

Copyright Undertaking

This work is protected by copyright, with all rights reserved.

By reading and using the work, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the work.
2. The reader will use the work for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this work are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

Undoing a Prosodic Sound Change in Progress:
Tone Merger Reversal in Hong Kong Cantonese
via Phonetic Imitation

LUO Jin 罗锦 1310

BA (Hons) in Chinese and Bilingual Studies
中文及雙語學（榮譽）文學士

Supervisor: Dr Yao Yao 姚瑤博士

Department of Chinese and Bilingual Studies
中文及雙語學系

The Hong Kong Polytechnic University
香港理工大學

2017

Abstract

Phonetic imitation means people sound more similar to the speech they hear (Babel, 2009), and whether phonetic imitation could operate on the reversed side of sound change is important in understanding linguistic behavior in individual and community levels (Babel, McAuliffe, & Haber, 2013). This thesis uses Hong Kong Cantonese tone mergers (Mok, Zuo, & Wong, 2013) to approach this question from the less studied lexical tones. The study investigates whether young HKC speakers can unmerge T3/T6 (mid-level and low-level tones) in a phonetic imitation task, and how the behavior is conditioned by social factors, including gender, age, and language attitude.

32 young Hong Kong Cantonese native speakers (16F, 16M) joined the study. They completed a four-block speech production task, using Cantonese monosyllables in T3 or T6: reading visual stimuli (B1) -- shadowing audio stimuli with clearly distinguished T3 and T6 (B2) -- shadowing (B3) -- reading (B4). Followed were an Implicit Association Test (IAT) (Greenwald, Nosek, & Banaji, 2003) and an explicit attitude survey (Pantos & Perkins, 2012) on attitudes toward “lazy accent”.

Results showed significant T3-T6 tonal distance increase in all three blocks during and after audio exposure, which meant subjects imitated and reversed the tone merger. Males were more merged than females initially, and they imitated more and earlier than females. Language attitudes and perceived age correlated with imitation in some blocks, but were not significant in a mixed effects model. Individual differences were discovered in both production and perception through the imitation and IAT tasks.

This research is among the first studies on phonetic imitation of tones, and sociolinguistics of Cantonese tone mergers. It contributes to the variation of phonetic imitation, and the flexibility and individual differences in production and perception of sound change. Generally, the study has multifold implications in phonetics, sociolinguistics, and psycholinguistics.

Acknowledgments

Reading acknowledgments in theses has always been a fun part in the literature review process. They tell me something about the stories behind the numbers and figures, and the authors as real human beings, who expressed strikingly universal emotions regardless of the topics they worked on. I have always wondered what mine would be like, and here it is: I hope this would be as interesting as the ones I have read, and I know no matter how long this is, it would never be enough to express my sincere gratitude to all the people supporting me along the way, both in this project and my undergraduate career.

First and foremost, my supervisor Dr Yao Yao is the person who made this entire work possible. Two years ago when I was a naïve sophomore wondering what research would be like, you gave me the first chance of assisting in research and attending group meetings, and taught me the very basics of research with great patience. Ever since the start of this project, you spent countless hours guiding me through each step, pushing me to find out answers for every detail, encouraging me to present at conferences which I never thought I could, and teaching me how to do good research and how to live life to its full with yourself as an inspiring example. Thank you so much, Dr Yao. My appreciation is beyond words.

I met Dr Stephen Politzer-Ahles after the experiments were finished, and he turned out to be such an important person in shaping this thesis. You walked me through my data when I knew nothing about statistics, and were always willing to offer help whenever I needed. Much ideas in this research were implicitly influenced by your endlessly interesting psycholinguistics course. You believed in me when I did not have faith in myself, and from your constant encouragement, I

gained the strengths to move forward. You are one of the kindest person I ever know. Steve, I cannot thank you enough.

Without Dr Cheung Kwan Hin and Dr Karl Neergaard, there would have been much more stumbles in the research process. Dr Cheung had never taught me courses, but you kindly discussed the project with me over summer, went over the whole stimuli list in detail, and pointed out a fatal mistake that could have messed all the data. Many ideas on experimental design and data analysis are accredited to Karl: you listened to my puzzles over the experiment and my future, saved me from executing a very much flawed paradigm, and inspired me to always work towards high standards. Thank you both for keeping this study on the right track!

I am grateful for having my friends and research groupmates, (Drs-to-be) Bhamini Sharma, Zhou Mi, Liu Chang and Dr Sun Jiaming, with whom linguistics has always been fun. We rushed for slides before the biweekly meetings, ate lots of Indian food together, complained about the confusing mixed effects models, and encouraged each other in the long thesis writing processes. I am glad that we have gotten this far, and we are all in this together!

This project has benefitted from the support and comments of many more people. I thank Dr Roxana Fung, Dr Chen Si, Dr Katherine Chen, Professor William Shi-Yuan Wang and Akshay Maggu for stimulating my ideas, Miss Annie Tsoek and Miss Xu Ruqing for helping me on experimental design, Mr Joe Chan and Mr Tong for always keeping their doors open when I needed technical assistance, my friend Windy Lui for giving me constructive feedback after being the pilot subject, and all the experiment participants for the support during the humid summer days. It was wonderful presenting at the 21st International Conference on Yue Dialects, the University of Chicago Linguistic Research Forum, and the Joint Postgraduate

Student Symposium on Language, Culture and Cognition: comments from the audiences all made this piece of work a better one.

It is almost the end of my undergraduate journey as I write up this thesis, and I would like to express my sincere gratitude to everyone that has walked by me during these four years. To my family: thank you for the unfailing support to my dreams, however wild they might be, and I sincerely hope what I have been doing have made you proud. To my friends, especially Zhuo, Huixin, Kexin, Siyu, Sze Long and Christine: thank you for standing by me, for the long hours over lunch, phone calls connecting the north and south of China, and messages across the Pacific Ocean, when we tried to figure out together what the future holds for us. To the amazing teachers and staff at Department of Chinese and Bilingual Studies, especially Ms Valentina Chan, Dr Li Dechao, Dr Doreen Wu, Dr Angel Chan, Dr Cindy Ngai, Miss Wendy Liu, Ms Bonnie Seid and many more: you supported me not only in linguistics research, but also in other adventures that have shaped who I am today and who I want to become in the future. To professors at the Department of Computing, especially Dr Cao Yixin, Dr Vincent Ng and Dr Dennis Liu, who provided great help and understanding when I was often the only student in class struggling with the very basics: the skills and mindsets I learned from your courses are assets that I can take with me wherever I go. To other people that I had the luck to know at PolyU, including Ksenia who inspired my passion for writing again, Dr Raymond Kwok who gave me the wonderful opportunity to explore Hong Kong indigenous culture, Dr Alex Cockain who pointed the direction for me when I was obsessed with sociology, and many others that I could not possibly list all here. To all the inspiring souls I have encountered in Beijing and New York, Singapore and Russia, especially Dr Koh Keng We for providing me with the motivation to learn

Cantonese and to study migration, Nastya for our friendship despite the language barrier, and Professor Irina Sekerina for being amazingly helpful in supporting my future pursuits. I am beyond fortunate to have known you all.

Finally, this work is dedicated to you, the fragrant harbor, the Pearl of the Orient. The ability to speak Cantonese is one of the best gifts I could take with me beyond my stay in here, and working on Cantonese added much more fun to the research process. The language and the city were totally unfamiliar to me four years ago, but today, they are part of my identity, that connects to something deep-down about myself and all the people related to here. I feel so lucky to be among the people touched by the ever-shining lights across the Victoria Harbor, and to have shared these years of my life full of adventures with you. You have been accompanying me throughout my college life, and have witnessed my growth into an independent adult. You are my endless source of inspiration. Hong Kong, this one is for you.

Table of Contents

Abstract	i
Acknowledgments	iii
Table of Contents.....	vii
List of Tables and Figures	ix
Chapter 1. Introduction.....	1
Chapter 2. Background.....	3
2.1. Phonetic imitation.....	3
2.1.1. Overview	3
2.1.2. Influencing factors.....	4
2.2. Hong Kong Cantonese and the sound change	6
2.2.1. The “lazy accent” sound change.....	7
2.2.2. Tone mergers	9
2.3. Implicit Association Test (IAT)	12
2.4. Summary	14
Chapter 3. Research questions, hypotheses and predictions	15
Chapter 4. Methodology	16
4.1. Overview	16
4.2. Stage 1: Speech production task.....	17
4.2.1. Materials	17
4.2.2. Procedure	20
4.3. Stage 2: Implicit Association Test	21
4.3.1. Materials	21
4.3.2. Procedure	22

4.4. Stage 3: Explicit language attitude survey	23
4.5. Other enquiries	24
4.6. Data analysis.....	24
4.6.1. Speech production	24
4.6.2. IAT scores	26
4.6.3. Explicit attitude scores	26
4.6.4. Mixed effects modelling.....	26
Chapter 5. Results.....	28
5.1. Speech production: overall	28
5.2. Speech production: by gender	31
5.3. Speech production: by perceived age	36
5.4. Language attitudes.....	38
5.5. Speech production: by language attitudes	41
5.6. Speech production: individual differences	42
5.7. Mixed Effects Models	45
5.8. Summary	47
Chapter 6. Discussion.....	49
Chapter 7. Conclusion	53
References	55
Appendices	63
Appendix 1. Speech production task stimuli list.....	63
Appendix 2. Implicit Association Test stimuli list	65

List of Tables and Figures

Table 1: Typical features of the "lazy accent" sound change	8
Table 2: Cantonese tones with examples, from Mok et al. (2013).....	10
Table 3: Sequence of an Implicit Association Test on 2000 Election in Greenwald et al. (2003).....	13
Table 4: Speech production task procedure.....	20
Table 5: Implicit Association Test procedure	22
Table 6: Means and Standard Deviations (SDs) of averaged (T3-T6) tonal difference in blocks	28
Table 7: Results of t-tests comparing baseline block B1 with subsequent blocks B2, B3 and B4.....	29
Table 8: Means and SDs of imitation scores	30
Table 9: Means of T3 and T6 values in each block.....	30
Table 10: Means and SDs of averaged (T3-T6) tonal difference in blocks, separated by gender	32
Table 11: Results of t-tests comparing B1 with B2, B3 and B4, separated by gender	33
Table 12: Means and SDs of imitation scores, separated by gender	33
Table 13: Means of T3 and T6 values in each block, separated by gender ...	34
Table 14: Correlation among imitation scores, young age-lazy accent association, and perceived age of the model talker	37
Table 15: IAT stimuli confusion of each participant	38
Table 16: Means, SDs, maximum and minimum values of the three language attitude scores	39

Table 17: Means, SDs, maximum and minimum values of the three language attitude scores centered to -1~1	39
Table 18: Correlation of two versions of IAT scores.....	41
Table 19: T-test of two versions of IAT scores	41
Table 20: Correlation among imitation scores and language attitude scores	41
Table 21: Results of the first mixed effects model, using pitch values from each production	45
Table 22: Results of the second mixed effects model, using tonal differences in each block.....	46
Figure 1: F0 (fundamental frequency) contours of Cantonese tones uttered by a male speaker, and lexical tones of Cantonese, from Peng and Wang (2005)	10
Figure 2: Cantonese tone charts with only six unchecked tones, from Peng (2006)	11
Figure 3: Bar plot of average (T3-T6) tonal difference change across blocks	28
Figure 4: Box plot of average (T3-T6) tonal difference change across blocks	29
Figure 5: Bar plot of imitation scores.....	30
Figure 6: Bar plot of T3 and T6 mean values in each block	31
Figure 7: Bar plots of female and male tonal differences across blocks	32
Figure 8: Bar plot of imitation scores, separated by gender.....	33
Figure 9: Bar plot of female T3 and T6 values in each block	35

Figure 10: Bar plot of male T3 and T6 values in each block, together with the stimuli (model talker's production)	35
Figure 11: Opinion on young age-lazy accent association	36
Figure 12: Perceived age of the model talker	37
Figure 13: Line chart of first IAT scores and explicit attitude scores	40
Figure 14: Line chart of second IAT scores and explicit attitude scores	40
Figure 15: Female individual tonal difference change from B1 to B4	43
Figure 16: Male individual tonal difference change from B1 to B4	44

Chapter 1. Introduction

Everything in the world is always changing. Language changes all the time, and so does our speech. Language changes when the differences are agreed by a linguistic community; our speech changes whenever we speak to someone new. Humans are often fascinated by the rule, something that is not easily changed, underlying the constant variation; and the rule that counts for both individuals and communities, could probably reveal the answers to what makes us humans, and what is unique about language.

Sound change is always an interesting topic for linguists. Members in a speech community accept certain variations in progress, and agree on a new norm after the sound change is complete. This community-level process must be achieved by individuals in the group. Possible alternations in speech within a single person could take place through speaking and listening to people around. These people influence, and are also influenced by, the individual. One important phenomenon in the personal interactions is phonetic imitation.

No matter whether one realize it or not, people tend to adjust their speech patterns according to who they talk to. If they become more similar-sounding, the behavior is called phonetic imitation (Babel, 2009). When phonetic imitation meets sound change, it becomes a question whether people who have went down the path of sound change would still imitate the unchanged speech. Answers to this question will tell us about the flexibility in language systems and mental processes.

Humans communicate because of our social nature. What we talk not only reflects the message we want to send, but also reveals our identities and relationships that position us in this society. Language, therefore, is not only a

mental product related to thinking; it is also highly social in nature. Sociolinguistics and psycholinguistics are both important dots that connect us to the reality of human languages.

To know the general mechanism of language, we should not only focus on the well-studied languages like English. Other languages that have distinct characteristics than English, could provide important evidence on the overall picture. Lexical tone, for example, is a feature that English does not have but Chinese has. Almost no phonetic imitation literature had focused on tones, which inspired the current research in hope to tackle this issue.

This article explores phonetic imitation of sounds undergoing change, and the social factors mediating variation of this behavior. Using the merging tones in Hong Kong Cantonese (HKC), this study hopes to answer how flexible can the ongoing sound change be, and how socially variable is the phonetic imitation process. Phonetic imitation papers had never focused tones, and HKC tone mergers were not tackled from sociolinguistics or phonetic imitation perspectives. This research expects to fill the gaps at the intersection of disciplines, and to provide multifold implications on the social and psychological mechanisms in language change.

The following parts of this thesis will be organized as follows. Chapter 2 reviews three bodies of literature: phonetic imitation, HKC sound change, and Implicit Association Test (IAT). Research questions, hypotheses and predictions are given in Chapter 3. Chapter 4 introduces the methodology of this study that includes speech production task, an IAT and a language attitude survey. Chapter 5 presents all results from acoustic analyses and the attitude tests. Finally, Chapters 6 and 7 give discussion and conclusion respectively.

Chapter 2. Background

2.1. Phonetic imitation

2.1.1. Overview

Phonetic imitation is the process in which people sound more similar to each other during conversational interactions (Babel, 2009). This phenomenon is also known as phonetic convergence or phonetic accommodation.

Having been extensively documented, it was discovered that phonetic imitation could take place in both social interactions (Pardo, 2006) and asocial laboratory environments (Namy, Nygaard, & Sauerteig, 2002; Babel, 2012). Sound features that had been imitated include vowel (Babel, 2012), fundamental frequency (F0) (Babel & Bulatov, 2011), voice onset time (VOT) (Abrego-Collier, Grove, Sonderegger, & Yu, 2011), duration and speaking rate (Pardo, 2013).

Phonetic imitation has implications in multiple disciplines. The phenomenon accounts for inter-person speech variability, social attitudes, and cognitive mechanisms of imitation (Babel, 2009). In broader perspectives, phonetic imitation helps to explain diachronic sound change through synchronic speech variation (Babel et al., 2013), and provides evidence to the human imitation behavior in general (Babel, 2009). Though Trudgill (2008) argued that phonetic imitation is a purely automatic behavior, it was found to be phonetically and socially selective by Babel (2009). On both sociolinguistic and phonetic sides, there have been numerous studies exploring the factors that modulate this behavior.

2.1.2. Influencing factors

The basis for phonetic imitation to take place is the initial acoustic mismatch, or interlocutor language distance, between people in speech interactions (Babel, 2009; Kim, Horton, & Bradlow, 2011). Though it is likely for speech from different individuals to be different from each other, it was discovered that imitation could only take place when there is phonetic space to do so (Babel, 2009; Babel, 2012). Main factors affecting the extent of imitation could be categorized into phonetic and social ones.

Phonetic properties that could influence imitation include acoustic characteristics and usage frequency. Though the results were different, Babel (2009) and Pardo (2010) found selective imitation towards certain vowels. Moreover, Goldinger (1998) found that low frequency words are more likely to be imitated, which resulted in the choice of infrequent words in many following phonetic imitation studies (Pardo et al., 2013).

Social dimensions are highly complicated in phonetic imitation. Generally, phonetic imitation could be interpreted as a device to manage social distance (Babel et al., 2013), which accounts for many aspects in social interactions. Findings from previous literature could be categorized into the following parts:

- (1) Characteristics of the person being imitated, or the model talker in laboratory experiments. Both information given to the imitators and subjective perception towards the talker by the imitator were found to have effects. Clear traits of the talker that have been discussed in previous research include gender (Namy et al., 2002), nationality (Babel, 2010), race and relevant facial priming (Babel, 2009; Babel & Russell, 2015), and

positive or negative conditions assigned to model talker's background (Babel et al., 2013). Expectations and subjective attitudes on the talker that affected imitation include perceived sexual orientation of the talker (Abrego-Collier et al., 2011), subjective attitude towards the talker (Yu, Abrego-Collier, Baglini, Grano, & Martinovic, 2011), and voice attractiveness judgments (Babel, McGuire, King, & Satterwhite, 2011).

(2) Characteristics of the imitator themselves. Gender (Namy et al., 2002) and personality traits of openness and autistic-like traits (Yu, Abrego-Collier, & Sonderegger, 2013) were found to be associated with phonetic imitation. This aspect was less explored compared to talker variability.

(3) Conversational context and interpersonal relationship. In experiments of conversational interactions, participants were often assigned different roles, and it was found that roles influence imitation and its direction (Pardo, 2006; Pardo, Hoshino, & Sowemimo-Coker, 2013). Moreover, strength of relationship was discovered to modulate the process between people that know each other in real life (Pardo, Gibbons, Suppes, & Krauss, 2012).

Despite the flourishing of phonetic imitation literature, lexical tones had not been included in the discussion. As a language that is known for its tonal complexity and has been undergoing sound change in tones, Hong Kong Cantonese (HKC) can provide the field with new possibilities in prosodic features. The next section introduces the background of HKC, its recent sound change and its tonal system in detail.

2.2. Hong Kong Cantonese and the sound change

Hong Kong Cantonese (HKC) is the variety of Cantonese spoken in Hong Kong. 88.1% of Hong Kong's population aged between 6 to 65 speak Cantonese as their first language (Census and Statistics Department, Hong Kong Special Administrative Region, 2016). Cantonese belongs to the Yue Dialect group of Chinese, and is spoken mostly in southern China and overseas Chinese communities (Chen, 1998). Cantonese is not mutually intelligible with Mandarin, the Chinese dialect with most speakers and one of the official languages of Hong Kong since 1997.

Though located near Dongguan and Bao'an (now Shenzhen), HKC was derived from Cantonese spoken in Guangzhou (previously known as Canton), capital city of Guangdong Province in south China, despite the geographical distance. Due to the migration waves predominantly from areas near Guangzhou in mid-20th Century and the social unrest in 1960s, the British colonial government of Hong Kong promoted the use of Guangzhou-based Cantonese, the language spoken by the majority of its population, and united people originally from various dialect groups under this language (Ding, 2010; Bolton, 2011). Studies in 1960s found that Hong Kong and Guangzhou Cantonese share the same phonological system, but the two varieties started to diverge in 1990s (Cheung, 2002). Sound changes in HKC in the last two decades are particularly prevalent among young people, and this new style of speech is often referred to as "lazy accent", or *laan jam* in Cantonese (Cheung, 2002).

2.2.1. The “lazy accent” sound change

“Lazy accent” is the colloquial term frequently used in media and public discourse to describe the sound changes in HKC from around 1990s. Scholars have given it other names in academic language, such as “innovative” and “colloquial” forms (Chen, 1998), reflecting the nature of this newly emerged variety: young-dominant, and mostly spoken in informal contexts. The concept was established in contrast to the “correct pronunciation”, which was actively promoted by professor and television personality Ho Man-Wui (Ho, 1989). Originated from its phonetic and sociolinguistic features, this term has rich meaning in both aspects.

Phonetically, sound change has been going on in initials, finals and tones in HKC syllables, the general trend of which is reduction in the numbers of phonological categories in all three aspects (Cheung, 2002). Evidence from diachronic and synchronic research supported the variations in specific segments (To, Mcleod, & Cheung, 2015). Chen (1998) presented two types of variation: colloquial-citation and reading pronunciation variations, the first of which has the same meaning as “lazy accent” is used for current discussion. Table 1 provides a summary of sound changes from Chen (1998), Cheung (2002), Ding (2010) and To et al. (2015).

Type	“Correct”	“Lazy”
Initial	[n]	[l]
	[ŋ]	Ø
	[k ^w]	[k]
	[ŋ]	[m]
Final	[ŋ]	[n]
	[k]	[t]
	[m]	[n]
Tone	High rising (35) and low rising (13) tones	High rising (35) tone only

Table 1: Typical features of the "lazy accent" sound change

Sociolinguistically, negative reading can be implied from its name “lazy accent” already, which infers the origin of this accent to be “not making enough efforts in pronunciation”. The Cantonese pronunciation before 1990s, or even earlier to the norms described in books written hundreds of years ago, was constructed as “correct pronunciation”, in opposition to “lazy accent” which symbolized deviation from the standard (Chen, 1998). This linguistic ideology of Cantonese standardization was mainly promoted by Ho Man Wui through media campaigns, and had triggered numerous debates for and against this idea. In the qualitative research conducted by Chen (1998), it was found that older people frequently complained about the “lazy accent” in young people, teachers and broadcasters, and the whole movement of Cantonese standardization was closely related to the anxiety in Hong Kong’s identity and the position of Cantonese after the reunification with mainland China.

The contact of Cantonese with other languages, including Mandarin, English and other Chinese dialects, was said to be one of the main reasons that caused this ongoing sound change (Chen, 1998; Ding, 2010; To et al., 2015). Language policy of the current Hong Kong government is known as “trilingualism and biliteracy”, that calls for Hong Kong people’s proficiency in oral Cantonese, Mandarin and English, as well as written Chinese and English (Bolton, 2011). Besides extensive language contacts, phonetic factors were also projected as one factor of the sound change, that speakers conserved articulation efforts and thus assimilated sounds together (To et al., 2015). Though studies on “lazy accent” mainly paid attention to initials and finals in HKC syllables (Cheung, 2002; Ding, 2010; To et al., 2015), an increasing body of research has been studying the assimilation of tone types, which is usually referred to as HKC tone mergers.

2.2.2. Tone mergers

HKC has one of the most complicated tonal inventories among Chinese languages, consisting of 6 unchecked and 3 checked lexical tones in 6 tonal contours. Table 2 from Mok, Zuo and Wong (2013) and Figure 1 from Peng and Wang (2005) give an overview on the tonal system. From here on, T1-T9 will be used to refer to the specific tones in short.

Register	Lexical tones			Allotones
High	T1 high-level [ji55] ‘clothes’	T2 high-rising [ji25] ‘chair’	T3 mid-level [ji33] ‘idea’	T7 high-stopped [jik5] ‘benefit’ T8 mid-stopped [jak3] ‘eat’
Low	T4 low-falling [ji21] ‘suspicious’	T5 low-rising [ji23] ‘ear’	T6 low-level [ji22] ‘two’	T9 low-stopped [jik2] ‘wing’

Table 2: Cantonese tones with examples, from Mok et al. (2013)

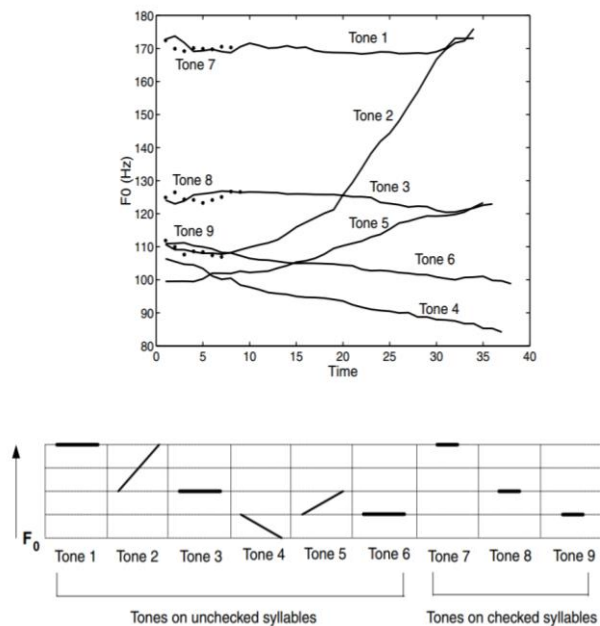


Figure 1: F0 (fundamental frequency) contours of Cantonese tones uttered by a male speaker,
and lexical tones of Cantonese, from Peng and Wang (2005)

Tones in this highly sophisticated system have been discovered to be merging in both perception and production. One of the earliest reviews on HKC tonal changes is So (1996), which discussed the variation in T1, the reported merger of T3 and T5, and the tonal contours of two rising tones, T2 and T5. In the 21st century, HKC tone mergers have attracted a lot of research attention. Most studies focused on the six unchecked tones that have almost the same time durations (Peng, 2006), as the three checked tones still fall into the levels of unchecked tones T1, T3 and

T6 respectively, which the current study follows. According to the analysis of Peng (2006) measuring the height and slope of tones, two rising tones T2/T5, and two level tones T3/T6 seem to be merging, as illustrated by Figure 2.

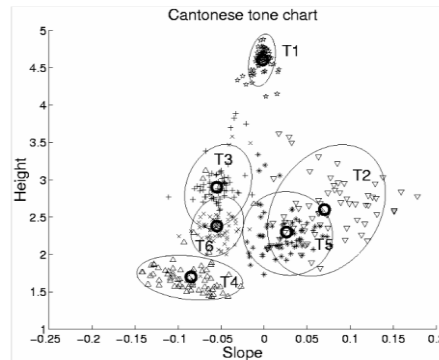


Figure 2: Cantonese tone charts with only six unchecked tones, from Peng (2006)

Researchers have not yet reached consensus on which tones are merging. The most explored pair is T2/T5, the high and low rising tones, as discussed in Bauer, Cheung and Cheung (2003) and Fung and Wong (2011). T3/T6 merger has also been observed in multiple studies (Peng, 2006; Mok, Zuo & Wong, 2013). Other pairs that were said to have been merging include T4/T6 (Fung, Kung, Law, Su, & Wong, 2012) and T3/T5 (Fung & Wong, 2010), though some studies did not discover the merging of those pairs (Peng, 2006).

Previous literature on HKC tone mergers was mainly in phonetic (Mok et al, 2013) and psycholinguistic perspectives (Fung et al, 2012), but sociolinguistic studies were far from sufficient. The only sociolinguistic study found on Cantonese tone mergers is Luo and Durvasula (2014) on T2/T5 merger in Guangzhou Cantonese. Investigating the social tendency of this sound change, young people predominantly merged the two rising tones rather than older people. Subjects from

middle social class merged less than others; no gender difference was discovered in the merging behavior (Luo & Durvasula, 2014).

With interests in extending the sociolinguistic discussions in this phenomenon, Implicit Association Test (IAT), which is a paradigm in social psychology, is used in this study (Greenwald, McGhee, & Schwartz, 1998). IAT has been frequently put into practice in phonetic imitation studies, and the next section would explore this method in detail.

2.3. Implicit Association Test (IAT)

As its name suggests, Implicit Association Test (IAT) was designed to measure the implicit associations, attitudes, or biases of people. The method was first proposed by Greenwald et al. (1998), and has been frequently used in social psychology research, as well as other fields in academia. Each IAT tests human association of two target concepts with two attributes through the following steps: first, assign one left key and one right key on the keyboard for subjects to respond, and ask subjects to categorize the stimuli into one of the two targets; second, ask subjects to categorize the stimuli into one of the two attributes; third, combine targets and attributes, that one target and one attribute shares one key for response, and subjects classify both attributes and targets in this step; fourth, do targets' categorization only, but reverse the left-right sequence of that in step one; fifth, combine targets and attributes again, following the key assignment of targets in step four. Participants' implicit attitudes are calculated from the different reaction times in two combined blocks, three and five. Greenwald et al. (1998) tested the validity

of this method with universally associated concepts, personal preferences, and social biases that people sometimes deny.

The IAT algorithm was improved by Greenwald, McGhee, and Schwartz (2003) with a 7-block experimental sequence (Table 3), D-transform data treatment method, and exclusion criteria of trials and respondents. They suggested subjects that make too many errors or are too fast/slow in their trials, and the individual trials that are too fast or too slow, should be excluded from the dataset (Greenwald et al., 2003). For IAT designs that allow error correction, latency for the second trial can be treated as an error penalty, and can be added to the latency in the first trial as the overall reaction time for the stimulus (Greenwald et al., 2003).

Block	No. of trials	Function	Items assigned to left-key response	Items assigned to right-key response
1	20	Practice	George Bush images	Al Gore images
2	20	Practice	Pleasant words	Unpleasant words
3	20	Practice	Pleasant words + Bush items	Unpleasant words + Gore items
4	40	Test	Pleasant words + Bush items	Unpleasant words + Gore items
5	20	Practice	Al Gore images	George Bush images
6	20	Practice	Pleasant words + Gore images	Unpleasant words + Bush images
7	40	Test	Pleasant words + Gore images	Unpleasant words + Bush images

Table 3: Sequence of an Implicit Association Test on 2000 Election in Greenwald et al. (2003)

With rising popularity of the method, numerous studies have been published on the factors influencing IAT results (Cui & Zhang, 2004). One of the most observed one is the sequence of congruent and incongruent combinations: whether the first target-attribute combination aligns with the subjects' implicit attitude affects nearly all IAT studies. Ways to reduce this effect include counterbalancing (Greenwald et al., 2003), and doubling the number of trials for the reversed target block (Nosek, Greenwald, & Banaji, 2005). Scholars have also argued on the influence from word properties and subject individual differences (Rothermund &

Wentura, 2001). Factors that had no experimental effect on IAT data include the number of target stimuli (Cui & Zhang, 2004), specific left and right keys assigned (Greenwald et al., 1998), and error feedback (Karpinski & Hilton, 2001). Prevalence of the IAT paradigm has also lead to arguments challenging the method, including the argument on whether the results could really reflect implicit attitudes of experiment participants themselves, or rather the social values that subjects were frequently exposed to and got more familiar with (Cui & Zhang, 2004).

IAT has been applied to research in various topics. The method has generally been proven useful, by studies on attitudes toward migration policies (Perez, 2010), race (Greenwald, Banaji, & Nosek, 2015), and language varieties (Pantos & Perkins, 2012).

2.4. Summary

The phonetic and sociolinguistic dynamics of HKC tonal mergers-in-progress and the “lazy accent” sound change provide abundant research opportunities. Though the three bodies of literature reviewed in this chapter are rich by themselves, works at the intersection of phonetic imitation, sound change, and social attitudes are very limited to date (Babel et al., 2013; Pantos & Perkins, 2012); research on the unique merger reversal phenomenon in sound change mostly focused on vowels (Babel et al., 2013; Yao & Chang, 2016), and lexical tones have not received much attention. This study aims to fill in the current gap, and the research details are described in the following chapters.

Chapter 3. Research questions, hypotheses and predictions

The central research question of this thesis is whether lexical tones can be phonetically imitated, and whether tonal mergers-in-progress can be reversed in speech accommodation. Two merging tones in Hong Kong Cantonese, the mid-level tone (T3) and low-level tone (T6), are being investigated. This study is also interested in how imitation is influenced by imitator variability. The hypothesis is that phonetic imitation does operate on lexical tone mergers, with variations by social factors. If the hypothesis is true, participants' pitch difference between T3 and T6 tokens will increase after exposure to an unmerged model talker, and the extent of imitation will vary according to subject's gender, language attitude, and perceived age of the model talker.

Chapter 4. Methodology

4.1. Overview

Subjects were voluntarily recruited through flyers around The Hong Kong Polytechnic University campus, and filled in a language background survey before being invited to the experiment. 32 subjects joined the study, half were male and half were female. All of them were born and raised in Hong Kong, and identified Cantonese as their sole native language and the only language spoken at home. Most subjects reported having never lived outside Hong Kong for more than one year, except two male subjects: one lived in the United Kingdom from 16 to 20, the other lived in Guangdong Province, China from 0 to 2. Participants were either students or staff at the university, and were all aged between 18 to 25. No speech or hearing disorders were reported.

The whole experiment had three stages, and needed 35 minutes to complete. All participants completed the three stages. They sat alone inside a sound-proof booth of the linguistics laboratory at Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University for the whole period, in front of a monitor controlled by a Lenovo ThinkCenter MT-M 10A8-A1FGHC Desktop computer. The experimenter was outside the booth but inside the laboratory, and monitored the participant through a camera placed inside the booth, to make sure the experiment has been properly done. Oral instructions given by the experimenter were all in Cantonese. All subjects received 50 Hong Kong Dollars cash as compensation.

The study was approved by Human Subjects Ethics Application Review System at The Hong Kong Polytechnic University.

4.2. Stage 1: Speech production task

4.2.1. Materials

The critical stimuli were 24 Cantonese monosyllables, balanced in tone category (T3 or T6), word frequency (high or low), and phonotactic structure of the syllable (full set – the syllable can be combined with all 6 tones in Cantonese; or semi-full set – the syllable can be combined with both T3 and T6, but not with all the other 4 tones). Frequency thresholds were decided as more than 5,500, and less than 700 occurrences out of 138,000,000 characters in *A Chinese Character Database: With Word-formations phonologically disambiguated according to the Cantonese Dialect* (Research Centre for the Humanities Computing, Chinese University of Hong Kong, 2003), referring to the standards of Kuo et al. (2013). Some characters slightly exceeding the frequency threshold were used because of constraints in other aspects that will be explained later in this section. This design followed the studies of Mok and Wong (2010), Fung and Wong (2010) and Fung and Wong (2011).

Fillers were 24 monosyllables from tones other than T3 and T6, and were randomly mixed with critical stimuli during the experiment. They were balanced in tone categories (T1, T2, T4 or T5) and frequency (high or low). Lists of the critical stimuli and fillers can be found in Appendix 1.

All 48 monosyllabic characters were embedded in a same carrier sentence:

<i>Chinese</i>	而	家	講	—	字
<i>IPA</i>	ji21	ka55	kəŋ25	—	tsi22
<i>Jyutping</i>	ji4	ga1	gong2	—	zi6
<i>English</i>	Now	reading	—	character	

The carrier sentence was adapted from that in Yu (2007) and Mok et al. (2013). As T3 and T6 are both level tones that differ only in pitch level but not contour, perception of the two categories is related to the prosodic context, which made it particularly important to embed stimuli into a carrier sentence.

Selection of the stimuli was done through exhaustion of all Cantonese syllables in the 138,000,000-character database introduced earlier in this section (Research Centre for the Humanities Computing, Chinese University of Hong Kong, 2003). The following steps were performed in choosing the critical stimuli:

- (1) Find all syllables without -p, -t or -k endings (to avoid checked tones that are shorter in duration) in Cantonese, that have legal word formations with both T3 and T6.
- (2) Check whether any T3 or T6 character has only one possible pronunciation. If none of the characters in either T3 or T6 does, exclude that syllable. Attention was paid to tones in particular; it was made sure that only one tone could be associated with one character.
- (3) Make sure the T3 and T6 are recognizable characters. The list was checked by a Hong Kong Cantonese native speaker.
- (4) Check whether high and low frequency words are both available, in both T3 and T6. Frequency thresholds were introduced earlier in this section.

(5) Exclude characters similar in shape, and syllables overlapping in initials or finals, in order to avoid unnecessary priming. Characters fulfilling the frequency requirement but having phonological or orthographical similarities with other characters were replaced by those that are unique in shape and sound with slight sacrifice on the frequency requirement.

Fillers were chosen in similar manners as critical items: each syllable has both high and low frequency monophonic characters that are recognizable in one tone category. Syllables chosen for fillers were presented in one or two tonal categories, and syllable types in tone distribution were not taken into consideration.

The model talker is a male native Hong Kong Cantonese speaker in his sixties. The recording was done through Audacity 2.1.0 (Audacity Team, 2015) through a German Maestro GMH C 8.100D headset in a sound-treated booth. Each stimulus was embedded in the carrier sentence, and each full sentence was read by the model talker for five times. Productions with more stable level pitch contour, and the ones with clear distinction of T3 and T6 were selected (average T3-T6 pitch difference = 18 Hz). Before programming the final stimulus files in the experiment, background noises before and after the sentences were removed, and chosen productions were averaged in time duration and intensity through Praat (Boersma & Weenink, 2016). This audio recording process resulted in 24 productions, and each production is mapped to a pair of homophonic characters visually, one in high frequency and one in low frequency.

4.2.2. Procedure

Participants wore a JVC Stereo Headphone HA-D610 and a DPA Microphone d:fine 88 Dual-Ear Directional Headset, and had their voices recorded by a Marantz Solid State Recorder PMD661 MKII. The experiment was programmed in OpenSesame (Mathôt et al, 2012). The 4-block design followed that of Babel et al. (2013). Details see Table 4.

Block	Instructions	Stimuli	Number of trials	Steps and timing for each trial
B1: Baseline	Read out the sentence on the screen in the clearest and most natural way possible	Visual only (48 stimuli in characters embedded in the carrier sentence)	2 practice trials randomly selected from the 24 fillers + 48*1 fully randomized trials = 50	Fixation dot: 500 ms Stimuli: 4,500 ms
B2: Shadowing-1	Repeat the sentence heard in the clearest and most natural way possible (without explicit instruction to imitate)	Audio only (24 stimuli embedded in the carrier sentence in .wav sound file read by the model talker)	2 practice trials randomly selected from the 24 fillers + 24*2 fully randomized trials = 50	Fixation dot: 500 ms Headset picture: 100 ms Headset picture + playing the stimuli: 2,500 ms Picture of a microphone: 4,500 ms
B3: Shadowing-2	Same as B2	Same as B2	50, same as B2	Same as B2
B4: Post-exposure	Same as B1	Same as B1	50, same as B1	Same as B1

Table 4: Speech production task procedure

Pace of the experiment was controlled by the program according to the timing details in Table 4, and participants could not monitor the process by keyboard or mouse. They were given optional chances to take short breaks between each block, and all of them had a brief rest before moving on to Stage 2.

4.3. Stage 2: Implicit Association Test

4.3.1. Materials

Following the improved IAT algorithm by Greenwald et al. (2003), two sets of stimuli were picked: targets – “lazy accent” and “correct pronunciation”, and attributes – positive and negative. As “lazy accent” is very much a concept in sound rather than text, audio stimuli of “correct” and “lazy” accents were used with visual attribute words as attributes. Similar design of combining audio targets and visual attributes was used by Pantos and Perkins (2012).

Targets in this experiment were ten productions of five disyllabic words, each word pronounced in both “correct” and “lazy” ways. The words were chosen from examples in each of the five categories of typical “lazy accent” listed in an educational online podcast *Speak Proper Cantonese* (Radio Television Hong Kong, 2008), as the ones chosen by a public education program are expected to be the most common. All five categories have appeared in academic literature (Chen, 1998; Cheung, 2002; Ding, 2010). To eliminate the potential confounding factors, only the first syllables of the five disyllabic words had “correct”/“lazy” versions, and on only one phoneme for each.

Attributes, which are positive and negative words shown in texts for this experiment, consists of 10 disyllabic words, half positive and half negative. The number of attributes match with that of the audio stimuli. Words were chosen from Yuen, Chan, Lai, Kwong and T'sou (2004), because their corpus included written Chinese data from Hong Kong newspapers, which fits linguistic background of participants. Lists of targets and attributes are presented in Appendix 2.

A 30-year-old female native speaker of Cantonese, who was born and raised in Hong Kong and had language teaching background, produced the audio target stimuli. She was instructed to read the words in the most “correct” and “lazy” ways she could do, while preserving the naturalness of speech. Recording was done in a similar manner as that of the speech production task.

4.3.2. Procedure

This IAT was designed according to Greenwald et al. (2003) and Pantos and Perkins (2012). Details see Table 5.

Block	No. of trials	Function	Left key – “E”	Right key – “I”
B1	20	Attribute Practice	Positive	Negative
B2	20	Target Practice	Correct	Lazy
B3	20	Combined Practice	Positive Correct	+ Negative + Lazy
B4	40	Combined Test	Positive Correct	+ Negative + Lazy
B5	40	Target Practice	Lazy	Correct
B6	20	Combined Practice	Positive + Lazy	Negative Correct
B7	40	Combined Test	Positive + Lazy	Negative Correct

Table 5: Implicit Association Test procedure

The experiment was built by OpenSesame (Mathôt et al, 2012). Participants wore the same headphone as stage 1, and sat in front of the same computer. They were instructed to respond as correctly and quickly as possible. One chance was given for correction immediately following the error trial. As there were only two possible choices for each trial, no further chances would be given if the second trial is wrong again. Participants were given chances to take short skippable breaks between each block. To minimize the effect of appearing sequence of congruent and incongruent combinations, trial number of B5 was increased to 40 (Nosek et al., 2005). In blocks 3, 4, 6 and 7 where targets and attributes both appeared, audio and visual stimuli appeared alternately to avoid the possible effect on reaction times when subjects do not know what to expect next. Stimuli were pseudo-randomized by Mix (van Casteren & Davis, 2006) before being manually modified in the OpenSesame program.

4.4. Stage 3: Explicit language attitude survey

In this last stage, subjects were asked to complete a questionnaire that explicitly asked their attitudes toward “correct” and “lazy” accents on a 1-7 scale. The survey was conducted through Google Form. Some questions were adopted from Chen (1998), and others came from frequently discussed aspects in the public. The scaled scores did not point to the same direction across questions: 6 of them had higher scores associated with positive attitudes towards “lazy accent”, and 8 of them were in the opposite direction. In addition to the 14 attitudinal items, one question was on the extent to which they agree with the statement that “younger people obviously

speak in lazy accent more than older people”. In the end, participants were asked about whether they had received formal training on Cantonese pronunciation, including learning Jyutping (Cheung et al., 1993).

4.5. Other enquiries

Between stages 1 and 2, participants were asked about how old they thought the model talker was, and whether they encountered any unrecognizable characters. Subjects gave a 10-year-old range on the perceived age, and the upper threshold was coded and later put in the statistical model. Possible values of the perceived age include 30, 40, 50 and 60.

Between stages 2 and 3, subjects were enquired on whether they had any difficulty distinguishing the “correct” and “lazy” productions. These questions were asked directly in person, the answers of which were manually recorded by the experimenter.

4.6. Data analysis

4.6.1. Speech production

Phonetic imitation is quantified through acoustic measurements in this study, and perceptual analysis was not conducted due to time and funding constraints. From the original sound recording, pitch values were obtained following the steps below:

- (1) Segment each production of the whole sentence from the single sound file for each participant that is around 20 minutes long, and save each segment as a separate sound file in .wav. Only critical stimuli went through this step, and fillers were skipped. Mispronounced productions were excluded from further analyses.
- (2) Annotate the tone-bearing vowel area of critical monosyllabic stimulus using Praat (Boersma & Weenink, 2016).
- (3) Extract the average pitch from 20%-80% of the vowel interval, with pitch floors set to 40 Hz for males and 75 Hz for females, using a Praat script adopted from Lennes (2011). Reason for not using the whole duration of the vowel interval is to minimize influence from neighboring consonants. As extra low pitch levels (lower than 75 Hz for male and 100 Hz for female, which are the default pitch floors) were found in participant speech, the 40 Hz and 75 Hz floors were set to include those low-pitch productions.

From data for each trial, one averaged pitch value was calculated from the productions of one tonal category, in one trial of one person. If there is no mispronunciation, this value was averaged from 12 data points. There are 8 such values for one participant.

Tonal difference scores, obtained from subtracting T3 and T6 average pitch values, represent the degrees of T3/T6 merger. There is one score per person per block. In total, there are $32 \times 4 = 128$ values.

Imitation scores are the differences between one shadowing or post-exposure block, and the baseline block. A positive imitation score means imitation, because it shows an increase in T3/T6 difference. One subject has three imitation scores (B2-B1, B3-B1, B4-B1).

4.6.2. IAT scores

IAT scores were calculated in SPSS, following D measurement in Greenwald et al. (2003), with the D1 (built-in error penalty to record latency of error trials) syntax of Greenwald (2007). IAT scores range from -2 to 2, where 2 means being most positive towards “correct pronunciation” and -2 means being most positive towards “lazy accent”. The scores were later centered on a -1 to 1 scale to compare with the explicit scores.

4.6.3. Explicit attitude scores

The one explicit attitude per person is on 1-7 scale, where 7 means most positive towards “correct pronunciation” and 1 means being most positive towards “lazy accent”. Before being averaged to one score, the 6 questions that had higher scores associated with positive attitudes towards “lazy accent” were transformed to the opposite direction. The scores were later centered on a -1 to 1 scale to compare with the implicit scores.

4.6.4. Mixed effects modelling

Two linear mixed effects models were constructed in R (R Core Team, 2016), using packages “lme4” (Bates, Maechler, Bolker, & Walker, 2015), “lmerTest” (Kuznetsova, Brockhoff, & Christensen, 2016) and BOBYQA algorithm for optimization (Powell, 2009).

The first model used raw pitch data points from each trial. Out of 3072 trials, 26 were mispronounced and therefore resulted in 3046 data points. The dependent

variable is the pitch value. Independent variables are block (block 1 is the intercept, and there is one variable for each subsequent block that are dummy coded), tone category (T3 or T6, deviation coded), block and tone interactions, and subject gender (male or female, deviation coded). Random variables are by subject and by item, and random slopes are set as 1.

The second model used the 128 tonal difference scores as the dependent variable. Independent variables are block, gender, implicit language attitude, explicit language attitude, one score (1-7) that is subject's answer to "do you agree that lazy accent is predominantly spoken by young people", and subject's perceived age of the model talker. The random variable is by subject, and the random slope is 1.

Chapter 5. Results

5.1. Speech production: overall

Generally, the (T3-T6) tonal difference was increased during and after exposure to the unmerged model talker. Table 6 shows the means and standard deviations (SDs) of averaged (T3-T6) in each block, from which an increasing trend can be discovered. Figure 3 and Figure 4 visualize the data through bar plot and box plot. The error bars are standard errors.

	B1	B2	B3	B4
Mean (Hz)	6.576	8.666	8.669	9.017
SD	4.190	4.304	3.707	4.429

Table 6: Means and Standard Deviations (SDs) of averaged (T3-T6) tonal difference in blocks

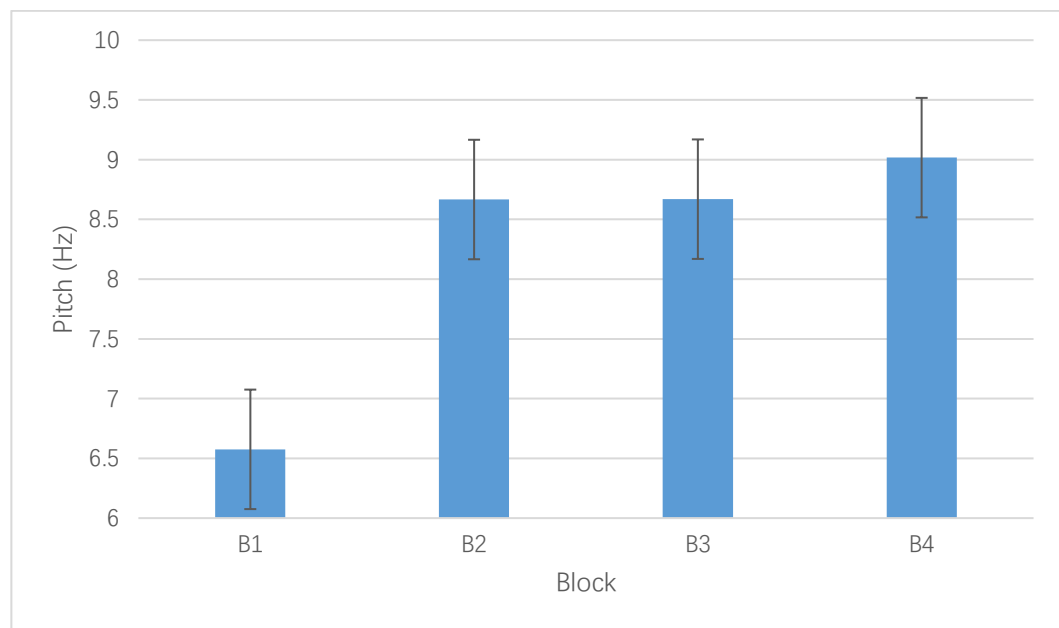


Figure 3: Bar plot of average (T3-T6) tonal difference change across blocks

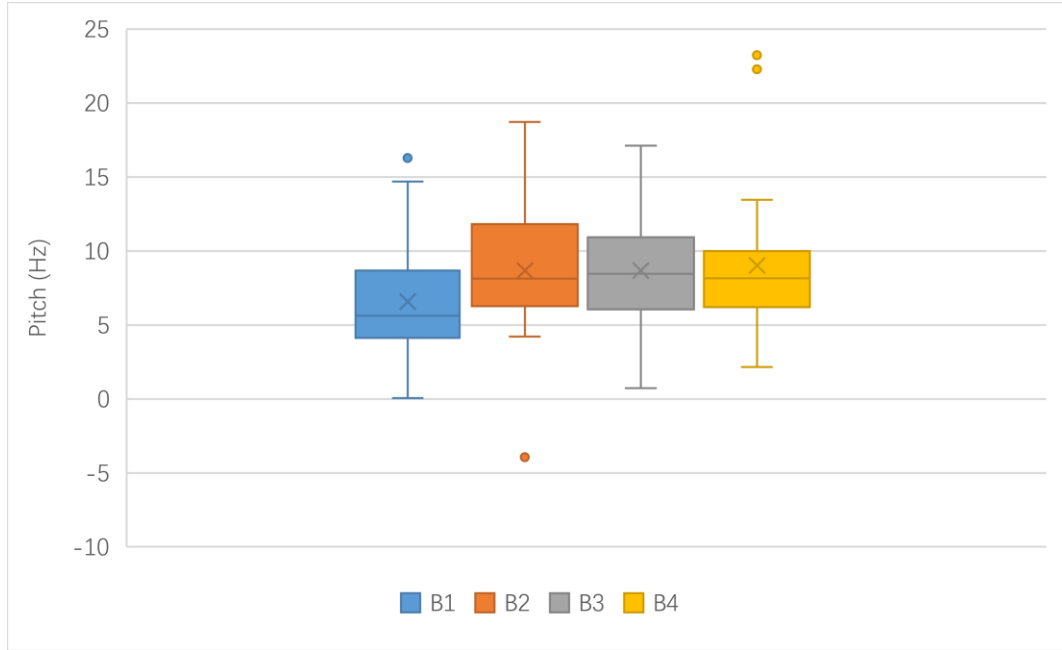


Figure 4: Box plot of average (T3-T6) tonal difference change across blocks

Paired t-tests were performed to compare the tonal difference scores in baseline and subsequent blocks. The values in B2, B3 and B4 were all significantly larger than the values in B1. Table 7 shows the t and p statistical values.

	B1/B2	B1/B3	B1/B4
t	-2.589	-2.835	-3.603
p (one-tailed)	.007	.003	.0005
p (two-tailed)	.014	.007	.001

Table 7: Results of t-tests comparing baseline block B1 with subsequent blocks B2, B3 and B4

The imitation scores (3 per person) are from subtracting the baseline tonal difference from subsequent blocks. If the subject imitated at a certain block, the corresponding imitation score is larger than 0. Table 8 and Figure 5 describe the imitation scores, further confirming that imitation had taken place.

	B2-B1	B3-B1	B4-B1
Mean (Hz)	2.091	2.094	2.441
SD	4.567	4.177	3.832

Table 8: Means and SDs of imitation scores

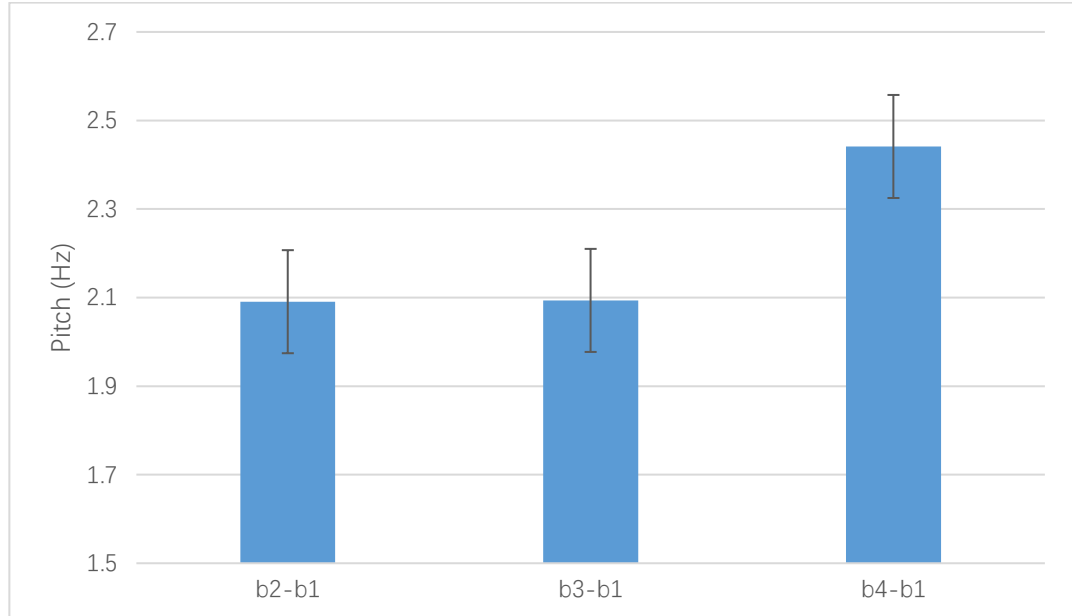


Figure 5: Bar plot of imitation scores

By separately plotting averaged T3 and T6 values, it was found that the increase in T3-T6 tonal distance mainly came from height rising of T3. Although T6 had also increased, the change was not as big as T3. Table 9 and Figure 6 show the change of T3 and T6 average values across blocks.

	B1T3	B1T6	B2T3	B2T6	B3T3	B3T6	B4T3	B4T6
Mean	145.52	138.94	150.62	141.95	152.24	143.57	150.13	141.12
n	2	7	4	7	6	7	8	1
(Hz)								

Table 9: Means of T3 and T6 values in each block

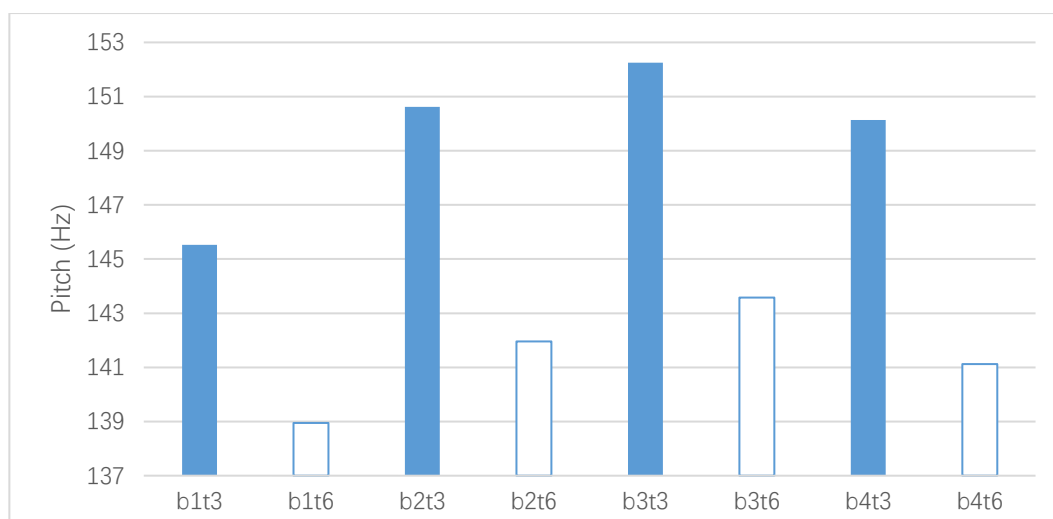


Figure 6: Bar plot of T3 and T6 mean values in each block

From data obtained from all subjects, it was discovered that the tonal difference of T3 and T6 significantly increased during and after shadowing the model talker, though the increase was not big in number and still had a distance from the model talker (18 Hz difference). Tonal distance became even larger in the post exposure block, when subjects were not directly influenced by the model talker. Though T3 and T6 both raised their pitch heights, T3 raised more and mainly contributed to the tonal distance increase.

5.2. Speech production: by gender

Male and female participants exhibited very different baseline production and imitation patterns. Table 10 show the means and standard deviations of females and males, and Figure 7 show the averaged trend of change.

	B1	B2	B3	B4
Mean (Hz) - female	8.570	8.842	9.556	10.333
SD - female	4.080	5.349	4.348	5.406
Mean (Hz) - male	4.581	8.490	7.782	7.701
SD - male	3.339	3.100	2.796	2.759

Table 10: Means and SDs of averaged (T3-T6) tonal difference in blocks, separated by gender

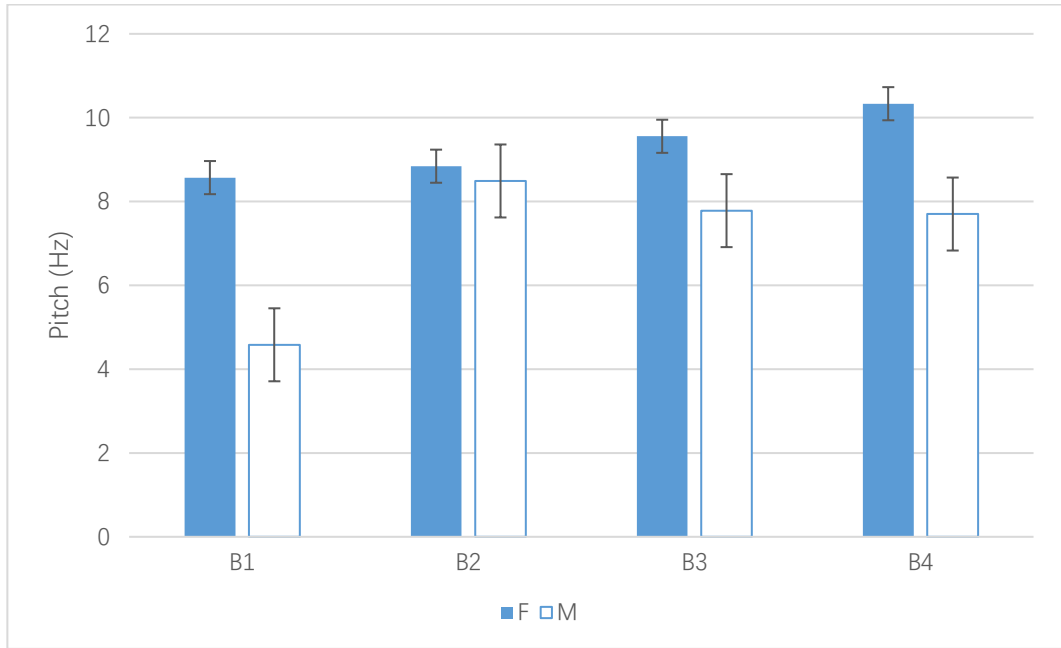


Figure 7: Bar plots of female and male tonal differences across blocks

Similar to the overall data, paired t-tests were performed comparing values in B2, B3 and B4 versus B1. None of the subsequent blocks were significantly different from baseline for females, while they were all significant for males. Table 11 shows the statistical values.

	B1/B2	B1/B3	B1/B4
t - female	-.249	-.879	-1.581
p (one-tailed) - female	.403	.197	.067
p (two-tailed) - female	.806	.393	.135
t - male	-3.797	-3.511	-4.055
p (one-tailed) - male	.0008	.0016	.0005
p (two-tailed) - male	.002	.003	.001

Table 11: Results of t-tests comparing B1 with B2, B3 and B4, separated by gender

The imitation scores describe the data in another way, as shown in Table 12 and Figure 8.

	B2-B1	B3-B1	B4-B1
Mean (Hz) - female	0.272	0.986	1.763
SD - female	4.369	4.487	4.461
Mean (Hz) - male	3.909	3.201	3.119
SD - male	4.119	3.647	3.077

Table 12: Means and SDs of imitation scores, separated by gender

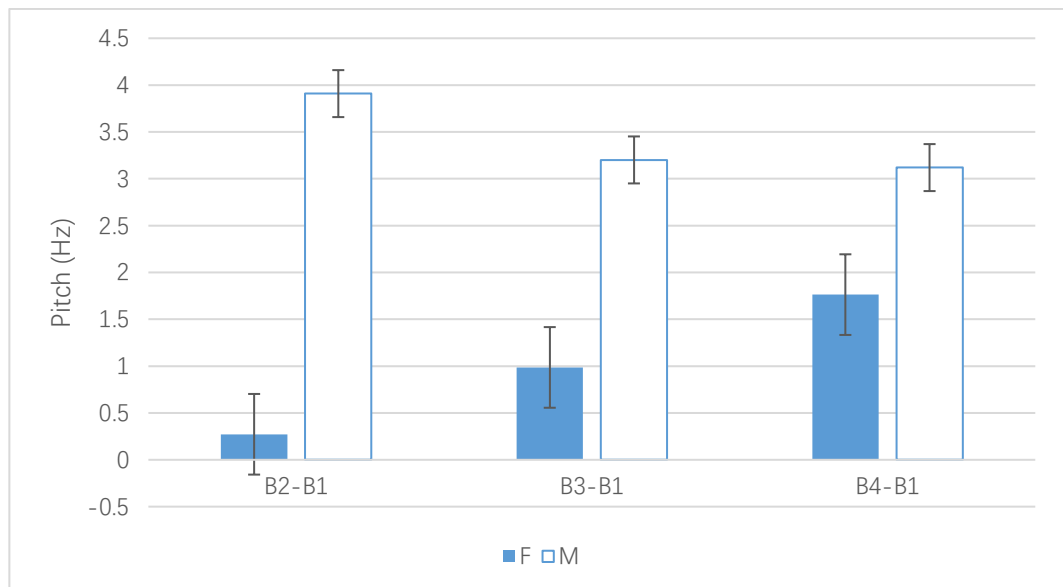


Figure 8: Bar plot of imitation scores, separated by gender

From the above illustrations, males imitated more and earlier compared to females. In fact, females barely imitated in B2, the first shadowing block. On the other hand, females had a bigger tonal difference in the baseline block, which might be a factor contributing their less imitation.

Again, T3 and T6 were separately analyzed, shown in Table 13, Figure 9 and Figure 10. Both males and females show a larger increase in T3 average pitch. Figure 10 plotted male participants with the male model talker for a direct comparison. This was not done for female participants as female pitch is always higher.

	B1T3	B1T6	B2T3	B2T6	B3T3	B3T6	B4T3	B4T6
Female mean	189.52 2	180.95 2	196.55 4	187.71 2	198.85 4	189.29 7	195.09 0	184.75 7
Female SD	12.420	12.261	12.220	11.980	12.527	13.564	13.305	13.038
Male mean	101.52 3	96.941	104.69 3	96.203	105.63 9	97.856	105.18 6	97.486
Male SD	10.206	9.707	8.770	8.232	8.596	7.473	10.850	9.065

Table 13: Means of T3 and T6 values in each block, separated by gender

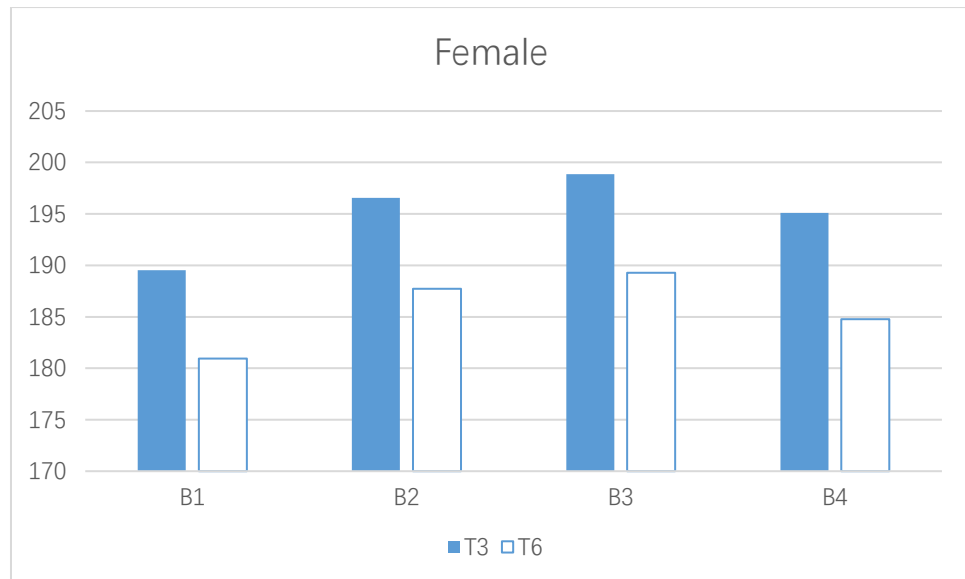


Figure 9: Bar plot of female T3 and T6 values in each block

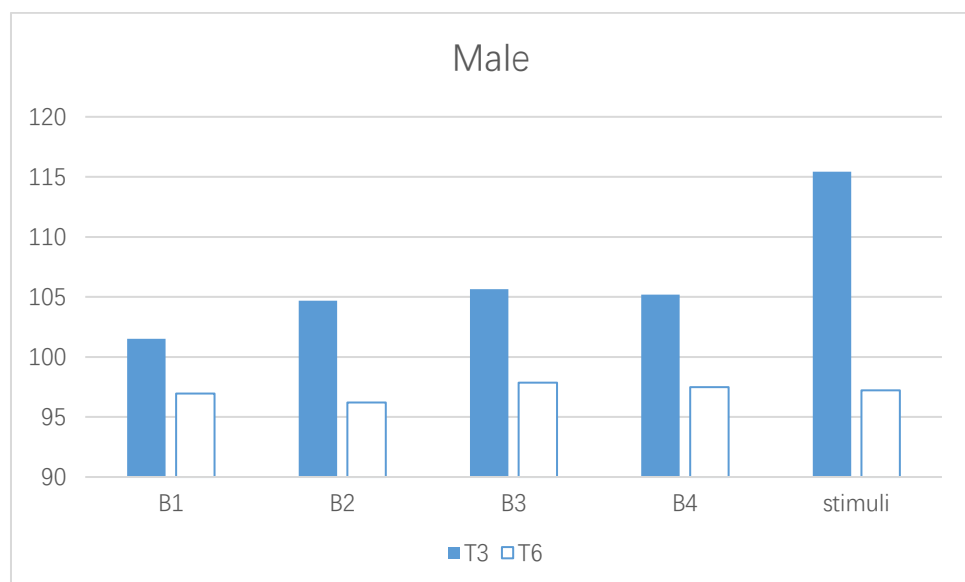


Figure 10: Bar plot of male T3 and T6 values in each block, together with the stimuli (model talker's production)

The above analyses showed a big gender difference in the original degree of merging T3 and T6, and in the pattern of imitation. In general, males imitated more and earlier, and the imitation was statistically significant. Though females still imitated, the tonal difference increase was most obvious in the post-exposure block,

and none of the subsequent blocks were significantly different from the baseline block.

5.3. Speech production: by perceived age

There are two additional variables in this category. One is the answer to “do you agree that lazy accent is predominantly spoken by young people” from the explicit attitude questionnaire, on a 1-7 scale (only integers are possible). Another is subject’s perceived age of the model talker: subjects gave an interval of 10 years old for their estimates of the model talker’s age (e.g. 20-30, 30-40), and the upper thresholds were coded and resolved into the variable used here. Figure 11 and Figure 12 show the distributions of those two variables.

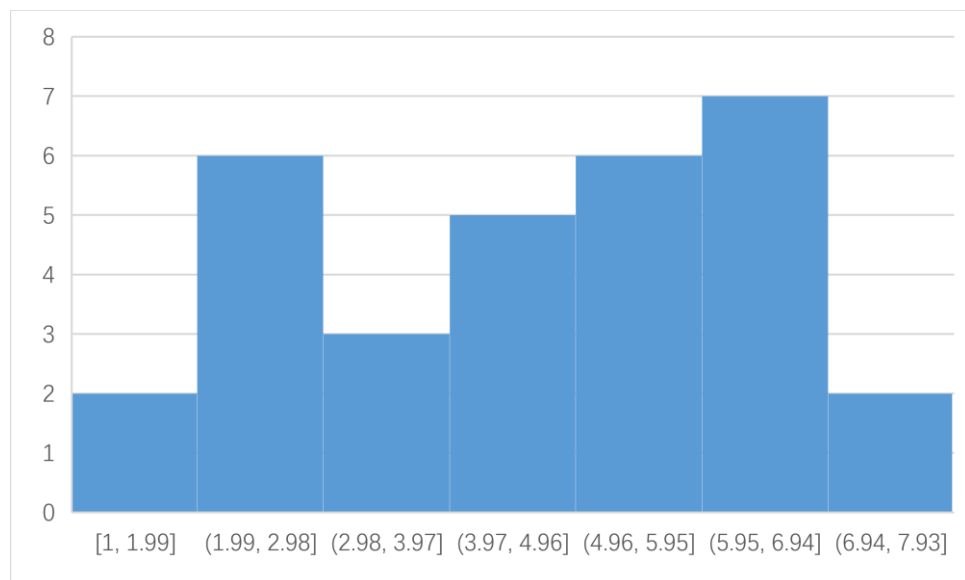


Figure 11: Opinion on young age-lazy accent association

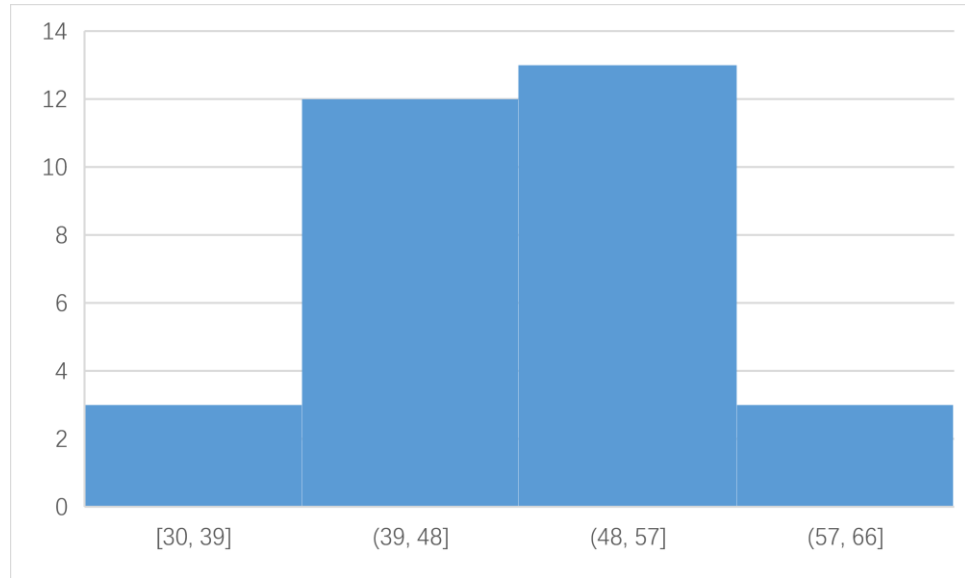


Figure 12: Perceived age of the model talker

Correlation analyses were performed among the abovementioned two age-related variables, and the imitation scores – the shadowing and post-exposure blocks' differences in (T3-T6) from the baseline block. The correlation coefficients are listed in Table 14.

	b2-b1	b3-b1	b4-b1	young-lazy	perceived age
b2-b1	1				
b3-b1	0.37073	1			
b4-b1	0.504281	0.616158	1		
young-lazy	-0.05477	0.04206	0.01139	1	
perceived age	0.315472	0.302966	-0.0378	-0.0592	1

Table 14: Correlation among imitation scores, young age-lazy accent association, and perceived age of the model talker

The perceived age had a moderate positive correlation with the imitation scores of the two shadowing blocks, but not the post-exposure block; the opinion on young age-lazy accent association was not correlated with the imitation scores.

5.4. Language attitudes

From participant responses after the IAT test, it was found that many participants had difficulty telling the correct/lazy pronunciations apart for certain stimuli. The confusion was largely individually different. In Table 15 below, the confusing stimuli for each subject are marked as “n”. Because of experimenter mistakes, the record of four subjects, 104, 112, 115 and 216 were missing.

Subject	女 (n/l)	士 (gw/g)	國 (ng/0)	家 (ng/n)	牙 (k/t)
101					n
102		n			n
103				n	n
104					
105		n			n
106		n			n
107					
108			n		
109					n
110		n			n
111				n	n
112					
113			n		n
114			n	n	
115					
116		n			
201	n	n	n	n	n
202	n				n
203	n		n		
204		n			n
205	n	n	n	n	n
206				n	n
207	n				
208				n	n
209		n			
210					n
211					
212					n
213		n			
214		n			
215				n	n
216					

Table 15: IAT stimuli confusion of each participant

It was not expected that the subjects had perceptual confusion towards these typical “lazy accent” segments. Confusion might have caused delay in response time, so trials on their self-reported confusing stimuli were removed for a new round of IAT score calculation. Therefore, two sets of IAT scores were generated: version 1, with all trials inclusive; and version 2, without confusing trials. The explicit attitude scores were calculated following the method described previously, and no such problem was detected. The descriptions of these three scores are listed in Table 16.

	IAT v.1 (-2~2)	IAT v.1 (-2~2)	explicit (1~7)
Mean	.224	.311	4.449
SD	.363	.403	.809
Max	.831	1.049	6.071
Min	-.483	-.450	3.143

Table 16: Means, SDs, maximum and minimum values of the three language attitude scores

All three scores were centered to a -1 to 1 scale for direct comparison. Table 17, Figure 13 and Figure 14 describe the values.

	IAT v.1 (-1~1)	IAT v.1 (-1~1)	explicit (-1~1)
Mean	.112	.156	.150
SD	.181	.202	.270
Max	.415	.524	.690
Min	-.241	-.225	-.286

Table 17: Means, SDs, maximum and minimum values of the three language attitude scores centered to -1~1

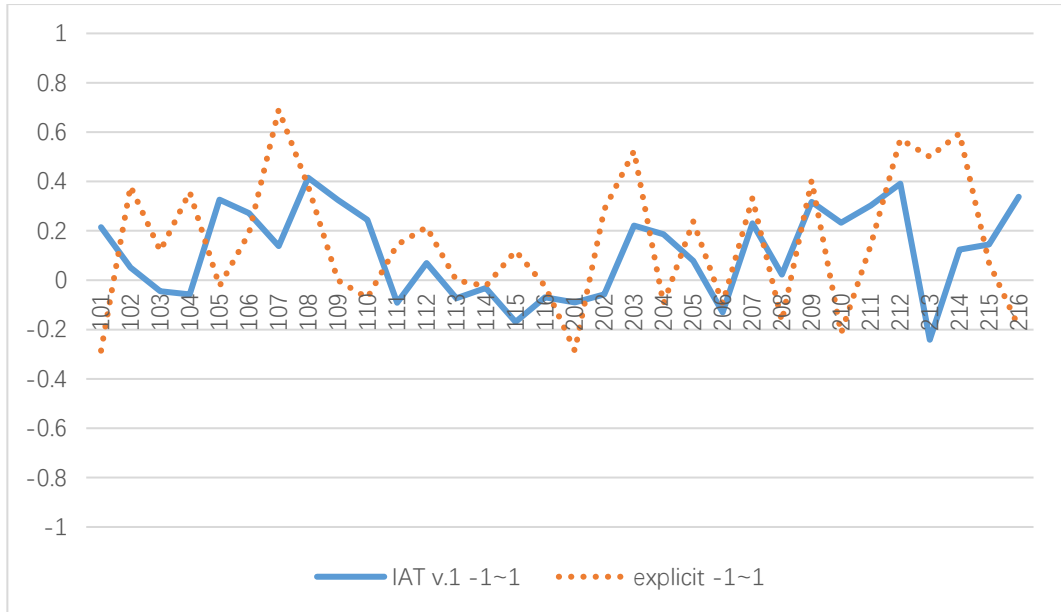


Figure 13: Line chart of first IAT scores and explicit attitude scores

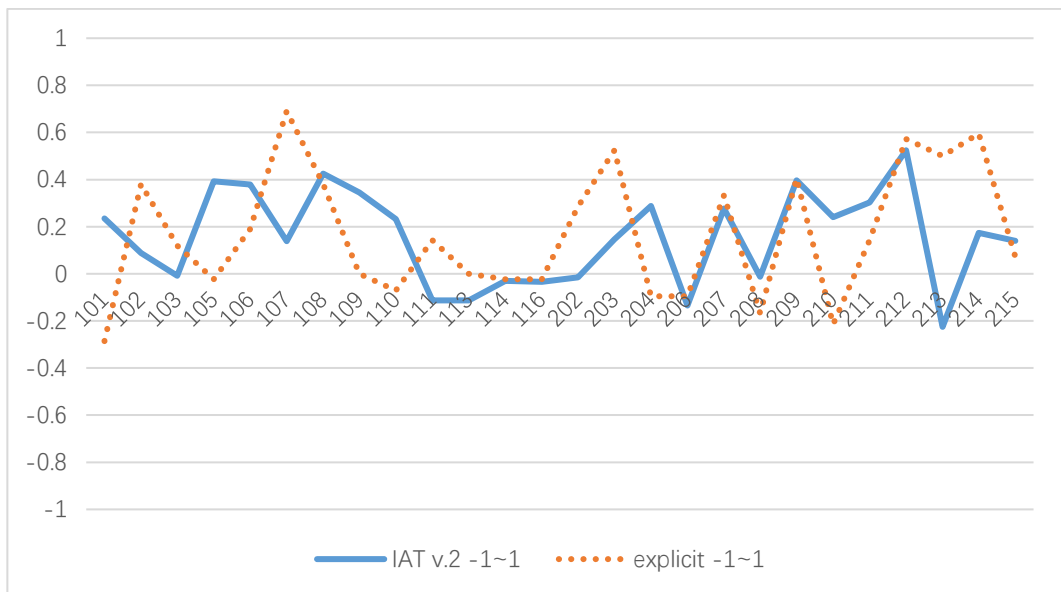


Figure 14: Line chart of second IAT scores and explicit attitude scores

Correlation analyses and paired t-tests among the three scores were performed. According to Table 18 and Table 19, implicit and explicit scores don't have strong correlation, while not being significantly different in their means.

	IAT v.1 -1~1	IAT v.2 -1~1
explicit -1~1	.077	.135

Table 18: Correlation of two versions of IAT scores

	IAT v.1 / explicit	IAT v.2 / explicit
t	-.676	-.179
p (one-tailed)	.252	.430
p (two-tailed)	.504	.859

Table 19: T-test of two versions of IAT scores

5.5. Speech production: by language attitudes

The three language attitude scores scaled from -1 to 1 were correlated with the imitation scores, the shadowing and post-exposure blocks' differences in (T3-T6) from the baseline block. The correlation coefficients are listed in Table 20.

	B2-B1	B3-B1	B4-B1
IAT v.1 -1~1	-.010	-.093	.070
IAT v.2 -1~1	-.263	-.321	.256
explicit -1~1	.067	.325	-.102

Table 20: Correlation among imitation scores and language attitude scores

The IAT score after excluding confusion trials had moderate negative correlation with imitation scores in the two shadowing blocks B2 and B3, but positive correlation with that in the post-exposure block B4. Interestingly, the explicit attitude score positively correlated with imitation score in B3, but did not show obvious correlation with B2 and B4. As the original IAT score might have

been affected by stimuli confusion, it is not surprising to see very weak correlation with any of the imitation scores.

5.6. Speech production: individual differences

Phonetic imitation could be an individually varied behavior. Figure 15 and Figure 16 visualized the (T3-T6) change from B1 to B4 for each subject.

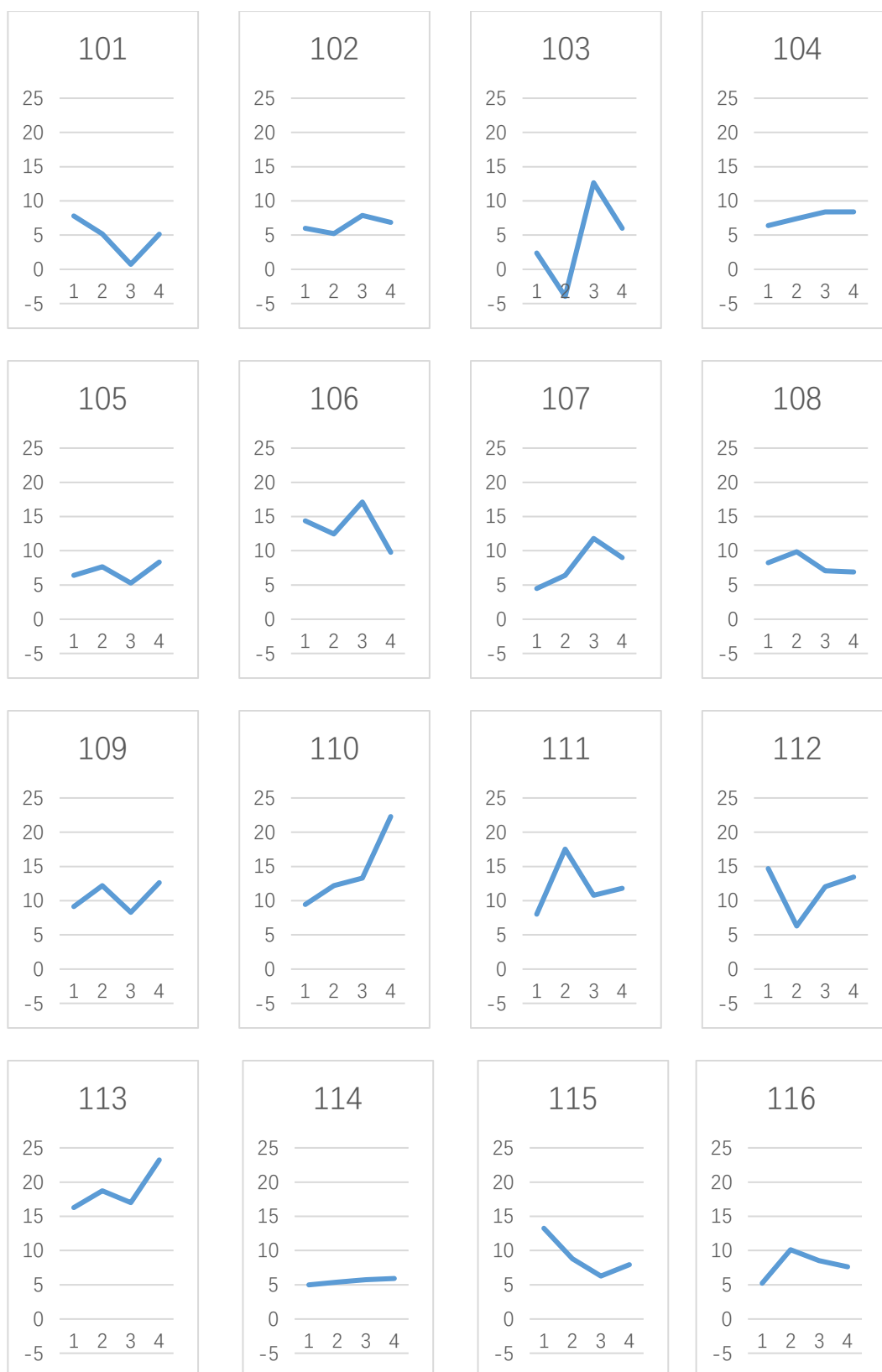


Figure 15: Female individual tonal difference change from B1 to B4

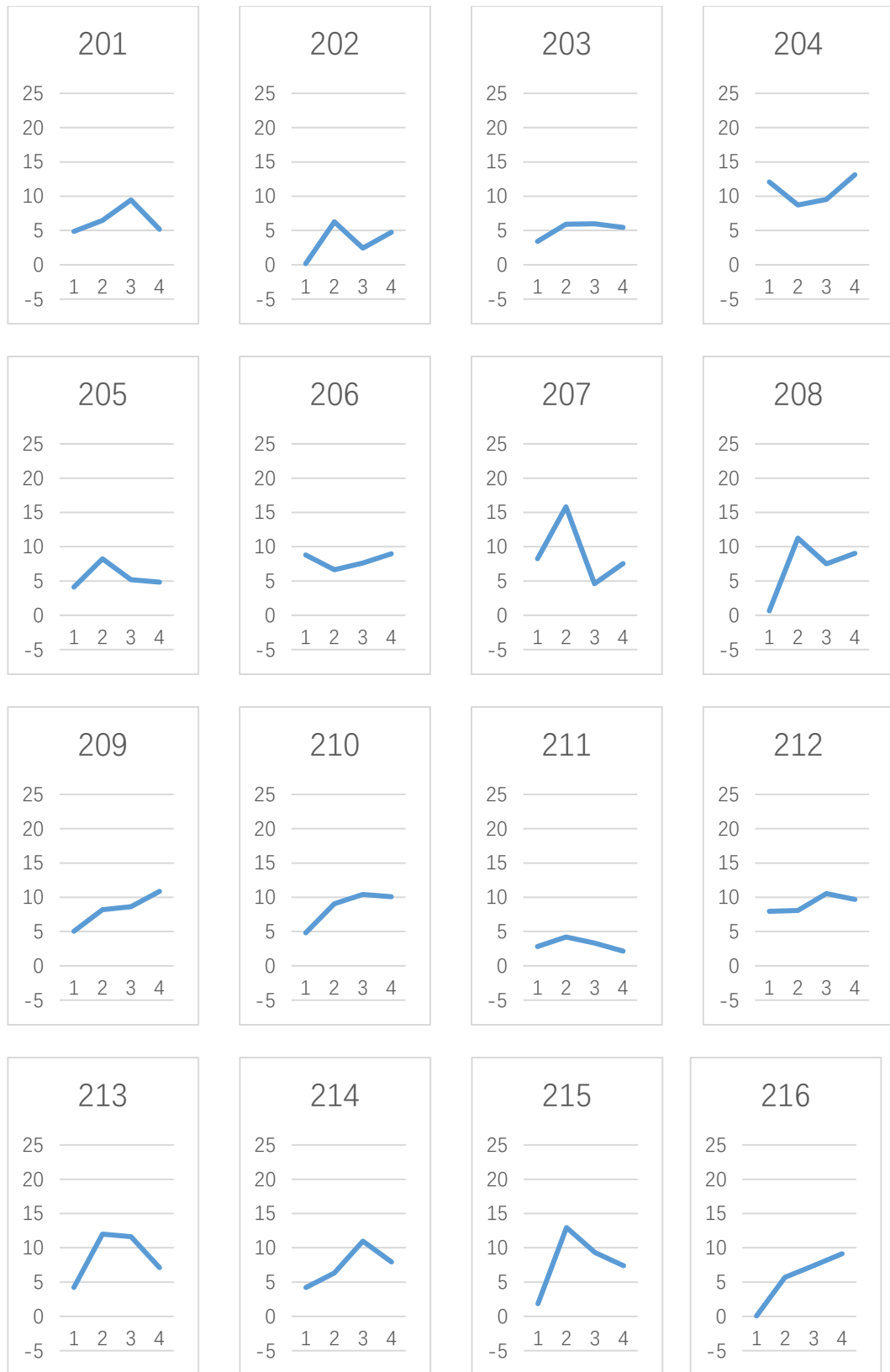


Figure 16: Male individual tonal difference change from B1 to B4

Collectively, the participants imitated the model talker, and the imitation was statistically significant. Looking closer into individual data, there are some whose initial T3/T6 differences were already larger than 10 Hz (106, 112, 113, 115, 204); some of them distinguished T3/T6 even more while some did not. Some subjects even had smaller T3/T6 differences after exposure (101, 112, 115); some had largely different productions in B2 and B3, the two shadowing blocks (103, 207); and some, in line with the overall trend, showed obvious increase in T3/T6 difference (107, 110, 113, 202, 208, 209, 210, 214, 215, 216). The general pattern of male participant is more consistent with the overall result than females, possibly because the imitation is mainly driven by them, as discussed earlier in this chapter.

5.7. Mixed Effects Models

The output of the two mixed effects models described are listed below.

	t	p
block2	10.354	< 2e-16
block3	14.436	< 2e-16
block4	8.697	< 2e-16
tone	-4.081	.001
gender	-24.391	< 2e-16
block2:tone	-2.726	.006
block3:tone	-2.771	.006
block4:tone	-3.145	.002

Table 21: Results of the first mixed effects model, using pitch values from each production

	t	p
block2	3.154	.002
block3	2.885	.005
block4	3.396	.001
gender	-.443	.662
IAT v.1	.239	.814
IAT v.2	-.194	.848
explicit language attitude	-.557	.584
opinion on young-lazy	1.650	.115
accent association		
perceived age of model talker	.291	.774

Table 22: Results of the second mixed effects model, using tonal differences in each block

The models verified again that T3/T6 difference did increase for the participant pool during and after exposure to the model talker. Factors including gender, language attitude, and age perception, are not significant predictors in the amount of imitation through the blocks, though they did show trend differences described in this chapter, especially gender.

As word frequency was only a valid predictor in the first and fourth blocks where visual stimuli were shown, a smaller model with half data points were built, and no significant effect of word frequency was found.

5.8. Summary

Results from all subjects showed that, young participants in the experiment imitated the model talker's speech pattern by enlarging their T3/T6 tonal differences. The tonal difference increase is statistically significant for all the three shadowing and post-exposure blocks, and the effect is robust as shown through several statistical methods. The averaged tonal difference in Block 4 after exposure has a 37.1% increase from the baseline production. This increase was achieved by raising the pitch level of T3. On absolute numbers, though, the increase is not large (around 2 Hz), which still has a distance from the model talker's tonal difference of 18 Hz.

Male and female subjects behaved very differently. Males contributed more to the overall imitation trend, and imitated the most in the first shadowing block B2. On the other hand, females just imitated slightly, and the largest imitation was found in the post-exposure block B4. Males also showed a more consistent pattern of increasing the tonal difference when looking at individual trends.

Participants' perception on the model talker's age correlated positively with the imitation in shadowing blocks B2 and B3, though the effect was not carried on to post-exposure block B4. Age was not a significant predictor in the mixed effects model of tonal differences.

Implicit and explicit language attitude tests showed interesting individual variations in Cantonese segment perception: most participants had difficulty distinguishing one or more pairs of segments undergoing sound change, and the specific pairs they had difficulty in were individually varied. The implicit and explicit scores were not significantly different in means, but not strongly correlated.

The explicit score positively correlated with the imitation score in the second shadowing block B3, but not with other blocks. The implicit and explicit language attitude scores were neither significant in the mixed effects model of tonal differences.

Analyzing deeper into every single subject, the dataset showed individual difference in this phonetic imitation behavior. Males showed more consistent increase in T3/T6 difference, while females had more varied patterns.

Chapter 6. Discussion

The main objective of this research is to investigate whether young Hong Kong Cantonese (HKC) speakers could unmerge the mid-level tone T3 and the low-level tone T6 after exposure to unmerged speech, and under which conditions. 32 young HKC native speakers participated in the study. Each of them completed a speech production experiment which exposed subjects to an unmerged, 60-year-old model talker, an Implicit Association Test (IAT), and an explicit language attitude questionnaire. The results showed that participants imitated the model talker, and increased their (T3-T6) tonal difference during and after the speech exposure. Tonal difference increase was mainly achieved by T3 raising its position. Male and female subjects showed different imitation patterns: males imitated more and faster than females. Age prediction on the model talker, and implicit and explicit language attitude towards the “lazy accent” variety, both correlated with imitation in some but not all the blocks. There were individual differences in both phonetic imitation patterns and perception of Cantonese segments undergoing sound change.

Significance of all three exposure and post-exposure blocks comparing to the baseline block provided evidence for the prediction, that T3 and T6 can be unmerged due to phonetic imitation of the model talker’s speech pattern. Broadly, the results extended the present literature of phonetic imitation, that this phenomenon is also applicable to tones in tonal languages; and confirmed the prevalence of merging T3 and T6 among young people in Hong Kong. The imitation effect was larger and appeared earlier than the study of Babel et al. (2013) on vowel mergers in New Zealand English, probably because HKC tone mergers remain largely unstable and prone to change, in both the speech community at large and across individuals (Mok et al., 2013). The post-exposure block B4 had an even

larger difference with the baseline block B1 than the second shadowing block B3, which was not expected in the predictions. This outcome echoes with that of Babel et al. (2013), but not Babel (2012). The reason why the effect was maintained and even strengthened after removal of speech stimulation deserves to be studied deeply in the future.

Female and male participants showed very different patterns throughout the imitation process, while in the mixed effects model of tonal differences, gender was not a significant predictor. The role of gender in phonetic imitation has been under debate with varied results obtained. Namy et al. (2002) found that females were perceived as imitating more than males, and female subjects imitated more towards the male model talker rather than the female one. This study only had perceptual judgments on stimuli's overall similarity, but not specific features or acoustic analysis. In Babel and Bulatov (2011) on fundamental frequency in phonetic imitation, listeners judged females to imitate more than males towards the male model talker, while acoustic analysis showed that males imitated more actually. Babel et al. (2013) had few male participants, who were primary contributors to the overall imitation towards the male model talker rather than the female majority. The current study aligns with the findings of Babel and Bulatov (2011) and Babel et al. (2013), that males imitate more than females in this Cantonese lexical tone pair. It could be because male participants formed a closer social distance with the male model talker than the females, or because males had merged T3/T6 more than females. These results support the view of Eckert (1989), that no simple conclusion can be drawn on gendered speech being more or less conservative.

Implicit and explicit language attitudes both correlated with imitation scores in some blocks, but the overall effects were not significant. The perceptual confusions

discovered in this IAT led to two sets of IAT scores and low correlation with the explicit attitudes. Although the self-reported confusing trials were subsequently removed before generating a new set of scores, being unclear at some of the stimuli might have affected overall reaction time. Some of the previous research did find correlation between degree of imitation and implicit attitude. Babel (2009) tested the effect of implicit attitude towards African and Caucasian Americans, and Babel et al. (2013) found that the more positive New Zealanders were towards Australia, the more they imitated the Australian model talker. On the other hand, IAT scores should not be interpreted in overly simple ways, because they might have shown familiarity towards the stimuli rather than attitude. Although subjects did have varied attitudes toward “lazy accent”, a lot of them had scores around the middle of the scale, showing that they might not have strong attitudes; therefore, their scores could change should the questions were worded in other ways, which added uncertainty to the data obtained. Nevertheless, the correlations in some of the blocks have preliminary implications in language attitude influence on phonetic imitation, and future work needs to be done along this direction.

Perceived age positively correlated with imitation in B2 and B3, but did not correlate with the post-exposure block B4. This is probably because, in B4 the subjects were no longer exposed to the model talker, and therefore the direct trait of age was not presented in B4. The possible effect of age can go bi-directionally: the older the subjects perceived the model talker to be, the wider social distance will it create (as the subjects were all young people), but in turn, the accreditation on the “standardness” of the speech will also be higher. From the results in B2 and B3, it seemed that the second approach of converging to the model talker because of speech standardness was adopted by most of the participants.

Discoveries on the big individual differences in “lazy accent” segments’ perception were not expected before the experiment. It was assumed that, with the efforts of promoting “correct” instead of “lazy” pronunciation in public, young HKC speakers would be aware of the differences in both perception and production, while consciously preferring to speak in a more comfortable way. It turned out that many distinguished the typical initials and finals of “lazy accent” in neither perception nor production, let alone the tone categories that are more invisible in the whole sound change. It was interesting that, though most participants reported having difficulty in differentiating the -k and -t syllable codas, their perception of other words varied greatly among individuals. The seemingly imperfect results from IAT could be meaningful not only for scholars studying the rapid HKC sound change, but also for IAT users: the results implied that IAT would have the best performance when all stimuli equally “stood out” (Rothermund & Wentura, 2001), and would probably generate unsatisfactory results if the understandability of some stimuli largely differ from the others.

Chapter 7. Conclusion

This study investigated phonetic imitation of an ongoing merger-in-progress, using Hong Kong Cantonese merging tones T3 and T6 as examples. The experiment included a speech production task involving reading and shadowing, and two language attitude tests. Results showed significant increase of T3-T6 tonal difference, which means that the participant imitated the model talker despite the reversed direction in the sound change process. Variation in imitation patterns according to social factors including gender, language attitude, and age was discovered.

The findings have important implications in several directions. First, sound change on a community level could be flexible in individual levels through communication, as phonetic imitation operated on the reversed direction of sound change. Second, this is one of the first research on phonetic imitation of lexical tones. The fact that tones were indeed imitated is meaningful to the universality of imitation behavior in varied speech communities. Third, this is among the first studies approaching Hong Kong Cantonese tone mergers through sociolinguistics. The variation patterns showed both similarities and differences in sound change of other languages, and the social nature of HKC deserves to be discussed. Fourth, results from both the speech production task on tones, and the IAT on segments, showed perception-production asymmetry and its large individual differences in the current HKC “lazy accent” sound change at large. These results contribute to the studies of phonetics, sociolinguistics, and psycholinguistics.

The present study is not an end in exploring the interesting questions on sound change and its variations. There are several directions that future studies can pursue.

First, a perceptual judgment task can be added (Pardo et al., 2013). Perceptual and acoustic measurements could show different results (Babel et al., 2013), and it would be interesting to do this comparison in this study. Second, another group of participants can use a different experiment with dual tasks, or the same experiment after a certain period. Dual tasks in between the productions would involve either verbal or spatial interference, and the objective is to see whether the phonetic imitation effect could still happen when subsequent imitation is interrupted. Longitudinal experiments were rarely done in phonetic imitation studies, while they are important in defining the time limit of phonetic imitation effect. Third, use another group of participants with a female model talker, to further explore the gender dynamics in this phonetic imitation context specifically. Fourth, as there is perceptual confusion on even the most typical “lazy accent” segments in HKC, new paradigms for language attitude retrieval need to be designed to dig deeper into its potential effect. Fifth, the individual differences in HKC segment perceptual confusion is very interesting. Separate experiments can be done to explore the possible relationship between perception and production in an ongoing sound change.

The world still changes constantly. This study, like other research, is an effort to discover the unchanged patterns underlying the changes. Specifically, this research has peeked into the big picture of the social and psychological nature of human speech in variance. One certain thing is, what we speak does not just mean a string of utterances produced by our own articulators. We speak for our community, our social awareness, and our minds that unite us across the linguistic borders.

References

- Abrego-Collier, C., Grove, J., Sonderegger, M., & Yu, A. C. L. (2011). Effects of speaker evaluation on phonetic convergence. In *Proceedings of the 17th International Congress of the Phonetic Sciences* (pp. 192-195).
- Audacity Team. (2015). Audacity® Version 2.1.0. Audio editor and recorder. Retrieved from <http://www.audacityteam.org/>
- Babel, M. E. (2009). *Phonetic and social selectivity in speech accommodation* (Doctoral dissertation). University of California, Berkeley.
- Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society*, 39(4), 437-456.
- Babel, M., & Bulatov, D. (2011). The role of fundamental frequency in phonetic accommodation. *Language and Speech*, 55(2), 231-248.
- Babel, M., McGuire, G., King, J., & Satterwhite, A. (2011). Vocal aesthetics, prototypicality, and stereotypicality. In *Proceedings of the New Ways of Analyzing Variation 40* (NWAV40), Washington, D.C..
- Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, 40(1), 177-189.
- Babel, M., McAuliffe, M., & Haber, G. (2013). Can mergers-in-progress be unmerged in speech accommodation?. *Frontiers in psychology*, 4, 1-4.
- Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America*, 137(5), 2823-2833.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.

- Bauer, R. S., Cheung, K.-H., & Cheung, P.-M. (2003). Variation and merger of the rising tones in Hong Kong Cantonese. *Language Variation and Change*, 15(2), 211-225.
- Boersma, P. & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.22. Retrieved from <http://www.praat.org/>
- Bolton, K. (2011). Language policy and planning in Hong Kong: Colonial and post-colonial perspectives. *Applied Linguistics Review*, 2(1), 51-71.
- Census and Statistics Department, Hong Kong Special Administrative Region. (2016). *Thematic Household Survey Report No. 59: Pattern of smoking, Personal computer and Internet penetration, Use of language*. Retrieved from <http://www.statistics.gov.hk/pub/B11302592016XXXXB0100.pdf>
- Chen, H. Y. (1998). *Norms of pronunciation and the sociolinguistics of Cantonese in Hong Kong* (Master of Philosophy thesis). University of Hong Kong.
- Cheung, H.-N. S. (2002). Cantonese of the 21st Century: Formation of a New Phonological System in the Hong Kong Language. *Journal of Jinan University (Philosophy & Social Science Edition)*, 24(2), 25-40.
- Cheung, K. H., Chin, A., Chung, P., Lee, C., Lun, C., Man, V., & Tang, S. W. (1993). *Linguistic Society of Hong Kong Jyutping Word List*. Retrieved from <http://www.lshk.org/jyutping>
- Cui, L., & Zhang, G. (2004). The Studies of the Implicit Association Test (IAT). *Psychological Science*, 27(1), 161-164.
- Ding, P. S. (2010). Phonological change in Hong Kong Cantonese through language contact with Chinese topolects and English over the past century. In Millar, R. M. (Ed.), *Marginal Dialects: Scotland, Ireland and Beyond*. Aberdeen: Forum

- for Research on the Languages of Scotland and Ireland* (198-218). Aberdeen, UK: The University of Aberdeen.
- Eckert, P. (1989). The whole woman: Sex and gender differences in variation. *Language variation and change*, 1(3), 245-267.
- Fung, R. S. Y., & Wong, C. S. P. (2010). Mergers and near-mergers in Hong Kong Cantonese tones. Paper presented at *Tone and Intonation in Europe (TIE4)*, Stockholm, Sweden.
- Fung, R. S. Y., & Wong, C. S. P. (2011). Acoustic analysis of the new rising tone in Hong Kong Cantonese. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 716-718).
- Fung, R., Kung, C., Law, S. P., Su, I. F., & Wong, C. (2012). Near-merger in Hong Kong Cantonese tones: A behavioural and ERP study. Paper presented at *The 3rd International Symposium on Tonal Aspects of Languages (TAL 2012)*, Nanjing, China.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, 105(2), 251.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology*, 74(6), 1464-1480.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An Improved Scoring Algorithm. *Journal of personality and social psychology*, 85(2), 197-216.
- Greenwald, A. G. (2007). Laboratory software for the IAT. Retrieved from http://faculty.washington.edu/agg/iat_materials.htm

- Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the Implicit Association Test can have societally large effects. *Journal of personality and social psychology*, 108(4), 553-561.
- Ho, M. W. (1989). *The Correct Pronunciation of Cantonese: Demonstration*. Hong Kong: Publications (Holdings) Ltd.
- Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology*, 81(5), 774-788.
- Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, 2(1), 125-156.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). Package ‘lmerTest’. *R package version*, 2.0-32.
- Kuo, W.-J., Yeh, T.-C., Lee, C.-Y., Wu, Y.-T., Chou, C.-C., Ho, L.-T., Hung, D. L., Tzeng, O. J. L., & Hsieh, J.-C. (2003). Frequency effects of Chinese character processing in the brain: an event-related fMRI study. *Neuroimage*, 18(3), 720-730.
- Lennes, M. (2011). *SpeCT - The Speech Corpus Toolkit for Praat*. Retrieved from http://www.helsinki.fi/~lennes/praat-scripts/public/collect_pitch_data_from_files.praat
- Luo, Q., & Durvasula, K. (2013). Social Factors in the Rising Tone Merger in Guangzhou Cantonese. Paper presented at *The 22nd Annual Conference of International Association of Chinese Linguistics & the 26th North American Conference on Chinese Linguistics*, University of Maryland-College Park, MD.

- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior research methods*, 44(2), 314-324.
- Mok, P. P.-K., & Wong, P. W.-Y. (2010). Production of the merging tones in Hong Kong Cantonese: preliminary data on monosyllables. In *Proceedings of Speech Prosody 2010*, Chicago, IL.
- Mok, P. P. K., Zuo, D., & Wong, P. W. Y. (2013). Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese. *Language variation and change*, 25(2013), 341-370.
- Namy, L. L., Nygaard, L. C., & Sauerteig, D. (2002). Gender Differences in Vocal Accommodation: The Role of Perception. *Journal of Language and Social Psychology*, 21(4), 422-432.
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin*, 31(2), 166-180.
- Pantos, A. J., & Perkins, A. W. (2013). Measuring Implicit and Explicit Attitudes Toward Foreign Accented Speech. *Journal of Language and Social Psychology*, 32(1), 3-20.
- Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, 119(4), 2382-2393.
- Pardo, J. S. (2010). Expressing oneself in conversational interaction. In Morsella, E. (Ed.), *Expressing oneself/expressing one's self: Communication, cognition, language, and identity* (183-196). Hove, England: Psychology Press/Taylor & Francis.

- Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics*, 40(1), 190-197.
- Pardo, J. (2013). Measuring phonetic convergence in speech production. *Frontiers in psychology*, 4, 559.
- Pardo, J. S., Jordan, K., Mallari, R., Scanlon, C., & Lewandowski, E. (2013). Phonetic convergence in shadowed speech: the relation between acoustic and perceptual measures. *Journal of Memory and Language*, 69(3), 183-195.
- Peng, G. (2006). Temporal and Tonal Aspects of Chinese Syllables: A Corpus-based Comparative Study of Mandarin and Cantonese. *Journal of Chinese Linguistics*, 34(1), 134-154.
- Peng, G., & Wang, W. S.-Y. (2005). Tone recognition of continuous Cantonese speech based on support vector machines. *Speech Communication*, 45(1), 49-62.
- Pérez, E. O. (2010). Explicit evidence on the import of implicit attitudes: The IAT and immigration policy judgments. *Political Behavior*, 32(4), 517-545.
- Phillips, B. S. (1984). Word frequency and the actuation of sound change. *Language*, 320-342.
- Powell, M. J. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06*, University of Cambridge, Cambridge, UK.
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Radio Television Hong Kong. (2008). *Speak Proper Cantonese*. Retrieved from <http://rthk.hk/elearning/speakpropercantonese/>

Research Centre for the Humanities Computing, Chinese University of Hong Kong.

(2003). *A Chinese Character Database: With Word-formations phonologically disambiguated according to the Cantonese Dialect*. Retrieved from <http://humanum.arts.cuhk.edu.hk/Lexis/lexi-can/>

Rothermund, K., & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test (IAT). *Zeitschrift für experimentelle Psychologie*, 48(2), 94-106.

So, L. K. H. (1996). Tonal changes in Hong Kong Cantonese. *Current Issues in Language & Society*, 3(2), 186-189.

To, C. K., Mcleod, S., & Cheung, P. S. (2015). Phonetic variations and sound changes in Hong Kong Cantonese: Diachronic review, synchronic study and implications for speech sound assessment. *Clinical linguistics & phonetics*, 29(5), 333-353.

Trudgill, P. (2008). Colonial dialect contact in the history of European languages: On the irrelevance of identity to new-dialect formation. *Language in Society*, 37(02), 241-254.

van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior research methods*, 38(4), 584-589.

Yao, Y., & Chang, C. B. (2016). On the cognitive basis of contact-induced sound change: Vowel merger reversal in Shanghainese. *Language*, 92(2), 433-467.

Yu, A. C. L. (2007). Understanding near mergers: The case of morphological tone in Cantonese. *Phonology*, 24(2007), 187-214.

Yu, A. C. L., Abrego-Collier, C., Baglini, R., Grano, T., Martinovic, M., Otte III, C., Thomas, J., & Urban, J. (2011). Speaker attitude and sexual orientation

affect phonetic imitation. *University of Pennsylvania Working Papers in Linguistics*, 17(1), 26.

Yu, A. C. L., Abrego-Collier, C., & Sonderegger, M. (2013). Phonetic imitation from an individual-difference perspective: Subjective attitude, personality and “autistic” traits. *PloS one*, 8(9), 1-13.

Yuen, R. W. M., Chan, T. Y. W., Lai, T. B. Y., Kwong, O. Y., & T'sou, B. K. Y. (2004). Morpheme-based derivation of bipolar semantic orientation of Chinese words. In *Proceedings of the 20th International Conference on Computational Linguistics* (p. 1008). Association for Computational Linguistics.

Appendices

Appendix 1. Speech production task stimuli list

Critical: (Transcription in Jyutping and IPA; numbers are occurrences in the Cantonese word database)

Syllable	T3 high freq	T3 low freq	T6 high freq	T6 low freq
Full tone set				
ji [ji]	意 15534	懿 48	義 20250	貳 11
sin [sin]	線 19039	扇 574	善 2546	擅 163
fu [fu]	富 3882	褲 324	負 4148	赴 452
Semi-full tone set				
bou [pou]	報 8349	怖 158	部 28458	捕 1461
gei [kɛi]	記 7653	寄 904	技 6690	忌 178
zoeng [tʃœŋ]	障 1160	帳 450	象 13709	匠 242

Filler: (Transcription in Jyutping and IPA; numbers are occurrences in the Cantonese word database)

Tone	Syllable	High freq	Low freq	Tone	Syllable	High freq	Low freq
T1	soeng	商	霜	T4	jau	由	悠
	[sœŋ]	6948	289			21176	322
	jing	英	櫻		jing	形	螢
	[jɪŋ]	4797	171			20035	42
	gaai	階	佳		ming	明	鳴
	[kai]	9631	408			17766	654
T2	soeng	想	賞	T5	jau	友	誘
	[sœŋ]	14661	699			3088	349
	ci [tsh̥i]	此	恥		lei	理	李
		16270	269			25814	4744
	zyu	主	煮		lou	老	鹵
	[tsy]	38619	525			12882	206

Appendix 2. Implicit Association Test stimuli list

Targets: (Transcription in Jyutping and IPA)

Test item	Stimuli	“Correct”	“Lazy”
[l] & [n] (initial)	女士	neoi [nøy]	leoi [løy]
[kʷ] & [k] (initial)	過去	gwo [kʷɔ]	go [kɔ]
[ŋ] & Ø (initial)	牙齒	ngaa [ŋa]	aa [a]
[ŋ] & [n] (final)	朋友	pang [pʰaŋ]	pan [pʰan]
[k] & [t] (final coda)	學校	hok [hɔk]	hot [hɔt]

Attributes:

Positive	Negative
誠實 Honest	虛偽 Hypocritical
聰明 Clever	愚蠢 Stupid
優秀 Good	惡劣 Bad
善良 Kind	殘暴 Cruel
英勇 Brave	懦弱 Weak