they would be influenced by other dialects. Q12 was asked to find out if they were bilingual speakers using both Cantonese and Mandarin at home. The subsequent questions from 15 to 23 checked the participants' language learning history, including the ages they started to learn Mandarin and English, the length of studying English and their experience of studying English with English native teachers as well as the methods used in studying English outside school. The last five questions (Q24-Q28) investigated the participants' preferences for English varieties and the ability to distinguish between them.

The answers to Q1-7, Q9-12 and the question about the hometown were considered to be essential constraints in selecting appropriate participants in the research so as to avoid the influence of other factors.

3.5.3 Testing instrument

The possible difficulties experienced by the two groups of Chinese dialect/language speakers have been discussed in the previous chapter and the hypotheses presented in section 3.1. The testing instrument was then designed to provide data to test the hypotheses. A well-designed and well-constructed testing instrument is necessary in order to achieve adequate data collection and to allow the required analysis. The instrument included three sections involving word translation, Chinese reading and carrier sentence reading.

The participants had distinct levels of exposure to English through schooling from middle school to university, and their store of vocabulary was also expected to differ greatly. Moreover, English textbooks can vary from city to city and even from school to school. To obtain useful and valid data, English words were selected as test tokens with great care from all of their textbooks so as to guarantee that the words used were already in their vocabularies, and thereby to further guarantee their ability to pronounce

English words. In order to gain valid data, this was especially important for participants at beginner level. Therefore, English words were chosen as test tokens based on the English textbooks used by participants at beginning level. However, the participants had not studied all of the vocabulary in the Year Nine English textbooks when the data were collected because they had only recently been moved from Year eight to Year nine one month previously. This means that the English textbooks for Year eight used by participants in both cities were consulted for word selection.

3.5.3.1 Word translation

The test was designed to detect the successful acquisition of English phonology in segments, syllable structure and also stress for speakers of the two different dialects/languages. Specific words were chosen to test each hypothesis. To avoid the influence of reading (orthography) on the participants' pronunciation of the words (Bassetti and Atkinson, 2015), all test words were translated into Chinese and presented in Chinese characters. There were 174 test tokens in this section, excluding repeated words for different targets in the test, covering the investigation of segments of consonant and vowel phonemes, syllable structure and stress. The specific procedures for data collection through this instrument can be seen in below section 3.5.4 Procedures of data collection.

Test tokens 1-95 for differences in consonants

Test tokens 1-69 were randomly arranged to detect differences in pronunciation of the consonant segments /v, w, ʃ, ʒ, tʃ, dʒ, h, θ, ð/ and then to test hypotheses H1.1, H1.2, H1.4, H1.6 and H1.7, while test tokens 70-95 were used for the consonant segments /n, l, ɹ/ and to test to hypotheses H1.3, H1.5 and H1.7.

It was also important to look at phonemes in different positions; that is, different phonetic realisations when consonants occur in different phonological positions in English. Taking /t/ as an example, occurring word initially its realisation is [tʰ], but [t] and [ʔ/t̚] respectively when occurring word-medially and word finally. Also, as mentioned in the literature review in section 2.3.1, L1 transfer of syllable constraints also brings the difficulties to L2 segments' acquisition (Zampini, 2008). Thus, it is interesting to see how L2 English speakers treat these phonemes occurring in different positions. To detect all potential consonant singleton differences produced by the different groups of speakers, tokens for each consonant singleton occurring in initial, medial and final positions were selected. Apart from consonants in some specific positions with phonotactic restrictions, such as /h/ and /w/ not occurring in the final position and /ʒ/ seldom occurring in initial and final positions, all consonant phonemes in all possible positions were covered. Examples of test tokens of these targeted consonants are illustrated below (see Appendix 3.1 and Appendix 3.2 for full English and Chinese versions of the test tokens):

1. 兽医 'vet'	5. 孩子 'child'	9. 他们 (主格) 'they'
2. 湿的 'wet'	6. 日本 'Japan'	10. 很，非常 'very'
3. 船 'ship'	7. 希望 'hope'	11. 想知道 'wonder'
4. 通常 'usually'	8. 认为 'think'	12. 鞋 (单数) 'shoe'

…

Test tokens 96-138 for differences in vowels

This part explored the differences generated by speakers of different dialects/languages in 12 vowel phonemes /i, ɪ, u, ʊ, ɜ, ə, ɔ, ɒ, ɑ, ʌ, æ, ɛ/ and five diphthongs /eɪ, aɪ, aʊ, əʊ(oʊ), ɔɪ/, to test hypotheses H2.1 to H2.10. Three tokens were assigned to each of the 17 vowels, thus totalling 51 tokens; however, 43 tokens were eventually selected, as shown in Appendix 3.1 or Appendix 3.2, because eight were repeated. For example, *ship* would have been tested for the lax vowel /ɪ/ but had also been used to test the consonant /ʃ/. Therefore, these repeated items were excluded from this part.

Test tokens 139-162 for differences in syllable structure

This subgroup of test tokens was designed to test hypotheses H3.1 to H3.4 concerning syllable structures. Tokens were selected based on the following patterns of structures in Table 3.7 with three tokens per pattern for structures 1 to 4 and two tokens for structure 5. 24 tokens in all were adopted in this part, excluding 4 repeated ones.

Table 3.7 Syllable structure and pattern

No.	Structure	Pattern
1	/(C)+j+V+(C)/ (mostly in RP)	C+/j/+V+(C)
2	/(C)+w+V+(C)/	C+/w/+V+(C)
3	$CVC_{N:}$	CV/m/
4	$CVC_{P:}$	CV/p/ CV/t/ CV/k/
5	$(C)V(V)C_{N:}$	(C)/aɪn/ (C)/eɪn/ (C)/aʊn/ (C)/əʊn/ (C)/ɔɪn/

Test tokens 163-174 for differences in stress

As mentioned previously, most Chinese words are disyllabic. To achieve comparability, disyllabic English words were also used as tokens. In addition, both Chinese dialects/languages have iambic stress patterns (although stress in Cantonese does not seem to be very clear) according to the discussion in Chapter 2. Disyllabic tokens had to be selected with stress on the first syllable but not on the second, so that it would be easy to see the influence of an iambic stress pattern in the dialects/languages if stress was placed on the second syllable by the participants. It would not be so easy to see an L1 influence if the second syllable of tokens was stressed. So, L2 disyllabic English words were selected as test tokens to test suprasegmental differences and to detect any influence of L1 dialect/language stress patterns. These words included 8 compounds and 4 monomorphemic words, and the results were used to test hypothesis H4.

3.5.3.2 Chinese reading

Harbinese speakers tend to pronounce /w/ as [v] in Mandarin while Cantonese speakers always substitute /l/ for /n/, as mentioned in chapter 2. In keeping with other studies, a Chinese reading test was designed to target the three consonants /n, l, w/ for further confirmation of the claim certain consonants are treated differently depending on the first dialect/language of Chinese (please see section II in Appendix 3.1and Appendix 3.2).

Test tokens 175-192 for differences in /n, l, w/ in Chinese

18 Chinese words as tokens were carefully selected to target these three consonants and the target consonants in these 18 words were expected to be pronounced the same in both Chinese dialects/languages. For example, 牛奶 'cow milk' was used to target the onset /n/ in 奶 'milk' which is pronounced the same in Harbinese [**n**ai3] and Guangzhou Cantonese [**n**a:i1]. Conversely, the onset of the first character 牛 'cow' in 牛奶 is pronounced differently as [n] in Harbinese and [ŋ] in Cantonese; thus, it cannot be used for this comparison. In these tokens, /n/ occurred nine times in seven words, /l/ five times in five words and /w/ eight times in six words. Therefore, the sounds /n/, /l/ and /w/ were respectively targeted with frequencies of nine, five and eight tokens.

3.5.3.3 Carrier sentence

The use of carrier sentences is a method to help expand the length of children's utterances and to aid in language learning. However, in the present research carrier sentences were used to check how participants produced the target sounds in running speech instead of in isolation. Deterding, Wong, and Kirkpatrick (2008:156-157) mention that Hong Kong Cantonese speakers often omit final consonants in their English speech during interviews. Also, according to my previous classroom teaching observations, some Guangzhou Cantonese learners of English pronounce English words in isolation correctly without difficulty, but when they speak in English or read English texts, they make mispronunciations. For example, Cantonese speakers can correctly pronounce words with word-final consonant stops in isolation but will omit the stops in their speech because of the influence of the Cantonese non-release of these plosives. Therefore, the use of carrier sentences was intended to compensate for the deficiencies when using isolated words in the word translation test. The carrier sentence used in this task was *I think she said* ____ *yesterday*, which was carefully designed to avoid the influence of adjacent words on the test tokens. Anything following the word is likely to have an effect. It is true that [j] in *yesterday* can be combined with previous coda to be a consonant cluster with a glide [j], but participants or English learners would not combine the previous coda and [j] and pronounce something like [‘tjestədeɪ] or [‘njestədeɪ] because *yesterday* is an easy word for them to pronounce. Thus, the influence of [j] in *yesterday* on previous sound is little. After careful consideration, *yesterday* was determined to be an appropriate word put after the blank in this sentence.

Ideally, all of the test tokens in the word translation section should have been tested with the exception of those for stress because stress is influenced by various linguistic factors in a sentence. However, allowing for the burdensome workload for the participants because of an enormous number of test tokens, it would have taken more than one hour for each participant to complete the whole test without a break, which would have led to physical and mental fatigue and may have resulted in resistance to continued participation in the data collection. Therefore, those tokens which may have resulted in different pronunciations for isolated words compared to those in connected speech as mentioned in the literature were checked, i.e. the tokens covered the consonant segments and syllable structure of $CVC_{N\&P.}$ Finally, a total of 107 tokens was checked in this task, the specific tokens can be seen section III in Appendix 3.1 and Appendix 3.2).

In summary, each participant had to produce 299 tokens in total, including 174 tokens in the translation task, 18 tokens in the Chinese reading task and 107 in the carrier sentence task.

3.5.4 Procedures of data collection

Data collection should be conducted in a quiet place for audio recording without the interference of noise; however, quiet surroundings could not be guaranteed in the schools and universities because they had no audio laboratories and the noise coming from roads, corridors, teaching in other classrooms or activity during break intervals could not be prevented even if a special room was arranged. Efforts were made to ensure high quality recordings under these difficult conditions.

The schedule for meeting each participant for data collection was arranged and an appointment was made to collect data from participants one by one. A participant information sheet and consent form was provided to guarantee their knowledge of the research and their right to withdraw at any time without giving reasons. When they had finished reading the information sheet and consent form, their signatures were obtained to record their permission to use their data.

After that, a questionnaire was provided to obtain basic information on the participants' personal background and their language or dialect learning and use. When the questionnaire was completed, it was swiftly checked to see whether they had omitted to answer some questions due to carelessness or whether they needed further assistance with some questions. After this, the questionnaires were collected. The procedure took approximately five to ten minutes.

Next, in the main section of data collection, the participants produced the speech data. All test tokens were shown using PowerPoint software with one token per slide. The complete set of slides consisted of three parts to correspond to the three test tasks. The tokens shown in Chinese characters with numbers in front of them were placed in the middle of the slides in a font size of 100-120 and in white set off against a black background. For difficult words, some clues were given below the main test tokens to help them pronounce them in case they could not recall and produce them. These clues included indications of the word class of a token or its initial letter, or a blank spacing in a phrase or a short sentence using the targeted token. In the carrier sentence task, the sentence *I think she said ____ yesterday* was always placed below the retested tokens but these were still represented in Chinese as before.

The audio recorder was set to be ready before data collection was carried out. Subsequently, around two or three minutes of guidance and instructions were provided for the participants to ensure that they were familiar with the testing procedure and materials so as to yield valid and high-quality data. Participants were required to utter each token in English with its number in front of it at a normal speed and to leave a second-long pause between every utterance. It should be mentioned that the Chinese reading test differed from the first test in which participants were required to pronounce tokens in English; before they proceeded to take this test, participants were instructed to pronounce the Chinese tokens in their own dialect/language. Also, the participants saw the instruction slide before they could continue. In the third part, participants were required to fluently read the carrier sentence and supply the test words at a natural speed and in a natural voice.

In the process of recording, a reminder was given to participants to use another word if they could not pronounce the targeted one even though it had a similar meaning, or they were required to make another attempt if it was pronounced mistakenly. For example, the targeted tokens *choice* and *away* were mistakenly produced by participants as *choose* and *leave*. In this case, they were asked to make another attempt and provided with a clue, such as it being a noun or an adverb, to help them to produce the correct targeted tokens. After the first part was completed, a short break was given to alleviate tiredness. Refreshments and water were provided during the break and gifts such as large chocolate bars brought from the UK or a set of nice stationery were given after the completion of data collection as compensation for their time. The entire sound production stage lasted from 30 minutes up to one hour, mainly depending on participants' familiarity with the vocabulary. All of the data were collected from each participant by me individually.

3.6 Data analysis

Without valid scientific methods to analyse data, there would be no valid, reliable and accurate outcomes; thus, the results would not contribute to testing hypotheses and answering research questions. In this study, I was not only interested in errors and accuracy rates, which would allow hypotheses to be tested using quantitative analysis, but also the types of errors made which would allow the use of qualitative analysis to determine what types of errors participants committed in production in different dialects/languages. Therefore, both quantitative and qualitative analysis was used. To guarantee the quality of the analysis, a sufficiently extensive period was taken for the data analysis. It has to be mentioned that I began analysis as soon as I had data from my first participant to check that things were proceeding as I had hoped; thus, no pilot study was carried out in the study.

The analysis was conducted using various tools and methods including International Phonetic Alphabet (IPA) transcription of the tokens, auditory transcriptions with analysis using Praat software for audio recordings, and SPSS and Excel were used for statistical analysis and chart plotting. Also, a native RP English speaker with linguistics knowledge was recruited to transcribe a small set of my data to compare against my own transcriptions in order to enhance their reliability and validity.

To protect participants' privacy, all of their data were analysed anonymously and confidentially. All questionnaire and recorded data were labelled and represented using a code consisting of a city name, a school level and a number. For example, data from the Harbin middle school were labelled from HM1 to HM15, with H standing for Harbin, M for middle school level, and 1-15 for each participant in the group. The number was assigned to each participant in accordance with the sequence in which they participated in data collection to maintain consistency and avoid mismatches between participants and their recordings and corresponding questionnaires. Similarly, data from Harbin high school and Harbin university were referred to as HH1-15 and HU1-15. The same applied to the Guangzhou Cantonese group, where GM1, GH1 and

GU1 respectively refer to the first participants from middle school, high school and university of Guangzhou.

3.6.1 Analysis of questionnaires

The questionnaire data was analysed using the statistical software SPSS. Answers to 28 questions (see Appendix 2.1 Questionnaire (English version)) as well as demographic information were converted into a total number of 52 variables and appropriate values set for each variable in the variable view panel of SPSS Statistics Editor. All variables were arranged based on their sequential occurrence in the questionnaire. For example, the first data item in the questionnaires is name, which was placed as the first variable with the selection of 'string' for type and appropriate value for width for text-input questions, while gender as the second variable with 'numeric' selected as type with two fixed response choices encoded as 1 for male and 2 for female. Multiple choices answers were also encoded numerically with 1 for choice A, 2 for B and 3 for C and so on. All names were represented by codes.

After the appropriate settings were given for all variables, data from the questionnaires were input into cases in SPSS one by one and group by group starting from low schooling level to high for the Harbin group first and then Guangzhou. All the aforementioned essential questions (hometown, Q1-7, Q9-12) in the questionnaire of data collection were tested one by one with a certain hierarchy as essential constraints. Descriptive statistics and case summary reports were generated by SPSS to easily check the participant's violation of any selection criteria. Those who violated the criteria were eliminated from the data to maximally ensure the least influence from the extraneous factors.

'Hometown' was set as the first constraint so that participants not coming from Harbin and Guangzhou were excluded immediately because this research considers only native dialect speakers from these two cities. Two participants were excluded because they did not come from Guangzhou, as seen in Tables 3.8 and 3.9 clearly indicates the two specific violators. Therefore, data from GU-15 and GM-13 were deleted.

Table 3.8 Summary of hometown

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Harbin	45	50.0	50.0	50.0
	Guangzhou	43	47.8	47.8	97.8
	Others	2	2.2	2.2	100.0
	Total	90	100.0	100.0	

Table 3.9 Summary of hometown with participants' IDs

Hometown				ID
	Others	1		GU-15
		2		GM-13
		Total	N	2
	Total	N		90

Answers to Q1 were also used to exclude participants who were not in the targeted school years. One participant at university was excluded because s/he was a fourth year student rather than third year, as shown in Tables 3.10 and 3.11. Therefore, the data for HU-13 was deleted.

Table 3.10 Summary of University level

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	D. Year 3	29	32.2	96.7	96.7
	D. Year 4	1	1.1	3.3	100.0
	Total	30	33.3	100.0	
Missing	System	60	66.7		
Total		90	100.0		

Table 3.11 Summary of university level with participants' ID

Hometown	Harbin	Q1D	D. Year 4	1		HU-13
				Total	N	1

Subsequently, the results under Q2 and Q3 were checked and it was found that another four participants did not meet the requirements because they answered 'No' to questions 2 and 3 as seen in Table 3.12, which means that they were not born and raised in the relevant cities or were born there but had not lived there for a long time. Therefore, these four data sets were also excluded.

Table 3.12 Summary of Q2 and Q3

Q2	B. No	Q3	A. Yes	1		HU-5
				2		GU-8
				Total	N	2
			B. No	1		HU-7
				2		HU-9
				Total	N	2
			Total	N		4

Q4 was asked so as to exclude those who had lived outside of their hometown for more than three years and Q5 was to further ensure that there was no influence of other dialects on participants because they had lived in the targeted cities in the most recent three years, which was crucial. Therefore, those answering question 4 with A or B and 'Yes' to Q5 were eligible participants. Those with other choices had to be excluded because they were considered inappropriate for the research. From the analysis of data, one participant out of 83 violated Q4, as seen in Table 3.13, but the remaining cases all met the criterion for Q5 and were retained.

Table 3.13 Summary of Q4

Q4	A. 0 years	Total N	79
	B. 1-3 years	Total N	3
	E. More than 10 years	1	HM-1
		Total N	1
	Total	N	83

Q6 and Q7 were used to exclude those whose parents were not born and raised in their hometowns, but their parents not having been born but raised there was not considered to violate the criteria if they had acquired the local dialect naturally when they had moved to this area at a young age even though they were not born there. From Table 3.14, it can be seen that 11 participants did not meet one of these two criteria and eight violated both. These eight participants had to be excluded because their parents were not native local dialect speakers and the participants could have been influenced by their dialects. Smith, Durham and Fortune (2007) mentioned that some language features in adult speech can be evidently found in their children's speech at early stage of their language acquisition. They also thought input from the primary caregiver plays a crucial role in this process. Parents as first teachers and mostly as the crucial caregivers in China, it cannot be ensured their children will not learn some language features from their parents. In addition, it could be that in China participants may learn and speak some their parents' dialects for easy and intelligible communication with their grandparents. Therefore, to guarantee no influence from other dialects/languages except Mandarin, the participants whose parents violated either criterion should have been excluded in theory, but if one parent spoke the local dialect as a native speaker and the participants and their parents were living in these dialect-dominant cities, this was considered to be sufficient. Moreover, according to the statistical analysis, the first dialect of these participants was the local dialect. Therefore, it was reasonable not to exclude the 11 participants having one parent not born and raised in the target cities. Now, 74 sets of data are valid so far.

Table 3.14 Summary of Q6 and Q7

Q6	A. Yes	Q7	A. Yes	Total	N	45
			D. Not born here but raised here	Total	N	7
			B. No	Total	N	7
			Total	N		59
	B. No	Q7	A. Yes	Total	N	3
			B. No	1		HM-2
				2		HM-10
				3		HH-2
				4		HH-12
				5		GM-12
				6		GM-15
				7		GH-11
				8		GU-14
				Total	N	8
			D. Not born here but raised here	Total	N	1
			Total	N		12
	D. Not born here but raised here	Q7	A. Yes	Total	N	6
			D. Not born here but raised here	Total	N	4
			B. No	Total	N	1
			Total	N		11
	Total	N				82

Q9 and Q10 concerned the dialects parents used with participants and the first dialect they used in their daily lives were also essential constraints. Therefore, those whose parents spoke non-target dialects with them in their daily lives and those who did not speak this as their first dialect were excluded. From Table 3.15, 71 participants met the criteria except GH-15, GH-5 and GH-9 whose parent spoke other dialects with them in their daily lives while they were growing up. Thus, these three were excluded. All participants passed the test in Q10 with all of them speaking the target dialects as their first dialect.

Table 3.15 Summary of Q9

Q9Mo	A. Harbin Dialect	Q9MoC	Q9Fa	A. Harbin Dialect	Q9FaC		Total N		36
						Total	N		36
				Total	N				36
	B. Guangzhou Dialect	Q9MoC	Q9Fa	B. Guangzhou Dialect	Q9FaC		Total N		35
						+3Hakka sometimes	1		GH-15
							Total	N	1
						Total	N		36
				None of the above	Q9FaC	Mandarin	1		GH-5
							2		GH-9
							Total	N	2
						Total	N		2
				Total	N				38
	Total	N							74

Q11 was asked to exclude those who could speak other dialects and Q12 to see whether the target dialects were mainly used at home. If participants could speak another dialect or they spoke both at home and in their daily lives, this would increase the influence from non-target dialects. However, as mentioned, it is not realistic for Chinese, and especially youngsters to not speak Mandarin. Therefore, if the responses to Q11 mentioned any other dialect except Mandarin, those data were excluded. If participants spoke more than one dialect/language at home, it would mean they were bilingual or multilingual. They should also be excluded. According to the results in Table 3.16, three participants did not meet the criterion for Q11 and had to be excluded. They were HU-15, who understood the Shanghai dialect, HH-13, who could not speak any other Chinese dialect besides Harbinese but could speak the Sibe or Xibo language[7], and GM-5, who was a Cantonese speaker but could speak the Northeastern dialect. From

[7] Sibe (Xibo) language: is a Tungusic language spoken by the Sibe people ethnic minority Chinese, living in Xinjiang and the Northeastern part of China.

Table 3.17, data from GH-13 and GU-13 were deleted because they were bilingual, speaking both Cantonese and Mandarin at home.

Table 3.16 Summary of Q11

Hometown	Harbin	Q11	Understand SHD but can't speak	1		HU-15
				Total	N	1
			No	Total	N	32
			No, but XiboLan	1		HH-13
				Total	N	1
			Not really	Total	N	1
			Yes, M-16ys	Total	N	1
			Total	N		36
	Guangzhou	Q11	No	Total	N	5
			No but M	Total	N	1
			NortheasternD for 1w	1		GM-5
				Total	N	1
			M-8ys	Total	N	1
			M-9ys	Total	N	1
			M-14ys	Total	N	4
			No，M-9ys	Total	N	1
			M-11ys	Total	N	3
			M-10ys	Total	N	3
			M-12ys	Total	N	2
			No，M-10ys	Total	N	1
			No, M-15ys	Total	N	4
			No，M-16ys	Total	N	3
			No，M-14ys	Total	N	1
			No，M-12ys	Total	N	1
			No，CanAccentM	Total	N	1
			M	Total	N	1
			Total	N		35
	Total	N				71

Table 3.17 Summary of Q12

Hometown	Harbin	Q12	Harbinese	Total	N	25
			HarbinNortheastD	Total	N	3
			M	Total	N	4
			M(Harbinese)	Total	N	2
			Total	N		34
	Guangzhou	Q12	GZC	Total	N	24
			CAN	Total	N	8
			CAN-FA-2hs,M-MO-2hs	1		GH-13
				Total	N	1
			GZC+M	1		GU-13
				Total	N	1
			Total	N		34
	Total	N				68

Eventually, 66 participants, presented in the Table 3.18 below, were judged to be appropriate participants for further analysis of their recorded speech data.

Table 3.18 Participant information

School Years	Harbin Group				Guangzhou Group			
	Number	Age range	Age mean	Std. Deviation	Number	Age range	Age mean	Std. Deviation
Year 9	12	14-16	14.42	0.669	11	14-15	14.18	0.405
Year 11	12	16-17	16.33	0.492	10	16-17	16.60	0.516
Year 3 in Uni	10	20-22	20.80	0.632	11	20-21	20.18	0.405
Total	34	14-22	16.97	2.702	32	14-21	17.00	2.578

3.6.2 Analysis of sound recordings

Prior to the beginning of the sound analysis, a series of editing operations was conducted. The raw sound recordings of the 66 participants who met the criteria were listened to and edited with great care one by one. For subsequent convenience in the sound analysis, the irrelevant parts of each recording were edited out; for example, of interruptions caused by the bell at recess, the entrance of teachers, the giving of instructions before or during data collection, or hints given to help participants produce the correct tokens. The resulting audio clips were then compiled together into a single complete recording for each participant. Afterwards, according to the purposes of the word translation, Chinese reading and carrier sentence instruments and the different methods used in the analysis of data, the recordings were cut and classified in four sections to make the sound analysis easier. Tokens numbered from 1 to 162 were into the first section analysed to answer hypotheses relevant to segments and syllable structure named according to their original code plus the number 1. For example, HM3-1 which means the data for analysis of segments and syllable structure from the third participant from Harbin middle school. Tokens from 163 to 174 were used to analyse stress, and were labelled with the number 2 like HM3-2. The other 18 tokens from 175 to 192 were included in a sound clip filed with the number 3 to correspond to the Chinese reading task. The final part of the recordings formed the last section 4 for the carrier sentence. In the process of editing, one recording (GM11) was found to include background noise too loud to allow analysis; therefore, this data also had to be excluded. Ultimately, there were 65 participants whose sound production data was analysed, please see the new participant information in Table 3.19 below:

Table 3.19 Participant information

School Years	Harbin Group				Guangzhou Group			
	Number	Age range	Age mean	Std. Deviation	Number	Age range	Age mean	Std. Deviation
Year 9	12	14-16	14.42	0.669	11	14-15	14.20	0.422
Year 11	12	16-17	16.33	0.492	10	16-17	16.60	0.516
Year 3 in Uni	10	20-22	20.80	0.632	11	20-21	20.18	0.405
Total	34	14-22	16.97	2.702	32	14-21	17.10	2.561

Various methods were adopted in analysing the sound recordings. Firstly, all 174 tokens were transcribed in IPA seen in Appendix 3.4. The tokens with English and American pronunciations were represented in two versions but all were transcribed based on RP and GA. Although there were 174 tokens, some were used to test not only one point. For example, the third token *ship* was used to test the consonant segment /ʃ/ and the vowel segment /ɪ/. Thus, the number of items tested is 187 in all.

In theory, acoustic analysis would be the most scientific and appropriate tactic in the present research. However, it is impossible to process such an enormous amount of data in this way in a limited time. Also, phoneticians said it is impossible to use acoustic method to analyse such a large-scale study with so many participants and so much data. Therefore, auditory analysis was employed as the main approach in this study plus acoustic analysis using Praat[8] for part of the data. In addition, an RP English native speaker with linguistic knowledge was recruited as an assistant to guarantee the quality and accuracy of transcriptions. The learners were grouped into three levels for data collection as mentioned above, but it has to be mentioned here that all data were grouped together to represent the results of each target and level (i.e. amount of English exposure) was not treated as an independent variable.

3.6.2.1 Analysis of sound recordings for consonants, vowels and syllable structure

All recordings in section 1 were transcribed on a sheet using auditory transcription. Some tested segments could not be easily transcribed using the auditory method, such as /θ/ in the token *think* (see Figure 3.1 and Figure 3.2 below), which was pronounced as [s], [f] or [θ] by participants. In these cases Praat was used to identify what had been produced by segmenting that specific part of the word to determine acoustic features such as waveforms or spectrograms, or to plot a vowel by measuring F1 (the formant) and F2 (the second formant).

[8] Praat: is sound analysis software designed and developed by Paul Boersma and David Weenink in 1995 (2018).

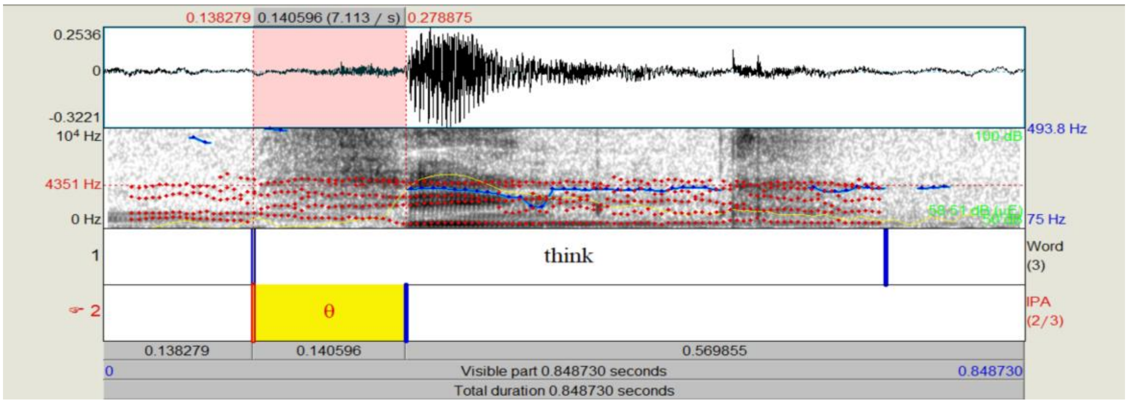


Figure 3.1 Acoustic spectrographic analysis of [θ] of 'think' in Praat

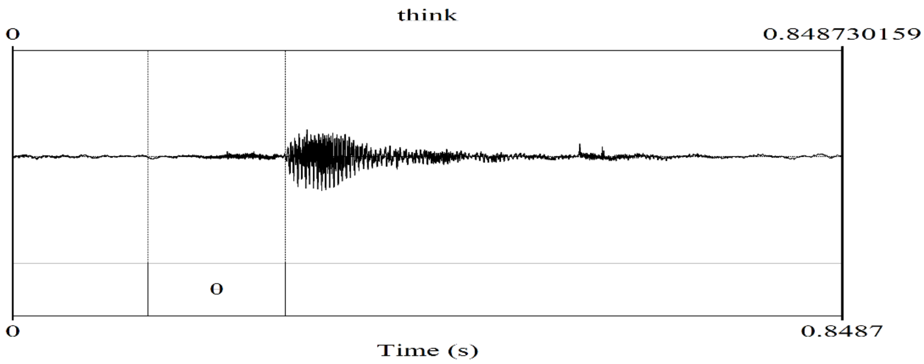


Figure 3.2 Waveform of the labelled [θ] of 'think' in Praat

Meanwhile, tokens with the tested vowel segments were annotated in TextGrid with four labelled tiers giving the name of words, IPA, target and actual production for acoustic analysis to test the hypotheses. An example of [ɪ] in the token 'ship' is shown in Figure 3.3 below.

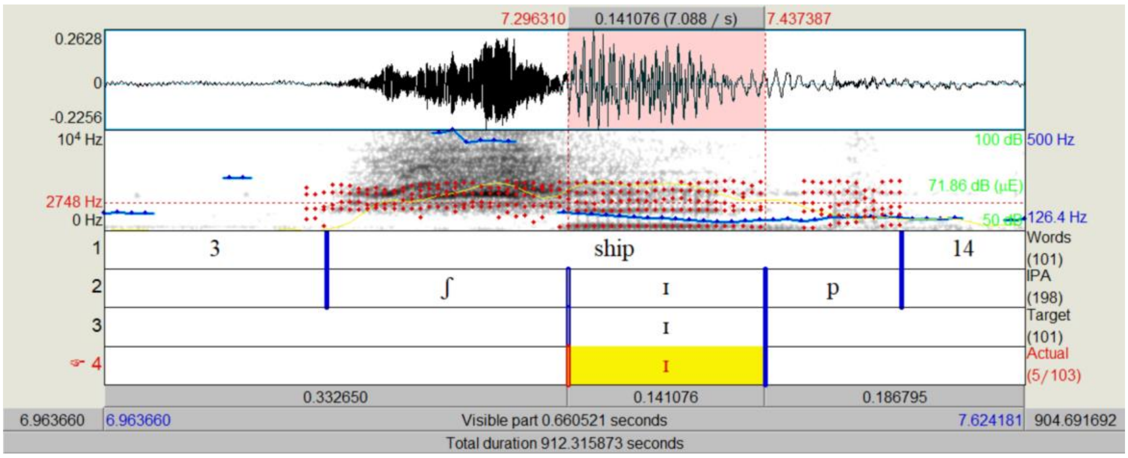


Figure 3.3 Acoustic spectrographic analysis of [ɪ] of 'ship' in Praat

When auditory transcriptions and vowel segmentations were completed, the transcriptions of speech produced in section 1 were input into Excel to form a dataset with the sequential order of tokens. Two targets in one token were each coded with its sequential number and sequential number plus a decimal for easy location in the transcription sheet, such as 2 and 2.1 for the two targets /w/ and /ɛ/ in the second token *wet*. Subsequently, the types of realisation of each targeted segment and syllable structure were summarised and their corresponding percentages calculated to see what different realisations there were and their frequencies of production among the two different groups of dialect speakers when they pronounced the tokens in English.

According to different research questions, different statistical calculations were conducted and distinct charts were plotted to test the research hypotheses H1-H6 by comparing the ratios of response types in the Harbinese and Cantonese groups.

In addition, small sets of representative data were transcribed by an assistant, who was a native RP English speaker, whose own transcriptions were then compared with mine to check their accuracy. Representatives from each level of participants were selected by comparing whether the assistant's average accuracy rate approached the average accuracy rates for that level because these accuracy rates can more typically indicate the real results of level. After calculations and comparisons, six participants seen in Table 3.20 were selected as representatives of their levels.

Table 3.20 Representatives selected from each level

Group Level	Group Medial Accuracy Rate	Group Average Accuracy Rate	Representatives	Rep's Average Accuracy Rate
HM	64%	62%	HM-7	63%
HH	71%	69%	HH-14	70%
HU	67%	66%	HU-3	66%
GM	73%	71%	GM-14	72%
GH	76%	74%	GH-10	74%
GU	73%	70%	GU-10	71%

After the completion of transcriptions by the native English assistant, her transcriptions and mine were statistically compared using an excel analysis of repetition rate. The higher the agreement rate, the more reliable the transcription is. According to the literature, it is acceptable if the degree of interjudge transcription agreement is greater than 85% (Pye, Wilcox, and Siren, 1988:19; Shriberg and Lof, 1991:25). The average degree of transcription agreement between my assistant and me reached 90.5%, which indicates that my auditory transcriptions are reliable. To pay utmost attention to the quality of transcription, given the small proportion of difference between our transcriptions, they were listened to many times and Praat was frequently used to check them to give final results. For example, the token /ð/ in [wɪð] *with* /ð/ was pronounced [θ] and transcribed as [θ] in my transcription but was transcribed as [θf] by the assistant. Therefore, it was checked with care using Praat to see if there was a labiodental fricative [f] sound after the dental fricative [θ]. As shown in Figure 3.4, no fricative [f] sound was detected in the spectrograph and waveform. Thus, the problem of this difference in transcriptions was solved and the final result was declared as [θ] for the correct transcription of this target consonant /ð/ in the token *with* rather than [θf] as transcribed by the assistant.

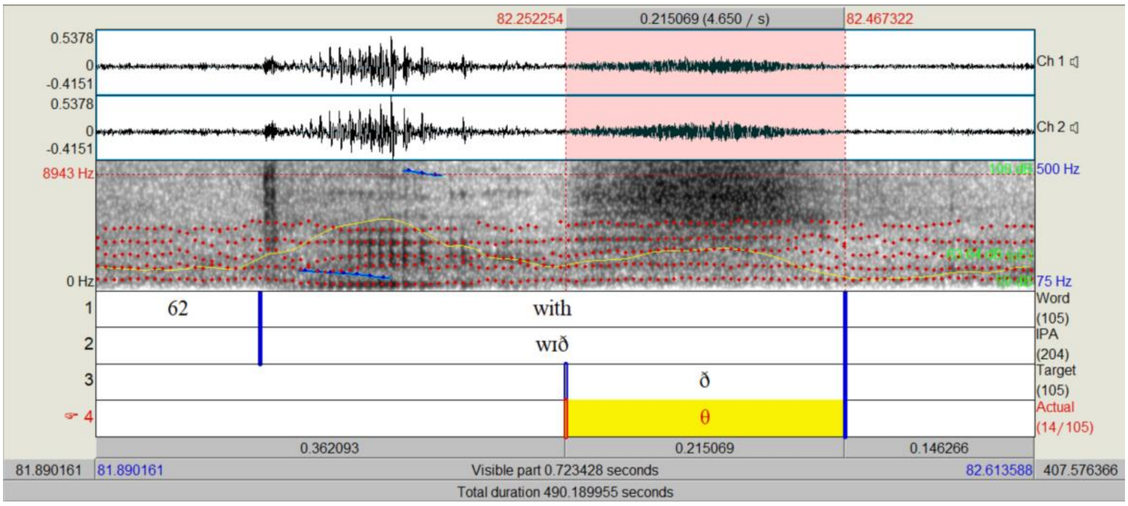


Figure 3.4 Acoustic spectrographic analysis of /ð/ of 'with' in Praat

3.6.2.2 Analysis of sound recordings for stress

As mentioned in the literature, English lexical stress is associated with many acoustic correlates, including pitch (or fundamental frequency f0), intensity, duration, and vowel quality, among which pitch is the most significant phonetic parameter (Duanmu, 2000:144). In theory, all these relevant acoustic cues should be measured; however, the tokens to test stress patterns were not designed with controlled vowels and surroundings. Therefore, it was not possible to compare the quality and duration of different vowels or the same vowels with different prevocalic and postvocalic consonants which may influence vowel duration and quality. Thus, these two acoustic correlates were not used to measure the stress patterns. Intensity as an acoustic cue for stress is debatable, but is still regarded as a possible correlate by some researchers (Fry, 1955; Beckman, 1986). Average intensity was measured as Zhang, Nissen and Francis did (2008:4499), but was not considered as a major parameter compared with pitch (fundamental frequency f0) in this research because the sound data was not collected in a quiet laboratory but in schools with uncontrolled background noise which would influence values of intensity. The average f0 and intensity within syllables were both measured by Praat in Hz and dB respectively and the pitch range was set to 75-300 Hz for male speakers and 100-500 Hz for female speakers, as mentioned by Zhang et al. (2008:4503).

All the recordings in section 2 were analysed acoustically and labelled for syllables with a boundary as a transition to separate disyllabic tokens into two syllables named S1 and S2. To guarantee that accurate values of fundamental frequency were obtained, the pitch contours had to be checked and modified and two principles had to be noted when modifying them as mentioned by Xiong (2004:44, 47). He thought that, firstly, the trajectory of the variation in pitch contour for the human voice is smooth, and any sudden large-scale fluctuations and jumps may lead to erroneous values of fundamental frequency. Secondly, alternation in voicing, such as from a voiced segment to an unvoiced one within a very short period (less than 40 milliseconds) generally indicate that the f0 value is wrong because alternations cannot be made within such a short time. Therefore, the pitch contour of each token was checked with care and those with irregular curves were manipulated using PitchTier in Praat to conform to the rules of sound movement.

Two Praat scripts were exploited to extract the value of the average f0 and average intensity. One script was borrowed from Xiong (2004:157-160) and adapted to extract the average f0 and the other was from Kawahara (2010) to extract average intensity. These scripts can be seen in Appendix 4.1 and Appendix 5.1. All recordings were annotated in TextGrid and PitchTier with the sound files to meet the requirements for the extraction of average values of f0 and the intensity of each syllable in the tokens. All the results for average f0 and intensity were obtained in txt format, which was converted into an Excel dataset. Subsequently, the values of average f0 and the intensity of the first syllable in a token were compared with those of the second syllable in the same token to see which had a higher pitch and a stronger intensity. If values of pitch and intensity for syllable one were both greater than those for syllable two, syllable one was stressed, and vice versa. However, if the values were in contradiction, that would mean that syllable one was higher in pitch than syllable two but less strong in intensity. As mentioned above, the significant acoustic cue for stress is pitch (f0), which is taken as the main criterion in assessing the stress placement of the tokens. Afterwards, the percentage rate of stress placement on syllables one and two was calculated to test hypothesis H4.

3.6.2.3 Analysis of sound recordings for Chinese reading

This analysis was used to confirm the observation that phonemes such as /n, l, w/ were treated differently by participants using the two first dialects.

Firstly, the Chinese tokens 165 to 192 were transcribed in IPA in Harbinese and Cantonese versions respectively as criteria, by then the third set of sound recording transcriptions were analysed and acoustically checked.

After the task of transcription was completed, a new Excel file was created with input from all of the transcription results. The accuracy rates of the consonants /n, l, w/ in each dialect/language group were calculated to see if participants treated these consonants differently according to their dialects.

3.6.2.4 Analysis of sound recordings for carrier sentences

This analysis was designed to examine whether or not the production of isolated words in section one was the same in connected speech. A small set of data was analysed to detect if the participants produced words differently in different environments. Typical data included in this small set of data were the same as those selected and transcribed by the English native assistant. All tokens supplied by participants in the carrier sentences were selected from the tokens in the translation section for consonant segments and syllable structure. If the transcriptions of the targeted tokens in carrier sentences and in section one are the same, this indicates that words are not produced differently in isolation and in connected speech. To enhance the reliability of this analysis, this comparison was conducted by the same transcriber, the RP English native speaker. Also, it was conducted soon after the completion of the transcriptions of the first section in order to maintain her consistent judgement conditions.

The transcriptions of carrier sentences and their counterparts in section one data were input into an Excel spreadsheet to compare the two. Agreement rates were calculated to determine if the phonological environment affected pronunciation. After the calculations, the average transcription agreement rate was 91.2%. Table 3.21 shows that there were slight differences in transcription between carrier sentences and their section one data counterparts, but these differences might be due to by the transcriber's inconsistent judgements due to tiredness or physical discomfort, or might result from incorrect pronunciations or from the different actual pronunciations, or might be due to all these reasons combined. However, as said in the literature, it would be unrealistic to expect 100 per cent agreement and it is generally unnecessary to achieve this. Because there is no significant difference between pronunciations in isolation and in connected speech, the part three data for carrier sentences did not need to be analysed.

Table 3.21 Transcription agreement rates for carrier sentence

	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Mean
Same	89.8%	95.4%	88.9%	90.7%	90%	93%	91.2%
Different	10.2%	4.6%	11.1%	9.3%	10%	7%	8.8%

3.6.3 Summary of questionnaires on language background with 65 participants

The full results from the analysis questionnaires data from all 65 participants along with the question titles and multiple choice responses as well as the frequency and percentage rates of answers for each group and overall are listed in Appendix 6.1 for details.

The first three items in Appendix 6.1 indicate the participants' basic demographic information such as gender, age and hometown. It can be seen that 34 participants came from Harbin and 31 from Guangzhou. No participant chose 'others', which means that all these participants are valid for research. Also, the total number of participants is 65 which exactly matches the number mentioned above. It can also be seen that in both groups there are more female participants, with 21 female in the Harbin group and 18 from Guangzhou. The gender ratio between males and females for all participants is 40 : 60 per cent. This imbalance in gender distribution is not an issue because there is no involvement in the sociolinguistic research in this thesis. The ages of participants range from 14 to 22 years old across the three levels. For the middle school level, both groups have more participants aged 14 years, accounting for 25 per cent of the total. The ages of high school participants are between 16 and 17, but Harbin has more participants at age 16 while Guangzhou has more participants aged 17. Six out of ten Harbin participants at university level are aged 21 while 9 out of 11 from Guangzhou are aged 20. Most of the Guangzhou university participants are one year younger than their counterparts in the Harbin group.

Question 1 was used to assess the participants' educational level and to confirm that they were appropriate participants in this research. The educational level options in Q1 range from primary school to doctoral postgraduate. My research only covers the three levels of middle and high school and university. From the responses to Q1, we can see that the participants came from the targeted three educational levels: 22 from middle schools and 22 from high schools. Also, these 22 participants included 12 Harbinese participants and 10 Cantonese participants from years 9 and 11 of middle and high schools respectively. There were 21 participants with 10 participants from year 3 at a Harbin university and 11 from the same year at a Guangzhou university. The proportions of participants at different levels were 34%, 34% and 32%.

All of the responses to the questions 2 and 3 were 'yes', showing that all were born and raised in the targeted cities and had spent most of their lives here. Q4 was designed to discover how long they had lived outside of their hometowns if they had not stayed there for some time. The results indicate that all participants in the research had stayed in their respective hometowns without leaving them for even a certain period such as one to three years. Q5 was designed to identify if they had not live in their hometowns in the most recent three years if participants answered they lived outside of their hometowns for a while in Q4. This is because the influence of a recent period would be greater. Fortunately, the results show that all participants had lived in their hometowns for the last three years. The results from Q1 to Q5 demonstrate that these participants do not violate the constraint imposed and that they were valid participants so far.

To avoid the influence of their parents, Q6-Q7 were used to gather the basic information of their parents' backgrounds, such as if they were born and raised in the targeted dialect speaking areas. The four choices for these questions are respectively 'yes' or 'no' and 'born here but not raised here' and 'not born here but raised here'. Those participants who answered 'yes' and 'not born here but raised here' were still acceptable participants. A person raised in a place from a young age will still likely acquire the locally dominant dialect even if s/he was not born there (Smith, Durham and Richards, 2013; Fischer, 1958). However, those who answered 'no' (born elsewhere) or 'born here but raised elsewhere' were excluded because their parents were not the local dialect speakers. Although they would have acquireed the local dialect if they lived in an area for long enough, the dialect they acquired would not be as native as local speakers. In addition, they may only use the acquired dialect at work but use their own L1 dialects at home, which would influence the pronunciation of their children (i.e. the participants in the study). Another reason to exclude these participants is that it is also the fact that in China their parents may teach them to speak the parents' dialects in case they can communicate with their grandparents. To ensure that there is no any impact, as mentioned in above section 3.6.1, those who chose 'no' or 'born here but not raised here' were excluded. From the results, it can be seen that 31 fathers and 29 mothers in the Harbin group were born and raised in Harbin while 23 fathers and 19 mothers were born and raised in Guangzhou, and thus those participants and their parents were definitely native local dialect speakers. However, the Guangzhou group included fewer whose parents were born and raised in the targeted native dialect speaking areas than the Harbin group. Two Harbin participants had fathers not born in Harbin but raised there, while seven Guangzhou participants fitted the same pattern. Both groups had five mothers not born but raised in the targeted cities. These participants also met the requirements. However, the results show that each group had one father who was not born and raised there and the Guangzhou group had seven mothers not born and raised in Guangzhou. In theory, participants with these parents should have been excluded to avoid the influence of their linguistic background, but it was considered acceptable to retain them in the sample because they all had one parent speaking the native dialect, and they and their parents lived in the dialect-dominant area with the dominant dialects as their first. As seen in the table, participants answering 'no' account for a small proportion of the whole, with 3% for fathers and 11% for mothers. Also, Guangzhou group had more participants than the Harbin group who had one parent not speaking the local native dialect.

Question 8 to 14 were used to understand the situation with the language or dialect use of the participants and their parents. Q8 was used to discover which dialects the participants' parents spoke. The choices provided were 'Harbinese', 'Guangzhou

Cantonese', and 'none of the above'. From the results, all Harbinese participants' parents spoke Harbinese, but 29 mothers (94%) and 30 fathers (97%) out of the respective 31 parents spoke Guangzhou Cantonese, which means that 2 mothers and 1 father from the Cantonese group spoke other dialects. The Cantonese participants would be influenced if their parents spoke other dialects, but Q8 was not as essential as Q9 and Q11. Because whether or not the participants would be influenced depends on whether their parents spoke their own dialects with them in their daily lives. Although their parents may speak other dialects, if the participants were not exposed to the dialects with their parents, the participants would not be influenced. Therefore, compared to Q9 and Q11, Q8 was not essential. Q9 was used to determine which dialects the participants' parents used at home with them when they were growing up. From the results, we can see that all mothers and fathers from each group used the targeted dialects with their children (the participants in this research) during their childhood. This demonstrates that the participants are appropriate participants so far.

From the above results, it can be noted that there would be no influence from the parents on the participants to speak non-targeted dialects. However, it is still necessary to know what dialects their parents' first dialects were. Thus, Q10 was asked. The results show that all participants from the Harbin and Guangzhou groups spoke their own dialects as first dialects. Q10 also asked Harbinese participants for their ideas on the relationship between Harbinese and Mandarin. In fact, Harbinese is a variety of the Northeastern dialect which is one of the varieties of the Mandarin dialect group, as mentioned in Chapter 2. Due to the closeness of Harbinese to standard Mandarin, Harbinese speakers may have been confused and thought that what they spoke was just standard Mandarin and they did not realise there were differences. The results under choice A for Harbinese in Q10 confirms this. A majority of Harbin participants (30 out of 34, 88%) thought that Harbinese was the same as standard Mandarin. Three (9%) thought Harbinese and standard Mandarin were different and one participant thought they were quite similar. However, even though the Harbin participants had different perceptions of the relationship between Harbinese and standard Mandarin, this did not affect their participation in the research into English phonology acquisition.

Q11 was an open question given to participants to learn if they could speak any other dialects apart from their first dialects. As mentioned earlier, it is unrealistic to expect that Chinese people who were born and educated in China would not speak and understand the official language of Mandarin. Therefore, those who gave a response of “Mandarin” would not be excluded, but only those who could speak other dialects would be excluded, in order to guarantee that the participants were not being influenced by other dialects. The responses are summarised in Q11 in Appendix 6.1. Almost all (97%) of Harbin participants answered ‘no’, that is, they could not speak any other dialects besides Harbinese, but one participant answered ‘yes, Mandarin’ and that he had learned it for 16 years. This reflected the different understandings of the relationship between Harbinese and Mandarin. From the summary of results in Q11, we can see that 15 Guangzhou participants (48%) answered ‘yes’ and 16 (52%) answered ‘no’. Even if they answered ‘yes’, they noted that the other dialect they could speak was Mandarin but no others. And if they answered ‘no’, they also mentioned that they could not speak other dialects except Mandarin. It is observed that although they adopted two different strategies to answer the question, the meanings expressed were the same; that none of them could speak other dialects except official Mandarin. It is interesting to note that one Cantonese participant answered ‘no’, but emphasised that she could speak Mandarin with a Cantonese accent. Thus Cantonese was his/her first dialect and L1 transfer occurred when speaking Mandarin. The results also show that the duration of learning Mandarin for Cantonese participants varied from 8 to 16 years. These subjects are all acceptable participants who violated no constraints.

Q12 was asked to find out if participants were bilingual speakers at home or single dialect speakers. Although it was impossible to eliminate participants who can speak Mandarin, its influence can be reduced to the minimum if they at least prefer to use their first dialect at home. From the results, 74% of participants in each group responded that they spoke their first dialects at home. Meanwhile 26% of Cantonese participants said that they used ‘Cantonese’. This implies that all Cantonese participants spoke Cantonese as their first dialect at home despite using the general name ‘Cantonese’ instead of the specific ‘Guangzhou Cantonese’. The Harbin participants answered that they spoke the ‘Harbin Northeastern dialect’ (9%), ‘Mandarin’ (12%), and ‘Harbinese Mandarin’ (6%)[9]. As someone speaking that dialect, I understand that, in fact, they meant to express the same thing, but used different terminology due to the unclear understanding of the relationship between Harbinese and Mandarin which could be confirmed according to the answers to questions Q10 and Q11. However, it is also true to say that Harbinese is the Harbin Northeastern dialect and it is indeed a variety of the Mandarin dialect group. Therefore, it can be concluded that all participants from both groups did use the targeted dialects as their first dialects at home.

Q13 was asked to determine for how long each day on average they used their first dialects. The answers range from a minimum of two hours to a maximum of 24 hours per day. The largest proportion of participants at 15% used their first dialects for eight hours per day. For the Harbin group, 10 hours and 12 hours per day were spent speaking Habinese by 12% of participants respectively, and most of the Harbin participants spoke Harbinese for less than 12 hours a day. However, 3 participants (9%) responded with 24 hours per day, which seems unreasonable because a person cannot continue speaking throughout the whole day. It is likely that they meant that they always used Harbinese in their daily lives whenever they needed to speak. Unlike the Harbin group whose most popular average length of first dialect use was 12 hours or less, participants in the Guangzhou group spoke their first dialect Cantonese less (8 hours for the largest proportion of 23% of Cantonese participants). In the Guangzhou group, the maximum duration of first dialect use per day was 14 hours for only one participant. From these results, the time participants spent daily in speaking their first dialects differs, with Harbin participants using their first dialect for longer per day than did Guangzhou participants.

[9] 74% of participants in each group responded in this way, meaning that the remaining participants should make up 26% each. However, the sum for these three items is 27% due to rounding up.

Q14 was asked to discover how long the participants would spend using other dialects each day. From the results for Q11, no participants spoke any other dialects or languages apart from their first dialect and Mandarin. Therefore, the answers to Q14 indicate the time they spoke Mandarin. 82% of the Harbin participants did not think they could speak any other dialects, so might think that Mandarin and Harbinese are the same. Other Harbin participants spoke Mandarin for between 7 and 12 hours daily. Guangzhou participants spoke Mandarin for 0 to 10 hours per day (less than Harbinese speakers), with the largest proportion among Cantonese speakers (23%) for 6 hours.

The results from Q8 to Q14 on language and dialect use demonstrate that all participants met the strict requirements used in the research to avoid the influence of other dialects.

Q15 to Q23 were designed to probe the participants' language learning history including Mandarin and English, and experience of learning English with English native speakers.

Q15 was asked to identify when participants started to learn Mandarin. This question was particularly aimed at Cantonese participants, because Mandarin is very different from Cantonese and it would represent an L2 dialect. The results in Q15 indicate that Harbin participants started to learn Mandarin much earlier than their Guangzhou counterparts. Almost all Harbin participants (97%) had started to learn Mandarin before starting school. Among these, 50% started aged one and most of the rest began learning it before the age of three. This implies that the Mandarin in their minds was the same as the first Chinese dialect they started to learn, which was Harbinese. Meanwhile participants from Guangzhou started to learn Mandarin much later. Most (52%) started to learn it at primary school. The others started in the preschool period, but most of them only started to learn Mandarin after the age of three.

Q16 was asked to discover when they started learning English. It seems that this was much later than beginning to learn Mandarin. Most participants (80%) started to learn English at primary school, including 74% of the Harbin group and 87% of the Guangzhou group. Among these, 38% of Harbin participants started learning English at the age of six and 21% were aged seven. Ages of six and nine were the most common among Guangzhou participants for starting to learn English (45% and 32% respectively). The Harbin participants started learning English in preschool, which was earlier than the Guangzhou group, with the earliest at the age of three but also at four years old for one Guangzhou participant. Participants who started to learn English in preschool mostly started at the age of five in both groups. In the Harbin group, one participant only started to learn English in middle school at the age of 12, but no participants in the Guangzhou group started studying English later than the age of ten. It can be seen from the results that the span for Harbin participants starting to learn English is wider than that for Guangzhou participants. Also, the Harbin group mostly started to learn English at six or seven but in Guangzhou at six or nine.

Q17 asked about the duration of participants' English learning. Their responses varied from 7 years to 15 years across the three levels. A majority had had English education for 9 years or more (97% from Harbin and 95% from Guangzhou), but between 9 and 11 years included a large proportion of Harbin participants while 9-12 years applied to a big percentage of the Guangzhou group.

Q18 was asked to determine if participants could speak any languages besides Chinese and English in case they would be influenced by other languages. From the results, 94% of Harbin participants and 87% of Guangzhou participants had never learned any other languages and could not speak them, but two participants from Harbin had respectively learned a little Japanese and Russian at the age of 15 and a little Korean at 14 while two from Guangzhou had learned a little Spanish and Japanese at 11 years old and above for a couple of months and another two had studied French and German for 1-2 years. These participants were not considered invalid because they had not studied these L3 or L4 languages before puberty and anyway did not study them for long.

3.7 Validity and reliability

Validity and reliability have to be assured because they are essential measures of the quality of research and the accuracy of scientific experiments. Before the data was collected, all aspects of validity and reliability were taken into consideration. For example, the choice of tokens was carefully considered to suit the English ability of participants at low levels from both regions. Also, schools were chosen prudently so that they were comparable in the two regions to reduce any discrepancies caused by economic and geographical differences. Strict constraints such as the hometown and main dialect use were set to guarantee the validity of the findings.

Various research methods and strategies were adopted in this research to guarantee validity and reliability. Not only were the tokens transcribed in IPA (see Appendix 3.4), but also both auditory analysis and acoustic analysis were employed to analyse the sound recordings (see 3.6.2 Analysis of sound recordings). To further increase the credibility of the results, a native RP English speaker was recruited to conduct spot-checks of transcriptions. Scripts, calculations and statistical data were also repeatedly tested and checked. To summarise, validity and reliability were maximally guaranteed to the best of my knowledge in this research.

3.8 Summary

This chapter introduced the research hypotheses for this thesis and described how methods were designed to test them. It then detailed the procedures by which participants were recruited and how the data was collected and analysed. The next chapter presents the results.

Chapter 4. Results

This chapter introduces results of the analysis of data, including questionnaire responses, the production of English segments, syllable structure, stress and the Chinese reading. The results are described in detail and demonstrated with figures and tables and used in the testing of the research hypotheses along with statistical analysis to ground their support or rejection.

4.1 Results of the partial questionnaires

I assumed that the participants in my sample would have been exposed to both British English and North American English, which may include native speakers of English. Therefore in the questionnaire survey ten questions were asked regarding the English varieties to which they could have been exposed. These questions were part of the questionnaire which also served to include participants who were only speakers of Harbinese or Cantonese (see Chapter 3).

After statistical analysis using SPSS, the results related to English varieties and the methods used to learn English are illustrated below after a summary of the questions with their response options. Full details can be found in Appendix 6.1.

Q19 concerned if participants had experience of studying English with English native speakers. 65% of Harbin participants and 68% of Guangzhou participants had never studied English with native English speakers. The rest had had this experience including 12 Harbin and 10 Guangzhou participants.

Learning English with native speakers of English may influence the learners' pronunciation, but this also depends on other factors such as at what period participants started to learn with native English speakers, how long the participants learnt with them, if participants had a strong motivation to learn with them, and whether or not the native English teachers had a strong motivation to teach and correct the participants' pronunciation. Therefore, Q20 was designed to elicit more detailed information on the participants' learning of English with native English teachers. Participants were required to disclose at what age and for how long they had learned English with native English teachers. The results show that most participants with experience of learning English with native speakers were at primary school or secondary school, starting between the ages of seven to thirteen. The ages of Harbin participants studying English with native English teachers were mostly concentrated between 7 and 8 years old and between 12 and 13 years while the Cantonese were mostly between 9 and 11. Three participants had studied English with native speakers of English during the preschool period at between 4 and 6 years old. The duration of their study span with native English teachers varied from less than one year to up to eight years.

It is very difficult to determine if these participants were influenced by native English teachers. To take my own experience as an example, I have been learning English for over 20 years, since I was 12, and spent four years learning it as a major at university in China. Within that period and beyond, I attended oral English lessons with native English teachers for 2 or 3 years for one hour per week. Moreover, I had experience studying at an English programme in an English speaking country for more than two years. However, my English pronunciation started to change slightly only after I started to study linguistics, when my supervisor pointed out my issues with pronunciation. Since then, I have actively corrected my pronunciation with strong awareness and motivation in recent years.

Q21 was asked to discover information about their native English teachers and the participants' learning history with them. During the learning of English, participants may have different native English teachers with different English accents or because they come from different countries where different varieties of English are used. To assess the influence of native English teachers, it needs to be known where participants' native English teachers came from and the length of time participants learned English with them. In addition, whether or not participants were influenced by their native

English teachers could also be assessed by examining their production of specific English tokens which vary according to English accents. For example, if participants learned English with British English native teachers, these specific English tokens should be pronounced more like British English. Although no research questions and hypotheses were formulated for this variable, it is discussed together with the results and findings later.

The responses were classified according to the countries from which the participants' English native teachers came. It can be seen from the results that most of the native English speaking teachers came from the US (respectively 42% and 40% for the Harbin and Guangzhou groups), followed by Canada (25%) and the UK (8%) for the Harbin group and the UK (20%) and New Zealand (10%) in the Guangzhou group. Two Harbin participants (17%) had learned English with more than one English native teacher from different countries. In addition, one participant in each group mentioned that s/he could not remember where his/her English teacher came from. The Canadians speak English very similar to general American English compared with British English, so teachers from the US and Canada were regarded as speakers of GA generally. Therefore, it can be seen that both Harbin and Guangzhou participants had more exposure to the American variety of English with their English native teachers.

Their length of studying English with native speakers was then divided into six categories from less than six months up to one year, two years and three years as well as more than three years. The results show that 50% of Harbin participants studied English with native speakers for less than one year, 25% for two years and 17% for three years while 40% of Guangzhou participants had studied for between six months and one year, 30% for two years, and 20% for more than three years.

Q22 was asked to learn which teacher had the most influence on the participants if they had more than one native English teacher from different countries. It was also used to learn which variety of English those English teachers spoke to the participants. 25% of

Harbin participants and 20% of Guangzhou participants thought there was no influence from their English native teachers. Also, 17% of Harbinese and 30% of Cantonese participants did not mention if their English native teachers influenced them or not, or who influenced them the most. This seems to suggest that they were not much influenced by their native English teachers because they did not take the question seriously. However, Harbin participants thought the teachers from Canada were most influential, accounting for 25%, followed by teachers from the US (17%) and the UK (17%). It can be seen that US teachers had the most influence on the Guangzhou participants (20%), followed by teachers from the UK and Canada, at 10% each. It can be inferred from the results that they were more exposed to the American variety of English (42% in Harbin and 30% in Guangzhou) than the British variety of English (17% and 10% respectively in the two groups).

Q23 was designed to learn how participants learned English outside of class and how long they spent on it per week; for example, in watching English films for five hours. Methods of learning English outside the class adopted by participants include attending English training classes, watching English television programmes and films, reading English books or newspapers, listening to the radio or English music and memorising English words. Participants adopted many ways to learn English outside of school. From the results, it can be seen that watching English TV series and films was the most popular method used to learn English outside school education, for 53% of Harbin participants and up to 67% of Guangzhou participants, followed by listening to English music or radio by 33% and 48% of the Harbin and Guangzhou groups respectively. Attending extra English courses was another way for participants to learn English in their spare time with more Harbin participants (24%) learning English in this way than Guangzhou (13%). Reading English books or newspapers was used by the smallest proportions of participants (15% Harbin and 3% Guangzhou) to learn English. One participant (3%) in each group responded that their only exposure to English was learning at school.

Irrespective of the methods adopted, the duration participants spent weekly in extra English learning outside school was also calculated. The time participants spent ranged from zero to 11 hours. 24% of Harbin participants spent two hours every week and 26% three. 88% of Harbin participants spent less than four hours per week but only a small proportion of Harbin participants (12%) spent eight hours or ten hours weekly on this. The length Guangzhou participants spent on extra English learning per week varied more widely with two (19%), three (19%), four (16%), and five hours (13%) spent. Fewer Guangzhou participants spent above five hours per week but one spent up to 11 hours. The average length spent on learning English outside of class was calculated, and the results show that the Harbin group spent three hours on average per week while for the Guangzhou group an average of approximately 3.6 hours was spent weekly in this way.

From responses to Q15-23, the participants' language learning history could be mapped in detail, including information on their extra English learning outside school, and information about those who had studied English with native speakers as well as the latter's background information.

The last five questions in the questionnaire (Q24-Q28) were designed to discover the participants' preference for English varieties, their knowledge of them and their ability to distinguish between them.

Q24 asked the participants about their preferences for British or American TV series and films in order to determine which variety of English participants were most exposed to. The results indicate that up to 86% of Harbin and 68% of Guangzhou participants preferred American TV series and films. Therefore, it is more likely that participants were exposed to American accents, especially for the Harbinese.

Q25 was asked to find which variety of English participants preferred. The results are inconsistent with those for Q24. Harbin participants were split in half in preferring

American or British varieties of English. However, more Guangzhou participants (68%) preferred British English while preferring American TV series and films as indicated in Q24. It can be inferred from these results that there is no close correlation between preference for American or British media and preference for English variety.

Q26 was further asked to discover what variety of English participants thought they used. The choices provided for participants were ‘British, American, Chinglish and other types’; where Chinglish is a type with a strong influence of Chinese on English pronunciation. In the Harbin group, half of the participants thought they spoke English with a strong Chinese accent, followed by 35% speaking the American variety of English and 15% speaking the British variety. British English was spoken by 52% of Guangzhou participants, followed by Chinglish (29%) and American English (13%), as well as other types of English with 6%.

However, were they correct about the variety of English they thought they spoke? Q27 was designed to test if they could distinguish between varieties of English. Their responses indicate that they did not have a clear understanding of differences between English varieties: 79% of Harbin and 68% of Guangzhou participants answered ‘No’ to the question ‘Can you distinguish American English from British English in pronunciation?’ To those who answered ‘Yes’, Q28 asked how they distinguished these two English varieties. Their answers were categorised in three types: ‘can distinguish’, ‘know a little’, and ‘can’t distinguish’. Most participants could not distinguish between American and British English (86% of the Harbin group and 60% Guangzhou). Only one Harbin participant (14%) had a little knowledge of how to distinguish them, two Guangzhou participants (20%) knew a little and another two (20%) knew the differences.

4.2 Results of hypotheses relating to consonants

This section presents the outcomes of testing the hypotheses relating to consonants which are categorised as ‘identical’, ‘similar’ and ‘new’. Moreover, the results of statistical tests are presented to show if each hypothesis is supported or rejected.

The hypotheses are tested by analysing the data gained using the test instruments of word translation and Chinese reading applied with 65 participants as seen in Table 4.1. There were 34 appropriate participants from Harbin and 31 from Guangzhou. Data from these 65 appropriate participants is used not only to test hypotheses relating to consonants, but also for vowels, syllable structures, and stress as presented below. Hypotheses are tested by analysing the participants' production of target linguistic tokens.

Table 4.1 Number and ratio of participants from Harbin and Guangzhou

Hometown	Number	% of total
Harbin	34	52%
Guangzhou	31	48%

4.2.1 Category of 'identical' in consonants

4.2.1.1 H1.1 and results

> H1.1 The English consonant phoneme /h/ is identical in Cantonese but similar to /x/ in Harbinese, so Harbinese speakers will make errors.

The English phoneme /h/ never occurs in the final position of a word; therefore, only initial and medial positions were checked. The English glottal consonant /h/ exists in Cantonese but it is absent from Harbinese; however, Harbinese has the similar voiceless velar consonant /x/. Therefore, the Harbinese /x/ is phonemically similar to English /h/. It can be seen from Table 4.2 and Figure 4.1 that the voiceless glottal fricative /h/ was not difficult to learn for both the participants, especially when occurring in the medial position, with accuracy rates up to 97%. The Harbinese pronounced it slightly worse than the Cantonese, when /h/ occurred in the initial position. The only error type made for this consonant was substitution of /h/ by [x], which implies an influence of Harbinese or Mandarin, even on Cantonese speakers. It can be seen that Harbinese speakers exhibited more transfer than the Cantonese, but it represents the first language transfer for Harbinese speakers but not for Cantonese speakers.

Table 4.2 Production of English consonant /h/

Cons		Harbin		Guangzhou	
/h/	Type	Initial	Medial	Initial	Medial
	Correct	87%	97%	98%	98%
	Sub-x	13%	3%	2%	1%
	*Error[10]				1%

(Cons: consonant; Sub: substitution)

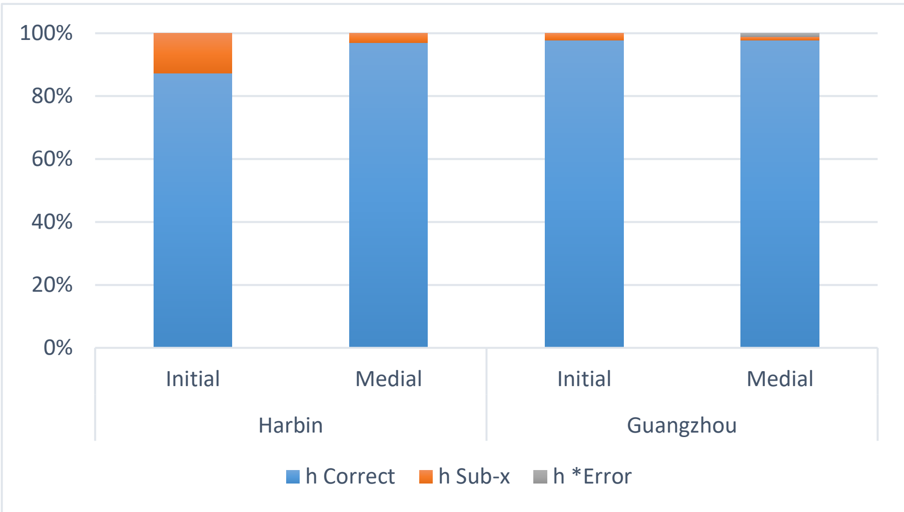


Figure 4.1 Production of English consonant /h/

Table 4.3 T-test results for English phoneme /h/ in Harbin and Guangzhou

Hometown		N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
h	Harbin	34	92.16	11.037	1.893	-2.248	0.028
	Guangzhou	31	97.85	9.371	1.683		

[10] Although great care was made to avoid mistakes such as participants pronouncing an unrelated word instead of the target token when recording, there were still some mistakes made like this in the data. Such mistakes are indicated by an asterisk symbol * with the word Error in all tables here and below.

From the above t-test results for the English consonant /h/, it can be seen that the accuracy rate for /h/ in the Cantonese group is higher than that in the Harbinese group although both are very high. Therefore, it seems that the hypothesis H1.1 is supported by the results in that Harbinese speakers make more errors. However, the disparities between the inaccuracy rates are slight. The hypothesis should be answered in a scientific way. Therefore, a statistical test was employed to support or reject the hypothesis 1.1 in a scientific result instead of the direct confirmation by the number of ratios of accuracy rates being taken for granted.

The statistical test Table 4.3 indicates that the average accuracy rate of /h/ in Guangzhou group is 97.85, which is also higher than their counterpart group (92.16). Moreover, the t-test table reveals that there is significant difference between two groups in the production of the English phoneme /h/ [p=0.028 (<0.05)] although both groups have achieved high accuracy rates.

From the results of the t-test of English phoneme /h/, it can be summarised that hypothesis H1.1 is supported by the results and even by the statistical test that Harbinese speakers made errors with English phoneme /h/ and the difference between the two dialect speaking groups is significant although the difference is not large. These findings confirm that 'identical' is easier. However, it does not show how difficult 'similar' is because the Harbinese group also achieves a high accuracy rate despite making some errors.

4.2.1.2 H1.2 and results

> H1.2 The English /w/ is identical in Cantonese but similar in Harbinese because Harbinese has /w/ with two allophones [w] and [v], so the English /w/ is phonemically identical but allophonically different in Harbinese. Thus, the Harbinese will make errors with /w/ .

In fact, /w/ exists in both Chinese dialects, but it is not completely identical to the English /w/ in Harbinese which has two allophones and speakers often mispronounce

/w/ as [v]. Again /w/ can only occur in the initial and medial positions in English words; thus, the pronunciation of /w/ in these two positions as checked.

From Table 4.4 and Figure 4.2, it can be seen that it was not a challenge for both participants to pronounce /w/ and both groups demonstrate high accuracy rates. However, the Cantonese group performance (approximately 90%) was much better in pronouncing English /w/ than the Harbinese (approximately 80%). Substitution was the only error type made. Table 4.4 indicates that Harbinese speakers preferred [v] as the substitute when replacing /w/. Cantonese speakers also used [v] to replace /w/ but at a rather lower percentage. Apart from this substitution, [ɹ] was also used by Cantonese speakers when /w/ occurred word-initially.

Table 4.4 Production of English consonant /w/

Cons		Harbin			Guangzhou		
	Type	Initial	Medial	Final	Initial	Medial	Final
	Correct	83%	81%		92%	96%	
/w/	Sub-v	16%	19%		5%	4%	
	Sub-ɹ				2%		
	*Error	1%					

(Cons: consonant; Sub: substitution)

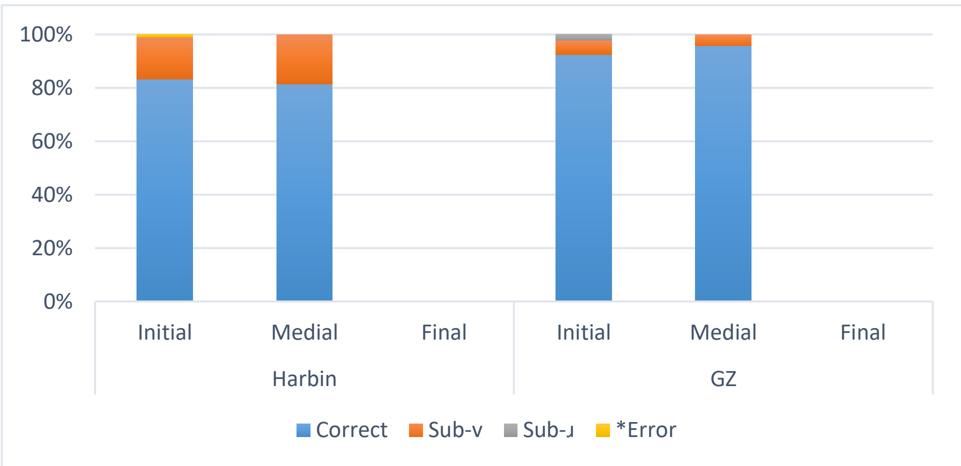


Figure 4.2 Production of English consonant /w/

The results indicate that Cantonese speakers performed better than Harbinese speakers in pronouncing /w/, which supports the hypothesis H1.2 that Harbinese speakers will make errors with /w/. In addition, the result of the statistical tests for /w/ calculated using the mean value in all positions also confirm that hypothesis H1.2 is supported because a significant difference between Harbinese and Guangzhou Cantonese speakers [p=0.008 (<0.05)] is shown in /w/ (see Table 4.5). However, it does not seem that the similarity with phonemically identical and allophonically different sounds leads to much difficulty for Harbinese speakers, because their accuracy rates were still high at approximately 80%.

Table 4.5 T-test results for English phoneme /w/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
w	Harbin	34	82.35	21.297	3.652	-2.777	0.008
	Guangzhou	31	94.09	11.823	2.124		

4.2.1.3 H1.3 and results

> H1.3 The English /n/ and /l/ [l] exist in Harbinese and Cantonese but are sometimes neutralised in Cantonese, thus the English /n/ and /l/ [l] are identical with Harbinese but only similar in Cantonese. Therefore, the Cantonese will make errors.

The consonant /n/ exists in both dialects/languages. As mentioned in the literature, Cantonese has cases of neutralisation between /n/ and /l/. From the Figure 4.3, we can see that neither groups of participants experienced serious production issues, especially in initial and medial positions. No mistakes were made in the medial position by any participant. Small errors were made when /n/ occurred initially, as shown in Table 4.6, where a /k/ sound was inserted in front of the initial /n/ by some Harbinese speakers but only in about 1% of cases. This mistake seems to be influenced by the orthography of the word ‘knife’. Therefore, it can be assumed that Harbinese speakers have no difficulty in pronouncing an initial /n/. However, the results confirm that the Cantonese may neutralise /n/ with /l/ , but this only constitutes a very small proportion of cases. /n/ was pronounced by 1% of Cantonese participants as [l] and by 2% as the nasalised lateral [l̃]. Accuracy rates decreased, however when /n/ occurred finally. Also, different errors types were made. In the Harbinese group, insertion and deletion were used, while for the Cantonese, substitution was still dominant, followed by deletion and insertion. 21% of Harbinese participants inserted a schwa [ə] in front of /n/, which conforms to the pronunciation of /n/ as [ən] in their first dialect. Another 15% preferred the deletion of the final /n/, which might result from the nasalisation of the preceding vowel and thus omitting the final nasal sound /n/. In the Cantonese group, 11% also deleted the final /n/, but the error type most often used was substitution and the substitute most selected was the velar nasal [ŋ] (12%), followed by the bilabial nasal [m] (4%), which never occurred in the Harbinese group. Another result was that a small proportion of Cantonese speakers inserted an alveolar stop after /n/ as in like [nt] and [nd].

Table 4.6 Production of English consonant /n/

Cons		Harbin			Guangzhou		
/n/	Types	Initial	Medial	Final	Initial	Medial	Final
	Correct	99%	100%	65%	97%	100%	71%
	Sub-l				1%		
	Sub-l̃				2%		
	Sub-m						4%
	Sub-ŋ						12%
	Ins-ən			21%			
	Ins-nd						1%
	Ins-nt						1%
	Ins-kn	1%					
	Del			15%			11%

(Cons: consonant; Sub: substitution; Ins: insertion; Del: deletion)

In summary, /n/ was not a great challenge for participants when it occurred initially and word-medially although some Cantonese speakers neutralised the initial /n/ with /l/. However, when /n/ occurred finally, both groups of participants had some problems. The insertion (of a schwa) was the main error among Harbinese speakers but it was substitution ([ŋ]) for Cantonese. Deletion was used by both groups as a second strategy to tackle the final /n/.

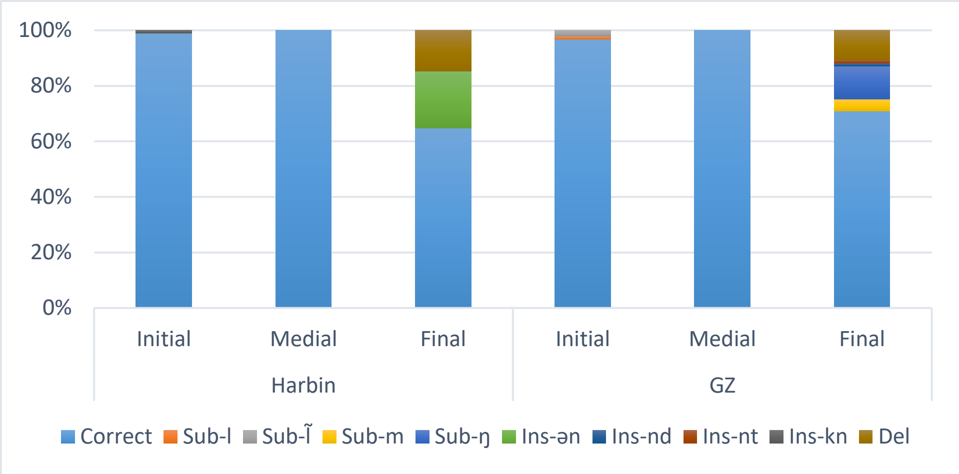


Figure 4.3 Production of English consonant /n/

The English /l/ has two allophones: one is the clear lateral [l] occurring in the initial position of a syllable and the other is dark lateral [ɫ] which often occurs in the coda. Both Chinese dialects/languages have [l] but not [ɫ]. A hypothesis was only proposed for the clear /l/ [l]. Due to neutralisation between /n/ and /l/, the /l/ was also checked. It can be seen from Figure 4.4 and Table 4.7 that /l/ can be pronounced perfectly when it occurs initially, especially among the Harbinese group who made no errors. A small proportion of the Cantonese mispronounced the initial /l/ as [n] or the nasalised /l/, that is, [l̃], which further confirms that Cantonese speakers occasionally neutralised /n/ and /l/. This means that /n/ and /l/ are substituted for each other. However, although Cantonese speakers made a small proportion of errors, generally speaking, there was no particular challenge for participants in pronouncing the initial /l/.

Table 4.7 Production of English consonant /l/ [l]

Cons		Harbin			Guangzhou		
/l/ [l]	Types	Initial	Medial	Final	Initial	Medial	Final
	Correct	100%			97%		
	Sub-l̃				2%		
	Sub-n				1%		

(Cons: consonant; Sub: substitution)

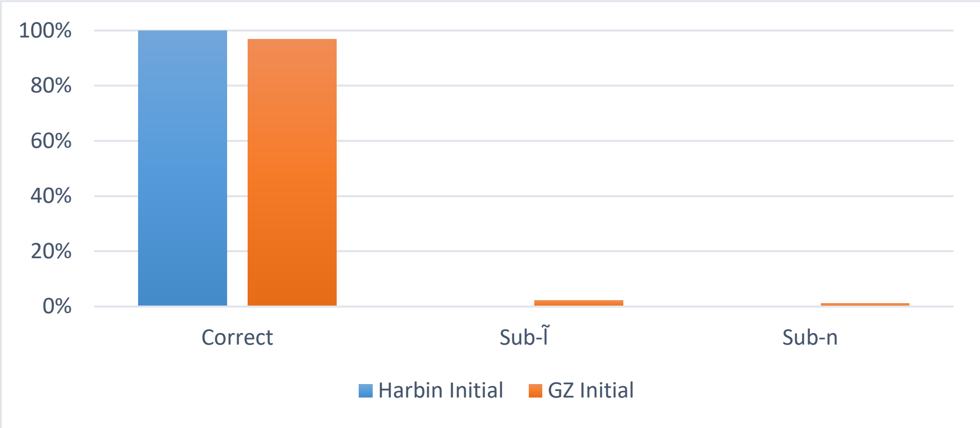


Figure 4.4 Production of English consonant /l/

Harbinese speakers performed better than Cantonese speakers in pronouncing /n/ and /l/ in the initial positions, and so it seems true that Cantonese speakers make more errors with these two consonants. However, Harbinese participants made more errors with the final /n/ on average. These results indicate that it is true that 'identical' is easier because Harbinese speakers exhibited perfect production. However, it does not seem that 'similar' for the Cantonese group is difficult. The results clearly show that Cantonese speakers also had high accuracy rates, especially in the initial position at up to 97%. They made errors, but these accounted for only very small proportions. The t-test results of /n/ calculated using average value in all positions and /l/ in initial postion in Table 4.8 show that there is no significant difference here between Harbinese and Cantonese speakers [p=0.553 for /n/ and p=0.083 for /l/ [l], both are greater than 0.05], further confirming that similarity being difficult is not true in this context. Therefore, hypothesis H1.3 that Cantonese speakers will make more errors is rejected, although they did indeed make a few errors.

Table 4.8 T-test results for English phonemes /n/ and /l/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
n	Harbin	34	87.91	8.807	1.510	-0.596	0.553
	Guangzhou	31	89.25	9.288	1.668		
l [l]	Harbin	34	100.00	0.000	0.000	1.793	0.083
	Guangzhou	31	96.77	10.017	1.799		

The English /w/, /n/ and /l/ [l] all exist in Harbinese and Cantonese but the literature mentions that Cantonese speakers tend to neutralise /n/ and /l/ in their first dialect and Harbinese speakers confuse /w/ and /v/. Thus, to further confirm the literature, these three segments were checked using their respective dialects in the second instrument of Chinese reading in this research and the results are shown in Table 4.9 and Figure 4.5 to confirm that the participants also have some difficulties with these sounds in their L1s. The results of these three segments are shown.

As seen in Figure 4.5, the Harbinese speakers pronounced /n/ and /l/ in their own dialect with 100% accuracy, which implies that Harbinese speakers can contrast /n/ with /l/ and never neutralise them. This result corresponds to the results for the production of the English /n/ and /l/ [l] in hypothesis H1.3, indicating that Harbinese speakers can pronounce these two sounds and no negative transfer to L2 English occurs. However, the neutralisation of /n/ and /l/ does happen among 71% of Cantonese speakers who neutralised /n/ into /l/ and only 29% can contrast them. Also, Cantonese speakers are not likely to pronounce /l/ as [n]. Although this happened occasionally, the percentage occurrence of this case was only at 1%. This implies that Cantonese speakers tend to neutralise /n/ to /l/ but rarely the reverse. However, the results for the production of /n/ and /l/ [l] as tested with hypothesis H1.3 do not show a serious problem for Cantonese speakers in the production of the English /n/ and /l/ [l]. They did indeed exhibit some negative transfer from their $L1_C$ but only in 3% of cases. This indicates that the neutralization of the Cantonese /n/ and /l/ is indeed transferred to L2 English but it accounts for a smaller percentage of cases compared to in Cantonese. There is a clear tendency to neutralise the Cantonese /l/ to /n/ but rarely vice versa. However, the neutralization in English does not clearly show this tendency. They both are transformed but mostly pronounced as a nasalised lateral [l̃]. These findings accord with those in the literature concerning the Cantonese /n/ and /l/.

Figure 4.5 shows that /w/ is indeed a serious issue for Harbinese speakers, who frequently substituted /w/ with [v] in up to 88% of cases. However, the Cantonese participants made no such substitution, with 100% accuracy. This further confirms that the Harbinese indeed have a problem in pronouncing /w/. We can see that the negative transfer existed in Harbinese speakers in the pronunciation of the English /w/ but the error rates in English were lower than non-standard rates in Harbinese. Meanwhile, Cantonese speakers did not make errors in the Chinese test with /w/ but made some errors in the English production of /w/ when pronouncing /w/ as [v]. The reason for this is unclear.

The results for the /n/, /l/ and /w/ pronounced in their own dialects/languages confirm the findings in the literature, in that the L1s do influence the production of L2 English.

Table 4.9 Production of segments /n, l, w/ in Chinese dialects

Segments	Types	Harbin	Guangzhou
n	Correct	100%	29%
	l		71%
l	Correct	100%	99%
	n		1%
w	Correct	12%	100%
	v	88%	

Figure 4.5 Production of segments /n, l, w/ in Chinese dialects

4.2.2 Category of 'similar' in consonants

4.2.2.1 H1.4 and results

> H1.4 The English /v/ is new in Cantonese but exists as an allophone in Harbinese where it is similar. Therefore, according to the degree of difficulty for the category of 'similar', it is hypothesised that Harbinese speakers will make errors.

Table 4.10 and Figure 4.6 show that the results for the English consonant /v/ produced by Harbin and Guangzhou participants. It can be seen that the Harbinese can pronounce /v/ better than Guangzhou Cantonese speakers with higher accuracy rates in all positions with the exception of the medial position where Cantonese participants achieved a slightly higher accuracy rate. Both groups can pronounce the word-initial /v/ better (with the accuracy rate of 68% from Harbin and 53% from GZ) than when it occurs in other positions. It seems that it is most difficult for both groups of participants to produce the word-final /v/ because the accuracy rates are then very low, at 28% in the Harbin group and 16% in the GZ group.

Substitution was the only error type made when /v/ occurred word-initially. 32% of Harbinese participants used [w] to replace /v/ while up to 46% of Cantonese participants did this. In addition, only GZ Cantonese participants substituted [ɹ] for /v/, accounting for only 1%. It seems that this case is indirectly influenced by [w] because [ɹ] was substituted for /w/, as indicated in the results for /w/ described above H1.2. [w] worked as a transition in this context for the substitution of the [ɹ] for /v/.

When /v/ occurred in a medial position, substitution was still the most common error type for both groups, followed by deletion and insertion only in the GZ group. Two substituted consonants were used by each group but [w] was the most frequently used, followed by substitution with [ʋ] used by only 1% of Harbin participants and with /f/ by 2% of GZ Cantonese speakers. Deletion was the second most popular error type for both groups (9% Harbin and 13% GZ group). Insertion was only used by GZ participants accounting for 1% of cases, where a schwa was often inserted after /v/.

When /v/ occurred in a final position, more errors were made by both groups and different strategies were used to generate the pronunciation of the word-final /v/. Substitution was still the most common error, but the substituted consonants used by each group were different, as clearly seen in Figure 4.6. Harbin participants were consistent in using [w] to replace /v/ in 64% of cases in the word-final position, but GZ participants substituted /v/ with [f] in up to 81% of cases. [f] was also used by Harbin participants but only in 1% of the cases. The GZ group also used [s] as a substituted for /v/ in 1% of cases. For the Harbin group, substitution plus insertion and insertion were other error types Harbinese participants made in the pronunciation of the word-final /v/, but Cantonese speakers did not prefer these errors. The vowel [ə] or [u] were often inserted by Harbinese speakers after the substituted consonant [w], so it can be seen that this error is a variant developing from substitution, where a vowel was inserted after the most frequently-used substitute which was [w]. Apart from these error types, deletion was also used by both groups of participants in a low percentage of 2% for cases.

Table 4.10 Production of English consonant /v/

Cons		Harbin			Guangzhou		
/v/	Types	Initial	Medial	Final	Initial	Medial	Final
	Correct	68%	47%	28%	53%	49%	16%
	Sub-w	32%	43%	64%	46%	33%	
	Sub-f			1%		2%	81%
	Sub-s						1%
	Sub-ʋ		1%				
	Sub-ɹ				1%		
	Ins-vu			1%			
	Ins-vɤ			1%			
	Ins-və					1%	
	S+I-wu			2%			
	S+I-wə			1%			
	Del		9%	2%		13%	2%
	*Error					1%	

(Cons: consonant; Sub: substitution; Ins: insertion; S+I: substitution+insertion; Del: deletion)

Disparities in the pronunciation of /v/ in the two groups across the three different positions can be easily seen in Figure 4.6. Harbinese speakers could pronounce /v/ comparatively better than Cantonese speakers. Substitution was the most popular error made in the pronunciation of the consonant /v/ across positions, and [w] was the most frequently-used substitute by Harbinese speakers everywhere. However, [w] was the most popular substitute adopted by Cantonese speakers in the initial and medial positions, while [f] was the preferred substitute for word-final /v/. In addition, deletion was also a solution used by both groups of participants to tackle the problem of /v/ in a medial position.

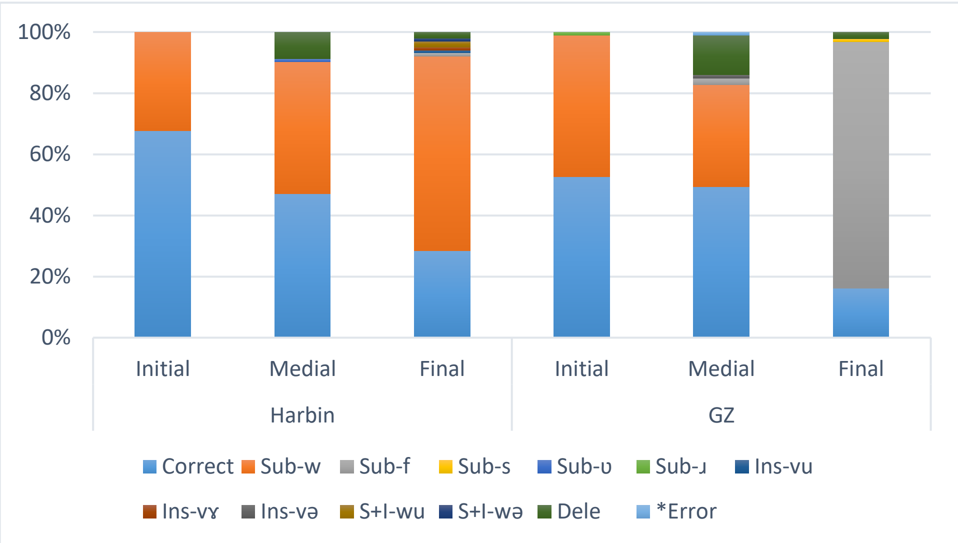


Figure 4.6 Production of English Consonant /v/

From the results for the English consonant /v/ shown in Table 4.10 and Figure 4.6, we can see that hypothesis H1.4, that Harbinese speakers will make more errors, cannot be supported by the results. Moreover, the t-test results of /v/ calculated using the mean value in all positions shown in table 4.11 confirm that there is no significant difference between the two groups ($P = 0.247$, $P>0.05$). In fact, both groups of participants made errors, but the Harbinese group pronounced /v/ better than the Cantonese group. It seems that the allophone [v] in Harbinese may play a role in helping Harbinese speakers pronounce [v] better to some extent, compared with the complete absence of /v/ in Cantonese. However, it cannot be confirmed that Harbinese speakers are better than

Cantonese speakers in pronouncing the English /v/, but the hypothesis that only the Harbinese but not Cantonese will make errors is definitely incorrect. Thus, the hypothesis H1.4 is rejected because not only do Harbinese speakers make errors, but so do Cantonese speakers. Fleges' idea that 'new' is easy cannot be confirmed here, because the English /v/ is new in Cantonese and Cantonese speakers' production is worse than that of Harbinese speakers. On the other hand, the low accuracy rates for both groups suggest the CA idea that the 'new' is difficult.

Table 4.11 T-test result for the English phoneme /v/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
v	Harbin	34	47.71	19.827	3.400	1.170	0.247
	Guangzhou	31	40.50	28.625	5.141		

4.2.2.3 H1.5 and results

> H1.5 The English /ɹ/ is similar to the Harbinese retroflex /ɻ/ but absent in Cantonese. According to the difficulty scale, Harbinese speakers will make errors but Cantonese speakers will not.

The English /ɹ/ does not exist in either Chinese dialects/languages, but Harbinese has a phonemically similar retroflex consonant /ɻ/. As mentioned in Chapter 2, /ɹ/ does not occur finally in RP but does in GA. In addition, /ɹ/ is a typical feature used to distinguish between British English (RP) and American English (GA). Therefore, /ɹ/ used in the final position here was checked. If it was pronounced by participants, this means that those participants pronounced it using an American English. If not, the variety they spoke was the non-rhotic British variety. In addition, it has to be mentioned again that all test tokens were given to participants in Chinese and they were asked to give the words in English as a precaution to reduce the influence of orthography (see Chapter 3, section 3.5). Of course, this does not mean that participants were not recruiting their knowledge of how a word was spelled.

From Table 4.12 and Figure 4.7, it can be clearly seen that although /ɹ/ exists in neither Chinese dialects/languages, it was not challenging for participants to pronounce it. Their accuracy rates in the initial and medial positions were high at 80% and above. When it occurred in the initial position, substitution was the most popular error type. The retroflex approximant consonant /ɻ/ was used by 20% of Harbinese participants to replace the English /ɹ/. Interestingly, the substituted consonants most often used to replace /ɹ/ by Cantonese speakers were [w] (4%) and [l] (2%). Only 1% of Cantonese participants pronounced /ɹ/ as [ɻ], and this small proportion might have been influenced by Mandarin. It needs to be mentioned that /ɹ/ was the onset rather than coda in three word-medial tokens because it is not pronounced in RP. Apart from 2% of the errors made by Harbinese speakers where unrelated words were pronounced, no other Harbinese participants made any errors in pronouncing the medial /ɹ/. However, besides the 2% of unrelated mistakes, Cantonese speakers also made other errors. The most common error was still substitution and [w] (5%) and [l] (2%) were used as substitutes, while the other error type was deletion (3%). These errors were never made by Harbinese.

English varieties can be distinguished according to the pronunciation of /ɹ/ as a coda. The results for the English varieties used by speakers of the two Chinese dialects/languages can be seen in Figure 4.7. More than 80% of Harbinese speakers pronounced the postvocalic /ɹ/ (orange bars). It can be said that a majority of the Harbinese speakers spoke American English in this context. Meanwhile less than 40% of Cantonese speakers pronounced the /ɹ/, indicating that most of them spoke the British variety of English (approximately 60%). A smaller percentage of Harbinese speakers (19%) spoke a British variety of English.

Table 4.12 Production of English consonant /ɹ/

Cons		Harbin			Guangzhou		
	Types	Initial	Medial	Final	Initial	Medial	Final
	Correct	80%	98%		91%	87%	
	A			81%			38%

/ɹ/	B			19%			62%
	Sub-w				4%	5%	
	Sub-l				2%	2%	
	Sub-ɻ	20%			1%		
	Del					3%	
	*Error		2%		1%	2%	

(Cons: consonant; A: American English; B: British English; Sub: substitution; Del: deletion)

In summary both groups of participants could easily pronounce English /ɹ/. Some errors were made, where /ɹ/ was the most frequent substitute for [ɻ] among Harbinese participants but as [w] or [l] were more common among Cantonese speakers. In addition, Harbinese speakers tended to speak the American variety of English but Cantonese speakers favoured British English.

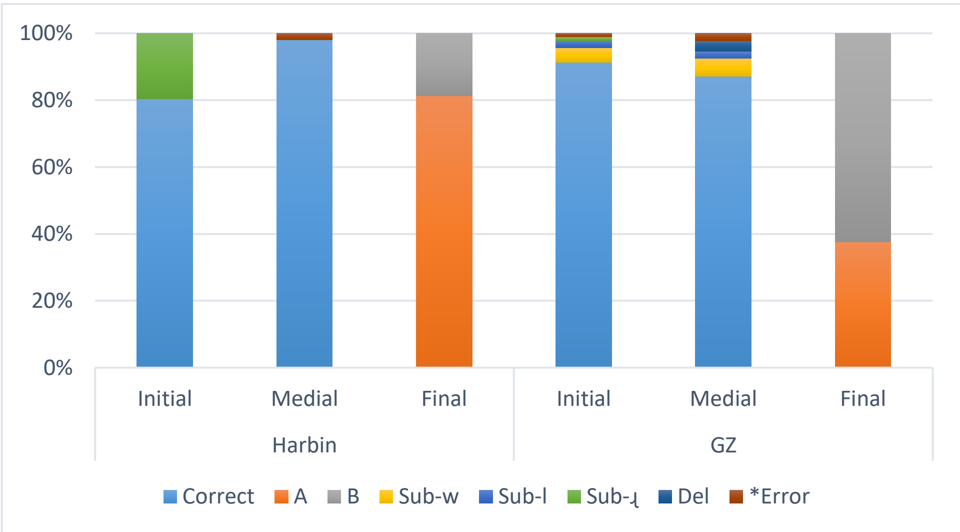


Figure 4.7 Production of English consonant /ɹ/

Hypothesis H1.5 that Harbinese will make errors cannot be supported by the results for the production of /ɹ/ even though the accuracy rate for /ɹ/ in the initial position pronounced by Cantonese speakers is higher than by in Harbinese speakers. However, when /ɹ/ occurred in the medial position, the Harbinese group had higher accuracy rates than the Cantonese. In addition, the statistical analysis for /ɹ/ calculated using average value in all positions reveals that, despite the differences between the two dialect groups, they are not statistically significant [p=0.508 (>0.05)], as shown in Table 4.13.

Therefore, hypothesis H1.5 that the Harbinese will make more errors is rejected. Although accuracy rates in the Harbinese group are numerically lower than that in Cantonese group, Flege's SLM-based prediction that 'similar' is difficult cannot be supported because the difference is not significant.

Table 4.13 T-test results for English phoneme /ɹ/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
ɹ	Harbin	34	89.22	14.142	2.425	-0.665	0.508
	Guangzhou	31	89.25	19.976	3.588		

It can be concluded that Harbinese speakers prefer to speak the American variety of English while Cantonese speakers use the British variety. The t-test was applied to determine if this conclusion is statistically significant. Table 4.14 shows whether or not /ɹ/ is rhoticised significantly differently between Harbinese and Cantonese speakers. The results of /ɹ/ in final position reveal that there is a strongly statistically significant tendency for Harbinese speakers to prefer GA and for Cantonese speakers to favour the British variety of English as represented by p-values (<0.001) for both varieties.

Table 4.14 T-test results for English phoneme /ɹ/ rhoticised in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
ɹA	Harbin	34	81.37	26.197	4.493	6.332	0.000
	Guangzhou	31	37.63	29.493	5.297		
ɹB	Harbin	34	18.63	26.197	4.493	-6.332	0.000
	Guangzhou	31	62.37	29.493	5.297		

4.2.2.4 H1.6 and results

H1.6 The English /dʒ, tʃ, ʃ/ are phonemically similar to the retroflex /ʈʂ, ʈʂʰ, ʂ/ in Harbinese. However, /dʒ, tʃ, ʃ/ are comparatively allophones in Cantonese, so they are allophonically identical but phonemically different in Cantonese. According to the criteria of degree of difficulty in the category of 'similar', it is hypothesised that a) both should be difficult but b) the Cantonese will make errors.

The voiced postalveolar affricate /dʒ/ was not a difficult English consonant to be pronounced by either group of participants, as seen in Table 4.15 and Figure 4.8, especially when it occurred in initial and medial positions. Cantonese speakers pronounced /dʒ/ in these positions better than Harbinese speakers but worse in the final position. Both groups adopted substitution as their main error type with the pronunciation of /dʒ/ across positions, except for in the final position. When /dʒ/ was in the initial position, more mistakes were made by Harbinese speakers but both had high accuracy rates, at respectively 88% and 97%. Different error types were made by participants. The Harbinese used substitution, where the voiceless unaspirated retroflex [ʈʂ] was the most preferred substitute accounting for 11%, while the Cantonese used insertion and a glide [j] was always inserted after /dʒ/. In the medial position, [ʈʂ] was still the most frequently-used substitute among Harbinese speakers although various other substitutes were also rarely used (together representing 1%), while [ʒ] was the only substitute used by Cantonese speakers in this position, making up 1% of cases.

When /dʒ/ occurred in the final position, substitution plus insertion was the main error made by participants from both groups. As Harbinese participants did in initial and medial positions, they first substituted /dʒ/ with the voiceless aspirated retroflex [ʈʂ], and then a special retroflex vowel was inserted. This vowel exists in Mandarin/Harbinese and is marked with the symbol [ʅ] by Chinese linguists; however, the retroflex approximant [ɻ] was adopted in the present research because [ʅ] is still controversial. Some researchers like Duanmu (2000:26) regard it as a retroflex consonant rather than a special vowel, but the symbol /r/ he uses is not a retroflex; therefore, the retroflex approximant [ɻ] is used in this study to represent the traditionally-alleged retroflex vowel [ʅ] in Mandarin/Harbinese Chinese. In addition, the symbol [ʅ] cannot be found in the IPA chart, and thus [ɻ] is a good choice. Therefore, when /dʒ/ occurrs finally, a syllable [ʈʂɻ] composed of the substitute [ʈʂ] for /dʒ/ plus the insertion of [ɻ] was pronounced by Harbinese speakers in order for it to conform to the rule of the well-formedness of the syllable [ʈʂɻ] in Harbinese. It is clear that $L1_H$ transfer to L2 English occurred. However, Cantonese speakers substituted /dʒ/ with different consonants and vowels. They replaced /dʒ/ with the voiceless aspirated and unaspirated alveolo-palatal affricates [tɕ] and [tɕ^h] first and then a [y] sound was inserted following the substitution. Thus, when /dʒ/ occurred finally, /dʒ/ was pronounced by 16% of Cantonese speakers as the syllable [tɕy] or [tɕ^{h}y]. It needs to be mentioned that a very small proportion of Cantonese participants pronounced the final /dʒ/ as the retroflex syllable [tʂɻ̩], as the Harbinese did. This could imply that some of the Cantonese participants were influenced by Mandarin, but not in significant numbers because this only accounted for 1% of cases. Besides the method of substitution plus insertion, the results also show that Cantonese speakers used substitution as their second commonest error type used (14%) to replace /dʒ/ occurring in the final position. However, the most frequently used substitute was the voiceless counterpart [tʃ].

Figure 4.8 demonstrates that participants performed well in pronouncing the English consonant /dʒ/. Substitution was the error type adopted most often but the complex type of substitution and insertion was the preferred type in the final position.

Table 4.15 Production of English consonant /dʒ/

Cons		Harbin			Guangzhou		
	Types	Initial	Medial	Final	Initial	Medial	Final
	Correct	88%	91%	78%	97%	99%	69%
	Sub-tʂ	11%	4%	1%			
	Sub-tɕ	1%					
	Sub-ʒ			1%		1%	
	Sub-tʃ						14%
	Sub-dɹ		1%				
/dʒ/	Sub-z		1%				
	Sub-ɻ		1%				
	Sub-dʒ^w		1%				
	Ins-dʒj				3%		
	Ins-dʒɻ						1%
	S+I-tʂɻ̩			20%			1%
	S+I-tɕy̥						8%

	S+I-tɕ^{h}y̥						8%
	*Error		1%				

(Cons: consonant; Sub: substitution; Ins: insertion; S+I: substitution+insertion)

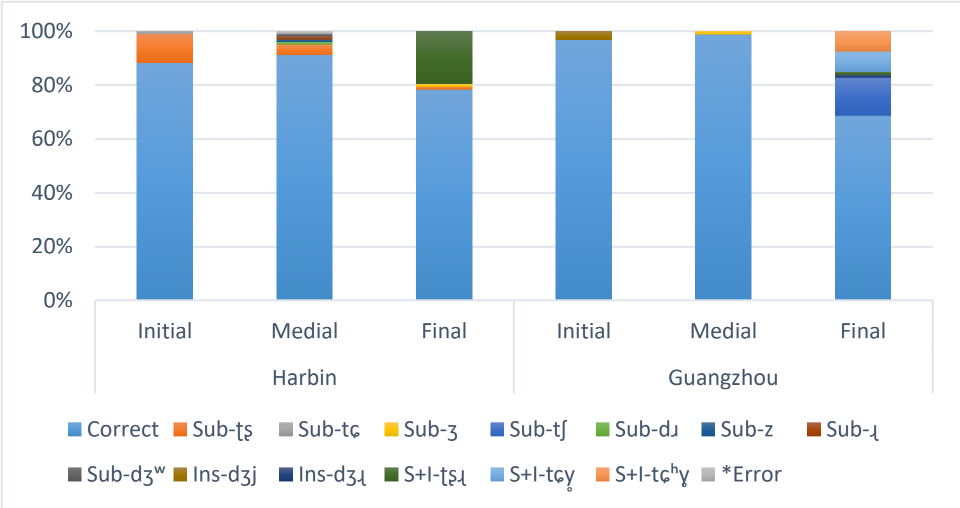


Figure 4.8 Production of English Consonant /dʒ/

From Table 4.16 and Figure 4.9, it can be seen that participants from both dialect areas could pronounce /tʃ/ with high accuracy especially Cantonese speakers. The error types made in different positions differed between groups, and so were the substitutes used. Table 4.16 demonstrates that substitution was the only type of error made by participants when /tʃ/ occurred word-initially and medially. When it occurred in the final position, the most frequent error type was substitution plus insertion. However, the substitutes and inserted segments used were specific to each group as for /dʒ/. Participants from both groups had no difficulty in producing the initial /tʃ/, although they made a few small errors. Harbinese participants preferred to insert the voiced labial-velar approximant [w] after /tʃ/ and coarticulate them as [tʃ^{w}] as a substitute replacing /tʃ/, while GZ participants preferred [tɹ] instead. When /tʃ/ occurred word-medially, the top three substitutes Harbinese speakers used were [tʃ^{w}] (10%), [tɹ] (8%), and [ʈʂ^{h}] (6%) but [tɹ] was the most popular substitute for Cantonese speakers, used in 4% of cases. The errors made with /tʃ/ occurring in the final position among both groups of participants were substitution plus insertion. The final /tʃ/ was always substituted with [ʈʂ^{h}] (14%) by the Harbinese and with [tɕ^{h}] (10%) by the Cantonese participants. A different segment was also inserted after the respective substitutions with [ɻ] inserted by Harbin participants and [y] by GZ participants, which are the same strategies used into tackling /dʒ/ in the final position.

Table 4.16 Production of English consonant /tʃ/

Cons		Harbin			Guangzhou		
	Types	Initial	Medial	Final	Initial	Medial	Final
	Correct	96%	73%	86%	98%	91%	88%
	Sub-$tɕ^h$	1%					1%
	Sub-$tʃ^w$	2%	10%			1%	
	Sub-t		2%				
	Sub-ts					1%	
/tʃ/	Sub-ts^{hw}		1%				
	Sub-$tʂ^h$		6%				
	Sub-tɹ	1%	8%		2%	4%	
	Sub-ɾ		1%				
	Sub-f						1%
	S+I-$tɕ^h$y						10%
	S+I-$tʂ^h$ɻ			14%			
	*Error					2%	

(Cons: consonant; Sub: substitution; S+I: substitution+insertion)

It can be seen from Figure 4.9 that neither group had much difficulty in pronouncing the English consonant /tʃ/ and the Cantonese group pronounced it slightly better, especially in the word-medial position. Substitution and substitution and insertion were the most common error types when /tʃ/ occurred initially and medially, and finally respectively.

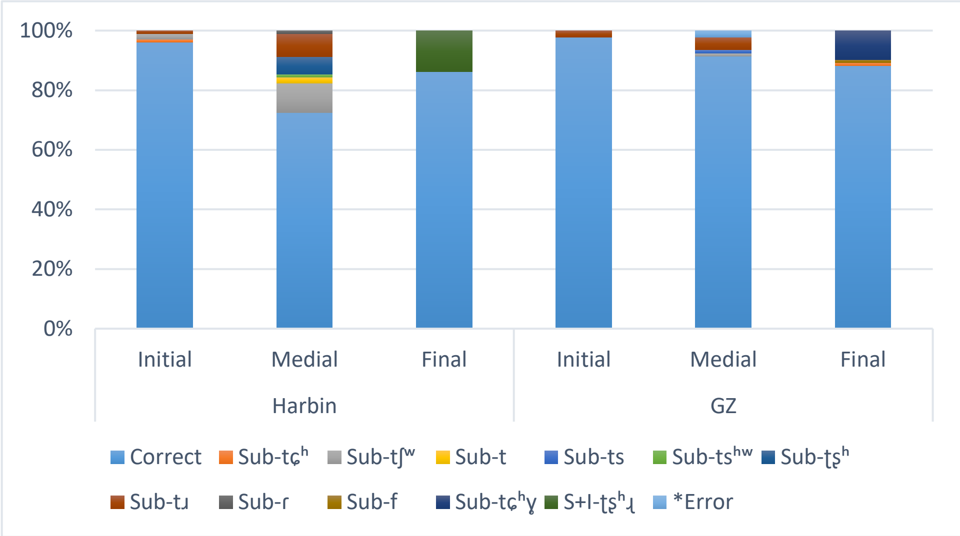


Figure 4.9 Production of English consonant /tʃ/

It can be seen from Table 4.17 and Figure 4.10 that both dialect groups of participants had no problem in the pronunciation of the voiceless postalveolar fricative /ʃ/ consonant, because they achieved very high accuracy rates up to more than 95% across positions except for the final position for Harbin participants with a comparatively low accuracy rate of 86%. Harbin and Guangzhou participants could pronounce the initial /ʃ/ perfectly well with only very minor errors accounting for 1% of cases in each group. In fact, these made in the Harbin group were because participants produced an unrelated word rather than the targeted one, whereas 1% of GZ participants pronounced it as the voiceless alveolar-palatal fricative [ɕ]. When it occurred in the medial position of a word, /ʃ/ was mispronounced by a small proportion of participants and substitution was adopted. The Harbin participants adopted [s], [tʃ] or [t] as substitutes for /ʃ/ but GZ participants only substituted /ʃ/ with its voiced counterpart [ʒ]. In the word-final position, substitution plus insertion was the most frequently used error type and both groups inserted a vowel after the substituted consonant. However, the substitutes and inserted vowels used were different from those for /dʒ/ and /tʃ/. 12% of Harbinese participants substituted the voiceless retroflex fricative [ʂ] for /ʃ/ plus an insertion of [ɻ]. However, the Cantonese participants preferred the voiceless alveolo-palatal fricative [ɕ] and aspirated affricate [tɕ^{h}] as substitutes and a [y] sound was inserted and devoiced after the substituted consonant, as when they tackled the final /dʒ/ and /tʃ/.

Table 4.17 Production of English consonant /ʃ/

Cons		Harbin			Guangzhou		
/ʃ/	Types	Initial	Medial	Final	Initial	Medial	Final
	Correct	99%	96%	86%	99%	97%	96%
	Sub-s		1%				
	Sub-tʃ		1%	2%			
	Sub-t		1%				
	Sub-ɕ				1%		
	Sub-ʒ					3%	
	S+I-ʂɻ			12%			
	S+I-ɕy̥						2%
	S+I-tɕʰy̥						2%
	*Error	1%	1%				

(Cons: consonant; Sub: substitution; S+I: substitution+insertion)

Figure 4.10 shows that /ʃ/ was also not a difficult consonant for both dialect groups of participants to pronounce. A few minor errors were made when it occurred in medial and final positions, and especially in the final position for Harbinese speakers. Substitution was the most common error type in the medial position and substitution plus insertion was most frequently used in the final position.

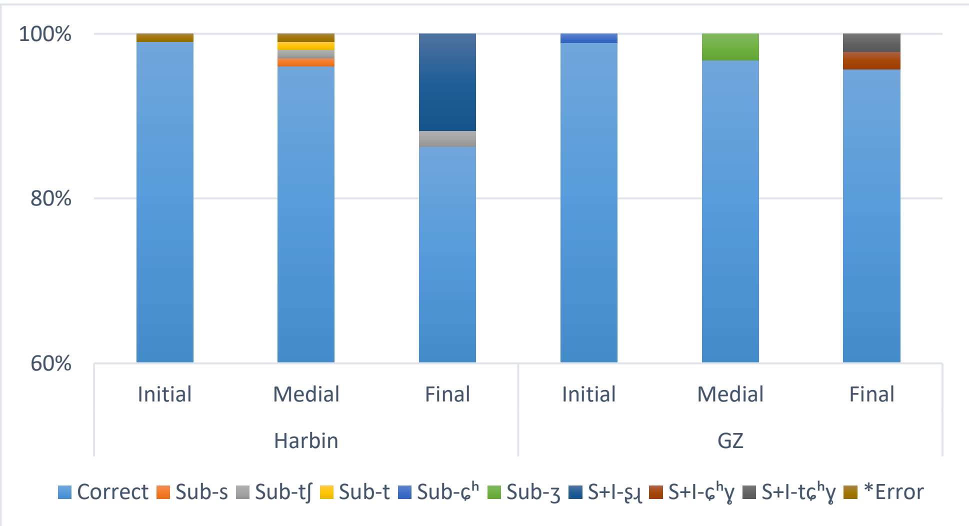


Figure 4.10 Production of English consonant /ʃ/

The results for these last three consonants show that Cantonese participants performed better in pronouncing them than Harbinese participants overall. It seems that the results contradict the prediction according to degree of difficulty in the sense that the 'similar' as 'allophonically identical and phonemically different' is more difficult than the 'similar' as 'phonemically similar'. However, the results indicate that the existence of allophones may play a role in helping the Cantonese speakers. According to Flege's SLM (1995), 'similar' is difficult; that is, 'similar' leads to more errors or lower accuracy rates. $L1_C$ and $L1_H$ are both similar to the L2 English /dʒ, tʃ, ʃ/ but similar in a different way. Thus, both should have made errors and achieved low accuracy rates, but the actual results do not reflect this. Therefore, it can be claimed that Flege's idea that 'similar' is difficult does not work in this case. However, the statistical results calculated using the average value of /dʒ, tʃ, ʃ/ in all positions seen in Table 4.18 indicate that the differences between $L1_H$ and $L1_C$ are not statistically significant for the L2 English /dʒ/ and /ʃ/ [p=0.508 and p=0.129 respectively] but are significant to /ʃ/ [p=0.027]. Therefore, which type of similarity is more difficult cannot be absolutely confirmed here because the differences found may not be statistically significant. In addition, the Cantonese achieved better results than the Harbinese for all three targets. This may suggest that the similarity as 'allophonically identical and phonemically different' is easier than the type which is 'phonemically similar'. That is, the more similarity between the L2 and L1 in this context, the easier it may be for L2 learners to learn.

From the production and t-test result in tables 15-18, it can be claimed that hypothesis H1.6a) that pronunciation will be difficult for both groups is rejected, because both achieved high accuracy rates. The hypothesis H1.6b) that Cantonese will make errors is also rejected because the Cantonese production was better although not statistically significantly so for the segments /dʒ/ and /ʃ/.

Table 4.18 T-test results table for English phonemes /dʒ/, /tʃ/ and /ʃ/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
dʒ	Harbin	34	85.95	14.552	2.496	-0.665	0.508
	Guangzhou	31	88.17	12.149	2.182		
tʃ	Harbin	34	84.97	15.444	2.649	-2.269	0.027
	Guangzhou	31	92.47	10.496	1.885		
ʃ	Harbin	34	93.79	10.303	1.767	-1.540	0.129
	Guangzhou	31	97.13	7.008	1.259		

4.2.3 Category of 'new' in consonants

4.2.3.1 H1.7 and results

> H1.7 The English /θ, ð, ʒ/ and (/l/) [ɫ] are absent from both dialects/languages. According to Flege's SLM, 'new' is easy; therefore, a) neither will make errors with these segments; b) if they do make errors, there will be no differences between the two dialects/languages.

The voiceless interdental fricative /θ/ does not exist in either dialect/language. From Figure 4.11, we can see that both groups of participants had difficulties pronouncing this English consonant but the Cantonese performed comparatively better than the Harbinese. Table 4.19 presents their accuracy rates across positions, which were below 50% except for the initial position in the Cantonese group who achieved 60% accuracy rate. From Table 4.19, it can be seen that substitution was the error type most used. However, it varied slightly between the medial and final positions. Although substitution was the main error type for both groups across positions, Cantonese speakers also used deletion in the medial position while Harbinese speakers used more types in final position, including substitution plus insertion, insertion, and deletion. Figure 4.11 indicates the specific different errors clearly. The long orange bar suggests that Harbinese speakers often mispronounced /θ/ as the voiceless alveolar fricative [s] everywhere in all positions. However, /θ/ was frequently pronounced as [f] by Cantonese speakers across positions as seen in the large dark blue bars in Figure 4.11. Apart from these main errors, a small proportion of Cantonese participants pronounced

/θ/ as [s] as did the Harbinese, which might be influenced by Mandarin. However, these errors are insignificant in frequences compared with the main error [f]. The figure shows that, in the final position, besides the main error [s], [sɿ] was only made by Harbinese speakers in 9% of cases. which was indeed developed from substitution [s] plus the special Chinese vowel [ɿ]. Like the special Harbinese/Mandarin retroflex vowel [ʅ] which was often inserted after the English /dʒ, tʃ, ʃ/ by Harbinese speakers when /dʒ, tʃ, ʃ/ occurred finally, as mentioned in the section 4.2.2.4, this special vowel [ɿ] occurred after /t͡s, t͡sʰ, s/ in Harbinese and Mandarin. The reason why the special vowel occurred in this context was due to the influence of $L1_H$ being exerted on L2 English. When the final /θ/ was pronounced as [s], [ɿ] was often inserted by some Harbinese participants after the [s] to conform to the convention in Harbinese that [sɿ] should be pronounced as an individual syllable. It needs to be mentioned that this special Harbinese vowel [ɿ] as with [ʅ], is controversial. Chao (1968:24) and Duanmu (2000:36-7) think that Chinese apical vowels are voiced syllabic consonants which are extended from the preceding consonants to the syllabic position (Lin, 2007:72). To maintain consistency with /ʅ/ represented by the symbol [ɻ̩], I follow Lee and Zee (2003) and Lin (2007:72) in replacing the controversial Harbinese vowel [ɿ] with a syllabic approximant [ɹ̩]. From Table 4.19, we can see that, besides [ɹ̩] inserted after [s], other vowels such as [a] or [ə] were also inserted, but they only accounted for small proportions of insertions. Therefore, [s] is the most widespread error among Harbinese speakers and [f] is among Cantonese speakers when pronouncing the English /θ/.

Table 4.19 Production of English consonant /θ/

Cons		Harbin			Guangzhou		
	Types	Initial	Medial	Final	Initial	Medial	Final
	Correct	46%	45%	50%	60%	47%	49%
	Sub-s	52%	53%	36%	5%	8%	4%
	Sub-z		2%			1%	
	Sub-t͡s	1%		1%			
	Sub-d				1%	2%	
	Sub-d̪	1%					
/θ/	Sub-f				32%	39%	46%

	Sub-ʒ					1%	
	Ins-θɹ̩			1%			
	S+I-sɹ̩			9%			
	S+I-sa			1%			
	S+I-sə			1%			
	Del			1%		2%	
	*Error				1%		

(Cons: consonant; Sub: substitution; Ins: insertion; S+I: substitution+insertion; Del: deletion)

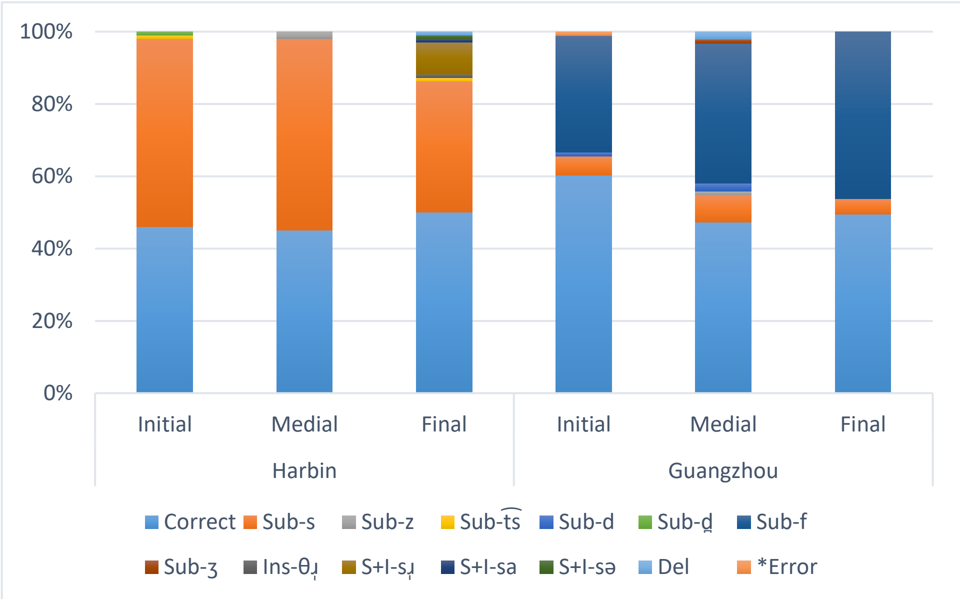


Figure 4.11 Production of English consonant /θ/

From Figure 4.12, it can be seen that both dialect/language speakers have considerable difficulty in pronouncing the voiced interdental fricative /ð/, but Harbinese speakers had much better performance, pronouncing it with higher accuracy than Cantonese speakers across all positions. The accuracy rates achieved by Harbinese speakers ranged from 40% to 62% while those of Cantonese speakers were very low, with the highest at 26%. Substitution was the most frequently used method to substitute for /ð/, with various substitutes used by participants across positions as seen in Table 4.20. Other types of errors were substitution plus insertion, insertion and deletion. They were only used in non-initial positions by considerably smaller proportions of participants.

From Figure 4.12, we can see that participants used different substitutes as solutions when /ð/ occurred in different positions. For the initial /ð/, only substitution was used. Various substitutes were used by Harbinese speakers, but among these errors [d] and [t͡s] were the two most frequently pronounced, respectively accounting for 26% and 23%, followed by [d̪] (5%) and [z] (5%). However, only two substitutes were used by Cantonese participants. Up to 91% pronounced the initial /ð/ as [d] and another 5% as the dentalised alveolar stop [d̪]. When /ð/ occupied a medial position, the accuracy rate of the Harbinese group pronouncing correctly (64%) was over double that of the Cantonese group (26%). In this position, it was often pronounced by Harbinese speakers as [t͡s] (24%) and [z] (10%) rather than [d] which only made up 4% of cases. [d] and [d̪] were still the most preferred substitutes used by Cantonese speakers respectively accounting for 37% and 16% of cases. A small percentage of Cantonese participants pronounced it as the flap [ɾ] (6%) and voiceless labio-dental fricative [f] (4%). Apart from these errors, a special feature which occurred in this position in the Cantonese group was that /ð/ was developed into a pattern of a fricative [f]/[ð]/[θ] plus a stop [d]/[d̪], although the proportion of cases is fairly small as seen in Table 4.20. When /ð/ occurred finally, the Harbinese combined to use [t͡s] (12%) and [z] (10%) as the most popular substitutes along with [θ] (9%) and [s] (8%). Another error type of substitution plus insertion was also frequently used where /ð/ could be substituted by [t͡s] and then a vowel was inserted to form a [t͡si] syllable. This can be regarded as the further development of the substitution [t͡s]. Interestingly, the Cantonese adopted different substitutes to replace the final /ð/. The most common two substitutes were [f] and [θ] accounting for 39% and 30% respectively, followed by the infrequently used substitute [s] (3%). Other types of errors were rarely used in the final position, such as substitution plus insertion and deletion.

Table 4.20 Production of English consonant /ð/

Cons		Harbin			Guangzhou		
	Types	Initial	Medial	Final	Initial	Medial	Final
	Correct	40%	62%	48%	3%	26%	26%
	Sub-f					4%	39%

/ð/	Sub-θ			9%		1%	30%
	Sub-s			8%		1%	3%
	Sub-z	5%	10%	10%		1%	
	Sub-d	26%	4%		91%	37%	
	Sub-d̪	5%			5%	16%	
	Sub-t		1%				
	Sub-ɾ					6%	
	Sub-l	1%					
	Sub-t͡s^h			1%			
	Sub-t͡s	23%	24%	12%		2%	
	Sub-dʒ					1%	
	Ins-ðd̪					1%	
	S+I-t͡sɹ̩			10%			
	S+I-sɹ̩			2%			
	S+I-θd̪					1%	
	S+I-fs						1%
	I+S-fd					2%	
	Del						1%
	*Error			1%			

(Cons: consonant; Sub: substitution; Ins: insertion; S+I: substitution+insertion; I+S: insertion+substitution; Del: deletion)

In summary, the English /ð/ was a difficult English consonant for both groups, but the Harbin group achieved better pronunciation accuracy than the Cantonese. Harbinese speakers preferred [t͡s] (dark green colour in Figure 4.12) across positions to substitute for /ð/ or [d] in the initial position, but Cantonese speakers preferred [d] (light green colour) in initial and medial positions and [f] (orange) and [θ] (grey) in the final position.

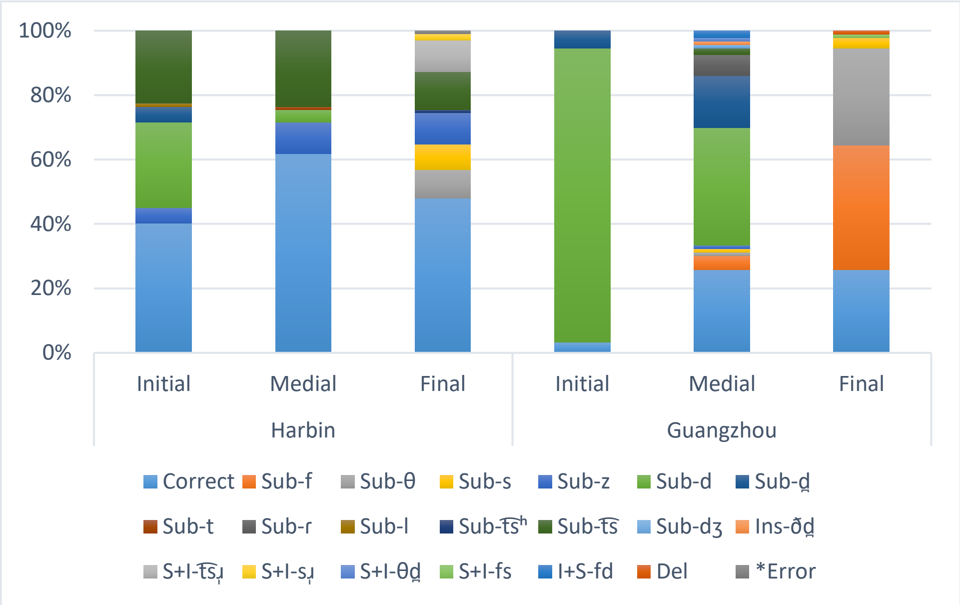


Figure 4.12 Production of English consonant /ð/

The voiced postaveolar fricative /ʒ/ was only checked in the medial position because it is very rare in other positions and such English words would be unknown to the participants. From Table 4.21 and Figure 4.13, we can see that there was a huge disparity between the two groups. It was a rather difficult consonant for Harbinese speakers to pronounce with an accuracy rate of only 7% while it was easier for the Cantonese whose accuracy rates reached up to 67%. Substitution was the main error type for both groups of participants. Various substitutes were used to replace the targeted /ʒ/. [ɹ] was most frequently used by Harbinese speakers (39%), followed by [ɻ] (14%), [dʒ] (11%) and [ʃ] (8%), while [ʃ] was the most preferred substitute for Cantonese speakers, followed by [dʒ] (6%) and others with low proportions. Substitution was the only strategy used by Cantonese speakers but Harbinese speakers also used deletion and substitution plus insertion. Deletion accounted for 14% of cases as the second most commonly used method to tackle the pronunciation of /ʒ/.

Table 4.21 Production of English consonant /ʒ/

Cons		Harbin			Guangzhou		
	Types	Initial	Medial	Final	Initial	Medial	Final

/ʒ/	Correct		7%			67%	
	Sub-ɹ		39%			3%	
	Sub-dʒ		11%			6%	
	Sub-ɻ		14%				
	Sub-ʃ		8%			20%	
	Sub-tɕ		1%				
	Sub-s		1%			3%	
	Del		14%				
	S+I-jiɹ		1%				
	S+I-sj		1%				
	*Error		4%				

(Cons: consonant; Sub: substitution; Del: deletion; S+I: substitution+insertion)

It can be concluded that both Cantonese and Harbinese speakers made errors with the English /ʒ/ but Cantonese speakers achieved much better results. Substitution was the only method Cantonese speakers used and the most highly-used substitute was [ʃ] while substitution and deletion were the two most common error types among Harbinese speakers and the top substitutes they used were the rhotic approximants [ɹ] and [ɻ].

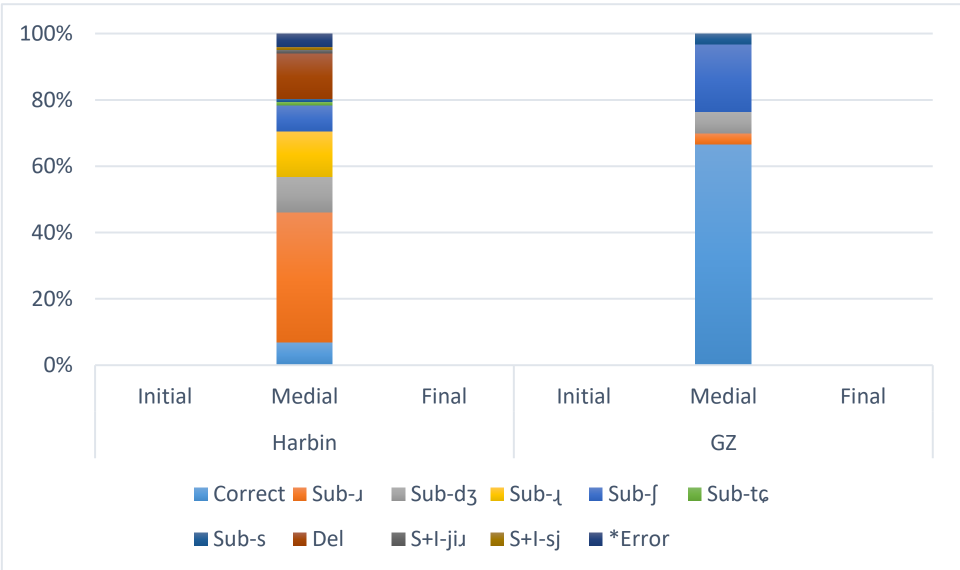


Figure 4.13 Production of English consonant /ʒ/

Two sets in total of six tokens, *always*, *soldier*, *world,* and *hole*, *smell*, *school* were used to check English dark [ɫ]. The first set was used to test word-meidal [ɫ] and the second set to test word-final [ɫ]. To be more precise, the medial [ɫ] was word-internal, and for the first two words, in the coda of the first of two syllables and the coda of the third word.

When /l/ occurred word-medially and finally as [ɫ], accuracy rates decreased dramatically especially for the Harbinese group whose accuracy rate was only around 10%, as shown in Table 4.22. For this dark lateral [ɫ] which does not exist in Chinese, both groups of participants adopted deletion as the most popular error, especially among the Harbin, where 82% of participants deleted the medial /l/ and 65% the final /l/. Similarly, high percentages of the Cantonese group also used this method (47%) for the medial /l/ and 51% for final the /l/. Substitution was also a common error type for both groups, and [w] was the most frequently used substitute among many. The percentages using the substitute [w] for /l/ in the final position were higher than in the medial position, as clearly seen in the grey bars in Figure 4.14. Another method of the insertion of a schwa was also used, but only rarely. In this case, /l/ was converted from the dark /l/, [ɫ], into a clear /l/, [l].

Table 4.22 Production of English consonant /l/ [ɫ]

Cons		Harbin			Guangzhou		
/l/ [ɫ]	Types	Initial	Medial	Final	Initial	Medial	Final
	Correct		13%	9%		41%	19%
	Del		82%	65%		47%	51%
	Sub-w		4%	25%		6%	29%
	Sub-ɫ̃						
	Sub-n						
	Sub-ɔ			2%			
	Sub-oʊ					2%	
	Ins-lə		1%			2%	1%
	Meta-sloʊd i					1%	

(Cons: consonant; Del: deletion; Sub: substitution; Ins: insertion; Meta: metathesis)

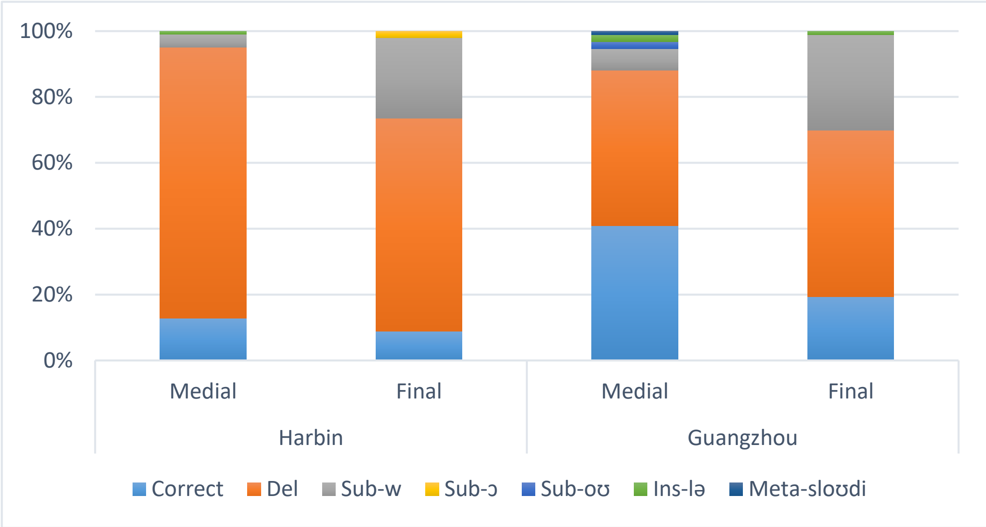


Figure 4.14 Production of English consonant /l/ [ɫ]

The above results for these three English consonants and the English allophone [ɫ] which are absent from both Chinese dialects/languages reveal that hypothesis H1.7a that neither group will make errors with these segments is rejected because both made many errors when pronouncing /θ, ð, ʒ/ and [ɫ]. Moreover, H1.7b, that their dialect/language will make no difference is also rejected because the errors made do differ. Statistical tests calculated using the 'average' in all positions were of the significance of these differences between the two groups. The results confirm the differences between the two groups in pronouncing the consonants /ð/, /ʒ/ and [ɫ] are strongly statistically significant with all p-values less than 0.01 [p=0.000 (<0.05)] as seen in Table 4.23, while there is no statistically significant difference for the consonant /θ/ [p=0.574 (P>0.05)]. The t-test results further confirm that the two groups exhibited significant differences in pronouncing /ð/, /ʒ/ and [ɫ] although no difference was significant for the segment /θ/. However, the errors made with /θ/ show that the two groups made different errors. It can be claimed from the above statement that Flege's idea that 'new' is easy is not supported at all by the results in this context. On the other hand, 'new' is confirmed to be very difficult.

Table 4.23 T-test results for English phonemes /θ/, /ð/, /ʒ/ and /l/ [ɫ] in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
θ	Harbin	34	47.06	33.509	5.747	-0.565	0.574
	Guangzhou	31	52.33	40.895	7.345		
ð	Harbin	34	50.00	29.367	5.036	5.258	0.000
	Guangzhou	31	18.28	18.492	3.321		
ʒ	Harbin	34	6.86	15.953	2.736	-9.085	0.000
	Guangzhou	31	66.67	33.334	5.987		
l [ɫ]	Harbin	34	10.7856	12.22879	2.09722	-5.484	0.000
	Guangzhou	31	30.1094	15.76389	2.83128		

4.2.4 Summary of the results of hypotheses relating to consonants

From the results for consonants presented above, it can be summarised that both groups have more difficulty in English consonants /θ, ð, ʒ/ and (/l/) [ɫ]/ which are absent in L1s. Also, more errors made by both groups occur in final positions, especially to the consonants /v/, /n/ and /dʒ/. Hypotheses H1.1 and H1.2 in the category of 'identical' were supported, which confirms that 'identical' is easier than 'similar' except H1.3, whose results did show 'identical' is easier but there was no statistically significant difference. However, in turn, H1.3 can imply that 'similar' is not difficult because it had quite similar results with the category 'identical' which was supported to be easier. All hypotheses from H1.4 to H1.6 in the category of 'similar' were rejected, which means 'similar' being difficult cannot be supported. The two predictions in H1.7 in the category of 'new' were both rejected. Both groups made not only errors but also different errors. The results of H1.7 confirm that 'new' is not easy but very difficult for both groups.

4.3 Results of hypotheses relating to vowels

This section presents the results of the testing of hypotheses relating to vowels. The hypotheses were supported or rejected according to statistical tests conducted to show if the differences between the Harbin and Guangzhou groups were significant or not.

4.3.1 Category of 'identical' in vowels

4.3.1.1 H2.1 and results

> H2.1 The English diphthongs /eɪ, aɪ, aʊ, oʊ (GA), ɔɪ/ exist in Cantonese but are only similar in Harbinese because Harbinese has a different /ɑu/ and lacks /ɔɪ/. Harbinese English learners will therefore make errors with these diphthongs.

English diphthongs are covered by Hypothesis 2.1. Differences between the two dialects/languages with English diphthongs occur with /au/ and /ɔɪ/, where Harbinese has a similar /au/ but does not have /ɔɪ/. In this section, the results for five English diphthongs pronounced by the two groups are based on the difference: group one has no difference between two dialects/languages (/aɪ/, /eɪ/, and /oʊ/) and group two has the difference (/ɑu/ and /ɔɪ/).

The English diphthong /aɪ/ exists in Harbinese and Cantonese but in Harbinese it is not pronounced as open and long as in English, without gliding from one vowel to the other with considerable movement. The first vowel /a/ in the Harbinese /aɪ/ is pronounced shorter than that in English and the movement in gliding from one vowel to the other is quicker.

It can be seen from Table 4.24 that both groups of participants did not always have the difficulty in pronouncing /aɪ/, especially for the Cantonese who achieved high accuracy rate of 86%. Although the accuracy rate of Harbinese group was 30% lower than the Cantonese group, 38% of Harbinese participants pronounced /aɪ/ as [aɛ], which is very close to /aɪ/. The difference between them is that the second vowel of /aɪ/ was not pronounced as high as it should be but lowered to [ɛ]. Fewer Cantonese had made this error (13%). The error of diphthongal substitution was the most frequent among both groups of speakers, especially by Harbinese speakers. Apart from this error, the diphthong was often contracted into a close monophthong by Harbinese participants and also lengthened such as [æ:] and [ɛ:]. However, Cantonese participants did not have this problem, but they sometimes inserted a nasal [n] after /aɪ/.

The diphthong /eɪ/ also exists in both Chinese dialects/languages. Table 4.24 indicates that participants could pronounce this diphthong easily with very high accuracy rates reaching 97% for Harbinese speakers and 99% for Cantonese. Only one error type was made, which was the contraction of the diphthong into a monophthong. The specific errors made with /eɪ/ were [ɛ:] by Cantonese and [e:] by Harbinese speakers. The preference for the substitution of a diphthong by a lengthened monophthong can again be seen here.

Rows A and B in Table 4.24 and afterwards again represent the American and British varieties of English. This diphthong /əʊ/ occurs in British English while its counterpart in American English is /oʊ/. The two varieties of Chinese have the American diphthong /oʊ/. However, the results in Table 4.24 indicate that different varieties of English were used by both dialect/language speakers. Both groups had more than half of the participants pronounced the American diphthong. However, more Cantonese participants articulated the British /əʊ/ than Harbinese (25%). Participants also made errors with this diphthong. The most frequent was the contraction of the diphthong to a monophthong, and /oʊ/ was reduced to the lengthened monophthongs [o:] and [ɔ:] respectively accounting for 15% and 3% of cases among Harbinese speakers out 2% and 1% in the Cantonese group. In addition, smaller percentages of participants substituted /əʊ/ or /oʊ/ with other diphthongs such as [ɤʊ] and [aɔ].

Table 4.24 Production of English vowels /aɪ/, /eɪ/ and /əʊ (oʊ)/

Vowels	Types	Harbin	Guangzhou
aɪ	Correct	56%	86%
	aɛ	38%	13%
	æ:	4%	
	ɛ:	1%	
	aɪn		1%
	*Error	1%	
eɪ	Correct	97%	99%
	ɛ:		1%
	e:	3%	

əʊ/oʊ A-oʊ B-əʊ	A	55%	59%
	B	25%	37%
	o:	15%	2%
	ɔ:	3%	1%
	ɤʊ	1%	
	ɤo	1%	
	aɔ		1%

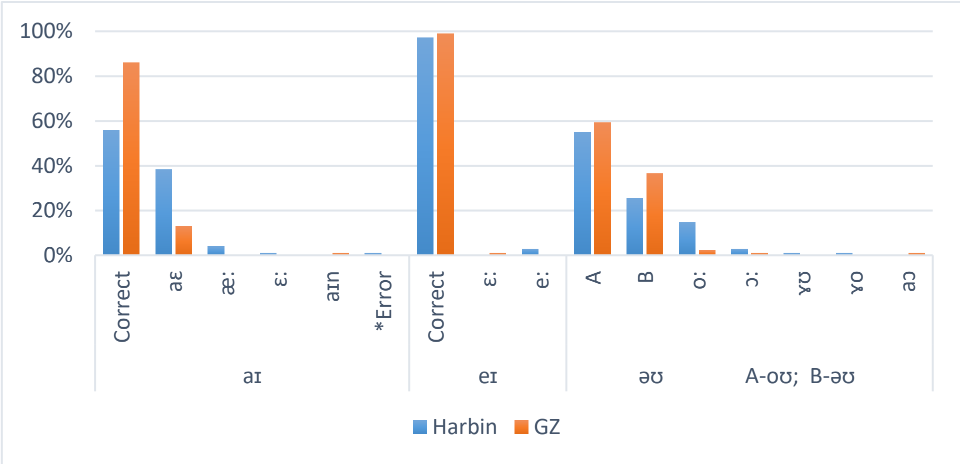


Figure 4.15 Production of English vowels /aɪ/, /eɪ/ and /əʊ (oʊ)/

From Figure 4.15, it can be clearly seen that there was no difficulty for participants in pronouncing these three diphthongs which exist in both dialects/languages, except for /aɪ/ in Harbinese. The t-test results for these diphthongs shown in Table 4.25 confirm that there are no statistical differences between two dialects/languages with the diphthongs /eɪ/ [p=0.457 (p>0.05)] and /əʊ (oʊ)/ [p=0.548 for /oʊ/; p=0.118 for /əʊ/ (p>0.05)] but there was a significant difference for the diphthong /aɪ/ [p=0.004 (p<0.05)]. The t-test results for the English varieties indicate no significant difference between Harbinese and Cantonese, and therefore the hypothesis that different dialect/language speakers prefer a specific variety of English cannot be supported.

Table 4.25 T-test results for English phonemes /aɪ/, /eɪ/ and /əʊ (oʊ)/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
aɪ	Harbin	34	55.8821	48.35391	8.29263	-2.989	0.004
	Guangzhou	31	86.0216	31.94235	5.73702		
eɪ	Harbin	34	97.0588	12.62712	2.16553	-0.749	0.457
	Guangzhou	31	98.9248	5.98624	1.07516		
oʊ-A	Harbin	34	54.9018	30.57648	5.24382	-0.605	0.548
	Guangzhou	31	59.1403	25.40165	4.56227		
əʊ-B	Harbin	34	25.4900	29.65512	5.08581	-1.585	0.118
	Guangzhou	31	36.5584	26.32518	4.72814		

The results for the remaining English diphthongs are presented in this section. The diphthongs /aʊ/ and /ɔɪ/. /aʊ/ exists in Cantonese, but Harbinese has a similar one which is /ɑʊ/. The Harbinese /ɑʊ/ is different from English /aʊ/, as it is composed of a low back unrounded /ɑ/ rather than a low central unrounded /a/ plus a high back rounded vowel /ʊ/. It can easily be seen from the results in Table 4.30 and Figure 4.18 that there was a massive disparity in accuracy rates between the Harbinese and Cantonese groups with this English diphthong /aʊ/. Cantonese speakers achieved rather excellent results pronouncing /aʊ/ with accuracy rates up to 92% while Harbinese speakers had much lower accuracy at 39%. The errors participants made can be categorised as diphthongal substitutes or monophthongal substitutes. A large number of Harbinese participants pronounced English /aʊ/ as the diphthongal Harbinese [ɑʊ], accounting for 19% of cases. This implies that L1H exerted a great influence on the L2 /aʊ/. Only small number of Cantonese speakers (2%) also made the same error. They might have been influenced by Mandarin. Another diphthongal error participants made is [ɑɔ], which was pronounced by lowering the second vowel of /aʊ/ to the place of /ɔ/ among 12% of the Harbinese group while the Cantonese group had only 2%. The error [ɑɔ] is very close to [ɑʊ], and may have developed from it. Thus, [ɑʊ] ([ɑɔ]) was regarded as the error type for participants mispronouncing /aʊ/. Another error is [oʊ] which was rare and only made by Harbinese speakers. Various monophthongs were also used by Harbinese participants such as [ɔ], [o] and [a] but [ɔ] was the only frequently used one, accounting for 25% while the others were rare. However, Cantonese speakers only used [ɔ] in 2% of cases to substitute for /aʊ/. It can be concluded that diphthongal substitution was the main error type for participants in pronouncing /aʊ/, especially for Harbinese speakers, followed by monophthongal substitutions where [ɔ] was the most popular substitute used by both dialect/language speakers.

The English diphthong /ɔɪ/ exists in Cantonese but is absent in Harbinese. Table 4.26 and Figure 4.16 show that it is not difficult for Cantonese participants to pronounce this diphthong, with a high accuracy rate of 86%. However, Harbinese participants achieved comparatively poor results with an accuracy rate of 67%, where /ɔɪ/ was frequently pronounced by 18% of them as well as by 4% of Cantonese speakers. This is similar to the case with /aɪ/, which implies that participants preferred to lower the ending vowel from /ɪ/ to [ɛ], especially among the Harbinese speakers. Another common substitute for /ɔɪ/ was [aɪ], as produced by 6% of Harbinese and 5% of Cantonese participants. In addition, another 3% of Cantonese participants pronounced /ɔɪ/ as [aɛ] with a lowered [ɪ] in [aɪ]. From Table 4.26, it can be seen that Harbinese participants adopted various substitutes for /ɔɪ/ but the Cantonese group made errors only between [ɔɛ] and [aɪ] as well as [aɛ], which is a variant of [aɪ]. Apart from the common substitutes mentioned above, some Harbinese participants inserted a glide [w] in front of the substituted diphthongs such as in [wɔɛ] and [waɪ]. Moreover, they also used a monophthong to replace the /ɔɪ/ but this accounted for a very small proportion. However, Cantonese speakers never used these two methods.

Table 4.26 Production of English vowels /aʊ/ and /ɔɪ/

Vowels	Types	Harbin	Guangzhou
aʊ	Correct	39%	92%
	ɑʊ	19%	2%
	ɑɔ	12%	2%
	oʊ	2%	
	ɔ	25%	2%
	o:	1%	
	a	2%	
	*Error	1%	1%

ɔɪ	Correct	67%	85%
	ɔɛ	18%	4%
	oɛ	1%	
	aɪ	6%	5%
	aɛ		3%
	ɛ	1%	
	ɔ:	1%	
	woɪ	2%	
	waɪ	2%	
	wɔɛ	2%	
	wɛ	1%	
	*Error		2%

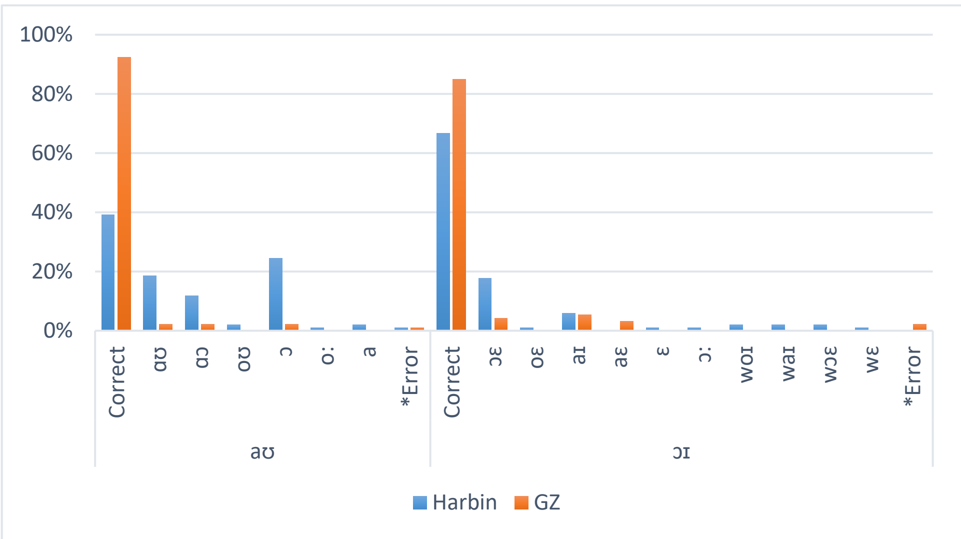


Figure 4.16 Production of English vowels /aʊ/ and /ɔɪ/

The results show speakers whose dialect/language has identical diphthongal phonemes achieve better pronunciation than those who do not or who only have phonemically similar ones. Moreover, the t-test results in Table 4.27 further confirm that these differences are statistically significant {/aʊ/ [p=0.000 (<0.05)] ; /ɔɪ/ [p=0.018 (<0.05)]}.

Table 4.27 T-test results for English phonemes /aʊ/ and /ɔɪ/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
aʊ	Harbin	34	39.2156	32.27784	5.53560	-8.228	0.000
	Guangzhou	31	92.4732	18.67817	3.35470		
ɔɪ	Harbin	34	66.6668	37.60561	6.44931	-2.453	0.018
	Guangzhou	31	84.9471	20.79614	3.73510		

Figure 4.15 and Figure 4.16 clearly indicate that Cantonese speakers achieved better overall results in pronouncing these five English diphthongs although the differences between the dialects/languages with the English diphthongs /eɪ/ and /əʊ (oʊ)/ were not so significant. Hypothesis H2.1 that Harbinese speakers would make errors with English diphthongs is supported overall. However, the results of the T-tests do not confirm that there are significant differences between the two dialects/languages with the English /eɪ/ and /əʊ (oʊ)/, while the differences with the English /aɪ/, /aʊ/ and /ɔɪ/ are statistically significant. Therefore, hypothesis H2.1 concerning English diphthongs as a whole can be supported where Harbinese speakers make more errors than Cantonese. So, it is proved again that 'identical' is easy. However, the cases of /eɪ/ and /əʊ (oʊ)/ also show that similarity does not necessarily lead to difficulty, which conflicts with Flege's idea that 'similar' is difficult.

4.3.2 Category of 'similar' in vowels

4.3.2.1 H2.2 and results

> H2.2 Harbinese lacks /ɔ/ while Cantonese has this English phoneme but with two allophones. Thus, English /ɔ/ is phonemically identical but allophocially different in Cantonese but new in Harbinese. It is hypothesised that Cantonese speakers will make errors.

English /ɔ/ is a mid-low back rounded vowel which exists in Cantonese but not in Harbinese. However, Cantonese /ɔ/ has two allophones while English has only one, thus, Cantonese is similar to English in /ɔ/ because they are not completely identical. Table 4.28 and Figure 4.17 show that the Cantonese speakers had no difficulty pronouncing this vowel, with a high accuracy rate at up to 88%, but the Harbinese speakers had a lower accuracy rate around 64%. To the errors made, Harbinese speakers were particularly likely to substitute /ɔ/ with diphthongs such as [ɑʊ] or [ɑɔ], respectively accounting for 9% and 12% of cases, and other diphthongs such as [ʊɔ] and [əɔ] making up tiny proportions. It seems that [ɑɔ] might be developed from [ɑʊ] where the second vowel segment did not reach a position as high as [ʊ]. However, although Cantonese participants also pronounced /ɔ/ as diphthongs sometimes such as [ɑɔ], this only accounted for 1% of cases. In fact, Cantonese speakers preferred to use monophthongs as substitutes for diphthongal substitutes as seen in Table 4.28. The most frequently used substitute by Cantonese speakers was the low back rounded vowel [ɒ], comprising 5% of cases. Apart from the main method of diphthongal substitution, Harbinese speakers also used monophthongal substitute but the monophthong was often rhoticised, as shown [ɔ˞] in Table 4.28, in up to 10% of cases, but this was rare among Cantonese participants. In addition, Harbinese speakers also pronounced /ɔ:/ as [ɒ], as did Cantonese speakers, but this error comprised only a small proportion of the table.

Table 4.28 Production of English vowel /ɔ/

Vowels	Types	Harbin	Guangzhou
ɔ:	Correct	64%	88%
	ɑʊ	9%	
	ɑɔ	12%	1%
	ɔ˞	10%	2%
	ɒ	3%	5%
	ɜ:		1%
	a		2%
	a:w	1%	
	ʊɔ	1%	
	əɔ	1%	

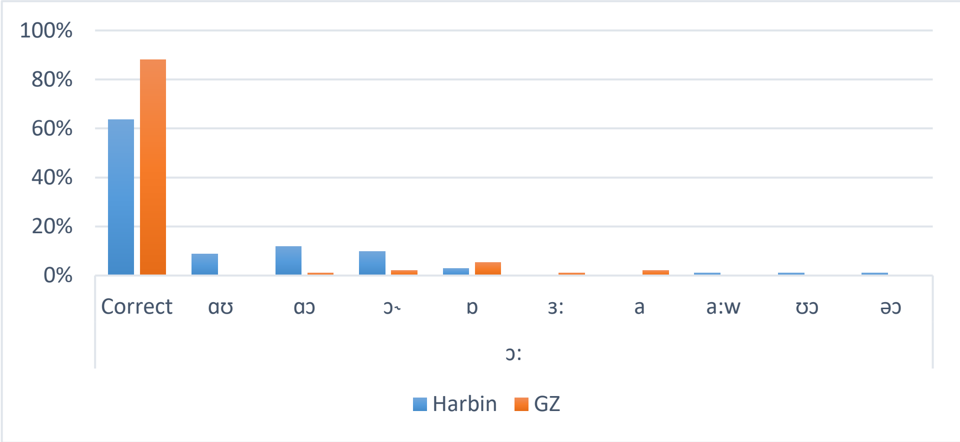


Figure 4.17 Production of English vowel /ɔ/

The above results show that /ɔ/ was not difficult to pronounce among Cantonese participants but was more difficult for the Harbinese. Therefore, it can be concluded that hypothesis H2.2 that Cantonese speakers will make errors is rejected. The T-test results for the English phoneme /ɔ/ in Table 4.29 demonstrate that these is a statistically significant difference between the two dialects/languages with the vowel /ɔ/ [p=0.001 (<0.05)]. This confirms that hypothesis H2.2 is rejected and that 'similar' (phonemically identical but allophonically different) is easy and 'new' is difficult in this context.

Table 4.29 T-test results for English phoneme /ɔ/ in Harbin and Guangzhou

Hometown		N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
ɔ:	Harbin	34	63.7259	34.19924	5.86512	-3.634	0.001
	Guangzhou	31	88.1729	18.35447	3.29656		

4.3.2.2 H2.3 and results

H2.3 The English /ɛ/ is phonemically identical and allophonicially different in Cantonese but allophonically identical and phonemically different in Harbinese. Thus, the relation is 'similar' vs. 'similar'. Based on the rules of degree of difficulty, the type of similarity with 'allophonically identical' is predicted to be more difficult than that with 'phonemically identical'. Therefore, it is hypothesised that a) both should be difficult but b) Harbinese will make more errors.

The English vowel /ɛ/ is a low front vowel, which is identical in Cantonese but Cantonese has two allophones. It means Cantonese is phonemically identical but allophonically different from English in /ɛ/, i.e. they two are similar. However, Harbinese has only an allophne [ɛ] with unidentical phoneme, so Harbinese is allophonically identical with English in /ɛ/ but phonimically different. The relation of each Chinese variety to English is 'similar' vs. 'similar', but they just belong to different types of similarity.

As seen in the bar chart in Figure 4.18, the Cantonese group achieved excellent performance in pronouncing /ɛ/ with a a accuracy rate nearly at 90%, but the Harbinese group achieved accuracy of less than 40%, as shown in Table 4.30. Harbinese participants were likely to mispronounce /ɛ/ as diphthongs such as [aɪ] cover 50% of cases or much more rarely, [aɛ]. It can be observed that the difference between the errors [aɪ] and [aɛ] among Harbinese speakers is that they did not raise their tongue front part as high as the second vowel segment [ɪ] requires. Unlike Harbinese speakers, only 1% of Cantonese participants pronounced /ɛ/ as the diphthong [aɪ]. The second most common substitute among Harbinese participants was the low front vowel [æ] (6%) but it was the most frequent substitute among Cantonese speakers at 8%. Another error was the high front lax vowel [ɪ] (3% and 2%). It seems that participants tended to lower or raise the place of articulation when mispronouncing /ɛ/. In addition, it seems that Harbinese speakers preferred to substitute /ɛ/ with a diphthong, but Cantonese speakers used a monophthong instead.

Table 4.30 Production of English vowel /ɛ/

<table>
<tr><th>Vowels</th><th>Types</th><th>Harbin</th><th>Guangzhou</th></tr>
<tr><td rowspan="5">ɛ</td><td>Correct</td><td>38%</td><td>89%</td></tr>
<tr><td>aɪ</td><td>51%</td><td>1%</td></tr>
<tr><td>aɛ</td><td>2%</td><td></td></tr>
<tr><td>æ</td><td>6%</td><td>8%</td></tr>
<tr><td>ɪ</td><td>3%</td><td>2%</td></tr>
</table>

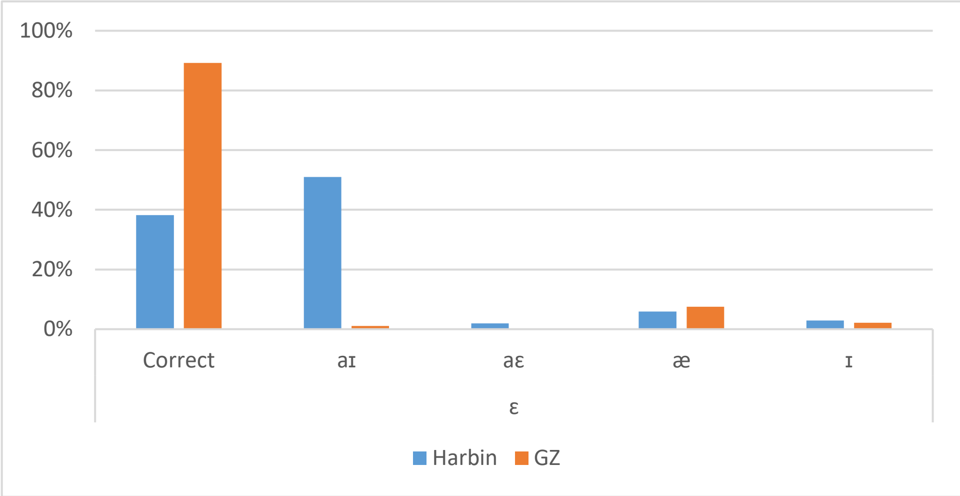


Figure 4.18 Production of English vowel /ɛ/

It can be concluded that the participants could achieve good performance in pronouncing this vowel which does exist in L1, such as Cantonese. The results further reveal that an L2 vowel as a phonemically identical but allophonically different sound in the L1 can be pronounced much better than an allophonically identical but phonemically different sound in the L1. Hypothesis H2.3a), that English /ɛ/ is difficult for both groups of Chinese speakers, is rejected by the results, because the results indicate that Cantonese speakers did not have difficulty in pronouncing /ɛ/. However, the low accuracy rate in Harbinese group indeed represent Harbinese speakers have certain difficulty. The results of the statistical test for the English phoneme /ɛ/ in Table 4.31 further confirm that Cantonese did not have difficulties but Harbinese did, because the difference between the two groups of participants is highly statistically significant [$p=0.000$ ($p<0.05$)]. The hypothesis H2.3b) that Harbinese speakers will make more errors was supported. It seems to be contradictory to the results of 'similar' because one shows that 'similar' is easy while the other seems to support Flege's SLM that 'similar' is difficult. However, the seemingly contradictory results indeed confirm that the difficulty among different types of similarity is different. It can be seen from this case that the type 'phonemically identical but allophonically different' of similarity is much easier than the similar type 'allophonically identical but phonemically different'. As to the two types of 'similar' that one has difficulty but the other does not, it needs to be discussed in detail in chapter 5 discussion to see why it happened in this way.

Table 4.31 T-test result for English phonemes /ɛ/ in Harbin and Guangzhou

Hometown		N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
ɛ	Harbin	34	38.2356	35.89482	6.15591	-6.997	0.000
	Guangzhou	31	89.2474	21.75168	3.90672		

4.3.2.3 H2.4 and results

> H2.4 Harbinese is phonemically identical but allophonically different from the English phoneme /ə/, while Cantonese does not have this phoneme. Thus the Harbinese will make errors.

The English /ə/ is a phoneme in Harbinese but Harbinese has four allophones while English has only one. Thus, it is phonemically identical to but allophonically different from Harbinese. However, it does not exist at all in Cantonese. Therefore, the English /ə/ is similar in Harbinese and new in Cantonese. According to Flege’s SLM, it is predicted that the Harbinese will make errors. However, we can see from Table 4.32 and Figure 4.19 that although Cantonese does not have this vowel, Cantonese speakers can pronounce it more accurately than Harbinese with an accuracy rate of 94%. The Harbinese participants also preferred to rhoticise the schwa as the r-coloured vowel [ɚ] in 7% of cases. Although Cantonese speakers also made this type of error, it only accounted for 2% of cases, possibly reflecting Mandarin influence. Some Harbinese speakers (4%) substituted /ə/ with a low vowel [a], whereas a few Cantonese participants substituted its counterpart central long vowel [ɜ] or a back high vowel [u]. Harbinese participants also inserted the lateral consonant [l] after /ə/ to form a syllable such as [əl], but [l] was also vocalised as [w] after /ə/ by some Harbinese speakers. However, no Cantonese speakers made this mistake.

Table 4.32 Production of English vowel /ə/

Vowels	Types	Harbin	Guangzhou
ə	Correct	84%	94%
	ɚ	7%	2%
	əw	2%	
	əl	3%	
	ɜ		2%
	u		2%
	a	4%	

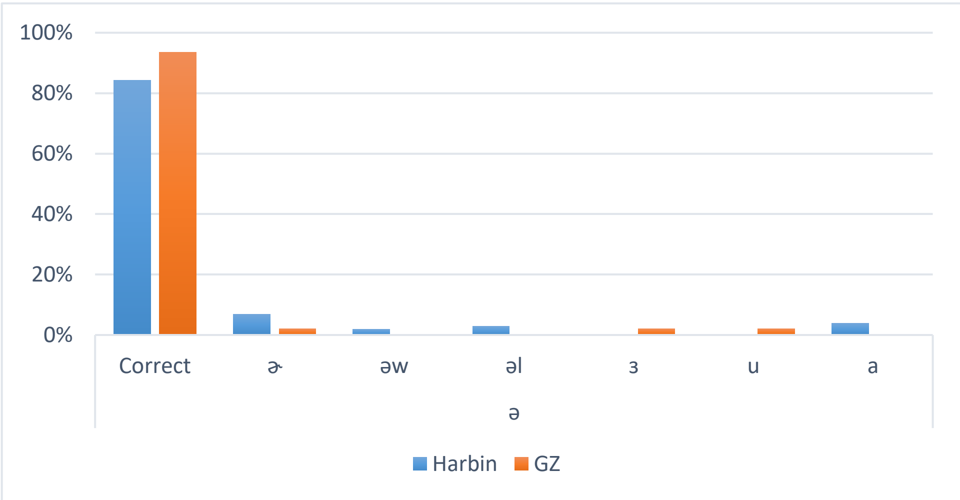


Figure 4.19 Production of English vowel /ə/

From the above results, it seems that hypothesis H2.4 is supported because the Harbinese speakers exhibited worse production than the Cantonese even though the English /ə/ exists in Harbinese and is lacking in Cantonese. Moreover, the results of the statistical tests for /ə/ shown in Table 4.33 confirm that the difference between the two dialects/languages is statistically significant with a p-value of 0.046 ($p < 0.05$). Therefore, the hypothesis H2.4 that the Harbinese will make errors but the Cantonese will not is completely supported. This agrees with Flege's idea that 'similar' is difficult. Why this result supports Flege's idea that similarity is difficult while other such hypotheses are rejected, and whether or not there are other influences which lead to these results needs to be researched and discussed further.

Table 4.33 T-test results for English phoneme /ə/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
ə	Harbin	34	84.3147	20.49056	3.51410	-2.039	0.046
	Guangzhou	31	93.5487	15.91462	2.85835		

4.3.2.4 H2.5 and results

H2.5 The English vowel phoneme /ɑ/ is allophonically identical but phonemically different in Harbinese. Thus, the English /ɑ/ is similar in Harbinese but new in Cantonese. According to Flege's SLM, the 'similar' is more difficult than the 'new', and so Harbinese speakers will make errors and the Cantonese will not.

The English /ɑ:/ is a low back unrounded vowel which occurs in Harbinese as an allophone but does not exist in Cantonese. It can be seen from Table 4.34 and Figure 4.20 that both groups of participants had problems in pronouncing it. Their accuracy rates were approximately 40% and most participants pronounced it as the slightly centralised front low vowel [a] (59% in each group). It was also occasionally pronounced by both groups of participants as the low front [æ]. Rarely, 1% of Harbinese speakers pronounced /ɑ/ as the rhoticised mid-high back unrounded vowel [ɤ˞].

Table 4.34 Production of English vowel /ɑ/

Vowels	Types	Harbin	Guangzhou
ɑ:	Correct	38%	40%
	a	59%	59%
	æ	2%	1%
	ɤ˞	1%	

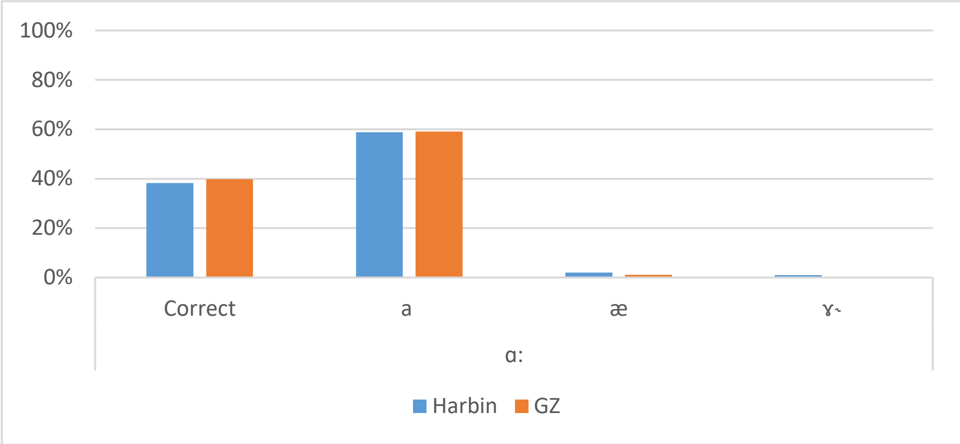


Figure 4.20 Production of English vowel /ɑ/

It seems that the hypothesis is supported by the results in that the Cantonese group achieved higher accuracy than the Harbinese group. However, Figure 4.20 clearly indicates that there is no big differences between the two dialects/languages in pronouncing this English /ɑ:/ with accuracy and without error. However, the Cantonese accuracy only slightly higher by 2%. The results of the statistical testing of the English phoneme /ɑ:/ confirm that there is no significant difference between the two dialects/languages as shown in Table 4.35 [p=0.879 (>0.05)]. Therefore, hypothesis H2.5 that Harbinese speakers make errors but Cantonese speakers will not must be rejected. Both groups have difficulties leading to the low accuracy rates. This implies that 'new' is difficult, but that the type of 'similar' with sounds which are 'allophonically identical but phonemically different' is also difficult because the two groups achieved similar accuracy rates.

Table 4.35 T-test results for English phoneme /ɑ:/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
ɑ:	Harbin	34	38.2353	41.94983	7.19434	-0.152	0.879
	Guangzhou	31	39.7845	39.83283	7.15419		

4.3.2.5 H2.6 and results

H2.6 Harbinese does not have the English /ʌ/ but Cantonese has the phonemically similar /ɐ/, so Cantonese speakers will make errors with /ʌ/.

The mid-low central vowel /ʌ/ in English does not exist in Harbinese but it is very similar to the Cantonese vowel represented by the IPA symbol [ɐ] which is also a mid-low central vowel. From the bar chart in Figure 4.21 it can easily be seen that there was a large disparity in accuracy rates between the two groups and the Harbinese participants had greater difficulty in pronouncing /ʌ/. From Table 4.36, 35% of Harbinese speakers pronounced /ʌ/ correctly but the percentage was nearly double in the Cantonese group (63%). When errors were made, /ʌ/ was predominantly pronounced as the front low vowel [a] by speakers of both dialects/languages. Also, this represented 62% of Harbinese errors which was double the proportion of Cantonese errors. A small percentage of participants from both groups pronounced /ʌ/ as [ɔ] (4% of Cantonese errors; 2% of Harbinese errors). Other substitutions for /ʌ/ were rare, such as [æ] pronounced by 1% of Harbinese speakers and [ɑ] by 1% of Cantonese speakers.

It seems that both groups had major problems in pronouncing /ɑ:/ and /ʌ/, except for the use of /ʌ/ by Cantonese participants. The Cantonese group could pronounce /ʌ/ much more accurately than the Harbinese group. Also, [a] was the most frequently used substitute to replace /ɑ:/ and /ʌ/ among both groups and especially in the Harbinese group.

Table 4.36 Production of English vowel /ʌ/

Vowels	Types	Harbin	Guangzhou
ʌ	Correct	35%	63%
	a	62%	31%
	æ	1%	
	ɔ	2%	4%
	ɑ		1%

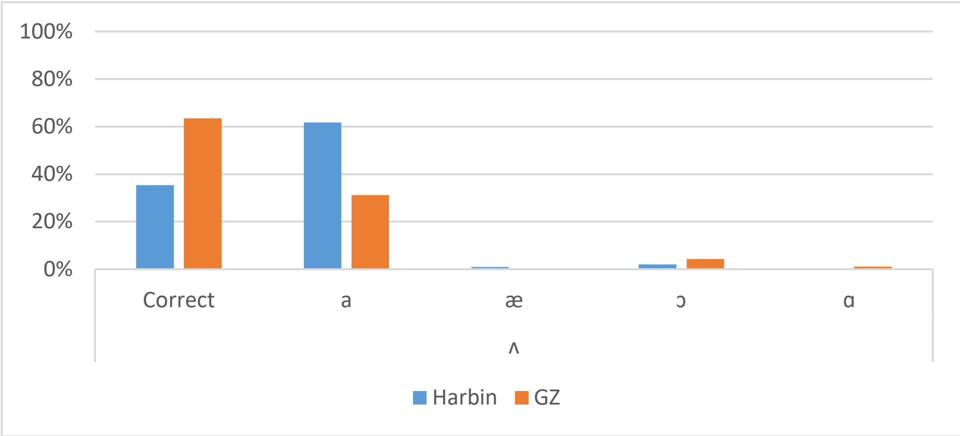


Figure 4.21 Production of English vowel /ʌ/

It can be concluded that both groups had difficulties in pronouncing the English /ʌ/, but the Cantonese group achieved much better results. Thus, hypothesis H2.6, that Cantonese speakers will make errors but Harbinese speakers will not cannot be supported. Moreover, the results of the statistical tests for the English phoneme /ʌ/ shown in Table 4.37 show a very significant disparity between speakers of the two dialects/languages [p=0.001 (p<0.05)]. This further confirms that the hypothesis H2.6 must be rejected. Harbinese speakers made serious errors pronouncing /ʌ/; meanwhile, Cantonese speakers also made errors, but far fewer than the Harbinese. This result supports the idea that 'new' for the Harbinese group is difficult while 'phonemically similar' is much easier, which does not support Flege's SLM-based predictions.

Table 4.37 T-test results for English phoneme /ɑ:/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
ʌ	Harbin	34	35.2929	27.14452	4.65525	-3.542	0.001
	Guangzhou	31	63.4410	35.85521	6.43979		

4.3.2.6 H2.7 and results

H2.7 Harbinese does not have the English /ɒ/ but Cantonese has the phonemically similar /ɔ/, so Cantonese speakers will make errors with /ɒ/.

The vowel /ɒ/ does not exists in both dialects/languages, but Cantonese has the phonemically similar /ɔ/. Moreover, as mentioned above, /ɒ/ only exists in British English, so we can not only trace the acquisition of /ɒ/ in the two dialects/languages, but also test which variety of English participants prefer to pronounce, either the British English [ɒ] or the American English [ɑ].

The results in Table 4.38 and Figure 4.22 show that a majority of participants pronounced the vowel in the British way but a very small number of participants used the American. Harbinese speakers had a much lower accuracy rate than the Cantonese. Although there is no /ɒ/ in Cantonese, it seems that Cantonese speakers had no difficulty and their accuracy rate was 84%. The most frequently used error by Cantonese speakers was to use [ɔ]. Similarly, [ɔ] was also the most frequently pronounced substitute by 29% of Harbinese speakers. The second most commonly used substitute for /ɒ/ was the high-mid back unrounded vowel [ɤ] (14% in the Harbinese and 3% in Cantonese groups) which exists in Harbinese as an allophone of /ə/. Apart from these two popular substitutes, Harbinese speakers also used other back vowels to substitute for /ɒ/, such as [ɔ˞] and [o] but they accounted for a very small percentages. In addition, the mid central vowels [ə] and [ɜ] were also used to replace the /ɒ/ by both groups, but the Harbinese used the rhotised [ɝ] instead of [ɜ]. All these central substitutes represented very small percentages. Besides these monophthongal substitutions, diphthongal substitutions were adopted by Harbinese but not Cantonese speakers, such as [ɑɔ] (5%) and [ɑʊ] (3%).

Table 4.38 Production of English vowel /ɒ/

Vowels	Types	Harbin	Guangzhou
ɒ A-ɑ; B-ɒ	A	3%	1%
	B	37%	84%
	ɔ	29%	10%
	ɔ˞	1%	
	o	2%	
	ɤ	14%	3%
	ɑɔ	5%	

	ɑʊ	3%	
	ə	4%	1%
	ɝ	2%	
	ɜː		1%

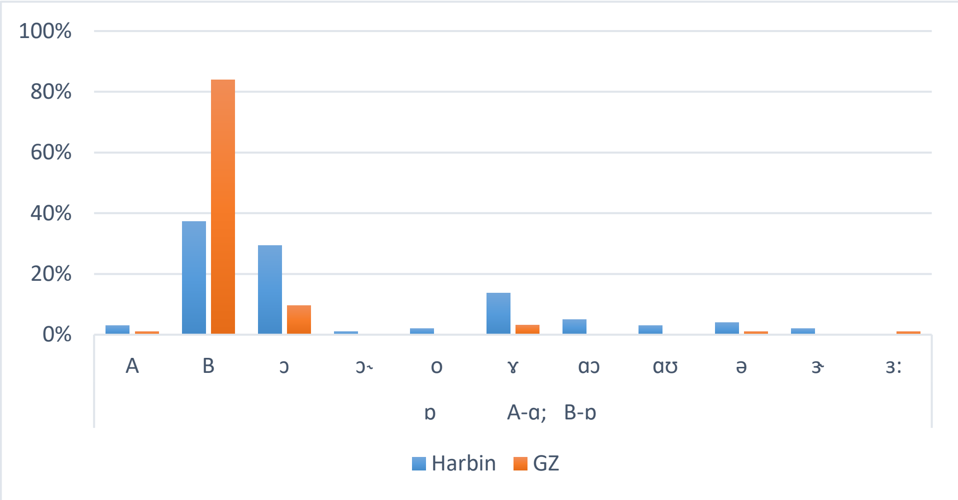


Figure 4.22 Production of English vowel /ɒ/

From the above results, it can be concluded that hypothesis H2.7 cannot be supported because Cantonese speakers do not have difficulty in pronouncing /ɒ/. Instead, Harbinese speakers have serious problems. Moreover, the results of the t-test for the vowel /ɒ/ shown in Table 4.39 indicate that the difference in the /ɒ/ (British variety) between the two dialects/languages is statistically significant [p=0.000, (p<0.05)]. Therefore, H2.7 must be rejected. This indicates again that ‘new’ is more difficult than ‘similar’ and the results disagree with the predictions of Flege’s SLM.

From the results, it can also be seen that both groups of participants predominantly preferred the British English variety of /ɒ/ to the American variety /ɑ/. Moreover, the t-test results for the British /ɒ/ show that Cantonese speakers have a tendency to use the British variety more than Harbinese speakers with a significant difference between the two dialects/languages [p=0.000, (p<0.05)] (see Table 4.39). However, no statistically significant difference is found in the t-test results for American English /ɑ/ [p=0.457, (p>0.05)].

Table 4.39 T-test results for English phoneme /ɒ/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
ɑA	Harbin	34	2.9409	12.62559	2.16527	0.749	0.457
	Guangzhou	31	1.0752	5.98624	1.07516		
ɒB	Harbin	34	37.2550	31.53299	5.40786	-6.643	0.000
	Guangzhou	31	83.8719	24.14555	4.33667		

4.3.2.7 H2.8 and results

> H2.8 Cantonese does not have the English /ɜ/ or /ɝ/ but Harbinese has the phonemically similar /ə/ (/ɚ/), so Harbinese speakers will make errors with the English /ɜ/ (/ɝ/).

The English vowel /ɜ/ is a central vowel like schwa. It is a vowel in the British variety of English because the letter *r* in orthographical representations of the vowel /ɜ/ such as *ur, er* and *ir* is never pronounced. On the other hand, /ɜ/ is represented by /ɝ/ in American English because the postvocalic rhotic *r* is pronounced. Neither Chinese dialect/language has the English vowel /ɜ/, but Harbinese has a similar one, /ə/ or /ɚ/. Thus, based on Flege's SLM, the hypothesis is proposed that Harbinese speakers will make errors with /ɜ/. This vowel is considered not only to test this hypothesis, but also to see which English variety the two groups of participants would use.

From Figure 4.23, it can be clearly seen that an overwhelming majority of participants on both sides pronounced /ɜ/ with the rhotic, especially the Harbinese group at 98%, as shown in Table 4.40. Only 2% of Harbinese speakers pronounced it as British English. However, the Cantonese group has more participants (18%), which pronounced it in the British variety. The Cantonese made more errors than the Harbinese and they used [ɔ:] to substitute for /ɜ/, but this did not happen in the Harbinese group. These errors are not significant because they comprise very small percentages.

Table 4.40 Production of English vowels /ɜ/ and /ə/

Vowels	Types	Harbin	Guangzhou
ɜ: A-ɝ; B-ɜ:	A	98%	80%
	B	2%	18%
	ɔ:		1%
	ɔ:ɜ		1%

It can be concluded that there were no problems for either groups of participants in pronouncing /ɜ/, even though it is lacking in both dialects/languages. Also, speakers of both dialects/languages preferred to pronounce /ɜ/ in the American variety of English /ɝ/, especially the Harbinese group.

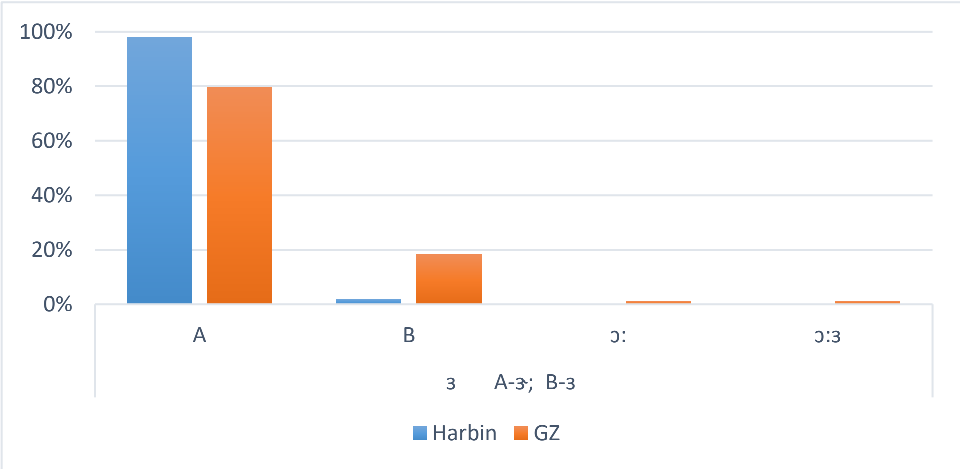


Figure 4.23 Production of English vowel /ɜ/

From the above results, it can be concluded that hypothesis H2.8, that the Harbinese speakers will make errors, is rejected. Harbinese speakers had no difficulty in pronouncing the vowel and they achieved a perfect accuracy rate. Flege's idea that 'similar' is difficult did not apply; however, 'new' is difficult also did not apply to the Cantonese group. This conclusion is rather similar to that for /ə/. A possible explanation is that the influence of Mandarin for Cantonese speakers means that they have acquired /ə/ in that variety, presumably at a young age. Like Harbinese speakers, Cantonese speakers also produced /ɜ/ accurately.

The results of the t-test for the English varieties of /ɜ/ and /ɝ/ as shown in Table 4.41 indicates that differences in the selection of English varieties between the two dialects/languages are statistically significant. Therefore, it can be claimed that although a majority of speakers in both groups preferred to speak the American variety for this English target, significantly, more Harbinese speakers statistically preferred to speak the American variety [p=0.006, (p<0.05)] compared to Cantonese speakers. In addition, although a small number of speakers in both groups spoke the British variety for this target, the statistical test shows that Cantonese speakers spoke the British variety significantly more [p=0.007, (p<0.05)]. This further indicates that Harbinese speakers significantly preferred American English while Cantonese speakers had a tendency to speak British English.

Table 4.41 T-test results for English variety selection for /ɜ/ and /ɝ/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
ɝA	Harbin	34	98.0394	7.96029	1.36518	2.945	0.006
	Guangzhou	31	79.5700	34.07799	6.12059		
ɜB	Harbin	34	0.9803	5.71605	0.98029	-2.881	0.007
	Guangzhou	31	17.2042	30.87756	5.54577		

4.3.2.8 H2.9 and results

H2.9 Harbinese does not contrast tense-lax but Cantonese does. Both Chinese dialects/languages have the English tense vowels /i, u/ while Harbinese does not have the lax counterparts /ɪ, ʊ/ but Cantonese has [ɪ, ʊ] as allophones. Therefore, each Chinese varietyis similar to English in a different way. According to Flege's SLM where 'similar' is difficult, it is predicted that both are difficult but the Cantonese will make more errors.

Both Chinese dialects/languages have the tense vowels /i, u/ as mentioned above. Whether or not the tense and lax vowel contrast in English can be acquired depends on the participants acquiring the English lax vowels /ɪ, ʊ/, which are allophones in Cantonese, but are lacking in Harbinese. The results for these vowels would be represented in the tense and lax groups as follows.

The sounds /i, u/ are two high tense vowels in English. They both exist in the two dialects/languages. From Table 4.42 and Figure 4.24, it can be clearly seen that, although /i/ exists in both dialects/languages, the accuracy rates were not very high, especially for the Cantonese group at only 32%. The Harbinese group had double the accuracy rate of the Cantonese at approximately 66%. The participants tended to mispronounce /i/ as its counterpart lax vowel [ɪ]. All errors for /i/ made by Cantonese participants belonged to this type. However, besides the substitution of [ɪ] used by 32% of Harbinese participants, a very small number of participants pronounced /i/ instead as the diphthong [eɪ].

The results indicate that the back tense vowel /u/ was not difficult for speakers of either Chinese variety to pronounce although they did make some errors. The accuracy rates of the tense /u/ were higher than those of tense /i/, reaching nearly 76% on average. The Harbinese group was slightly more accurate than the Cantonese with /u/. The main error involved confusion between /u/ and /ʊ/, where the tense /u/ was often mispronounced as the lax /ʊ/, which was the same for the tense vowel /i/. Apart from the erronerous /ʊ/ participants substituted most for /u/, Harbinese speakers also pronounced /u/ as a diphthong; for example, /ʊə/, /ʊɔ/ or /oʊ/. Among these diphthongal substitutions, /ʊə/ was most frequently pronounced, comprising 11% of cases. This error might be caused by confusion of the correct pronunciations of the target word *pool* and the orthographically similar word *poor*. Interestingly, this type of error never occurred in the Cantonese group. Besides these two popular substitutions, some Harbinese participants also replaced /u/ with other vowels such as [ɔ:] and [o:] as seen in Table 4.42, but the errors accounted for very small percentages.

Table 4.42 Production of English vowels /i/ and /u/

Vowels	Types	Harbin	Guangzhou
i	Correct	66%	32%
	ɪ	32%	68%
	eɪ	2%	
u	Correct	77%	75%
	ʊ	8%	24%
	ʊə	11%	
	ʊɔ	1%	
	ɒ		1%
	ɔ:	1%	
	o:	1%	
	oʊ	1%	

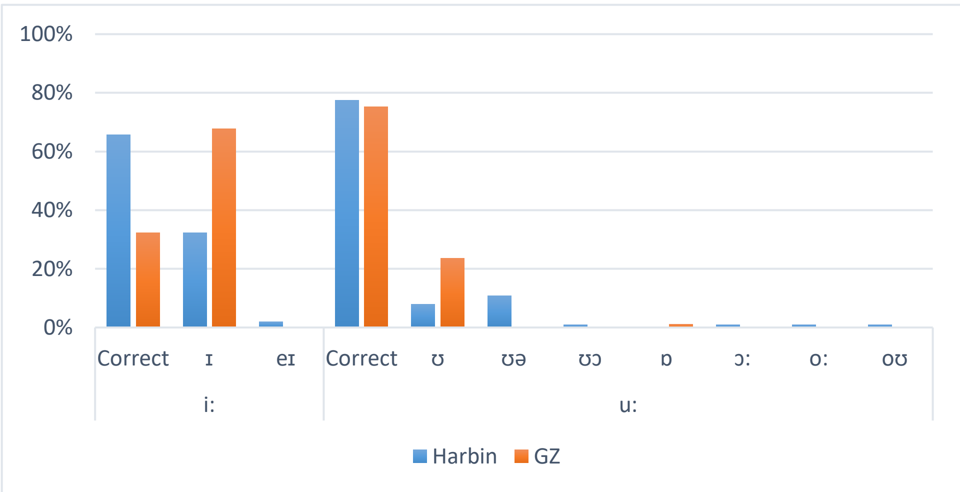


Figure 4.24 Production of English vowels /i/ and /u/

The two high English lax vowels /ɪ/ and /ʊ/ are allophones in Cantonese but do not exist in Harbinese. Lax vowels would be the key for participants to acquire the tense-lax contrast. For the English lax high front vowel /ɪ/, the results show that a large percentage of participants pronounced it correctly, respectively comprising 70% in the Harbinese group and 85% in Cantonese, as shown in Table 4.43 and Figure 4.25. The Cantonese speakers achieved better results than the Harbinese. In the same way as /i/ was mispronounced as its counterpart [ɪ] by participants, /ɪ/ was pronounced as [i] by 15% of Cantonese participants and 26% of Harbinese participants.

Substitution with /i/ was the commonest error. However, 3% of Harbinese participants pronounced /i/ as a diphthong [eɪ] but the Cantonese never did.

It can be seen from the results in Table 4.43 that both groups had difficulties in pronouncing the high back lax vowel /ʊ/ because the accuracy rates were not very high, but the Cantonese group were more accurate with rates of 59% versus 42%. The erroneous use of [u] was the main substitute among the Harbinese and the only substitution made by the Cantonese. The error rate of the substitution of [u] for /ʊ/ was 53% in the Harbinese group and 41% in the Cantonese. Apart from these substitutions Harbinese speakers also used [ʊə] and [a] but they accounted for very small percentages.

Table 4.43 Production of English vowels /ɪ/ and /ʊ/

Vowels	Types	Harbin	Guangzhou
ɪ	Correct	70%	85%
	i	26%	15%
	eɪ	3%	
	*Error	1%	
ʊ	Correct	42%	59%
	u	53%	41%
	ʊə	2%	
	a[11]	3%	

[11] [a] seems to be an interesting error here. It is worth noting that it could be influenced by the mispronunciation of spelling of other words in their minds. The test token is 'bush' presented in Chinese. Participants may think 'u' in 'bush' has the same sound of 'u' in the 'bus' in their minds. Also, /a/ was often used to substitute for /ʌ/ in 'bus'. That might be the reason why [a] was produced here.

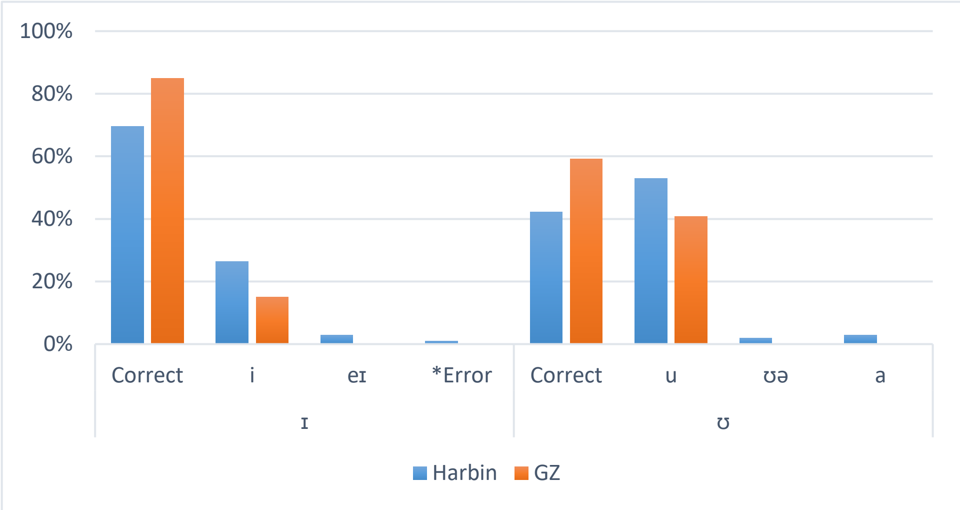


Figure 4.25 Production of English vowels /ɪ/ and /ʊ/

From the above results, we can conclude that both groups of participants made errors mainly with tense and lax vowels. Tense vowels were often substituted by their lax counterparts and vice versa. This was particularly prominent in the Cantonese group. Apart from this error type, the Harbinese group also used diphthongal substitutions but not used as often as substitutions between tense and lax pairs. Although tense vowels exist in both dialects/languages, it seems that they also made errors with them. It is salient in the tense vowel /i/ for Cantonese speakers.

The acquisition of lax vowels plays a key role in acquiring tense-lax contrast. Therefore, whether or not the hypothesis tested was supported or rejected, the results for the lax vowels are significant. It can be seen from Figure 4.25 that Cantonese speakers had better pronunciation of lax vowels than Harbinese speakers. Thus, it can be said that the predictions of H2.9 were not exactly rejected. The prediction of both being difficult was supported but the prediction that Cantonese will make more errors was rejected since for both groups, their error frequencies were comparable'. The results of the T-test for /ɪ, ʊ/ indicate that there are statistically significant differences between the two dialects/languages [$p=0.008$ for /ɪ/ and $p=0.022$ for /ʊ/ ($p<0.05$)], as shown in Table 4.44. This proves that the Harbinese speakers have more difficulties in pronouncing the lax vowels /ɪ, ʊ/ than the Cantonese, which leads to the further rejection of hypothesis H2.9 that the Cantonese will make errors. Although Cantonese speakers have more difficulties in pronouncing the tense /i/ than the Harbinese group, the t-test results also show a significant difference. The Harbinese group have more difficulties with both lax vowels while the Cantonese only have so many difficulties with one tense vowel. Therefore, hypothesis H2.9 that speakers of both dialects/languages make errors and that the Harbinese make comparatively more errors is rejected. This indicates that, in this context, ‘similar’ is not difficult but ‘new’ is.

Table 4.44 T-test results for the English phonemes /i, u/ and /ɪ, ʊ/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
i	Harbin	34	65.6862	36.22423	6.21240	3.728	0.000
	Guangzhou	31	32.2577	35.98802	6.46364		
u	Harbin	34	77.4532	15.82702	2.71431	0.425	0.673
	Guangzhou	31	75.2697	24.29454	4.36343		
ɪ	Harbin	34	69.6091	22.27489	3.82011	-2.759	0.008
	Guangzhou	31	84.9468	22.50718	4.04241		
ʊ	Harbin	34	42.1565	27.59743	4.73292	-2.350	0.022
	Guangzhou	31	59.1403	30.68428	5.51106		

4.3.3 Category of ‘new’ in vowels

4.3.3.1 H2.10 and results

> H2.10 The English vowel phoneme /æ/ is absent from both dialects/languages. According to Flege’s SLM, ‘new’ is easy; therefore, a) neither will make errors with these segments; b) if they do make errors, there will be no differences between the two dialects/languages.

Neither Chinese variety has the English low front vowel /æ/, and the results in Figure 4.26 show that they had problems pronouncing this vowel. Although a certain number pronounced /æ/ correctly, the accuracy rates for both groups were less than one-third. From Table 4.45 it can be seen that the errors made included several different substitutes. Among these, the most frequently used substitute was the mid-low front vowel [ɛ] which was adopted by 45% of Harbinese participants and up to 75% of Cantonese participants. It seems that the participants preferred to use a similar segment in their dialects/languages to replace it, which is most evidence in the result for the Cantonese. Approximately one-fifth of Harbinese participants pronounced /æ/ as the diphthong [aɪ] but no Cantonese participants made this sort of error. Apart from these errors, there were 1% of participants from each group who substituted /æ/ with the diphthong [eɪ] and 1% of Harbinese participants who replaced it with [a].

Table 4.45 Production of English vowel /æ/

Vowels	Types	Harbin	Guangzhou
æ	Correct	31%	24%
	ɛ	45%	75%
	aɪ	22%	
	a	1%	
	eɪ	1%	1%

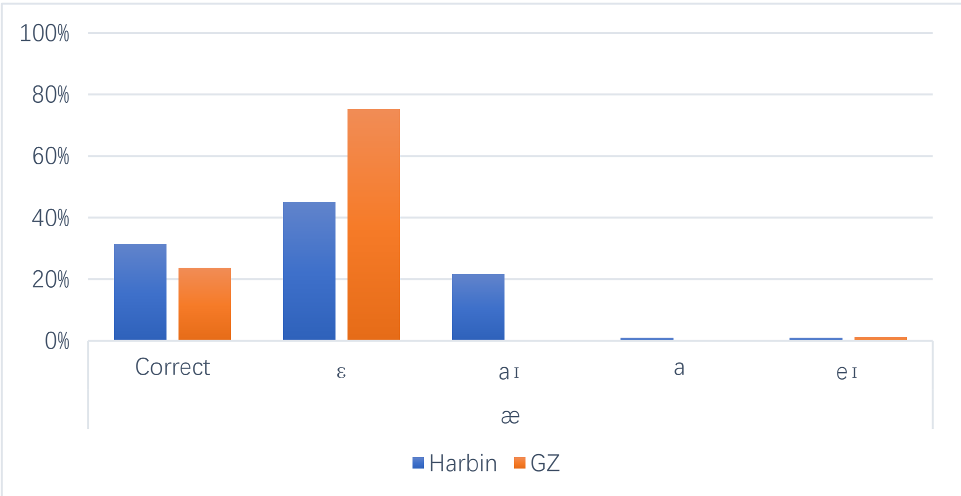


Figure 4.26 Production of English vowel /æ/

The results show that both groups had great difficulty in pronouncing /æ/ and made serious errors, which means that hypothesis H2.10a, neither will make errors, is rejected. Moreover, they made different types of errors, and therefore hypothesis H2.10b that there will be no differences in errors, is also rejected. However, the results of the statistical test for /æ/ in Table 4.46 do not show a significant difference between the two dialects/languages in accuracy rate [p=0.387, (p>0.05)]. Therefore, hypothesis H2.10 is rejected but not confirmed supported. There are some differences between the two groups although it is difficult for both groups. The findings contradict the suggestion from Flege's SLM that 'new' is easy, because here 'new' is obviously difficult.

Table 4.46 T-test results for English phoneme /æ/ in Harbin and Guangzhou

Hometown		N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
æ	Harbin	34	31.3724	37.55280	6.44025	0.870	0.387
	Guangzhou	31	23.6555	33.54790	6.02538		

4.3.4 Summary of the results of hypotheses relating to vowels

It can be seen from the results above that both groups have difficulties in English vowels /æ, ɑ, ʌ/ which are new in both dialects/languages. Besides these vowels, Harbinese have difficults in more English vowels such as /ɔ, ɒ, ɪ, ʊ, ɛ, ɔ, ɔɪ/ which are also absent in Harbinese while Cantonese have difficulties in English vowels /i, u, ɪ, ʊ/ although Cantonese has tense phonemes /i, u/ and lax allophones [ɪ, ʊ]. The results indicate that the Cantonese group has a better performance in pronunciation of English vowels than the Harbinese group.

Substitution is the most common error type for both groups, followed by insertion but it is rarely used. Harbinese are apt to substitute monophthongs with diphthongs which happened more frequently than Cantonese. It seems to be difficult for Harbinese to distinguish /ɛ, æ and aɪ/ but /ɛ, æ/ for Cantonese. English diphthongs are often monophthongised by both groups but more in Harbinese.

The hypothesis H2.1 in identical category was supported, which provides evidence that 'identical' is easy. Hypotheses in the category of similarity from 2.2 to 2.9 with rejected results mean that most hypotheses support 'new' is difficult but not 'similar' except H2.4. As to the different types of similarity, it can be summarised that the type of 'phonemically identical and allophonically different' seems to be easier than the type of 'allophonically identical and phonemically different' which is easier than the type of 'phonemically similar' and 'phonemically similar' is easier than 'new'.

4.4 Results of hypotheses relating to syllable structure

Chinese has a simpler syllable structure than English. This section shows the results for some specific English syllable structure which are identical or similar to either of the two Chinese dialects/languages. The results can indicate the degree of influence of different Chinese dialects/languages on the acquisition of English syllable structure.

4.4.1 Category of 'identical' in syllable structure

4.4.1.1 H3.1 and results

> H3.1 Cantonese does not have C+glide onset but Harbinese and English have in the syllable /CC_GV(C)/. Thus, Cantonese but not Harbinese speakers will make errors with English syllables with glides such as in /CjV(C)/ (mostly in RP) and /CwV(C)/.

The onsets /Cj/ and /Cw/ exist in Harbinese and English syllable structure but do not exist in Cantonese as mentioned in the previous chapters. Thus, tests were concluded to see if Cantonese speakers can pronounce English with C+glide onset in the syllable structure. Thus, to the pattern /CC_GV(C)/, only onset /CC_G/ was regarded as the main target and analysed. The glides /j/ and /w/ in the two English syllable patterns are constrained according to their surrounding environment. For example, the vowel after /j/ in /CjV/ must be /u/ or /ʊ/ in most cases and the consonants before /w/ in /CwV/ are limited to alveolar stops or fricatives and velar stops in English such as /t/, /d/, /s/, /k/ and /g/.

It can be clearly seen from Figure 4.27 that neither group of participants had difficulties in pronouncing English words with these two syllable structures. The accuracy rates reached as high as 96% as shown in Table 4.47. For the pattern /CjV/, two types of errors were made. This glide /j/ was often pronounced as the high front lax vowel [ɪ] by certain participants, especially in the Cantonese group. All errors made by Cantonese speakers were of this type. The other type involved not the glide /j/ but its neighbouring surrounding vowel. Harbinese speakers pronounced /j/ correctly but the vowel was changed from /u/ or /ʊ/ to [əʊ] or [ə] in [Cjəʊ] and [Cjə]. If the consonants and constrained vowels were not taken into consideration, it could be said that the Harbinese group achieved a remarkably good performance, producing this syllable structure with only 2% of errors.

The accuracy rates for the syllable pattern /CwV/ were lower at 88% in the Harbinese and 90% in the Cantoense group. Various types of errors were made, especially by the Harbinese group. Errors with /CwV/ can be classified into three types. The first is the substitution of /w/ with [ɹ]. Most errors made by the Cantonese speakers belonged to this type, accounting for approximately 8%. Fewer Harbinese speakers made these errors (2%). Another error type is the insertion of a vowel between the consonant and /w/. This was only made by the Harbinese, accounting for 3%. Vowels inserted were [ə] and [u], with the former used more frequently. The third type of error is substitution plus insertion. It seems that it combines the first two types, in that /w/ was first substituted by [ɹ] and then the vowel [ə] or [u] was inserted between the consonant and [ɹ]. [ə] is the most frequently epenthesied vowel in this syllable pattern. From Table 4.47, we can see epenthesis was frequently used by Harbinese speakers while substitution of [ɹ] for /w/ was often used by the Cantonese group.

Table 4.47 Production of English syllable structure /CjV/ and /CwV/

Patterns	Types	Harbin	Guangzhou
CjV	Correct	96%	94%
	Sub-Cɪw	2%	4%
	Sub-Cɪu		2%
	njəʊ	1%	
	bjə	1%	

CwV	Correct	89%	90%
	Sub-gɹɪ	2%	5%
	Sub-kɹɛ		3%
	Ins-gəwɪ	2%	
	Ins-guwɪ	1%	
	S+I-gəɹɪ	4%	
	S+I-guɹɪ	1%	1%
	*Error	1%	

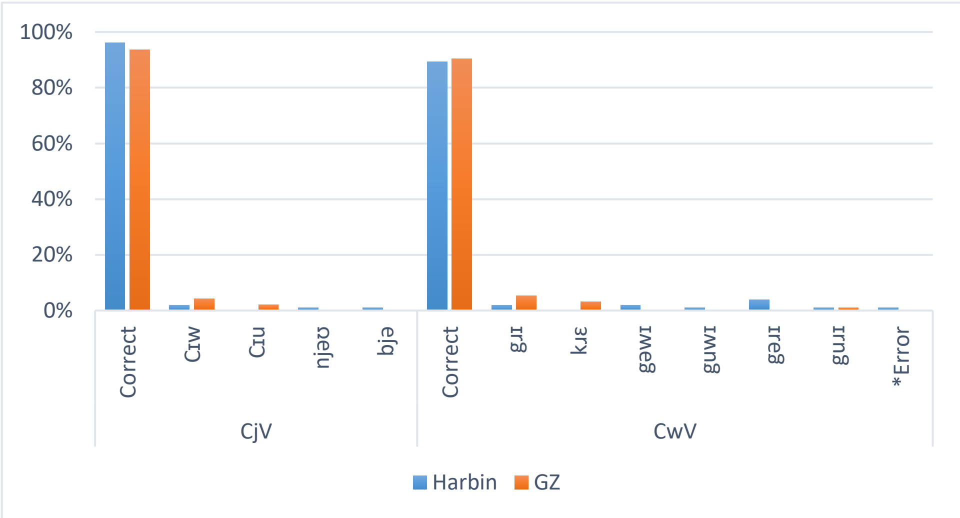


Figure 4.27 Production of English syllable structure /CjV/ and /CwV/

From the above results, it is concluded that hypothesis H3.1 cannot be supported because the English syllable structure with /Cj/ or /Cw/ onset is not very difficult for either group of participants who achieved high accuracy rates. Although there are differences between two dialects/languages, they do not seem very important. The results of the t-test in Table 4.48 for these English syllable structures show a statistically significant difference [p=0.493 for /CjV/; p=0.640 for /CwV/ (p>0.05)]. This further confirms that the hypothesis H3.1 must be rejected. The results indicate that not only the ‘identical’ is easy, but also the ‘new’ is easy in this case because there is no significant difference between the two groups. This seems to contradict other findings where the ‘new’ is difficult. Thus, this conclusion should be further discussed.

Table 4.48 T-test results for English syllable structure /CjV/ and /CwV/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
CjV	Harbin	34	96.0788	10.90008	1.86935	0.690	0.493
	Guangzhou	31	93.5484	18.09349	3.24969		
CwV	Harbin	34	89.2165	17.82849	3.05756	-0.470	0.640
	Guangzhou	31	90.3232	17.62411	3.16538		

4.4.1.2 H3.2 and results

> H3.2 Harbinese does not allow /m/ in the coda position while Cantonese does. Therefore, Harbinese but not Cantonese speakers will make errors with the English syllable structure /CVC_N/ (/CVm/).

Cantonese allows /m/ to occur after a vowel in a syllable, but this is forbidden in Harbinese. This hypothesis was proposed to test the influence of the two dialects/languages on the acquisition of the English syllable structure /CVm/. Figure 4.28 suggests that there is no great challenge for either group of participants, especially the Cantonese who achieved notably high accuracy rates up to 95% although minor errors were made. However, the accuracy rate in the Harbinese group was somewhat lower at 88%. Table 4.49 shows the details of errors made. Insertion was the most popular error used by Harbinese participants, making up 9% of the total. The vowel /u/ was often inserted by Harbinese speakers after the coda /m/ to form the syllable. However, deletion was the only error type used by Cantonese participants, accounting for 5%. Harbinese speakers also occasionally deleted /m/ in around 2% of cases. It was rare even for Harbinese speakers to substitute /n/ for /m/ (1%).

Table 4.49 Production of English syllable structure /CVm/

Patterns	Types	Harbin	Guangzhou
CVm	Correct	88%	95%
	Ins-mu	9%	
	Del	2%	5%
	Sub-n	1%	

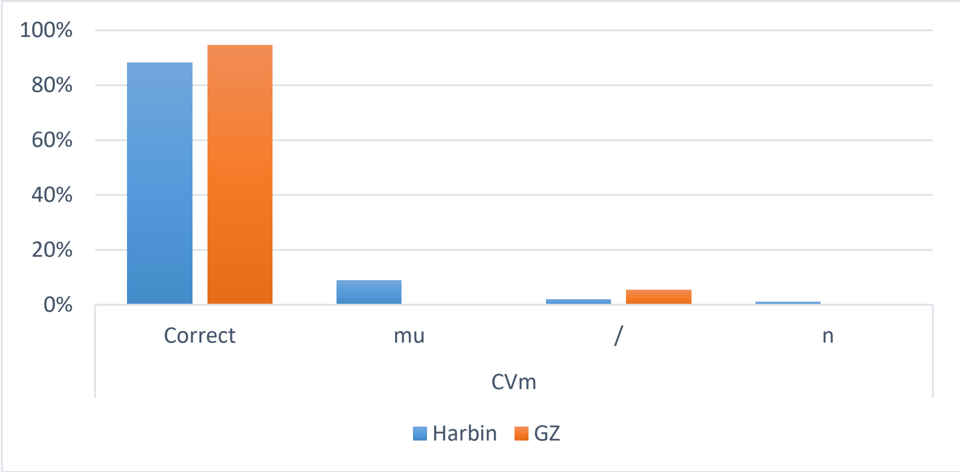


Figure 4.28 Production of English syllable structure /CVm/

From the results we can see that hypothesis H3.2 seems to be supported, because the Cantonese group achieved higher accuracy. However, the statistical result of English syllable structure /CVm/ in Table 4.50 does not show a significant difference between the two dialects/languages, [p=0.185, (p>0.05)]. Although there is a slight difference between the two dialects/languages, it is not significant. Therefore, hypothesis H3.2 has to be rejected. This confirms that 'identical' is easy but also indicates that 'new' is not difficult either in this context. Whether or not other factors may have influence, this result requires discussion.

Table 4.50 T-test results for English syllable structure /CVm/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
CVm	Harbin	34	88.2356	24.45687	4.19432	-1.344	0.185
	Guangzhou	31	94.6242	12.46136	2.23813		

4.4.2 Category of 'similar' in syllable structure

4.4.2.1 H3.3 and results

H3.3 Cantonese has a similar syllable /CVC$_P$/ (/CVp, t, k/) to English but the syllable structure is new in Harbinese, so Cantonese speakers but not Harbinese speakers will make errors with the English syllable pattern /CVC$_P$/ (/CVp, t, k/).

Cantonese has the English syllable pattern /CVp,t,k/ but Harbinese does not. However, the Cantonese syllables are not exactly the same as the English ones, because the final stops are unreleased in Cantonese. This hypothesis tests whether or not the 'similar' between English and Cantonese will be more difficult to learn than the 'new' to Harbinese speakers.

Figure 4.29 clearly shows that the Cantonese group achieved excellent results in pronouncing these English syllable patterns. The accuracy rates for these patterns for the Cantonese group are above 90% compared to less than 80% in the Harbinese group, as seen in Table 4.51. The errors and error types are distinct and typical for each dialect/language. The Harbinese only chose epenthesis as the solution. They inserted the vowel [ə] or [u] after the stops. The vowel [ə] was mainly inserted, with [u] only inserted after the stop /p/ in a few cases. Substitution was the principal and typical error type for the Cantonese group, as they frequently substituted stops with the corresponding unreleased ones [p̚, t̚, k̚]. This error is mentioned by many researchers investigating Cantonese speakers' acquisition of English. However, the results show that while these are the main errors among Cantonese speakers, they are still rather uncommon. One interesting error type made by Cantonese speakers with this English syllable structure was where the alveolar and velar stops were pronounced as their voiced counterparts, such as /t/ in /CVt/ as [d] and /k/ in /CVk/ as [g], but this was also rare. From the results, it can be observed that the /p/ in the pattern /CVp/ seems to be more frequently unreleased by Cantonese participants compared with other patterns. There is a further exception for the pattern /CVk/ where substitution plus insertion was used by Cantonese speakers but again only rarely. The vowel /ə/ was inserted after the stop, which might be influenced by Mandarin.

Table 4.51 Production of English syllable structure patterns /CVp,t,k/

Patterns	Types	Harbin	Guangzhou
CVp	Correct	75%	91%
	Ins-pə	22%	
	Ins-pu	2%	

	Sub-p̚		9%
	*Error	1%	
CVt	Correct	79%	97%
	Ins-tə	21%	
	Sub-t̚		2%
	Sub-d		1%
CVk	Correct	79%	97%
	Ins-kə	21%	
	Sub-k̚		2%
	S+I-g̥ə		1%

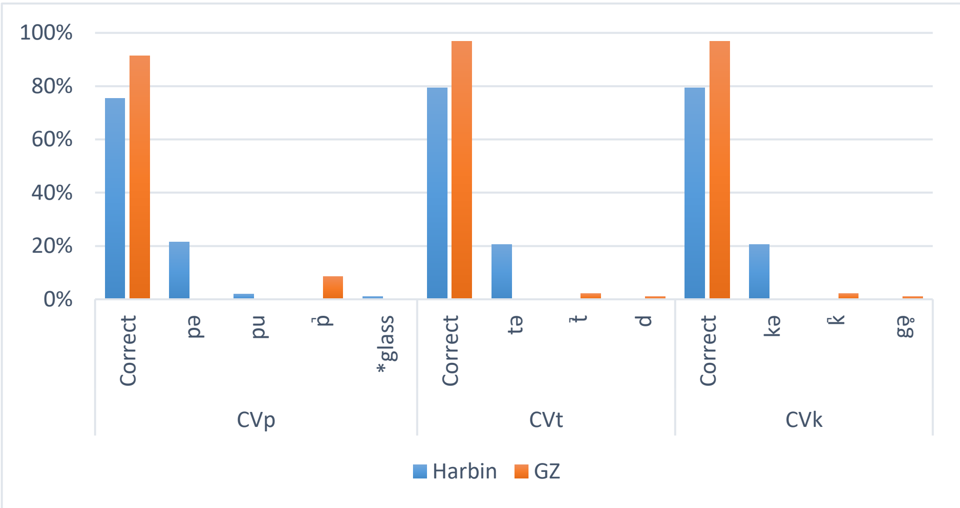


Figure 4.29 Production of English syllable structure patterns /CVp,t,k/

From the above results, it can be concluded that the Cantonese group outperformed the Harbinese in the pronunciation of the English /CVp,t,k/ syllable structure and each group makes its typical error types. Therefore, hypothesis H3.3, that the Cantonese will make errors but not the Harbinese, is rejected. Moreover, the t-test results in Table 4.52 show that the differences between the two dialects/languages with these patterns are statistically significant [p=0.029 for /CVp/, p=0.008 for /CVt/ and p=0.015 for /CVk/]. Therefore, the rejection of hypothesis H3.3 is confirmed to be rejected. Here ‘similar’ does not result in difficulties as claimed by Flege’s SLM (1995). On the contrary, ‘new’ is more difficult when compared to ‘similar’.

Table 4.52 T-test results for English syllable structure /CVp,t,k/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
CVp	Harbin	34	75.4906	36.98113	6.34221	-2.256	0.029
	Guangzhou	31	91.3984	17.14313	3.07900		
CVt	Harbin	34	79.4121	34.82983	5.97327	-2.783	0.008
	Guangzhou	31	96.7745	10.01690	1.79909		
CVk	Harbin	34	79.4121	38.50278	6.60317	-2.537	0.015
	Guangzhou	31	96.7745	10.01690	1.79909		

4.4.3 Category of 'new' in syllable structure

4.4.3.1 H3.4 and results

> H3.4 Harbinese and Cantonese prohibit the pattern of a diphthong plus a nasal coda, especially with /n/, such as in /(C)VVC$_N$/ in their L1s. Therefore, according to Flege's SLM, 'new' is easy and it is predicted that a) neither will make errors with the syllable structure /(C)VVC$_N$/; b) if they do make errors, there will be no differences between the two dialects/languages.

The syllable structure of a diphthong plus a nasal /n/ has often been found to be a problem for Chinese English learners. Although this syllable structure is absent in both dialects/languages, it is worth determining if there are differences between the speakers of different grouops. Five patterns formed by the combination of each of the five English diphthongs plus a /n/ were tested and the results are shown below.

The first pattern is /aɪn/. Figure 4.30 shows that both groups had problems but the Cantonese group pronounced tokens with this syllable pattern much better than the Harbinese. The accuracy rate for the Cantonese was 66%, double that of the Harbinese group, who made far more errors and various error types compared with the Cantonese. The errors can be classified as four types producing one segment, two, three or four. The type with one segment means that the three segments of this syllable pattern are reduced to only one segment left in a pattern. Generally, this one segment would be a vowel from this pattern, a substituted vowel, or a vowel with nasalisation, for example [æ̃]. 3% of Harbinese participants pronounced /aɪn/ as a nasalised monophthong [æ̃] while this never occurred in the Cantonese group. The error type with two segments can be obtained from the remaining segments after the deletion of any vowel or consonant segment in the pattern /aɪn/, or after the substitution of the diphthong /aɪ/ with a monophthong plus /n/, or after substitution of the diphthong with two vowels without /n/. This error type was most frequent among both groups of participants seen in Table 4.53. The deletion of the vowel segment /ɪ/ and the nasal consonant /n/ from /aɪn/ are the most popular strategies. Among the errors in this type, [an] was the most frequently pronounced by Harbinese participants accounting for 29%, followed by [aɪ] (12%) and [aɛ̃] (4%), while [aɪ] was the commonest error among Cantonese participants (27%). Other errors in this type made by Cantonese participants were [aɛ] and [aɛ̃], respectively representing 3% and 2%. The third type of error is the three-segment type where the same number of segments is kept but one segment or segments were substituted; for example in [aɛn] and [aɪm] in which one segment was substituted. The fourth type was only made by the Harbinese, who inserted a vowel into the pattern to form a two-syllable pattern such as [aɪən] and [aɪɪn] where a [ə] or [ɪ] is inserted before the nasal.

To deal with the pronunciation of the English syllable pattern /aɪn/, both groups preferred a two-segment pattern which conforms more to the syllable structure in Chinese which only allows a diphthong or a single vowel plus a nasal. /aɪn/ was mostly pronounced as [an] and [aɪ] by Harbinese participants and [aɪ] by Cantonese speakers.

Table 4.53 Production of English syllable structure /aɪn/

Pattern	Types	Harbin	Guangzhou
aɪn	Correct	32%	66%
	æ̃	3%	
	aɪ̃	1%	
	aɛ̃	4%	2%
	aɪ	12%	27%
	aɛ	1%	3%

	an	29%	
	æn	1%	
	ɛn	1%	
	aɛn	6%	
	aɪm		2%
	aɪən	4%	
	aɪɪn	1%	
	aɪæn	1%	

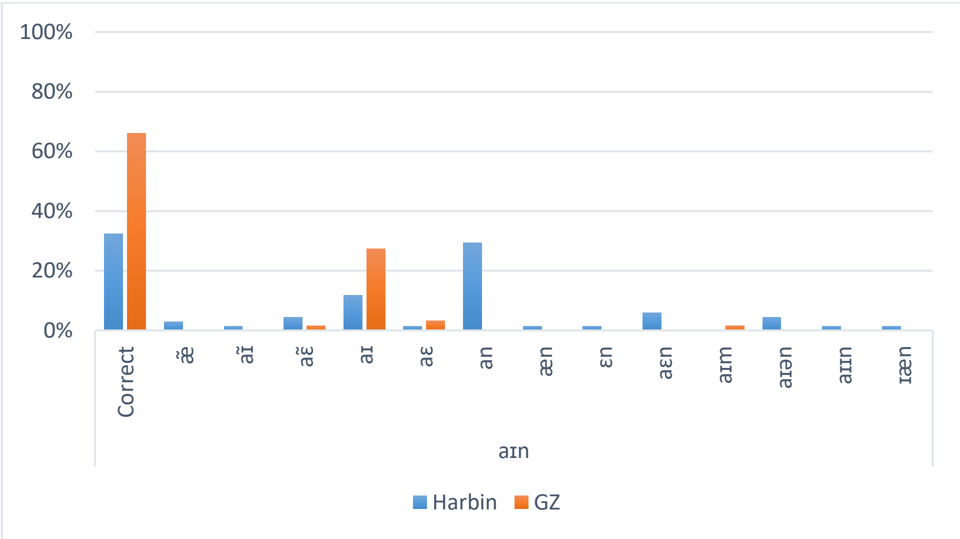


Figure 4.30 Production of English syllable structure /aɪn/

The second syllable pattern tested is /eɪn/. Figure 4.31 indicates that both groups of participants made errors, but the Cantonese group performed better than the Harbinese group. From the results in Table 4.54, the Harbinese group made more types of errors than the Cantonese of different types as was found with the /aɪn/ pattern. The one-segment error type was only made by Harbinese participants who substituted /eɪn/ with a monophthong, which may be lengthened or nasialised such as [ɛ:], [e:] and [ɪ̃] as shown in the Table 4.54. This type of error represented a very small proportion of the Harbinese cases. The two-segment type of error was frequently used by both groups, but the solutions used by each group differed. The Harbinese preferred to pronounce /eɪn/ as [in] (10%), followed by [eɪ] (9%), [eɪ̃] (7%) and [ɪn] (6%). However, 19% of Cantonese participants mispronounced /eɪn/ as [eɪ]. It seems that Harbinese participants used a dual strategy to adapt this syllable pattern to the two-segment pattern, but they preferred to retain the nasal consonant when deleting it since even when the nasal was deleted, the preceding vowel was sometimes nasalised such as in [eĩ]. However, Cantonese participants used the opposite strategy, preferring deletion of the nasal, such as in [eɪ]. The third error type was the three-segment pattern where either segment in the pattern /eɪn/ could be substituted. This type was more often found in the Cantonese group where the alveolar nasal /n/ was substituted with the velar nasal /ŋ/ forming a new sound [eɪŋ]. This was the second most common error made by a large proportion of Cantonese participants. Meanwhile only a small number of Harbinese participants inserted a vowel into the pattern /eɪn/ to form a four-segment error type.

To tackle the pronunciation of the English syllable pattern /eɪn/, the two-segment type of error was most frequently used by the Harbinese and Cantonese. However, Cantonese speakers also used a three-segment pattern, while [in] and [eɪ] were the errors most frequently made by Harbinese participants whereas [eɪ] and [eɪŋ] were most frequent among the Cantonese.

Table 4.54 Production of English syllable structure /eɪn/

Pattern	Types	Harbin	Guangzhou
eɪn	Correct	41%	61%
	ɛ:	1%	
	e:	1%	
	ĩ	1%	
	aɪ	3%	
	eɪ	9%	19%
	aĩ	1%	
	eĩ	7%	2%
	ɪn	6%	
	in	10%	
	en	3%	3%
	ən	3%	
	ɪŋ	3%	3%
	aɪn	4%	2%

	eɪŋ		10%
	eɪən	3%	
	eɪɪn	1%	

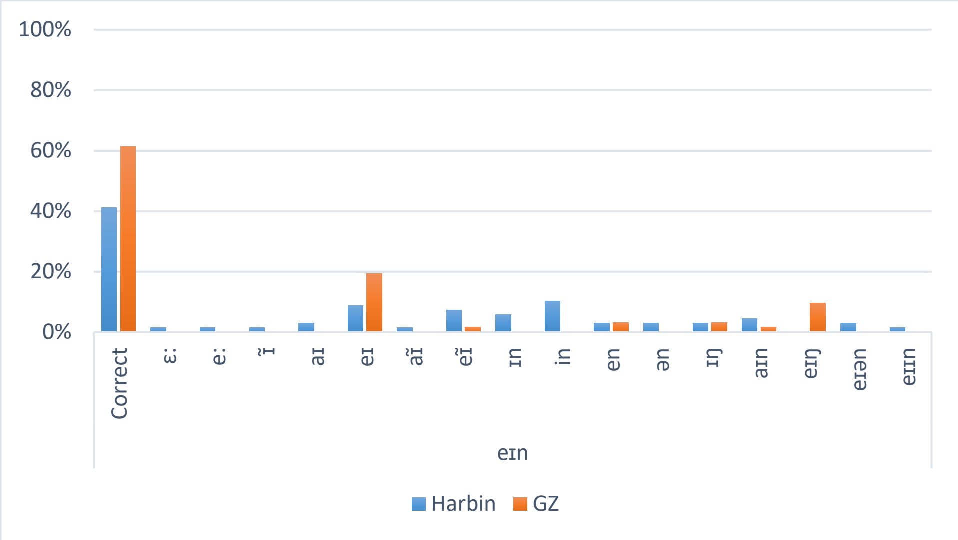


Figure 4.31 Production of English syllable structure /eɪn/

The third pattern with this syllable structure is /ɔɪn/. For this pattern, the results in Figure 4.32 indicate that both groups made many errors but the Cantonese group performed better than the Harbinese who achieved a substantially lower accuracy rate of less than 10% while for the Cantonese group it reached 42%, as seen in Table 4.55. Both groups made various errors classified into three types. The two-segment error type, was a rather popular syllable pattern for both groups. It was the main strategy for the Cantonese group but the second commonest among the Harbinese. The /ɪ/ was often deleted from the diphthong /ɔɪ/ in the syllable pattern /ɔɪn/ by both groups forming the erroneous [ɔn] (16% in the Harbinese group and 18% in the Cantonese). In addition, /n/ was deleted especially by the Cantonese group (15%). Harbinese participants often lowered the height of /ɪ/ in the diphthong /ɔɪ/ to /ɛ/, and thus, the erroneous [ɔɛ] was produced by 7% of the Harbinese participants. The three-segment type was most frequently used by Harbinese participants with the error [ɔən] constituting to 49% of cases. It can be observed here that the Harbinese participants preferred to substitute [ɪ] in /ɔɪn/ with a schwa [ə]. This is not because /ɔə/ is a diphthong in Harbinese but because [ən] is pronounced for the letter *n* in Harbinese. That is the reason why the [ɪ] in /ɔɪn/ was often substituted with [ə]. The three-segment type was the second commonly used error type for Cantonese participants, and /ɔɪn/ was also often pronounced by the Cantonese participants as [ɔɛn] (8%) and [aɪn] (6%). Two errors further made by a few Harbinese participants. A glide [w] was inserted preceding the vowel, such as in [ʷan] and [ʷaɪn] is four-segment error type was rare, and only occurred in the Harbinese group.

To solve the problem of the English syllable pattern /ɔɪn/, the three-segment pattern was the most common choice followed by the two-segment pattern for the Harbinese group, while the reverse was true for the Cantonese group. [ɔən] and [ɔn] were most common among the Harbinese, while [ɔn] and [ɔɪ] were the top two wishes made by the Cantonese.

Table 4.55 Production of English syllable structure /ɔɪn/

Pattern	Types	Harbin	Guangzhou
ɔɪn	Correct	9%	42%
	ɔn	16%	18%
	ɔɪ	3%	15%
	ɔɛ	8%	2%
	aʊ		2%
	aɪ	1%	
	ɔən	49%	3%
	ɔɛn	3%	8%
	ɔan	3%	3%
	ɔɪm		2%
	oən	1%	
	aɪn		6%
	aən	1%	
	ʷan	3%	
	ʷaɪn	1%	

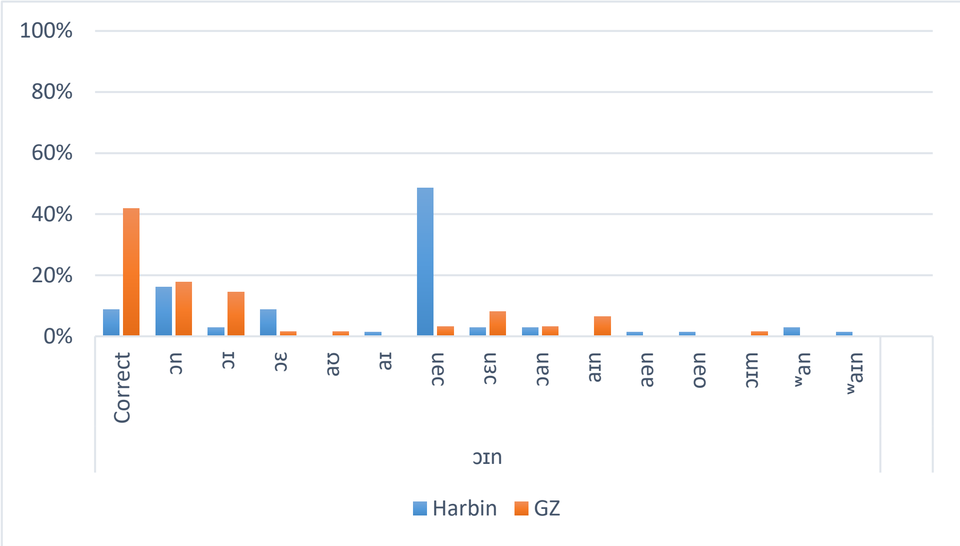


Figure 4.32 Production of English syllable structure /ɔɪn/

The fourth English syllable pattern is /aʊn/. Figure 4.33 and Table 4.56 show that both groups achieved rather low accuracy rates, the Cantonese group (29%) performed better than the Harbinese (13%). Three types of errors were made. The one-segment type was made rarely, and only by the Harbinese where the pattern /aʊn/ was reduced to a nasalised /a/, giving [ã]. The two-segment type was the most popular error type which both groups of participants often used for the English pattern /aʊn/. Among this type, /aʊn/ was often changed to [aŋ] by up to 68% of Harbinese participants and into [a:ŋ] with a lengthened [a] by 27% of Cantonese participants. In addition, [a:m] and [a:n] were errors frequently made by the Cantonese group, respectively accounting for 19% and 10%. From these three errors, it seems that Cantonese speakers had difficulty in final nasals because the nasal /n/ in /aʊn/ was often pronounced as [ŋ] or [m]. However, Harbinese participants showed no evidence of issues with final nasals with the pattern /aʊn/. The three-segment type was not commonly used although both groups made errors of this type.

For the English syllable pattern /aʊn/, both groups frequently used two-segment errors to replace /aʊn/. [aŋ] was the most frequent error made by Harbinese speakers while [a:ŋ], [a:m] and [a:n] were prevalent among Cantonese speakers.

Table 4.56 Production of English syllable structure /aʊn/

Pattern	Types	Harbin	Guangzhou
aʊn	Correct	13%	29%
	ã	1%	
	aʊ		5%
	aɔ	1%	
	aɔ̃		3%
	aŋ	68%	
	a:ŋ		27%
	a:n		10%
	a:m		19%
	ɔŋ	6%	2%
	ɔn	4%	
	ʌŋ		2%
	aʊŋ	3%	2%
	aʊm		2%
	ɔən	1%	
	oən	1%	

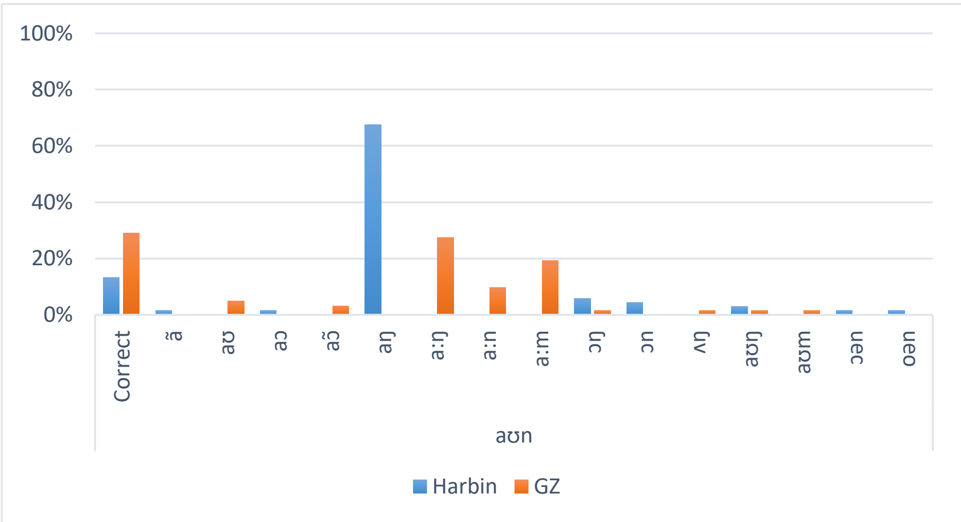


Figure 4.33 Production of English syllable structure /aʊn/

The British English variety of /oʊn/ is /əʊn/ and its counterpart in American English is /oʊn/. Thus, this English syllable pattern was not only checked to see how participants tackled it but also to see what varieties of English the participants used. Figure 4.34 indicates that both groups had low accuracy rates but the Cantonese group performed better. Also, speakers of both groups pronounced this syllable pattern in the American way (21% of the Harbinese speakers and 40% of Cantonese), as shown in Table 4.57. Only 2% of Cantonese speakers pronounced it in the British variety as /əʊn/ but none of the Harbinese speakers did. Apart from the two varieties which are correct pronunciations, all other productions are errors which can be classified into the same four error types as for /aɪn/. The one-segment type error was very rare, only occuring in the Harbinese group. The two-segment error type was commonly used by both groups. Many errors were made in this type and [ʊŋ] was the most common such mistake among the Harbinese, accounting for 15% while [ɔŋ] was the most common by Cantonese participants (18%). Numerous three-segment errors were also made by both groups. The Harbinese often pronounced /əʊn/ or /oʊn/ as [oʊŋ] (12%), [oən] (10%), or [ʊən] (9%), while 15% of Cantonese participants always used [oʊŋ] or /oʊn/ instead. The four-segment error type only occurred in the Harbinese group, and was one of the most common error types where [oʊən] or [əʊən] are the most common such errors, representing 16% and 4% of cases respectively.

It can be concluded from the above results for the English syllable pattern /əʊn/ or /oʊn/ that both groups made errors and the Cantonese group made more than the Harbinese. For whom the four-segment error type was most common, followed by the two-segment and three-segment types. However, two-segment errors were the commonest among the Cantonese followed by the three-segment type, but no Cantonese one or four-segment error types occurred. Among the different types of errors, the most common among the Harbinese were [oʊən], [ʊŋ], [oʊŋ] and [oən], while [ɔŋ], [oʊŋ] and [ɔn] were most frequent among the Cantonese.

Table 4.57 Production of English syllable structure /əʊn/ (/oʊn/)

Patterns	Types	Harbin	Guangzhou
	A	21%	40%
	B		2%
	ɔ:	1%	

əʊn/oʊn A-oʊn B-əʊn	əʊ		2%
	ɔ:w		2%
	on	1%	
	ɔn	4%	8%
	ɔŋ	3%	18%
	əŋ		3%
	ʊŋ	15%	6%
	ãm		2%
	oən	10%	
	ʊən	9%	
	ɔən	3%	
	oʊŋ	12%	15%
	oʊm		2%
	ɔŋm		2%
	əʊən	4%	
	oʊən	16%	

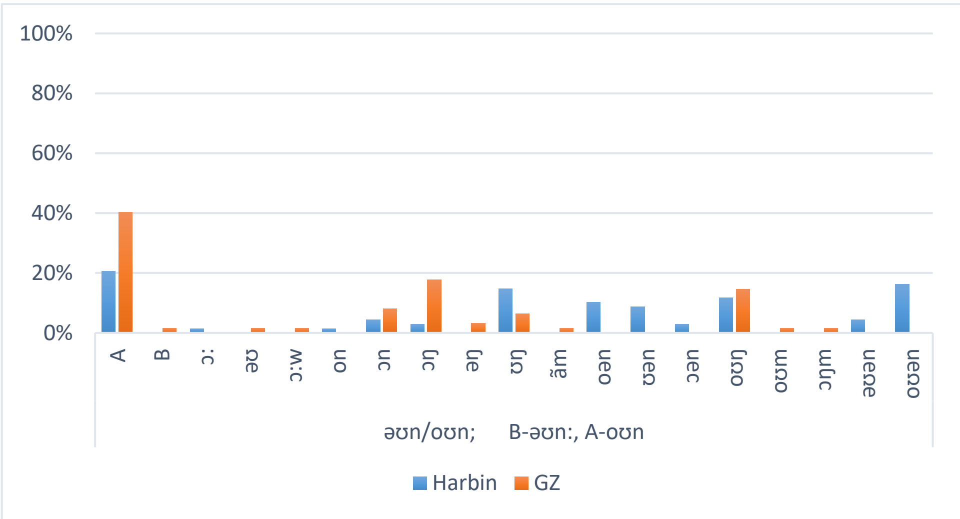


Figure 4.34 Production of English syllable structure /əʊn/ (/oʊn/)

For convenience and a clear illustration of the results, the results for each syllable pattern of /VVn/ are combined in Figure 4.35 showing only the most typical errors in each group. It can be clearly seen that both groups made numerous errors, especially for the patterns /ɔɪn/, /aʊn/ and /əʊn (oʊn)/. Overall, the Cantonese group achieved better pronunciation of all of these five syllable patterns. Therefore, hypothesis H3.4a is rejected because speakers of both groups made errors but the errors were different so hypothesis H3.4b that there will be no differences between dialects/languages is also rejected. Moreover, the results of the statistical tests for these syllable patterns in Table 4.58 show that the differences between the two dialects/languages are statistically significant except for the patterns /aʊn/ and /eɪn/ because whose p-values are greater than 0.05. However, the p-value of /eɪn/ is 0.056 which is only slightly lighter than 0.05. Thus, the difference between the two dialects/languages even with the pattern /eɪn/ is very close to being statistically significant. In addition, the p-value of /aʊn/ is 0.084, which can be interpreted as fairly close to statistical significance. Therefore, it can be concluded that, overall, there are statistically significant differences between the two dialects/languages with all of these syllable patterns, and hypothesis H3.4 must be rejected.

From the results, the syllable structure with these five specific patterns is very difficult for both groups who achieved low accuracy rates. This finding indicates that the 'new' is difficult and not easy which contradicts the prediction from Flege's SLM that the 'new' is easy.

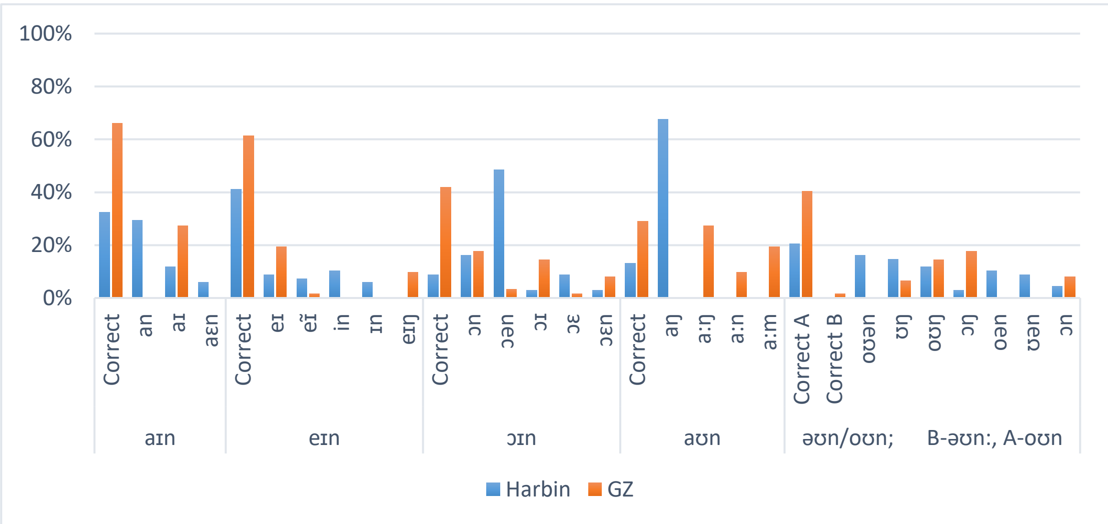


Figure 4.35 Productions of English syllable structure /VVn/ between Harbin and Guangzhou

Table 4.58 T-test results for the English syllable structure /VVn/ in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
aɪn	Harbin	34	32.3529	36.68530	6.29148	-3.391	0.001
	Guangzhou	31	66.1290	43.56431	7.82438		
eɪn	Harbin	34	41.1765	35.82489	6.14392	-1.953	0.056
	Guangzhou	31	61.2903	46.02477	8.26629		
ɔɪn	Harbin	34	8.8235	22.93135	3.93269	-3.820	0.000
	Guangzhou	31	41.9355	43.00538	7.72399		
aʊn	Harbin	34	13.2353	30.91735	5.30228	-1.759	0.084
	Guangzhou	31	29.0323	40.36127	7.24910		
oʊn A	Harbin	34	20.5882	35.07059	6.01456	-2.072	0.042
	Guangzhou	31	40.3226	41.67204	7.48452		

4.4.4 Summary of the results of hypotheses relating to syllable structure

From the results above, it can be seen that both groups have different difficulties in English syllable structure. To syllable structure with C+glide onset i.e. /Cj, Cw/, substitution is the most common error type for both groups, but the Harbinese group also use insertion which is rarely used by Cantonese. To syllable structure with codas /m/ and stops /p, t, k/, Harbinese often inserted a vowel after codas while Cantonese tend to unrelease the stop codas and delete nasal coda /m/. It seems that both groups have problems in English syllable structure of a diphthong plus /n/. Both groups tend to reduce three segments to two segments except /ɔɪn/ in Harbinese which was often to keep three segments with substitutions. From the results, it can be seen that Harbinese have more error types and specific errors than Cantonese.

Hypotheses in the identical category on syllable structure were both rejected. They did not support ‘identical’ is easy but indicated that ‘identical’ is easy and ‘new’ is also easy. Hypothesis H3.3 in the category of similarity was rejected, which means ‘similar’ is easy but ‘new’ is difficult. Hypotheses H3.4 in the category of ‘new’ were both rejected to indicate that ‘new’ is difficult. It seems that Flege’s SLM could be applied to comparison in syllable structure although the hypotheses in identical category were not completely supported. The reason for this rejection will be discussed in chapter 5 in detail.

4.5 Results of hypothesis relating to stress

4.5.1 Category of 'similar' in stress

4.5.1.1 H4 and results

> H4 English has various stress patterns with trochaic feet but Harbinese has two main stress patterns mainly with the iambic foot and with the trochaic foot as mentioned in the literature review. Harbinese is similar to some extent to English in the prosodic template for stress. However, Cantonese has no stress, so it is new in Cantonese. Therefore, the stress pattern should be classified into the category of 'similar', and, it is hypothesised that Harbinese speakers will make errors with English stress patterns with iambic feet but the Cantonese will not.

The acoustic correlates of English stress considered in this research are pitch and intensity, without considering duration and vowel quality because these variables were not controlled in the test tokens for vowels and their surrounding environments, as mentioned in chapter 3. The average pitch and intensity of syllables one and two of disyllabic tokens were computed and contrasted to see which syllable was stressed.

Table 4.59 indicates that the frequency and percentage of stress applied to the first and second syllables. It also shows the number and percentage of disagreement between pitch and intensity and shows the total number and percentage.

Data from 65 participants, including 34 Harbinese and 31 Cantonese participants for 12 English disyllabic words were checked to determine the use of stress, giving a total of 780 words: 408 pronounced by Harbinese speakers and 372 pronounced by Cantonese. In the Harbin group, 171 words (42%) were accented on the first syllable and 40 words (10%) stressed on the second syllable. Of the 408 English words spoken by Harbinese participants, for 197 (48%) it was not possible to decide which syllable was stressed according to pitch and intensity. However, 79% of disyllabic English tokens were stressed in the first syllable by Cantonese speakers and only three tokens were stressed in the second syllable (1%). In the Cantonese group, the percentage of disagreement about English words in terms of pitch and intensity reached 20%.

Neglecting data for which stress placement was undecided, Figure 4.36 clearly shows that the first syllable was most often stressed by a majority of participants in both groups. However, Cantonese participants were more likely to place stress on the first syllable than the Harbinese who were more likely to stress the second syllable.

Table 4.59 Frequency and percentage of stress assignment based on pitch and intensity

Types	Harbin		Guangzhou		Total	
	No.	%	No.	%	No.	%
F	171	42%	294	79%	465	60%
S	40	10%	3	1%	43	6%
Others	197	48%	75	20%	272	35%
Sum	408	100%	372	100%	780	100%

F: first syllable; S: second syllable; Others: those with disagreement between pitch and intensity.

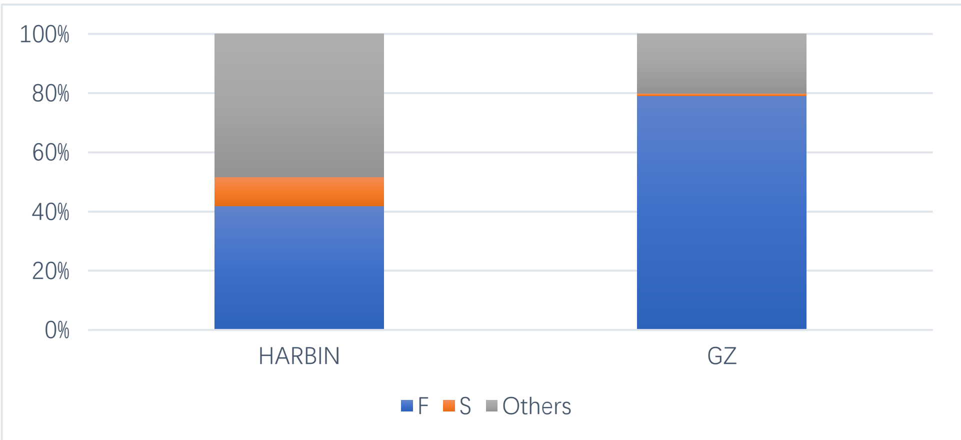


Figure 4.36 Frequency and percentage of stress assignment based on pitch and intensity

Of the 780 English words tested for stress assignment, 272 could not be classified as F or S due to disagreement between pitch and intensity values. After removing these words how stress was placed by the two groups was calculated. Table 4.60 and Figure

4.37 show the new results for frequency and percentage of stress placed on the first and the second syllable by the Harbin and Guangzhou groups. The percentage of stress placed on the first syllable by both groups far exceeds stress on the second syllable. The ratios were 81% and 99% respectively for Harbin and Guangzhou stressing on the first syllable versus 19% and 1% for the second syllable.

Table 4.60 Frequency and percentage of stress assignment based on pitch and intensity excluding others

Types	Harbin		Guangzhou		Total	
	No.	%	No.	%	No.	%
F	171	81%	294	99%	465	92%
S	40	19%	3	1%	43	8%
Sum	211	100%	297	100%	508	100%

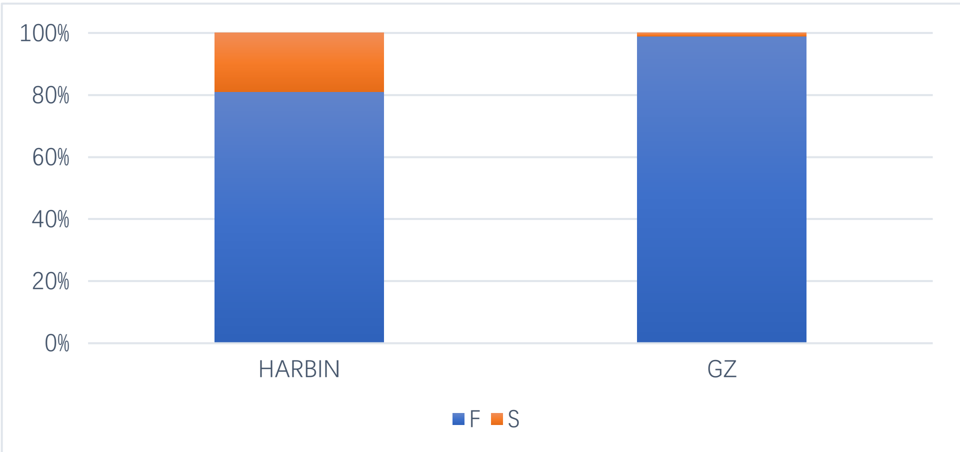


Figure 4.37 Frequency and percentage of stress assignment based on pitch and intensity excluding others

Compared with other correlates of stress, pitch (f0) is regarded as the most significant cue by Duanmu (2000:144, cited by Lin, 2007:224). However, an agreement could not often be reached between pitch and intensity values as indicators of stress in this research. Thus, another method was used to determine stress placement by using pitch as the only criterion regardless of intensity. Stress assignment was then performed based on the new method, followed by another recalculation. Table 4.61 and Figure

4.38 show the results after recalculating the frequency and percentage of the stress placed on first and second syllables.

Now, Harbinese speakers placed stress on the first syllable 263 times (64%) and on the second syllable 145 times up (36%). However, Cantonese speakers assigned the stress to the first syllable (91%) much more often than on the second syllable (9%). Figure 4.38 clearly reveals that Harbinese speakers were more likely to place stress on the second syllable in disyllabic English words than Cantonese speakers and Harbinese speakers did this four times as often as Cantonese speakers.

Table 4.61 Frequency and percentage of stress assignment based on pitch only

Types	Harbin		Guangzhou		Total	
	No.	%	No.	%	No.	%
F	263	64%	338	91%	601	77%
S	145	36%	34	9%	179	23%
Sum	408	100%	372	100%	780	100%

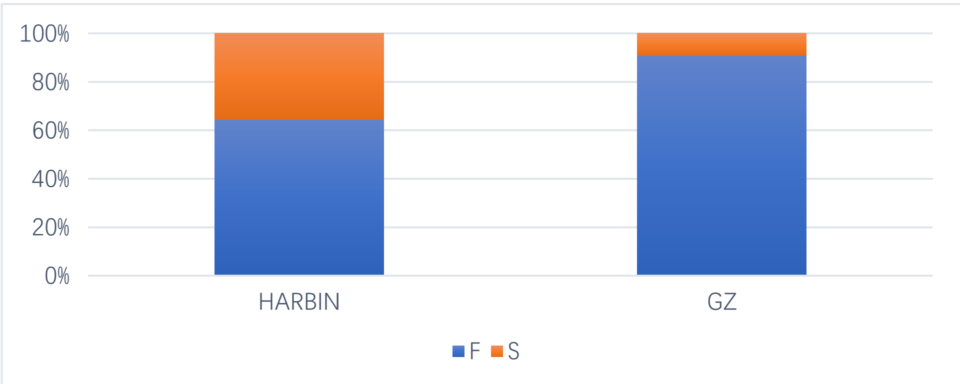


Figure 4.38 Frequency and percentage of stress assignment based on pitch only

The results of the statistical test of stress assignment in the two dialects/languages were conducted, and show significant differences in stress on the first and the second syllables between the Harbinese and Cantonese groups [p=0.000 for S1, p=0.000 for S2 (p<0.05)]. Hypothesis H4.1 that the Harbinese will make errors with the English stress pattern with trochaic feet can therefore be supported. The results show that Harbinese speakers have a tendency to place the stress on the second syllable forming an iambic foot, which may result from L1 transfer. Thus, they have more difficulties than Cantonese speakers. This indicates that here, 'similar' is difficult and 'new' is easy.

Table 4.62 T-test results for English stress assignment in Harbin and Guangzhou

	Hometown	N	Mean	Std. Deviation	Std. Error Mean	t	p (Sig)
S1	Harbin	34	64.4597	18.03836	3.09355	-7.030	0.000
	Guangzhou	31	90.8606	11.85479	2.12918		
S2	Harbin	34	35.5403	18.03836	3.09355	7.030	0.000
	Guangzhou	31	9.1394	11.85479	2.12918		

The test tokens for stress are in two categories: compound disyllabic English words and monomorphemic disyllabic English words. Whether or not Harbinese speakers and Cantonese speakers assign stress using the same methods with the different types of disyllabic English words was then tested.

Table 4.63 and Figure 4.39 present the results for stress assignment on the syllables of compound disyllabic English words. The numbers of Harbinese and Cantonese participants are respectively 34 and 31 and each read eight compound disyllabic English words; thus, 520 compound disyllabic English words were pronounced, 272 by the Harbin group and 248 by the Guangzhou group. From Table 4.63, it can be seen that the ratios of stress placed on the first and second syllables in the Harbin and Guangzhou groups are respectively 60% vs 40% and 90% vs 10%. This implies that Harbinese speakers were three times more likely to assign stress to the second syllable than Cantonese speakers.

Table 4.63 Frequency and percentage of stress assignment to compound disyllabic English words

Types	Harbin		Guangzhou		Total	
	No.	%	No.	%	No.	%
F	162	60%	222	90%	384	74%

S	110	40%	26	10%	136	26%
Sum	272	100%	248	100%	520	100%

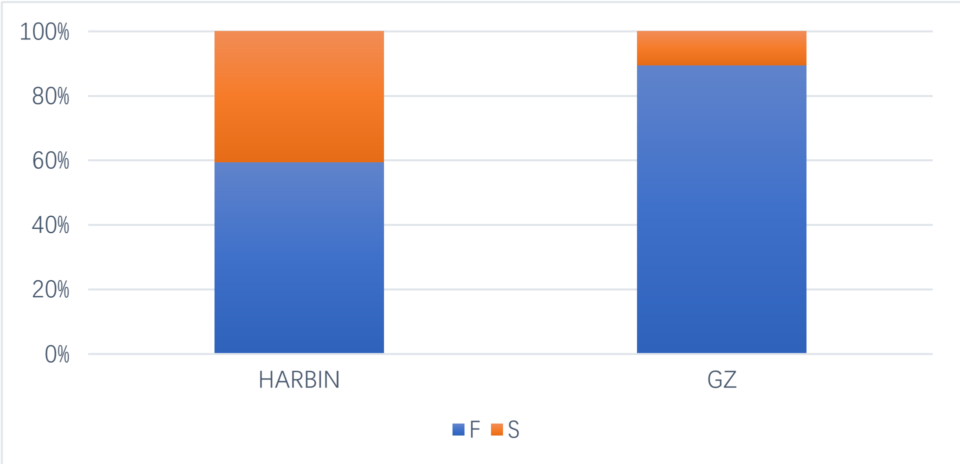


Figure 4.39 Frequency and percentage of stress assignment to compound disyllabic English words

Table 4.64 and Figure 4.40 demonstrate the results for stress placement on the syllables of monomorphemic disyllabic English words. Four tokens from this category were tested with each participant. Thus, the total numbers of disyllabic monomorphemic words pronounced by Harbinese speakers is 136, with 124 pronounced by Cantonese speakers, giving a total of 260.

For this category of disyllabic English words, 74% of Harbinese speakers placed stress on the first syllable and 26% on the second, while 94% of Cantonese users stressed the first syllable with only 6% stressing the second. The Harbinese speakers assigned stress to the second syllable much more often than Cantonese speakers.

Table 4.64 Frequency and percentage of stress assignment to monomorphemic disyllabic English words

Types	Harbin		Guangzhou		Total	
	No.	%	No.	%	No.	%

F	101	74%	116	94%	217	83%
S	35	26%	8	6%	43	17%
Sum	136	100%	124	100%	260	100%

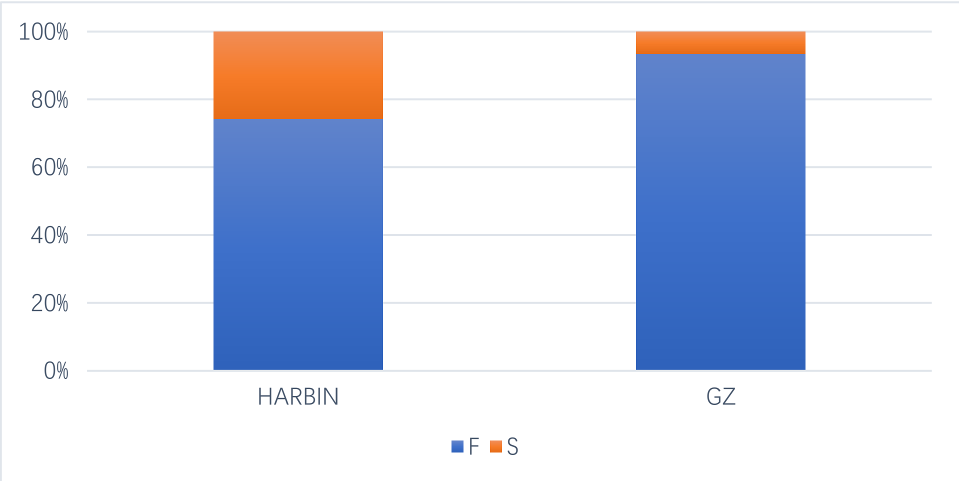


Figure 4.40 Frequency and percentage of stress assignment to monomorphemic disyllabic English words

To clearly see the differences in stress assignment between compound and monomorphemic disyllabic English words, the results in the tables and figures for each are combined in Table 4.65 and Figure 4.41, clearly demonstrating the lower frequency of stress assignment on the second syllable in both groups with the monomorphemic category. This implies that stress is more likely to be placed on the second syllable by both groups in compound disyllabic English words. Irrespective of the type of disyllabic English word, the figure confirms that Harbinese speakers have more inclination to stress the second syllable than Cantonese speakers.

Table 4.65 Comparison of stress assignment of compound and monomorphemic disyllabic English words

Types	Compound		Monomorphemic	
City / Syllable	Harbin	Guangzhou	Harbin	Guangzhou
F	60%	90%	74%	94%
S	40%	10%	26%	6%

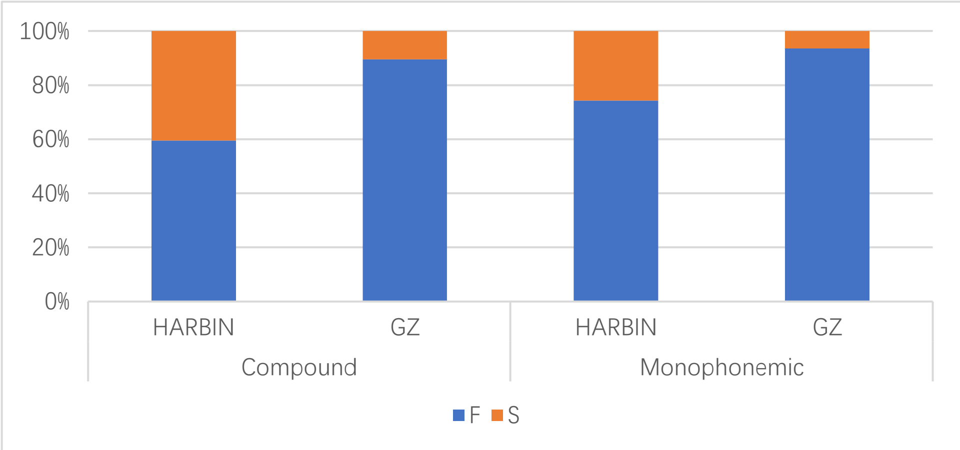


Figure 4.41 Comparison of stress assignment of compound and monomorphemic disyllabic English words

4.6 Results of hypothesis relating to length of exposure

4.6.1 H5 and results

L1 based errors will decrease with increased exposure to English.

It is quite common for Chinese English learners to be exposed to instructed L2 English learning in schools. This research aims to determine how the L1 influences reacquisition of L2 phonology, and also to see if the length of exposure to the L2 influences its production in the L2 classroom settings. This section presents the results for the English production of L2 English learners from two-dialect/language speaking areas in terms of three educational levels of middle school, high school and university. The average accuracy rates of each level in the Harbin and GZ groups are shown in Table 4.66 and Figure 4.42.

Figure 4.42 compares the average accuracy rates of the English production of the different levels in the two groups. It clearly illustrates a fluctuating developmental trend between the educational levels. With the increase of length of exposure from middle school to high school, the figure shows that there is a positive trend of increasing accuracy from 62% to 69% in the Harbin group and from 71% to 74% in the GZ group (see also Table 4.66). The accuracy rates of productions at high school level reaches the summit (69% and 74%) for both high schools within their groups. The development performance then declines slightly to 66% in the Harbin group and 70% in the GZ group at university level. The extent of decline at university level differs between the groups. For the Harbin group, it is still better of the middle school level (the ratio between them is 66% vs. 62%). While performance at university level in the GZ group is slightly lower even than the middle school level (their ratios are at 70% vs. 71%). Thus, it can be concluded that both groups have similar developmental trends but the specific trajectory within each group is slightly distinct. Both groups perform best at high school level.

From the above results, hypothesis H5, which states that with increased exposure to English, L1 based errors decrease is rejected. Error rates decreased, only when comparing middle school and high school levels.

Table 4.66 Results for average accuracy rates at each level for Harbin and Guangzhou

Levels	Harbin	Guangzhou
MidS	62%	71%
HS	69%	74%
Uni	66%	70%
Average	66%	72%

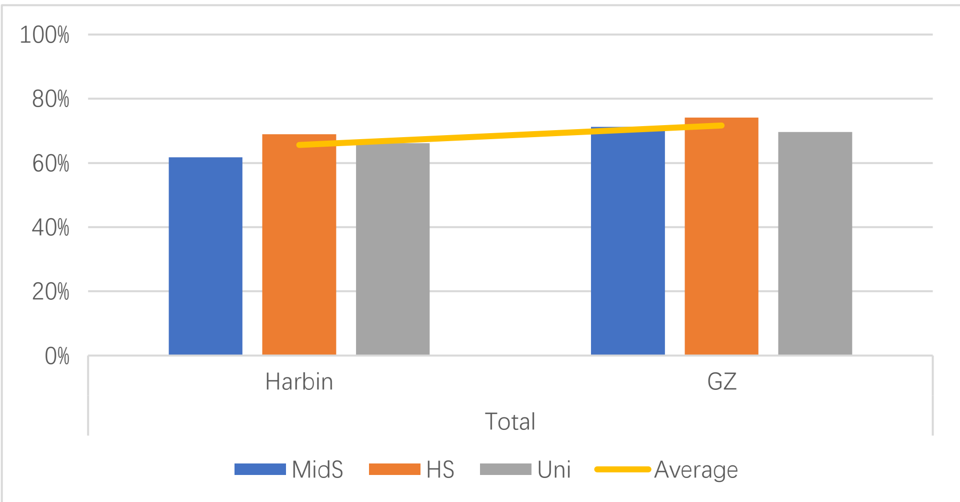


Figure 4.42 Results of average accuracy rates at each level for Harbin and Guangzhou

From the t-test results in Table 4.67 of the comparisons between levels in each group, it can be seen that difference only between middle school and high school in the Harbin group shows statistical significance [p=0.026, (p<0.05)]. However, there are no significant differences for the GZ group or between the other levels in the Harbin group. Therefore, this further confirms that hypothesis H5 cannot be supported. It means the length of exposure does not have a significant influence on L2 learners' pronunciation in this L2 classroom environment.

Table 4.67 T-test results of average accuracy rates of each level in Harbin and Guangzhou

Hometown		N	Mean	Std. Deviation	Std. Error	F	Sig.	Multiple Comparisons
Harbin	B. Middle School	12	60.1433	7.13839	2.06068	3.807	0.033	B vs. C (p=0.026)
	C. High School	12	69.0508	10.01162	2.89011			B vs. D (p=0.507)
	D. University	10	63.9570	5.64854	1.78623			C vs. D (p=0.305)
	Total	34	64.4088	8.57973	1.47141			
Guangzhou	B. Middle School	10	74.3640	8.59634	2.71840	0.260	0.773	B vs. C (p=0.787)
	C. High School	10	76.7800	9.03748	2.85790			B vs. D (p=0.995)
	D. University	11	74.6964	6.79055	2.04743			C vs. D (p=0.829)
	Total	31	75.2613	7.94963	1.42780			

Figure 4.43 compares each level for the two groups, showing that each level within the Cantonese group has a higher accuracy rate than its corresponding level in the Harbinese group. suggesting that Cantonese speakers have acquired better English than Harbinese speakers at all levels.

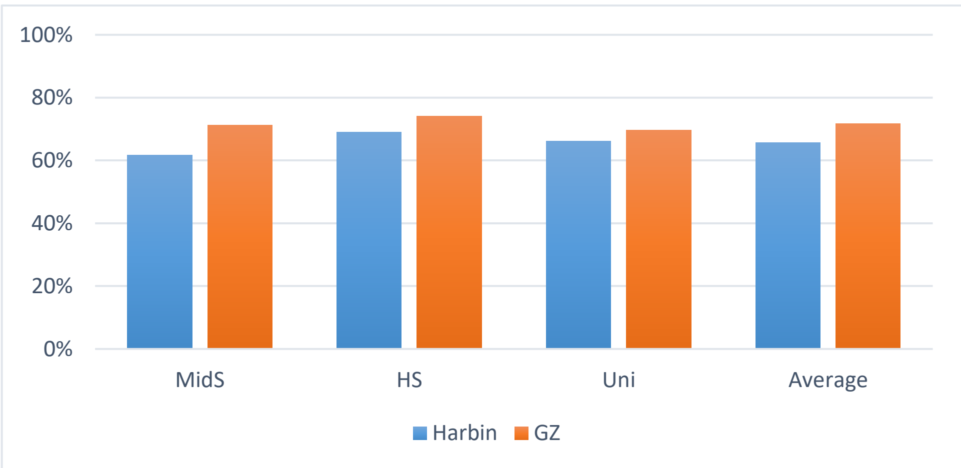


Figure 4.43 Comparison of average accuracy rates from Harbin and Guangzhou at each level

Table 4.68 indicates that the differences at each level in two groups are statistically significant except for at high school level [p=0.074, (p>0.05)]. This implies that the participants' achievements from high schools in the two cities in L2 English phonology acquisition are similar. However, overall accuracy rates in the groups are significantly different [p<=0.001]. Cantonese speakers can acquire English phonology better than Harbinese speakers.

Table 4.68 T-test results for average accuracy rates in Harbin and Guangzhou at each level

Levels		N	Mean	Std. Deviation	Std. Error	F	P(Sig.)
B. Middle School	Harbin	12	60.1433	7.13839	2.06068	18.000	0.000
	Guangzhou	10	74.3640	8.59634	2.71840		
	Total	22	66.6073	10.53035	2.24508		
C. High School	Harbin	12	69.0508	10.01162	2.89011	3.546	0.074
	Guangzhou	10	76.7800	9.03748	2.85790		
	Total	22	72.5641	10.15006	2.16400		
D. University	Harbin	10	63.9570	5.64854	1.78623	15.340	0.001
	Guangzhou	11	74.6964	6.79055	2.04743		
	Total	21	69.5824	8.22314	1.79444		
Total Average	Harbin	34	64.4088	8.57973	1.47141	27.818	0.000
	Guangzhou	31	75.2613	7.94963	1.42780		
	Total	65	69.5846	9.87015	1.22424		

4.7 Summary of results of English production

4.7.1 Production of English consonants

The error types and specific errors with their percentage of prediction of the above consonants across positions are displayed in Table 4.69 to allow a comparison of the two groups.

Table 4.69 shows that five categories of errors were employed by participants in two groups to cope with the production of these consonants. The categories are substitution, deletion, insertion, a complex one involving substitution plus insertion, and metathesis. Substitution is the most frequently used method and metathesis is the rarest. For consonants occurring in different positions, different categories of errors were made.

Substitution:

Substitution can be found everywhere except in some positions for specific consonants, such as /ʃ/, /tʃ/ in the final position, /dʒ/ in the final position in the Harbin group, /dʒ/ in the initial position in the GZ group, and the /n/ in the initial and final positions in the Harbin group, as well as /l/ in the medial and final positions.

Deletion:

Deletion can be found in the medial and final positions, but it was used primarily in the medial position except for the consonants /n/ and /l/. Deletion was a common strategy adopted by both groups to tackle the medial and final /l/ and it was also used to cope with the final /n/. It also occasionally occurred in the final positions of some other consonants. Apart from the predominant role in the production of /l/, deletion was secondary to the commonest strategy of substitution in the medial position.

Insertion:

Insertion was used by small proportions of the participants in all positions but more substantially in the medial and final positions and rarely in the initial positions. The Harbinese group, often used it in final positions but it can also be often found in both medial and final positions in the Cantonese group. This was not the main error type for either group except for the final /n/ in the Harbin group.

Substitution plus insertion:

This complex category was often used in the medial and final positions but mainly in the final position, especially among the Harbin group. It is a specially used as the main error type for the consonants /ʃ/, /tʃ/ and /dʒ/ by both groups. It is also a common method with other consonants but is not the most common error type.

Metathesis:

Metathesis rarely occurred in the production of English consonant segments. It only appeared once in the Cantonese group. The syllable structures of the test token

/səʊldʒə/ ‘soldier’ is $C_1V_2C_3C_4V_5$ but here it was metathesised as [sloʊdi] with a C_3 medial /l/ being placed after the initial consonant C_1 /s/ forming the consonant cluster /sl/ in a new sequence of $C_1C_3V_2C_4V_5$.

Initial position:

Substitution was most frequently used in the initial position except for /n/ in the Harbin group and /dʒ/ in the GZ group where consonants were epenthesised before and after. For example, in the Harbin group, a /k/ was inserted preceding the /n/ in /n/ in [naɪf] ‘knife’ in forming [knaɪf]. In the GZ group, a glide /j/ was often inserted after an initial /dʒ/ such as in /dʒʌmp/ ‘jump’ and /dʒəpæn/ ‘Japan’ which were pronounced as [dʒjʌmp] and [dʒjɛpæn].

Medial position:

The dominant error type in the word medial position was substitution, which is the same as in the initial position except for /l/ whose deletion was the most frequent choice for both groups of participants. The next most common strategy was frequently deletion, especially by the Cantonese group. Apart from with the consonant segment /l/, deletion was used in the medial position for the consonants /v/, /θ/, and /ɹ/ by the Cantonese speakers while it was used for /v/ and /ʒ/ by Harbinese speakers. Insertion was another method used in this position but mostly by Cantonese speakers for the consonants /v/, /ð/, and /l/, while it was used only once by Harbinese speakers for the consonant /l/. The complex method of substitution plus insertion was also used but infrequently. It was used for /ʒ/ by Harbinese speakers and /ð/ by Cantonese speakers. In the medial position, metathesis was adopted once by Cantonese speakers only in the medial position of /l/. It seems that, apart from the substitution and deletion used as the main solutions by both groups, the Cantonese speakers were more likely to use insertion and the other methods than Harbinese speakers when consonant segments occurred in the medial positions.

Final position:

In the final position, various error categories appear but the most significant are substitution and substitution plus insertion. Deletion and insertion were also used but not so frequently except for the consonants /n/ and /l/. Substitution plus insertion seems to be a significant feature arising in the final position.

Table 4.69 Summary of error categories and specific errors with percentages of English consonant segments

Cons	Harbin			Guangzhou		
	Initial	Medial	Final	Initial	Medial	Final
/v/	**Sub:** w- 32%	**Sub**: w-43% ʋ-1% **Del**: 9%	**Sub**: w-64% f-1% **S+I:** wu-2% wə-1% **Del**: 2% **Ins:** vɤ-1% vu-1%	**Sub**: w-46% ɹ-1%	**Sub**: w-33% f-2% **Del**: 13% **Ins:** və -1%	**Sub**: f-81% s-1% **Del**: 2%
/w/	**Sub:** v-16%	**Sub:** v-19%		**Sub:** v-5% ɹ-2%	**Sub:** v-4%	
/ʃ/	Correct	**Sub:** s-1% tʃ -1% t -1%	**S+I:** ʂɻ̩-12% **Sub:** tʃ -2%	**Sub:** ɕ-1%	**Sub:** ʒ-3%	**S+I:** ɕy̥-2% tɕʰy̥-2%
/ʒ/		**Sub:** ɹ-39% ɻ-14% dʒ-11% ʃ-8% tɕ-1% s-1% **Del**: 14% **S+I:** jiɹ-1% sj-1%			**Sub:** ʃ-20% dʒ-6% ɹ-3% s-3%	
/tʃ/	**Sub:**	**Sub:**	**S+I:**	**Sub:**	**Sub:**	**S+I:**

	tʃʷ-2% tɕʰ-1% tɹ-1%	tʃʷ-10% tɹ-8% ʈʂʰ-6% t-2% tsʰʷ-1% ɾ-1%	ʈʂʰɻ-14%	tɹ-2%	tɹ-4% tʃʷ-1% ts-1%	tɕʰy-10% **Sub:** tɕʰ-1% f-1%
/dʒ/	**Sub:** ʈʂ-11% tɕ -1%	**Sub:** ʈʂ-4% dɹ-1% ɻ -1% z-1% **dʒʷ**-1%	**S+I:** ʈʂɻ-20% **Sub:** ʈʂ-1% ʒ-1%	**Ins:** dʒj-3%	**Sub:** ʒ-1%	**Sub:** tʃ-14% **S+I:** tɕy̥-8% tɕʰy̥-8% ʈʂɻ-1% **Ins:** dʒi-1%
/h/	**Sub:** x-13%	**Sub:** x-3%		**Sub:** x-2%	**Sub:** x-1%	
/θ/	**Sub:** s-52% dz-1% d̪-1%	**Sub:** s-53% z-2%	**Sub:** s-36% dz-1% **S+I:** si-9% sə-1% sa-1% **Ins:** θi -1% **Del**: 1%	**Sub:** f-32% s-5% d-1%	**Sub:** f-39% s-8% d-2% z-1% ʒ-1% **Del**: 2%	**Sub:** f-46% s-4%
/ð/	**Sub:** d-26% dz-23% z-5% d̪-5% l-1%	**Sub:** dz-24% z-10% d-4% t-1%	**Sub:** dz-12% z-10% θ-9% s-8% ts-1% **S+I:** dzi-10% si-2%	**Sub:** d-91% d̪-5%	**Sub:** d-37% d̪-16% ɾ-6% f-4% dz-2% θ-1% s-1% dʒ-1% **S+I:** fd-2% θd̪-1% **Ins:** ðd̪-1%	**Sub:** f-39% θ-30% s-3% **S+I:** fs-1% **Del**: 1%
/n/	**Ins:** kn-1%	Correct	**Ins:** ən-21% **Del**: 15%	**Sub:** l̃ -2% l-1%	Correct	**Sub:** ŋ -12% m-4% nd-1%

						nt-1% **Del**: 11%
/l/	Correct	**Del**: 82% **Sub:** w-4% **Ins:** lə-1%	**Del**: 65% **Sub:** w-25% ɔ-2%	**Sub:** l̃ -2% n-1%	**Del**: 47% **Sub:** w-6% oʊ-2% **Ins:** lə-2% **Meta:** sloʊdi-1%	**Del**: 51% **Sub:** w-29% **Ins:** lə-1%
/ɹ/	**Sub:** ɻ-20%	Correct	A-81% B-19%	**Sub:** w-4% l-2% ɻ-1%	**Sub:** w-5% l-2% **Del**: 3%	B-62% A-38%

4.7.2 Production of English vowels

To conveniently compare differences in vowel production between the two groups and to summarise the features of their vowel production, results for all error types and specific errors with vowels are combined in Table 4.70.

It can be seen that Harbinese speakers are prone to use the American variety of English more than Cantonese speakers, and that Cantonese speakers produce British English more. In addition, the variety used depends on specific vowels, for example, an overwhelming majority of both groups of speakers pronounced /ɒ/ in British English rather than the American /ɑ/, while /ɜ/ was pronounced as [ɝ] by most speakers in both groups in the American way. I think it is influenced by the features of each dialect/language and also English instruction in China.

The error types and errors with monophthongs are summarised in Table 4.70, showing that Cantonese and Harbinese speakers are apt to substituting monophthongs with other monophthongs, while in addition Harbinese speakers also substituted these monophthongs with diphthongs. Although diphthongal substitution was also used by Cantonese speakers, this only occurred very rarely. Insertion was also used by

Harbinese speakers but only for the specific sound /ə/ occurring finally.

It can also be concluded that speakers of both dialects/languages have difficulties in contrasting some vowels, Harbinese speakers especially confuse vowels more often. The main errors of confusions among groups of vowels are listed below:

1. The Cantonese group often made errors with /ɔ/ and /ɒ/ whereas the Harbinese group often made errors with /ɔ/, /ɒ/, /ɑʊ/ and /aʊ/. They may have some variants of these vowels such as [ɑɔ] is the variant of /ɑʊ/.

2. Errors made by Cantonese participants were always between /ɛ/ and /æ/ but Harbinese participants often erred between /ɛ/, /æ/ and /aɪ/.

3. Errors were often made by Cantonese speakers between tense and lax such as /i/ & /ɪ/ and /u/ & /ʊ/ while Harbinese speakers made errors not only between tense & lax but also with diphthongs such as, /i/ & /ɪ/ & /eɪ/ and /u/ & /ʊ/ & /ʊə/.

4. A low central /a/ was often substituted for /ʌ/ and /ɑ/ by speakers of both groups.

For diphthongs, the Harbinese group tended to use vowel reduction more than the Cantonese group, especially for diphthongs such as /eɪ/, /əʊ (oʊ)/, /aʊ/. Also, both groups seemed prone to lower the second segment of diphthongs from the high-mid to low-mid position, especially among the Harbinese group. For example, /aɪ/, /ɔɪ/ and /aʊ/ were lowered to [aɛ], [ɔɛ] and [aɔ]. Harbinese speakers often inserted a glide [w] before a specific diphthong /ɔɪ/ or its substitutions.

Table 4.70 Summary of error categories and specific errors with percentages of English vowel segments

Vowels	Harbin	Guangzhou
/ə/	**Substitution:** ɚ-7%	**Substitution:** ɚ-2%

	a-4% **Insertion:** əl-3% əw-2%	ɜ-2% u-2%
/ɜ:/ A-ɝ; B-ɜ:	**Variety:** A-98% B- 2%	**Variety:** A-80% B-18% **Substitution:** ɔ:-5% **Diphthong:** ɔ:ɜ-1%
/ɔ/	**Diphthong:** ɑɔ-12% ɑʊ-9% ʊɔ-1% əɔ-1% **Substitution:** ɔ˞-10% ɒ-3% **Insertion:** a:w-1%	**Substitution:** ɒ-5% ɔ˞-2% a-2% ɜ:-1% **Diphthong:** ɑɔ-1%
/ɒ/ A-ɑ; B-ɒ	**Variety:** A-3% B- 37% **Substitution:** ɔ-29% ɤ-14% ə-4% ɝ-2% o-2% ɔ˞-1% **Diphthong:** ɑɔ-5% ɑʊ-3%	**Variety:** A-1% B- 84% **Substitution:** ɔ-10% ɤ-3% ɜ:-1% ə-1%
/i/	**Substitution:** ɪ-32% **Diphthong:** eɪ-2%	**Substitution:** ɪ-68%
/ɪ/	**Substitution:** i-26% **Diphthong:** eɪ-3%	**Substitution:** i-15%
/u/	**Substitution:** ʊ-8%	**Substitution:** ʊ-24%

	ɔ:-1% o:-1% **Diphthong:** ʊə-11% ʊɔ-1% oʊ-1%	ɒ-1%
/ʊ/	**Substitution:** u-53% a-3% **Diphthong:** ʊə-2%	**Substitution:** u-41%
/ɛ/	**Diphthong:** aɪ-51% aɛ-2% **Substitution:** æ-6% ɪ-3%	**Substitution:** æ-8% ɪ-2% **Diphthong:** aɪ-1%
/æ/	**Substitution:** ɛ-45% a-1% **Diphthong:** aɪ-22% eɪ-1%	**Substitution:** ɛ-75% **Diphthong:** eɪ-1%
/ɑ:/	**Substitution:** a-59% æ-2% ɤ-1%	**Substitution:** a-59% æ-1%
/ʌ/	**Substitution:** a-62% ɔ-2% æ-1%	**Substitution:** a-31% ɔ-4% ɑ-1%
/aɪ/	**Substitution:** aɛ-38% **Monophthong:** æ:-4% ɛ:-1%	**Substitution:** aɛ-13% **Insertion:** aɪn-1%
/eɪ/	**Monophthong:** e:-3%	**Monophthong:** ɛ:-1%
/əʊ(oʊ)/ A-oʊ B-əʊ	**Variety:** A-55% B- 25% **Monophthong:** o:-15% ɔ:-3% **Substitution:**	**Variety:** A-59% B- 37% **Monophthong:** o:-2% ɔ:-1% **Substitution:**

	ɤʊ-1% ɤo-1%	aɔ-1%
/aʊ/	**Substitution:** ɑʊ-19% ɑɔ-12% oʊ-2% **Monophthong:** ɔ-25% a-2% o:-1%	**Substitution:** ɑʊ-2% ɑɔ-2% **Monophthong:** ɔ-2%
/ɔɪ/	**Substitution:** ɔɛ-18% aɪ-6% oɛ-1% **Insertion+Sub:** waɪ-2% wɔɛ-2% woɪ-2% wɛ-1% **Monophthong:** ɛ-1% ɔ-1%	**Substitution:** aɪ-5% ɔɛ-4% aɛ-3%

4.7.3 Production of English syllable structure

The results for errors in the production of English syllable structure are combined in Table 4.71.

1. For syllable structure with onset C+glide /Cj/ and /Cw/, substitution was the most frequent error type for both groups. /j/ was substituted by [ɪ] and /w/ by [ɹ]. However, insertion was often used by the Harbinese group more than the Cantonese group. The vowel [ə] was frequently inserted, followed by [u], but the Cantonese group only ever used [u].

2. For syllable structures with /m/ and stops as codas, different substitutions were used. Harbinese speakers often epenthesised a vowel after the consonants. The inserted vowels were often a schwa after stop consonants or [u] after the bilabial nasal [m] and sometimes after the stop [p]. However, Cantonese speakers adopted substitution instead. They used unreleased stops to substitute for each corresponding stop.

3. For syllable structures with a diphthong plus /n/, the Harbinese group had more difficulties than the Cantonese, making more errors of more types. Cantonese speakers often changed the syllable structure into a two-segment structure. They also ranked the three-segment structure but with changes to some segments but this method was less used than the two-segment structure. However, in the Harbinese group, the error type depended on the specific diphthongs involved. They often used the two-segment structure more for the patterns /aɪn/, /eɪn/ and /aʊn/ and the three-segment structure more for the patterns /ɔɪn/ and /oʊn/. The Harbinese group also used the four-segment and one-segment structures, but the Cantonese group never made these types of error with English syllable structures with a diphthong plus /n/.

Table 4.71 Summary of error categories and specific errors with percentages of English syllable structure

Patterns	Harbin	Guangzhou
/CjV/	**Substitution:** Cɪw-2% Vowel changes: njəʊ-1% bjə-1%	**Substitution:** Cɪw-4% Cɪu-2%
/CwV/	**S+I:** gəɹɪ-4% guɹɪ-1% **Substitution:** gɹɪ-2% **Insertion:** gəwɪ-2% guwɪ-1%	**Substitution:** gɹɪ-5% kɹɛ-3% **S+I:** guɹɪ-1%
/CVm/	**Insertion:** mu-9% **Deletion:** 1% **Substitution:** n-1%	**Deletion:** 5%
/CVp/	**Insertion:** pə-22% pu-2%	**Substitution:** p̚-9%

/CVt/	**Insertion:** tə-21%	**Substitution:** t̚-2% d-1%
/CVk/	**Insertion:** kə-21%	**Substitution:** k̚-2% **Insertion:** gə̥-1%
/aɪn/	**Two-segment:** an-29% aɪ-12% aɛ̃-4% aĩ-1% aɛ-1% æn-1% ɛn-1% **Three-segment:** aɛn-6% **Four-segment:** aɪən-4% aɪɪn-1% aɪæn-1% **One-segment:** æ̃ -3%	**Two-segment:** aɪ-27% aɛ-3% aɛ̃-2% **Three-segment:** aɪm-2%
/eɪn/	**Two-segment:** in-10% eɪ-9% eĩ-7% ɪn-6% aɪ-3% ən-3% en-3% ɪŋ-3% aĩ-1% **Three-segment:** aɛn-6% **Four-segment:** aɪən-4% aɪɪn-1% aɪæn-1% **One-segment:** ɛ: -1% e:-1% ĩ-1%	**Two-segment:** eɪ-19% en-3% ɪŋ-3% eĩ-2% **Three-segment:** eɪŋ-10% aɪn-2%
/ɔɪn/	**Three-segment:** ɔən-49%	**Two-segment:** ɔn-18%

	ɔɛn-3% ɔan-3% ʷan-3% oən-1% aən-1% **Two-segment:** ɔn-16% ɔɛ-8% ɔɪ-3% aɪ-1% **Four-segment:** ʷaɪn-1%	ɔɪ-15% ɔɛ-2% aʊ-2% **Three-segment:** ɔɛn-8% aɪn-6% ɔən-3% ɔan-3% ɔɪm-2%
/aʊn/	**Two-segment:** aŋ-68% ɔŋ-6% ɔn-4% aɔ-1% **Three-segment:** aʊŋ-3% ɔən-1% oən-1% **One-segment:** ã-1%	**Two-segment:** a:ŋ-27% a:m-19% a:n-10% aʊ-5% aɔ̃-3% ɔŋ-2% ʌŋ-2% **Three-segment:** aʊŋ-2% aʊm-2%
/əʊn (oʊn)/ A-oʊn B-əʊn	**Variety:** A-21% **Three-segment:** oʊŋ-12% oən-10% ʊən-9% ɔən-3% **Two-segment:** ʊŋ-15% ɔn-4% ɔŋ-3% on-1% **Four-segment:** oʊən-16% əʊən-4% **One-segment:** ɔ:-1%	**Variety:** A-40% B-2% **Two-segment:** ɔŋ-18% ɔn-8% ʊŋ-6% əŋ-3% əʊ-2% ɔ:w-2% ãm-2% **Three-segment:** oʊŋ-15% oʊm-2% ɔŋm-2%

4.8 Summary

In testing the proposed hypotheses, this chapter has described in detail the results of L2 English production from the two dialect/language groups. In addition, statistical calculations were conducted to confirm statistical significance of any differences between speakers of the two dialects/languages. The hypotheses have all been supported or rejected.

In this section, a summary is presented of all research hypotheses along with the findings concerning their support and rejection, see Tables from 4.72 to 4.76 below. It can be concluded from the list that the hypotheses belonging to the category of 'identical' with English segments were mostly supported to be easy except for H1.3 concerning consonants which was rejected. However, the hypotheses on syllable structure in the category of 'identical' were both rejected. Interestingly, the hypotheses about the 'similar' category were all rejected except for hypothesis H4.1 concerning stress and hypotheses H2.3b) and H2.4 which were supported, which means Flege's SLM that 'similar' being difficult cannot be mostly supported. For the category of 'New' being easy, all hypotheses were rejected. To the different types of 'similar', it can be concluded that different types of similarity indeed have different degree of difficulty. The rankings can be seen below:

Similar 'phonemically identical but allophically different' (H2.3)

is easier than

Similar 'allophonically identical but phonemically different' (H1.6)

is easier than

Similar 'phonemically similar' (H2.6)

is easier than

'new'.

Table 4.72 Hypotheses for consonants were based on the following

Category	No.	Harbinese ($L1_H$)	Cantonese ($L1_C$)	English (L2)	Predictions	Results of hypotheses
Identical	H1.1	/x/	/h/	/h/	Harbinese will make errors	Supported.
	H1.2	/w/	/w/	/w/	Harbinese	Supported.

		[w] and [v]			will make errors	
	H1.3	/n/ and /l/[l]	Neutralise /n/ and /l/[l]	/n/ and /l/[l]	Cantonese will make errors	Rejected. Harbinese is better, Canontese is also good
Similar	H1.4	[v] /w/	/	/v/	Harbinese will make errors	Rejected. ‘Similar’ is not more difficult than ‘new’. ‘Allophonically identical’ seems to play a role but is not significant.
	H1.5	/ɻ/	/	/ɹ/	Harbinese will make errors	Rejected. Can’t show that Harbinese make errors. ‘Similar’ is difficult can’t be proven.
	H1.6	/ʈʂ, ʈʂ^{h}, ʂ/	[tʃ, tʃ^{h}, ʃ] /ts, ts^{h},s/	/dʒ, tʃ, ʃ/	a. Both should be difficult. b. Cantonese will make more errors.	a. and b. both Rejected. Cantonese have higher accuracy rate than Harbinese. Two ‘similar’ types: the one with ‘allophonically identical’ seems easier than ‘phonemically similar’.
New	H1.7	/	/	/θ, ð/, /ʒ/, (/l/) [ɫ]	a. both will not make errors b. no differences	Both rejected. ‘New’ is confirmed to be difficult and not easy.

Table 4.73 Hypotheses for vowels

Category	No.	Harbinese ($L1_H$)	Cantonese ($L1_C$)	English (L2)	Predictions	Results of hypotheses
Identical	H2.1	/eɪ, aɪ, ɑu, oʊ (GA)/	/eɪ, aɪ, aʊ, oʊ (GA), ɔɪ/	/eɪ, aɪ, aʊ, oʊ (GA), ɔɪ/	Harbinese will make errors	Supported. Proven that 'identical' is easy.
Similar	H2.2	/	/ɔ/ [ɔ], [o]	/ɔ/	Harbinese will make errors	Rejected. 'Similar' (phonemically identical but allophonically different) is easy and 'new' is difficult.
	H2.3	[ɛ] /a/	/ɛ/ [ɛ], [e]	/ɛ/	a. Both should be difficult b. Harbinse will make more errors.	a. Rejected (but Harbinese had difficulty). b. Supported. 'Similar' in Cantonese is easy but difficult in Harbinese. Similar type of 'phonemically identical but allophonically different' is much easier than similar type of 'allophonically identical but phonemically different'.
	H2.4	/ə/ [ə,e,o,ɤ]	/	/ə/	Harbinese will make errors	Supported. 'Similar' is difficult. Not shown that 'new' is difficult.
	H2.5	[ɑ] /a/	/	/ɑ/	Harbinese will make	Rejected. 'New' is

					errors	difficult but 'similar' also difficult.
	H2.6	/	/ɐ/	/ʌ/	Cantonese will make errors	Rejected. 'New' is difficult while 'phonemically similar' is much easier than 'new'
	H2.7	/	/ɔ/	/ɒ/	Cantonese will make errors	Rejected. 'New' is more difficult than 'similar' and the result disagrees with Flege's SLM Varieties: /ɒ/,/ɑ/ both prefer British /ɒ/ and Cantonese significantly prefer British over Harbinese.
	H2.8	/ɚ/	/	/ɜ/-/ɝ/	Harbinese will make errors	Rejected. 'Similar' is easy, which contradicts Flege's idea that 'similar' is difficult. 'New' is difficult is not proven. More discussion needed!
	H2.9	/i, u/ /	/i, u/ [ɪ, ʊ]	/i, u/ /ɪ, ʊ/	Cantonese will make errors with tense-lax /i, u/ and /ɪ, ʊ/	Rejected. 'Similar' is not difficult but 'new' is. It seems that 'allophonically similar' is

						better.
New	H2.10	/	/	/æ/	a. both will not make errors b. no differences	Rejected. Proven that 'new' is difficult.

Table 4.74 Hypotheses on syllable structure are based on the following

Category	**No.**	**Harbinese ($L1_H$)**	**Cantonese ($L1_C$)**	**English (L2)**	**Predictions**	**Results of hypotheses**
Identical	H3.1	/(C)jV(C)/ (mostly in RP) or /(C)wV(C)/	/	/(C)jV(C)/ (mostly in RP) or /(C)wV(C)/	Cantonese will make errors	Rejected. 'Identical' is easy, and 'new' is also easy.
	H3.2	/	/CVm/	/CVm/	Harbinese will make errors	Rejected. 'Identical' is easy but 'new' is also easy. No significant difference. Needs to be discussed.
Similar	H3.3	/	/CVp̚,t̚, k̚/	/CVp, t, k/	Cantonese will make errors	Rejected. Similar is easy but new is difficult.
New	H3.4	/	/	/(C)VVn/	a. both will not make errors b. no differences	Rejected. 'new' is difficult which disagrees with Flege's SLM

Table 4.75 Hypothesis on stress is based on the following

Category	**No.**	**Harbinese ($L1_H$)**	**Cantonese ($L1_C$)**	**English (L2)**	**Predictions**	**Result of hypothesis**
Similar	H4	Two stress patterns with trochaic but main iambic foot	/	Various stress patterns with trochaic foot	Harbinese will make errors	Supported. 'Similar' is difficult and 'new' is easy, contradicting the findings from the above hypotheses, thus, needs to be discussed to discussed further.

Table 4.76 Hypothesis on L2 exposure length is based on the following

No.	**L2 Exposure length**	**Predictions**	**Result of hypothesis**
H5	Increase	Errors decrease	Rejected

Chapter 5. Discussion

Chapter 4 has presented the results for the English production by speakers of $L1_H$ and $L1_C$. This chapter discusses the results in light of the existing literature, starting with the application of Lado's CAH and Flege's SLM to predict patterns and to account for the results. I also discuss to what extent the present results corroborate those of others. I then turn to other possible influences on the patterns found in the data.

5.1 L1s' influence on L2

Various researchers since Weinreich (1953) and Lado (1957) (e.g. also Anderson 1987; Corder 1967; Ellis, 1994; Fisiak 1978, 1991; Gass, 1979; Mukattash 1981; Odlin, 1989; Rasier and Hiligsmann 2007; Young-Scholten 1985) have claimed that the L1 is the most important influence on the acquisition of an L2 phonology. However, very few studies have compared the influence of two varieties of the same language. Doing so allows us to see how each variety influence L2 phonology and what differences in L2 production may arise. From the results in Chapter 4, we can see an obvious and pervasive influence of L1 Chinese varieties on L2 English production. These appear for consonants, vowels, syllable structure and stress, even though the comparative status of stress in the two varieties is not clear. Besides the influence of L1, an influence of Standard Mandarin on English phonology is also observed. This is in line with previous findings (Mahmoud, 2000; Mohammed, 1992; Mukattash, 1981; Scott and Tucker, 1974; Tadros, 1966) that a standard variety also spoken by learners will also influence speakers' acquisition of a second language phonology.

5.1.1 Consonant results in comparison to previous studies

From the summary of results for consonants, we find many results conforming to the findings mentioned in Chapter 2. The results for the production of the English /v/ supports the finding that Mandarin/Harbinese speakers often substitute /v/ with [w] (Shi, 1996; Yin and Li, 2014). Liu and Wang's (2007) findings that Cantonese speakers often substitute /v/ with [w] or [f] are also supported. The results for the production of

/v/ in my research show that Cantonese speakers indeed made these errors, but Liu and Wang (2007) did not mention in what contexts their subjects substituted /v/ with [w] and /v/ with [f]. The different substitutions are constrained by the position in the syllable. My results clearly show that, when /v/ occurs word-initially and medially, the English /v/ sound is substituted by [w]. It is also substituted by [f] if /v/ occurs word-finally. My findings strongly support Chan and Li's (2000) discoveries on the production of the English /v/ by Cantonese speakers in Hong Kong.

The results for the production of English /w/ show that [v] is often used by Mandarin/Harbinese speakers to substitute for /w/, which confirms the findings by Kong and Wang (2001) and Zhang (2012). However, it is found in my research that a small number of Cantonese speakers substituted /w/ with [v] or [ɹ] . This has not been mentioned in the literature. As for the English /w/ and /v/, Gong and Tian (2008) and Chen (2010) claimed that mutual substitution is found among English learners in Northeast China. This can be confirmed by the results for /v/ and /w/, and further indicates that Harbinese speakers confuse /v/ and /w/, as mentioned by Shi (1996).

Chen (2010) and Yan and Yang (2011) mentioned that Northeastern dialect speakers have difficulty in distinguishing the apical fricative and affricative sounds /s/, /ts^{h}/, /ts/ from the postalveolar fricative and affricative sounds /ʂ/, /ʈʂ^{h}/, /ʈʂ/ in northeastern dialects; thus L1 transfer can be seen in their English production of /s/, /ts/, /dz/ and /ʃ/, /tʃ/, /dʒ/. This transfer cannot be seen in my research because Harbinese speakers do not have this contrast difficulty in their variety although they are from Northeastern China. However, it can be seen that the English /ʃ/, /tʃ/, /dʒ/ are often replaced by the similar Harbinese retroflex phonemes /ʂ/, /ʈʂ^{h}/, /ʈʂ/. Moreover, this often happens when these English consonants occur word-finally. Kong and Wang (2001) found that the Mandarin [ʂ] and [ɻ] are often substituted for the English /ʃ/ and /ʒ/ by Mandarin speakers, but my research found that /ʃ/ is substituted by the Mandarin [ʂ] when occurring word-finally, but it is pronounced accurately by my participants when occurring word-initially and medially. Moreover, [ɻ] is not the most frequently used substitute to replace the English /ʒ/. In my research, Harbinese speakers prefer to use [ɹ], followed by [ɻ]. Chan and Li (2000) and Liu and Wang (2007) found that Cantonese speakers often used [s] to substitute for /ʃ/, but the results here do not support this. Rather, /ʃ/ is often substituted by [ɕ] or [tɕʰ] among Guangzhou Cantonese speakers. Moreover, [ɕ], [tɕʰ], [tɕ] are frequently used by Guangzhou Cantonese speakers to substitute for the English /ʃ/, /tʃ/, /dʒ/ occurring word-finally. This differs from the findings of Chan and Li (2000) for Hong Kong Cantonese speakers who often substitute [s], [ts], [dz] for /ʃ/, /tʃ/, /dʒ/. These differences in production between Guangzhou Cantonese speakers and Hong Kong Cantonese speakers may be caused by the influence of Mandarin on Guangzhou Cantonese speakers, but this requires more investigation.

Transfer from the neutralisation of /n/ and /l/ in Cantonese to the production of the L2 English /n/ and /l/ has been mentioned by many researchers (Chen, 2013; Liao, 2014; Liu and Wang, 2007). This was also found in the present research, but it only accounts for a small percentage of errors made and only occurs with the initial /n/ and /l/. Liao's (2014) findings that [n] substituting for /l/ is more frequent than [l] substituting for /n/ cannot be supported in my research because /n/ and /l/ are both produced as the nasalised lateral, i.e. [l̃] instead. Liu and Wang (2007) mentioned that [l] is substituted for /ɹ/ by Cantonese speakers. My study confirms this, but it is not a commonly occurring substitution. Rather, [w] is the most frequent substitution, followed by [l]. The English /ɹ/ is often replaced by the Mandarin retroflex [ɻ] in Harbinese production. This is not mentioned in the literature either.

Among the consonants, the English interdental fricatives are worthy of consideration because they are less common in other world's languages (Wells, 1982:96), and thus are marked. As claimed by Eckman (1977), marked sounds will lead to difficulty in acquisition. The results of the production of interdental fricatives in this research can support Eckman's Markedness Differential Hypothesis. The results show that speakers of both varieties have low accuracy rates with English interdental fricatives. However, speakers do not always substitute the same L1 consonants for interdental fricatives. This so-called differential substitution has attracted great attention amongst researchers and the results of my study corroborate the existence of different L1-based substitutions.

To begin with, in research on L1 English acquisition, the interdental fricatives /θ/ and /ð/ are only acquired later by native English children (Wells, 1982:96), before which they typically use [f]. Variants are also used by native English speakers in some varieties. For example, [f] and [v] are substituted for /θ/ and /ð/ by speakers of London Cockney (Wells, 1982:328) and by Australian English speakers (Horvath, 1985:96-97, cited by Dubois and Horvath, 1998:248). [f] is also reported to be used for /θ/ in southern American dialects, but only in the final position (Morgan, 1971:277-278, ibid.); for example, in central Kentucky (Davis, 1971, ibid.). These fricatives have been found to be difficult for L2 English learners but with different substitution patterns found by researchers. Speakers of Japanese, German, French and Egyptian Arabic all use [s] and [z] substitutions (Kohmoto, 1965, Ritchie, 1968, and James, 1984, cited by Weinberger, 1997:269; Lombardi, 2003:225), and Thais, and Russians use [t] and [d] for /θ/ and /ð/ (Weinreich, 1968, cited by Weinberger, 1997:269; Lombardi, 2003:225) while the Polish and Slovakians (Lombardi, 2003:225) use [f] for /θ/.

For Chinese, as noted by Weinberger (1997:269), Fonda (1984) found that Mandarin speakers substituted [t] for /θ/ in the onset position and [f] in the coda position. Unlike Fonda, Rau, Chang and Tarone (2009:582) found that Chinese Mandarin speakers substituted [s] most frequently for /θ/. /θ/ and /ð/ were found by Shi (1996) to be often substituted by [s] and [d] by Mandarin speakers, and Chen (2010) and Zhang (2012) found that speakers of Northeastern varieties often substituted [s] and [ts] for /θ/. Wang (2015:87) agreed with Shi's findings (1996), but besides [s] and [d] as the most frequently used substitutions for /θ/ and /ð/ she discovered that Mandarin speakers also substituted [t] for /θ/ and [z] as well as [l] for /ð/. The results in my research agree with the findings of Rau et al. (2009), Shi (1996), Chen (2010), Zhang (2012) and partly support Wang's (2015) findings, but refute Fonda's (1984). The substitution [t] never occurred in the production of Harbinese or Cantonese speakers across all positions, and [f] was also used by Cantonese speakers. It is unclear whether or not Fonda collected the data from participants who speak Mandarin as an L1, or who could speak Mandarin along with other Chinese dialects.

The /θ/ and /ð/ in my study were most frequently substituted with [s] and [ts]/[d] by Harbinese speakers. It needs to be noted that [d] was only used to replace the initial /ð/, but [s] and [ts] were used across all positions. Patterns of the substitution of /θ/-/ð/ by Cantonese speakers have been found to differ by different researchers. So, [s]-[z] are substituted for /θ/-/ð/ (Liu and Wang, 2007), [s] and [f] are (but [s] is more frequently) for /θ/ and [d] for /ð/ (Yang, 1997; Liu and Guo, 2013; Zhang, 2017), [d] and [z] are for /ð/ (Zhang, 2017). Substitutions are also made by Hong Kong Cantonese speakers with [f] for /θ/ (Deterding, Wong and Kirkpatrick, 2008; Peust, 1996, cited by Rau, Chang and Tarone, 2009:582), and with [t]/[f] for /θ/ and [d]/[f] for /ð/ (Chan and Li, 2000). My findings do not completely agree with these studies. [s] was found to be a substitute for /θ/ in my study but only accounts for a small percentage of errors. The most frequently used substitution was [f] for /θ/ across all positions, which agrees with the findings of Deterding, Wong and Kirkpatrick (2008) for Hong Kong Cantonese. Meanwhile [d] was most frequently substituted for /ð/ in my study, but never when occurring word-finally. When /ð/ did occur word-finally, it was substituted by [f] (39% of the cases) or [θ] (30%) in my research. My findings largely support Chan and Li's concerning substitutions in Hong Kong Cantonese but one important difference is that Guangzhou Cantonese speakers never use [t] to substitute for /θ/. The present research also supports Li's claim (2006) that new consonants in English such as /θ/, /ð/, /ʒ/ are difficult for L1 Chinese to learn and they make many mispronunciations. Both dialect/language groups in the present study achieved low accuracy rates in pronouncing them, which supports Lado's CAH (1957) in that L2 sounds which are new to the L1 are difficult for L2 learners to acquire.

It can be seen from the production results of consonants that the findings of the present research support Corder's claim (1967) that errors made by L2 learners are not only caused by L1 transfer but are also due to universals such as markedness. In contrast, the findings also contradict the CAH's claim that all errors can be explained by L1 transfer. Larsen-Freeman and Long (1991) also thought that positive transfer will occur when an L2 is similar to the L1, but negative transfer will happen when they are different. The results do not refute their idea but it should be noted that not all errors are caused by negative transfer.

For interdental English consonants, the literature on the acquisition of English by Chinese learners (Zhang, 2012; Yang, 1997; Liu and Guo, 2013; Zhang, 2017) mentions that /θ/ and /ð/ are substituted by Harbinese speakers with [s] and [t͡s], which exist in Harbinese, and by Cantonese speakers with [f] and [d]. It can be confirmed that the present results for the production of /θ/ and /ð/ support those findings. The errors mentioned in the literature do occur in the production in this research, and the influence of L1 dialects on L2 English is evident. However, interestingly, the results of this study indicate that the substitution patterns are complex. Not only are the substitutions mentioned above produced, but others were also produced, especially the voiced interdental fricative /ð/. In the Cantonese group, the most frequently pronounced substitution is with [d], as mentioned in the literature, and [d̪] is the second most frequent substitution when /ð/ occurs word-initially and medially. However, when /ð/ occurs word-finally, the most frequent substitute is /f/, followed by /θ/. Unlike with the Cantonese group, the Harbinese group use different substitutions. Besides [t͡s] mentioned in the literature, another substitute for /ð/ is [z]. These two occur across all positions. However, when /ð/ occurs word-initially in the Harbinese group, it is more difficult to see that there are two sets of main substitutions. One set is [d] and [d̪], which is the same as the substitutions used in the Cantonese group; the other set is [t͡s] and [z]. The most frequently used substitute for /ð/ is [d] (26%), followed by [t͡s] (23%), and [z] and [d̪] (both 5%).

From these errors, we can see that the substitutions with [t͡s] and [f] must be the outcome of the direct influence from $L1_H$ and $L1_C$ respectively, because these items exist in their own dialects/languages: [t͡s] in Harbinese and [f] in Cantonese. Although [f] also exists in Harbinese, it does not occur in word-final positions, which implies that certain underlying grammatical rules constrain the use of [f] as the optimal candidate in Cantonese production for the word-final /ð/. In addition, substitution with [z] in the Harbinese group and with [d̪] in both groups implies that L2 English learners are acquiring and approaching the target language but have not yet completely acquired its sounds. That is to say, they are still in the process of learning the second language. They are at an in-between, or intermediate stage, and neither at the destination nor the starting line. The use of [z] and [d̪] here can be regarded as outcomes of the development process for speakers of each dialect/language.

The reason for this is that [z] does not exist in Harbinese but it does in English and [d̪] is not a sound in either dialect/language but has a dental feature closer to the English /ð/. [z] and [d̪] have features more similar and with a closer distance to the target /ð/ than do [t͡s] and [d]. This development can be clearly shown in the comparison of the phonological features of these errors and the English /ð/ in Table 5.1. The Harbinese group use two sets of substitutes in the word-initial position, only one of which is the same as the substitute in the Cantonese group; therefore, the differential substitutions in the two dialects/languages is demonstrated here.

In the Harbinese group, the development can be seen from the substitution of [t͡s] with three features different from /ð/ to error [z] with only one distinct feature. Likewise, it can be seen in the Cantonese group. The common error [d] has two features which are different from /ð/ but [d̪], which may be upgraded from [d], has only one feature distinct from /ð/. The observed phenomena that can be explained if a certain proportion of speakers are in the midst of a developmental process of acquisition.

Table 5.1 Comparison of English /ð/ and substitutions by phonological feature

	Harbinese Production		English sound	Cantonese Production	
Features	Error 1	Error 2		Error 1	Error 2
	[t͡s]	[z]	/ð/	[d]	[d̪]
1.[Syllabic]	−	−	−	−	−
2.[Consonantal]	+	+	+	+	+
3.[Sonorant]	−	−	−	−	−
4.[Coronal]	+	+	+	+	+
5.[Anterior]	+	+	+	+	+
6. [Distributed]	−	−	+	−	+
7.[Continuant]	−	+	+	−	−
8.[Strident]	+	+	−	−	−
9.[Lateral]	−	−	−	−	−
10.[Delayed release]	+	−	−	−	−
11.[Nasal]	−	−	−	−	−
12.[Voice]	−	+	+	+	+
13.[High]	−	−	−	−	−
14.[Back]	−	−	−	−	−

Table 5.2 Comparison of English /θ/ and substitutions by phonological feature

	Harbinese Production	English sound	Cantonese Production
Features	Error		Error
	[s]	/θ/	[f]
1.[Syllabic]	−	−	−
2.[Consonantal]	+	+	+
3.[Sonorant]	−	−	−
4.[Coronal]	+	+	−
5.[Anterior]	+	+	+
6.[Distributed]	−	+	−
7.[Continuant]	+	+	+
8. [Strident]	+	−	−
9.[Lateral]	−	−	−
10.[Delayed release]	−	−	−
11.[Nasal]	−	−	−
12.[Voice]	−	−	−
13.[High]	−	−	−
14.[Back]	−	−	−

Actually, both dialects/languages have [t͡s] and [d]; however, the Harbinese speakers favour [t͡s] and the Cantonese speakers prefer [d]. From the comparison of the features in Table 5.1, it can be seen that Harbinese and Cantonese speakers have specific preferences for phonological features in perceiving sounds. In the perception of the English /ð/, the [+strid] feature is important to the Harbinese group while the [-cont] feature is significant to the Cantonese group. It is the reason why Harbinese pronounced /ð/ as [t͡s] or [z] while Cantonese as [d] or [d̪]. We can also understand it in the other way: the feature [-strid] to English /ð/ was paid less salient by Harbinese speakers and they feature [+cont] by Cantonese speakers. Thus, it is the reason why they both did not produce English /ð/ correctly. It is the differences between these two main phonological features that lead to the different productions of /ð/ by speakers of the two groups. Questions might be asked concerning how important stridency and continuance are, and why the [+strid] feature is crucial to Harbinese speakers (or [-strid] is less salient) while the feature [-cont] is to Cantonese speakers (or [+cont] is less salient). This is a subject for further research. Possibly, the rankings of constraints underlying each dialect/language may differ. It might be that feature stridency plays a more dominant role in the constraint hierarchy in Harbinese, while continuance is supreme in the Cantonese group, and thus different optimal productions are generated in each dialect/language. These questions cannot be answered scientifically at present. The problem needs more discussion, but a review of previous linguistic research may provide ideas to help explain the results of the present study.

Although researchers are very interested in the reasons why different substitutions are used by L2 English learners speaking different languages, the differentiate substitution has been a conundrum for decades (Weinberger, 1997:305). Weinreich (1953/1968, cited by Weinberger, 1997:264) claimed that the CAH method was beneficial to teachers but it could not answer the differential substitution question. He thought that structural analysis offered the most promising possibility. Despite various studies using structural solutions, no one could properly explain the phenomenon (Weinberger, 1997:305). Weinberger (1997:265, 304-305) argued that transfer theory was useful in explaining L2 errors. However, he maintained that differential substitution was not a simple process of over-transfer. Instead, he proposed Underspecification Theory as an appropriate approach to explain the differential substitution. He (1997:306) thought the acquisition of new phonology is a learning process where only a minimal amount of information is required by learners. For example, it is sufficient for learners to know key features such as [-sonorant] and [voice] which lead to the different substitutions of /θ/. Lombardi (2003:227) holds a different opinion although she agrees that Weinberger's analysis was a great breakthrough at the time. However, Weinberger's method was not universally accepted and he did not clearly explain how learners can achieve the correct underspecified representations. Moreover, the requirement for an L2-specific mechanism makes feature pruning more complicated. Therefore, Lombardi adopted Optimality Theory (OT) to explain [s] and [t] substitutions for /θ/. She argued that OT can give a satisfactory and precise explanation of these differential substitutions. She explained (2003:225, 247) that highly ranked markedness leads to [t] substitution, which represents an action of universal substitution, while highly ranked faithfulness results in [s] substitution which is an indication of transfer. In an investigation of substitution for /θ/ and /ð/ by Dutch speakers under the framework of Optimal Theory (OT), Wester, Gilbers and Lowie (2007:477) predicted substitutions from acoustic, phonological and segmental points of view. They predicted that [f] and [v] would be the substitutions via an the acoustic comparison of [f] and [s] with /θ/, with the results for [f] more acoustically similar to /θ/ than [s]. According to the most phonologically distinctive features, they predicted that [s] and [z] are phonologically the most similar sounds to /θ/ and /ð/ because only the [Strident] feature distinguishes between these two pairs. In addition, according to segmental markedness where /θ/ and /ð/ are more highly marked while [t] and [d] are the least marked, [t] and [d] are predicted segmentally to be the least marked sounds. Based on these characteristics, authors conducted an OT analysis of the substitutions of interdental fricatives (Wester et al., 2007:485-489). They explained that the reasons for different substitutions for /θ/ were because the substitution with [f] is dominated by the constraint of maximal acoustic CORRESPONDENCE;

while the ranking for the [s] substitution is CORRESPONDENCE CONTINUANCY >> SEGMENTAL MARKEDNESS >> CORRESPONDENCE STRIDENCY; and the ranking for [t] is SEGMENTAL MARKEDNESS >> CORRESPONDENCE CONTINUANT >> CORRESPONDENCE STRIDENCY. The tableau in Tables 5.3 and Table 5.4 show the constraint rankings in Wester et al. (2007) .

Table 5.3 Tableau of Dutch substitution of [s] for /θ/

Input: /θ/	CORRESPONDENCE CONTINUANCY	SEGMENTAL MARKEDNESS	CORRESPONDENCE STRIDENCY
[θ]		**!	
[s] √		*	*
[t]	*!		

Table 5.4 Tableau of Dutch substitution of [t] for /θ/

Input: /θ/	SEGMENTAL MARKEDNESS	CORRESPONDENCE CONTINUANCY	CORRESPONDENCE STRIDENCY
[θ]	**!		
[s]	*!		
[t] √		*	*

It should be noted that the constraints for segmental markedness are different. A hierarchy representing these from the least to the most constrained is deduced in Table 5.5 from a tableau used by Wester et al. (2007:487).

Table 5.5 Hierarchy of segmental markedness

t	0 violation
s, d	1 violation
θ, f, z	2 violations
ð, v	3 violations

Wester et al. thought that the constraint CORRESPONDENCE VOICE dominates the constraint SEGMENTAL MARKEDNESS for the [d] substitution for /ð/. Dutch speakers prefer ([t]/[d]) stops initially and the fricative ([f]) finally (Wester et al., 2007:489). They mentioned that "HCODA prefers fricatives to plosives syllable-finally and HONS prefers plosives to fricative syllable initially" (ibid.) and the ranking is SYLLABLE-FINAL DEVOICING >> HCODA, CORRESPONDENCE VOICE. However, if [d] is a syllable-initial substitution, they claimed that SYLLABLE-CORRESPONDENCE VOICE dominates while SEGMENTAL MARKEDNESS and HONS dominate if [t] is the syllable-initial substitution.

From these proposals, it seems that the results of my research could go some way to explaining why the Harbinese use [s] and the Cantonese use [f] to substitute for /θ/. The main reason would be the different underlying rankings in L2 acquisition. As in Wester et al. (2007:485-489), the constraint rankings could be used to explain that the Harbinese [s] substitution is identical with the ranking for Dutch speakers, so that for Harbinese CORRESPONDENCE CONTINUANCY is ranked higher than SEGMENTAL MARKEDNESS which is higher than the constraint CORRESPONDENCE STRIDENCY. However, Wester et al. did not fully explain [f] substitution, but only claimed that it requires highly ranked phonetic CORRESPONDENCE. Therefore, the Cantonese [f] substitution cannot be explained in terms of the findings of Wester et al. However, differences between Harbinese and Cantonese influence can be observed from the feature analysis of English /θ/ in Table 5.2, where [+strid] and [-cor] are crucial features for both groups to pronounce /θ/ successfully. Nevertheless, for Harbinese speakers the salient feature is that it is [+cor] and they do not have its [-strid] feature as salient, whereas for Cantonese speakers it is the feature [-strid] which is salient and so accurately produced and they do not have [+cor] as salient. Therefore, differential substitutions were produced by the two groups of speakers. If we were to apply Optimality Theory (Smolensky and Prince, 1993), we can see that the different constraint rankings on features in the minds of different dialect/language speakers might eventually lead to the differential substitution of the interdental voiceless /θ/.

However, when /ð/ occurs word-initially, unlike with the simplicity of Dutch substitutions for /ð/, Harbinese and Cantonese have the most frequent substitute of [d], as Dutch does; however, Harbinese has another frequently used [ts] substitution and it occurs across all positions. Meanwhile, Harbinese has [d] as substitute only in the word-initial position while Cantonese uses it frequently word-initially and word-medially. Apart from the Harbinese [ts] substitution occurring in all positions, [z] is also frequently used word-medially and word-finally. Unlike in Harbinese, Cantonese uses [f] and [θ] frequently in word-final positions. Obviously, the substitutions made by Harbinese and Cantonese speakers in my research are very complex. Previous findings cannot fully explain the reasons for these differential substitutions. Moreover, due to time limitations, attempts to explain the results for /θ/ and /ð/ must be deferred. From the previous research on differential substitution, OT seems to be the most appropriate tool to find solutions. Thus, OT will be used to conduct detailed future research on the differential substitutions of /θ/ and /ð/ in different positions.

5.1.2 Influence of the Chinese varieties on L2 vowels

The influence of the L1 on L2 vowels can also be seen and differences in production occurred between speakers of Harbinese and Cantonese. There is no /ɔ/ in Harbinese, and it is found that Harbinese speakers often confused /aʊ/ and /ɔ/, and often mutually substituted one for the other. Moreover, the Harbinese /ɑʊ/ is similar to the English /aʊ/. L1 transfer can be detected again in the present research when the Harbinese/Mandarin /ɑʊ/ is often used to substitute for the English /aʊ/ and /ɔ/. This finding is in line with those of (Shi, 1996) where /ɔ/ is often pronounced as /aʊ/ by Mandarin/Harbinese speakers.

The English /æ/ is absent in both dialects/languages, which leads to difficulties for their speakers. Shi (1996) and Zhang (2012) claimed that [ai] is used to substitute for /æ/ by Mandarin speakers but Kong and Wang (2001) found that [ɛ] is often used . My research findings are slightly different, in that Harbinese speakers confused English /æ/, /ɛ/ and /ai/ and mutually substituted one for another. Cantonese speakers could easily distinguish monophthongs from diphthongs, and thus they did not have this problem. However, it was found that they could not recognise differences between /æ/ and /ɛ/, which agrees with the findings of Zhang (2017). The [ɛ] vowel was often found in my research to replace the English /æ/ by 75% of Cantonese participants, but this contradicts Zhang's finding (2017) that 54% of his subjects pronounced /ɛ/ as [æ]. The reason why Harbinese and Cantonese speakers have different difficulties in pronouncing /æ/ may result from the different influences of their L1, since Cantonese /ai/ contrasts with the length of /a/ but the Harbinese /ai/ does not. Thus, Cantonese speakers do not confuse /æ/ and /ɛ/ with /ai/, but Harbinese speakers do. Chen (2010) claims that Mandarin diphthongs are produced with a rapid transition from one vowel to but English diphthongs are pronounced more slowly with clear transitions gliding from one vowel to the other. This may be one reason why Harbinese speakers in this study had difficulties in distinguishing /ai/ from /æ/ or /ɛ/. Similarly, it was also found that Harbinese speakers confused /i/, /ɪ/ and /eɪ/ but the Cantonese only confused /i/ with /ɪ/.

Cantonese has more vowels similar to English than does Harbinese. Liu and Wang (2007) claimed that the effect of negative transfer to English vowels is comparatively less significant. Yang (1997) agreed, claiming that Cantonese speakers have fewer difficulties in learning English vowels than Northern Mandarin speakers. Seen from the perspective of vowel production, the present research confirms their findings.

5.1.3 Influence of the Chinese varieties on L2 syllable structure

Harbinese does not allow any consonants as coda except for the two nasals /n/ and /ŋ/. Thus, $L1_H$ transfer was used in pronouncing English words with closed syllables. This study has found that a vowel is often inserted after the coda in closed syllables. This finding supports those of Shi's research (1996) on English syllable structure from Mandarin speakers. It is found that the vowels mostly inserted by Harbinese speakers before the coda are [ə] or [u] (with [ə] used more than [u]), which further confirms Chen's (2010) findings that English learners in the Northeast often insert a [ə] or [u] after a coda of closed syllables. Another reason why Harbinese speakers prefer to insert vowels after closed syllable could be $L1_H$ transfer since Harbinese (Mandarin and

Northeastern dialects) prefers open syllables (Shi, 1996; Chen, 2010).

Cantonese allows the unreleased plosives /p/, /t/, /k/ as codas. Chan and Li (2000) and Jia (2011) found that the transfer of $L1_C$ tend to be exerted on English and Cantonese speakers often unrelease the plosive codas in English syllables. The results in my research confirm their findings. However, the present research does not indicate that a very large percentage of such errors were made. My results show that the bilabial stop /p/ is much more likely to be unreleased than the other two.

Syllable structure with a diphthong plus a /n/, or, /VVn/, is not allowed in either dialect/language. Many researchers (Kong and Wang, 2001; Gong and Tian, 2008) mention that Mandarin/Northeastern dialect speakers have difficulties in pronouncing English words with this syllable structure. Chan and Li (2000) also pointed out that Cantonese speakers have problems pronouncing English words with this pattern, such as [meɪn] 'main'. The present research found that syllable structures with /VVn/ are indeed difficult for L2 Chinese learners of English. Both groups achieved low accuracy rates, especially the Harbinese group. This supports the findings of Kong and Wang (2001), Gong and Tian (2008), and Chan and Li (2000). Kong and Wang (2001) that /aʊn/ is often replaced by [ɑŋ], which is influenced by L1 transfer because Mandarin/Northeastern dialect only allows syllable structures such as /Vn/ or /VV/. Their findings can be confirmed in that up to 68% of Harbinese participants pronounced /aʊn/ as [ɑŋ] in my study. From the production results, errors with the syllable structures /Vn/ and /VV/ account for most errors, except for /ɔɪn/ in the Harbinese group who preferred [ɔən] more often. It is clear that L1 syllable structure influences the acquisition of L2 English phonology. Chan and Li (2000) also gave another explanation for this syllable structure being difficult, which is because it is the most marked. This finding once again supports Lado's assumption that the 'new' is difficult.

5.1.4 Influence of the Chinese varieties on L2 English stress

The influence of the L1 on L2 English stress can be seen clearly from the results shown in Figure 5.1. Harbinese speakers tend to assign stress to the second syllable in disyllabic English words to form an iambic foot, which is quite similar to the stress pattern in Harbinese. Although stress is not a very typical feature of Harbinese/Mandarin, Duanmu (2000:Ch6; 2007:Ch6) claimed that Mandarin does have stress and the main stress pattern is in the iambic foot as mentioned in the literature review.

From Figure 5.1, we can see that $L1_H$ exerted a great influence on Harbinese speakers' stress assignment, with more than one-third of English tokens stressed on the second syllable. However, a small proportion of Cantonese speakers assigned stress to the second syllable. Why did they place stress similarly to Harbinese speakers? It might be because they were influenced by Standard Mandarin. As mentioned many times above, Harbinese is one of the varieties of Mandarin. That Harbinese speakers have a tendency to place stress on the second syllable agrees with the findings of Xu and Song (2008:66), Gao and Deng (2009:12), Ma and Wang (2013:150), and Yuan and Wu (2017:87). Their research on stress placement of disyllabic English words and polysyllabic words found that Chinese university L2 English learners prefer to stress the second syllable of English words. The main reason for such stress errors can be attributed to L1 transfer. Ma and Wang (2013:149) thought that it may be influenced by the Chinese syllable rhythm and cited Li's research to confirm this. Li (2012, cited by Ma and Wang ibid.) conducted research into Mandarin word stress from a phonological perspective and found that the foot in Mandarin is iambic. The results of my study further confirm this. Moreover, these researchers also think that misplacements of word stress may be caused by intralingual factors such as the complexity of the rules of English stress assignment. The present study confirms their findings and clearly identifies the transfer exerted by the L1 on English stress placement.

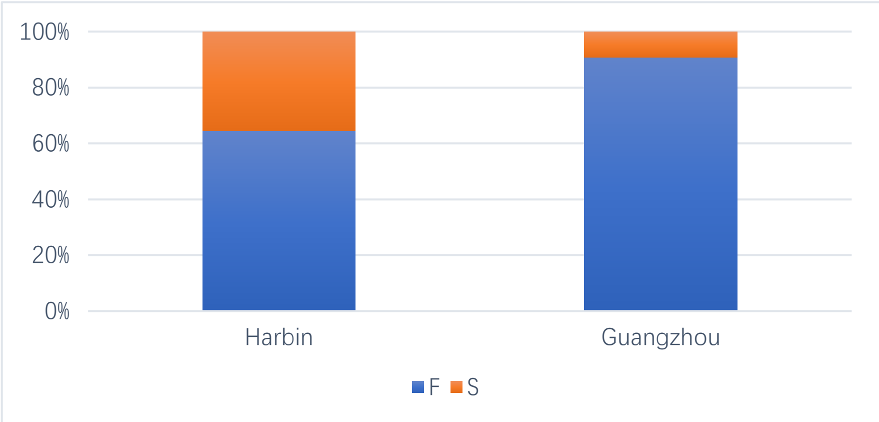


Figure 5.1 Frequency and percentage of stress assignment based on pitch only

The narrow focus on stress in this research needs to be explained. The main reason is that stress assignment is an important issue for Chinese English learners, as mentioned by many researchers (Xu and Song, 2008; Gao and Deng, 2009; Ma and Wang, 2013; Yuan and Wu, 2017). In addition, I wonder if all Chinese people speaking different Chinese dialects/languages have the same problem when they learn a second language. This has not been addressed in previous research. Chinese English learners are usually considered as a homogeneous group, disregarding dialect/language differences as well as the different influences on learners because of their dialects/languages. Research is needed on this point. Stress is known to be a very complicated aspect of phonology. Such a study of stress would make a doctoral thesis by itself. However, time is limited and stress is not the main research focus here. For these reasons, I have concentrated only on a narrow focus in the research.

From the above discussion on segments, syllable structure and stress, it can be concluded that distinct influences of regional varieties on L2 phonology exist even when speakers also know the standard variety. Moreover, this influence is felt in every aspect of L2 phonology. I now turn to discussion of evidence for the influence of Mandarin.

5.1.5 Influence of Mandarin

Mandarin is the official language in China used in schools, universities, the media and elsewhere for official purposes. The issue of the influence of a standard language on a regional variety often arises in studies of L2 acquisition. For example, even though Modern Standard Arabic is spoken by no one as a native speaker (Cowan, 1968; Mahmoud, 2000:129), researchers find its influence widely in L2 acquisition. Arabic learners of English always speak a variety of Arabic (abbreviated as NSA for non-standard Arabic) and also have to learn Modern Standard Arabic (MSA). There are numerous differences at all linguistic levels between MSA and NSA, as mentioned by Thompson-Panos and Thomas-Ruzie (1983; cited by Mahmoud, 2000:129). Many examples are given by researchers to demonstrate the influence of the MSA on the English acquisition of NSA learners of English (Scott and Tucker, 1974; Mukattash, 1981; Mohammed, 1992; Tadros, 1966; Mahmoud, 2000) although some of these studies have tested the influence of MSA only in syntactic and lexical domains of English because NSA is used in speaking and MSA for writing.

However, unlike Arabic, alongside regional Chinese dialects/languages acquired from birth and used is speaking in daily life, the Chinese standard variety of Mandarin is learnt and used for both speaking and writing. Therefore, the influence of Mandarin on English should be readily detectable in learners of English in the phonological domain as well. As mentioned above, the results show some influence of Mandarin in the production of L2 English, and this is more noticeable in Cantonese speakers. It is not detectable in Harbinese speakers because Harbinese is a variety very close to Standard Mandarin, and is even considered to be a variety of Mandarin which some claim is closer to Standard Mandarin than other varieties, including the Beijing dialect (Nie, 2005a:67; 2005b:35). It is therefore difficult to distinguish between the influence on English of standard Mandarin or Harbinese Mandarin. This section therefore mainly considers the influence of Mandarin in general on Cantonese speakers.

Such influence can be detected not only in the results for the production of consonants and vowels but also syllable structure and stress as in examples in Tables 5.8 and 5.9. It can be seen that there was a great influence of Mandarin on Cantonese speakers'

English consonants. The alveo-palatal fricative /ɕ/ and alveo-palatal affricates /tɕ/ and /tɕʰ/ are three consonant phonemes in Mandarin which do not exist in Cantonese, but they occurred in the errors made by Cantonese speakers in the substitution of the English consonants /ʃ/, /tʃ/ and /dʒ/. Interestingly, this often happened when the English consonants /ʃ/, /tʃ/ and /dʒ/ occurred word-finally, but also with /ʃ/ in the initial position. Another evident influence is the effect of retroflex consonants in Mandarin which can be found in the errors with /dʒ/ and /ɹ/ such as pronouncing /ʈʂ/ and /ɻ/. However, these retroflexes are absent in Cantonese. Cantonese also has the same glottal /h/ as English, but a small number of Cantonese speakers pronounced the English /h/ as [x], which exists in Mandarin.

Apart from the four retroflex consonants, Mandarin also has a rhoticised schwa /ɚ/ as a single syllable. These are typical Mandarin segmental features which do not occur in Cantonese. The examples of vowels and syllable structure in Table 5.7 represent the clear influence of Mandarin on Cantonese speakers in their pronunciation of English vowels. Not only the schwa, but also the vowel /ɔ/ was rhoticised by Cantonese speakers as for Harbinese speakers. However, Harbinese speakers exhibited a larger percentage of vowel rhoticisation than Cantonese speakers. In addition to the influence on vowel rhoticisation, the substitution of [ɤ] for the English /ɒ/ and of [ɑʊ] for the English diphthong /aʊ/ are not sounds in Cantonese, but they are in Mandarin.

The influence of Mandarin on Cantonese speakers in English syllable structure was also visible. A vowel was often inserted by Mandarin speakers after the coda in English closed syllables due to the preference for open syllables, as mentioned in the literature (Shi, 1996:90). Moreover, the inserted vowels /ə/ or /u/ are often observed in examples given by other researchers (Dou, 2010:155). A typical error of the English syllable structure /CVk/ produced by Cantonese speakers is that the coda /k/ is unreleased. However, vowel insertion can be seen after the coda /k/ of /CVk/ in Table 5.7. This is a typical error often made by Mandarin speakers (Shi, 1996; Dou, 2010). Mandarin speakers tend to insert vowels after codas, which may be constrained by the preference of open syllables in Mandarin (ibid.) or by the perceiving of codas as onsets. Due to Mandarin as a lingua franca in China, Cantonese speakers who can also speak fluent Mandarin made the same error as Harbinese Mandarin speakers did in this research. This can imply that the influence of Mandarin was also apparent in Cantonese speakers. In addition, the influence of Mandarin on Cantonese efforts at English syllable structure can be see in the production of the English consonants /l/ and /dʒ/ in Table 5.6. Vowel epenthesis was implemented after these consonants when the consonants occurred word-finally. We cannot be sure if the insertion of the vowel [y] was influenced by Mandarin because it never happens with Mandarin/Harbinese speakers, but it is certain that the insertion of a retroflex vowel represented here by the syllabic retroflex approximant [ɹ̩] would have been influenced by Mandarin.

From these examples, we can see that Mandarin does exert an influence on L2 English, although the influence is relatively insignificant compared to the influence of the L1 native dialect. This also supports Mahmoud's conclusion (2000:134) that the influence of Arabic on Arabs learning English represents interlingual transfer not only from NSA but also from MSA; moreover, the influence of the native dialect than from the standard one.

Table 5.6 Examples of influence of Mandarin on English consonants of Cantonese speakers

Consonant segment	Guangzhou		
	Initial	Medial	Final
/ʃ/	***ɕ***-1%		***ɕ***y̥-2% ***tɕʰ***y̥-2%
/tʃ/	-	-	***tɕʰ***y-10% ***tɕʰ***-1%
/dʒ/	-	-	***tɕ***y̥-8% ***tɕʰ***y̥-8% ***t̺ʂɻ̩***-1%
/h/	***x***-2%	***x***-1%	-
/l/	-	l***ə***-2%	l***ə***-1%
/ɹ/	-	***ɻ̩***-1%	-

(The segments influenced by Mandarin are represented in bold and italics)

Table 5.7 Examples of influence of Mandarin on English vowels and syllable structure of Cantonese speakers

Vowel and syllable structure	Guangzhou
/ə/	***ɚ*** - 2%
/ɔ/	***ɔ˞*** - 2%
/ɒ/	***ɤ*** - 3%
/aʊ/	***aʊ*** - 2%
CVk	g***ə*** - 1%

(The target items influenced by Mandarin are represented in bold and italics)

5.2 Application of Lado's CAH and Flege's SLM[12]

Research on the acquisition of L2 phonology has been conducted over many decades. Various important theoretical frameworks or models have emerged and played a substantial role in the development of knowledge of L2 acquisition. As mentioned in the literature review, Lado's CAH (1957) is one such theory. The CAH (Lado 1957) claimed that items 'similar' between the L1 and L2 are easy to learn but it is difficult if they are different, that is, 'new'. Although Lado's CAH has been challenged, Contrastive analysis is still a useful method applied in different fields of research into L2 acquisition by many researchers (Anderson, 1987; Fisiak, 1978, 1991; Nemser, 1961; Politzer, 1968; Stockwell and Bowen, 1965; Wardhaugh, 1970; Yarmohammadi, 2002). However, Flege's Speech Learning Model (1995) is another useful model widely used in a large body of empirical research into the acquisition of L2 sounds. The SLM (Flege, 1995) claims that L2 sounds which are similar to L1 sounds are more difficult to learn than those that are new in the L1. As mentioned in Chapter 3, the present research has combined CAH and the SLM, which hold completely opposite ideas on theoretical grounds, as a way of compensating for the deficiencies of the CAH. Moreover, this research has also investigated the production of English

[12] It is worth pointing out that Lado's CAH has an extensive range of applications, which can be used not only in the comparison of phonology, syntax, and lexis but also various fields in linguistics and even translation (Toury 1995, Venuti 2000, cited by Kramsch, 2007:245), computer corpora (Ajmer, Alenberg and Johansson 1996; Johansson and Oksefjell 1998, ibid.), and L2 classroom teaching (Blyth 1995; Belz 2003, ibid.). However, Flege's SLM has comparatively narrow application suitable for the phonetic level of L2 speech.

targets by participants speaking different Chinese dialects/languages to test these two models and to determine which model better predicts difficulty for L2 learners and how different dialects/languages influence their production.

The questions asked in this study concerning which category is more difficult than the other, 'similar' or 'new', and which theoretical idea is more suitable for my research, CAH, SLM or both. Seen from the summary of results of all hypotheses, they both have been answered. Many of the results of my research support Lado's CAH (1957) in that 'similar' is easy but 'new' is difficult. On the other hand, the results suggest that Flege's SLM that 'similar' is more difficult than 'new' (Flege, 1995) does not apply very often. The results do not support the assertion that 'similar' is difficult and 'new' is easy, because most hypotheses where difficulty was predicted with 'similar' items were refuted and all of the hypotheses predicting no difficulties with the category of 'new' were also rejected. It can be seen from the actual pronunciation results that 'new' leads to much higher error rates than 'similar' for speakers of both dialects/languages, which means that 'new' for L2 learners, especially in the context of classroom, is more difficult than 'similar'. These findings show that, in researching L2 acquisition and predictions of the general difficulties L2 learners may have, contrastive analysis is still more successful although not all errors result from L1 interference and not all difficulties can be accounted for by the CAH. CA is a very important and effective tool in research into L2 teaching and learning, especially in L2 classroom. In addition, the present findings are in line with Ouyang's research (2018) on the acquisition of Chinese EFL learners' dental fricative, where Lado's CAH was borne out by her results which showed the more different L2 sounds are for L1 learners, the more difficult they are to acquire.

This does not mean that Flege's SLM should be completely dismissed for L2 learners at lower proficiency levels. It may play a role in solving some L2-specific phonetic problems. The findings of the present research support the importance of phonetic perception between differences of L1 and L2 (Best, 1993; Flege, 1995). Flege (1995) claimed that phonetic perception influences the establishment and acquisition of listeners' L2 sounds, as shown by the results relevant to H1.7 concerning /l/ [ɫ] in this research (see section 4.2.3.1). Although both Harbinese and Cantonese have the same phoneme /l/ as in English, they do not have the allophonic [ɫ]. Due to the absence of a phonetic perception of the allophone [ɫ], this phonetic sound cannot easily be produced. This is perhaps the reason why both Harbinese and Cantonese speakers made many errors in producing this sound. The error types for /l/ [ɫ] in this study, such as deletion or vocalization, also confirm Flege's first hypothesis in the SLM (1995:238) that "learners perceptually relate positional allophones in the L2 to the closest positionally defined allophone (or "sound") in the L1". Neither Harbinese and Cantonese have the allophone [ɫ] in the coda position; thus, the vocalisation of the dark /l/ or deletion (often occurring when the preceding vowel of /l/[ɫ] is /u/ or similar) were used to establish a relationship to the closest sounds in their L1s.

Flege's SLM also has limitations in assigning difficulty to the category of 'similar'. This can be seen from the testing of hypothesis H3.3 in the present research concerning the syllable structure of the word-final stops $CVC_{/p,\ t,\ k/}$. The results indicate that Cantonese speakers significantly outperformed Harbinese speakers in producing these stops, which supports the results of similar research conducted by Flege and Wang (1989:299, 311, 312) on the word-final English /t/-/d/ contrast in Mandarin, Shanghainese and Cantonese. They found that Cantonese speakers were significantly much better than Mandarin speakers in perceiving the English /t/-/d/ contrast in the word-final position. However, the hypothesis H3.3 that Cantonese will make errors was rejected. This indeed supports the above-mentioned findings in previous research. However, it disagrees with Flege's SLM where the category of 'similar' is more difficult than the 'new' (1995). This finding shows that Flege's SLM model does not apply to all sounds similar in the L2 and L1. That is to say, similar sounds have some degree of similarity. That is perhaps the reason why so many hypotheses concerning the 'similar' category were rejected in this research. Fortunately, Flege did not apply his SLM model in the research he and Wang (1989) conducted on word-final stops, or else their results may also not have supported his model.

I would like to argue that Flege's SLM is likely to be an effective and beneficial tool in accounting for subtle L1-L2 differences, especially at the phonetic level. In addition, it may work well to account for errors highly experienced L2 learners continue to make. As Flege has mentioned (1995:238), his participants were highly experienced learners, unlike the participants in my study who were at beginner or intermediate levels and also in L2 classroom. This is probably one reason why Flege’s model does not account for errors made by learners in this study. The participants recruited in Flege’s research were mostly highly experienced L2 learners or bilingual speakers (Flege, 1995:238, 264; Piske, 2007). In the review of Flege’s SLM, Guo (2006:73) mentioned several problems that should be attended to when applying SLM. One is that participants must be suitable and must not be beginners or intermediate level learners. Flege (1995:238) also mentions that “bilinguals who have spoken their L2 for many years are not beginners.” Highly experienced L2 learners may have established new categories for L2 sounds absent in their L1, but they may still have foreign accents due to the indistinguishability of the L2 sounds closest to their L1; so that the most similar sounds are difficult to fully learn and produce in the longer term.

Whether ‘similar’ is difficult and ‘new’ is easy or ‘similar’ is easy and ‘new’ is difficult is likely to depend on stage of learning. For L2 learners at the beginner and intermediate stage, and even sometimes at advanced stages, new sounds still tend to be difficult. The results in the present study provide evidence of this. Flege (1995:265) also mentioned that even highly experienced L2 learners may produce new sounds inaccurately. For highly advanced second-language learners, ‘new’ perhaps is not as difficult any more because they may have established a category of new sounds in the L2; however, ‘similar’ at this moment is likely to be more difficult due to the lack of contrast in closely similar sounds. As posited by Flege (1995), L2 learners will successfully acquire new sounds earlier than similar sounds. In this case, the closest similar sounds are in the nuances between two languages. It is most relevant that phonological relationships between the L1 and L2 stated in Major's Ontogeny Phylogeny Model (2001) are that as exposure length increases, L2 increases, L1 influence decreases and developmental factors increase and then finally decrease when the learners' phonology becomes more target-like. At the early stage, similarity effect declines. In addition, it can be said that the focus of contrast in CAH and SLM differs. CAH mainly emphasises general differences between two languages, and SLM focuses on nuances.

Therefore the choice between the CAH and SLM in terms of 'similar' or 'new' being easier or more difficult has no resolution. Each party has shown evidence to support their views.

Through the use of the two major ideas in the present research, it can be seen that the CAH can still account for general differences between L1 and L2 although it cannot account for all of them. Moreover, it works well in L2 classroom. Whereas, the SLM does not work in this L2 classroom environment and it seems to work well in mapping out subtle phonetic differences between L1 and L2.

Finally in this discussion, I need to explain why I did not adopt the criteria of classification used in Flege's SLM. This is because, if my classification had been conducted using Flege's criteria (Flege, 1997:17) for the 'identical' category, there would be no members of this category in the research, because no sounds are exactly the same in every respect between the varieties of Chinese and English. As mentioned by researchers (Flege, 1997:21; Davenport and Hannahs, 2010:41-42; Disner, 1983; Collins and Mee, 1984), phonetically, the English [i] would not occur in exactly the same place as the Dutch [i] or German [i] in the vowel space. This also applies to the classification of syllable structures. Although different languages share the /CVC/ syllable, phonotactic constraints come into play and, in this case, languages cannot be regarded as completely 'identical' either.

Similarly, the category of 'new' would not exist in my research either. As in the definition of 'new' by Delattre (1964, 1969, cited by Flege, 1997:16) requiring that some L2 sounds differ 'radically' from any L1 sound, the 'new' in the present study was defined as L2 sounds that are absent in the L1 phonemic and phonetic inventories. For example, English interdental fricatives were classified in the category of 'new' in this research because they are absent in both Chinese dialects/languages. However, it may be questioned why they cannot be sorted into 'similar' because the voiceless /θ/ is perceived as similar to the Harbinese /s/ by Harbinese speakers and similar to the Cantonese /f/ by Cantonese listeners. This appears to make sense. However, if I also put /θ/ into the 'similar' category, the whole target would be in only one category of 'similar' and nowhere else. Moreover, the criteria for the category of 'similar', which as Chapter 4 shows are problematic seem to be even more so. Flege's model neither explores different types of 'similar', nor does it predict the degree of difficulty for different degrees or types of 'similar'. Flege claims in his SLM (1995) that the closer L2 sounds are to L1 sounds, the more difficult the L2 sounds will be to acquire by L2 learners. However, what L2 sounds can be regarded as the closest sounds to L1, and how to assess the degree of similarity between L2 and L1 remain problems.

5.3 Inconsistency of results for hypotheses in the same categories

The results in this study have been used to answer the hypotheses but, as seen in the summary of hypotheses, those in the category of 'identical' are mostly supported except for H1.3, H3.1 and H3.2. Most results in the 'identical' category can generally prove that it is easier if L2 sounds are identical to L1 ones. This supports the claim of Weinreich (1953, 2010) that identical L2 and L1 sounds lead to the positive transfer in the production of the L2. It also agrees with James' idea (1984, cited by Flege, 1997:17) that L2 speech errors are generally attributed to similar and new rather than identical sounds.

However, the hypotheses concerning the category of 'similar' being difficult are mostly rejected except for H2.3b), H2.4 and H4.1. This means L2 sounds which are similar to L1 sounds were not difficult to acquire for the participants in this research. The findings demonstrate that the most difficult sounds in L2 are still sounds new to the learners' L1. This indicates that Flege's SLM cannot be supported in these parts of this research but Lado's CAH is strongly supported. The results conform to those of other studies concerning the influence of Chinese (typically Mandarin) on English phonology where similar sounds lead to positive transfer for L2 English learners, while different or new sounds lead to negative transfer (Shi, 1996; Yang, 1997; Kong and Wang, 2001; Li, 2006; Wang, 2007; Yan, 2007; Gong and Tian, 2008; Peng, 2009; Yu, 2009; Liu and Wang, 2007; Li and Chen, 2007; Li, 2008; Li, 2009; Chen, 2010; Yin and Li, 2014; Yan and Yang, 2011; Jia, 2011; Zhang, 2012; Chen, 2013; Yang, 2014; Zhang, 2017). Moreover, such research argues that new sounds are the most difficult for L2 learners and lead to the most errors.

Why are hypotheses such as H1.3, H3.1 and H3.2 in the category of 'identical' not supported? Also, why are hypotheses such as H2.3b), H2.4 and H4.1 in the category of 'similar' not rejected like most of the other such hypotheses? These interesting questions are now discussed.

The hypotheses supported or rejected in the category of 'identical' are discussed first. Most hypotheses in this category are supported; that is, it is found that identical sound are easy to learn. However, H1.3 concerning the English /n/ and /l/[l] is rejected. Also, H3.1 and H3.2 in this category concerning syllable structure are both rejected.

The results for H1.3 in Chapter 4 show why H1.3 is rejected. H1.3 predicted that Cantonese speakers will make errors because Harbinese is identical to English in /n/ and /l/[l] but Cantonese is only similar because Cantonese speakers tend to neutralise /n/ and /l/. However, the results for the Cantonese production of these two English sounds did not show any difficulties. Moreover, the t-test did not indicate significant differences with respect to the errors in English made by Harbinese and Cantonese. Therefore, the hypothesis was rejected. In this case, we can see that although it is rejected, this does not mean that identical sounds are difficult. The idea is rejected that similar sounds in Cantonese and English lead to difficulties, as in hypothesis H1.3. This rejection further confirms that 'similar' is easy.

H3.1 and H3.2 concerning syllable structure in the category of 'identical' are both rejected. Despite this, it can still be proven that 'identical' is easy. However, an issue which arises is that the results showed that 'new' is also easy because no significant differences were found. These results are different from others in the 'identical' category. There may be some other factors leading to the rejection of the hypotheses. Therefore, these two rejected hypotheses H3.1 and H3.2 in the category of 'identical' are worth discussing further.

From the results for H3.1 in Table 5.8, we can see that there are no /Cj/ and /Cw/ onsets in the Cantonese syllable structure, but Cantonese participants pronounced these tokens with these onsets in the syllables just as well as Harbinese participants did. What are the reasons for this? The answer could be the influence of $L1_C$ on these syllable structure patterns. As mentioned in chapter 3, the glide /j/ is subject to a phonotactic constraint in English syllable structures where it is always combined with the vowels /u/ or /ʊ/ to form /ju/ or /jʊ/. It is rare to find a /j/ plus any other vowels in English. Thus, the majority of English words with the glide /j/ are combined with the vowels /u/ or /ʊ/. However, the combination /ju/ or /jʊ/ is similar to the diphthong /iu/ in Cantonese. Moreover, in Cantonese, /iu/ can be combined with all Cantonese consonants except the velar nasal /ŋ/ and labial-velar approximant /w/ as well as the labialised velar stops /k^w/ and /k^{hw}/, which are quite similar to the English /ju/ or /jʊ/. Although /j/ in /ju/ or /jʊ/ and /i/ in /iu/ work differently, they are very similar in terms of perception. This is why Cantonese participants achieved high accuracy when pronouncing tokens with the onset /Cj/ in English, because they were influenced by $L1_C$.

The influence of $L1_C$ can also be seen in the high accuracy rate for the syllable structure with the /Cw/ onset in the Cantonese group. The English /w/ is also subject to a phonotactic constraint. A limited number of consonants can be preceded by /w/ to form a consonant cluster. /k/ and /g/ are two cosonants out of the limited number which can combine with /w/ in English., The tokens selected in this research to test the syllable structure pattern with /w/ have /k/ and /g/ preceding the glide; that is, having /kw/ and /gw/ as onsets. However, Cantonese has two phonemic labialised velar stops /k^{hw}/ and /k^{w}/. They are pronounced like the English /kw/ and /gw/, although English contrasts voicing while Cantonese contrasts aspiration and the /w/ glide in English takes a segmental position while it is coarticulated with consonants in Cantonese. The positive transfer of the $L1_C$ does help Cantonese speakers pronounce the syllable patterns with the onset /Cw/ with more accuracy. This would be the reason why Cantonese participants pronounced the words with this syllable structure very well. However, whether or not English words with other consonants preceding /w/, which are not similar to Cantonese segments, would also be pronounced well requires further research.

Table 5.8 Production of English syllable structures /CjV/ and /CwV/

Patterns	Types	Harbin	GZ
CjV	Correct	96%	94%
	Sub-Cɪw	2%	4%
	Sub-Cɪu	-	2%
	njəʊ	1%	-
	bjə	1%	-
CwV	Correct	89%	90%
	Sub-gɹɪ	2%	5%
	Sub-kɹɛ	-	3%
	Ins-gəwɪ	2%	-
	Ins-guwɪ	1%	-
	S+I-gəɹɪ	4%	-
	S+I-guɹɪ	1%	1%
	*Error	1%	-

Another rejected hypothesis in the category of 'identical' was H3.2 concerning /CVm/. Cantonese allows /m/ as a syllable coda whereas Harbinese does not. Therefore, it was predicted that the Harbinese would make errors. The actual results support this hypothesis but the statistical test shows no significant difference ($p=0.185$), and thus, it was rejected. If there is no such syllable structure /CVm/ in Harbinese, why can its speakers acquire this syllable structure ending with /m/ with a high accuracy rate? It is not influenced by $L1_H$ because the L1 does not allow /m/ as a coda. There must be some other phonological mechanism which helps Harbinese speakers acquire it.

Researchers (Shi, 1996; Dou, 2010) have found that Mandarin/Harbinese speakers tend to epenthesise a vowel after the final coda to form a syllable due to the preference for open syllables in Mandarin. Perhaps Harbinese participants regarded /m/ as the onset of a new syllable /mu/, which can also be seen in the errors in Table 5.9. It may be used as a strategy to complete this syllable structure.

It may also be possible because /m/ is more universal and unmarked compared to other nasals. Stemberger and Stoel-Gammon (1991) claimed that the first sounds acquired by children were produced in the labial site of articulation, which are the oral and nasal stops (/b, p/ and /m/). Hume (2003:10) found from Smith's data of 2,500 words in Sri Lankan Portuguese Creole (1978) that [m] is the least marked place of articulation because it is the most frequently used. She also found (2003:6) that the labial nasal occurs almost as frequently as the coronal nasal, but the dorsal nasal fell significantly behind. Harbinese has the dorsal nasal as the most marked as a coda in the syllable structure; thus, it would not be difficult to acquire a syllable structure with a final /m/. From this evidence, this seems a reasonable explanation of why Harbinese speakers had a high accuracy rate in producing English words with the syllable structure having /m/ as coda.

Table 5.9 Production of English syllable structure /CVm/

Patterns	Types	Harbin	GZ
CVm	Correct	88%	95%
	Ins-mu	9%	-
	Del	2%	5%
	Sub-n	1%	-

In the category of 'similar', all the hypotheses were rejected except for hypotheses H2.3b) regarding /ɛ/, H2.4 concerning /ə/ and H4 about stress assignment. Most results in this category indicate that 'similar' is not difficult, in disagreement with Flege's model. However, why do these three findings for H2.3b), H2.4 and H4 support Flege's SLM? Does this support the idea that similar sounds are difficult to learn?

H2.3 is a special hypothesis including two sub-hypotheses a) and b) because the relation between each dialect/language and English is 'similar' vs. 'similar'. Thus, it was predicted to be difficult for both groups as H2.3a). Based on the predictions on the degree of difficulty in Chapter 3, the similar type with 'allophonically identical but phonemically different' was predicted more difficult than other types of similarity, thus, H2.3b) was predicted that Harbinese (with [ɛ] /a/) make more errors than Cantonese (with /ɛ/ [ɛ, e]). Although H2.3b) was supported by the results, it confirms that different types of similarity have different difficulty in the degree and confirms the prediction that 'allophonically identical but phonemically different' similar type is more difficult than 'phonemically identical'.

The results for the production of the English /ə/ support hypothesis H2.4. This seems to agree with Flege's idea that 'similar' is difficult and 'new' is easy. However, this would contradict the rejection of other hypotheses. The errors in the production of the English /ə/ are shown in Table 5.10. It can be seen /ə/ was over-rhoticised by some participants in both groups but Harbinese speakers tended to do it more. It has been found (Liu, 2013:126; Cui, 2014:190) that more Chinese words are rhoticised in Northeastern dialects, including Harbinese than in Standard Mandarin. Therefore, we can see that the error of the rhoticised [ɚ] is obviously influenced by the L1 Harbinese dialect and that the Cantonese group would be influenced by Mandarin. Harbinese/Mandarin has r-suffixed rhyme phenomena (rhoticisation) and also has /ɚ/ as a single syllable for some Chinese words (Chao, 1968:47-50, cited by Li and Thompson, 1989:9; Lin, 2007:80-81). This error was the main error among Harbinese speakers. It seems that the error /ɚ/ always happens when the schwa forms an open syllable. The epenthesis of the dark /l/ or /w/ sound which is developing from the vocalisation of /l/ was another error only seen in the Harbinese group, often occurring when /ə/ was word-final. This might be caused by confusion in Harbinese speakers about the pronunciation of the token *camer**a*** [kæməɹ**ə**] with that of the English word cam***el*** [kæm**əl**] as [kæməɹ**əl**] or [kæməɹ**əw**]. Because of the limited number of tokens with /ə/ in the final position in this research, it cannot be certain if the position involved plays a role in influencing production. It may be influenced by the phonotactic constraint in the $L1_H$ dialect. Transfer from L1 seems to be the main reason to explain why H2.4 is supported while others are not. The /ə/ is new in Cantonese, but the results show that Cantonese speakers did not have difficulty in pronouncing it, which implies that 'new' is easy. A possible explanation could be that Cantonese speakers have already acquired the phoneme /ə/ from Mandarin because they had studied Mandarin from a young age. In addition, due to the fact that Cantonese do not have rhoticisation, which helps them avoid the overgeneralization in rhoticising /ə/, the occurrence of [ɚ] is avoided as an error. Therefore, this may be the reason why Flege's SLM seems to be supported.

Table 5.10 Production of English vowel /ə/

Vowels	Types	Harbin	GZ
ə	Correct	84%	94%
	ɚ	7%	2%
	əw	2%	-
	əl	3%	-
	ɜ	-	2%
	u	-	2%
	a	4%	-

Hypothesis H4 on stress is different from H2.4. Although the hypothesis in H4 that Harbinese will make errors with English stress placement is supported, this reason is not because of the similarity of stress patterns in Harbinese and English leading to difficulties for Harbinese speakers, as claimed by Flege's SLM where 'similar' is difficult. Instead the influence of $L1_H$ causes it. As mentioned in the literature review, Harbinese has a main stress pattern with an iambic foot (70%) but it also has a stress pattern with a trochaic foot (30%) which is a similarity with the English stress pattern of a main trochaic foot. Although there is a similarity in stress between Harbinese and English, the fact is that there are more differences than similarities. The main stress pattern with an iambic foot in Harbinese led to the Harbinese speakers tending to place the stress on the second syllable of English disyllabic words, and this caused the illusion that similarity increased the difficulty as predicted by H4.1, so that H4.1 was supported. This implies that the more similar the two languages are, the easier it may be for L2 learners to learn; and the more different they are, the more difficult learning is. This supports Lado's CAH (1957) again.

As far as the above discussions are concerned, the reasons why the hypotheses H1.3, H3.1 and H3.2 in the category of 'identical' and H2.4 and H4 in the category of 'similar' lead to distinct findings compared with the results of other hypotheses is interference from the L1.

5.4 Different degrees of difficulty within the category of 'similar'

Target sounds for the category 'similar' represent a big group in this research and it is complex. As indicated in the criteria for classification in Chapter three, the category of 'similar' is further classified into three types of similarity of sounds. Moreover, the three different types of similarity were also predicted to relate to different degrees of difficulty. However, it is unclear if the type of similarity which is allophonically identical but phonemically different should be more difficult than the type which is phonemically identical but allophonically different, or the type which is phonemically similar. This is an interesting question discussed in detail following, because there is no published account of the degree of difficulty associated with different types of similarity.

All of the hypotheses were formulated based on the three types of similarity to determine the actual levels of difficulty experienced. It was found that only two hypotheses, H1.6 and H2.3, included two types of similarity and all of the others considered only one type. That means that only H1.6 and H2.3 can be used to compare different types of similarity to see which is more difficult, while all of the other hypotheses compare one type of similarity with 'new' or 'identical'.

Each type of 'similar' is discussed in this section. The production results show that the speakers of a dialect/language with sounds which are allophonically identical but phonemically different in English all achieve better performance than the dialect/language with sounds which are 'new' in English or achieve worse production results than the other dialect/language with completely identical sounds in English. The results also support the idea that phonemic similarity is easier to learn than 'new' but more difficult than completely identical. In addition, the results for another type of similarity, which is phonemically identical but allophonically different also show that 'similar' is more difficult than 'identical' as in hypothesis H1.2. However, H2.4 is the only example which contradicts these findings. Otherwise, the results for all three types of similarity suggest that the category 'similar' is easier than the category of 'new' but not easier than the category of 'completely identical', as the CAH proposes (Lado, 1957).

The testing of H1.6 compares two types of 'similar', which are allophonically identical and phonemically similar. The production results show that speakers of the dialect/language with allophones identical to English sounds outperform speakers of the other dialect/language with only phonemic similarity to English sounds. This implies that the type of similarity which is allophonically identical is easier to learn than the type which is phonemically similar. However, the results of t-test show statistical significance for only one target but not for the other two. Therefore, whether 'allophonically identical' is really easier than 'phonemically similar' is uncertain. Further work is needed.

Fortunately, the comparison between the similar types of 'allophonically identical' and 'phonemically identical' can be seen in H2.3. Production results show that the group with phonemes identical to English surpass the other group with allophones identical. It reflects that the similar type of 'phonemically identical but allphonically different' is much easier than the similar type of 'allophonically identically but phonemically different'. From H1.6 and H2.3, the different difficulties assigned to different types of similarities are clear to see. It is the similar type of 'phonemically identical' is easier than the type of 'allophonically identical', and which is easier than the type of 'phonemically similar'.

In looking at the summary in section 4.9, some hypotheses have the relation with 'phonemically similar' vs. 'new' such as H2.6 and 2.7, which can be used to check the difficulty between the similar type of 'phonemically similar' and 'new'. It is shown from the results of H2.6 and H2.7 that the similar type of 'phonemically similar' is easier than 'new'. Therefore, the difficulty assigned to different categories 'identical', 'similar' including different types of similarity, and 'new' can be generally concluded as below:

'Identical'

is easier than

'Similar'

including

Similar 'phonemically identical but allophically different' (H2.3)

is easier than

Similar ‘allophonically identical but phonemically different’ (H1.6)

is easier than

Similar ‘phonemically similar’ (H2.6)

is easier than

‘New’.

However, it is very difficult to assess whether ‘allophonically identical’ is more difficult than ‘phonemically similar’ or vice versa when comparing the production results between ‘allophonically identical’ & ‘new’ and ‘phonemically similar’ & ‘new’, although H1.6 gave the connection. For example, the results of H2.5 with ‘allophonically identical’ indicate that ‘new’ is difficult but ‘similar’ (allophonically identical) is also difficult whereas, the results of H2.6 show that ‘new’ is difficult but ‘phonemically similar’ is much easier than ‘new’. However, H1.6 gave the results to show ‘similar’ (allophonically identical) is easier than ‘similar’ (phonemically similar). Therefore, the conclusion for the rankings of difficulty above needs to be further researched in the future.

5.5 Effect of length of exposure on production accuracy

Hypothesis 5 was proposed to see if production accuracy increases with the increase of exposure length. The results cannot support this hypothesis. A fluctuating trend is observed starting with low accuracy to high and then to low with the highest accuracy achieved at high school level. According to general thinking on language development, it may be thought that the more exposure to a L2 and the longer time spent learning the language, the better results learners should achieve. However, this seems not to be the case. The results in this study show that university leaners scored lower accuracy rates than high school students. The accuracy rates at the university in Guangzhou are even lower than at Guangzhou middle school. Although there is no statistically significant difference between university and middle school, the results suggest that middle schoolers pronouncing the targets outperformed university participants in the Cantonese group. This finding disagrees with those of a previous study conducted by Mayo (2003) on the influence of length of exposure to the L2 in classroom. He conducted the research on the influence of length of exposure on L2 English in grammaticality judgement tasks and found that there was a significant increase in correct judgements over time, concluding that "the longer the exposure to the L2, the better performance becomes" (Mayo, 2003:106). Perhaps, the present finding may be caused by the different research focus and the targets. Moreover, the use of different methods may also have led to the findings. This research focused on the phonology and tested its production but Mayo focused on grammar and tested L2 English learners' judgements. My findings also do not agree with those of Ouyang in her 2018 study of the acquisition of dental fricatives by Chinese EFL learners where it was found that participants in the higher level group performed better than those at lower levels in both perception and production.

On the contrary, it seems that my findings support the research on immigrants, that length of exposure does not have a significant effect on the degree of L2 foreign accent. Although exposure length (length of residence/LoR) is found insignificant in L2 immigrant settings, researchers (Flege, Munro and MacKay, 1995; Riney and Flege, 1998; Meador, Flege and MacKay, 2000) concluded that it depends on whether learners are in an early phase of L2 learning. That is, they found from their studies that LoR might show a significant effect if learners are just starting, so in a sense exposure does have an effect. But it then seems to cease to make a difference since no effect is found for highly experienced learners (those with a long LoR). My participants are beginners and low-intermediate level learners who are learning the L2 as a foreign language in the classroom not highly experienced learners. There should therefore be an effect of exposure. However, my findings did not show an effect. There are likely factors responsible for these results. Below is a discussion to attempt to speculate on what underlies these findings.

In explaining the apparently counterintuitive findings in the present study, it is not difficult to understand that accuracy at high school is better than at middle school, because this would conform to the rules of cognitive development. However, the key difficulty in explaining the results is that the developmental trajectory of accuracy rates goes downhill from high school to university. Although the results do not show a statistical significance, this surface results can also suggest some problems between them in the real situation. Thus, it is worth figuring out. This may be due to various factors.

Firstly, as the most important stage of education, the period of high school for students' future is as important as the critical period is to L2 acquisition. This is because, at the end of the third year of high school, students have to take entrance exams for universities and colleges and this determines their potential career paths. The ability to enrol in a top university depends entirely on scores in national exams. In order to achieve high marks and to study at a top university, students devote an incredible amount of time and energy to their high school education. This devotion also involves support from teachers and schools as well as parents. It can be said that this stage of learning is the most intense, the most important, and the toughest stage of a student's life, but it is also the most fruitful stage enabling the accumulation of sufficient knowledge and to lay a solid foundation for further study at university, especially in the third year of high school. This is why I had to change my original plan to collect data from third-year high-school students to second-year ones, so as to avoid distracting the third-year students

Secondly, the time spent in school, and on English lessons is also considerable. It is not uncommon for students to spend 10 to 12 hours a day from early morning until late evening for five or six days per week studying at school. English, along with Maths and Chinese, is one of the key subjects for the entrance exams for university. The workload allocated to English learning in Harbin and Guangzhou high schools is the same, including five intensive English sessions plus at least one tutorial session per week, each taking 40 minutes according to information collected from the English teachers in the schools sampled. They may have one more or two tutorials per week, depending on the year the students are currently in. During high school, the difficulty of the English taught increases compared with the English they learned in middle school. Training in listening and speaking skills also increases in importance and amount.

Thirdly, with longer exposure to English as well as intensive study in English, it is understandable that performance in pronouncing English is better than among middle schoolers. In the process of the accumulation and development of knowledge, the frequency of repeated corrections by teachers of mispronunciations of English sounds in high school may play a key role in helping them to achieve better results.

Finally, during this intense period of high school study, students strive to fulfil their university dreams with clear objectives and strong motivation. All of these factors may eventually result in the best results occurring at high-school level.

Performance at university may be lower than at high school because, firstly, attitudes towards and motivations for learning English at university are different compared to high school and middle school. Compared with life at high school, university students enjoy a more relaxed and comfortable life. They have less pressure than before. High school English is focused on intensive teaching but university English is based more on student self-learning (Shen, Ding and Liu, 2015:116). English class sizes grow from 40-50 students at high school to around 100 and even up to 200 at university. The ratio of teachers to students falls. English lectures at university generally adopt teacher-centred teaching and there is little interactive communication in large classes. All of these changes in English learning and teaching at university level may lead to changes in students' learning attitudes and motivation, as pointed out by Shen, Ding and Liu (2015:117).

The English syllabus at university requires that teachers help students to lay a solid foundation of language knowledge, strengthen the ability to listen and speak, and improve the ability to use English to communicate (Shen et al., 2015). Although university students are required to pass the national college English tests band 4 or 6 to demonstrate their proficiency in English, they need not pay particular attention to pronunciation because CET-4 and CET-6 test only their proficiency in listening, reading, writing and translation. The speaking test is separate from CET-4 and CET-6. Also, its criteria are set very high, and only those who can achieve accuracy rates above 80% are qualified to attend the speaking test. Therefore, the opportunity for the improvement of English pronunciation and speaking skills is not given the same priority at university as at high school.

In addition, it is possible that English education in China has lagged behind, by current standards, and is affected by when these university participants first started to learn English. That is to say, the English education in China has improved, recently especially in pronunciation, in high school English education. This would mean that the improvements are already being seen in high school participants but the university participants were too old to have benefitted from this. The results in this research may confirm this. For example, in the literature it is mentioned that Cantonese speakers always unrelease the codas /p, t, k/, however, the results for the syllable structure /CVp, t, k/ in Table 5.11 below show a rise in pronunciation accuracy from university to middle school. This type of error is no longer typical for high schoolers any more, and especially for Cantonese middle schoolers, due to improving educational standards.

Table 5.11 Accuracy rates of /CVp, t, k/ at different levels

Syllable structure	University	HighSchool	MidSchool
CVp	82%	97%	97%
CVt	94%	97%	100%
CVk	97%	97%	97%

Another possible reason is that there are differences in quality between high school students and university students. The high school participants recruited for the present study were at key city schools who might then go to better universities to study compared with their counterparts at university sampled in the present study. This might partly account for the university participants having lower accuracy rates than the high schoolers.

So improvement in English education in China might be responsible for the better performance of high schoolers from both cities in pronunciation than their counterparts at university, and a similar explanation may apply to the better results of Cantonese middle schoolers than Cantonese university learners in this study, even though difference is not statistically significant. an improvement might be implied. The following section looks at answers to the question of why Cantonese middle schoolers outperformed Cantonese university students while Harbinese middle schoolers did not do better than Harbinese university learners.

The differences in performance between the two middle schools might result from the differences in education policy and disparities in the quality of education in the two cities. From a comparison of documents published by the official bureaux in the two cities, it seems that the differences in policy for high school entrance exams in English between the cities may have played a role in causing the disparity in performance. Besides the test on grammar, reading, and writing, the English exam for high school entrance in Guangzhou has a speaking and listening test, the scores of which are included in the overall score for high schools entrance , as stated in the English Reform Programme issued by the Guangzhou Admissions and Examination Committee Office in 2015. However, since 2014 Harbin has removed the listening test from the English exam for high school entrance. Instead, it is combined with a speaking test as a communication skills proficiency test (Harbin Education Bureau, 2014). Unlike Guangzhou's requirement, the listening and speaking test in Harbin neither scored nor counted as part of the overall score for entrance to high school. To achieve high scores and to enrol in the best high schools, teachers and students will pay more attention to tests which they can score in rather than tests without scores. This is highly likely to lead to different results between the two middle schools, with the consequence that Guangzhou middle schoolers have comparatively higher accuracy rates than Guangzhou university participants while the situation in Harbin middle school is the opposite. The different policies may also lead to imbalances in the quality of education between the two middle schools. My experience when collecting data also shows a disparity between the middle schools which I discovered when communicating with the English teachers. For the listening and speaking test, middle school English teachers not only attach great importance to the pronunciation of students, but also to the sources they can utilise to help them. I was informed by an English teacher in the Guangzhou middle school that their English learners use an app focusing on English listening and speaking called Ets100, which is a test preparation product for English at the elementary education stage which is equipped with voice and accuracy evaluation technology (Ets online). This app is accompanied by tutorial books for teachers and students to help improve the teaching of pronunciation and to help English students improve their pronunciation. However, the English teachers at Harbin middle school had never heard of these resources. Therefore, they did not use them to guide their students who would not have the chance to improve their pronunciation by using Ets100.

From the results of the present study, we can see that Cantonese English learners outperformed Harbinese English learners from middle school level to university level. Also, the statistical tests indicate significant differences between them except at high school level, where the difference is marginally non-significant ($p=0.074$). Possible explanations of why the Cantonese groups perform better than the Harbinese at all levels are given next.

Firstly, there might be regional disparities in economic and educational development. Guangzhou is the central city of the highly developed Pearl River Delta Economic

Zone, and its provincial annual GDP in 2018 was CNY 972.778 billion (National Bureau of Statistics of China, NBSC, 2019), which is equivalent to USD $138.968 billions[13]. Harbin is the provincial capital of Heilongjiang and it is an important regional central city of the Northeast; however, its economy is relatively backward. Heilongjiang's annual GDP for 2018 was CNY 1636.162 billion (NBSC, 2019), approximately equal to USD $233.737 billion. The economic gap in GDP between the two provinces is more than six-fold. The level of economic development in a region will affect the local level of educational development (Yue, 2003:38; Huang and Yang, 2012:155). Huang and Yang (2012:155) claimed that the differences in levels of educational input caused by imbalances in economic development have become the main factor in the formation of regional inequalities in education. As discussed above in terms of English teaching requirements, policy-making differences for English exams between the two cities are good examples of regional differences in education. The educational policy in Guangzhou has strengthened training in listening and speaking skills in attempting to match international standards, while that in Harbin is weaker in the training and development of practical English communication skills.

The fact that students from Guangzhou perform better than those from Harbin across all educational levels might also be due to different amounts of exposure to English outside of school. According to responses to Question 24 in the questionnaire on ways of learning English outside school, and the time they spent doing this, the results indicate that the Cantonese group spent an average of 3.6 hours while the Harbinese group devoted 3 hours per week to extracurricular English learning.

It is possible that, from a linguistics perspective, the results might be caused by the effect of L1 sound inventories. Wang (2009:117) mentioned in her research that negative transfer of Cantonese vowels to L2 English does not seem obvious, except in contrasts in vowel length. Cantonese has more phonemically identical and similar

---

[13] Assuming an exchange rate of 1 USD is equal to 7 CNY.

segments to the inventory of English than does Harbinese. It is likely to be easier for Cantonese learners of English to acquire English phonology. The results for identical and similar segments confirm this.

5.6 English variety

As mentioned in the literature, Standard English as a dominant language is widely used in L2 teaching and learning across the world. It has two main varieties: British English, or RP, and American English, or GA. Attention has been paid by researchers to attitudes towards Standard English varieties internationally (Starks and Paltridge, 1996; McKenzie, 2004). Zhang and Hu (2008) and Xu, Wang, and Case, (2010) have conducted Chinese research on this theme. However, little study has examined which variety of English is used by Chinese learners of English. I assume that a substantially large number of people think that Chinese English learners speak a Chinese style of English which could be one of the varieties of global English like Singaporean English and Indian English. It cannot be denied that Chinese speakers of English speak English with Chinese accents because English is a foreign language which is not frequently used in Chinese people's daily lives. However, what varieties of English do their pronunciations approximate to? Do different Chinese dialects/languages influence their speakers' use of English varieties? The present investigation of pronunciation by different dialect/language speakers of certain segments can indirectly help to answer these questions.

As Griner (2014) mentioned, Chinese learners of English do mix pronunciations of British English and Amerian English. The results in Table 5.12 confirm this. The production of participants of each dialect/language includes elements of both British and American varieties of English. This means that they mix the English varieties in their English production. This can result in Chinese English or Chinglish. However, apart from these two standard varieties of English (RP and GA), other pronunciations are considered as errors in this research because mispronunciations cannot be accepted as a legitimate variety of local non-standard English. Moreover, English is a foreign language, therefore, the correct pronunciation in L2 teaching and learning is very important.

This research contains four English segments involving two English varieties. The consonant /ɹ/ is a typical representative used to distinguish American English from British English when occurring post-vocalically. From the results, we can see that 81% of the Harbinese speakers used American English in the pronunciation of rhotics, while fewer Cantonese speakers did so. Meanwhile, a large proportion of Cantonese speakers pronounced the /ɹ/ in the British way without rhotics. Moreover, this difference was statistically significant according to the results of the t-test reported in chapter 4. However, the conclusion that Harbinese speakers speak GA and Cantonese speakers mainly use RP cannot be proven because the results for /ɜ/ show that a very large percentage of Cantonese speakers pronounced /ɜ/ as /ɝ/ with rhotics. Moreover, the results for the English diphthong /əʊ/ do not show a large percentage of Cantonese speakers speaking British English. However, it can be concluded that Cantonese speakers are more likely to speak the British variety than Harbinese speakers. From the results in Table 5.12, we can assume that Harbinese speakers are generally more apt to speaking American English, although the English /ɒ/ is an exception.

From the comparison of English and the two Chinese dialects/languages, we can assume that Harbinese speakers are likely to speak the American variety of English, as is the case with Cantonese speakers, because similar or identical sounds exist in their respective dialects/languages. For example, Harbinese has the r-suffixation phenomenon, which results in rhotics being pronounced by Harbinese speakers, and it has also the diphthong /oʊ/. However, it is not clear why more Harbin speakers pronounce /ɒ/ in the British way. It might sound slightly similar to the Harbinese /ɑʊ/, or possibly they were taught British pronunciation in schools at an early stage, as mentioned by Griner (2014), and this pronunciation has become rooted in their minds as prior knowledge of the English sound system. This could also explain the case of speakers of Cantonese which do not have r-suffixation, and therefore a large group of the Cantonese participants used British English, in the same way they did for the English /ɒ/. The English /ɒ/ is similar to the Cantonese /ɔ/. This is why Cantonese speakers may be more likely to speak British English than Harbinese speakers. However, /ɜ/ is an exception for Cantonese speakers and the reason for its use is not clear, it might be influenced by Mandarin where /ɚ/ is a single phoneme as a syllable in some Chinese characters such as *er* 'son', and /ɚ/ is similar to /ɝ/.

Table 5.12 English segments with two varieties pronounced by Harbin and Guangzhou participants

Segments	Varieties	Harbin	Guangzhou
/ɹ/	A-ɹ	81%	38%
	B-/	19%	62%
/ɜ/	A-ɝ	98%	80%
	B-ɜ	2%	18%
/ɒ/	A-ɑ	3%	1%
	B-ɒ	37%	84%
/əʊ/	A-oʊ	55%	59%
	B-əʊ	25%	37%

(A: American English; B: British English)

From the above discussion, we can see that both groups mix English varieties in their English production but Harbinese speakers are more likely to speak American English while Cantonese speakers are more apt to speaking the British variety. It is reasonable to doubt whether they actually know how to distinguish between the two English varieties. If they cannot clearly distinguish the differences, it would mean there may be another reason to explain their tendency to speak one of the specific varieties of English. I conclude that the variety of L2 English is produced that has more sounds similar to the dialect of their L1.

The results from the analysis of questionnaire data confirm this point. Questions 24 and 25 in the questionnaires were mainly designed to discover attitudes towards English varieties, and 85% of Harbinese participants preferred to watch American TV series while the percentage for the Cantonese group was 68%. The preference for

American films and TV series could mean they may be more exposed to American English. That may be another reason why /ɜ/ is usually pronounced in the American way by the Cantonese group. When participants were asked which variety of English they preferred to speak, the percentage of preference for the British variety increased. The ratio for American and British English in the Harbinese group was 50% vs. 50%, but 68% of the Cantonese group preferred British English. This result seems to accord with the conclusion that the Cantonese group achieved higher accuracy rates in pronouncing British English than the Harbinese group.

When asked what variety of English they thought they spoke, most respondents said that they thought they spoke Chinese English. Q27 was asked to determine if they could distinguish American English from British English, 79% of Harbinese participants and 68% of Cantonese participants answered that they could not. For the participants who answered they could, a further question was asked about the differences between British and American English, but few could indeed demonstrate that they knew the difference. This indicates that the participants from these two dialects/languages cannot speak in a particular variety of English as they expected. So, why does their production represent a trend towards the use of a variety of English? I still think the main reason is that their dialect/language has more similar sounds with one variety of L2 English, and so their production is closer to that variety of English.

As mentioned above, Harbinese has the r-suffixation phenomenon; therefore, this is possibly the reason why a majority of Harbinese speakers pronounce /ɹ/ and /ɝ/. However, due to the lack of this r-suffixation feature in Cantonese, it may be wondered why a large group of Cantonese speakers pronounced /ɝ/. There may be many factors causing this. As explained above, it may be influenced by Mandarin. Also, there might be the influence of English teachers from the USA. After all, the results of the questionnaire survey indicate that participants had more American English teachers. However, this factor perhaps does not play a vital role because the answers from the participants do not imply that they were greatly influenced by their native English teachers. Moreover, the number who had learned English with native English teachers is only a small proportion of the whole group. Therefore, another factor which possibly plays a significant role is the greater exposure to American English in films, TV series and music. It can be seen from the results of questionnaires that they have more exposure to these media forms.

Finally, it is concluded here that the learners could not describe the differences between American and British English; therefore, their actual English production mixes them together. However, from the perspective of accuracy, Harbinese speakers are more prone to speak American English and Cantonese speakers are apt to speaking British English. As can be seen from the above discussion, this is most likely explained according to which dialect/language has more sounds of one variety of English, whose speakers' pronunciation then tends towards that variety of English.

5.7 The pedagogical implications of the findings

This research has contrasted two Chinese dialects/languages and standard Mandarin and compared them with English phonology and has discussed the errors and error types found. The present research has pedagogical significance in providing a good guide for English pronunciation teaching especially in Harbin and regions speaking Mandarin similarly to Harbinese and Guangzhou and areas where dialects similar to Guangzhou Cantonese are spoken. The findings in this research demonstrate the general difficulties experienced in English pronunciation in terms of consonant, vowel segments and syllable structure, and reveal the general errors and error types English learners frequently made under the influence of their own dialect/language. It can help teachers better understand the impact of these dialects on English learning as well as the difficulties students have in English pronunciation. It also provides assistance and guidance for English teachers in solving students' pronunciation problems with the targeted segments and positions and in carrying out targeted training. It not only provides pedagogical guidance to English teachers but also to English learners. It can assist English learners to be aware of which segments and positions they may have difficulties with and make errors, where they may not otherwise realise this, so that they can clearly understand and be able to reduce the influence of dialects and improve their English pronunciation. In addition, the thesis can provide guidance for teaching writers of materials for English pronunciation, and supporting materials such as the Ets100 app.

From this research, we can see that the influence of L1 on the L2 acquisition of English phonology is very significant. Therefore, it needs to be noted that research on English teaching and learning in China cannot be conducted only using a frame of Standard Mandarin or of the general term Chinese, but it is necessary to pay more attention to the influence of specific Chinese dialects/languages.

This research has discussed the use of English varieties by speakers of different Chinese dialects/languages. The findings reveal that English learners cannot distinguish between British and American English. However, much research on Chinese English learners' attitudes towards English varieties has been conducted (Zhang and Hu, 2008; Xu, Wang and Case, 2010). However, it may be pointless to know what their attitudes are if learners cannot distinguish between the English varieties. They definitely cannot acquire the sound system of such a variety, let alone speak only that variety of English. Therefore, educating learners on the differences between GA and RP should focus primarily on phonology.

This research has some significance in terms of guidance for education administrators. The findings in this research show the consequences of different education policies and the disparity in education levels that may exist. Therefore, this research indicates how important the policies adopted by educational authorities are. I also recommend that all educational authorities should carry out a series of detailed investigations with experts in this field before they design and implement important new policy decisions. The findings also reveal that there may be regional differences in educational levels. Educational authorities should be aware of this, especially those in less developed economic areas, and take action to maximise resource assignment to education in relevant local areas and to provide opportunities for teachers to develop more advanced teaching methods and a better mastery of specialised knowledge.

Therefore, this could also help high proficiency L2 learners to identify their problems, and to shorten the perceived phonetic distance between subtle distinctions in L2 and L1 sounds, thus further improving their L2 pronunciation towards an L1 native-like level.

I believe contrastive analysis is more appropriate to help L2 learners and teachers to understand the general difficulties learners may have in L2 learning by contrasting phonemes in the L1 and L2. Moreover, the application of contrastive analysis could be more suitable for the study of L2 learners at low proficiency levels. However, for high proficiency L2 learners, SLM would be a good theoretical model to predict the difficulties which might arise due to subtle differences between L2 sounds closest to L1 sounds. SLM would be very helpful to further improve the accuracy of pronunciation for L2 learners. Therefore, this model is more suitable for research on L2 learners with high levels of proficiency and experience.

Chapter 6. Conclusions

6.1 Contributions of the research

This thesis has examined the influence of two different Chinese dialects/languages on the acquisition of L2 English phonology and explored how speakers of these different dialects/languages acquire English phonology. The study started with an introduction to the linguistic background, and primarily the phonetics and phonology, of two Chinese dialects/languages, Harbinese and Guangzhou Cantonese. By determining the differences between the two Chinese dialects/languages and the differences between each dialect/language and English, the chapter describes how the research sounds were chosen. Afterwards, the thesis briefly reviewed the history of research of the second language acquisition of phonology and highlighted two influential theoretical frameworks, the Contrastive Analysis Hypothesis and Speech Learning Model. This was followed by a review of previous research on the L1's influence on L2 acquisition and two Chinese varieties' influence on English.

The research instrument underpinning this thesis includes a word translation test with 174 tokens for participants from a middle school, a high school and a university in each city of Harbin and Guangzhou. After the collection of data from 90 participants, that from 65 subjects was considered valid and used in this thesis, and analysed by auditory and acoustic methods, to answer questions mentioned in the introduction chapter and to discover the nature of the influence of each dialect/language on the acquisition of English in terms of segments, syllable structure, and stress.

The results confirm the claims of previous researchers, including Anderson (1987); Corder (1967); Ellis (1994); Fisiak (1978, 1991); Gass, (1979); Lado (1957); Odlin (1989); Rasier and Hiligsmann (2007); Weinreich (1953); Young-Scholten (1985), that the influence of the L1 on L2 is significant when it comes to phonology. It was discovered that the influence of the Chinese dialects/languages on the production of English phonology was profound, but the influence of each dialect/language on the phonological acquisition of English differed. It was also found from the results that there appears to be an influence of Mandarin on the Cantonese learners of English, although this was not significant. The thesis tested the predictions of Lado's CAH (1957) and Flege's SLM (1995) according to the results of English production from Harbinese and Guangzhou Cantonese speakers. It was found that the results support Lado's CAH in that sounds similar in both the L1 and L2 are easier to acquire than new sounds. Flege's SLM, on the other hand, is shown in this thesis not to work perfectly. The hypotheses proposed based on the difficulty predicted due to L1-L2 similarity from Flege's SLM were mostly rejected and only a few were supported. From the results, we can see that Lado's Contrastive Analysis is still effective in predicting the difficulties L2 learners may experience, although it cannot predict all of the difficulties. Moreover, not all difficulties can be accounted for by L1 transfer. These findings provide an opportunity for L2 learners and teachers to find the main problems experienced and to try to correct the errors made when they are acquiring a second language. Although Flege's SLM, which states that 'similar' sounds are difficult, does not seem to apply in this research, it could still be supported. It is proposed here that, if and only if similar sounds between the L2 and L1 are nuanced, they would be difficult for L2 learners to contrast and distinguish between, because the results indicate that L1-L2 sounds which are phonemically similar, phonemically identical but allophonically different, or allophonically identical but phonemically different were easy to learn. However, further research is needed to confirm this. Moreover, based on the results of the study, I proposed that the difficulty for 'similar' and 'new' can be mutually transformed, but it depends on the learner's proficiency level in the L2. If L2 learners are at an advanced level, or are as experienced as were the participants in Flege's (1995:238) research, the 'similar' could be difficult for them to contrast with nuanced sounds, but it should be easy for L2 learners at other levels. In this study the difficulties with different types of 'similar' have been predicted based on Flege's SLM, but, unfortunately, the results cannot clearly represent the relationships of difficulty between types of 'similar' because there are no direct comparisons between every two types of 'similar'. However, a tendency can be seen that similarity in the sense of 'allophonically identical' is easier than when being 'phonemically similar'. Further research is again required.

Through analysing the acquisition of English phonology by participants with different proficiency levels, the influence of the length of exposure to English could be determined, and it was found that error rates may not decrease over time. The accuracy rates of the high school participants in each city were higher than those studying at university in each city. Even the results for the Guangzhou middle school students were better than those at university in Guangzhou. One possible explanation for this is that English education in China is improving. In addition, it was discovered that the participants from Guangzhou generally had better English pronunciation than those from Harbin at the high school level although this was not statistically significant. This may be attributed to the fact that Cantonese has more similarities to English than does Harbin in segments, and especially in vowels, or to possible differences in the standard of education between the two cities, as suggested overall above.

This thesis also looked at the use of English varieties by the speakers in the sample and discovered that the Cantonese subjects tended to speak British English while the Harbinese tended to speak American English. However, after analysing the questionnaire data, it seems that this conclusion is premature because they reported that they did not clearly know the differences between British and American English. Thus, it can only be stated that they have a tendency to speak an English variety which is similar to their own variety of Chinese.

This study has found that different errors were made by Harbinese and Cantonese speakers. The most interesting finding here concerns the differential substitutions of /θ/ and /ð/. In the present study it was found that /θ/ and /ð/ were most frequently substituted with [s] and [ts]/[d] by Harbinese speakers, with [s] and [ts] occurring across all positions but [d] only initially. For Guangzhou Cantonese speakers, [f] was found to be most frequently substituted for /θ/ across all positions, and [d] was most frequently substituted for /ð/ when /ð/ occurred word-initially and word-medially. However, [f] or [θ] were most frequently substituted for the word-final /ð/.

The thesis makes contributions to the field of second language acquisition by describing the sources of errors in the domain of phonology. The findings also have pedagogical implications for L2 learners and teachers who speak varieties which are not standard.

6.2 Limitations

This study, like all research, has some limitations which need to be noted by future researchers. The testing instrument was designed with care and the tokens were carefully selected. However, there are still some flaws with the tokens used. Moreover, because there was only a limited amount of time to analyse a large amount of data, the analysis of consonant clusters had to be omitted from the research. Another limitation concerns participant recruitment. The original plan was to collect data from students at different school levels; that is, from middle school, high school and university. Each consecutive level was meant to have a gap of three years. This was not possible because high school students in the final year are preparing for their national university entrance exams and they were not allowed to participate. Therefore, data were collected from year-two high school students rather than year-three, leaving an unequal span between levels. Finally, conditions during data collection were not always strictly controlled. Due to the limited facilities in the schools, no phonetic laboratories were available which would have guaranteed the quality of recordings without the presence of extraneous noise. Instead data collection was conducted in a room where noise could sometimes be heard from the streets or classrooms nearby.

6.3 Future research

Some questions raised by the research in this thesis still require further research. The differential substitution of /θ/ and /ð/ would be a fascinating research focus. However, due to the complex nature of the errors found in the Cantonese and Harbinese groups, existing research is not sufficient to explain why dialect/language speakers used one type of substitution rather than another. Therefore, further research needs to be conducted to explain the differential substitutions of /θ/ and /ð/ made by Harbinese and Guangzhou Cantonese speakers, and especially the substitutions of the voiced interdental fricative /ð/.

Future research could also consider the consonant clusters which unfortunately had to be excluded from the present analysis. Both Chinese dialects/languages lack clusters (except consonants with glides in Harbinese) in their syllable structures, and thus it would be very interesting to see if the production of subjects in both dialect/language groups differs. If so, it would further confirm the influence of L1 dialect on the L2. In addition, if different specific error types are found for Harbinese and Cantonese speakers, this again has implications for teaching.

Research which takes into account the varieties of English used by L2 English learners is scarce. The results of the present study show different tendencies for the English varieties used by each dialect/language-speaking group, but these differences were not statistically significant. Research could be carried out to determine the influence of the use of English varieties by different Chinese dialect/language-speaking groups of L2 English learners.

This thesis classifies the targets used into three categories of ‘identical’, ‘similar’ and ‘new’. The category of ‘similar’ was composed of three different types. There are no published findings reporting a hierarchy of difficulty among these three types. In this thesis, only two hypotheses compare different types of ‘similar’ together and all other hypotheses compare one type of ‘similar’ with ‘new’ or ‘identical’. Although a difficulty ranking has been found from the results for the categories of ‘identical’, ‘similar’ and ‘new’ and different types of similarities, it also requires more research to prove whether the finding works.

This thesis combined CAH and SLM which hold opposite ideas to check degree of difficulty of ‘similar’ vs. ‘new’. From the results the prediction of Flege’s SLM that ‘similar’ is more difficult to learn than ‘new’ is rejected and instead all three types of ‘similar’ are found to be easier than the ‘new’. In this case, I posited that the degree of difficulty to the categories ‘similar’ and ‘new’ can be mutually transformed from each other, depending on the L2 learner’s level of proficiency. Since most of the results support the prediction of Lado’s (1957) CAH that ‘similar’ is easy, I then posit that ‘similar’ could be more likely to be difficult in cases of nuanced differences between the L1 and the L2, which would then support Flege’s (1995) SLM. More relevant research is needed. In addition, the question of whether the SLM indeed applies more to higher-level than lower-level learners also requires further investigation, looking at identical vs. similar vs. new by the year levels within each group.

This research focuses only on production by L2 learners, but it would be very interesting if their perception were also considered in order to see how different dialect speakers perceive English phonology and what differences there may be between dialects/languages.

This thesis compares the production of English by Harbinese and Cantonese speakers who were recruited in three separate levels, based on year in school/university. I carried out a general analysis for target vs. non-target performance for these year-groups; however, differences between subjects at different levels in terms of segments, syllable structure and stress were not examined in detail. Therefore, questions of how speakers at different proficiency levels acquire English phonology, what error features they may have for that specific level, and what difficulties speakers at different levels experience are still open.

References

Aijmer, K., Altenberg, B. and Johansson, M. (eds.) 1996. *Languages in contrast. Papers from a symposium on text-based cross-linguistic studies, Lund 4–5 March 1994* [Lund Studies in English 88]. Lund University Press.

Anderson, J. 1987. ‘The markedness differential hypothesis and syllable structure difficulty’. *Interlanguage Phonology: The Acquisition of a Second Language Sound System*, 279-291.

Archangeli, D. and D. T. Langendoen. 1997. *Optimality Theory*. USA: Blackwell Publishers.

Barrie, M. 2003. ‘Contrast in Cantonese Vowels’. *Toronto Working Papers in Linguistics*, 20, 1-19.

Bassetti, B., and Atkinson, N. 2015. ‘Effects of orthographic forms on pronunciation in experienced instructed second language learners’. *Applied Psycholinguistics*, 36(1), 67-91.

Best, C. T. 1993. ‘Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development’. In *Developmental Neurocognition: Speech and Face Processing in the First Year of Life,* 289-304. Dordrecht: Springer.

Best, C. T. 1994. ‘The emergence of native-language phonological influences in infants: A perceptual assimilation model’. *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, 167(224), 233-277.

Best, C. T., McRoberts, G.W. and Sithole, N. M. 1988. ‘Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants’. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 345.

Belz, J. 2003. Identity, deficiency, and first language use in foreign language education. In C. Blyth (ed.), *The sociolinguistics of foreign language* *classrooms: contributions of the native, the near-native, and the non-native speaker*. Boston: Heinle and Heinle, 209–50.

Bialystok, E. 1997. The structure of age: in search of barriers to second language acquisition'. *Second language Research*, 13, 116-137.

Blyth, C. 1995. 'Redefining the boundaries of language use: the foreign language classroom as a multilingual speech community'. In C. Kramsch (ed.), *Redefining the boundaries of language study*. Boston: Heinle and Heinle. 145–84.

Boersma, P. and Weenink, D. 2018. Praat: doing phonetics by computer [Computer program]. Version 6.0.37, retrieved 14 March 2018 from http://www.praat.org/.

Broselow, E., Chen, S., and Wang, C. 1998. 'The emergence of the unmarked in second language phonology'. *Studies in Second Language Acquisition*, 20: 261–280.

Brown, A. 1991. *Pronunciation Models*. Singapore: Singapore University Press.Burstall, C. 1975. 'Factors affecting foreign language learning: A consideration of some recent research findings'. *Language Teaching and Linguistics Abstracts*, 8, 5–25.

Chan, A. Y., and Li, D. C. 2000. 'English and Cantonese phonology in contrast: Explaining Cantonese ESL learners' English pronunciation problems'. *Language Culture and Curriculum*, 13(1), 67-85.

Chao, Y. R. 1968. *A grammar of spoken Chinese*. Univ of California Press.

Chen, P., Teng H. and Song, C. 2010. 'Influences of Hubei Dialects in English Pronunciation'. *Journal of Wuhan Institute of Technology*, 32 (2), 78-80.

Chen, W. Y. 2013. 'The Negative Transfer of Cantonese on English Consonants /r/ and /l/'. *Journal of Yangtze University*, 36 (3), 88-89.

Chen, Y. 2010. 'The Negative Transfer of Northeastern Dialect on English Learners' English Phonology and Strategies'. *Time Education*, 3, 12-13.

Chen, Y. 2011. *Practical Cantonese Guide*. Guangzhou: World Publishing Corporation.

Collins, B. and Mees, I. M. 1984. *The sounds of English and Dutch*. Brill Archive.

Cook, V. 1999. ‘Using SLA research in language teaching’. *International Journal of Applied Linguistics*, 9(2), 267–84.

Cook, V. 2004. *The English Writing System.* London: Arnold.

Corder, S. P. 1967. ‘The significance of learners’ errors’. *International Review of Applied Linguistics*, 5, 161-170.

Corder, S. P. 1971. ‘Idiosyncratic dialects and error analysis’. *IRAL-International Review of Applied Linguistics in Language Teaching*, 9(2), 147-160.

Cowan, W. 1968. ‘Notes toward a definition of modern standard Arabic’. *Language Learning*, 18, 29–34.

Cucchiarini, C. 1996. ‘Assessing transcription agreement: methodological aspects’. *Clinical Linguistics & Phonetics*, 10(2), 131-155.

Cui, H.N. 2014. ‘Analysis of the Phonetic Features of the Northeast Accent’. *Contemporary Literature*, 2014(1), 190.

Davenport, M. and Hannahs, S. J. 2013. *Introducing Phonetics and Phonology*. Routledge.

Davis, L. M. 1971. ‘Some social aspects of the speech of Blue-Grass Kentucky’. In Williams, J. V. & Burke, V. M. (eds) *A various language: Perspectives on American dialects*. New York: Holt, Rinehart and Winston. 335–340.

Delattre, P. 1964. ‘Comparing the consonantal features of English, German, Spanish, and French’. *International Review of Applied Linguistics*, 2, 155-203.

Delattre, P. 1969. ‘An acoustic and articulatory study of vowel reduction in four languages’. *International Review of Applied Linguistics*, 7, 295-323.

Deterding, D., Wong, J. and Kirkpatrick, A. 2008. ‘The pronunciation of Hong Kong English’. *English World-Wide*, 29(2), 148-175.

Disner, S. F. 1983. WPP, No. 58: Vowel Quality: The relation between Universal and Language-specific Factors.

Dou, L. L. 2010. ‘The Analysis of College Freshmen Phonetic Leaning Error and its Implications to College English Teaching’. *Journal of Qiqihar Junior Teachers’ College*, 3, 155-157.

Duanmu, S. 2000. *The Phonology of Standard Chinese.* New York: Oxford University Press.

Duanmu, S. 2007. *The Phonology of Standard Chinese Second Edition.* New York: Oxford University Press.

Dubois, S. and Horvath, B.M. 1998. 'Let's tink about dat: Interdental fricatives in Cajun English'. *Language Variation and Change*, 10(3), 245-261.

Dulay, H., Burt, M. and Krashen, S. 1982. *Language Two*. Oxford: Oxford University Press.

Eckman, F. R. 1977. 'Markedness and the contrastive analysis hypothesis'. *Language learning*, *27*(2), pp.315-330.

Edwards, J.G.H. and Zampini, M.L. (eds) 2008. *Phonology and second language acquisition* (Vol. 36). John Benjamins Publishing.

Ellis, R. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Ethnologue. 2020. Last accessed on 10th April 2020 via https://www.ethnologue.com/guides/most-spoken-languages.

Fan, R. L. and He, S. Z. 2012. 'Language Transfer Transcendence of SLA: from Contrastive Analysis to Interlanguage'. *Journal of Qinghai Normal University (Philosophy and Social Sciences)*, 34(3), 111-115.

Fatemi, M. A., Sobhani, A. and Abolhassani, H. 2012. 'Difficulties of Persian Learners of English in Pronouncing Some English Consonant Clusters'. *World Journal of English Language*, 4, 69-75.

Fischer, J.L. 1958. 'Social influences on the choice of a linguistic variant'. *Word*, *14*(1), pp.47-56.

Fisiak, J. (ed.). 1991. *Further Insights into Contrastive Analysis* (Vol. 30). John Benjamins Publishing.

Fisiak, J., Lipińska-Grzegorek, M. and Zabrocki, T. 1978. *An Introductory English-Polish Contrastive Grammar*. Państ. Wydaw. Naukowe.

Flege, J. E. 1988. 'Factors affecting degree of perceived foreign accent in English sentences'. *Journal of the Acoustical Society of America*, 84, 70-79.

Flege, J. E. 1992. ‘The intelligibility of English vowels spoken by British and Dutch talkers’. *Intelligibility in Speech Disorders: Theory, Measurement, and Management*, 1, 157-232.

Flege, J. E. 1995. ‘Second language speech learning: Theory, findings, and problems’. In W. Strange (ed.) *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press. 233–277.

Flege, J. E. 1997. ‘English vowel production by Dutch talkers: more evidence for the “similar” vs.“new” distinction’. *Second-language speech*, 11-52.

Flege, J. E. 1998. ‘The role of subject and phonetic variables in second language learning’. In *Papers from the 34th annual meeting of the Chicago linguistic Society, Vol. II*: *The panels* (K. Gruber, D. Higgins, K. Olsen & T. Wysocki, editors). Chicago: Chicago Linguistic Society.

Flege, J. E. and Fletcher, K. L. 1992. ‘Talker and listener effects on degree of perceived foreign accent’. *The Journal of the Acoustical Society of America*, *91*(1), pp.370-389.

Flege, J. E. and Wang, C. 1989. ‘Native-language phonotactic constraints affect how well Chinese subjects perceive the word-final English/t/-/d/contrast’. *Journal of phonetics*, 17(4), 299-315.

Flege, J. E., Frieda, E. M., and Nozawa, T. 1997. ‘Amount of native-language (L1) use affects the pronunciation of an L2’. *Journal of Phonetics, 25,* 169–186.

Flege, J. E., MacKay, I. R. A., and Meador, D. 1999. ‘Native Italian speakers’ perception and production of English vowels’. *Journal of the Acoustical Society of America, 106*, 2973–2987.

Flege, J. E., Munro, M. J. and MacKay, I. R. A. 1995. ‘Factors affecting strength of perceived foreign accent in a second language’. *Journal of the Acoustical Society of America*, 97, 3125-3134.

Flege, J. E., Schirru, C. and MacKay, I. R. 2003. ‘Interaction between the native and second language phonetic subsystems’. *Speech communication,* 40(4), 467-491.

Flynn, C.-Y.-C. 2001. *Intonation in Cantonese* (Doctoral dissertation, University of London (School of Oriental and African Studies).

Flynn, C.-Y.-C. 2003. *Intonation in Cantonese*, München: LINCOM Europa.

Fry, D. B. 1958. ‘Experiments in the perception of stress’. *Language and Speech,* 1(2). 126-152.

Fonda, C. 1984. ‘An analysis of pronunciation mistakes and second language strategy in the case of Italian and mandarin speaking learners of English’, in: M. P. R. van den Broecke - A. Cohen (eds.)*, Proceedings of the Tenth International Congress of Phonetic Sciences.* Dordrecht: Foris.

Fry, D. B. 1958. ‘Experiments in the perception of stress’. *Language and speech*, *1*(2), 126-152.

Gao, L. and Deng, Y.C. 2009. ‘A Study on the Stress Displacement of English Words for Chinese College Students:A Corpus-based Study’. *Foreign Language World,* 3, 10-16, 44.

Gass, S. 1979. ‘Language transfer and universal grammatical relations’. *Language learning*, *29*(2), pp.327-344.

Gass, S. M., Behney, J. and Plonsky L. 2013. *Second language acquisition: An introductory course, 4th ed*. Routledge.

Gong C. and Tian, Y. 2008. ‘The Influence of Northeastern Dialect on English and Chinese Phonological Acquisition’. *Shaanxi Education,* 10: 30-31.

Google Map. 2019. Last accessed in July 2019 via: https://www.google.com/maps/dir/Harbin,+Heilongjiang,+China/Guangzhou,+Guangdong,+China/@33.1133676,101.5471518,4z/data=!4m14!4m13!1m5!1m1!1s0x5e4364f8a6641461:0x5e7c92735aa02cd5!2m2!1d126.534967!2d45.803775!1m5!1m1!1s0x3402f895a35c2bc7:0xe59e075adeae415!2m2!1d113.264385!2d23.12911!5i2.

Gottardo, A., Yan, B., Siegel, L.S. and Wade-Woolley, L. 2001. ‘Factors related to English reading performance in children with Chinese as a first language: More evidence of cross-language transfer of phonological processing’. *Journal of Educational Psychology,* 93(3), 530-542.

Gowrie, M. S. S., Saravanakumar, A. R. and Subbiah, S. 2012. ‘Incorporation of Pronunciation Skills in English Teaching for Beginners’. *Golden Research Thoughts*, 1 (7), 1-3.

Griner, B. D. 2014. *Practical Tips for Teaching English Vowels to Chinese Speakers*. 2014 Hawaii University International Conferences on Education and STEM.

Groves, J. M. and Mair, V. H., 2008. Language or Dialect - or Topolect? A comparison of the Attitudes of Hong Kongers and Mainland Chinese towards the Status of Cantonese.

Guangzhou Admissions and Examination Committee Office. 2015. 2016 Guangzhou Senior High School English Exam Reform Program. Last accessed on 28 July 2019 via: https://www.liuxue86.com/k_广州中考英语改革方案/.

Gui, M. C. 2005. *The Phonology of Guangzhou Cantonese.* München: Lincom Europa.

Guion, S. G., Flege, J. E., Akahane-Yamada, R. and Pruitt, J. C. 2000. ‘An investigation of current models of second language speech perception: The case of Japanese adults’ perception of English consonants’. *The Journal of the Acoustical Society of America*, *107*(5), 2711-2724.

Guo, Y.B. 2006. ‘A Review of Flege's Speech Learning Model’. *Journal of Taiyuan Institute of Education*, 24, 71-73.

Hammer, C. S., Komaroff, E., Rodriguez, B., Lopez, L., Scarpino, S. and Goldstein, B. G. 2012. ‘Predicting Spanish–English bilingual children’s language abilities’. *Journal of Speech, Langauge, and Hearing Research*, 55, 1251-64.

Hansen, J. G. 2004. ‘Developmental sequences in the acquisition of English L2 syllable codas’. *Studies in Second Language Acquisition*, 26: 85–124.

Harbin Education Bureau, 2014. Harbin High School Entrance Exams Guide Book (English), last accessed on 28 July 2019 via https://news.21cnjy.com/A/130/204/V62755.shtml.

Hashimoto, O. K. Y. (1972). *Studies in Yue dialects 1: Phonology of Cantonese, (Vol. 1)*. Cambridge University Press.

Hattem, J. 2009. ‘Substitution Patterns for English Inter-dental Fricatives by L1 Latin American Spanish-Speakers’. Retrieved from Rutgers Optimality Archive website accessed via :*http://andromeda.rutgers.edu/~hattem/Academic%20 Research/Sample%20Research%20Paper%20Substitution_Patterns_for_EnglishInter_Dental_Fricatives_(David_J_Hattem)_FINAL_PAPER. pdf*.

Horvath, B. M. 1985. ‘Variation in Australian English’. The Sociolects of Sydney. *Cambridge Studies in Linguistics London*, (45), 1-200.

Hou, M.J. 2012. ‘Contrastive Analysis, Error Analysis and Interlanguage Theory in Second Language Acquisition Research’. *Journal of Jishou University (Social Sciences Edition)*, 32(3), 159-162.

Howie, J. M. 1976. *Acoustical studies of Mandarin vowels and tones (Vol. 18)*. Cambridge University Press.Huang, X. and Yang, J. 2012. ‘China’s Regional Education Inequality Decomposition and Improvement Measures’. *Journal of Beijing Institute of Technology (Social Sciences Edition)*, 14(1), 150-156.

Hume, E. 2003. ‘Language Specific Markedness: The Case of Place of Articulation’. *Studies in Phonetics, Phonology and Morphology*.

Hung, T. T. N. 2000. ‘Towards a Phonology of Hong Kong English’. *World Englishes,* 19 (3), 337-356.

Hyman, L. 1977. ‘On the Nature of Linguistic Stress’, in Larry Hyman (ed.), *Studies in Stress and Accent, Southern California Occasional Papers in Linguistics*, 4, 37-82.

James, A. 1984. ‘Phonic transfer: the structural bases of interlingual assessments’, in: M.P.R. van den Broecke - A. Cohen (eds.), *Proceedings of the Tenth International Congress of Phonetic Sciences*. Dordrecht: Foris.

James, A.1984. ‘Syntagmatic segment errors in non-native speech’. *Linguistics*, 22, 481-505.

James , C. 1980. *Contrastive Analysis*. Harlow Essex: Longman.

Jia, S. 2011. ‘An Experimental Study on the Negative Transfer of Voiceless Plosive Endings: A Case of Cantonese EFL Learners’. *Journal of Guangdong university of Education,* 31 (4). 74-79.

Jiang, X. L. 2015. 'Comparative Study of the Phonetic and Phonological Differences between Two Varieties (Harbin and Shenyang) of Northeastern Dialect'. MA Dessertation, Newcastle University.

Johansson, S. and Oksefjell, S. (eds.) 1998. *Corpora and cross-linguistic research: theory, method, and case studies*. Amsterdam: Rodopi.

Kager, R. 1999. *Optimality Theory.* Cambridge: CUP.

Kawahara, S. 2010. '*Script of get_intensity_minmax*', last accessed on 5 February 2019 via: http://user.keio.ac.jp/~kawahara/scripts/get_intensity_minmax.praat.

Klein, W. 1986. *Second Language Acquisition*. Cambridge: Cambridge University Press.

Kohmoto, S. 1965. *Applied English Phonology*. Tokyo: Tanaka Press.

Kong, Y. F. and Wang, X. M. 2001. 'The Influence of Mandarin and Dialect on English Pronunciation Teaching'. *Journal of Da Tong Vocational Technical College*, 15(1), 51-52.

Kramsch, C. 2007. 'Re-reading Robert Lado, 1957, Linguistics across Cultures. Applied linguistics for language teachers'. *International Journal of Applied Linguistics*, *17*(2), 241-247.

Kuhl, P. K. and Iverson, P. 1995. 'Linguistic experience and the "perceptual magnet effect"'. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 121-154.

Kuo, Y.C., Rosen, S. and Faulkner, A., 2008. 'Acoustic cues to tonal contrasts in Mandarin: Implications for cochlear implants'. *The Journal of the Acoustical Society of America*, *123*(5), 2815-2824.

Lado, R. 1957. *Linguistics across cultures. Applied linguistics for language teachers*. University of Michigan Press.

Larsen-Freeman, D. and Long. M. 1991. *An Introduction to Second Language Acquisition Research.* London: Longman.

Lee, W. and Zee, E. 2003. 'ILLUSTRATIONS OF THE IPA: Standard Chinese (Beijing)'. *Journal of the International Phonetic Association,* 33, 1.

Li, B. 1998. 'The Generation, Fundamental Principles and Application of Optimality Theory'. *Modern Foreign Languages,* 3, 72-91.

Li, C. N. and Thompson, S. A. 1989. *Mandarin Chinese: A Functional Reference Grammar*. California: University of California Press.

Li, L. 2006. 'The influence of Chinese on English Phonological Teaching and Strategies'. *Journal of Basic English Education,* 8 (2), 3-7.

Li, T. 2008. 'The Investigation and Analysis of English Phonology and Mandarin Phonology in Cantonese-Speaking Areas'. *Academic Exchange*, 168 (3), 145-147, 192.

Li, X. 2009. 'The English Phonological Acquisition of English Learners in Cantonese Context and Learning Strategies'. *Economic and Social Development,* 7(2), 191-195.

Li, X. D. 2012. 'From the Phonology to Learn Putonghua Term Stress Analyses'. *The Guide of Science & Education*, 1, 81-82.

Li, Y. L. and Chen, Z. P. 2007. 'Analysis on the Common English Pronunciation Mistakes of Cantonese Students'. *Journal of Guangdong university of Technology (Social Sciences Edition),* 7(3), 66-68.

Liang, M. 2010. 'A Brief Analysis on the Features of Singlish Influenced by Dialects'. *Around Southeast Asia*, 15, 84-87.

Liao, Y. Y. 2014. 'Discussion on Guangdong students' pronunciation errors of English consonants /n/ and /l/'. *Modern Enterprise Education*, 7, 393.

Lightbown, P. M. and Spada, N. 2006. *How languages are learned (3rd ed.)*. Oxford: Oxford University Press.

Lin, H. 2001. *A Grammar of Mandarin Chinese.* Muenchen: Lincom Europa.

Lin, Y. H. 2007. *The sounds of Chinese.* Cambridge: Cambridge University Press.

Liu, J. and Guo, R. 2013. 'A study of English pronunciation errors in practice of Cantonese-speaking English majors of Dongguan University of Technology'. *Journal of Dongguan University of Technology*, 20(2), 70-74.

Liu, F. 1924. *Szu Sheng Shih Yen Lu*. Ch'un Yi Shanghai.

Liu, L. Z. 2013. ‘Discussion on the phonetic features of Heilongjiang dialect’. *Journal of Heilongjiang College of Educatioin*, 32(10), 124-126.

Liu, P. and Y. Wang. 2007. ‘Transference on the acquisition of English pronunciation from Yue dialect’. *Journal of Harbin University*, 28(7), 119-122.

Lombardi, L. 2003. ‘Second language data and constraints on Manner: Explaining substitutions for the English interdentals’. *Second Language Research’* 19: 225–250.

Ma, H. X and Wang, L. 2013. ‘Analysis of word stress displacement for Chinese learners of English’. *Journal of Hubei Normal University (Philosophy and Social Science)*, 33(4), 148-150, 156.

Ma, Q. 2015. *What Is Phonology?*. Shanghai: Shanghai Foreign Language Education Press.

Ma, Q. and B. Chen. 2004. ‘The constraints in OT: Typology and definition’. *Journal of PLA University of Foreign Languages*, 27, 20-25.

Mahmoud, A. 2000. ‘Modern Standard Arabic vs. Non-Standard Arabic: Where do Arab students of EFL transfer from?’. *Language culture and curriculum*, *13*(2), 126-136.

Major, R. C. 1986. ‘The ontogeny model: Evidence from L2 acquisition of Spanish’. *Language Learning*, *36*(4), 453-504.

Major, R. C. 2013. *Foreign Accent. The Encyclopedia of Applied Linguistics*, Edited by Carol A. Chapelle. UK: Blackwell Publishing Ltd.

Matthews, S. and Yip, V. 1994. *Cantonese: A Comprehensive Grammar*. USA & Canada: Routledge.

Mayo, M. P. G. 1999. ‘Grammaticality judgements, metalinguistic awareness and the age factor in EFL’. In P. Gallardo and E. Llurdá (eds) *Proceedings of the XXII International AEDEAN Conference*. Lleida: Universitat de Lleida. 561–566.

Mayo, M. D. P. G. 2003. ‘Age, length of exposure and grammaticality judgements in the acquisition of English as a foreign’. *Age and the Acquisition of English as a Foreign Language*, 4, 94.

McAllister, R., Flege, J. E. and Piske, T. 1999. 'The acquisition of Swedish long vs. short vowel contrasts by native speakers of English, Spanish and Estonian'. In *Proceedings of the 14th International Congress of Phonetic Sciences* (Vol. 1, 751). Berkeley, CA: University of California.

McCarthy, J. 2003. 'Comparative markedness [short version]'. *Theoretical Linguistics.*

McCarthy, J. J. and Prince, A., 1995. 'Faithfulness and reduplicative identity'. *Linguistics Department Faculty Publication Series*, p.10.

McCarthy, J. J. and Prince, A., 1999. 'Faithfulness and identity in Prosodic Morphology'. The Prosody-Morphology Interface, ed. by René Kager, Harry van der Hulst and Wim Zonneveld, 218-309.

MCCD. 2014. 'Multi-function Chinese character database: With archaic script forms, etymologies, English-Chinese lookup, & word-formations phonologically disambiguated according to the Cantonese dialect'. Last accessed on 2 December 2016 via http://humanum.arts.cuhk.edu.hk/Lexis/lexi-mf/.

McKenzie, R. M. 2004. 'Attitudes of Japanese nationals towards standard and non-standard varieties of Scottish English speech'. *East Asian Learner*, 1, 16–25.

Meador, D., Flege, J. E. and MacKay, I. R. A. 2000. 'Factors affecting the recognition of words in a second language'. *Bilingualism*: l*anguage and Cognition*, 3, 55-67.

Meng, G. Z., An, H. Y. and Zheng, W. F. 2001. *Mandarin Proficiency Testing Guide Book.* Harbin: Heilongjiang Education Press

MEPRC (Ministry of Education of the People's Republic of China). 2001. '*The ministry of education's guidance on actively promoting English courses in primary schools'.* Last accessed on 5 April 2018 via: http://old.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe711/200407/665.html.

Mitchell, R., and Myles, F. 2004. *Second Language Learning Theories (2nd ed.).* London: Arnold.

Mitchell, R., Myles, F. and Marsden, E. 2019. *Second language learning theories*. Routledge.

Mohammed, A. M. 1992. 'Error-based interlingual comparisons as a learner-centred technique of teaching grammar'. *PhD thesis*, University of Salford.

Morgan, L. C. 1971. 'North Carolina accents'. In Williams, J. V. and Burke, V. M. (eds) *A various language: Perspectives on American dialects*. New York: Holt, Rinehart and Winston. 268–279.

Moyer, A. 1999. 'Ultimate attainment in L2 phonology'. *Studies in Second language Acquisition*, 21, 81-108.

Mukattash, L. 1981. 'Wh-questions in English: A problem for Arab students'. *IRAL International Review of Applied Linguistics in Language Teaching*, 19, 317-325.

National Bureau of Statistics of China. 2019. Provincial Annual GDP Data for 2018. Last accessed on 5 August 2019 via: http://data.stats.gov.cn/search.htm?s=2018 年广东 GDP, and from http://data.stats.gov.cn/search.htm?s=2018 年黑龙江 GDP.

Nemser, W. 1961. *Hungarian Phonetic Experiments*. (Research and Studies in Uralic and Altaic Languages Project 32.) New York. American Council of Learned Societies. 1961b. 'The interpretation of English stops and interdental fricatives by native speakers of Hungarian'. *Dissertation Abstracts M, 111*, 262-267.

Nemser, W. 1971. 'Approximative systems of foreign language learners'. *IRAL-International Review of Applied Linguistics in Language Teaching*, *9*(2), 115-124.

Nie, Z. P. 2005. 'A Brief Account of Heilongjiang dialect'. *Journal of Harbin University*, 26, 109-113.

Nie, Z. P. 2005. 'Discussion on Harbin dialect dictionary'. *Journal of Qiqihar Junior Teachers' College*, 90(2), 35-39.

Nie, Z. P. 2005. 'On two problems in Harbin dialect dictionary'. *Journal of Daqing Normal University*, 25(1), 67-69.

Niu, J. 2011. ‘The Influence of the Dialect of English Phonetic Learning - Based on Henan Dialect as a Case Study’. *Minzu University of China.* [MA Dissertation].

Odlin, T., 1989. Language Transfer: Cross-linguistic Influence in Language Learning. Cambridge University Press, Cambridge.Oller Jr, J. W. and Ziahosseiny, S. M. 1970. ‘The contrastive analysis hypothesis and spelling errors’. *Language learning*, *20*(2), 183-189.Ouyang, T. 2018. ‘An empirical study of English voiceless fricatives by native Mandarin speakers’. MA Dissertation, Hunan University.

Oyama, S. 1976. ‘A sensitive period for the acquisition of a nonnative phonological system’. *Journal of Psycholinguistic Research, 5,* 261–283.

Panyajirawut, W. 2009. *Factors influencing confidence in Speaking English: A case study of senior Thai undergraduate students at Srinakharinwiroj University*. Bangkok: Thammasat University.

Paolillo, J. C. and Das, A., 2006. Evaluating language statistics: The ethnologue and beyond. *Contract report for UNESCO Institute for Statistics*.

Peng, Y. 2009. ‘Analysis on the positive transfer of Standard Chinese in English pronunciation learning’. *Journal of Hunan University of Arts and Science,* 34(6), 134-136.

Peust, C. 1996. Sum: th-substitution. *The Linguist List* 7.1108. Available from http://linguistlist.org/issues/7/7-1108.html.

Ping, W. 2001. ‘The influence of Henan dialect on the learning of English pronunciation and intonation’. *Journal of Henan Vocational-Technical Teachers' College,* 3, 66-67.

Piske, T. 2007. ‘Implications of James E. Flege's research for the foreign language classroom’. In O. Bohn, and M. J. Munro (eds) *Language experience in second language speech learning. In honor of James Emil Flege.* 301-314.

Piske, T. 2008. ‘Phonetic awareness, phonetic sensitivity and the second language learner’. *Encyclopedia of language and education,* 1912-1923.

Piske, T., Flege, J. E., MacKay, I. R. A., and Meador, D. 2002. 'The production of English vowels by fluent early and late Italian-English bilinguals'. *Phonetica, 59*, 49–71.

Piske, T., MacKay, I.R. and Flege, J.E., 2001. 'Factors affecting degree of foreign accent in an L2: A review'. *Journal of phonetics*, *29*(2), pp.191-215.

Politzer, R. L. 1968. 'An experiment in the presentation of parallel and contrasting structures'. *Language Learning*, *18*(1-2), 35-43.

Prince, A. and Smolensky, P. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar. Manuscript, Rutgers University and University of Colorado, Boulder.*

Pye, C., Wilcox, K. A. and Siren, K. A. 1988. 'Refining transcriptions: the significance of transcriber 'errors'. *Journal of Child Language*, *15*(1), 17-37.

Rasier, L. and Hiligsmann, P. 2007. 'Prosodic transfer from L1 to L2. Theoretical and methodological issues'. *Nouveaux cahiers de linguistique française*, *28*(2007), 41-66.

Rau, D.V., Chang, H.H.A. and Tarone, E.E., 2009. 'Think or sink: Chinese learners' acquisition of the English voiceless interdental fricative'. *Language Learning*, *59*(3), pp.581-621.

Riney, T. J. and Flege, J. E. 1998. 'Changes over time in global foreign accent and liquid identifiability and accuracy'. *Studies in Second language Acquisition*, 20, 213-244.

Roach, P. 2004. 'British English: Received pronunciation'. *Journal of the International Phonetic Association*, 34 (2), 239–245.

Roach, P., Hartman, J. and Setter, J. 2006. '*Daniel Jones' English pronouncing dictionary*. Cambridge University Press.Schachter, J. 1974. 'An error in error analysis 1'. *Language learning*, *24*(2), 205-214.Scott, M. and Tucker, G. 1974. 'Error analysis and English-language strategies of Arab students'. *Language Learning,* 24, 69–7.

Ritchie, W. 1968. 'On the explanation of phonic interference', *Language Learning*, 18: 183-197.

Scovel, T. 2000. ‘A critical review of the critical period hypothesis’. *Annual Review of Applied Linguistics,* 20, 213–23.

Seliger, H. W., Krashen, S. D. and Ladefoged, P. 1975. ‘Maturational constraints in the acquisition of second language accent’. *Language Sciences,* 36, 20-22.

Selinker. L. 1969. Language transfer. *General Linguistics* 9:67-92.

Selinker, L. 1972. ‘Interlanguage’. *IRAL-International Review of Applied Linguistics in Language Teaching*, *10*(1-4), 209-232.

Selinker, L. 1992. *1992: Rediscovering Interlanguage*. London: Longman.

Selkirk, E., and Shen. T. 1990. ‘Prosodic Domain in Shanghai Chinese’. *The Phonology-Syntax Connection. University of Chicago Press, Chicago/London*, 145-161.

Shen, X., Ding, J., and Liu, J. 2015. ‘Study on English teaching and learning connection between senior school and college -New thinking under background of college entrance examination reform’. *Research in Teaching,* 38(3), 115-120.

Shi, G. J. 1996. ‘An Analysis of the Influence of Mandarin on the Pronunciation and Intonation of English Beginners’. *Journal of Harbin College,* 4, 91.

Shriberg, L. D. and Lof, G. L. 1991. ‘Reliability studies in broad and narrow phonetic transcription’. *Clinical Linguistics & Phonetics*, *5*(3), 225-279.

Singleton, D. and Lengyel, Z. (eds) 1995. *The Age Factor in Second Language Acquisition*. Clevedon: Multilingual Matters.

Smith, Ian. 1978. *Sri Lanka Creole Portuguese Phonology*. Vanchiyoor, Trivandrum, India: Dravidian Linguistics Association.

Smith, J., Durham, M. and Fortune, L., 2007. ‘"Mam, ma troosers is fa'in doon!" Community, caregiver and child in the acquisition of variation in Scottish dialect’. *Language Variation and Change*, *19*(1), pp.63-99.Smith, J., Durham, M. and Richards, H. 2013. ‘The social and linguistic in the cquisition of sociolinguistic norms: Caregivers, children, and variation’. *Linguistics*, *51*(2), pp.285-324.

Smolensky, P. and Prince, A., 1993. ‘Optimality Theory: Constraint interaction in generative grammar’. *Optimality Theory in phonology*, *3*.

Song, J. and Yu, X. 2007. *Introduction to Cantonese*. Harbin: Harbin Institute of Technology Press.

Song, J. and Yu, X. 2012. *Introduction to the Cantonese*. Harbin: Harbin Institute of Technology Press.

Starks, D. and Paltridge, B. 1996. ‘A note on using sociolinguistic methods to study non-native attitudes towards English’. *World Englishes*, 15 (2), 217-224.

Stemberger, J. P. and Stoel-Gammon, C. 1991. ‘The underspecification of coronals: Evidence from language acquisition and performance errors’. In *The special status of coronals: Internal and external evidence*, 181-199. Academic Press.

Stockwell, R. and Bowen, J. 1965. *The Sounds of English and Spanish*. Chicago: University of Chicago Press.

Strange, W. 1995. ‘Phonetics of Second-Language Acquisition: past, present, future’. *Proc. ICPhS, 95*, 76-83.

Tadros, A. A. 1966. ‘An analysis of the interference errors in the written English of Sudanese students’. MA Dessertation, University of Khartoum.

Thomas, M. 2013. ‘History of the study of L2 acquisition’ In J.R. Herschensohn, and M. Young-Scholten, (ed.) *The Cambridge handbook of second language acquisition*. Cambridge: Cambridge University Press. 26-45.

Thompson, I. 1991. ‘Foreign accents revisited: the English pronunciation of Russian immigrants’. *Language learning*, 41, 177-204.

Thompson-Panos, K. and Thomas-Ruzie, M. 1983. ‘The least you should know about Arabic: Implications for the ESL writing instructor’. *TESOL Quarterly* 17, 609–23.

Toury, G. 1995. *Descriptive translation studies and beyond*. Amsterdam: Benjamins.Třísková, H., 2011. ‘The structure of the Mandarin syllable: Why, when and how to teach it’. *Archiv orientální*, *79*(1), 99-134.

Venuti, L. (ed.) 2000. *The translation studies reader*. London: Routledge.

Wan, B. 2008. ‘Differences of syllable structure between Cantonese and Mandarin-On Mandarin teaching in Hong Kong’. *Journal of Shaanxi Normal University*, 37, 43-50.

Wang, J. 2002. *On Tone Sandhi and Neutral Tone in the Tianjin Dialect within the Framework of Optimality Theory*. China Academic Journal Electronic Publishing House.

Wang, J. Y. 2015. ‘A Brief Analysis of the Substitution of English Dental Fricatives for Mandarin Learners’. *Course Education Research*, 2015(35), 87.

Wang, L. and Liu, X. M. 2017. ‘Acoustic experimental study on the influence of Northeastern dialect on learning English monophthongs’. *Culture and Education*, 12, 200-202.

Wang, M., Koda, K., and Perfetti, C. A. 2003. ‘Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners’. *Cognition*, 87(2), 129-149.

Wang, X. 2007. ‘The impact of Chinese on English Phonology’. *Journal of Liaoning Administration College*, 9(5), 212-213.

Wang, Y. 2009. ‘A brief analysis of the negative transfer of Cantonese dialect pronunciation in English pronunciation acquisition’. *Crazy English Teachers Version*, 2009(4), 116-117, 160.

Wardhaugh, R. 1970. ‘The contrastive analysis hypothesis’. *TESOL quarterly*, 123-130.

Weinberger, S. H. 1997. ‘Minimal segments in second language phonology’. *Second language speech: Structure and process*, 263-312.

Weinreich, U. 1953. *Languages in contact.* New York: Linguistic Circle of New York.

Weinreich, U. 1953. *Languages in Contact: Findings and Problems.* The Hague: Mouton.

Weinreich, U. 1968. *Languages in Contact*. The Hague: Mouton.

Weinreich, U. 2010. *Languages in contact: Findings and problems* (No. 1). Walter de Gruyter.

Wells, J. C. 1982. *Accents of English*. Vols. 1-3. Cambridge: Cambridge University Press.

Wester, F., Gilbers, D. and Lowie, W., 2007. Substitution of dental fricatives in English by Dutch L2 speakers. *Language Sciences*, *29*(2-3), 477-491.

Whitman, R. L. 1970. ‘Contrastive analysis: problems and procedures’. *Language Learning*, 20(2), 191-197.

Whitman, R. L. and Jackson, K. L. 1972. ‘The unpredictability of contrastive analysis’. *Language learning*, 22(1), 29-41.

Whorf, B. L. 1941. ‘Languages and logic’. *Foundations of Cognitive Psychology*, 707-715.

Wiltshire, C. R. and J. D. Harnsberger. 2006. ‘The influence of Gujarati and Tamil L1s on Indian English: A preliminary study’. *World Englishes*. 25(1), 91-104.

Wu, Z.J. 1986, *The Spectrographic Album of Mono-syllables of Standard Chinese*. Beijing: Social Science Press.

Xiong, Z. 2004. ‘Script of syllable pitch extraction’. *Praat user manual*. Last accessed on 5 February 2019 via: http://paslab.phonetics.org.cn/index.php/author/xiongziyu/.

Xu, C. F. and Song, J. J. 2008. ‘An analysis of the stress displacement for University English learners’. *Neijiang Keji*, 5, 65-66.

Xu, W., Wang, Y. and Case, R. E. 2010. ‘Chinese attitudes towards varieties of English: A pre-olympic examination’. *Language Awareness*, 19(4), 249-260.

Yan, B. 2007. ‘Brief talk about the transfer effects of Chinese on English phonology. *Journal of Hubei University of Economics*, 4(9), 142-143.

Yan, Y. and Yang, X. 2011. ‘The impact of Northeastern dialect on English pronunciation of university students and teaching strategies. *China Education Innovation Herald*, 1, 163

Yang, C. 2010. ‘Phonological studies on Liaoning dialect’. *Journal of Liaoning Normal University*, 33, 93-99.

Yang, W.Y. 1997. 'An Analysis of the special obstacles in the English pronunciation of Guangdong students'. *Journal of Guangzhou Normal University (Social Science Edition)*, 3, 101-105.

Yang, X. 2014. 'The study on the negative transfer of Northeastern dialect on English pronunciation learning and teaching strategies'. *Journal of Inner Mongolia Normal University*, 27 (5), 124-126.

Yarmohammadi, L. 2002. 'A contrastive analysis of Persian and English'. Tehran: Payame Noor.

Yavas, M. 2016. *Applied English Phonology*. John Wiley & Sons.

Yin, L. Y. and Li, Z. Y. 2014. 'Negative transfer of Heilongjiang dialect to English pronunciation learning and countermeasures'. *Chinese Off-Campus Education*, 36, 132-133.

Yin, S. 1998. *Harbin Dialect Files*. Shanghai: Shanghai Education Press.

Yip, M. 1992. 'Prosodic morphology in four Chinese dialects'. *Journal of East Asian Linguistics*, 1:1-35.

Young-Scholten, M. 1985. 'Interference reconsidered: The role of similarity in second language acquisition'. *Selecta* 6, 6–12.

Yu, L. Q. 2009. 'The influence of speaking good Mandarin on learning English pronunciation'. *Journal of Chuzhou University*, 11(2), 112-113.

Yuan, S. H. and Wu, J. 2017. 'An empirical study on Chinese college students' English word stress misplacement'. *Journal of Changshu Institute of Technology (Philosophy & Social Sciences)*, 6, 83-88, 99.

Yue, C. J. 2003. 'The influence of regional differences of economic development level on the allocation of educational resources'. *Education and Economics*, 1, 35-41.

Zampini, M., 2008. 'L2 speech production research'. *Phonology and second language acquisition*, 219-249.

Zee, E. 1999. 'Chinese (Hong Kong Cantonese)'. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*, Cambridge: Cambridge University Press

Zhan, B. and Gan, Y. 2012. *Guangfu Dialect*. Guangzhou: Jinan University Press.

Zhan, B., Zhou, W., Li, R., Shan Z., and Chen. H. 2004. *Guangzhou Cantonese Pronunciation Dictionary*. Guangzhou: Guangdong People's Press.

Zhang, J. 2013. ‘The study on the transfer of Anhui dialectal initials on English consonants’. *Journal of Jiamusi Education College,* 127(5), 242, 248.

Zhang, J. 2017. ‘Research on influences of Cantonese on English pronunciation and intonation study and countermeasures from the perspective of language transfer theory’. *Journal of Changsha University*, 31(3), 155-158.

Zhang, W. and Hu, G. 2008. ‘Second language learners' attitudes towards English varieties’. *Language Awareness*, 17(4), 342-347.

Zhang, Y. 2012. ‘The impact of Northeastern dialect on the English pronunciation and intonation of university students’. *Mass Literature,* 28, 162-163.

Zhang, Y., Nissen, S. L. and Francis, A. L. 2008. ‘Acoustic characteristics of English lexical stress produced by native Mandarin speakers’. *The Journal of the Acoustical Society of America*, *123*(6), 4498-4513.

Appendix 1.1 Participant information form (English version)

Participant Information Form

Dear participants,

I am a PhD student in Newcastle University, now conducting research on *How Different Chinese Dialect speakers acquire English phonology* as my PhD thesis. To achieve the answer to this question, I need to collect data from speakers who speak these dialects as their first dialect. Therefore, I need your help and support. Now, I am formally inviting you to be one of my participants and to work with me. You will fill in a questionnaire and complete translation of words from Chinese to English and then a piece of Chinese reading task. Don't worry! A certain amount of time and guidance will be given to familiarise you with the words before you carry it out. Your speech productions will be audio recorded for future analysis as my thesis data. I assure you that all the information you provide will only be used as data for my study, and it will be reported anonymously in my thesis. All the information provided by you is significant and precious for my research. Hereby, I express my sincere and endless gratitude to you for your help and support.

Newcastle University requires that all researchers should have formal consent from the participants who will participate in any research. For this reason, please sign your name and the date below to indicate your consent.

Thank you very much!

Yours sincerely,
Xinliang Jiang

Appendix 1.2 Participant information form (Chinese version)

参 与 者 知 情 书

亲爱的参与者：

您好！我是纽卡斯尔大学的一名在读博士研究生，目前正在研究不同汉语方言者如何习得英语音系。为解答此问题，我需要从说这些方言的母语者那里搜集数据。因此，我需要您的支持与帮助。现在，我正式邀请您成为研究的参与者，与我共同参与此项研究。您将需要填写调查问卷，把汉语词汇译成英语以及阅读汉语字词。不用担心！在开始之前，您会获得一定的指导并有时间来熟悉这些词汇。您的回答将被录音，并经分析后，作为我的论文数据。我向您保证，您提供的所有信息仅用作研究之用，并将以匿名的形式出现在我的论文中。您提供的数据对我的研究来说意义重大且宝贵。为此，我诚挚地感谢您所给予的支持与帮助！

纽卡斯尔大学规定所有研究者需获得参与者给予的正式书面同意书。因此，请您在下方签署您的姓名及日期以表同意。

非常感谢！

此致

敬礼！

蒋新亮

Appendix 1.3 Consent form (English version)

Consent Form

Project title: Chinese EFL Learners' Acquisition of Phonology: A Comparative Analysis of the Influence of Two Dialects (Harbin Northeastern Dialect and Guangzhou Cantonese)

Consent to the use of all data I provide.

I have read the statement provided for the above research project and I have had the opportunity to ask questions. I consent to participate in this research project. I understand that all the data will be kept confidential and I will be anonymous in the research report. I also know that the data gathered from this project will be used for the purposes stated in the Participant Information Form.

I understand that participation is voluntary and that withdrawal from the project is possible at any time without needing to give a reason.

Signature of participant: ________________________

Date: ___________________

Appendix 1.4 Consent form (Chinese version)

同　意　书

研究项目标题：中国英语学习者的音系习得：对比分析哈尔滨东北方言和广州粤方言的影响

同意使用我所提供的所有数据。

我已阅读上述研究项目所述要求，并有机会提出疑问。本人同意参与这项研究项目，知道所有数据将被保密并以匿名形式进行研究报告，也知道此项目所搜集的数据将用于参与者知情书中所述用途。

我知道参与是自愿行为，而且知道可以随时退出研究项目而无需任何理由。

参与者签名：________________________　　日期：__________________

Appendix 2.1 Questionnaire (English version)

QUESTIONNAIRE ABOUT EFFECTS OF CHINESE DIALECTS ON THE ACQUISITION OF ENGLISH PHONOLOGY

All responses will be treated anonymously.

Background Information:

Name: ________________________ Gender: Male/ Female __________

Age: _________________________ Hometown: ___________________

Please tick "√" the best answer and answer the questions.

1. Education Level:
   - A. Primary School — Year (1, 2, 3, 4, 5/6)
   - B. Middle School — Year (6/7, 8, 9)
   - C. High School — Year (10, 11, 12)
   - D. University (Bachelor) — Year (1, 2, 3, 4)
   - E. Master
   - F. PhD

2. Were you born and raised in your hometown?
   A. Yes B. No

3. Have you stayed in your hometown for most of your life?
   A. Yes B. No

4. How long have you not lived in your hometown?
   A. 0 years B. 1-3 years C. 4-6 years D. 7-10 years
   E. More than 10 years

5. Have you lived in your hometown for the past 3 years?
   A. Yes B. No

6. Was your father also born and raised in your hometown?
   A. Yes B. No
   C. Born here but not raised here D. Not born here but raised here

7. Was your mother also born and raised in your hometown?
   A. Yes B. No
   C. Born here but not raised here D. Not born here but raised here

Language/Dialect Information:

8. What dialects do your parents speak?
   Mother:___________________________ Father:___________________________

   A. Harbin dialect
   B. Guangzhou Cantonese
   C. None of the above What is that? ______________________

9. Which dialect(s) did your parents use in the home when you were growing up?
   Mother:___________________________ Father:___________________________

   A. Harbin dialect
   B. Guangzhou Cantonese
   C. None of the above What is that? ______________________

10. Which Chinese dialect is your first dialect (your mother tongue)?
   A. Harbin dialect If you choose A, do you think Harbin dialect is the same as Mandarin? _______
   B. Guangzhou Cantonese
   C. None of the above What is that? ______________________

11. What other Chinese dialects can you speak, if any? How long have you learned/used them?
   (For example, *Hakka---5 years*)
   ____
   ____

12. Which dialect/language do you speak at home?
   ____

13. How many hours per day, on average, do you speak this dialect?
   ____

14. How many hours per day, on average, do you speak other dialects?
   (For example, *Mandarin---5 hours*)
   ____
   ____

15. When did you start to learn Mandarin?
   A. Preschool At what age? Aged ______
   B. Primary School Year (1, 2, 3, 4, 5/6)
   C. Middle School Year (6/7, 8, 9)

16. When did you start to learn English?
   A. Preschool At what age? Aged ______
   B. Primary School Year (1, 2, 3, 4, 5/6)
   C. Middle School Year (6/7, 8, 9)

17. How long have you learned English for, so far?

____

18. What other languages can you speak, if any (besides Chinese and English)?

____

19. Did you learn English from foreign teachers?
    A. Yes   B. No   If you choose 'No', please skip to Q23.

20. During what time periods have you learned English from foreign teachers?
    (For example, *aged 12-16; aged 21-present; 6 months in 2016*)

____

21. Where did your foreign teachers come from? If you have had more than one foreign teacher, please list all of them in detail (e.g. giving name, country of origin, duration and place you have learned English with her/him).
    (For example, *Kate – UK – 5 months -- in school/outside school*)

____

____

22. If you have had many foreign teachers who influenced you the most?

____

23. How did you learn English outside of class? How many hours per week, on average, did you spend?
    (For example, *watching English movies --- 5 hours*)

____

24. Which do you prefer, American films/TV series or British films/TV series?
    A. British   B. American

25. Which variety of English do you like?
    A. British   B. American

26. Which variety of English do you think you speak?
    A. British   B. American   C. Chinglish   D. Others

27. Can you distinguish between English accents and American accents?
    A. Yes   B. No   If you choose 'Yes, please answer 28.

28. How do you distinguish British English from American English, that is, what are the differences between them?

____

Thank you for your help!

Appendix 2.2 Questionnaire (Chinese version)

Newcastle University

有关汉语方言对英语音系习得影响的调查问卷

所有答复将匿名处理！

背景信息：

姓名: ______________________ 性别: 男 / 女

年龄: ______________________ 家乡: ______________________

请在最佳答案处划"√"并回答问题。

1. 教育程度：
   A. 小学 （1, 2, 3, 4, 5/6）年级
   B. 中学 （6/7, 8, 9） 年级
   C. 高中 （10, 11, 12） 年级
   D. 大学 （1, 2, 3, 4） 年级
   E. 硕士
   F. 博士

2. 您是在您的出生地长大的吗？
   A. 是 B. 不是

3. 您是大部分时间都生活在您的家乡吗？
   A. 是 B. 不是

4. 您有多长时间没在您的家乡生活？
   A. 0 年 B. 1-3 年 C.4-6 年 D. 7-10 年 E.10 年以上

5. 近 3 年，您一直在家乡生活吗？
   A. 是 B. 没有

6. 您的父亲也在您的家乡出生并长大吗？
   A. 是 B. 不是
   C. 在这出生，但不在这长大 D. 不在这出生，但在这长大

7. 您的母亲也在您的家乡出生并长大吗?
   A. 是 B. 不是
   C. 在这出生，但不在这长大 D. 不在这出生，但在这长大

语言/方言信息：

8. 您的父母说什么方言？

母亲：___________________________ 父亲:___________________________

A. 哈尔滨话
B. 广州话
C. 以上都不是 那是什么方言：__________________________

9. 在您的成长过程中，您的父母在家说什么方言？

母亲：___________________________ 父亲:__________________________

A. 哈尔滨话
B. 广州话
C. 以上都不是 那是什么方言：__________________________

10. 哪种汉语方言是您的第一方言(母语)?

A. 哈尔滨话 如您选 A，您认为哈尔滨话和普通话相同吗？_______
B. 广州话
C. 以上都不是 那是什么方言：___________________________

11. 您还会说其它汉语方言吗？如果会，您学了/使用了多长时间？

(例如，*客家方言——5 年*)

____

12. 您在家说什么方言/语言？

____

13. 您平均每天说这种方言有多少个小时？

____

14. 您平均每天说其它方言有多少个小时？

(例如，*普通话——5 小时*)

____

____

15. 您是什么时候开始学习普通话的？

A.学前 几岁？ ___岁
B.小学 (1, 2, 3，4, 5/6） 年级
C.中学 (6/7，8, 9) 年

16. 您是什么时候开始学习英语的？

A.学前 几岁？ ___岁
B.小学 (1, 2, 3、4, 5/6） 年级
C.中学 (6/7，8, 9) 年级

17. 到目前为止，您学习英语多久？

____

18. 除汉语和英语外，您还会说哪种语言？

____

19. 您跟外籍教师学过英语吗？
A.学过　　B.没有　　如果您选择“没有”，请跳到第 23 题。

20. 您什么时候跟外籍教师学的英语？
(例如，*12-16 岁期间；21 岁到现在；2016 年学了 6 个月*）

____

21. 您的外籍教师来自哪个国家？如果您不止跟一个外籍教师学过英语，请您详细列出所有外籍教师的情况（如：姓名，国籍，您跟他/她学习英语的时长与地点）。
(例如，*Kate——英国——5 个月——校内/校外*）

____

____

22. 如果您跟很多外籍教师学过英语，哪一位对您的影响最大？

____

23. 您在校外是怎样学习英语的？平均每周花费多少小时？
(例如，*看英语电影---5 小时*)

____

24. 您更喜欢美国电影/美剧，还是英国电影/英剧？
A.英国　　B.美国

25. 您喜欢哪种英语？
A.英式英语　　B.美式英语

26. 您认为您说的英语属于哪种类型？
A. 英式英语　　B.美式英语　　C. 中式英语　　D.其它

27. 您能区分英式口音和美式口音吗？
A.能　　B.不能

28. 您如何区分英式口音和美式口音，也就是说，英式口音和美式口音有什么不同？

____

____

谢谢您的帮助!

Appendix 3.1 Research instrument --- Test Tokens (English version)

Test Tokens

I. Please interpret the following Chinese words into English.

1. 兽医
2. 湿的
3. 船
4. 通常
5. 孩子
6. 日本
7. 希望
8. 认为
9. 他们 (主格)
10. 很，非常
11. 想知道；奇迹
12. 鞋
13. 快乐，高兴(名词)
14. 选择(名词)
15. 笑话
16. 怎样，如何
17. 感谢
18. 那，那个
19. 蔬菜
20. 方式，方法；道路
21. 害羞
22. 休闲的；随便的
23. 机会
24. 跳
25. 她的
26. 理论
27. 那里
28. 每一的，每个的
29. 离开，远离(副词)
30. 亚洲
31. 自然
32. 想象
33. 也许
34. 生日
35. 父亲
36. 结束；在...之上
37. 总是
38. 机器
39. 姿态，手势
40. 危险(名词)
41. 使守规矩;行为动词
42. 地震
43. 没有(介词)
44. 发展(动词)
45. 三明治
46. 病人；耐心的
47. 世纪
48. 工程师
49. 在...后面
50. 方法，办法
51. 天气
52. 波浪；挥手
53. 矮树丛，灌木；布什
54. 看(电视，比赛)
55. 消息，留言
56. 地球
57. 和...;有...(介词）
58. 五
59. 完成
60. 触摸，接触
61. 大的，大量的
62. 嘴
63. 呼吸(动词)
64. 住，居住
65. 新鲜的
66. 教，教授
67. 年龄
68. 牙 (单数)
69. 给...穿衣服
70. 夜晚
71. 洞，孔
72. 正确的；右面的
73. 知道
74. 闻
75. 一行，一排
76. 小刀
77. 学校
78. 公路，马路，道路
79. 今晚
80. 搬运，抬
81. 国际的
82. 士兵
83. 外国的
84. 午夜
85. 世界
86. 几个；几个的
87. 好的，优良的
88. 生活
89. 数字
90. 自己的
91. 低的
92. 汽车
93. 计划
94. 灯光；清的
95. 地板

96．座位
97．绵羊
98．离开，留下(动词)
99．坐，坐下，就坐
100．池塘，(游泳)池
101．食物
102．二
103．拉 (推的反义词)
104．脚（单数）
105．鸟
106．护士
107．词，单词
108．关于，大约
109．相机
110．今天
111．促使，引起，造成
112．画 (动词)
113．小的
114．热的
115．盒子
116．在...之后
117．快速的 (f 开头)
118．班级
119．血，血液
120．来
121．切，砍
122．气体
123．胖的
124．坏的
125．让
126．床
127．男孩
128．噪音，噪声
129．外面的；在外面
130．现在
131．去
132．外套
133．我(主格)
134．领带
135．高的(h 开头)
136．英文第一个字母
137．支付，付款
138．五月

139．美丽的
140．一些
141．主要的
142．音乐
143．时间
144．大脑
145．新的
146．停止
147．城镇
148．语言
149．深的
150．在下面，向下(副词)
151．十分，相当
152．杯子
153．线；排成一行
154．问题，疑问
155．肉
156．石头
157．说(某种语言)
158．独自的，单独的
159．看，瞧(l 开头)
160．硬币
161．生病的
162．加入，参加

163．飞机（全称）
164．烟花
165．护照
166．明信片
167．雪人
168．市中心
169．角色扮演
170．牙刷
171．程序；节目
172．认识到，意识到
173．兴趣；利息
174．机器人

II. Please read the following words in your own dialect.

175．奶奶
176．娃娃
177．楼上
178．大年
179．位子
180．龙船
181．大连
182．青蛙
183．娘娘
184．李好
185．伟大
186．女老师
187．农船
188．温州
189．吕老师
190．你好
191．旺旺
192．牛奶

III. Please translate the following words into English, fill in the sentences, and read the entire English sentence fluently.

1-95,104, 120, 121, 140, 143, 146, 149, 152, 155, 157, 159, 161.

I think he said ________ yesterday.

Thank you for your help!

Appendix 3.2 Research instrument --- Test Tokens (Chinese version)

测试词汇

一、请把下列英语词汇译成汉语。

1. 兽医
2. 湿的
3. 船
4. 通常
5. 孩子
6. 日本
7. 希望
8. 认为
9. 他们 (主格)
10. 很，非常
11. 想知道；奇迹
12. 鞋 (单数)
13. 快乐，高兴(名词)
14. 选择(名词)
15. 笑话
16. 怎样，如何
17. 感谢
18. 那，那个
19. 蔬菜
20. 方式，方法；道路
21. 害羞的
22. 休闲的；随便的
23. 机会
24. 跳
25. 她的
26. 理论
27. 那里
28. 每一的，每个的
29. 离开，远离(副词)
30. 亚洲
31. 自然
32. 想象
33. 也许
34. 生日
35. 父亲
36. 结束；在...之上
37. 总是
38. 机器
39. 姿态，手势
40. 危险(名词)
41. 使守规矩；行为动词
42. 地震
43. 没有(介词)
44. 发展(动词)
45. 三明治
46. 病人；耐心的
47. 世纪
48. 工程师
49. 在...后面
50. 方法，办法
51. 天气
52. 波浪；挥手
53. 矮树丛，灌木；布什
54. 看(电视，比赛)
55. 消息，留言
56. 地球
57. 和...;有...(介词）
58. 五
59. 完成
60. 触摸，接触
61. 大的，大量的
62. 嘴
63. 呼吸(动词)
64. 住，居住
65. 新鲜的
66. 教，教授
67. 年龄
68. 牙 (单数)
69. 给...穿衣服
70. 夜晚
71. 洞，孔
72. 正确的；右面的
73. 知道
74. 闻
75. 一行，一排
76. 小刀
77. 学校
78. 公路，马路，道路
79. 今晚
80. 搬运，抬
81. 国际的
82. 士兵
83. 外国的
84. 午夜
85. 世界
86. 几个；几个的
87. 好的，优良的
88. 生活
89. 数字
90. 自己的
91. 低的
92. 汽车
93. 计划
94. 灯光；轻的
95. 地板

96．座位
97．绵羊
98．离开，留下(动词)
99．坐，坐下，就坐
100．池塘，(游泳)池
101．食物
102．二
103．拉 (推的反义词)
104．脚（单数）
105．鸟
106．护士
107．词，单词
108．关于，大约
109．相机
110．今天
111．促使，引起，造成
112．画 (动词)
113．小的
114．热的
115．盒子
116．在...之后；以后
117．快速的 (f 开头)
118．班级
119．血，血液
120．来
121．切，砍
122．气体
123．胖的
124．坏的
125．让
126．床
127．男孩
128．噪音，噪声
129．外面的；在外面
130．现在
131．去
132．外套
133．我(主格)
134．领带
135．高的(h 开头)
136．英文第一个字母
137．支付，付款
138．五月

139．美丽的
140．一些
141．主要的
142．音乐
143．时间
144．大脑
145．新的
146．停止
147．城镇
148．语言
149．深的
150．在下面，向下(副词)
151．十分，相当
152．杯子
153．线；排成一行
154．问题，疑问
155．肉
156．石头
157．说(某种语言)
158．独自的，单独的
159．看，瞧(l 开头)
160．硬币
161．生病的
162．加入，参加

163．飞机（全称）
164．烟花
165．护照
166．明信片
167．雪人
168．市中心
169．角色扮演
170．牙刷
171．程序；节目
172．认识到，意识到
173．兴趣；利息
174．机器人

二、请用你的母语方言读出下列汉字。

175．奶奶
176．娃娃
177．楼上
178．大年
179．位子
180．龙船
181．大连
182．青蛙
183．娘娘
184．李好
185．伟大
186．女老师
187．农船
188．温州
189．吕老师
190．你好
191．旺旺
192．牛奶

三、请把下面词汇翻译成英文，并填在句子里，然后读出整个英文句子。

1-95，104, 120, 121, 140, 143, 146, 149, 152, 155, 157, 159, 161.

I think he said ________ yesterday.

感谢您的帮助！

Appendix 3.3 Research instrument --- Answers to Test Token

Answers to Test Tokens

I. Please interpret the following Chinese words into English.

1.vet
2.wet
3.ship
4.usually
5.child
6.Japan
7.hope
8.think
9.they
10.very
11.wonder
12.shoe
13.pleasure
14.choice
15.joke
16.how
17.thank
18.that
19.vegetable
20.way
21.shy
22.casual
23.chance
24.jump
25.her
26.theory
27.there
28.every
29.away
30.Asia
31.nature
32.imagine
33.perhaps
34.birthday
35.father
36.over
37.always
38.machine
39.gesture
40.danger
41.behave
42.earthquake
43.without
44.develop
45.sandwich
46.patient
47.century
48.engineer
49.behind
50.method
51.weather
52.wave
53.bush
54.watch
55.message
56.earth
57.with
58.five
59.finish
60.touch
61.large
62.mouth
63.breathe
64.live
65.fresh
66.teach
67.age
68.tooth
69.clothe

70. night
71. hole
72. right
73. know
74. smell
75. row
76. knife
77. school
78. road
79. tonight
80. carry
81. international
82. soldier
83. foreign
84. midnight
85. world
86. several
87. fine
88. life
89. number
90. own
91. low
92. car
93. plan
94. light
95. floor

96. seat
97. sheep
98. leave
99. sit
100.pool
101.food
102.two
103.pull
104.foot
105.bird
106.nurse
107.word
108.about
109.camera
110.today
111.cause
112.draw
113.small
114.hot
115.box
116.after
117.fast
118.class
119.blood
120.come
121.cut
122.gas
123.fat
124.bad
125.let
126.bed
127.boy
128.noise
129.out
130.now
131.go
132.coat
133.I
134.tie
135.high
136.A
137.pay
138.May

139.beautiful
140.some
141.main
142.music
143.time
144.brain
145.new
146.stop
147.town
148.language
149.deep
150.down
151.quite
152.cup
153.line
154.question
155.meat
156.stone
157.speak
158.alone
159.look
160.coin
161.sick
162.join

163.airplane
164.firework
165.passport
166.postcard
167.snowman
168.downtown
169.role-play
170.toothbrush
171.program
172.realise
173.interest
174.robot

II. Please read the following words in your own dialect.

175. 奶奶
176. 娃娃
177. 楼上
178. 大年
179. 位子
180. 龙船
181. 大连
182. 青蛙
183. 娘娘
184. 李好
185. 伟大
186. 女老师
187. 农船
188. 温州
189. 吕老师
190. 你好
191. 旺旺
192. 牛奶

III. Please translate the following words into English, fill in the sentences, and read the entire English sentence fluently.

1-95,104, 120, 121, 140, 143, 146, 149, 152, 155, 157, 159, 161.

I think he said ________ yesterday.

Thank you for your help!

Appendix 3.4 Research instrument --- Transcriptions of Test Tokens in IPA

Newcastle University

Transcriptions of Test Tokens in IPA

I. Please interpret the following Chinese words into English.

1. [vεt]
2. [wεt]
3.[ʃɪp]
4.[ˈju:ʒʊəlɪ] [ˈjuʒəlɪ]
5.[tʃaɪld]
6.[dʒəˈpæn]
7.[həʊp][hoʊp]
8.[θɪŋk]
9.[ðeɪ]
10.[ˈvεɹɪ]
11.[ˈwʌndə/ɚ]
12.[ʃuː] [ʃu]
13.[ˈplεʒə/ɚ]
14. [tʃɔɪs]
15.[dʒəʊk] [dʒoʊk]
16.[haʊ]
17.[θæŋk]
18.[ðæt]
19.[ˈvεdʒtəbl]
20.[weɪ]
21.[ʃaɪ]
22.['kæʒjʊəl] ['kæʒʊəl]
23.[tʃɑːns][tʃæns]
24. [dʒʌmp]
25. [hɜː][hɝ]
26. ['θɪəɹɪ]['θiəɹi]
27. [ðεə][ðεɹ]
28. ['εvɹi]
29. [ə'weɪ]
30. [ˈeɪʃə][ˈeɪʒə]
31. ['neɪtʃə] ['neɪtʃɚ]
32. [ɪ'mædʒɪn]
33. [pəˈhæps] [pɚ'hæps]
34. ['bɜːθdeɪ] ['bɝθdeɪ]
35. ['fɑːðə] ['fɑðɚ]
36. ['əʊvə] ['oʊvɚ]
37. [ˈɔːlweɪz]['ɔlweɪz]
38. [mə'ʃiːn][mə'ʃin]
39. ['dʒεstʃə]['dʒεstʃɚ]
40. [ˈdeɪndʒə] ['deɪndʒɚ]
41. [bɪ'heɪv]
42.['ɜːθkweɪk]['ɝθ'kweɪk]
43. [wɪˈðaʊt] [wɪðˈaʊt]
44. [dɪ'vεləp]
45. ['sænwɪdʒ] ['sænwɪtʃ]
46. ['peɪʃ(ə)nt] [ˈpeɪʃənt]
47. [ˈsεntʃəɹɪ]
48. [εndʒɪ'nɪə][ˌεndʒɪ'nɪɹ]
49. [bɪ'haɪnd]
50. ['mεθəd]
51. ['wεðə] ['wεðɚ]
52. [weɪv]
53. [bʊʃ]
54. [wɒtʃ] [wɑtʃ]
55. ['mεsɪdʒ]
56. [ɜːθ] [ɝθ]
57. [wɪð]
58. [faɪv]
59. ['fɪnɪʃ]
60. [tʌtʃ]
61. [lɑːdʒ] [lɑɹdʒ]
62. [maʊθ]
63. [bɹiːð] [bɹið]
64. [lɪv]
65. [fɹεʃ]
66. [tiːtʃ] [titʃ]
67. [eɪdʒ]
68. [tuːθ] [tuθ]
69. [kləʊð] [kloʊð]

70. [naɪt]
71. [həʊl] [hoʊl]
72. [ɹaɪt]
73. [nəʊ] [noʊ]
74. [smεl]
75. [ɹəʊ] [ɹoʊ]
76. [naɪf]
77. [skuːl] [skul]
78. [ɹəʊd] [ɹoʊd]
79. [tə'naɪt]
80. ['kæɹɪ]
81. [ɪntə'næʃ(ə)n(ə)l]
    [,ɪntɚ'næʃnəl]
82. ['səʊldʒə] ['soʊldʒɚ]
83. ['fɒɹɪn] ['fɔɹən]
84. ['mɪdnaɪt]
85. [wɜːld] [wɝld]
86. [ˈsεvɹəl]
87. [faɪn]
88. [laɪf]
89. ['nʌmbə] ['nʌmbɚ]
90. [əʊn] [oʊn]
91. [ləʊ] [loʊ]
92. [kɑː] [kɑɹ]
93. [plæn]
94. [laɪt]
95. [flɔː][flɔɹ]

96. [si:t] [sit]
97. [ʃi:p] [ʃip]
98. [li:v] [liv]
99. [sɪt]
100.[pu:l] [pul]
101.[fu:d] [fud]
102.[tu:] [tu]
103.[pʊl]
104.[fʊt]
105.[bɜ:d] [bɝd]
106.[nɜ:s] [nɝs]
107.[wɜ:d] [wɝd]
108.[ə'baʊt]
109.['kæmərə]
110.[tə'deɪ]
111.[kɔ:z] [kɔz]
112.[drɔ:] [drɔ]
113.[smɔ:l] [smɔl]
114.[hɒt] [hɑt]
115.[bɒks] [bɑks]
116.['ɑ:ftə] ['æftɚ]
117.[fɑ:st] [fæst]
118.[klɑ:s] [klæs]
119.[blʌd]
120.[kʌm]
121.[kʌt]
122.[gæs]
123.[fæt]
124.[bæd]
125.[lɛt]
126.[bɛd]
127.[bɔɪ]
128.[nɔɪz]
129.[aʊt]
130.[naʊ]
131.[gəʊ] [goʊ]
132.[kəʊt] [koʊt]
133.[aɪ]
134.[taɪ]
135.[haɪ]
136.[eɪ]
137.[peɪ]
138.[meɪ]

139.['bju:tɪf(ə/ʊ)l][-rɪ-]
140.[sʌm]
141.[meɪn]
142.['mju:zɪk]
143.[taɪm]
144.[breɪn]
145.[nju:][nu]
146.[stɒp][stɑp]
147.[taʊn]
148.['læŋgwɪdʒ]
149.[di:p] [dip]
150.[daʊn]
151.[kwaɪt]
152.[kʌp]
153.[laɪn]
154.['kwɛstʃ(ə)n]
155.[mi:t] [mit]
156.[stəʊn] [stoʊn]
157.[spi:k]
158.[ə'ləʊn] [ə'loʊn]
159.[lʊk]
160.[kɔɪn]
161.[sɪk]
162.[dʒɔɪn]

163.['ɛəpleɪn] ['ɛɹpleɪn]airplane
164.['faɪəwɜ:k] ['faɪɚwɝk] firework
165.['pɑ:spɔ:t] ['pæspɔɹt] passport
166. ['pəʊs(t)kɑ:d] ['poʊst'kɑɹd] postcard
167. ['snəʊmæn] ['snoʊmæn] snowman
168.['daʊntaʊn] downtown
169.['ɹəul,plei]role-play
170.['tu:θbɹʌʃ] ['tuθbɹʌʃ] toothbrush
171.['pɹəʊgɹæm] ['pɹoʊgɹæm] program
172.['ɹɪəlaɪz] ['ɹiəˌlaɪz]realise
173.['ɪnt(ə)ɹɪst] ['ɪntɹəst] interest
174. ['ɹəʊbɒt] ['ɹoʊbɑt] robot

II. Please read the following words in your own dialect.

175. 奶奶 2n
[nai3nai0] [na:i5na:i5]
176. 娃娃 2w
[wa2wa2][wa:1wa:1]
177. 楼上 1l
[ləu2ʂɑŋ4][lɐu4sœŋ5/6]
178. 大年 1n
[da4njan2][da:i6nin4]
179. 位子 1w
[wei4tsz0][wei6tsi2]
180. 龙船 1l
[luŋ2ʈʂ^{h}uan2][luŋ4syn4]
181. 大连 1l
[da4ljan2][da:i6lin4]
182. 青蛙 1w
[tɕ^{h}iŋ1wa1][ts^{h}iŋ1wa:1]
183. 娘娘 2n
[njaŋ2njaŋ0][nœŋ4nœŋ4]
184. 李好 1l
[li3xɑu3][lei5hou2]
185. 伟大 1w
[wei3da4][wei5da:16]
186. 女老师 1n
[ny^{3}lɑu3ʂʅ][nœi5lou5si1]
187. 农船 1n
[nuŋ2tʃuan2][nuŋ4syn4]
188. 温州 1w
[wən1ʈʂəu1][wɐn1tsɐu1]
189. 吕老师 1l
[ly^{3}lɑu3ʂʅ1][lœi5lou5si1]
190. 你好 1n
[ni3xɑu3][nei5hou2]
191. 旺旺 2w
[waŋ^{4}waŋ^{4}][wɔ:ŋ6wɔ:ŋ6]
192. 牛奶 1n
[njəu2nai3][ŋɐu4na:i1/5]

Thank you for your help!

Appendix 4.1 Praat script of syllable pitch extraction

This script is extracted from Handbook for Praat of Ziyu, Xiong (2004:157-160) but adapted by Yang Long.

```
form 对话框
positive the_Index_of_Referenced_Tier_in_TextGrid 1
positive the_Number_of_Pitch_Points_in_a_Interval 1000
endform
dirPath$ = "C:\temp\"
Create Strings as file list... list 'dirPath$'\*.TextGrid
fileNum = Get number of strings
tierNum = the_Index_of_Referenced_Tier_in_TextGrid
pointNum = the_Number_of_Pitch_Points_in_a_Interval
for ifile to fileNum
select Strings list
fileName$ = Get string... ifile
newFileName$ = fileName$ - ".TextGrid"
textGridFileName$ = newFileName$ +".TextGrid"
textGridFileName$ = dirPath$ +textGridFileName$
pitchTierFileName$ = newFileName$    +".PitchTier"
pitchTierFileName$ = dirPath$+pitchTierFileName$
saveFileName$ =    newFileName$ +"_Pitch.txt"
saveFileName$ = dirPath$ +saveFileName$
filedelete 'saveFileName$'
fileappend 'saveFileName$' 文件名
fileappend 'saveFileName$' ,
fileappend 'saveFileName$' 标注内容
fileappend 'saveFileName$' ,
fileappend 'saveFileName$' 起点时间
fileappend 'saveFileName$' ,
fileappend 'saveFileName$' 末点时间
fileappend 'saveFileName$' ,
fileappend 'saveFileName$' 时长
fileappend 'saveFileName$' ,
fileappend 'saveFileName$' 最小值
fileappend 'saveFileName$' ,
fileappend 'saveFileName$' 最大值
fileappend 'saveFileName$' ,
fileappend 'saveFileName$' 平均值
fileappend 'saveFileName$' ,
```

```
Read from file... 'pitchTierFileName$'
select PitchTier 'newFileName$'
pitchPointNum=Get number of points
for pitchNum from 1 to pitchPointNum
pitchTime'pitchNum'=Get time from index... 'pitchNum'
endfor
pitchNum=pitchPointNum+1
pitchTime'pitchNum'=Get finishing time
Read from file... 'textGridFileName$'
select TextGrid 'newFileName$'
dd=Is interval tier... 'tierNum'
if dd=1
intervalNum= Get number of intervals... 'tierNum'
endif
if dd=0
intervalNum= Get number of points...    'tierNum'
endif
beginTime0=0
endTime0=0
for interNum from 1 to intervalNum
if dd=1
labeName'interNum'$= Get label of interval... 'tierNum' 'interNum'
beginTime'interNum'=Get starting point... 'tierNum' 'interNum'
endTime'interNum'=Get end point... 'tierNum' 'interNum'
selLength'interNum'=endTime'interNum'-beginTime'interNum'
endif
if dd=0
labeName'interNum'$= Get label of point... 'tierNum' 'interNum'
xuhao='interNum'-1
beginTime'interNum'=endTime'xuhao'
endTime'interNum'=Get time of point... 'tierNum' 'interNum'
selLength'interNum'=endTime'interNum'-beginTime'interNum'
endif
startPoint=1
pitchStartTime=0
pitchEndTime=0
startSel=0
sT=beginTime'interNum'
eT=endTime'interNum'
pitchStartTime'interNum'=0
pitchEndTime'interNum'=0
for pitchNum from startPoint to pitchPointNum
```

```
if pitchTime'pitchNum'>=sT and pitchTime'pitchNum'<=eT
startSel=startSel+1
if startSel=1
pitchStartTime'interNum'=pitchTime'pitchNum'
endif
pitchNextNum=pitchNum+1
pitchNextTime=pitchTime'pitchNextNum'
if pitchNextTime>eT
pitchEndTime'interNum'=pitchTime'pitchNum'
endif
startPoint=pitchNum-1
endif
endfor
endfor
select PitchTier 'newFileName$'
dianNum= pointNum

for interNum from 1 to intervalNum
if interNum =1
#fileappend 'saveFileName$' 基频起点
#fileappend 'saveFileName$' ,
#fileappend 'saveFileName$' 基频末点
#fileappend 'saveFileName$' ,
#for dian from 1 to dianNum
#fileappend 'saveFileName$' 点
#fileappend 'saveFileName$' 'dian'
#fileappend 'saveFileName$' -基频值
#fileappend 'saveFileName$' ,
#endfor

fileappend 'saveFileName$' 'newline$'
endif
fileappend 'saveFileName$' 'textGridFileName$'
fileappend 'saveFileName$' ,
lN$=labeName'interNum'$
fileappend 'saveFileName$' 'lN$'
fileappend 'saveFileName$' ,
bT=beginTime'interNum'
fileappend 'saveFileName$' 'bT'
fileappend 'saveFileName$' ,
eT=endTime'interNum'
fileappend 'saveFileName$' 'eT'
```

```
fileappend 'saveFileName$' ,
sL=selLength'interNum'
fileappend 'saveFileName$' 'sL'
fileappend 'saveFileName$' ,
pST=pitchStartTime'interNum'
pET=pitchEndTime'interNum'

if pST>0 and pET>0
#fileappend 'saveFileName$' 'pST'
#fileappend 'saveFileName$' ,
#fileappend 'saveFileName$' 'pET'
#fileappend 'saveFileName$' ,
a=pST
b=pET
c=('b'-'a')/('dianNum'-1)
select PitchTier 'newFileName$'
tempposition= 'a'

total=0
number=0

for pitchdata from 1 to 'dianNum'
pitchTemp= Get value at time... 'tempposition'
if number = 0
	maxV=pitchTemp
	minV=pitchTemp
else
	if maxV < pitchTemp
	maxV=pitchTemp
	endif
	if minV > pitchTemp
	minV = pitchTemp
	endif
endif

total=total+pitchTemp
number=number+1
#fileappend 'saveFileName$' 'pitchTemp'
#fileappend 'saveFileName$' ,
tempposition='tempposition'+'c'
endfor
meanValue=total/number
```

```
fileappend 'saveFileName$' 'minV'
fileappend 'saveFileName$' ,
fileappend 'saveFileName$' 'maxV'
fileappend 'saveFileName$' ,
fileappend 'saveFileName$' 'meanValue'
fileappend 'saveFileName$' ,
endif

fileappend 'saveFileName$' 'newline$'
endfor
endif
select TextGrid 'newFileName$'
Remove
select PitchTier 'newFileName$'
Remove
endfor
select Strings list
Remove
exit
```

Appendix 5.1 Praat script of syllable intensity extraction

This script is originally from Shigeto Kawahara (2010) but adapted by Yang Long.

This Praat script will get average intensity, minimal intensity, and maximal intensity (in dB) of all labelled intervals of all (or a specified set of) files in a folder.

To use, you specify a folder with wav.files, and base names if you want to analyse only a subset of files.

The script assumes that you already have labelled intervals. The textgrid files and sound files should have the same name.

Written by. Shigeto Kawahara # version 2/10/2010 http://user.keio.ac.jp/~kawahara/scripts/get_intensity_minmax.praat

```
form Get Intensity
	sentence Directory ./
	comment If you want to analyze all the files, leave this blank
	word Base_file_name
	comment The name of result file
	text textfile intensity_list.txt
endform

#Print one set of headers

fileappend "'textfile$'" File name'tab$'Interval name'tab$'Avg Int'tab$'Min Int'tab$'Min Int Time'tab$'Max Int'tab$'Max Int time'tab$'
fileappend "'textfile$'" 'newline$'

#Read all files in a folder

Create Strings as file list... wavlist 'directory$'/'base_file_name$'*.wav
Create Strings as file list... gridlist 'directory$'/'base_file_name$'*.TextGrid
n = Get number of strings


for i to n
clearinfo
#We first extract intensity tiers
```

```
	select Strings wavlist
	filename$ = Get string... i
	Read from file... 'directory$'/'filename$'
	soundname$ = selected$ ("Sound")
	To Intensity... 100 0

# We print out the file names
	labelline$ = "'soundname$'"tab$'"
	fileappend intensity_list.txt 'labelline$'

# We now read grid files and extract all intervals in them
	select Strings gridlist
	gridname$ = Get string... i
	Read from file... 'directory$'/'gridname$'
	int=Get number of intervals... 1

# We calculate intensity for all labeled intervals
for k from 1 to 'int'
	select TextGrid 'soundname$'
	label$ = Get label of interval... 1 'k'
	if label$ <> ""

		# calculates the onset and offset
		onset = Get starting point... 1 'k'
		offset = Get end point... 1 'k'

		#calculates the intensity values
		select Intensity 'soundname$'
		min_int = Get minimum... onset offset Parabolic
		min_time = Get time of minimum... onset offset Parabolic
		max_int = Get maximum... onset offset Parabolic
		max_time = Get time of maximum... onset offset Parabolic
		meanIntensity = Get mean... onset offset dB


		resultline$ =
"'label$"tab$"meanIntensity"tab$"min_int"tab$"min_time"tab$"max_int"tab$"max_ti
me"tab$'"
		fileappend "'textfile$'" 'resultline$'
	endif
endfor
```

```
fileappend "'textfile$'" 'newline$'
endfor

# clean up

select all
Remove
```

Appendix 6.1 Results of questionnaires with 65 participants

Questions	**Choices**	**Harbin (34) frequency/%**	**Guangzhou(31) frequency/%**	**Total (65) frequency/%**
Gender	Male	38% (13/34)	42% (13/31)	40% (26/65)
	Female	62% (21/34)	58% (18/31)	60% (39/65)
Age	14	24% (8/34)	26% (8/31)	25% (16/65)
	15	9% (3/34)	6% (2/31)	8% (5/65)
	16	26% (9/34)	13% (4/31)	20% (13/65)
	17	12% (4/34)	19% (6/31)	15% (10/65)
	20	9% (3/34)	29% (9/31)	18% (12/65)
	21	18% (6/34)	6% (2/31)	12% (8/65)
	22	3% (1/34)		2% (1/65)
Hometown	Harbin	100% (34/34)		52% (34/65)
	Guangzhou		100% (31/31)	48% (31/65)
	Others			
Q1	A. Primary School			
	B. Middle School	35% (12/34)	32% (10/31)	34% (22/65)
	Year 9	12	10	22
	C. High School	35% (12/34)	32% (10/31)	34% (22/65)
	Year 11	12	10	22
	D. University	29% (10/34)	35% (11/31)	32% (21/65)
	Year 3	10	11	21
	E. MA			
	F. PhD			
Q2	A. Yes	100% (34/34)	100% (31/31)	100% (65/65)
	B. No			
Q3	A. Yes	100% (34/34)	100% (31/31)	100% (65/65)
	B. No			
Q4	A. 0 years	100% (34/34)	100% (31/31)	100% (65/65)
	B. 1-3 years			
	C. 4-6 years			
	D. 7-10 years			
	E. more than 10 ys			
Q5	A. Yes	100% (34/34)	100% (31/31)	100% (65/65)
	B. No			
Q6	A. Yes	91% (31/34)	74% (23/31)	83% (54/65)
	B. No	3% (1/34)	3% (1/31)	3% (2/65)
	C. Born here but not raised here			
	D. Not born here but	6% (2/34)	23% (7/31)	14% (9/65)

	raised here						
Q7	A. Yes	85% (29/34)		61% (19/31)		74% (48/65)	
	B. No			23% (7/31)		11% (7/65)	
	C. Born here but not raised here						
	D. Not born here but raised here	15% (5/34)		16% (5/31)		15% (10/65)	
Q8		Mother	Father	Mother	Father	Mother	Father
	A. Harbinese	100% (34/34)	100% (34/34)			52% (34/65)	52% (34/65)
	B. GZ Cantonese			94% (29/31)	97% (30/31)	45% (29/65)	46% (30/65)
	C. None of the above			6% (2/31)	3% (1/31)	3% (2/65)	2% (1/65)
Q9		Mother	Father	Mother	Father	Mother	Father
	A. Harbinese	100% (34/34)	100% (34/34)			52% (34/65)	52% (34/65)
	B. GZ Cantonese			100% (31/31)	100% (31/31)	48% (31/65)	48% (31/65)
	C. None of the above						
Q10	A. Harbinese	100% (34/34)				52% (34/65)	
	1. Same	88% (30/34)					
	2. Different	9% (3/34)					
	3. Similar	3% (1/34)					
	B. GZ Cantonese			100% (31/31)		48% (31/65)	
	C. None of the above						
Q11	Yes, Mandarin			6% (2/31)		3% (2/65)	
	M-8years			3% (1/31)		2% (1/65)	
	M-9years			3% (1/31)		2% (1/65)	
	M-10years			10% (3/31)		5% (3/65)	
	M-11years			10% (3/31)		5% (3/65)	
	M-12years			6% (2/31)		3% (2/65)	
	M-14years			10% (3/31)		5% (3/65)	
	M-16years	3% (1/34)				2% (1/65)	
	No	97% (33/34)		13% (4/31)		57% (37/65)	

	No, but Mandarin		3% (1/31)	2% (1/65)
	No, M-9 years		3% (1/31)	2% (1/65)
	No, M-10 years		3% (1/31)	2% (1/65)
	No, M-12 years		3% (1/31)	2% (1/65)
	No, M-14 years		3% (1/31)	2% (1/65)
	No, M-15 years		10% (3/31)	5% (3/65)
	No, M-16 years		10% (3/31)	5% (3/65)
	No, CanAccentM		3% (1/31)	2% (1/65)
Q12	Harbinese	74% (25/34)		38% (25/65)
	H-NortheastD	9% (3/34)		5% (3/65)
	Mandarin	12% (4/34)		6% (4/65)
	Harbinese M	6% (2/34)		3% (2/65)
	Cantonese		26% (8/31)	12% (8/65)
	GZ Cantonese		74% (23/31)	35% (23/65)
Q13	2hs	6% (2/34)	6% (2/31)	6% (4/65)
	3hs	12% (4/34)	13% (4/31)	12% (8/65)
	3-4hs	3% (1/34)	3% (1/31)	3% (2/65)
	4hs	6% (2/34)	10% (3/31)	8% (5/65)
	5hs	9% (3/34)	6% (2/31)	8% (5/65)
	5-6hs		3% (1/31)	2% (1/65)
	6hs	6% (2/34)	13% (4/31)	9% (6/65)
	7hs	3% (1/34)	13% (4/31)	8% (5/65)
	7-8hs	3% (1/34)		2% (1/65)
	8hs	9% (3/34)	23% (7/31)	15% (10/65)
	8-10hs		3% (1/31)	2% (1/65)
	10hs	12% (4/34)	3% (1/31)	8% (5/65)
	12hs	12% (4/34)		6% (4/65)
	14hs		3% (1/31)	2% (1/65)
	15hs	3% (1/34)		2% (1/65)
	16hs	6% (2/34)		3% (2/65)
	18hs	3% (1/34)		2% (1/65)
	24hs	9% (3/34)		5% (3/65)
Q14	0hs	82% (28/34)	3% (1/31)	45% (29/65)
	M-2hs		13% (4/31)	6% (4/65)
	M-3hs		13% (4/31)	6% (4/65)
	M-4hs		6% (2/31)	3% (2/65)
	M-5hs		13% (4/31)	6% (4/65)
	M-5-6hs		3% (1/31)	2% (1/65)
	M-6hs		23% (7/31)	9% (7/65)
	M-7hs	3% (1/34)	10% (3/31)	6% (4/65)
	M-8hs	9% (3/34)	10% (3/31)	9% (6/65)

	M-10hs	3% (1/34)	6% (2/31)	5% (3/65)
	M-12hs	3% (1/34)		2% (1/65)
Q15	A. Preschool	97% (33/34)	48% (15/31)	74% (48/65)
	Age at 1	50% (17/34)	6% (2/31)	29% (19/65)
	Age at 2	12% (4/34)	3% (1/31)	8% (5/65)
	Age at 3	24% (8/34)	10% (3/31)	17% (11/65)
	Age at 4	9% (3/34)	23% (7/31)	15% (10/65)
	Age at 5	3% (1/34)	6% (2/31)	5% (3/65)
	B. Primary School	3% (1/34)	52% (16/31)	26% (17/65)
	Age at 6		42% (13/31)	20% (13/65)
	Age at 7	3% (1/34)	10% (3/31)	6% (4/65)
Q16	A. Preschool	24% (8/34)	13% (4/31)	18% (12/65)
	Age at 3	3% (1/34)		2% (1/65)
	Age at 4	6% (2/34)	3% (1/31)	5% (3/65)
	Age at 5	12% (4/34)	10% (3/31)	9% (7/65)
	Age at 6	3% (1/34)		2% (1/65)
	B. Primary School	74% (25/34)	87% (27/31)	80% (52/65)
	Age at 6	38% (13/34)	45% (14/31)	42% (27/65)
	Age at 7	21% (7/34)		9% (7/65)
	Age at 8	6% (2/34)	6% (2/31)	6% (4/65)
	Age at 9		32% (10/31)	15% (10/65)
	Age at 10	6% (2/34)	3% (1/31)	5% (3/65)
	Age at 11	3% (1/34)		2% (1/65)
	C. Middle School	3% (1/34)		2% (1/65)
	Age at 12	3% (1/34)		2% (1/65)
Q17	7 years		3% (1/31)	2% (1/65)
	8 years	3% (1/34)	3% (1/31)	3% (2/65)
	9 years	29% (10/34)	32% (10/31)	31% (20/65)
	10 years	15% (5/34)	6% (2/31)	11% (7/65)
	11 years	26% (9/34)	19% (6/31)	23% (15/65)
	12 years	3% (1/34)	16% (5/31)	9% (6/65)
	13 years	12% (4/34)	6% (2/31)	9% (6/65)
	14 years	3% (1/34)	3% (1/31)	3% (2/65)
	15 years	9% (3/34)	10% (3/31)	9% (6/65)
Q18	No	94% (32/34)	87% (27/31)	91% (59/65)
	A little Japanese & Russian at 15	3% (1/34)		2% (1/65)
	A little Korean at 14	3% (1/34)		2% (1/65)
	French-at 11 for 1y		3% (1/31)	2% (1/65)
	German-at 13 for 2ys		3% (1/31)	2% (1/65)

	Japanese-at 15 for 3ms		3% (1/31)	2% (1/65)
	A little Spanish- at 14 for 1m		3% (1/31)	2% (1/65)
Q19	A. Yes	35% (12/34)	32% (10/34)	34% (22/65)
	B. No	65% (22/34)	68% (21/34)	66% (43/65)
Q20	Age 4-6	0	10% (1/10)	5% (1/22)
	Age 5-13	8% (1/12)	0	5% (1/22)
	Age 6-14	0	10% (1/10)	5% (1/22)
	Age 7-12	0	10% (1/10)	5% (1/22)
	Age 7-8	8% (1/12)	0	5% (1/22)
	Age at 8	17% (2/12)	0	9% (2/22)
	Age 8-14	25% (3/12)	0	14% (3/22)
	Age 9-10	0	10% (1/10)	5% (1/22)
	Age 9-12	8% (1/12)	0	5% (1/22)
	Age 9-15	0	10% (1/10)	5% (1/22)
	Age 10-11	0	10% (1/10)	5% (1/22)
	Age 10-12	0	10% (1/10)	5% (1/22)
	Age 10-13	0	10% (1/10)	5% (1/22)
	Age 11-12	0	10% (1/10)	5% (1/22)
	Age 12-15	8% (1/12)	0	5% (1/22)
	Age 12-16	17% (2/12)	0	9% (2/22)
	Age 13-14	0	10% (1/10)	5% (1/22)
	Age 13-15	8% (1/12)	0	5% (1/22)
Q21	US	42% (5/12)	40% (4/10)	41% (9/22)
Countries	UK	8% (1/12)	20% (2/10)	14% (3/22)
	CA	25% (3/12)	0	14% (3/22)
	NZ	0	10% (1/10)	5% (1/22)
	Others	0	10% (1/10)	5% (1/22)
	More than one English native country	17% (2/12)	10% (1/10)	14% (3/22)
	Forget	8% (1/12)	10% (1/10)	9% (2/22)
Duration	less than 0.5 year	25% (3/12)	0	14% (3/22)
	0.5-1 year	25% (3/12)	20% (2/10)	23% (5/22)
	1 year	0	20% (2/10)	9% (2/22)
	2 years	25% (3/12)	30% (3/10)	27% (6/22)
	3 years	17% (2/12)	0	9% (2/22)
	greater than 3 years	0	20% (2/10)	9% (2/22)
	Forget	8% (1/12)	10% (1/10)	9% (2/22)
Q22	No influence	25% (3/12)	20% (2/10)	23% (5/22)

	US	17% (2/12)	20% (2/10)	18% (4/22)
	UK	17% (2/12)	10% (1/10)	14% (3/22)
	CA	25% (3/12)	10% (1/10)	18% (4/22)
	Others	0	10% (1/10)	5% (1/22)
	Not mentioned	17% (2/12)	30% (3/10)	23% (5/22)
Q23	Null	3% (1/34)	3% (1/31)	3% (2/65)
	Class	12% (4/34)	10% (3/31)	11% (7/65)
	Class+Radio&Musics	9% (3/34)	0	5% (3/65)
	Class+Series&Films	3% (1/34)	3% (1/31)	3% (2/65)
	Radio&Musics	6% (2/34)	13% (4/31)	9% (6/65)
	Radio&Musics + Books&Newspaper	38% (13/34)	35% (11/31)	37% (24/65)
	Series&Films + Radio&Musics	12% (4/34)	29% (9/31)	20% (13/65)
	Radio&Musics + Memorising	3% (1/34)	3% (1/31)	3% (2/65)
	Books&Newspaper	12% (4/34)	0	6% (4/65)
	Radio&Musics + Books&Newspaper	3% (1/34)	3% (1/31)	3% (2/65)
Duration	0	3% (1/34)	3% (1/31)	3% (2/65)
	1h	12% (4/34)	6% (2/31)	9% (6/65)
	1.5hs	6% (2/34)	0	3% (2/65)
	2hs	24% (8/34)	19% (6/31)	22% (14/65)
	2.5hs	6% (2/34)	6% (2/31)	6% (4/65)
	3hs	26% (9/34)	19% (6/31)	23% (15/65)
	3.5hs	3% (1/34)	3% (1/31)	3% (2/65)
	4hs	9% (3/34)	16% (5/31)	12% (8/65)
	5hs	0	13% (4/31)	6% (4/65)
	6hs	0	6% (2/31)	3% (2/65)
	8hs	9% (3/34)	3% (1/31)	6% (4/65)
	10hs	3% (1/34)	0	2% (1/65)
	11hs	0	3% (1/31)	2% (1/65)
	Avg duration	3hs/w	3.6hs/w	3.3hs/w
Q24	A. British	15% (5/34)	32% (10/31)	23% (15/65)
	B. American	85% (29/34)	68% (21/31)	77% (50/65)
Q25	A. British	50% (17/34)	68% (21/31)	58% (38/65)
	B. American	50% (17/34)	32% (10/31)	42% (27/65)
Q26	A. British	15% (5/34)	52% (16/31)	32% (21/65)
	B. American	35% (12/34)	13% (4/31)	25% (16/65)
	C. Chinglish	50% (17/34)	29% (9/31)	40% (26/65)

	D. Others	0	6% (2/31)	3% (2/65)
	Less Chinglish	0	6% (2/31)	3% (2/65)
Q27	A. Yes	21% (7/34)	32% (10/31)	26% (17/65)
	B. No	79% (27/34)	68% (21/31)	74% (48/65)
Q28	Can distinguish	0	20% (2/10)	12% (2/17)
	Know a little	14% (1/7)	20% (2/10)	18% (3/17)
	Can’t distinguish	86% (6/7)	60% (6/10)	71% (12/17)